THÈSE DE DOCTORAT

présentée par

Hamed AMINI

Spécialité : MATHÉMATIQUES APPLIQUÉES

Équipe d'acceuil : École Normale Supérieure - INRIA Rocquencourt (Projet TREC)

Epidemics and Percolation in Random Networks

soutenue le 24 Juin 2011, devant le jury composé de:

Venkat ANANTHARAM	University of California, Berkeley	Examinateur
François BACCELLI	INRIA - École Normale Supérieure	Directeur de thèse
Francis COMETS	Université Paris 7	Examinateur
Remco van der HOFSTAD	Eindhoven University of Technology	Rapporteur
Marc LELARGE	INRIA - École Normale Supérieure	Directeur de thèse
Laurent MASSOULIÉ	Technicolor Research	Rapporteur
Gilles SCHAEFFER	École Polytechnique	Examinateur

To my family

Remerciements

Ma reconnaissance revient principalement à mes directeurs de thèse Francois Baccelli et Marc Lelarge qui, au cours de ces quatre années, ont encadré mon travail sur un sujet passionnant. Je les remercie pour leur soutien et leur conseils judicieux. Merci de m'avoir fait découvrir le monde de la recherche.

Je tiens à remercier Remco van der Hofstad et Laurent Massoulié qui ont eu la lourde tâche de rapporteurs. Je les remercie infiniment pour la lecture et la correction de mon manuscrit et pour l'intérêt qu'ils ont portè à mon travail. Je remercie également les autres membres du jury, Venkat Anantharam, Francis Comets et Gilles Schaeffer de m'avoir fait l'honneur de participer à mon jury de thèse.

I would like to thank several people for the privilege to visit their group and work with them during my thesis. First, Moez Draief at Imperial College in London, then Nikolaos Fountoulakis at Max-Planck-Institut in Saarbrcken, and finally Yuval Peres for visiting his group at Microsoft research in Redmond. I'm grateful for fruitful discussion with the members of these respective groups.

L'application en finance d'une partie de mes travaux n'aurait pas été possible sans l'apport de Rama Cont et Andreea Minca, à qui je remercie également.

J'ai eu la chance de passer ma thèse au projet TREC où j'ai beaucoup apprécié la grande liberté offerte pour la recherche, dans une ambiance stimulante et chaleureuse. Je tiens à remercier tous les collègues du projet TREC et en particulier Ana, Anne, Bartek, Bruno, Calvin, Emilie, Furcy, François-Xavier, Frédéric, Florence, Giovanna, Justin, Mathieu, Minh-ahn, Nadir, Omid, Pierre, Tien, Van Minh et Yogesh.

Je remercie également Nathalie Abiola, Claudette Dessertaine, Christine Ferret, Joelle Isnard, Audrey Lemarechal et Hélne Milome pour leur gentillesse et leur aide dans les tâches administratives.

Les dernières lignes de ces remerciements vont à mes parents, ma soeur Hadis, à Pooran et à Andreea qui mont toujours soutenu durant cette thèse. Enfin, je remercie mille fois mon frère Omid pour son soutien et sa présence au moment difficile.

Preface

L'omniprésence des réseaux dans tous les aspects de la vie moderne et les problèmes associés ont incité de nombreuses recherches dans le but de comprendre et prédire le comportement de telles structures irrégulièrs, complexes, et évolutives. Les exemples de réseaux vont de l'Internet jusqu'aux interconnexions d'agents financiers ou bien aux réseaux neuronaux.

Doté de grande capacité de calcul et de la disponibilité de larges ensembles de données sur des réseaux réels (comme l'Internet, les réseaux de transport, réseaux de téléphone, réseaux coauteurs scientifiques, réseaux protéine-protéine biologiques, etc.), l'étude de réseaux complexes a émergé comme un domaine à croissance rapide.

La plupart des réseaux réels sont très grands, de plusieurs milliers de noeuds dans un réseau d'entreprises à plusieurs milliards de noeuds dans certains systèmes biologiques. En raison de leur complexité, leurs relations ne peuvent être décrites qu'en termes statistiques. L'une des observations les plus frappantes est la suivante : des réseaux avec des fonctions complètement différentes, comme par exemple des réseaux sociaux et des réseaux biologiques, partagent des caractéristiques communes. En particulier, nombre de ces systèmes sont des "petits-mondes", ce qui traduit le fait que la distance topologique moyenne dans le réseau (qui mesure le nombre moyen de liens à franchir sur le réseau pour aller d'un site à un autre) varie très lentement avec le nombre total de sites (typiquement comme un logarithme).

Une autre découverte particulièrement importante est le fait que la fréquence d'apparition de sites avec k voisins (sites de degré k) est une distribution en loi de puissance : si on définit la distribution des degrés par p_k (probabilité qu'un noeud choisi uniformément parmi tous les noeuds ait k voisins), il existe un paramètre γ tel que $p_k \sim k^{-\gamma}$. Ce résultat a permis l'identification d'une nouvelle catégorie de réseaux dits "sans-échelle", et différencie ces réseaux du graphe classique aléatoire proposé par Erdős et Rényi dans les années 60, dans lequel la distribution est

Poisson et donc homogène, dans le sens où le nombre de voisins de chaque noeud fluctue très peu autour d'une valeur moyenne.

Une question essentielle dans l'étude des réseaux est l'influence de la topologie d'un réseau sur sa réponse aux facteurs perturbatifs. Par exemple dans de nombreux domaines, il est essentiel de comprendre et caractériser la robustesse des réseaux, quand on supprime un certain nombre de noeuds, ou de liens. Un autre exemple consiste à étudier la propagation des épidémies sur les différents types de réseaux. Les épidémies sur des graphes permettent de modéliser de nombreux phénomènes dans les réseaux tels que la propagation de virus, de vers, de rumeurs, et aussi les faillites des agents financiers, pour en citer quelques-un. Les systèmes de particules en interaction (comme le processus de contact) ont permis de modéliser avec succès de tels phénomènes. Ces systèmes aléatoires ont deux caractéristiques principales : ils modélisent un grand nombre de particules sur un graphe, et l'état de chaque particule dépend des états des particules voisines sur le graphe. Ce genre de systèmes présente généralement une transition de phase entre une phase active où l'épidémie infecte une large proportion du réseau et une phase inactive où la contagion reste contenue. Le paramètre de contrôle permettant de changer de phase est la transmissibilité de l'infection, et on s'intéresse généralement à la détermination de sa valeur critique.

De nouveaux modèles mathématiques, visant à capturer les propriétés des systèmes réels organisés en réseaux, ont été développés. Ces modèles ont enrichi la théorie des graphes aléatoires par l'étude mathématique rigoureuse des lois régissant l'évolution de ces systèmes, dans le même esprit que dans les graphes d'Erdős-Rényi. Les réseaux complexes ont une structure inhomogène, découlant non seulement du fait que les caractéristiques structurelles des sommets peuvent s'écarter fortement de la moyenne, mais aussi du fait que les propriétés statistiques peuvent changer entre les différentes parties du réseau. Des phénomènes critiques comme la transition de phase induite par l'apparition d'une composante géante ont été bien étudiés sur les graphes d'Erdős-Rényi. Des progrès récents ont été réalisés sur des graphes aléatoires plus réalistes tels que les graphe aléatoires avec la distribution des degrés donnée (aussi connu comme le modèle de configuration) : e.g. l'émergence de la composante géante, l'existence et les propriétés du k-core, la percolation et les épidémies.

Voici une description rapide du contenu de cette thèse.

Chapitre 1. *Réseaux complexes et graphes aléatoires.* Dans ce chapitre, nous introduisons la théorie des graphes aléatoires et les réseaux complexes. Nous rappelons d'abord quelques

notions et résultats de la théorie des processus de branchement, pertinentes pour prédire les propriétés des graphes aléatoires. Nous donnons ensuite les résultats connus sur l'émergence d'une composante géante, k-core, et les distances dans les graphes d'Erdős-Rényi, et finissons le chapitre par rappeler des résultats connus sur le modéle de configuration.

Chapitre 2. *Percolation de premier passage, l'inondation et le diamètre.* Dans ce chapitre, nous considérons l'impact des poids sur les distances dans les graphes aléatoires dilués. Nous interprétons ces poids comme des retards, et les prenons comme des variables aléatoires exponentielles i.i.d.. Nous analysons le temps d'inondation défini comme le temps minimum nécessaire pour atteindre tous les noeuds à partir d'un noeud choisi d'une manière uniforme, et le diamètre correspondant au pire cas pour le temps d'inondations. Sous certaines conditions de régularité sur la séquence de degrés du graphe aléatoire, nous montrons que ces quantités croissent comme le logarithme de n, lorsque la taille du graphe n tend vers l'infini. Nous trouvons également la valeur exacte des préfacteurs. Ce resultat nous permet d'analyser un algorithme de transmission asynchrone dans les graphes aléatoires réguliers. Nous montrons que la version asynchrone de l'algorithme est plus performante que sa version synchronisée quand la taille du graphe est suffisament grande, il atteindra l'ensemble du réseau plus rapidement, même si le dynamique local est similaire en moyenne. Ce chapitre est fondé sur les articles [4, 7], en collaboration avec M. Draief et M. Lelarge.

Chapitre 3. *Percolation bootstrap, diffusion et cascades.* Dans ce chapitre, nous étudions la diffusion et les cascades dans les graphes aléatoires. Notre modèle de diffusion peut être considéré comme une variante d'un processus de croissance d'un automate cellulaire : supposons que chaque site puisse être dans l'un des deux états possibles, inactif ou actif. Les paramètres du modèle sont deux fonctions données $\theta : \mathbb{N} \to \mathbb{N}$ et $\alpha : \mathbb{N} \to [0, 1]$. Au début du processus, chaque noeud v de degré d_v devient actif avec une probabilité $\alpha(d_v)$ indépendamment des autres sommets. La présence de sommets actifs déclenche un processus de percolation : si un noeud v est actif, il reste actif pour toujours. Et s'il est inactif, il deviendra actif à condition qu'au moins $\theta(d_v)$ de ses voisins soient actifs. Dans le cas où $\alpha(d) = \alpha$ et $\theta(d) = \theta$, pour chaque $d \in \mathbb{N}$, notre modèle de diffusion est équivalent à ce qui est souvent appelé "percolation bootstrap". Notre résultat principal est un théorème qui nous permet de trouver la proportion finale des sommets actifs dans le cas asymptotique, c'est-à-dire lorsque $n \to \infty$. Ce chapitre est fondé sur l'article [1], en collaboration avec M. Draief et M. Lelarge, et l'article [2].

Chapitre 4. *Contagion dans les réseaux financiers.* Dans ce chapitre, nous réalisons une analyse asymptotique des cascades de défaut dans les réseaux financiers. En utilisant des techniques que nous avons développés dans les chapitres précédents,, nous obtenons une expression pour la taille asymptotique d'une cascade de défaut en fonction des caractéristiques du réseau. Ce résultat est utilisé pour obtenir un critère pour la résilience d'un réseau financier aux chocs de petite taille. Nos résultats soulignent le rôle joué par les expositions contagieuses, et montrent notamment que les noeuds qui sont étroitement liés et sur-exposés sont ceux qui contribuent le plus à la fragilité du réseau. Ces résultats asymptotiques concordent avec des simulations faites pour les réseaux dont les tailles sont réalistes, montrant la pertinence de l'étude des réseaux de taille infinie pour la régulation macro-prudentielle. Ce chapitre est fondé sur les articles [5, 6], en collaboration avec R. Cont et A. Minca.

Appendice A. Cette annexe a pour but de fournir un rappel aux principaux outils probabilistes dont nous nous servons tout au long de cette thèse.

Bibliographie personnelle

[1] H. Amini, M. Draief, and M. Lelarge. Marketing in a random network. *Proc. Network Control and Optimization*, LNCS 5425, p.17–25, 2009.

[2] H. Amini. Bootstrap percolation and diffusion in random graphs with given vertex degrees. *Electronic Journal of Combinatorics, 17: R25*, 2010.

[3] H. Amini. Bootstrap percolation in living neural networks. *Journal of Statistcal Physics*, 141:459–475, 2010.

[4] H. Amini, M. Draief, and M. Lelarge. Flooding in weighted random graphs. *Proc. SIAM Analytic Algorithmics and Combinatorics*, 2011.

[5] H. Amini, R. Cont, and A. Minca. Stress testing the resilience of financial networks. *To appear in International Journal of Theoretical and Applied Finance*, 2011.

[6] H. Amini, R. Cont, and A. Minca. Resilience to contagion in financial networks. *In phase of submission.*

[7] H. Amini and M. Lelarge. The diameter of weighted random graphs. *In phase of submission.*

Contents

Introduction

This section is devoted to the summary of the original contributions of this dissertation. We start this by fixing the notations and briefly explaining the random graph model that we will consider through this thesis.

Basic Notations.

Throughout the document, we let $\mathbb{R}$ and $\mathbb{N}$ denote the set of real and natural numbers, respectively. For nonnegative sequences x_n and y_n, we describe their relative order of magnitude using Landau's $o(.)$ and $O(.)$ notation: We write $x_n = O(y_n)$ if there exist $N \in \mathbb{N}$ and $C > 0$ such that $x_n \leq Cy_n$ for all $n \geq N$. Occasionally, we write $x_n = \Omega(y_n)$ to mean that there exists $N \geq 0$ and $C > 0$ such that for all $n \geq N$, $x_n \geq Cy_n$. If $x_n = O(y_n)$ and $x_n = \Omega(y_n)$, then we write $x_n = \Theta(y_n)$. If x_n converges to x as n goes to infinity, then we write $x_n \to x$, as $n \to \infty$. Also $x_n = o(y_n)$ means $x_n/y_n \to 0$, as $n \to \infty$. We write $x_n \sim y_n$ when $x_n/y_n \to 1$ as $n \to \infty$.

We let $\mathbb{P}(A)$ denote the probability of an event A, i.e., a measurable set defined on some probability space. We usually do not make explicit reference to the probability space since it is usually clear to which one we are referring. We say that an event A holds almost surely, and we write a.s., if $\mathbb{P}(A) = 1$. The random variables considered in this document take values in $\mathbb{R}$ or $\mathbb{R}^d$ for some $d \in \mathbb{N}$. The expected value of a real valued random variable X is denoted by $\mathbb{E}X$ or $\mathbb{E}[X]$, its variance by $\mathrm{Var}[X]$. The expected value of X conditional on A is written by $\mathbb{E}[X|A]$. The indicator function of an event A is of particular interest, it is denoted by $\mathbb{1}(A)$ and we have $\mathbb{E}\mathbb{1}(A) = \mathbb{P}(A)$.

We say that a sequence X_n of random variables

(i) converges *in distribution* to a limiting random variable X when

$$\lim_{n\to\infty} \mathbb{P}(X_n \leq x) = \mathbb{P}(X \leq x), \tag{1}$$

for every x for which $F(x) = P(X \leq x)$ is continuous. We write this as $X_n \xrightarrow{d} X$.

(ii) converges *in probability* to a limiting random variable X when, for every $\epsilon > 0$,

$$\lim_{n\to\infty} \mathbb{P}(|X_n - X| > \epsilon) = 0. \tag{2}$$

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

(iii) converges *almost surely* to a limiting random variable X when

$$\mathbb{P}(\lim_{n\to\infty} X_n = X) = 1. \tag{3}$$

We write this as $X_n \overset{a.s.}{\to} X$.

We consider the asymptotic case when $n \to \infty$ and say that an event holds w.h.p. (with high probability) if it holds with probability tending to 1 as $n \to \infty$. Similarly, we use o_p and O_p in a standard way. For example, if (X_n) is a sequence of random variables, then $X_n = O_p(1)$ means that "X_n is bounded in probability" and $X_n = o_p(n)$ means that $X_n/n \overset{p}{\to} 0$. For notational simplicity we will sometimes not show the dependency on n explicitly.

Configuration Model.

For $n \in \mathbb{N}$, let $(d_i)_1^n$ be a sequence of non-negative integers such that $\sum_{i=1}^n d_i$ is even. By means of the configuration model, we define a random multigraph with given degree sequence $(d_i)_1^n$, denoted by $G^*(n, (d_i)_1^n)$ as follows. To each node i, we associate d_i labeled half-edges. All half-edges need to be paired to construct the graph, this is done by randomly matching them. When a half-edge of a node i is paired with a half-edge of a node j, we interpret this as an edge between i and j. The graph $G^*(n, (d_i)_1^n)$ obtained following this procedure may not be simple, i.e., may contain self-loops due to the pairing of two half-edges of i, and multi-edges due to the existence of more than one pairing between two given nodes. However conditional on the multigraph $G^*(n, (d_i)_1^n)$ being a simple graph, we obtain a uniformly distributed random graph with the given degree sequence, which we denote by $G(n, (d_i)_1^n)$.

For $r \in \mathbb{N}$, let $u_k^{(n)} = |\{i,\ d_i = k\}|$ be the number of vertices of degree k and $m^{(n)}$ be the total degree defined by

$$m^{(n)} = \sum_{i=1}^{n} d_i = \sum_{k \geq 0} k u_k^{(n)}.$$

From now on, we assume that the sequence $(d_i)_1^n$ satisfies the following regularity conditions analogous to the ones introduced by Molloy and Reed in [128].

Condition 1. 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) $u_k^{(n)}/n \to p_k$ for every $k \geq 0$ as $n \to \infty$;

(ii) $\lambda := \sum_{k \geq 0} k p_k \in (0, \infty)$;

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

We refer to Chapter 1 for more on this model. Now, we are in position to announce our main results.

First Passage Percolation, Flooding and Diameter.

Let $G = (V, E)$ be a weighted graph, i.e., to each edge $e \in E$, a non-negative weight w_e is assigned. For any $a, b \in V$, a path between a and b is a sequence $\pi = (e_1, e_2, \dots e_k)$ where $e_i = (v_{i-1}, v_i) \in E$ and $v_i \in V$ for $i \in [1, k]$, with $v_0 = a$ and $v_k = b$. We write $e_i \in \pi$ to denote the fact that the edge e_i belongs to the path π. For $a, b \in V$, we define

$$\mathrm{dist}_w(a, b) = \min_{\pi \in \Pi(a,b)} \sum_{e \in \pi} w_e \,,$$

where the minimum is taken over all paths between a and b in the graph. The *weighted diameter* and the *weighted flooding time* are given by

$$\mathrm{diam}_w(G) = \max\{\mathrm{dist}_w(a, b),\ a, b \in V,\ \mathrm{dist}_w(a, b) < \infty\}, \tag{4}$$

$$\mathrm{flood}_w(G) = \max\{\mathrm{dist}_w(a, b),\ b \in V,\ \mathrm{dist}_w(a, b) < \infty\}, \tag{5}$$

where in (5) a is chosen uniformly at random in V.

Our main result (in Chapter 2) consist in precise asymptotic expressions for the weighted diameter and weighted flooding time of sparse random graphs on n vertices with i.i.d. exponentially distributed weights with parameter 1. To explain this, we need to introduce some extra notations.

We now consider a random graph $G(n,(d_i)_1^n)$, where the degree sequence $(d_i)_1^n$ satisfies Condition 1. Let us define $q=\{q_k\}_{k=0}^{\infty}$ the size-biased probability mass function corresponding to p, by

$$q_k=\frac{(k+1)p_{k+1}}{\lambda}\,, \tag{6}$$

and define ν to be its mean, i.e.,

$$\nu=\sum_{k=0}^{\infty}k\,q_k\ \in\big(0\,,\infty\big).$$

We further assume the following additional condition:

Condition 2. (iv) as $n\to\infty$, $\nu_n:=\frac{\sum_{i=1}^n d_i^{(n)}(d_i^{(n)}-1)}{m^{(n)}}\to\nu\in(1,\infty)$;

(v) for some $\tau>0$, $\Delta_n:=\max_{i\in V}d_i=O(n^{1/2-\tau})$;

(vi) $d_{\min}:=\min\{\,k\mid p_k>0\,\}$ is such that for $k<d_{\min}$; $u_k^{(n)}:=|\{i,\ d_i=k\}|=0$, for all n sufficiently large.

The condition $\nu>1$ is equivalent to the existence of a giant component in the configuration model, the size of which is proportional to n (see e.g., [110, 128]). We will assume this in the rest of this section, so $\nu>1$.

Let $\mathcal{X}_q$ be a Galton-Watson Tree (GWT) with offspring distribution q. Recall that the extinction probability of the branching process, that we denote by β, is the smallest solution in $[0,1]$ of the fixed point equation

$$\beta=G_q(\beta)\,, \tag{7}$$

where G_q is the generating function of the distribution q.

In addition, we introduce

$$\beta_*=G_q'(\beta)=\sum_{k=1}^{\infty}kq_k\beta^{k-1}\,. \tag{8}$$

Our main theorem in Chapter 2 is the following.

Theorem 1 (Theorem 2.2)**.** *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* Conditions 1 *and* 2*. Then we have*

$$\frac{\mathrm{diam}_w(G(n,(d_i)_1^n))}{\log n} \xrightarrow{p} \frac{1}{\nu-1}+\frac{2}{d_{\min}}\mathbb{1}_{(d_{\min}\geq 3)}+\frac{1}{(1-q_1)}\mathbb{1}_{(d_{\min}=2)}+\frac{2}{1-\beta_*}\mathbb{1}_{(d_{\min}=1)},$$

and

$$\frac{\mathrm{flood}_w(G(n,(d_i)_1^n))}{\log n} \xrightarrow{p} \frac{1}{\nu-1}+\frac{1}{d_{\min}}\mathbb{1}_{(d_{\min}\geq 3)}+\frac{1}{2(1-q_1)}\mathbb{1}_{(d_{\min}=2)}+\frac{1}{1-\beta_*}\mathbb{1}_{(d_{\min}=1)}.$$

In the particular case where G is a random r-regular graph with $r \geq 3$, we recover a result first proved in [52] concerning the weighted diameter.

We now consider the Erdős-Rényi random graph, $\mathrm{ER}(n,\lambda/n)$, on n vertices where there is an edge between two vertices with probability λ/n independently of everything else. We let $\lambda > 1$ to have a giant component. Then Conditions 1 and 2 hold a.s. (by conditioning on the vertex degrees, c.f. see Remark 1.23). In this case we have

$$q_k = e^{-\lambda}\frac{\lambda^k}{k!}. \tag{9}$$

And then $\nu = \lambda$ and $\beta_* = \lambda_*$ where λ_* is the solution $\lambda_* < 1$ to

$$\lambda_* e^{-\lambda_*} = \lambda e^{-\lambda}. \tag{10}$$

Applying Theorem 1 to this case we have :

Theorem 2 (Theorem 2.5)**.** *Let $\lambda > 1$ and assume $\mathrm{ER}(n,\lambda/n)$ with i.i.d. rate* 1 *exponential weights on its edges. Then*

$$\frac{\mathrm{diam}_w(\mathrm{ER}(n,\lambda/n))}{\log n} \xrightarrow{p} \frac{1}{\lambda-1}+\frac{2}{1-\lambda_*}, \quad \textit{and} \tag{11}$$

$$\frac{\mathrm{flood}_w(\mathrm{ER}(n,\lambda/n))}{\log n} \xrightarrow{p} \frac{1}{\lambda-1}+\frac{1}{1-\lambda_*}. \tag{12}$$

A lower bound for the weighted diameter in this case was given by Bhamidi, van der Hofstad and Hooghiemstra in [25]. In particular, the above theorem improves this, and gives the correct asymptotic.

Broadcasting in random regular graphs. Theorem 1 allows us to analyze an asynchronous randomized broadcast algorithm for random regular graphs. In continuous-time, we assume that each node is endowed with a Poisson process with rate 1 and that at the instants of its corresponding Poisson process a node wakes up and contacts one of its neighbors uniformly at random. In Section 2.3, we consider the well-studied push model. In this model, if a node i holds the message, it passes the message to its randomly chosen neighbor regardless of its state. As in the case of the standard discrete-time phone call model, we are interested in the performance of such an information dissemination routine in terms of the time it takes to inform the whole population. We denote this time by $\mathrm{ABT}(G)$ for asynchronous broadcast time

Our results show that the asynchronous version of the algorithm performs better than its synchronized version: in the large size limit of the graph, it will reach the whole network faster even if the local dynamics are similar on average.

Theorem 3 (Corollary 2.6)**.** *Let $G \sim \mathcal{G}(n,r)$ be a random r-regular graph with n vertices. Then w.h.p.*

$$ABT(G) = 2\left(\frac{r-1}{r-2}\right)\log n + o(\log n).$$

The classical randomized broadcast model was first investigated by Frieze and Grimmett [76]. Given a graph $G = (V, E)$, initially a piece of information is placed on one of the nodes in V. Then in each time step, every informed node sends the information to another node, chosen independently and uniformly at random among its neighbors. The question now is how many time-steps are needed such that all nodes become informed. Note that this model requires nodes to be synchronized.

Fountoulakis and Panagtotou in [74] have recently shown that in the case of random regular graphs, the process completes in $\left(\frac{1}{\log(2(1-1/r))} - \frac{1}{r\log(1-1/r)}\right)\log n + o(\log n)$ rounds w.h.p.

Figure 1 shows the comparison between results in [74] and our Theorem 3: in both cases, the time to broadcast is of the order of $\log n$ but the prefactors differ and are given by the two curves for various values of r. We see that the asynchronous version is always faster than the synchronized one. We refer to Section 2.3 for more discussions.

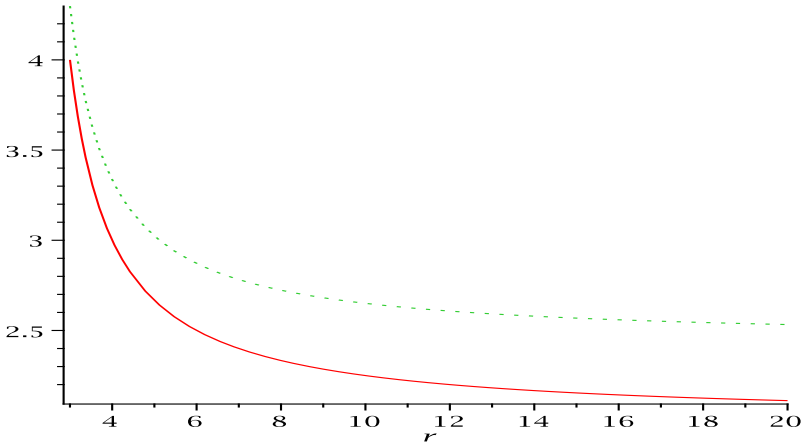

Figure 1: Comparison of the time to broadcast in the synchronized version (dashed) and with exponential random weights (plain)

Bootstrap Percolation, Diffusion and Cascades.

The diffusion model we consider in this section is a generalization of bootstrap percolation in an arbitrary graph (modeling a given network). Let $G = (V, E)$ be a connected graph. The threshold associated to a node i is $\theta(d_i)$, where d_i is the degree of i and $\theta : \mathbb{N} \to \mathbb{N}$ is a given fixed function. Assume that each node can be in one of two possible states: inactive or active. Let $\alpha : \mathbb{N} \to [0, 1]$ be a fixed given function. At time 0, each node i becomes active with probability $\alpha(d_i)$ independently of all the other vertices. At time $t \in \mathbb{N}$, the state of each node i will be updated according to a deterministic process: if a node i was active at time $t - 1$, it remains active at time t. Otherwise, i will become active if at least $\theta(d_i)$ of its neighbors were active at time $t - 1$.

In the case where $\alpha(d) = \alpha$ and $\theta(d) = \theta$, for each $d \in \mathbb{N}$, our diffusion model is equivalent to what is called *bootstrap percolation*. This model has a rich history in statistical physics, mostly in $G = \mathbb{Z}^d$ and finite boxes [4]

Let G be a graph with n nodes, i.e., $|V| = n$. Let A denote the adjacency matrix of G, with $A_{ij} = 1$ if $i \sim j$, and $A_{ij} = 0$ otherwise. The state of the network at time t can be described by the vector $(X_t(i))_{i=1}^n$: $X_t(i) = 1$ if the node i is active at time t and $X_t(i) = 0$ otherwise. Note that $X_0(i)$ is a Bernoulli random variable with parameter $\alpha(d_i)$. The evolution of this vector at time $t + 1$ follows the following functional equation, i.e., at each time step $t + 1$, each node v

applies:

$$X_{t+1}(i) \quad = \quad X_t(i) + (1 - X_t(i))\mathbb{1}\left(\sum_j A_{ij}X_t(j) \geq \theta(d_i)\right). \tag{13}$$

From the definition, $X_t(i)$ is non-decreasing. We let

$$\Phi^{(n)}(\alpha, \theta, t) := n^{-1}\sum_{j=1}^{n}\mathbb{E}[X_t(j)],$$

and $\Phi^{(n)}(\alpha, \theta) := \lim_{t\to\infty}\Phi^{(n)}(\alpha, \theta, t)$.

We now present our main result in Chapter 3. Consider a random graph $G(n, (d_i)_1^n)$ where the degree sequence $(d_i)_1^n$ satisfies Condition 1. Let D be a random variable with integer values and with distribution $\mathbb{P}(D = r) = p_r$, $r \in \mathbb{N}$. We define the function $f_{\alpha,\theta} : [0, 1] \to \mathbb{R}$ as follows

$$f_{\alpha,\theta}(y) := \lambda y^2 - y\,\mathbb{E}\left[\,\big(1 - \alpha(D)\big)\,D\,\mathbb{1}\left(\,\mathrm{Bin}(D - 1, 1 - y) < \theta(D)\right)\right], \tag{14}$$

where $\mathrm{Bin}(l, p)$ denotes a binomial variable with parameters l and p

$$\mathbb{P}\left(\mathrm{Bin}(l, p) = r\right) = \binom{l}{r}p^r(1 - p)^{l-r}.$$

Let $y^* = y^*_{\alpha,\theta}$ be the largest solution to $f_{\alpha,\theta}(y) = 0$, i.e.,

$$y^* := \sup\{\, y \in [0, 1] \mid f_{\alpha,\theta}(y) = 0\,\}.$$

Theorem 4 (Theorem 3.2)**.** *Consider a random graph $G(n, (d_i)_1^n)$ where the degree sequence $(d_i)_1^n$ satisfies* Condition 1. *Then we have with high probability (w.h.p.)*

1. *If $y^* = 0$, i.e., if $f_{\alpha,\theta}(y) > 0$ for all $y \in (0, 1]$, then w.h.p. $\Phi^{(n)}(\alpha, \theta) = 1 - o(1)$.*
2. *If $y^* > 0$ and furthermore y^* is not a local minimum point of $f_{\alpha,\theta}(y)$, then w.h.p.*

$$\Phi^{(n)}(\alpha, \theta) = 1 - \mathbb{E}\left[\,(1 - \alpha(D))\,\mathbb{1}\left(\,\mathrm{Bin}(D, 1 - y^*) \;<\; \theta(D)\right)\right] + o(1).$$

We now look at the diffusion with one initial active node. Let us call the following condition *the cascade condition*:

$$\mathbb{E}\left[\,D\,\right] \;< \mathbb{E}\left[\,D(D - 1)\mathbb{1}_{(\theta(D)=1)}\,\right].$$

The second theorem of Chapter 3 is the following.

Theorem 5 (Theorem 3.3)**.** *Consider a random graph $G(n,(d_i)_1^n)$ where the degree sequence $(d_i)_1^n$ satisfies Condition 1.*

- *If the cascade condition (defined above) is satisfied, then there exists w.h.p. a single node v which can trigger a global cascade, i.e., v can activate a strictly positive fraction of the total population w.h.p.*
- *If $\mathbb{E}[D] > \mathbb{E}\left[D(D-1)\mathbb{1}_{(\theta(D)=1)}\right]$, then for any $k = o(n)$, we have*

$$|\mathcal{C}(1,2,...,k)| = o_p(n),$$

where for $W \subseteq V$, $\mathcal{C}(W)$ denote the final set of active nodes when we start the diffusion with initial active nodes W.

We note that in the case where $\theta(d) = \theta d$, Watts [151] obtained the same condition by a heuristic argument validated through simulations. Our theorem provides as a very special case a mathematical proof of his heuristic results. We refer to Chapter 3 for more discussions.

Contagion in Financial Networks.

In Chapter 4, we perform an asymptotic analysis of default cascades in financial networks. Using analytical methods, we derive an expression for the fraction of defaulted nodes in the limit where the number of nodes is large, in terms of the empirical distribution of the in- and out-degrees and the proportion of contagious links in a financial network. This result is used to obtain a criterion for the resilience of a large network to macro-economic shocks. Given a macroeconomic stress scenario defined in terms of the magnitude of common shocks across balance sheet, our criterion yields a minimal capital ratio which guarantees stability of the system in the given stress scenario.

Interlinkages across balance sheets of financial institutions may be modeled by a weighted directed graph $\mathbf{g} = (\mathbf{v}, \mathbf{e})$ on the vertex set $\mathbf{v} = [1, \ldots, n]$, whose elements represent financial institutions. Denoting by $e(i,j)$ the exposure (in monetary units) of institution i to institution j, the interbank assets of i are given by $A_i = \sum_j e(i,j)$, whereas $L_i = \sum_j e(j,i)$ represents the interbank liabilities of i. In addition to these interbank assets and liabilities, a bank may hold other assets and liabilities (such as deposits). The net worth of the bank, given by its **capital**

c_i, represents its capacity for absorbing losses before it becomes insolvent. We define the ratio γ_i as

$$c_i = \gamma_i A_i.$$

We will refer to γ_i as "capital ratio" although technically it is the ratio of capital to interbank assets and not total assets. An institution is *insolvent* if its net worth is negative or zero, in which case we set $\gamma_i = 0$. The number of an institution's creditors is called its *in-degree*

$$d^-(i) = \#\{j \in \mathbf{v} \mid e(j,i) > 0\},$$

while the out-degree of a node i is the number of its debtors

$$d^+(i) = \#\{j \in \mathbf{v} \mid e(i,j) > 0\}.$$

In a network $(\mathbf{e}, \gamma)$, the set of initially insolvent institutions is represented by

$$\mathbb{D}_0(e,\gamma) = \{i \in \mathbf{v} \mid \gamma_i = 0\}.$$

If we denote by R_j the recovery rate for the debt of a market participant j, then j's default induces a loss equal to $(1-R_j)e(i,j)$ to its counterparty i. If this loss is greater than i's capital, then i defaults. The set of nodes which become insolvent due to their exposures to defaults is

$$\mathbb{D}_1(e,\gamma) = \{i \in \mathbf{v} \mid \gamma_i A_i < \sum_{j \in \mathbb{D}_0} (1-R_j)e(i,j)\},$$

and generally $\mathbb{D}_r$ represents the set of nodes defaulting in round r due to exposures to nodes defaulted in rounds $0, \ldots, r-1$.

It is easy to see that the process finishes at most after $n-1$ time steps if the network is of size n, and gives the increasing sequence of default sets

$$\mathbb{D}_0(e,\gamma) \subseteq \mathbb{D}_1(e,\gamma) \subseteq \cdots \subseteq \mathbb{D}_{n-1}(e,\gamma).$$

We assume in what follows that the recovery rate is constant over all nodes and equal to R. The final fraction of defaults at the end of the cascade process, denoted by $\alpha_n(\mathbf{e},\gamma)$, is a deterministic function of the exposure matrix and sequence of capital ratios:

$$\alpha_n(\mathbf{e},\gamma) = \frac{|\mathbb{D}_{n-1}(e,\gamma)|}{n}.$$

The financial network $(\mathbf{e}, \gamma)$ is embedded in a sequence of financial networks, indexed by their size $(\mathbf{e_n}, \gamma_\mathbf{n})$. The sequences of in- and out-degrees in these networks, also indexed by n, are denoted $\mathbf{d}_\mathbf{n}^+ = \{d_n^+(i)\}_{i=1}^n$ and respectively $\mathbf{d}_\mathbf{n}^- = \{d_n^-(i)\}_{i=1}^n$. Their empirical distribution is given by

$$\mu_n(j,k) := \frac{1}{n}\#\{i : d_n^+(i) = j, d_n^-(i) = k\},$$

and the total number of links in the network of size n by

$$m_n := \sum_i d_n^+(i) = \sum_i d_n^-(i).$$

We assume that the degree sequence $\mathbf{d}_\mathbf{n}^+ = \{d_n^+(i)\}_{i=1}^n$ and $\mathbf{d}_\mathbf{n}^- = \{d_n^-(i)\}_{i=1}^n$ are sequences of nonnegative integers satisfying the following regularity conditions analogues to Condition 1.

Condition 3. There exists a probability distribution μ on $\mathbb{N}^2$ such that:

1. The empirical proportion of nodes of degree (j,k) tends to $\mu(j,k)$:

$$\mu_n(j,k) \to \mu(j,k) \text{ as } n \to \infty;$$

2. Finite average degree property:

$$\exists \lambda \in (0,\infty), \qquad \sum_{j,k} j\mu(j,k) = \sum_{j,k} k\mu(j,k) =: \lambda;$$

3. $\sum_{i=1}^n d_n^+(i) = \sum_{i=1}^n d_n^-(i)$;

4. $\sum_{i=1}^n (d_n^+(i))^2 + (d_n^-(i))^2 = O(n)$.

The sequences of continuous exposures and capital ratios are mapped into discrete sequences representing the default threshold for each node. We denote by $\Sigma_i^\mathbf{e}$ the set of permutations of node i's debtors in a network $\mathbf{e}$. For each node i and permutation $\tau \in \Sigma_i^\mathbf{e}$, we define

$$\Theta(i,\tau,\mathbf{e}) := \min\{k \geq 0, \gamma_i \sum_{j=1}^{d^+(i)} e(i,j) < \sum_{j=1}^{k}(1-R)e(i,\tau(j))\}, \tag{15}$$

which represents the threshold function: conditional on the order τ in which the counterparties of i may default, this function determines how many defaults i's capital buffer can withstand before i defaults.

Let us define

$$p_n(j,k,\theta) := \frac{\#\{(i,\tau) \mid 1 \le i \le n,\ \tau \in \Sigma_i^{\mathbf{e_n}},\ d_n^+(i) = j,\ d_n^-(i) = k,\ \Theta(i,\tau,\mathbf{e_n}) = \theta\}}{n\mu_n(j,k)j!}.$$

We further assume the following additional condition:

Condition 4. There exists a function $p : \mathbb{N}^3 \to [0,1]$ such that for all $j,k,\theta \in \mathbb{N}$ $(\theta \le j)$

$$p_n(j,k,\theta) \overset{n\to\infty}{\to} p(j,k,\theta).$$

We say that a link is *contagious*, if it represents an exposure of a node larger than its capital. It is easy to see that $p_n(j,k,1)$ represents the proportion of 'contagious' links leaving nodes with degree (j,k). The limit $p(j,k,1)$ also represents the fraction of nodes with degree (j,k) that default when one counterparty defaults.

The financial network is modeled under incomplete information. Disclosure of counterparty identity is not required in our framework. This allows for valuable confidentiality. On the other hand, we require important information, as the exact composition of balance sheets: the size of all exposures and the connectivity of each node which determine the crucial characteristics of the network and its response to external shocks.

Let $\mathbb{G}_n(\mathbf{e_n}, \mathbf{d_n^+}, \mathbf{d_n^-})$ be the set of all weighted directed graphs with degree sequence $\mathbf{d_n^+}, \mathbf{d_n^-}$ such that, for any node i, the set of exposures is given by the non-zero elements of line i in the exposure matrix $\mathbf{e_n}$. On a probability space $(\Omega, \mathcal{A}, \mathbb{P})$, we define $\mathbf{E_n}$ as a random network uniformly distributed on $\mathbb{G}_n(\mathbf{e_n}, \mathbf{d_n^+}, \mathbf{d_n^-})$.

We endow the nodes in $\mathbf{E_n}$ with the capital ratios $\gamma_\mathbf{n}$. Then:

$$\forall i = 1 \dots n, \quad \{\mathbf{E_n}(i,j), \quad \mathbf{E_n}(i,j) > 0\} = \{\mathbf{e_n}(i,j), \quad \mathbf{e_n}(i,j) > 0\}, \quad (\mathbb{P} - a.s.)$$

$$\#\{j \in \mathbf{v},\ \mathbf{E_n}(j,i) > 0\} = d_n^+(j), \quad \text{and} \qquad \#\{j \in \mathbf{v},\ \mathbf{E_n}(i,j) > 0\} = d_n^-(i).$$

The quantity $\alpha_n(\mathbf{E_n}, \gamma_\mathbf{n})$ represents the size of the cascade generated by the default of

$$\mathbb{D}_0(\mathbf{E_n}, \gamma_\mathbf{n}) = \{i, \gamma_n(i) = 0\},$$

i.e., fraction of defaults in the network triggered by the default of nodes in $\mathbb{D}_0(\mathbf{E_n}, \gamma_\mathbf{n})$. The following theorems give the asymptotic behavior of this quantity.

Theorem 6 (Theorem 4.8)**.** *Define the function*

$$I(\pi) := \sum_{j,k} \frac{k\mu(j,k)}{\lambda} \sum_{\theta=0}^{j} p(j,k,\theta)\mathbb{P}(\mathrm{Bin}(j,\pi) \geq \theta), \tag{16}$$

where $\mathrm{Bin}(j,\pi)$ *denotes a binomial variable with parameters* j *and* π*.*

Consider a sequence of exposure matrices and capital ratios $\{(\mathbf{e_n})_{\mathbf{n}\geq\mathbf{1}}, (\gamma_\mathbf{n})_{\mathbf{n}\geq\mathbf{1}}\}$ *satisfying Conditions 3 and 4 and the corresponding sequence of random matrices* $(\mathbf{E_n})_{n\geq 1}$ *defined on* $(\Omega, \mathcal{A}, \mathbb{P})$ *as above. Let* π^* *be the smallest fixed point of* I *in* $[0,1]$*, i.e.*

$$\pi^* = \inf\{\pi \in [0,1] \mid I(\pi) = \pi\}.$$

1. *If* $\pi^* = 1$*, i.e. if* $I(\pi) > \pi$ *for all* $\pi \in [0,1)$*, then asymptotically all nodes default during the cascades*

$$\alpha_n(\mathbf{E_n}, \gamma_\mathbf{n}) \xrightarrow{p} 1.$$

2. *If* $\pi^* < 1$ *and furthermore* π^* *is a stable fixed point of* I *(*$I'(\pi^*) < 1$*), then the asymptotic fraction of defaults*

$$\alpha_n(\mathbf{E_n}, \gamma_\mathbf{n}) \xrightarrow{p} \sum_{j,k} \mu(j,k) \sum_{\theta=0}^{j} p(j,k,\theta)\mathbb{P}(\mathrm{Bin}(j,\pi^*) \geq \theta).$$

We define as *the resilience measure* the following function of the network's features, which takes values in $(-\infty, 1]$:

$$1 - \sum_{j,k} \frac{jk}{\lambda} \mu(j,k)p(j,k,1).$$

The second theorem of Chapter 4 is the following.

Theorem 7 (Corollary 4.10 - Theorem 4.11)**.** *Under the assumptions of Theorem 6:*

- *If the resilience measure is positive, i.e.*

$$1 - \sum_{j,k} \frac{jk}{\lambda} \mu(j,k)p(j,k,1) > 0, \tag{17}$$

then for every $\epsilon > 0$*, there exists* N_ϵ *and* ρ_ϵ *such that if the initial fraction of defaults is smaller than* ρ_ϵ*, then* $\mathbb{P}(\alpha_n(\mathbf{E_n}, \gamma_\mathbf{n}) \leq \epsilon) > 1 - \epsilon$ *for all* $n \geq N_\epsilon$*.*

- *If the resilience measure is negative, i.e.*

$$1 - \sum_{j,k} \frac{jk}{\lambda} \mu(j,k) p(j,k,1) < 0, \tag{18}$$

then there exists a connected set C_n of nodes representing a positive fraction of the financial system, i.e. $|C_n|/n \xrightarrow{p} c > 0$ such that, with high probability, any node belonging to this set can trigger the default of all nodes in the set: for any sequence $(\gamma_n)_{n\geq 1}$ such that $\{i, \gamma_n(i) = 0\} \cap C_n \neq \emptyset$,

$$\liminf_n \alpha_n(E_n, \gamma_n) \geq c > 0.$$

We illustrate then, in Chapter 4, how the resilience criterion may be used in a stress test framework, in which a macroeconomic shock is applied across balance sheets. As the magnitude of the shock increases, so does the proportion of contagious links, up to a critical threshold in which the resilience criterion is no longer satisfied and a *phase transition* ensues: a global cascade of defaults may then be triggered by the default of a single bank.

Chapter 1

Random Graphs and Complex Networks

Abstract. *In this chapter, we introduce the theory of random graphs and complex networks. We first recall some basic notations and results in the theory of branching processes, relevant for heuristic analysis and predicting the properties of random graphs. We give then the known results on the emergence of a giant component, k-core, and distances in the Erdős-Rényi random graphs, and end the chapter by a review of known results on the configuration model.*

Contents

1.1 Introduction

Pervasiveness of networks in all aspects of modern life and the associated problems, from viral attacks on the internet, marketing in social networks or government attempt to contain the default contagion in the financial market during the onset financial crisis, have spurred new research to help understand and predict the behavior of such irregular, complex and evolving structures of interrelated items. Empowered by computational capacity and the availability of large data-sets on real networks (like the internet, the World Wide Web, transportation networks, phone call networks, scientific coauthorship networks, protein-protein biological networks, etc.), the study of complex networks has emerged as a fast growing field, since the appearance at the end of the twentieth century of two seminal papers, the first one by Watts and Storogatz [152] on the small world model and the second one by Barabasi and Albert [17] on scale-free networks. What has been entitled as a new science of networks has seen much of the research carried in physics but also social sciences, finance and economics, computer science, telecommunications, natural sciences and mathematics.

Most real networks are very large, from several thousands of nodes in a network of firms to several billions of nodes in certain biological systems. Due to their complexity, their interrelations can only be described in statistical terms, and one of the most striking observations is that, networks different in functions such as social networks and biological networks share common features. The most striking one concerns the degree distribution, the degree of a node being the number of incident links. The degree distribution p_k, defined as the probability that a node chosen uniformly at random has degree k, exhibits in most such networks strong heterogeneity and has a power law tail, i.e., there exists a parameter γ such that $p_k \sim k^{-\gamma}$ for k larger than a given constant (e.g., see [2, 7, 132, 140]). This is known as the *scale-free* property and differentiates these networks from the classical Erdős and Rényi graphs, in which the distribution is Poisson and thus homogeneous. A related property is the *small world* property: the typical distance between any two nodes is small (one can quote the famous experiment of Milgram that asserted that the typical chain of acquaintances needed to link two people in the US is typically

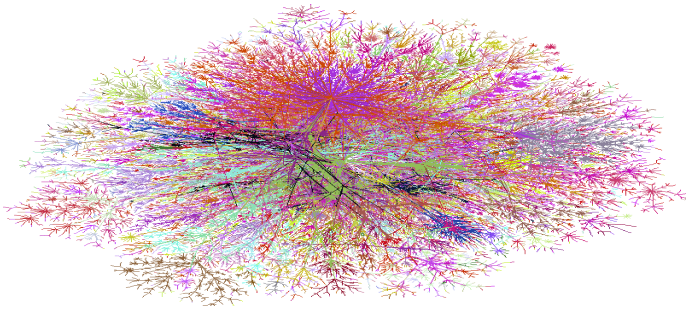

Figure 1.1: Internet topology (taken from http://www.cheswick.com/ches/map/)

equal six). The sparsity of links (they grow only linearly with the number of vertices) and clustering effect are other properties that characterize complex networks.

New mathematical models, aiming at capturing the properties of real systems organized as networks have been developed, built on the observed key characteristics. These models have enriched the theory of random graphs by the rigorous mathematical study of the laws governing the evolution of such systems, in the same spirit as in the Erdős-Rényi graph. One crucial area of study are critical phenomena, see e.g. [5, 28, 53]. Complex networks exhibit strong inhomogeneity in their structure stemming not only from the fact that structural characteristics of vertices may strongly deviate from the mean but also from the fact that statistical properties may change between different parts of the network. Critical phenomena like the phase transition induced by the emergence of a giant component have been well studied on Erdős-Rényi graphs. Recent progress has been made on corresponding mathematical studies on more realistic random graphs such as the random graph with fixed degree distribution (also known as the configuration model): the emergence of the giant component, the existence and properties of the k-core, percolation and epidemics.

The remainder of this chapter is organized as follows: we recall some important properties

of branching processes, relevant for heuristic analysis and predicting the properties of random graphs. We give then the known results on the emergence of a giant component and distances in the Erdős-Rényi graph, and end the chapter by a review of known results on the configuration model.

1.2 Galton-Watson Branching Processes

Many results in the theory of random graphs can be found heuristically by introducing a branching process that mimics the growth of the cluster. In this section, we recall some basic notations and results in the theory of branching processes, that we will use as an important tool throughout the thesis. For more information and proofs of the results, we refer to the books by Athreya and Ney [13], Harris [87], and Jagers [101].

The branching process model (also known as the Galton-Watson model) was introduced by Sir Francis Galton in 1873 to represent the genealogical descendants of individuals. This model is the simplest one describing the evolution in time of a population. Each individual independently gives birth to a random and identically distributed number of children. We denote the offspring distribution by $\{p_i\}_{i=0}^{\infty}$, where

$$p_i = \mathbb{P}\left(\text{the individual has } i \text{ children}\right). \tag{1.1}$$

Starting from one individual at generation 0 and denoting by Z_n the number of individuals at generation n, one easily see that

$$Z_{n+1} = \sum_{i=1}^{Z_n} \xi_{n,i}, \tag{1.2}$$

where $\xi_{n,i}$ is the number of children of the i-th individual of the n-th generation. By assumption $\{\xi_{n,i}\}_{n,i\in\mathbb{N}}$ is a doubly infinite array of i.i.d. random variables, distributed according to $\{p_k\}$, that is, for all $k \geq 0$, $\mathbb{P}(\xi = k) = p_k$. The reproduction generating function used by Watson is simply defined by

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

for $s \in [0,1]$. This function is increasing, convex, and concentrates all the information of the distribution of ξ. In particular, the expected number of children is

$$\lambda := \mathbb{E}[\xi] = \sum_{k\geq 0} k p_k = \phi'_\xi(1).$$

1.2.1 Extinction probability and limit theorems

One of the basic results of branching processes is that when $\lambda \leq 1$, the population dies out with probability one (unless $p_1 = 1$), while if $\lambda > 1$, there is a non-zero probability that the population will not become extinct. Let us denote by p_{ext} the extinction probability, i.e., the probability that after some finite n, $Z_n = 0$. The case where $p_1 = 1$ is uninteresting, so we always assume $p_1 < 1$.

Theorem 1.1 (Survival vs. extinction)**.** *For a branching process with i.i.d. offspring ξ, the extinction probability p_{ext} is the smallest solution in $[0, 1]$ of*

$$s = \phi_\xi(s). \tag{1.3}$$

In particular, the following regimes can happen:

(i) Subcritical regime: If $\mathbb{E}[\xi] < 1$, then $p_{ext} = 1$.

(ii) Critical regime: If $\mathbb{E}[\xi] = 1$ (and $p_1 < 1$), then $p_{ext} = 1$.

(iii) Supercritical regime: If $\mathbb{E}[\xi] > 1$, then $p_{ext} < 1$.

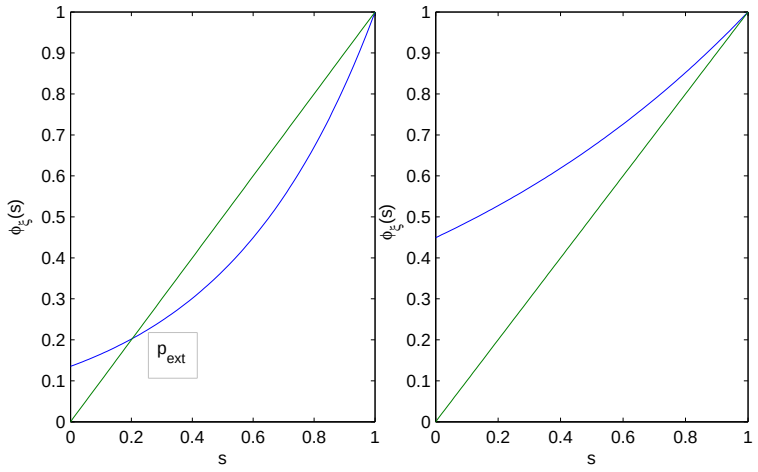

Figure 1.2: The generating functions for a supercritical (left) and subcritical (right) Galton-Watson processes are shown. In the subcritical case, 1 is the only root of $s = \phi_\xi(s)$ in $[0, 1]$. In the supercritical case, there is $p_{ext} \in [0, 1)$ such that $s = \phi_\xi(s)$.

One important result that will motivate the remainder of this section concerns the asymptotic size of the population:

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

If the process does not become extinct, then with probability 1 the size of the population grows to infinity. We investigate in what follows the nature of this divergence. It is easy to see, by a simple conditioning argument, that $\mathbb{E}Z_n = \lambda^n$. Intuitively, Z_n behaves like λ^n for large n. Doob's limit law [87, p. 13] characterizes more precisely the behavior of Z_n.

Theorem 1.3. *Let* λ *be finite. Then* $W_n = Z_n/\lambda^n$ *represents a martingale with* $\mathbb{E}W_n = 1$, *and* $W_n \to W$ *almost surely, as* $n \to \infty$, *where* W *is a nonnegative random variable.*

The distribution of W is not known. However, one can obtain accurate information on W, and the process behaves exactly as one expects (namely $\mathbb{E}W = 1$, $\mathbb{P}(W = 0) = p_{ext}$) if and only if Z's moment of order $x \log x$ is finite. This is made precise in the following theorem, due to Kesten and Stigum [118], which gives the asymptotic properties of supercritical Galton-Watson processes.

Theorem 1.4. *Let* Z_n *be the number of individuals in the* n*-th generation of a supercritical Galton-Watson process with progeny distribution* ξ. *The following statements are equivalent:*

1. $\mathbb{E}[\xi \log \xi] < \infty$;
2. $\lim_{n\to\infty} |W_n - W| = 0$;
3. $\mathbb{E}W = 1$;
4. $\mathbb{P}(W = 0) = p_{ext}$.

We have so far described the growth of the process when it does not die out. The next result treats the case of a supercritical branching process that goes extinct.

For $\mu > 0$, let $\mathrm{Poi}(\mu)$ denote a Poisson random variable with mean μ, i.e.,

$$\mathbb{P}(\mathrm{Poi}(\mu) = k) = e^{-\mu}\mu^k/k!. \tag{1.4}$$

Theorem 1.5 (Duality principle)**.** *A supercritical branching process conditioned to become extinct is a subcritical branching process. If the original offspring distribution is* $\mathrm{Poi}(\lambda)$ *with* $\lambda > 1$,

then the conditional one is $\mathrm{Poi}(\lambda_*)$*, where* $\lambda_* = \lambda\, p_{ext}$ *can be characterized as the solution* $\lambda_* < 1$ *to*

$$\lambda_* e^{-\lambda_*} = \lambda e^{-\lambda}.$$

One can give the following geometrical interpretation: a supercritical branching process conditioned to die out is a branching process with a generating function that is obtained by taking the graph of ϕ_ξ over $[0, p_{ext}]$ and rescaling it in the domain and range $[0, 1]$. In the next result, we take the graph of ϕ_ξ over $[p_{ext}, 1]$ and rescale to make the domain and range $[0, 1]$.

Theorem 1.6. *Consider a supercritical branching process with generating function* ϕ_ξ*. If we condition on nonextinction and look only at the individuals that have an infinite line of descent, then* $\hat{Z}_n$*, the number of individuals in generation* n*, is a branching process with offspring distribution* $\hat{\xi}$*, where*

$$\phi_{\hat{\xi}}(s) = \frac{\phi_\xi(s + (1 - p_{ext})s)}{1 - p_{ext}}.$$

1.2.2 Continuous-time Markov branching processes

In the discrete Galton-Watson process described in the previous section, each particle dies out after one unit of time. A natural generalization is to allow these lifetimes to be random variables. We consider a process $\{Z(t); t \geq 0\}$, where $Z(t)$ is the number of particles at time t. This process will in general not be Markovian unless the lifetimes are independent, exponentially distributed random variables.

Definition 1.7. On a probability space $(\Omega, \mathbb{F}, \mathbb{P})$, a stochastic process $\{Z(t, \omega); t \geq 0\}$ is called a *continuous time Markov branching process* if

(i) its state space is the set of non-negative integers;

(ii) it is a stationary Markov chain with respect to the fields $\mathbb{F}_t = \sigma\{Z(s, \omega); s \leq t\}$;

(iii) the transition probabilities $P_{ij}(t) = \mathbb{P}(Z_t = j | Z_0 = i)$ satisfy

$$\sum_{j=0}^{\infty} P_{ij}(t) s^j = \left[\sum_{j=0}^{\infty} P_{1j}(t) s^j \right]^i, \tag{1.5}$$

for all $i \geq 0$ and $|s| \leq 1$.

Properties (i) and (ii) say that $Z(t)$ is a continuous time Markov process on the integers, while (iii) characterizes the basic branching property.

We now consider the following Markov branching process. As before, a single ancestor is born at the origin at time 0. If a given particle is alive at a certain time, its additional life length is a random variable which is exponentially distributed with parameter 1. Upon death, it leaves k offsprings with probability $p_k, k = 0, 1, 2, \ldots$. As usual, each particle acts independently of other particles and the history of the process. The transition probabilities are determined as solutions of the Kolmogorov forward and backward equations (see [13], page 103). Let $\phi(s) := \sum_{k\geq 0} p_k s^k$. Then we have the following basic result (analogue to Theorem 1.1).

Theorem 1.8. *The extinction probability p_{ext} is the smallest root in $[0, 1]$ of the equation $\phi(s) = s$.*

The asymptotic behavior of $Z(t)$ for large t is very similar to that of the discrete time Galton-Watson process. It is easy to see that $\mathbb{E}Z(t) = e^{\lambda t}$. This suggests that the population grows at an exponential rate, and indeed, we have a martingale convergence confirming this Malthusian law of growth.

Theorem 1.9. *The family $\left\{Z(t)e^{-\lambda t}, \mathbb{F}_t; t \geq 0\right\}$ is a non-negative martingale and hence*

$$\lim_{t\to\infty} Z(t)e^{-\lambda t} = W \text{ exists a.s.} \tag{1.6}$$

Thanks to the additional structure of the continuous process, more can be said about the limit of the rescaled process compared to the discrete case. First, we note $\psi(u) = \mathbb{E}\left(e^{-uW}\right)$ satisfies the functional equation

$$\psi(u) = \int_o^\infty \phi\left(\psi\left(ue^{-\lambda y}\right)\right) e^{-y} dy. \tag{1.7}$$

We settle here the following theorem.

Theorem 1.10 (Harris [86]-Sevastyanov [144])**.** *Consider a supercritical continuous time branching process with generating function ϕ. Let $W = \lim_{t\to\infty} e^{-\lambda t}Z(t), \psi(u) = \mathbb{E}\left(e^{-uW}\right)$, and ψ^{-1} be the inverse function of ψ. Assume that $\sum_j p_j j \log j < \infty$. Then*

$$\psi^{-1}(x) = (1-x)\exp\left(\int_1^x \frac{\phi'(1)-1}{\phi(s)-s} + \frac{1}{1-s}\, ds\right), \quad p_{ext} < x \leq 1\,. \tag{1.8}$$

1.3 Erdős-Rényi Random Graphs

In this section we introduce the random graph model introduced by Erdős and Rényi in the late 1950's. The graph consists of n nodes, and each pair of nodes is independently connected with a fixed probability p. We denote by $\mathrm{ER}(n,p)$ the resulting random graph. This model was first studied in [62]. We refer to [12, 30, 42, 56, 90, 111] for more detailed references of the early literature on random graphs.

Despite the fact that $\mathrm{ER}(n,p)$ is the simplest imaginable model of a random network, a fascinating phase transition occurs as p increases. Phase transitions are well known in physics since they occur in various real phenomena, such as magnetism or the conductance properties of porous materials. Many models have been invented to describe and explain such phase transitions, and we shall see some examples in this thesis.

1.3.1 Emergence of the giant component

The Erdős-Rényi random graph $\mathrm{ER}(n,p)$ has vertex set $[n] = \{1,2,...,n\}$, and, denoting the edge between vertices $i,j \in [n]$ by (i,j), (i,j) is occupied (or present) with probability p, and absent or vacant otherwise, independently of other edges. For two vertices $i,j \in [n]$, we write $i \leftrightarrow j$ when there exists a path of occupied edges connecting i and j. By convention, we always assume that $i \leftrightarrow i$. For $v \in [n]$, we denote the connected component containing v (or cluster of v) by

$$C(v) := \{x \in [n] : v \leftrightarrow x\}. \tag{1.9}$$

We denote the size of $C(v)$ by $|C(v)|$. The largest connected component $\mathcal{C}_1$ is equal to any cluster $C(v)$ for which $|C(v)|$ is maximal, so that

$$|\mathcal{C}_1| = \max\{|C(v)|, v \in [n]\}. \tag{1.10}$$

Denote by $\mathcal{C}_2$ the second largest component.

The following theorem states that in the subcritical regime ($\lambda < 1$), all connected components are of logarithmic size; in the supercritical regime ($\lambda > 1$), a giant component appears, whose size scaled by n converges in probability as $n \to \infty$ to $1 - p_{ext}(\lambda)$, while other components remain of logarithmic size. In the critical regime ($\lambda = 1$), the largest critical cluster obeys a non-trivial scaling result, and it is of order $n^{2/3}$.

Theorem 1.11 ([29, 63, 107, 138]). *Consider the Erdős-Rényi random graph,* $\mathrm{ER}(n, \lambda/n)$*, on* n *vertices where there is an edge between two vertices with probability* λ/n *independently of everything else. Depending on the value of* λ*, the following regimes occur:*

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

$$\lim_{n\to\infty} \mathbb{P}\left(|C_1| \leq c \log n\right) = 1. \tag{1.11}$$

2. *Supercritical regime:* $\lambda > 1$*. Denote by* $p_{ext}(\lambda)$ *the extinction probability of a Galton-Watson branching process with* $\mathrm{Poi}(\lambda)$ *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:*

$$\mathbb{P}\left(\left|\frac{|C_1|}{n} - (1 - p_{ext}(\lambda))\right| \leq \delta, \text{ and } |\mathcal{C}_2| \leq c \log n\right) = 1. \tag{1.12}$$

3. *Critical regime:* $\lambda = 1$*. There exists a constant* $\alpha > 0$ *such that for all* $\omega > 1$ *and* n *sufficiently large,*

$$\mathbb{P}\left(\omega^{-1} n^{2/3} \leq |C_1| \leq \omega n^{2/3}\right) \geq 1 - \alpha/\omega. \tag{1.13}$$

Furthermore, Bollobás [29] found that the critical behavior extends throughout the regime where $\lambda = (1 \pm \epsilon)/n$ for $\epsilon = O(n^{-1/3})$, known as the critical window (or scaling window). Some more recent results in this area can be found in [124, 125, 8, 112, 3, 51].

We end this section by stating a central limit theorem in the supercritical regime, i.e., $\lambda > 1$, extending the law of large numbers for the giant component in the last theorem.

Theorem 1.12 ([137]). *Consider the Erdős-Rényi random graph,* $\mathrm{ER}(n, \lambda/n)$*, where* $\lambda > 1$*. Then,*

$$\frac{|C_1| - (1 - p_{ext})n}{\sqrt{n}} \xrightarrow{d} Z, \tag{1.14}$$

where Z *is a normal random variable with mean* 0 *and variance* $\sigma^2(\lambda) = \frac{p_{ext}(1-p_{ext})}{(1-\lambda p_{ext})^2}$.

1.3.2 Threshold for connectivity

In the last section we saw that, when the average degree λ of an Erdős-Rényi random graph is of constant order $\lambda > 1$, the graph contains a giant component of size of order n with high

probability. However, in that regime, this component's size is strictly less than n, so that the graph is disconnected. In this section we will investigate the question: How large does λ have to be, as a function of n, so that the probability that $\mathrm{ER}(n, \lambda/n)$ is connected (i.e., all vertices in the one component) tends to 1.

Let d_i be the degree of the vertex i. We have

$$\mathbb{P}\left(d_i = 0\right) = (1 - \lambda/n)^{n-1},$$

from which it easily follows that when $\lambda = c \log n$, we have $\mathbb{P}\left(d_i = 0\right) \sim n^{-c}$. Thus if $c < 1$, the number of isolated vertices $I_n = |\{i \in [n] : d_i = 0\}|$ has expectation

$$\mathbb{E}I_n = n\mathbb{P}\left(d_i = 0\right) \sim n^{1-c} \to \infty.$$

To show that the actual value of I_n is close to the mean, we note that if $i \neq j$,

$$\mathbb{P}(d_i = 0, d_j = 0) = (1 - \lambda/n)^{-1}\mathbb{P}\left(d_i = 0\right)\mathbb{P}\left(d_j = 0\right),$$

from which it is easy to conclude for $c < 1$ that

$$\mathrm{Var}(I_n) \sim n^{1-c} \sim \mathbb{E}I_n.$$

Using Chebyshev's inequality it follows that

$$\mathbb{P}\left(|I_n - \mathbb{E}I_n| > \omega(n)\mathbb{E}I_n\right) \leq \frac{1}{\omega(n)^2}.$$

We infer that for $c < 1$, with high probability there are about n^{1-c} isolated vertices, and hence the graph is not connected. Showing that the graph is connected is more complicated because we have to consider all possible ways in which the graph can fail to be connected. We refer to [30] for the proof of this.

Theorem 1.13. *Consider $G = \mathrm{ER}(n, \lambda/n)$ with $\lambda = \log n + c$. Then, the following limit holds.*

$$\lim_{n\to\infty} \mathbb{P}(G \text{ is connected }) = e^{-e^{-c}}. \tag{1.15}$$

A direct consequence of this theorem is that if the average degree λ asymptotically dominates $\log n$, i.e., if c is replaced by c_n such that c_n tends to infinity as n increases, then the graph is connected with high probability, i.e., with a probability that tends to one as n goes to infinity.

1.3.3 Emergence of the giant k-core

The k-core of the graph, denoted by Core_k, is by definition the maximal induced subgraph with minimum degree at least k, and in this section, we study the k-core of a random graph $\text{ER}(n, \lambda/n)$. The k-core of an arbitrary finite graph can be found by removing vertices of degree less than k, in an arbitrary order, until no such vertices exist. Note that the k-core may be empty. The question whether a non-empty k-core exists in a random graph has attracted a lot of attention over the last years. Here we recall the main theorem of [139].

For $\mu > 0$ and $j \in \mathbb{Z}^+$, let

$$\psi_j(\mu) := \mathbb{P}\left(\text{Poi}(\mu) \geq j\right).$$

Also let

$$\lambda_k := \min_{\mu>0} \mu/\psi_{k-1}(\mu); \tag{1.16}$$

and for $\lambda_k > 0$, denote by $\mu_k(\lambda)$, the largest solution to $\mu/\psi_{k-1}(\mu) = \lambda$, in $(0, 1]$.

In [139], Pittel, Spencer and Wormald discovered that for $k \geq 3$, $\lambda = \lambda_k$ is the threshold for the appearance of a nonempty k-core in the graph $\text{ER}(n, \lambda/n)$. Their strategy was to analyse an edge deletion algorithm that finds the k-core in the graph, showing that the corresponding random process is well approximated by the solution to a system of differential equations.

Theorem 1.14 (Pittel, Spencer and Wormald [139])**.** *For a fixed $\lambda > 0$, consider the random graph $\text{ER}(n, \lambda/n)$. Let $k \geq 2$ be fixed and let Core_k be the k-core of $\text{ER}(n, \lambda/n)$, and let $v(\text{Core}_k)$ and $e(\text{Core}_k)$ be the number of nodes and edges in the k-core, respectively.*

1. *If $\lambda < \lambda_k$ and $k \geq 3$, then Core_k is empty w.h.p. (with high probability) .*
2. *If $\lambda > \lambda_k$, then w.h.p. Core_k is non-empty, and*

$$\frac{v(Core_k)}{n} \xrightarrow{p} \psi_k(\mu_k(\lambda)), \tag{1.17}$$

$$\frac{e(Core_k)}{n} \xrightarrow{p} \frac{\mu_k(\lambda)^2}{2\lambda}. \tag{1.18}$$

Note that the first part does not hold for $k = 2$. Indeed in this case $\lambda_2 = 1$, and for $0 < \lambda < 1$, there is a positive limiting probability that the graph contains cycles (as already shown by Erdős and Rényi [64]). Thus a non-empty 2-core appears with positive probability. Nevertheless, in this case $e(\text{Core}_k) = O_p(1)$, and $v(\text{Core}_k) = O_p(1)$, so the 2-core is small (see Janson-Luczak [108]).

1.3.4 Distances and diameter

Having studied the existence of the giant component for $\lambda > 1$, we state now the known results on the typical distance between two points on the giant component.

Given a graph $G = (V, E)$, the distance $\mathrm{dist}(a, b)$ between two nodes a and b in V is the number of edges in E in the shortest path connecting these two vertices.

Theorem 1.15 (van den Esker, van der Hofstad and Hooghiemstra [65])**.** *Consider the Erdős-Rényi random graph,* $\mathrm{ER}(n, \lambda/n)$*, with* $\lambda > 1$*. Pick two vertices* a *and* b *independently at random from the giant cluster. Then*

$$\mathrm{dist}(a, b)/\log n \to 1/\log \lambda \tag{1.19}$$

in probability, as n *goes to infinity.*

This result follows from a result in [65]. It is very intuitive to see why such a result should be true. The branching process approximation of $\mathrm{ER}(n, \lambda/n)$ grows at rate λ^t, so the average distance is given by solving $\lambda^t = n$, i.e., $t = \log n/\log \lambda$.

The diameter of G, denoted by $\mathrm{diam}(G)$, is the maximum graph distance between any pair of connected vertices in V, i.e.,

$$\mathrm{diam}(G) := \max\{\mathrm{dist}(a, b) \mid \ a, b \in V,\ \mathrm{dist}(a, b) < \infty\}\,. \tag{1.20}$$

Note that the diameter of a graph is interesting from the point of view of applications, e.g., when the graph represents a network over which informations need to be transported; its value provides an upper bound on the time for the informations to go from any location u to any other location v, provided shortest paths between locations are used.

The main approach for understanding the diameter in Erdős-Rényi random graphs, is to compare the neighborhoods of a vertex of $\mathrm{ER}(n, \lambda/n)$ with the standard Poisson Galton-Watson branching process $\mathcal{Z}_\lambda = (Z_k)_{k\geq 0}$; this starts with a single particle in generation 0, and each particle in generation k, has a Poisson $\mathrm{Poi}(\lambda)$ number of children in the next generation, independently of the other particles and of the history.

A particle in the process $\mathcal{Z}_\lambda$ survives if it has descendants in all later generations; the process survives if the initial particle survives. If $\lambda > 1$, then the survival probability is the unique positive solution to $1 - s = e^{-\lambda s}$ (Theorem 1.1). Since particles in generation 1 survive

independently of each other, the number of such particles that survive has a $\mathrm{Poi}(s\lambda)$ distribution, the number of particles that go extinct has a $\mathrm{Poi}((1-s)\lambda)$ distribution, and these numbers are independent. It follows that conditioning on the extinction of the process, one obtains again a Poisson Galton-Watson process $\mathcal{Z}_{\lambda_*}$, with the dual parameter

$$\lambda_* = \lambda(1-s),$$

which may also be characterized as the solution $\lambda_* < 1$ to the equation

$$\lambda_* e^{-\lambda_*} = \lambda e^{-\lambda}.$$

Theorem 1.16 (Riordan and Wormald [141])**.** *Let $\lambda > 1$ be fixed, and let λ_* be the unique solution to $\lambda_* e^{-\lambda_*} = \lambda e^{-\lambda}$, $\lambda_* < 1$. Then*

$$\mathrm{diam}\,(\mathrm{ER}(n, \lambda/n)) = \frac{\log n}{\log \lambda} + 2\,\frac{\log n}{\log(1/\lambda_*)} + O_p(1). \tag{1.21}$$

As usual, we say that an event holds with high probability, or w.h.p., if its probability tends to 1 as $n \to \infty$. The above theorem simply says that, for any $K = K(n) \to \infty$, the diameter is w.h.p. within K of the sum of the first two terms. We remark that this theorem is a special case of a result of Fernholz and Ramachandran [69] for random graphs with a given degree sequence (see Theorem 1.29), and also of a result of Bollobás, Janson and Riordan [32, Section 14.2] for inhomogeneous random graphs with a finite number of vertex types. Earlier, Chung and Lu [41] also studied $\mathrm{diam}\,(\mathrm{ER}(n, \lambda/n))$, but their results were not strong enough to give the correct asymptotic form. Indeed, they conjectured that, under suitable conditions, the diameter is approximately $\log n/\log \lambda$, as one might initially expect.

The proof of the above theorem is given in [141]. However, the answer is easy to understand intuitively: typically, the size of the d-neighbourhood of a vertex (the set of vertices at distance d) grows by a factor of λ at each step (i.e., as d is increased by one). Starting from two typical vertices, taking $\log(\sqrt{n})/\log \lambda$ steps from each, the neighborhoods reach size about $\sqrt{n}$; at around this point, the neighborhoods are likely to overlap, so the typical distance between vertices is $\log n/\log \lambda$. The second term comes from exceptional vertices whose neighborhoods take some time to start expanding, or, equivalently, from the few very longest trees attached to (typical vertices of) the 2-core of $\mathrm{ER}(n, \lambda/n)$. It is well known that the trees hanging off the 2-core of $\mathrm{ER}(n, \lambda/n)$ have roughly the distribution of the branching process $\mathcal{Z}_{\lambda_*}$. Hence, some of these trees will have height roughly $\log n/\log(1/\lambda_*)$, and it turns out that the diameter arises from the distance between the leaves of two trees of (almost) maximal height attached to vertices in the 2-core at (almost) typical distance.

1.4 Configuration Model

After having recalled the basic results about the Erdős-Rényi random graphs, in this section, we present a detailed introduction to the random graphs with given fixed degrees (see e.g., [129, 56, 90, 131]), which will serve as our basic random graph model throughout this dissertation. The model covers a wide range of submodel arising in typical practical applications, like random regular graphs, $\mathrm{ER}(n, \lambda/n)$, the power law distributed graphs, exponentially distributed graphs, etc.

So ideally, we are interested in (uniformly chosen) random graphs having a prescribed degree sequence. But it is difficult to directly examine these random graphs, so instead, we introduce a model that produces a multigraph with the prescribed degrees, and which, when conditioned on simplicity of the multigraph, becomes uniform over all simple graphs with the prescribed degree sequence. This random multigraph is called the *configuration model* (or *'CM'*). The configuration model was originally developed by Bender and Canfield [19] and Bollobás [30] as a mean for generating a random graph with a prescribed sequence of vertex degrees (its earliest applications were in the study of random regular graphs).

For $n \in \mathbb{N}$, let $(d_i)_1^n$ be a sequence of non-negative integers such that $\sum_{i=1}^n d_i$ is even. By means of the configuration model, we define a random multigraph with given degree sequence $(d_i)_1^n$, denoted by $G^*(n, (d_i)_1^n)$ as follows. To each node i, we associate d_i labeled half-edges. All half-edges need to be paired to construct the multigraph, this is done by randomly matching them. When a half-edge of i is paired with a half-edge of j, we interpret this as an edge between i and j. The graph $G^*(n, (d_i)_1^n)$ obtained following this procedure may not be simple, i.e., may contain self-loops due to the pairing of two half-edges of i, and multi-edges due to the existence of more than one pairing between two given nodes. Note that $G^*(n, (d_i)_1^n)$ does not have exactly the uniform distribution over all multigraphs with the given degree sequence; there is a weight with a factor $1/j!$ for every edge of multiplicity j, and a factor $1/2$ for every loop, see e.g., [104]. However conditional on the multigraph $G^*(n, (d_i)_1^n)$ being a simple graph, we obtain a uniformly distributed random graph with the given degree sequence, which we denote by $G(n, (d_i)_1^n)$.

One specific example is when the degrees are all equal, in which case we speak of a random regular graph [156].

For $k \in \mathbb{N}$, let $u_k^{(n)} = |\{i : d_i = k\}|$ be the number of vertices of degree k, and $m^{(n)}$ be the

total degree defined by

$$m^{(n)} := \sum_{i=1}^{n} d_i = \sum_{k\geq 0} k u_k^{(n)}.$$

From now on, we will assume that we are given a sequence $(d_i)_1^n = (d_i^{(n)})_1^n$ for each $n \in \mathbb{N}$ satisfying the following regularity conditions analogous to the ones introduced by Molloy and Reed in [128]. For notational simplicity we will usually not show the dependence on n explicitly.

Condition 1.17. For each n, $\mathbf{d}^{(n)} = (d_i^{(n)})_1^n$ is a sequence of non-negative integers such that $\sum_{i=1}^n d_i$ 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: $u_k^{(n)}/n \to p_k$ 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: $\sum_{i=1}^n d_i^2 = O(n)$.

Let D_n be the random variable defined as the degree of a random (uniformly chosen) vertex in $G^*(n, (d_i)_1^n)$. Note in particular that

$$\mathbb{P}(D_n = k) = u_k^{(n)}/n. \tag{1.22}$$

Also remark that

$$\mathbb{E}D_n = n^{-1} \sum_{i=1}^{n} d_i^{(n)} = m^{(n)}/n. \tag{1.23}$$

Further, let D be the random variable with the distribution $\mathbb{P}(D = k) = p_k$.

Then Property (i) can be written as

$$D_n \xrightarrow{d} D. \tag{1.24}$$

In other words, D describes the asymptotic distribution of the degree of a random vertex in $G(n, (d_i)_1^n)$. Furthermore, Property (ii) is simply $\lambda = \mathbb{E}D \in (0, \infty)$, and Property (iii) can be written as

$$\mathbb{E}D_n^2 = O(1). \tag{1.25}$$

Remark 1.18. In particular, (1.25) implies that the random variables D_n are uniformly integrable, and thus, Condition 1.17 (i), in the form given in (1.24), implies that $\mathbb{E}D_n \to \mathbb{E}D$, i.e.,

$$\frac{2m^{(n)}}{n} = n^{-1}\sum_{i=1}^{n} d_i^{(n)} \to \lambda, \tag{1.26}$$

as $n \to \infty$ (see e.g., [84, Theorems 5.4.2 and 5.5.9]).

Theorem 1.19 (Janson [104])**.** *Consider a random graph $G^*(n,(d_i)_1^n)$ where the degree sequence $(d_i)_1^n$ satisfies Condition 1.17. Then*

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

As a corollary we obtain

Corollary 1.20. *Let $\mathbf{d}^{(n)} = (d_i)_1^n$ be a given fixed degree sequence satisfying* Condition 1.17*. Then, an event $\mathcal{E}_n$ occurs with high probability for $G(n,(d_i)_1^n)$ when it occurs with high probability for $G^*(n,(d_i)_1^n)$.*

Corollary 1.20 allows us to prove a property for uniform graphs with a given degree sequence by proving it for the configuration model with that degree sequence.

As in the case of Erdős-Rényi graphs, the growth of clusters can be approximated in the early stages by a branching process.

If we start with a given vertex x, then the number of neighbors (the first generation in the branching process) has distribution p_j. This is not true for the second generation. 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 is for $k \geq 1$ given by

$$q_{k-1} = \frac{kp_k}{\lambda}. \tag{1.27}$$

The $k-1$ on the left-hand side comes from the fact that we used up one edge connecting to the vertex.

Example 1.21. Consider the Poisson distribution $p_k = e^{-\lambda}\lambda^k/k!$, which is the asymptotic degree distribution for $\mathrm{ER}(n,\lambda/n)$. In this case we have

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

and q is again Poisson with mean λ. Conversely if $p = q$, we have $p_k = p_{k-1}\lambda/k$. Iterating this gives $p_k = p_0\lambda^k/k!$, so p_k is Poisson with mean λ.

Let ν denote

$$\nu := \sum_{k=0}^{\infty} k q_k = \frac{\mathbb{E}[D(D-1)]}{\mathbb{E}[D]} . \tag{1.28}$$

Define

$$\nu_n := \frac{\mathbb{E}[D_n(D_n-1)]}{\mathbb{E}[D_n]} = \frac{\sum_{i=1}^n d_i^{(n)}(d_i^{(n)}-1)}{m^{(n)}}. \tag{1.29}$$

Note that by Condition 1.17(iii), we have $\nu_n = O(1)$. This implies that $\nu < \infty$, since by Fatou's lemma

$$\nu \le \liminf_{n\to\infty} \nu_n.$$

We will sometimes need an extra assumption.

Condition 1.22. *As $n \to \infty$, $\nu_n \to \nu$. (Equivalently, $\mathbb{E}[D_n^2] \to \mathbb{E}[D^2]$.)*

This is clearly stronger than Condition 1.17 (iii). Assuming Condition 1.17, it is by (1.24) equivalent to uniform integrability of D_n^2. In particular this condition holds if $\sup_n \mathbb{E}\, D_n^{2+\epsilon} < \infty$ for some $\epsilon > 0$.

Remark 1.23. Our results in this thesis, can be also applied to some other random graphs models by conditioning on the vertex degrees. (This will work when the random graph conditioned on the degree sequence has a uniform distribution over all possibilities.) For example, for the Erdős-Rényi random graph $\mathrm{ER}(n, p_n)$, where every edge is present with probability p_n with $np_n \to \lambda \in (0, \infty)$, the Condition 1.17 (and the other conditions that we will consider through this thesis) holds in probability. Note that now $(d_i^{(n)})$ are random vertex degrees. As usual let $u_k^{(n)}$ be the (random) number of vertices with degree k. Indeed for $\mathrm{ER}(n, p_n)$ with $np_n \to \lambda \in (0, \infty)$, we have (see for example [109, Section 8])

(i) $u_k^{(n)} \xrightarrow{p} p_k$, for every $k \ge 0$ as $n \to \infty$;

(ii) for every $A \ge 1$, $\sum_k u_k^{(n)} A^k = \sum_{i=1}^n A^{d_i^{(n)}} = O_p(n)$.

Note that in this case, this holds with $p_k = e^{-\lambda}\lambda^k/k!$. By replacing the random graphs $\mathrm{ER}(n, p_n)$ by other random graphs $G(n, p_n)$ with the same distribution, we can assume that the random graphs are defined on a common probability space and that Properties (i) and (ii) above hold almost surely. It will be then easy to conclude that Conditions 1.17 and 1.22 hold a.s. In particular, Property (ii) which is for every $A \geq 1$, $\mathbb{E}[A^{D_n}] = \sum_k \frac{u_k^{(n)}}{n} A^k = O(1)$, implies uniform integrability of D_n (and all the powers D_n^s for fixed s; this can be for example obtained by observing that for $\epsilon > 0$, $\mathbb{E}\, D_n^{s(1+\epsilon)} \leq \mathbb{E}\, A^{D_n}$ for some large enough $A = A(s, \epsilon)$, and that $\mathbb{E}\, A^{D_n} = O(1)$ by assumption). Thus,

$$\frac{\sum_{i=1}^n d_i^{(n)}}{n} = \mathbb{E}[D_n] \to \mathbb{E}[D] = \sum_k k p_k = \lambda;$$

and $\lambda \in (0, \infty)$. Similarly, all higher moments converge, and Conditions 1.17 and 1.22 hold.

From the Skorokhod coupling theorem [113, Theorem 4.30] applied to the random sequences $\left(u_k^{(n)}\right)_{k=0}^{\infty}$, we may assume that the limit $u_k^{(n)} \to p_k$ in Property (i) above holds a.s., for every $k \geq 1$. We now "derandomize" Property (ii) above. By Property (ii), for every $j \geq 1$ and $k \geq 1$, we may choose $C_{k,j}$ increasing in j such that $\mathbb{P}(\sum_i k^{d_i^{(n)}} > C_{k,j} n) < 2^{-k}/j$. Now let $\mathbb{E}_j := \{\sum_i k^{d_i^{(n)}} \leq C_{k,j} n \text{ for every } k \geq 1\}$, and $\mathbb{E}_0 = \emptyset$. Note that $\mathbb{P}(\mathbb{E}_j) > 1 - 1/j$ for $j \geq 1$. Thus, condition on the event $\mathbb{E}_j$, Property (ii) holds uniformly. We apply the Skorokhod coupling theorem to $\left(u_k^{(n)}\right)_{k=0}^{\infty}$ conditioned on $\mathbb{E}_j \backslash \mathbb{E}_{j-1}$ for every $j \geq 1$ such that $\mathbb{P}(\mathbb{E}_j \backslash \mathbb{E}_{j-1}) > 0$; this shows that we can assume $u_k^{(n)} \xrightarrow{a.s.} p_k$ for every k on $\mathbb{E}_j \backslash \mathbb{E}_{j-1}$, and we only have to combine these pieces for $j \geq 1$. We refer to [109, Section 8], where this is done in more details for Erdős-Rényi random graphs.

The configuration model with i.i.d. degrees. We may also consider a model of random graphs given by a configuration model where the degrees are given by i.i.d. random variables having the distribution of a random variable D (in contrast to our earlier model with deterministic degree sequences). To explain this in more details, let us fix an integer n. Consider now an i.i.d. sequence $D_1, D_2, \ldots, D_n$. We will assume that $m_n = \sum_{j=1}^n D_j$ is even. If m_n is odd, then we add a half-edge to the n-th node, so that D_n is increased by one. This single half-edge will make hardly any difference in what follows, and we will ignore this effect (e.g., see [90]). We now specify the degree distribution. The probability mass function and the distribution function of

the nodal degree D are denoted by

$$p_j := \mathbb{P}(D = j), j = 0, 1, 2, \ldots, \quad \text{and}, \quad F(x) = \sum_{j=0}^{\lfloor x \rfloor} p_j, \tag{1.30}$$

where $\lfloor x \rfloor$ is the largest integer smaller than or equal to x. To construct the graph, we have n separate nodes and incident to node j, we have D_j half-edges. All half-edges need to be paired to construct the multigraph, this is done by randomly matching them. When a half-edge of i is paired with a half-edge of j, we interpret this as an edge between i and j. We denote the resulting random multigraph by $G(n, F)$. Note that here we have two sources of randomness: random degrees and random matching to construct the (multi-)graph.

Unless explicitly stated, in what follows we will be only working in the model $G(n, (d_i)_1^n)$; however, in few occasions we make reference to the model $G(n, F)$ for the existing results in the literature.

1.4.1 The giant component

The question of the existence of a giant component in $G(n, (d_i)_1^n)$ was answered by Molloy and Reed [129], who showed that a giant component exists w.h.p. if and only if (in the notation above) $\mathbb{E}[D(D-2)] > 0$.

Before we state the results, we start by introducing some notation. Let $G_p(x)$ be the probability generating function of D, i.e.,

$$G_p(x) := \mathbb{E}x^D = \sum_{k=0}^{\infty} p_k x^k, \tag{1.31}$$

and define further

$$h(x) := xG_p'(x) = \sum_{k=1}^{\infty} k p_k x^k, \text{ and} \tag{1.32}$$

$$H(x) := \lambda x^2 - h(x). \tag{1.33}$$

Note that $h(0) = 0$ and $h(1) = \lambda$, and thus, $H(0) = H(1) = 0$. Note also that

$$H'(1) = 2\lambda - \sum_{k=1}^{\infty} k^2 p_k = -\mathbb{E}D(D-2). \tag{1.34}$$

We recall that $v(G)$ and $e(G)$ denote the numbers of vertices and edges in the graph G, respectively; further, we let $v_k(G)$ be the number of vertices of degree k.

Theorem 1.24 (Molloy, Reed [129] - Janson, Luczak [110])**.** *Consider a random graph $G(n,(d_i)_1^n)$ where the degree sequence $(d_i)_1^n$ satisfies Condition 1.17. Assume that $p_0+p_2<1$. Let $\mathcal{C}_1$ and $\mathcal{C}_2$ be the largest and second largest components of $G(n,(d_i)_1^n)$.*

(i) If $\mathbb{E}[D(D-2)]=\sum_k k(k-2)p_k>0$, then there is a unique $\xi\in(0,1)$ such that $H(\xi)=0$, or equivalently $G_p'(\xi)=\lambda\xi$, and

$$v(\mathcal{C}_1)/n \xrightarrow{p} 1-G_p(\xi)>0, \tag{1.35}$$

$$v_k(\mathcal{C}_1)/n \xrightarrow{p} p_k(1-\xi^k), \text{ for every } k\ge 0, \tag{1.36}$$

$$e(\mathcal{C}_1)/n \xrightarrow{p} \frac{1}{2}\lambda(1-\xi^2), \tag{1.37}$$

while $v(\mathcal{C}_2)/n \xrightarrow{p} 0$ and $e(\mathcal{C}_2)/n \xrightarrow{p} 0$.

(ii) If $\mathbb{E}[D(D-2)]=\sum_k k(k-2)p_k\le 0$, then $v(\mathcal{C}_1)/n \xrightarrow{p} 0$ and $e(\mathcal{C}_1)/n \xrightarrow{p} 0$.

The same result holds for $G^*(n,(d_i)_1^n)$.

In the usual, more informal, language, the theorem shows that $G(n,(d_i)_1^n)$ has a giant component if and only if $\mathbb{E}[D(D-2)]>0$, i.e., $\nu>1$.

The case $p_0+p_2=1$ is much more exceptional. Note that in this case, $H(x)=0$ for all x. Quite different behaviours are possible in this case (e.g., see [110]).

(We refer to [114, 75, 88], for the results concerning the behavior of the largest component near the critical point $\nu=1$.)

Let $G_q(x)$ be the probability generating function of $\{q_k\}_{k=0}^\infty$, i.e.,

$$G_q(x):=\sum_{k=0}^\infty q_k x^k, \tag{1.38}$$

and let β denote the extinction probability of the branching process with distribution $\{q_k\}_{k=0}^\infty$, so that β is the smallest positive solution of the equation

$$\beta=G_q(\beta). \tag{1.39}$$

Set

$$\eta := \sum_{k=1}^{\infty} p_k \beta^k. \tag{1.40}$$

It is easy to see that η is the extinction probability of a branching process where the probability mass function of the first generation Z_1 is given by $\{p_k\}_{k=1}^{\infty}$, while the distribution of all further generation is given by $\{q_k\}_{k=0}^{\infty}$. We call such a process a delayed branching process.

We next state a duality principle for the configuration model. First define the dual probability distribution of $\{p_k\}_{k=1}^{\infty}$ as

$$\tilde{p}_k = \frac{p_k}{\eta} \beta^k. \tag{1.41}$$

Theorem 1.25. *The structure of the configuration model with degree sequence $(d_i^{(n)})_1^n$, formed by deleting the largest component $\mathcal{C}_1$ of $G(n, (d_i)_1^n)$, is the same as that of the configuration model with $\tilde{n} = n - |\mathcal{C}_1|$ vertices and with degree sequence $(\tilde{d}_i^{(n)})_1^{\tilde{n}}$, which are such that*

$$\lim_{n\to\infty} |\{i : 1 \le i \le \tilde{n}, \tilde{d}_i^{(n)} = k\}|/n = \tilde{p}_k,$$

and $\tilde{p}_k$ is defined in (1.41).

We refer to [129] for a proof of this. We now heuristically explain the form of the dual asymptotic degree distribution in the above theorem. Indeed the definition of $\{\tilde{p}_k\}_{k=1}^{\infty}$ can again be explained with help of a branching process. The $\{p_k\}$ sequence belongs to a branching process which will not become extinct with positive probability, which implies that the random graph with that degree sequence has a giant component. Thus, the removal of the giant component alters the accompanying branching process. As a result, with D_v being the degree of the vertex v, conditioned on v not belonging to the giant component, we would expect that

$$\tilde{p}_k = \mathbb{P}(D_v = k | v \text{ not in the giant component}) \sim \frac{\mathbb{P}\left(\{Z_1 = k\} \cap \{\text{extinction}\}\right)}{\mathbb{P}(\text{extinction})},$$

where Z_1 is the law of the first generation of the delayed branching process. The children in the first generation generate Z_1 independent branching processes, which all have to become extinct, and this happens with probability β^{Z_1}. The probability of extinction in the delayed branching process equals η, so that

$$\tilde{p}_k = \frac{p_k}{\eta} \beta^k.$$

1.4.2 The k-core

The existence of a large k-core in a random graph with a given degree sequence has been studied by several authors, see for example Cain and Wormald[38], Cooper [46], Fernholz and Ramachandran [68], and Janson and Luczak [108].

Let $k \geq 2$ be a fixed integer, and $\text{Core}_k^{(n)}$ be the k-core of the graph $G(n,(d_i)_1^n)$. We shall consider *thinning* of the vertex degrees in $G^*(n,(d_i)_1^n)$. Let D be the random variable with the distribution $\mathbb{P}(D=r)=p_r$, which is the asymptotic distribution of the vertex degrees in $G^*(n,(d_i)_1^n)$. For $0 \leq p \leq 1$, we let D_p be the thinning of D obtained by taking D points and then randomly and independently keeping each of them with probability p. For integers $l \geq 0$ and $0 \leq r \leq l$, let π_{lr} denote the binomial probabilities

$$\pi_{lr}(p) = \mathbb{P}\left(\text{Bin}(l,p)=r\right) = \binom{l}{r} p^r (1-p)^{l-r}. \tag{1.42}$$

Hence we have

$$\mathbb{P}\left(D_p = r\right) = \sum_{l=r}^{\infty} p_l \pi_{lr}(p). \tag{1.43}$$

We further define the functions

$$h(p) := \mathbb{E}\left[D_p \mathbb{1}(D_p \geq k)\right] = \sum_{r=k}^{\infty}\sum_{l=r}^{\infty} r p_l \pi_{lr}(p), \tag{1.44}$$

$$h_1(p) := \mathbb{P}\left(D_p \geq k\right) = \sum_{r=k}^{\infty}\sum_{l=r}^{\infty} p_l \pi_{lr}(p). \tag{1.45}$$

Note that D_p is stochastically increasing in p, and thus, both h and h_1 are increasing in p, with $h(0)=h_1(0)=0$. Note further that $h(1)=\sum_{r=k}^{\infty} r p_r \leq \lambda$, and $h_1(1)=\sum_{r=k}^{\infty} p_r \leq 1$, with strict inequalities unless $p_r=0$ for all $r=1,...,k-1$ or $r=0,1,...,k-1$, respectively.

Theorem 1.26 (Fernholz, Ramachandran [68] - Janson, Luczak [108])**.** *Consider a random graph $G(n,(d_i)_1^n)$ where the degree sequence $(d_i)_1^n$ satisfies Condition 1.17. Let $k \geq 2$ be fixed, and let $\text{Core}_k^{(n)}$ be the k-core of $G(n,(d_i)_1^n)$. Let $\hat{p}$ be the largest $p \leq 1$ such that $\lambda p^2 = h(p)$.*

(i) If $\hat{p}=0$, i.e., if $\lambda p^2 > h(p)$ for all $p \in (0,1]$, then $Core_k^{(n)}$ has $o(n)$ vertices and $o(n)$ edges w.h.p.

(ii) If $\hat{p} > 0$, and further suppose that $\hat{p}$ is not a local maximum point of the function $h(p) - \lambda p^2$. Then

$$v\left(\mathrm{Core}_k^{(n)}\right)/n \xrightarrow{p} h_1(\hat{p}) > 0, \tag{1.46}$$

$$v_j\left(\mathrm{Core}_k^{(n)}\right)/n \xrightarrow{p} \mathbb{P}\left(D_{\hat{p}} = j\right) = \sum_{l=j}^{\infty} p_l \pi_{lj}(\hat{p}),\ j \geq k, \tag{1.47}$$

$$e\left(\mathrm{Core}_k^{(n)}\right)/n \xrightarrow{p} h(\hat{p})/2 = \lambda \hat{p}^2/2. \tag{1.48}$$

The same result holds for $G^*(n, (d_i)_1^n)$.

1.4.3 Distances and diameter

In this section, we give an overview of the results on the average graph distance and diameter in the configuration model. We assume that $\nu > 1$, defined by (1.28), which is equivalent to the existence of a giant component in the configuration model (see Section 1.4.1).

In the the configuration model with i.i.d. degrees $G(n, F)$, where F is the distribution function of the integer-valued random variable D (the degree random variable), the following asymptotic for the distance between two uniformly chosen vertices a and b has been proved in [94], under the following extra assumption. Suppose that for some $\tau > 3$, there exists a constant $c > 0$ such that

$$1 - F(x) \leq c x^{-\tau+1}, \quad \text{for all } x \geq 1, \tag{1.49}$$

and that $\nu > 1$, where we recall that ν is given by

$$\nu = \frac{\mathbb{E}[D(D-1)]}{\mathbb{E}[D]}. \tag{1.50}$$

(Note that Assumption 1.49 implies that D has finite variance and so $\nu < \infty$. Also, the condition $\nu > 1$ is to ensure the existence of a giant component.)

Theorem 1.27 (van der Hofstad, Hooghiemstra and van Mieghem [94])**.** *Assume that Condition (1.49) is satisfied with $\tau > 3$, and let $\nu > 1$. Let a and b be two uniformly chosen vertices in the giant component of the graph $G(n, F)$, and so $\mathrm{dist}(a, b) < \infty$. Then*

$$\frac{\mathrm{dist}(a, b)}{\log n} \xrightarrow{p} \frac{1}{\log \nu}.$$

(We refer to [95] and [66], for the results concerning the typical distance in the case $\tau \in (2,3)$ and $\tau \in (1,2)$, respectively.)

We now explain the result concerning the diameter of $G(n,(d_i)_1^n)$.

Recall first the following basic result concerning Galton-Watson processes. Let $\mathcal{X}_q$ be a Galton-Watson tree with offspring distribution $\{q_k\}_{k=0}^{\infty}$. Then the extinction probability of the branching process, β, is the smallest solution of the fixed point equation

$$\beta = G_q(\beta).$$

We further define

$$\beta_* := G_q'(\beta) = \sum_{k=1}^{\infty} k q_k \beta^{k-1}. \tag{1.51}$$

Remark 1.28. Let $\mathcal{X}_q^+ \subseteq \mathcal{X}_q$ be the set of particles of $\mathcal{X}_q$ that survive (have descendants in all future generations). Note that in particular $\mathcal{X}_q^+$ contains the root of the branching process with probability $1-\beta$, and is empty otherwise. Denote by D the random variable with distribution $\mathbb{P}(D=r)=q_r$, and let D_p be the thinning of D obtained by taking D points and then randomly and independently keeping each of them with probability p (see Equation 1.43). Then the number of surviving children has the distribution $D_{1-\beta}$. Let D^+ denote the offspring distribution in $\mathcal{X}_q^+$. Furthermore, conditioning on a particular particle being in $\mathcal{X}_q^+$ means exactly the same thing as conditioning on at least one of its children surviving. Then we have

$$\begin{aligned}
\mathbb{P}(D^+=1) &= \mathbb{P}(D_{1-\beta}=1 | D_{1-\beta} \geq 1) \\
&= \frac{\sum_r q_r \beta (1-\beta)^{r-1}}{\beta} \\
&= G_q'(\beta) = \beta_*.
\end{aligned}$$

Hence the probability that $\mathcal{X}_q^+(k)$, the particles in generation k in $\mathcal{X}_q^+$, consists of a single particle, given that the hole process survives, is exactly β_*^k. This event corresponds to the branching process remaining *thin* for k generations.

The following asymptotic for the diameter of G_n, has been proved in [69]. This generalizes the result of Bollobás and Fernandez de la Vega [31] for the diameter of d-regular random graphs.

Theorem 1.29 (Fernholz and Ramachandran [69]). *Consider a random graph $G_n \sim G(n,(d_i)_1^n)$ where the degree sequence $(d_i)_1^n$ satisfies Conditions 1.17 and 1.22. Furthermore, assume that $u_1^{(n)} = 0$ when $p_1 = 0$, and that $u_2^{(n)} = 0$ when $p_2 = 0$. We have*

$$\frac{\operatorname{diam}(G_n)}{\log n} \xrightarrow{p} \frac{1}{\log \nu} + \frac{1}{|\log q_1|}\mathbb{1}(p_1 = 0, p_2 > 0) + \frac{2}{|\log \beta_*|}\mathbb{1}(p_1 > 0).$$

1.4.4 Site and bond percolation

Two important cases of disease diffusion on random graphs have been studied in depth. The first case is when all individuals are susceptible but the probability of an infected node to transmit the disease to a susceptible node is π. In the second case, only a fraction π of nodes is susceptible, but the disease is so contagious that if an individual gets infected all of their susceptible neighbors will become infected. The first model is known as bond percolation, where edges are deleted with probability $1-\pi$. The second is known as site percolation, and instead of the edges, it is the nodes that are randomly removed.

During the last decade, percolation theory has brought new understanding and techniques to a broad range of topics in physics, materials science, complex networks as well as in epidemiology (see e.g., [33, 82, 91, 116, 146]). Percolation is easy to study in Erdős-Rényi random graphs, since the result of retaining a fraction π of the edges or sites is another Erdős-Rényi random graph. Using the branching process heuristic, percolation occurs (there will be a giant component) if and only if the mean of the associated branching process is > 1. This observation is well known in the epidemic literature, where it is phrased "the epidemic will spread if the number of secondary infections caused by an infected individual is > 1".

We will consider percolation of the random graph given by the configuration model. We first generate a random graph $G^*(n,(d_i)_1^n)$ and then percolate it. Fountoulakis [72] and Janson [103] show that both for site and bond percolation in $G^*(n,(d_i)_1^n)$, if one denotes by $\tilde{n}$ the number of vertices in the resulting random graph, and then condition on its degree sequence $(\tilde{d}_i)_1^n$, then this obtained graph has the distribution of $G(\tilde{n},(\tilde{d}_i)_1^{\tilde{n}})$, the random graph with this prescribed degree sequence. This shows that one needs only to calculate the distributions of the degree sequence $(\tilde{d}_i)_1^{\tilde{n}}$, and finally apply known results to $G(\tilde{n},(\tilde{d}_i)_1^{\tilde{n}})$. Our presentation below mainly follows the original work of Janson [103].

Site percolation.

Given any graph $G = (V, E)$ and a probability $\pi \in [0, 1]$, we define the random graph $G_{\pi,v}$, the graph obtained from G by randomly deleting each vertex (together with all incident edges) with probability $1 - \pi$, independently of all other vertices.

Thus π denotes the probability to be kept in the percolation model. When, as in our case, the original graph G itself is random, it is further assumed that we first sample G and then proceed as above, conditionally on G.

The cases $\pi = 0, 1$ are trivial: $G_{1,v} = G$, while $G_{0,v} = \emptyset$, the empty graph with no vertices and no edges. We will thus consider $0 < \pi < 1$.

We consider the generalized site percolation model in which the probability depends on the degree of the vertex. Thus, if $\pi = (\pi_d)_0^\infty$ is a given sequence of probabilities $\pi_d \in [0, 1]$, let $G_{\pi,v}$ be the random graph obtained by deleting vertices independently of each other, with vertex $v \in G$ deleted with probability $1 - \pi_{d_v}$ where d_v is the degree of v in G.

Instead of deleting a vertex, if its degree is d, we first split it into d new vertices of degree one; the new vertices are colored in red (e.g., see[103]). Then we remove all red vertices. We note that the (random) splitting changes the number of vertices, but not the number of half-edges. Moreover, given the set of splittings, there is a one-to-one correspondence between configurations before and after the splittings, and thus, if we condition on the new degree sequence, the graph after splittings is still described by the configuration model. Furthermore, by symmetry, when removing the red vertices, all vertices of degree one are equivalent, so one simply has to remove the right number of vertices of degree one, chosen uniformly at random. We can thus obtain $G^*_{\pi,v}(n, (d_i)_1^n)$ as follows:

- with probability $1 - \pi(d_i)$, each vertex i is replaced by d_i new vertices of degree one. Let $(\tilde{d}_i)_1^{\tilde{n}}$ be the resulting random degree sequence, of length $\tilde{n}$ (the number of vertices after splittings), and let n_+ be the number of new vertices.
- Construct the random graph $G(\tilde{n}, (\tilde{d}_i)_1^{\tilde{n}})$.
- Finish by deleting n_+ randomly chosen vertices of degree one.

Let n_j be the number of vertices of degree j in $(d_i)_1^n$, and $\tilde{n}_j = \{i \le \tilde{n} : \tilde{d}_i = j\}$ be the

number of vertices of degree j after the splittings. Thus:

$$\sum_{j=0}^{\infty} \tilde{n}_j = \tilde{n}. \tag{1.52}$$

We let n_j^o be the number of vertices of degree j that are not split. Then

$$\begin{aligned} n_j^o &= \text{Bin}(n_j, \pi_j) \text{ (independent of each other)}, && (1.53)\\ n_+ &= \sum_{j=0}^{\infty} j(n_j - n_j^o), && (1.54)\\ \tilde{n}_j &= n_j^o,\ j \neq 1, && (1.55)\\ \tilde{n}_1 &= n_1^o + n_+. && (1.56) \end{aligned}$$

We consider a random graph $G^*(n, (d_i)_1^n)$ where the degree sequence $(d_i)_1^n$ satisfies Condition 1.17. Then by the law of large numbers and Condition 1.17 we have

$$\begin{aligned} n_j^o &= n_j\pi_j + o_p(n) = \pi_j p_j n + o_p(n), && (1.57)\\ n_+ &= \sum_{j=0}^{\infty} j(1-\pi_j)p_j n + o_p(n), && (1.58)\\ \tilde{n}_j &= \pi_j p_j n + o_p(n),\ j \neq 1, && (1.59)\\ \tilde{n}_1 &= (\pi_1 p_1 + \sum_{j=0}^{\infty} j(1-\pi_j)p_j)n + o_p(n), && (1.60)\\ \tilde{n} &= \sum_{j=0}^{\infty} (\pi_j + j(1-\pi-j))\, p_j n + o_p(n). && (1.61) \end{aligned}$$

We can rewrite the above equations as

$$\begin{aligned} \frac{\tilde{n}}{n} &\xrightarrow{p} \xi := \sum_{j=0}^{\infty} (\pi_j + j(1-\pi_j))\, p_j, && (1.62)\\ \frac{\tilde{n}_j}{\tilde{n}} &\xrightarrow{p} \widetilde{p}_j := \xi^{-1}\pi_j p_j,\ j \neq 1, && (1.63)\\ \frac{\tilde{n}_1}{\tilde{n}} &\xrightarrow{p} \widetilde{p}_1 := \xi^{-1}(\pi_1 p_1 + \sum_{j=0}^{\infty} j(1-\pi_j)p_j) && (1.64) \end{aligned}$$

We also have $\tilde{\lambda} := \sum_j j\widetilde{p}_j = \xi^{-1}\lambda$, and it is easy to see that Condition 1.17 holds, in probability, for the random degree sequence $((\tilde{d}_i)_1^{\tilde{n}})$. By Skorokhod coupling theorem [113, Theorem 4.30], we can then state that Condition 1.17 a.s. holds for $(\tilde{d}_i)_1^{\tilde{n}}$.

Let $G_{\widetilde{p}}$ be the probability generating function of $\{\widetilde{p}_k\}_{k=0}^{\infty}$. We obtain

$$\xi G_{\widetilde{p}}(x) = \sum_{j=0}^{\infty} \xi \widetilde{p}_j x^j = \sum_{j=0}^{\infty} \pi_j p_j x^j + \sum_{j=0}^{\infty} j(1-\pi_j)p_j x = \lambda x + \sum_{j=0}^{\infty} \pi_j p_j (x^j - jx). \tag{1.65}$$

In particular, if all $\pi_j = \pi$,

$$\xi G_{\widetilde{p}}(x) = \pi G_p(x) + (1-\pi)\lambda x, \tag{1.66}$$

where now $\xi = \pi + (1-\pi)\lambda$.

This leads to the following result for the asymptotic size of $\mathcal{C}_1$.

Theorem 1.30 (Janson [103]). *Consider the site percolation model $G^*_{\pi,v}(n,(d_i)_1^n)$ where the degree sequence $(d_i)_1^n$ satisfies Condition 1.17, and that $\pi = (\pi_d)_0^\infty$ with $\pi_d \in [0,1]$; suppose further that there exists $j \geq 1$ such that $p_j > 0$ and $\pi_j < 1$. Let $\mathcal{C}_1$ and $\mathcal{C}_2$ be the largest and the second largest components of $G^*_{\pi,v}(n,(d_i)_1^n)$, respectively.*

(i) If $\sum_{j=0}^{\infty} j(j-1)\pi_j p_j > \lambda$, then there is a unique $\xi = \xi_v(\pi) \in (0,1)$ such that

$$\sum_{j=1}^{\infty} j\pi_j p_j (1-\xi^{j-1}) = \lambda(1-\xi), \tag{1.67}$$

and we have

$$v(\mathcal{C}_1)/n \xrightarrow{p} \chi_v(\pi) := \sum_{j=1}^{\infty} \pi_j p_j (1-\xi^j) > 0, \tag{1.68}$$

$$e(\mathcal{C}_1)/n \xrightarrow{p} \mu_v(\pi) = (1-\xi)\sum_{j=1}^{\infty} j\pi_j p_j - \frac{1}{2}\lambda(1-\xi)^2, \tag{1.69}$$

while $v(\mathcal{C}_2)/n \xrightarrow{p} 0$ and $e(\mathcal{C}_2)/n \xrightarrow{p} 0$.

(ii) If $\sum_{j=0}^{\infty} j(j-1)\pi_j p_j \leq \lambda$, then $v(\mathcal{C}_1)/n \xrightarrow{p} 0$ and $e(\mathcal{C}_1)/n \xrightarrow{p} 0$.

As a corollary, in the standard case when all $\pi_d = \pi$, this leads to a simple criterion for the existence of the giant component (which has also been shown by Britton, Janson, Martin-Löf [36], and Fountoulakis [72] by different methods).

Corollary 1.31 (Britton, Janson, Martin-Löf [36] - Fountoulakis [72]). *Suppose that Condition 1.17 holds and $0 < \pi < 1$. Then there exists w.h.p. a giant component in $G^*_{\pi,v}(n,(d_i)_1^n)$ if and only if*

$$\pi > \pi_c := \frac{\mathbb{E}D}{\mathbb{E}D(D-1)}. \tag{1.70}$$

Bond percolation.

Given any graph $G = (V, E)$ and a probability $\pi \in [0, 1]$, we define the random graph $G_{\pi,e}$ the graph obtained from G by randomly deleting each edge with probability $1 - \pi$, independently of all other edges. Thus, π denotes the probability for each edge to be kept in the percolation model. The same ideas as in the case of site percolation may be applied.

Consider a random graph $G^*(n, (d_i)_1^n)$ where the degree sequence $(d_i)_1^n$ satisfies Condition 1.17. For any half-edge, with probability $1 - \sqrt{\pi}$, independently of all other half-edges, we transfer it to a new red vertex of degree one. This does not change the number of half-edges, and there is a one-to-one correspondence between configurations before and after the transferals. We finish by removing all red vertices and their incident edges. Since an edge consists of two half-edges, and every half edge is kept with probability $\sqrt{\pi}$, the resulting graph is equivalent to the bond percolation model $G^*_{\pi,e}(n, (d_i)_1^n)$ where edges are kept with probability π. This gives the following algorithm:

- For each vertex i, replace its degree d_i by an independent random degree $\tilde{d}_i = \mathrm{Bin}(d_i, \sqrt{\pi})$.
- Add $n_+ := \sum_{i=1}^n (d_i - \tilde{d}_i)$ new nodes with degree one to the sequence $(\tilde{d}_i)_1^n$, and let $\tilde{\mathbf{d}}_{\pi,e}$ be the resulting degree sequence and $\tilde{n} = n + n_+$ its length.
- Construct the random graph $G^*(n, \tilde{\mathbf{d}}_{\pi,e})$.
- Finish by deleting n_+ randomly chosen vertices of degree one.

Let $\tilde{n}_j = \{i \le \tilde{n} : \tilde{d}_i = j\}$ be the number of vertices of degree j after the transferals, and $\tilde{n}_{jl}$ be the number of vertices that had degree l before the transferals and j after. Hence $\tilde{n}_j = \sum_{l=j}^\infty \tilde{n}_{jl}$ for $j \neq 1$ and $\tilde{n}_1 = \sum_{l=1}^\infty \tilde{n}_{1l} + n_+$. A vertex of degree l will have after the transferals a degree with the binomial distribution $\mathrm{Bin}(l, \sqrt{\pi})$, and thus, the probability that it will become a vertex of degree j is the binomial probability $\pi_{lj}(\sqrt{\pi})$.

Then by independence of transferals at different vertices, and by the law of large numbers (using Condition 1.17), we have

$$\tilde{n}_{jl} = \pi_{lj}(\sqrt{\pi}) p_l n + o_p(n). \tag{1.71}$$

Furthermore, the number n_+ of new vertices equals the number of transferals, and thus, has

the binomial distribution Bin $\left(m^{(n)}, 1-\sqrt{\pi}\right)$. By analogy with the site percolation, we have

$$\frac{n_+}{n} \xrightarrow{p} (1-\sqrt{\pi})\lambda, \tag{1.72}$$

$$\frac{\tilde{n}}{n} \xrightarrow{p} \xi := 1 + (1-\sqrt{\pi})\lambda, \tag{1.73}$$

$$\frac{\tilde{n}_j}{\tilde{n}} \xrightarrow{p} \widetilde{p}_j := \xi^{-1} \sum_{l=j}^{\infty} \pi_{lj}(\sqrt{\pi}) p_l, \ j \neq 1, \tag{1.74}$$

$$\frac{\tilde{n}_j}{\tilde{n}} \xrightarrow{p} \widetilde{p}_1 := \xi^{-1} (\sum_{l=1}^{\infty} \pi_{l1}(\sqrt{\pi}) p_l + (1-\sqrt{\pi})\lambda). \tag{1.75}$$

Again it is easy to see that Condition 1.17 a.s. holds for $(\tilde{d}_i)_1^{\tilde{n}}$, and $G_{\widetilde{p}}$, the probability generating function of $\{\widetilde{p}_k\}_{k=0}^{\infty}$, satisfy

$$\xi G_{\widetilde{p}}(x) = G_p(1-\sqrt{\pi}+\sqrt{\pi}x) + (1-\sqrt{\pi})\lambda x. \tag{1.76}$$

This leads to the following result for the asymptotic size of $\mathcal{C}_1$.

Theorem 1.32 (Britton, Janson, Martin-Löf [36] - Fountoulakis [72] - Janson [103])**.** *Suppose that Condition 1.17 holds, and $0 < \pi < 1$. Let $\mathcal{C}_1$ and $\mathcal{C}_2$ be the largest and the second largest components of $G^*_{\pi,e}(n, (d_i)_1^n)$, respectively. Then there exists w.h.p. a giant component in $G^*_{\pi,e}(n, (d_i)_1^n)$, if and only if*

$$\pi > \pi_c := \frac{\mathbb{E}D}{\mathbb{E}D(D-1)}. \tag{1.77}$$

(i) If $\pi > \pi_c$, then there is a unique $\xi = \xi_e(\pi) \in (0,1)$, such that

$$\sqrt{\pi} G_p'(1-\sqrt{\pi}+\sqrt{\pi}x) + (1-\sqrt{\pi})\lambda = \lambda\xi, \tag{1.78}$$

and then

$$v(\mathcal{C}_1)/n \xrightarrow{p} \chi_e(\pi) := 1 - G_p(1-\sqrt{\pi}+\sqrt{\pi}\xi) > 0, \tag{1.79}$$

$$e(\mathcal{C}_1)/n \xrightarrow{p} \mu_e(\pi) = \sqrt{\pi}(1-\xi)\lambda - \frac{1}{2}\lambda(1-\xi)^2, \tag{1.80}$$

while $v(\mathcal{C}_2)/n \xrightarrow{p} 0$, and $e(\mathcal{C}_2)/n \xrightarrow{p} 0$.

(ii) If $\pi \leq \pi_c$, then $v(\mathcal{C}_1)/n \xrightarrow{p} 0$, and $e(\mathcal{C}_1)/n \xrightarrow{p} 0$.

We notice that the same criterion holds for the existence of a giant component for both site and bond percolation.

1.4.5 Random directed graphs with given vertex in- and out-degrees

We end this introductory chapter by considering a model of random directed graph constructed by the directed analogue of the configuration model. This model is indeed useful in real world applications, since many networks, and their interesting properties, are usually modeled in directed graph (in contrast to the undirected case), see for example [1, 6, 37, 89] and Chapter 4. An example of such a graph is the World-Wide Web (WWW), which is directed since every hyperlink between two pages goes in only one direction. However, the directed configuration model has not been considered in as much as details as the undirected case, and deserves more investigation. Our presentation below mainly follows the original work of Cooper and Frieze [47], all the definitions and results are from that paper (modified to be conform with our previous notations).

Let $\mathbf{d_n^+} = \{d_n^+(i)\}_{i=1}^n$, and $\mathbf{d_n^-} = \{d_n^-(i)\}_{i=1}^n$ be two sequences of non-negative integers such that $\sum_{i=1}^n d_n^+(i) = \sum_{i=1}^n d_n^-(i)$. The configuration model (CM) on n vertices with degree sequences $\mathbf{d_n^+}$ and $\mathbf{d_n^-}$ is constructed as follows:

We associate to each node i, two sets: $W_n^+(i)$ representing its out-going half-edges, and $W_n^-(i)$ representing its in-coming half-edges, with $|W_n^+(i)| = d_n^+(i)$, and $|W_n^-(i)| = d_n^-(i)$. Let $W_n^+ = \bigcup_i W_n^+(i)$, and $W_n^- = \bigcup_i W_n^-(i)$. A *configuration* is a matching of W_n^+ with W_n^-. To each configuration we assign a graph. When an out-going half-edge of node i is matched with an in-coming half-edge of node j, a directed edge from i to j appears in the graph. The configuration model is the probability space in which all configurations, as defined above, have equal probability. In other words, we choose the configuration at random, uniformly over all possible configurations. We denote the resulted random digraph with $G^*(n, \mathbf{d_n^+}, \mathbf{d_n^-})$.

It is quite easy to see that, conditional on the resulting multigraph being a simple graph, we obtain a uniformly distributed random digraph with the given degree sequence, which we denote by $G(n, \mathbf{d_n^+}, \mathbf{d_n^-})$.

We denote by m_n the total number of links

$$m_n := \sum_{i=1}^n d_n^+(i) = \sum_{i=1}^n d_n^-(i),$$

and we introduce the *empirical distribution of the degrees* as

$$\mu_n(j,k) := \frac{1}{n}\#\{i : d_n^+(i) = j, d_n^-(i) = k\}.$$

We let $\lambda_n := m_n/n$, and

$$\nu_n := \frac{1}{\lambda_n} \sum_{j,k} jk\mu_n(j,k). \tag{1.81}$$

We say that the degree sequences $\{\mathbf{d}_\mathbf{n}^+\}$ and $\{\mathbf{d}_\mathbf{n}^-\}$ are **proper** if they satisfy the following regularity conditions.

Condition 1.33. *For each $n \in \mathbb{N}$, $\mathbf{d}_\mathbf{n}^+ = \{(d_n^+(i))_{i=1}^n\}$ and $\mathbf{d}_\mathbf{n}^- = \{(d_n^-(i))_{i=1}^n\}$ are sequences of nonnegative integers with $\sum_{i=1}^n d_n^+(i) = \sum_{i=1}^n d_n^-(i)$, and such that, for some probability distribution $\mu(j,k)$, independent of n:*

(i) The empirical proportion $\mu_n(j,k)$ of vertices of degree (j,k) tends to $\mu(j,k)$: $\mu_n(j,k) \to \mu(j,k)$ as $n \to \infty$;

(ii) As $n \to \infty$; $\lambda_n \to \lambda \in (0,\infty)$, and $\nu_n \to \nu \in (0,\infty)$; where

$$\begin{aligned} \lambda &:= \sum_{j,k} j\mu(j,k) = \sum_{j,k} k\mu(j,k), \text{ and} \\ \nu &:= \sum_{j,k} jk\mu(j,k)/\lambda. \end{aligned}$$

(iii) The second moment property: $\sum_{j,k} j^2\mu_n(j,k) = O(1)$, $\sum_{j,k} k^2\mu_n(j,k) = O(1)$;

(iv) $\Delta_n := \max_{i\in V} (d_n^+(i), d_n^-(i)) \le n^{1/12}/\log n$;

(v) $\rho_n := \max\left(\sum_{i,j} \frac{i^2 j\mu_n(i,j)}{\lambda_n}, \sum_{i,j} \frac{j^2 i\mu_n(i,j)}{\lambda_n}\right) = o(\Delta_n)$, if $\Delta_n \to \infty$.

A directed graph is called strongly connected if there is a path from each vertex in the graph to every other vertex. The strongly connected components of a directed graph G are its maximal strongly connected subgraphs. The strong connectivity for spaces of sparse random directed graphs with prescribed degree sequence has been studied by Cooper and Frieze in [47]. Here, we summarize some of their results by following closely the presentation given in [47].

Consider the random graph $G(n, \mathbf{d}_\mathbf{n}^+, \mathbf{d}_\mathbf{n}^-)$ with the proper degree sequences, i.e., satisfying Condition 1.33. For a fixed vertex $a \in V_n = [1,n]$, the *fan-out* of a, denoted by $R_n^+(a)$, is the set of vertices $b \in V_n$ (including a) reachable from a by a directed path. Similarly, the *fan-in* of a, denoted by $R_n^-(a)$ is the set of vertices which can reach a, i.e., $b \in R_n^-(a)$ if and only if, there is a directed path from b to a.

Let us fix a small constant $\epsilon > 0$, and a large constant $C > 0$. We will say that a node a has a *large fan-out* if $|R_n^+(a)| > \epsilon n$, and that a has a *large fan-in* if $|R_n^-(a)| > \epsilon n$. We will say that a has a *small fan-out* if $|R_n^+(a)| < C\Delta_n^2 \log n$, and that a has a *small fan-in* if $|R_n^-(a)| < C\Delta_n^2 \log n$.

Let L_n^+ denote the set of vertices with a large fan-out and L_n^+ the set of vertices with a large fan-in in the random graph $G(n, \mathbf{d_n^+}, \mathbf{d_n^-})$.

We will first give an intuitive description of the model and then state the main result of [47]. As in the case of non directed graphs, the growth of clusters can be approximated in the early stages by a branching process.

Let

$$\mu_k^+ := \frac{1}{\lambda} \sum_j j\mu(j,k), \text{ and } \mu_j^- := \frac{1}{\lambda} \sum_k k\mu(j,k). \tag{1.82}$$

Indeed μ^+ is the distribution of the out-degree of the terminal vertex of a randomly chosen arc. Similarly μ^- is the distribution of the in-degree of the initial vertex of a randomly chosen arc. We infer ν may be written as

$$\nu = \sum_k k\mu_k^+ = \sum_j j\mu_j^-.$$

Let $\mathcal{X}^+$ be the independent branching process with a single initial node in which the probability distribution of the number of descendants of a node is μ^+. Let η^+ the probability that $\mathcal{X}^+$ continues indefinitely. Thus, $1 - \eta^+$, the extinction probability of $\mathcal{X}^+$, is given by the smallest positive solution of $x = \sum_k \mu_k^+ x^k$ (e.g., see Theorem 1.1) and satisfies

$$1 - \eta^+ = \sum_k \mu_k^+ (1 - \eta^+)^k. \tag{1.83}$$

If $\nu < 1$ then the smallest positive solution of the above fixed point equation is one (i.e., $\eta^+ = 0$), whereas if $\nu > 1$ there is a unique solution with $\eta^+ \in (0, 1)$.

If the root vertex $\emptyset$ has out-degree k then by doing a breadth first search from $\emptyset$, we see that the arcs of a small $R^+(\emptyset)$ will be approximately the union of k independent copies of the branching process $\mathcal{X}^+$. Thus a root vertex of out-degree k has probability $(1 - \eta^+)^k$ of a finite progeny. Thus, the probability that a randomly chosen vertex will have a small fan-out is approximately $1 - \pi^+$, where

$$1 - \pi^+ = \sum_{j,k} \mu(j,k)(1 - \eta^+)^k. \tag{1.84}$$

We should therefore expect that (as $n \to \infty$)

$$|L_n^+| \approx \pi^+ n.$$

Similarly, we expect that

$$|L_n^-| \approx \pi^- n,$$

where

$$1 - \pi^- = \sum_{j,k} \mu(j,k)(1-\eta^-)^k, \tag{1.85}$$

and $1 - \eta^-$ be the smallest positive value satisfying

$$1 - \eta^- = \sum_j \mu_j^-(1-\eta^-)^j. \tag{1.86}$$

If $\nu < 1$ then $\eta^+ = \eta^- = 0$, and hence $\pi^+ = \pi^- = 0$. This suggests that $\nu < 1$ implies that

$$L_n^+ = L_n^- = \emptyset.$$

For the number of vertices in a large fan-out note that: If a has a large fan-out and b has a large fan-in then it is very likely that there will be an arc directed from the fan-out of a to the fan-in of b which of course implies that b would be in the fan-out of a. Conversely, if b has a small fan-in then this is unlikely. Thus we expect that when $R_n^+(a)$ is large, then $L_n^- \subseteq R_n^+(a)$ and $|R_n^+(a) \backslash L_n^-| = o(n)$, and so

$$|R_n^+(a)| \approx \pi^- n.$$

Similarly, we expect that when $R_n^-(a)$ is large, then $L_n^+ \subseteq R_n^-(a)$ and $|R_n^-(a) \backslash L_n^-| = o(n)$, and so

$$|R_n^-(a)| \approx \pi^+ n.$$

Thus, we expect that w.h.p. $L_n^+ \bigcap L_n^-$ is contained in a strongly connected component (denoted by S_n). Indeed since any vertex in $L_n^+ \bigcap L_n^-$ must have a large fan-in and a large fan-out, then w.h.p. $L_n^+ \bigcap L_n^-$ induces a maximal strongly connected. The size of this strong component is approximately $(\pi^+ + \pi^- + \psi - 1)n$, where

$$\psi := \sum_{j,k} \mu(j,k)(1-\eta^-)^j(1-\eta^+)^k. \tag{1.87}$$

This can be explained as follows: Choose a random vertex. It has probability $\mu(j,k)$ of having in-degree j and out-degree k. Then the expression $(1-\eta^-)^j(1-\eta^+)^k$ is an estimate of the probability that all of the $i+j$ associated branching processes become extinct. Thus, $n(1-\psi)$ is a good estimate of $L_n^+ \bigcup L_n^-$.

The above giant strongly connected component will be unique. Every other strong component will be of size at most $C\Delta_n^2 \log n$ w.h.p. since every vertex not in S_n either has a small fan-out or a small fan-in.

Theorem 1.34 (Cooper and Frieze [47]). *Consider the random directed graph $G(n, \mathbf{d_n^+}, \mathbf{d_n^-})$, where the degree sequence satisfies Condition 1.33, i.e., the degree sequences are proper. Let the parameters π^+, π^-, ψ be defined as above.*

(i) If $\nu < 1$, then w.h.p. $L_n^+ = L_n^- = \emptyset$.

(ii) If $\nu > 1$, then there is a unique giant strongly connected component, with vertex set $S_n = L_n^+ \bigcap L_n^-$, and we have

$$\frac{|S_n|}{n} \longrightarrow \pi^+ + \pi^- + \psi - 1, \tag{1.88}$$

as $n \to \infty$.

Chapter 2

First Passage Percolation, Flooding, and Diameter

Abstract. *In this chapter, we study the impact of the edge weights on distances in diluted random graphs. We interpret these weights as delays, and take them as i.i.d exponential random variables. We analyze the edge flooding time defined as the minimum time needed to reach all nodes from one uniformly chosen node, and the edge diameter corresponding to the worst case edge flooding time. Under some regularity conditions on the degree sequence of the random graph, we show that these quantities grow as the logarithm of n, when the size of the graph n tends to infinity. We also derive the exact value for the prefactors.*

These allow us to analyze an asynchronous randomized broadcast algorithm for random regular graphs. Our results show that the asynchronous version of the algorithm performs better than its synchronized version: in the large size limit of the graph, it will reach the whole network faster even if the local dynamics are similar on average.

Contents

2.1 Introduction

In this chapter, we study the impact of the introduction of edge weights on the typical distances in a random graph and, in particular, on its diameter. Such weights can be thought of as economic costs, congestion delays, or carrying capabilities that can be encountered in real networks such as transportation systems and communication networks [126, Chapter 16]. Our main result consists of a precise asymptotic expression for the (edge) diameter of sparse random graphs on n vertices (under some standard regularity conditions) when the edge weights are i.i.d. exponential random variables of rate one.

The analysis of the asymptotics of typical distances in edge weighted graphs has received much interest by the statistical physics community in the context of *first passage percolation problems*. First-passage percolation (FPP) describes the dynamics of a fluid spreading within a random medium. This model has been mainly studied on lattices motivated by its subadditive property and its link to a number of other stochastic processes, see [83, 117, 85] for a more detailed discussion. First passage percolation with exponential weights has received substantial attention (see [21, 92, 93, 94, 25, 24, 102]), in particular on the complete graph, and, more recently, also on random graphs.

Driven by the distributed nature of modern network architectures, there has been intense research to devise algorithms to ensure effective network computation. Of particular interest is the problem of global node outreach, whereby some major event happening in one part of the network has to be communicated to all other nodes. In this context, gossip protocols have been identified as simple, efficient and robust mechanisms for disseminating and retrieving information for various network topologies. These mechanisms rely on simple periodic local operations between neighboring nodes [135].

Flooding corresponds to the most commonly used such process. A source node that first records the event notifies all the nodes within its reach. Subsequently, each of these neighbors forwards information to all of its neighbors and so on. If the underlying network is connected, such information will eventually reach all the nodes. The performance of this procedure can be evaluated in terms of the time it takes to complete. This in particular depends on the underlying network topology, namely, the existence of short paths between different vertices of the network. In practice, one may imagine that besides the network topology, there are other parameters that should be taken into account, such as, for example, the communication delays between nodes due to congestion. In this context, the spread of the information in the network can be thought of as a fluid penetrating the network reminiscent of the problem of first-passage percolation in a random medium. In this chapter, we will consider an asynchronous model in which each edge of the network is equipped with a random delay modeled by an exponential random variable with mean one.

One of the main motivations of our work comes from peer-to-peer networks. In particular, to motivate our random graph model, we recall that the most relevant properties of peer-to-peer networks are connectivity, small average degree, and approximate regularity of the degrees of the vertices. The random graph model considered in this chapter has these properties, and covers the classical $\mathcal{G}(n, r)$ model, which is the random graph model in which a graph is drawn uniformly at random from the set of n-vertex r-regular graphs, where r is a constant not depending on n. For this model of networks, we consider the push model for disseminating information. In this model, initially one of the nodes obtains some piece of information. Then every node which already has that information informs one other node chosen among its neighbors. The classical model goes iteratively and all nodes have the same clock. In each successive round, the nodes having the information choose independently and uniformly at random the neighbor they transmit to. In this chapter, we analyze an asynchronous randomized broadcast algorithm. Namely, nodes are not anymore assumed to be synchronized, so that each node has an independent Poisson clock. A node receiving the information will transmit it to a random neighbor at each tick of its own clock. When the graph of neighbors is a random r-regular graph, we show that the asynchronous version of the algorithm performs better than its synchronized version (see Section 2.3). To the best of our knowledge, our work is the first to study this model in an asynchronous version.

From a more theoretical point of view, our work contributes to the general theory of random graphs by providing new results for the weighted diameter of sparse random graphs. In [102], Janson considered the special case of the complete graph with fairly general i.i.d. weights on

edges, including the exponential distribution with parameter one. It is shown that, when n goes to infinity, the asymptotic distance for two given points is $\log n/n$, that the maximum distance if one point is fixed and the other varies is $2\log n/n$, and the maximum distance over all pairs of points (i.e., the weighted diameter) is $3\log n/n$. He also derived asymptotic results for the corresponding number of hops or hopcount (the number of edges on the paths with the smallest weight). More recently, a number of papers provide a detailed analysis of the scaling behavior of the joint distribution of the first passage percolation and the corresponding hopcount for the complete graph, e.g., [98, 21]. In particular, Bhamadi derives in [21], limiting distributions for the first passage percolation on both the complete graph and dense Erdős-Rényi random graphs with exponential and uniform i.i.d. weights on edges. This extends previous results by van der Hofstad et al. [93] exploring the link between the flooding time and first-passage percolation for both of these graphs with exponential edge-weights.

More closely related to the present work, Bhamidi et al. [24] study first passage percolation on random graphs with finite average degree, minimum degree greater than two and exponential weights, and derive explicit distributional asymptotic for the total weight of the shortest-weight path between two uniformly chosen vertices in the network. We compare their results to ours in the next section. The proofs will show that the analysis made in [24] is not sufficient to obtain results for the diameter. Indeed, we need to use large deviation techniques to control all the vertices and not only the uniformly chosen ones.

The remainder of the chapter is organized as follows. In the next section, we explain the model under consideration, introduce our notations, and state our main result together with applications to classical random graph models, namely, random regular graphs and Erdős-Rényi random graphs. Assuming these results, in Section 2.3, we restrict ourselves to the important class of random regular graphs and analyze a model for asynchronous randomized broadcast. We also compare in this section our results with the synchronized version. We then describe in Section 2.4 how to heuristically derive our main results using known properties of continuous-time Markov branching processes. This section is not technically required for the proof and is not written in a rigorous way. It is included to give some intuition for the proof of our main theorem which appears in the subsequent four sections. We prove the upper bound and the lower bound separately in Section 2.7 and Section 2.8, respectively.

2.2 Results

Let $G = (V, E)$ be a graph. Recall that for two vertices a and $b \in V$, a path between a and b is a sequence $\pi = (e_1, e_2, \dots e_k)$ where $e_i = \{v_{i-1}, v_i\} \in E$ and $v_i \in V$ for $i \in [k]$, with $v_0 = a$ and $v_k = b$. We write $e \in \pi$ if the edge $e \in E$ belongs to the path π, i.e., if $e = e_i$ for an $i \in [k]$. Recall that the distance between two nodes a and b in V, denoted by $\mathrm{dist}(a, b)$, is the number of edges in E in the shortest path connecting these two vertices. The diameter of G, denoted by $\mathrm{diam}(G)$, is the maximum graph distance between any pair of connected vertices in V, i.e.

$$\mathrm{diam}(G) = \max\{\mathrm{dist}(a, b),\ a, b \in V,\ \mathrm{dist}(a, b) < \infty\}\ .$$

For a graph G with vertex set V, the *flooding time* is defined by:

$$\mathrm{flood}(G) = \max\{\mathrm{dist}(a, b),\ b \in V,\ \mathrm{dist}(a, b) < \infty\}\ ,$$

where a is chosen uniformly at random in V. By an abuse of the notation, we also use $\mathrm{flood}(G)$ to denote the expectation of the random variable $\mathrm{flood}(G)$ over the choice of $a \in V$.

A weighted graph (G, w) is the data of a graph $G = (V, E)$ and a collection of weights w_e associated to each edge $e \in E$. We suppose that all the edge weights are non-negative. The weighted diameter and the flooding time of a weighted graph $\Big(G = (V, E), w = \{w_e\}_{e \in E}\Big)$ are defined similarly: For $a, b \in V$, the weighted distance between a and b is given by

$$\mathrm{dist}_w(a, b) = \min_{\pi \in \Pi(a,b)} \sum_{e \in \pi} w_e\ ,$$

where the minimum is taken over all the paths between a and b in the graph. The *weighted diameter* is given by

$$\mathrm{diam}_w(G) = \max\{\mathrm{dist}_w(a, b),\ a, b \in V,\ \mathrm{dist}_w(a, b) < \infty\}\ ,$$

while the *weighted flooding time* is

$$\mathrm{flood}_w(G) = \max\{\mathrm{dist}_w(a, b),\ b \in V,\ \mathrm{dist}_w(a, b) < \infty\}\ ,$$

where a is chosen uniformly at random in V.

From now on, we will consider a random graph $G(n, (d_i)_1^n)$, where the degree sequence $(d_i)_1^n$ satisfies Conditions 1.17 and 1.22 (of Chapter 1). We further assume the following two additional

conditions on the maximum and the minimum degree in the random graph $G(n,(d_i)_1^n)$. The first one is a slightly stronger condition than "the second moment property" in Condition 1.17. The second one merely says that the min-degree is well-defined at finite level.

Condition 2.1. (iv) For some $\tau > 0$, $\Delta_n := \max_{i\in V} d_i = O(n^{1/2-\tau})$;

(v) The *minimum degree* $d_{\min} := \min\{\, k \mid p_k > 0 \,\}$ is indeed the minimum degree for all n sufficiently large: For all $k < d_{\min}$, there is no $i \le n$ with $d_i = k$ provided that n is large enough, i.e., $u_k^{(n)} = 0$.

We recall that the size-biased probability mass function $q = \{q_k\}_{k=0}^{\infty}$ corresponding to $p = \{p_k\}_{k=0}^{\infty}$, is given by (c.f. Chapter 1)

$$q_k = \frac{(k+1)p_{k+1}}{\lambda}, \tag{2.1}$$

and its mean is denoted by ν, i.e., $\nu := \sum_{k=0}^{\infty} k q_k$.

The condition $\nu > 1$ is equivalent to the existence of a giant component in the configuration model, the size of which is proportional to n (see e.g. Theorem 1.24). We will assume that $\nu > 1$ in the rest of the chapter.

Since isolated vertices do not matter and can be removed from the graph, without loss of generality, let us also assume $u_0^{(n)} = 0$ too and consider thus the case $d_{\min} \ge 1$ in the rest of the chapter.

As for the analysis of the diameter of the random graphs, e.g., see Section 1.4.3, we need to introduce notations relevant to the branching process arising from the BFS (the (standard) breadth-first search) exploration of the graph. Let $G_p(z)$ and $G_q(z)$ be the probability generating functions of $\{p_k\}_{k=0}^{\infty}$ and $\{q_k\}_{k=0}^{\infty}$, respectively,

$$G_p(z) = \sum_{k=0}^{\infty} p_k z^k \qquad , \qquad G_q(z) = \sum_{k=0}^{\infty} q_k z^k \,. \tag{2.2}$$

Let $\mathcal{X}_q$ be a Galton-Watson Tree (GWT) with offspring distribution q. The extinction probability of the branching process, that we denote by β, is the smallest solution in $[0,1]$ of the fixed point equation $x = G_q(x)$, e.g., see Section 1.4.3. In addition, define $\beta_* = G_q'(\beta)$.

We can now announce our main theorem.

Theorem 2.2. *Consider a random graph $G(n,(d_i)_1^n)$ with i.i.d. rate one exponential weights on its edges. Suppose that the degree sequences $(d_i)_1^n$ satisfy* Conditions 1.17, 1.22, and 2.1. *We have*

$$\frac{\mathrm{diam}_w(G(n,(d_i)_1^n))}{\log n} \xrightarrow{p} \frac{1}{\nu-1} + \frac{2}{d_{\min}}\mathbb{1}_{(d_{\min}\geq 3)} + \frac{1}{(1-q_1)}\mathbb{1}_{(d_{\min}=2)} + \frac{2}{1-\beta_*}\mathbb{1}_{(d_{\min}=1)},$$

and

$$\frac{\mathrm{flood}_w(G(n,(d_i)_1^n))}{\log n} \xrightarrow{p} \frac{1}{\nu-1} + \frac{1}{d_{\min}}\mathbb{1}_{(d_{\min}\geq 3)} + \frac{1}{2(1-q_1)}\mathbb{1}_{(d_{\min}=2)} + \frac{1}{1-\beta_*}\mathbb{1}_{(d_{\min}=1)}.$$

This result is obviously consistent with the FPP asymptotic [24], where the typical weighted distance on random graphs with finite average degree is studied. In fact, the authors derive the following explicit distributional asymptotic for the minimum weight, dist_w, between two uniformly chosen vertices in the network, as well as for the *hopcount* (the number of edges in the shortest path connecting these two nodes).

Theorem 2.3 (Bhamidi, van der Hofstad and Hooghiemstra [24])**.** *Consider the configuration model $G(n,F)$ with i.i.d. degrees with common distribution function F of a random variable D (the degree random variable). Assume that F is non-degenerate, in the sense that it satisfies $F(x)=0$ for $x<2$ (so $d_{\min}\geq 2$), and that there exist $c>0$ and $\tau>3$ such that*

$$1-F(x)\leq cx^{-(\tau-1)},\ x\geq 0.$$

Let a,b be two uniformly chosen vertices in this graph. There exists a random variable V such that

$$\mathrm{dist}_w(a,b) - \frac{\log n}{\nu-1} \xrightarrow{d} V, \tag{2.3}$$

and the hopcount H_n between a and b satisfies the CLT

$$\frac{H_n(a,b)-\alpha\log n}{\sqrt{\alpha\log n}} \xrightarrow{d} Z, \tag{2.4}$$

where Z has a standard normal distribution, and $\alpha=\frac{\nu}{\nu-1}\in(1,\infty)$

(We refer to [24] and [23], for the results concerning the typical weighted distance and hopcount in the case $\tau\in(2,3)$ and $\tau\in(1,2)$, respectively.)

The appearance of the common factor $\frac{\log n}{\nu-1}$ in the above results is quite easy to understand at a heuristic level. Indeed, if one explores the neighborhood of a given vertex consisting of all vertices at (weighted) distance less than t, then this exploration process behaves like a continuous time Markov branching process (e.g., see Section 1.2.2 for a precise definition) which is known to grow exponentially fast like $e^{(\nu-1)t}$. In particular, at time $\frac{\log n}{2(\nu-1)}$, it reaches a size of the order $\sqrt{n}$. This means that if one considers two such exploration processes started separately from a and b, by that time the two processes should intersect with large probability. This explains why the typical weighted distance is of the order $\frac{\log n}{\nu-1}$. But of course, if one is interested in finding the diameter, one needs a more refined argument to control the behavior of all the vertices with respect to the exploration process. We give a more precise statement of the above heuristic in Section 2.4, deducing our main theorem at a heuristic level. When considering the weighted flooding time, we consider a case where one exploration process is started from a typical vertex whereas the other starting point is chosen in order to get a bad scenario in the sense that the exploration process started from this vertex grows slowly. Indeed, the bad scenario corresponds to a starting point having degree $d_{\min}$. This event gives the additional contribution of

$$\left(\frac{1}{d_{\min}}\mathbb{1}_{(d_{\min}\geq 3)}+\frac{1}{2(1-q_1)}\mathbb{1}_{(d_{\min}=2)}+\frac{1}{1-\beta_*}\mathbb{1}_{(d_{\min}=1)}\right)\log n \tag{2.5}$$

to the typical distance $\frac{\log n}{\nu-1}$. We refer to Section 2.8 for a formal treatment of this argument. Of course, to compute the weighted diameter, one has to consider a case where both starting points correspond to bad scenarios, obtaining then the total additional contribution of

$$\left(\frac{2}{d_{\min}}\mathbb{1}_{(d_{\min}\geq 3)}+\frac{1}{(1-q_1)}\mathbb{1}_{(d_{\min}=2)}+\frac{2}{1-\beta_*}\mathbb{1}_{(d_{\min}=1)}\right)\log n.$$

We see that if one is interested in passing the information between two typical vertices, this task can be achieved in time of the order $\frac{\log n}{\nu-1}$, and there is a price (in time) given by (2.5) to pay if one wishes to pass the information to everyone from a typical vertex and another price (in time), given by (2.5), to pay if one wishes to pass the information to everyone from a vertex in a worst case scenario.

We now instantiate our main theorem for some graphs of interest, namely random regular graphs, and sparse Erdős-Rényi random graphs.

Random regular graphs

Let $r \geq 3$ be a fixed integer, and let $G \sim \mathcal{G}(n,r)$ be a random r-regular graph with n vertices. The diameter of random regular graphs with exponential weights was studied in [52]. As a corollary of Theorem 2.2, we can recover their result, namely

Corollary 2.4. *Fix $r \geq 3$, and let $G \sim \mathcal{G}(n,r)$ be a random r-regular graph with n vertices and i.i.d. rate 1 exponential random variables on its edges. We have w.h.p.*

$$\frac{\mathrm{diam}_w(G)}{\log n} \xrightarrow{p} \frac{1}{r-2} + \frac{2}{r}, \text{ and} \tag{2.6}$$

$$\frac{\mathrm{flood}_w(G)}{\log n} \xrightarrow{p} \frac{1}{r-2} + \frac{1}{r}. \tag{2.7}$$

This result allows us to analyze an asynchronous randomized broadcast algorithm for random regular graphs in Section 2.3

Erdős-Rényi random graphs

Consider the Erdős-Rényi random graph $\mathrm{ER}(n,\lambda/n)$ on n vertices where each possible edge is chosen independently at random with probability λ/n. We assume $\lambda > 1$ to ensure the existence of a giant component with high probability. In this case we have

$$q_k = e^{-\lambda}\frac{\lambda^k}{k!}. \tag{2.8}$$

The Conditions 1.17, 1.22 and 2.1 both hold a.s. (by conditioning on the vertex degrees, c.f., Remark 1.23), and we have $\nu = \lambda$ and $\beta_* = \lambda_*$ where λ_* is the solution $\lambda_* < 1$ to the equation

$$\lambda_* e^{-\lambda_*} = \lambda e^{-\lambda}.$$

Applying Theorem 2.2 to this case, we obtain

Theorem 2.5. *Let $\lambda > 1$ be a real number, and consdier $\mathrm{ER}(n,\lambda/n)$ with i.i.d. rate one exponential weights on its edges. Then*

$$\frac{\mathrm{diam}_w(\mathrm{ER}(n,\lambda/n))}{\log n} \xrightarrow{p} \frac{1}{\lambda-1} + \frac{2}{1-\lambda_*}, \text{ and} \tag{2.9}$$

$$\frac{\mathrm{flood}_w(\mathrm{ER}(n,\lambda/n))}{\log n} \xrightarrow{p} \frac{1}{\lambda-1} + \frac{1}{1-\lambda_*}. \tag{2.10}$$

A lower bound for the weighted diameter in this case was given by Bhamidi, van der Hofstad and Hooghiemstra in [25]. Our theorem above improves their bound, and gives the correct asymptotic. It is worth mentioning that in [25], the authors derive a result similar to Theorem 2.3, i.e., they are able to remove the condition $d_{\min} \geq 2$ in Theorem 2.3 in the specific case of Erdős-Rényi random graphs.

2.3 Broadcasting in Random Regular Graphs

In this section we elaborate on another aspect of our result. By comparing our main Theorem 2.2 with Theorem 1.29 (in Chapter 1) in the case of random regular graphs, we see that the weighted flooding time (or diameter) is actually smaller than the graph distance flooding time (or diameter). With previous discussion, the heuristic explanation of this phenomenon is as follows. The random weights allow to introduce a non-zero variance that will result in a faster growth of the branching process (approximating the exploration process) compared to the constant weights. Indeed, even though the weights have an average of one, the weights with small values allow the branching process to grow faster than the case where all the weights are equal to one. Of course the variability of the weights has also a drawback when one looks at the worst case scenario which corresponds to the factors $\frac{\log n}{d_{\min}}$, in the case $d_{\min} \geq 3$. However in the case of random regular graphs, the advantages of variance exceeds its drawback, and the weighted flooding time is smaller than the graph-distance flooding time. Note that this will not be always true in the general case, e.g., when ν is much larger than $d_{\min}$. We now concentrate on one important practical implications of this phenomenon.

We consider the asynchronous analogue of the standard *phone call model* [136]. In continuous-time, we assume that each node is endowed with a Poisson process with rate one, and that at the instants of its corresponding Poisson process, a node wakes up and contacts one of its neighbors uniformly at random. We consider the well-studied push model. In this model, if a node i holds the message, it passes the message to its randomly chosen neighbor regardless of its state. Note that this may yield an unnecessary transmission (if the receiver had already the message or the transmitter does not hold the message). As in the case of the standard discrete-time phone call model, we are interested in the performance of such an information dissemination routine in terms of the time it takes to inform the whole population. We denote this time by $\mathrm{ABT}(G)$ for Asynchronous Broadcast Time.

We restrict ourselves to r-regular graphs, i.e., graphs where each node has degree $r \geq 3$, so that $p_r = 1$ in Condition 1.17. As shown below, the dynamic evolution of informed nodes corresponds to the flooding time with i.i.d. weights on edges distributed according to an exponential distribution with mean r. The symmetry of the graph is crucial to get this property, and this is the reason why we require G to be an r-regular graph. Our Theorem 2.2 allows us to analyze the asynchronous broadcast algorithm for these graphs, and we get the following corollary.

Corollary 2.6. *Let $G \sim \mathcal{G}(n, r)$ be a random r-regular graph with n vertices. Then w.h.p.*

$$\mathrm{ABT}(G) = 2\left(\frac{r-1}{r-2}\right)\log n + o(\log n).$$

A proof of this corollary is given in Section 2.9.

The classical randomized broadcast model was first investigated by Frieze and Grimmett [76]. Given a graph $G = (V, E)$, initially a piece of information is placed on one of the nodes in V. Then in each time step, every informed node sends the information to another node, chosen independently and uniformly at random among its neighbors. The question now is how many time-steps are needed such that all nodes become informed. Note that this model requires nodes to be synchronized. It was shown by Frieze and Grimmett [76] and Pittel [136] that for the complete graph K_n the number of steps needed to inform the whole population scales as $\log_2 n + \log n + o(\log n)$ with high probability. Fountoulakis et al. [73] proved that in the case of Erdős-Rényi random graphs $G(n, p_n)$, if the average degree, np_n, is slightly larger than $\log n$, then the broadcast time essentially coincides with the broadcast time on the complete graph. For any r-regular graphs it has been shown in [58] that this algorithm requires at least $\left(\frac{1}{\log(2-1/r)} - \frac{1}{r\log(1-1/r)}\right)\log n + o(\log n)$ rounds to inform all nodes of the graph, w.h.p. (the randomness comes here from the choice of the neighbor to which the information is pushed). Fountoulakis and Panagtotou in [74] have recently shown that in the case of random regular graphs, the process completes in $\left(\frac{1}{\log(2(1-1/r))} - \frac{1}{r\log(1-1/r)}\right)\log n + o(\log n)$ rounds w.h.p.

Note that if instead of independent Poisson clocks of rate one, we take a deterministic process with slots of size one, the situation becomes exactly the one studied in [74]. Hence locally, both processes behave similarly. When a node receive the information, it will need on average a time of $\left(r + \frac{r}{2} + \cdots + \frac{r}{r-1} + 1\right)$ to transmit it to all its neighbors (including possibly informed ones). Figure 2.1 shows the comparison between results in [74] and our Theorem 2.6. In both cases, the time to broadcast is of the order of $\log n$ but the prefactors differ and are given by the two curves for various values of r. We see that the asynchronous version is always faster than

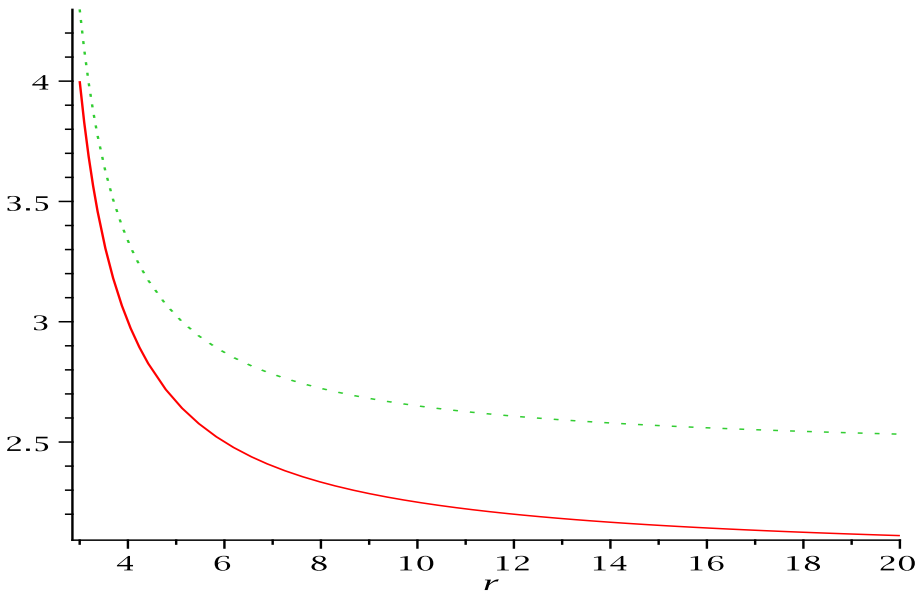

Figure 2.1: Comparison of the time to broadcast in the synchronized version (dashed) and with exponential random weights (plain)

the synchronized one. This result while surprising is in agreement with the previous discussion comparing our Theorem 2.2 with Theorem 1.29. The process of diffusion takes advantage of the variance of the exponential random delays and allows to broadcast the information faster in a decentralized and asynchronous way!

2.4 Branching Process Approximation

In this section, we provide heuristics for the behaviour of the diameter of the random graph $G_n \sim G(n,(d_i)_1^n)$, with minimum degree $d_{\min}$, endowed with i.i.d. Exp(1) (exponential one) edge-weights. To this end, we analyse the Markovian continuous-time branching process $Z(t)$ defined as follows (see also Section 1.2.2). The root has lifespan distributed according to Exp(1), and at her death, gives birth to D children where D is distributed according to p. The subsequent generations follow the same dynamics except that the offspring distribution is given by q defined in Equation (2.1). It is well known that $Z(t)e^{-(\nu-1)t}$ converges almost surely to a random variable W, e.g., see Theorem 1.9 (in Chapter 1).

In particular, Theorem 1.10 implies the following [24, Proposition C.4].

Proposition 2.7. *The limiting random variable W is given by*

$$W = \sum_{i=1}^{D} \widetilde{W}_i e^{-(\nu-1)\xi_i} ,$$

where D is distributed according to p, the ξ_i's are i.i.d. Exp(1)*, independent of the $\widetilde{W}_i$ which are i.i.d. with Laplace transform $\widetilde{\phi}(t) = \mathbb{E}\left(e^{-t\widetilde{W}}\right)$ whose inverse function is given by*

$$\widetilde{\phi}^{-1}(x) = (1-x)\exp\left(\int_1^x \frac{\nu-1}{G_q(s)-s} + \frac{1}{1-s}\, ds\right), \quad \beta < x \leq 1. \tag{2.11}$$

Given two nodes u and v in V, we grow two balls exploring their respective neighborhoods. We assume both of these processes follow a dynamics similar to the continuous-time branching process $Z(t)$. This assumption will be made rigorous later when proving our main result. Our heuristic is based on the rationale that if the two processes reach population sizes $\sqrt{n}$, then there is a high probability that the two balls will intersect. More precisely, we need to find x such that

$$\mathbb{P}\left(Z(x\log n) < \sqrt{n}\right) \approx n^{-1} , \tag{2.12}$$

and then we can approximate the diameter by $2x\log n$.

Using Proposition 2.7, it is reasonable to analyse $\mathbb{P}\left(W < n^{1/2}n^{-x(\nu-1)}\right)$. In what follows, we focus on deriving the behavior of $\mathbb{P}(W<\epsilon)$ for ϵ going to 0. Note that

$$\begin{aligned}\psi(s) := \mathbb{E}(e^{-sW}) &= \mathbb{E}\left(\left[\mathbb{E}\left(\exp\left(-s\widetilde{W}e^{-(\nu-1)\xi}\right)\right)\right]^D\right)\\ &\approx p_{d_{\min}}\left[\mathbb{E}\left(\exp\left(-s\widetilde{W}e^{-(\nu-1)\xi}\right)\right)\right]^{d_{\min}},\end{aligned}$$

where ξ is an Exp(1) random variable and $\psi(.)$ is the Laplace transform of W. Moreover,

$$\mathbb{E}\left(\exp\left(-s\widetilde{W}e^{-(\nu-1)\xi}\right)\right) = \int_0^\infty e^{-x}\widetilde{\phi}\left(se^{-(\nu-1)x}\right) dx .$$

Combining Equation (2.11) together with the Tauberian theorem [67, Section XIII.5], we infer that

$$\mathbb{P}(W<\epsilon) \approx \epsilon^{\alpha}, \quad \alpha = \begin{cases} \frac{d_{\min}}{\nu-1}, & \text{if } d_{\min}\geq 3,\\ \frac{2(1-q_1)}{\nu-1}, & \text{if } d_{\min}=2,\\ \frac{1-\beta_*}{\nu-1}, & \text{if } d_{\min}=1.\end{cases}$$

As a byproduct, we can see that the diameter of $G(n,(d_i)_1^n)$, with i.i.d. exponential one, can be approximated by $2x\log n$, where x, given by Equation (2.12), satisfies

$$\alpha(1/2 - x(\nu-1)) = -1.$$

We conclude that the diameter can be approximated by

$$\left(\frac{1}{\nu-1} + \frac{2}{d_{\min}}\mathbb{1}_{(d_{\min}\geq 3)} + \frac{1}{(1-q_1)}\mathbb{1}_{(d_{\min}=2)} + \frac{2}{1-\beta_*}\mathbb{1}_{(d_{\min}=1)}\right)\log n.$$

We will show in the following four sections that this is indeed the correct asymptotic. We end this section by giving now an overview of the proof.

There will be three different cases to consider depending on whether $d_{\min}\geq 3$, $d_{\min}=2$, or $d_{\min}=1$. It will be convenient to consider

$$s_n := \left(\frac{1}{d_{\min}}\mathbb{1}_{(d_{\min}\geq 3)} + \frac{1}{2(1-q_1)}\mathbb{1}_{(d_{\min}=2)} + \frac{1}{1-\beta_*}\mathbb{1}_{(d_{\min}=1)}\right)\log n, \tag{2.13}$$

which encodes all the three cases above.

When $d_{\min}=1$, the longest shortest path in a random graph is known to be between a pair of vertices a and b of degree one. Furthermore, this path consists of a path from a to the 2-core, a path through the 2-core, and a path from the 2-core to b (for the non-weighted case, see [69]). For this, we need to provide some preliminary results on the structure of the 2-core, this is done in the next section. We then consider in Section 2.6 a certain process, called exploration process, which consists in growing balls simultaneously from each vertex. The diameter will be the time the last pair of balls intersect. A precise treatment of the exploration process, resulting in information about the growth rates of the balls are given in this section. In addition, the section provides some necessary notations and definitions that will be used throughout the last three sections. Sections 2.7 and 2.8 form the heart of the proof. We first prove that the above bound is an upper bound for the diameter. This will consist in defining the two parameters α_n and β_n with the following significance. (i) Two balls of size at least β_n intersect almost surely, (ii) considering the growing balls centered at a vertex in the graph, the time it takes for the balls to go from size α_n to size β_n have all the same asymptotic for all the vertices of the graph, and the asymptotic is half of the typical weighted distance in the graph, and (iii) the time it takes for the growing balls centered at a given vertex to reach size at least α_n is upper bounded by $(1+\epsilon)s_n$ for all $\epsilon>0$ w.h.p . This will show that the diameter is w.h.p. bounded above by $(1+\epsilon)(\frac{1}{\nu-1}\log n + 2s_n)$, for all $\epsilon>0$. The last section provides the corresponding lower bound.

To obtain the lower bound, we show that w.h.p. (iv) there are at least two nodes with degree $d_{\min}$ such that the time it takes for the balls centered at these vertices to achieve size at least α_n is worst than the other vertices, and is lower bounded by $(1-\epsilon)s_n$, for all $\epsilon > 0$. And using this, we conclude that the diameter is w.h.p. bounded below by $(1-\epsilon)(\frac{1}{\nu-1}\log n + 2s_n)$, for all fixed $\epsilon > 0$, finishing the proof of our main theorem.

The actual values of α_n and β_n will be

$$\alpha_n := \log^3 n, \text{ and } \beta_n := 3\sqrt{\frac{\lambda}{\nu-1} n \log n}. \tag{2.14}$$

2.5 Structure of the Augmented 2-Core

Recall that the *k-core* of a given graph G is the largest induced subgraph of G with minimum vertex degree at least k.

Consider now a random graph $G_n \sim G^*(n,(d_i)_1^n)$. In the process of constructing a random graph G_n by matching the half-edges, the k-core can be found by successively removing the half-edge of a node of degree less than k followed by removing a uniformly random half-edge from the set of all the remaining half-edges until no such vertices (of degree less than k) remain. What remains at this time is the k-core. Since these half-edges are unexposed, the k-core edge set is uniformly random conditional on the k-core half-edge set (e.g., see Section 1.4.2).

From now on, we consider the case $k = 2$, and denote by $\tilde{G}$ the 2-core of a graph G. In particular applying Theorem 1.26 to the case $k = 2$, we have

$$\begin{aligned} h(\hat{p}) &:= \mathbb{E}[D_{\hat{p}} \mathbb{1}(D_{\hat{p}} \geq 2)] \\ &= \mathbb{E}[D_{\hat{p}}] - \mathbb{P}(D_{\hat{p}} = 1) \\ &= \lambda\hat{p} - \sum_l l\ p_l\ \hat{p}(1-\hat{p})^{l-1} \\ &= \lambda\hat{p}\ (1 - G_q(1-\hat{p})). \end{aligned}$$

Recall from Theorem 1.26 that we have to solve the equation $\lambda\hat{p}^2 = h(\hat{p})$, thus, we obtain $1-\hat{p} = G_q(1-\hat{p})$, and so $\hat{p} = 1-\beta$.

We conclude that the graph $\tilde{G}_n$ obtained from G_n has the same distribution as a random graph constructed by the configuration model on $\tilde{n}$ nodes with a degree sequence $\tilde{d}_1^{(n)}, ..., \tilde{d}_{\tilde{n}}^{(n)}$

satisfying the following properties. By Theorem 1.26,

$$\begin{aligned}\tilde{n}/n \xrightarrow{p} h_1(1-\beta) &:= \mathbb{P}[D_{1-\beta} \geq 2] \\ &= 1 - G_p(\beta) - (1-\beta)G_p'(\beta) \\ &= 1 - G_p(\beta) - \lambda\beta(1-\beta) > 0,\end{aligned}$$

and

$$\begin{aligned}|\{i, \tilde{d}_i^{(n)} = j\}|/n &\xrightarrow{p} \sum_{\ell=j}^{\infty} p_\ell \binom{\ell}{j} (1-\beta)^j \beta^{\ell-j},\ j \geq 2, \\ \sum_i \tilde{d}_i^{(n)}/n &\xrightarrow{p} \lambda(1-\beta)^2.\end{aligned}$$

It follows that the sequence $\{\tilde{d}_1^{(n)}, ..., \tilde{d}_{\tilde{n}}^{(n)}\}$ satisfies also the Condition 1.17, 1.22 and 2.1 for some probability distribution $\tilde{p}_k$ with mean $\tilde{\lambda}$ (which can be easily calculated from the two above properties).

Let $\tilde{q}$ be the size-biased probability mass function corresponding to $\tilde{p}$. We now show that $\tilde{q}$ and q have the same mean. Indeed, denoting by $\tilde{\nu}$ the mean of $\tilde{q}$, we see that $\tilde{\nu}$ is given by

$$\begin{aligned}\tilde{\nu} &:= \sum_k k\tilde{q}_k = \frac{1}{\tilde{\lambda}} \sum_k k(k-1)\tilde{p}_k \\ &= \frac{\sum_{k\geq 2} k(k-1) \sum_{\ell \geq k} p_\ell \binom{\ell}{k} (1-\beta)^k \beta^{\ell-k}}{\lambda(1-\beta)^2} \\ &= \frac{\sum_\ell p_\ell \sum_{k \leq \ell} k(k-1) \binom{\ell}{k} (1-\beta)^k \beta^{\ell-k}}{\lambda(1-\beta)^2} \\ &= \frac{\sum_\ell p_\ell \ell\ (\ell-1)}{\lambda} = \nu. \end{aligned} \tag{2.15}$$

To find the diameter in the case $d_{\min} = 1$, we also need to show that $\tilde{q}_1 = \beta_*$:

$$\begin{aligned}\tilde{q}_1 = \frac{2\tilde{p}_2}{\tilde{\lambda}} &= \frac{2\sum_{\ell \geq 2} p_\ell \binom{\ell}{2} (1-\beta)^2 \beta^{\ell-2}}{\lambda(1-\beta)^2} \\ &= \frac{1}{\lambda} G_p''(\beta) = G_q'(\beta) = \beta_*. \end{aligned} \tag{2.16}$$

We will also need the following relaxation of the notion of 2-core. Let $G = (V, E)$ be a graph. For a given subset $W \subseteq V$, define the W-augmented 2-core to be the maximal induced subgraph of G such that every vertex in $V \setminus W$ has degree at least two, i.e., the vertices in W

are not required to verify the minimum degree condition in the definition of the 2-core. The W-augmented 2-core of a graph G will be denoted by $\tilde{G}(W)$.

It is easy to see that the W-augmented 2-core of a random graph $G_n \sim G^*(n,(d_i)_1^n)$, denoted by $\tilde{G}_n(W)$, can be found in the same way as the 2-core, except that now the termination condition is that every nodes outside of W must have degree at least 2, since the half-edges adjacent to a vertex in W are exempt from this restriction. The conditional uniformity property thus evidently holds in this case as well, i.e., for any subset $W \subset V$, the W-augmented 2-core is uniformly random conditional on the W-augmented 2-core half-edge set. We will need the following basic result, the proof of which is easy and can be found for example in [69, Lemma A.7].

Lemma 2.8. *Consider a random graph $G_n \sim G(n,(d_i)_1^n)$ where the degree sequence $(d_i)_1^n$ satisfies Condition 1.17. For any subset $W \subset V(G_n)$, and any $w \in W$, there exists $C > 0$ (sufficiently large) so that we have*

$$\mathbb{P}\Big(e(\tilde{G}_n(W)) - e(\tilde{G}_n(W \setminus \{w\})) \le C \log n\Big) = 1 - o(n^{-1}).$$

Note that the above lemma implies (by removing one vertex from W at a time) that if $|W| = o(n/\log n)$, then w.h.p. the two graphs $\tilde{G}_n$ and $\tilde{G}_n(W)$ have the same degree distribution asymptotic.

2.6 First Passage Percolation in $G^*(n,(d_i)_1^n)$

We start this section by introducing some new notations and definitions. Before this, one remark is in order. In what follows, we will sometimes deliberately use the term "time" instead of the term "weighted distance". It will be clear from the context what we actually mean by this.

Let $\big(G = (V,E), w\big)$ be a weighted graph. For a vertex $a \in V$ and a real number $t > 0$, the t-radius neighborhood of a in the (weighted) graph, or the ball of radius t centered at a, is defined as

$$B_w(a,t) := \big\{\, b,\ \mathrm{dist}_w(a,b) \le t \,\big\}.$$

The first time t where the ball $B_w(a,t)$ reaches size $k+1 \ge 1$ will be denoted by $T_a(k)$, i.e.,

$$T_a(k) = \min\big\{\, t : |B_w(a,t)| \ge k+1 \,\big\}, \qquad T_a(0) = 0.$$

If there is no such t, i.e., if the component containing a has size at most k, we define $T_a(k) = \infty$. Note that there is a vertex in $B_w(a, T_a(k))$ which is not in any ball of smaller radius around a. When the weights are i.i.d. according to a random variable with continuous density, this vertex is in addition unique with probability one. We will assume this in what follows.

We use I_a to denote the size of the component containing a in the graph minus one, in other words,

$$I_a := \max\big\{\, |B_w(a,t)|, t \geq 0 \,\big\} - 1, \tag{2.17}$$

For an integer $i \leq I_a$, we use $\widehat{d}_a(i)$ to denote the forward-degree of the (unique) node added at time $T_a(i)$ in $B_w(a, T_a(i))$. Recall that the forward-degree is the degree minus one. Define $\widehat{S}_a(i)$ as follows.

$$\widehat{S}_a(i) := d_a + \widehat{d}_a(1) + ... + \widehat{d}_a(i) - i, \qquad \hat{S}_a(0) = d_a. \tag{2.18}$$

For a connected graph H, the tree excess of H is denoted by $\mathrm{tx}(H)$, which is the maximum number of edges that can be deleted from H while still keeping it connected. By an abuse of the notation, for a subset $W \subseteq V$, we denote by $\mathrm{tx}(W)$ the tree excess of the induced subgraph $G[W]$ of G on W. (If $G[W]$ is not connected, then $\mathrm{tx}(W) := \infty$.) Consider the growing balls $B_w(a, T_a(i))$ for $0 \leq i \leq I_a$ centered at a and define $X_a(i)$ as the tree excess of $B_w(a, T_a(i))$, i.e.,

$$X_a(i) := \mathrm{tx}\,(\, B_w(a, T_a(i))\,).$$

We extend the definition of X_a to all the integer values by setting $X_a(i) = X_a(I_a)$ for all $i > I_a$.

The number of edges crossing the boundary of the ball $B_w(a, T_a(i))$ is denoted by $S_a(i)$. A simple calculation shows that

$$S_a(i) = \widehat{S}_a(i) - 2X_a(i). \tag{2.19}$$

We now consider a random graph $G(n, (d_i)_1^n)$ with i.i.d. rate one exponential weights on its edges, such that the degree sequence $(d_i)_1^n$ satisfies Conditions 1.17, 1.22, and 2.1. One particularly useful property of the configuration model is that it allows one to construct the graph gradually, exposing the edges of the perfect matching one at a time. This way, each additional edge is uniformly distributed among all possible edges on the remaining (unmatched) half-edges.

We have the following useful lemma.

Lemma 2.9. *For any $k \le \frac{m^{(n)}-n}{2}$, we have*

$$\mathbb{P}\left(2X_a(k) \ge x \mid \widehat{S}_a(k), I_a \ge k\right) \le \mathbb{P}\left(\mathrm{Bin}\left(\widehat{S}_a(k), \sqrt{\widehat{S}_a(k)/n}\right) \ge x \mid \widehat{S}_a(k), I_a \ge k\right).$$

To prove this, we need the following intermediate result proved in [69, Lemma 3.2]. We recall here that for two real-valued random variables A and B, we say A is stochastically dominated by B and write $A \le_{st} B$ if for all x, we have $\mathbb{P}(A \ge x) \le \mathbb{P}(B \ge x)$ (e.g., see Section A.2). If C is another random variable, we write $A \le_{st} (B \,|\, C)$ if for all x, $\mathbb{P}(A \ge x) \le \mathbb{P}(B \ge x \,|\, C)$ a.s.

Lemma 2.10. *Let M be a set of m points (m is even number), i.e., $|M| = m$, and let F be a uniform random matching of elements of M. For $e \in M$, we denote by $F(e)$ the point matched to e, and similarly for $X \subset M$, we write $F(X)$ for the set of points matched to X. Now let $X \subset M$, $k = |X|$, and assume $k \le m/2$. We have*

$$|X \cap F(X)| \le_{st} \mathrm{Bin}(k, \sqrt{k/m}).$$

Proof. For $m = |M|$, and $k = |X|$, we let

$$f(m,k,x) := \mathbb{P}(|X \cap F(X)| \ge x).$$

We also define

$$g(k,p,x) := \mathbb{P}(\mathrm{Bin}(k,p) \ge x).$$

Hence to prove the lemma, we have to show that $f(m,k,x) \le g(k, \sqrt{k/m}, x)$ holds for all x.

For $k = 0, 1$, and any x and $m \ge 2k$, clearly $f(m,k,x) \le g(m, \sqrt{k/m}, x)$ holds, since for $k \le 1$, there cannot be any internally matched endpoints, i.e., $|X \cap F(X)| = 0$. We now proceed by induction on k. Assume that $f(m',k',x) \le g(m', \sqrt{k'/m'}, x)$ holds for all x, $k' < k$, and $m' \ge 2k'$. We will show that the inequality holds for k and any $m \ge 2k$.

Note that any given point $e \in X$ matches to another point in X with probability $\frac{k-1}{m-1}$. In this case, we have $k-2$ yet-unmatched points in X, and we have two internally matched points. Otherwise, we have $k-1$ remaining yet-unmatched points in X, and no internally matched points. In both cases, the total number of yet-unmatched points is $m-2$. Hence, we inductively

obtain

$$\begin{aligned} f(m,k,x) &= \frac{k-1}{m-1} f(m-2,k-2,x-2) + \frac{m-k}{m-1} f(m-2,k-1,x) \\ &\le \frac{k-1}{m-1} g(k-2,\sqrt{\frac{k-2}{m-2}},x-2) + \frac{m-k}{m-1} g(k-1,\sqrt{\frac{k-1}{m-2}},x) \\ &\le \frac{k}{m} g(k-2,\sqrt{\frac{k}{m}},x-2) + \left(1-\frac{k}{m}\right) g(k-1,\sqrt{\frac{k}{m}},x). \end{aligned}$$

We conclude that $f(m,k,x) \le g(k,\sqrt{k/m},x)$ by letting $p=\sqrt{k/m}$ in the following inequality

$$\begin{aligned} g(k,p,x) &= p\cdot g(k-1,p,x-1) + (1-p)\cdot g(k-1,p,x) \\ &= p^2\cdot g(k-2,p,x-2) + p(1-p)\cdot g(k-2,p,x-1) + (1-p)\cdot g(k-1,p,x) \\ &\ge p^2\cdot g(k-2,p,x-2) + (1-p^2)\cdot g(k-1,p,x). \end{aligned}$$

□

Proof of Lemma 2.9. Conditioning on all the possible degree sequences $\widehat{d}_a(1),\widehat{d}_a(2),\ldots,\widehat{d}_a(k)$, with the property that $d_a+\sum_{1\le i\le k}\widehat{d}_a(i)=\widehat{S}_a(k)$, the configuration model becomes equivalent to the following process: start from a and at each step $1\le i\le k$, choose a vertex a_i of degree $\widehat{d}_a(i)+1$ uniformly at random from all the possible vertices of this degree outside the set $\{a,a_1,\ldots,a_{i-1}\}$, choose a half-edge adjacent to a_i uniformly at random and match it with a uniformly chosen half-edge from the yet-unmatched half-edges adjacent to one of the nodes $a,a_1,\ldots,a_{i-1}$. And at the end, after a_k has been chosen, take a uniform matching for all the remaining $(m^{(n)}-2k)$ half-edges. Now the proof follows from Lemma 2.10 by the simple observation that, since $m^{(n)}-2k\ge n$,

$$\mathbb{P}\left(\mathrm{Bin}\left(\widehat{S}_a(k),\sqrt{\widehat{S}_a(k)/(m^{(n)}-2k)}\right)\ge x \mid \widehat{S}_a(k)\right) \le \mathbb{P}\left(\mathrm{Bin}\left(\widehat{S}_a(k),\sqrt{\widehat{S}_a(k)/n}\right)\ge x \mid \widehat{S}_a(k)\right).$$

□

In the sequel, we will also need to consider the number of vertices of forward-degree at least two in the (growing) balls centered at a vertex $a\in V$. Thus, for $i\le I_a$, define

$$\gamma_a(i) := \sum_{\ell=1}^{i} \mathbb{1}(\widehat{d}_a(\ell)\ge 2) = \left|\left\{b\in B_w(a,T_a(i)):\ b\ne a \text{ and } d_b\ge 3\right\}\right|, \tag{2.20}$$

and extend the definition to all integers by setting $\gamma_a(i) = \gamma_a(I_a)$ for all $i > I_a$. Note that $\gamma_a(0) = 0$.

Now define $\overline{T}_a(k)$ to be the first time where the ball centered at a has at least k nodes of forward-degree at least two. More precisely,

$$\overline{T}_a(i) \quad := \quad \min\Big\{T_a(\ell), \ \text{ for } \ell \text{ such that } \gamma_a(\ell) \geq k\Big\}. \tag{2.21}$$

As we explained before, the main idea of the proof of Theorem 2.2 consists in growing the balls around each vertex of the graph simultaneously so that the diameter becomes equal to twice the time when the last two balls intersect. In what follows, instead of taking a graph at random and then analyzing the balls, we use a standard coupling argument in random graph theory which allows to build the balls and the graph at the same time. We present this coupling in the next coming section.

2.6.1 The exploration process

Consider a random graph $G \sim G^*(n,(d_i)_1^n)$, with i.i.d. rate 1 exponential variables on its edges. Fix a vertex a in G, and consider the following continuous-time exploration process. At time $t = 0$, we have a neighborhood consisting only of a, and for $t > 0$, the neighborhood is precisely $B_w(a,t)$. We now give an equivalent description of this process. This provides a more convenient way for analyzing the random variables which are crucial in our argument, e.g., $S_a(k)$. The idea is that instead of taking a graph at random and then analyzing the balls, the graph and the balls are built at the same time. We will consider a growing set B and a list L of yet unmatched half-edges. Recall that in the usual way of constructing a random graph with given degree sequence, we match half-edges amongst themselves uniformly at random. In the following, by a matching, we mean a pair of matched half-edges.

- Start with $B = \{a\}$, where a has d_a half-edges. For each half edge, decide (at random depending on the previous choices) if the half-edge is matched to a half-edge adjacent to a or not. Reveal the matchings consisting of those half-edges adjacent to a which are connected amongst themselves (creating self-loops at a) and assign weights independently at random to these edges. The remaining unmatched half-edges adjacent to a are stored in a list L. (See the next step including a more precise description of this first step.)
- Repeat the following exploration step as long as the list L is not empty.

Given there are $\ell \geq 1$ half-edges in the current list, say $L = (h_1, \ldots, h_\ell)$, let $\Psi \sim \mathrm{Exp}(\ell)$ be an exponential variable with mean ℓ^{-1}. After time Ψ select a half-edge from L uniformly at random, say h_i. Remove h_i from L and match it to a uniformly chosen half-edge in the entire graph excluding L, say h. Add the new vertex (connected to h) to B and reveal the matchings (and weights) of any of its half-edges whose matched half-edge is also in B. More precisely, let d be the degree of this new vertex and $2x$ the number of already matched half-edges in B (including the matched half-edges h_i and h). There is a total of $m-2x$ unmatched half-edges, m being the total number of half-edges of the random graph G. Consider one of the $d-1$ half-edges of the new vertex (excluding h which is connected to h_i); with probability $(\ell-1)/(m-2x-1)$ it is matched with a half-edge in L and with the complementary probability it is matched with an unmatched half-edge outside L. In the first case, match it to a uniformly chosen half-edge of L and remove the corresponding half-edge from L. In the second case, add it to L. We proceed in the similar manner for all the $d-1$ half-edges of the new vertex.

Let $B(a,t)$ and $L(a,t)$ be respectively the set of vertices and the list generated by the above procedure at time t. Considering the usual configuration model and using the memoryless property of the exponential distribution, we have $B_w(a,t) = B(a,t)$ for all t. To see this, we can continuously grow the weights of the half-edges $h_1, \ldots, h_\ell$ in L until one of their rate 1 exponential clocks fire. Since the minimum of ℓ i.i.d exponential variables with rate 1 is exponential with rate ℓ, this is the same as choosing uniform half-edge h_i after time Ψ (recall that by our conditioning, these ℓ half-edges do not pair within themselves). Note that the final weight of an edge is accumulated between the time of arrival of its first half-edge and the time of its pairing (except edges going back into B whose weights are revealed immediately). Then the equivalence follows from the memoryless property of the exponential distribution.

Note that $T_a(i)$ is the time of the i-th exploration step in the above continuous-time exploration process. Assuming $L(a, T_a(i))$ is not empty, at time $T_a(i+1)$, we match a uniformly chosen half-edge from the set $L(a, T_a(i))$ to a uniformly chosen half-edge among all other half-edges, excluding those in $L(a, T_a(i))$. Let $\mathcal{F}_t$ be the σ-field generated by the above process until time t. Given $\mathcal{F}_{T_a(i)}$, $T_a(i+1) - T_a(i)$ is an exponential random variable with rate $S_a(i) = |L(a, T_a(i))|$ the size of the list consisting of unmatched half-edges in $B(a, T_a(i))$. In other words,

$$\big(T_a(i+1) - T_a(i) \,|\, \mathcal{F}_{T_a(i)}\big) = \mathrm{Exp}(S_a(i)),$$

this is true since the minimum of k i.i.d. rate one exponential random variables is an exponential of rate k.

Recall that $I_a = \min\{i,\, S_a(i) = 0\} \leq n-1$, and set $S_a(i) = 0$ for all $I_a \leq i \leq n-1$. We now extend the definition of the sequence $\widehat{d}(i)$ to all the values of $i \leq n-1$, constructing the sequence $(\widehat{d}_a(i))_{i=1}^{n-1}$ which will coincide in the range $i \leq I_a$ with the sequence $\widehat{d}_a(i)$ defined in the previous subsection. We first note that in the terminology of the exploration process, the sequence $(\widehat{d}_a(i))_{i\leq I_a}$ can be constructed as follows. At time $T_a(i+1)$, the half-edge adjacent to the $(i+1)^{\text{st}}$ vertex is chosen uniformly at random from the set of all the half-edges adjacent to a vertex out-side B, and $\widehat{d}(i+1)$ is the forward-degree of the vertex adjacent to this half-edge. Thus, the sequence $(\widehat{d}(i))_{i\leq I_a}$ has the following description. Initially, associate to each vertex j a set of d_j half-edges (corresponding the set of half-edges outside B and L). At step 0, remove the half-edges corresponding to vertex a. Subsequently, at step $k \leq I_a$, choose a half-edge uniformly at random among all the remaining half-edges; if the half-edge is drawn from the node j's half-edges, then set $\widehat{d}_a(k) = d_j - 1$, and remove the node j and all of its half-edges. Obviously, this description allows to extend the definition of $\widehat{d}_a(i)$ to all the values of $I_a < i \leq n-1$. Indeed, if $I_a < n-1$, there are still half-edges at step I_a+1, and we can complete the sequence $\widehat{d}_a(i)$ for $i \in [I_a+1, n-1]$ by continuing the sampling described above. In this way, we obtain a sequence $(\widehat{d}_a(i))_{i=1}^{n-1}$ which coincides with the sequence defined in the previous section for $i \leq I_a$.

We also extend the sequence $\widehat{S}_a(i)$ for $i > I_a$ thanks to (2.18). Recall that, we set $X_a(i) = X_a(I_a)$ for all $i > I_a$. It is simple to see that with these conventions, the relation (2.19) is not anymore valid for $i > I_a$ but we still have $S_a(i) \leq \widehat{S}_a(i) - 2X_a(i)$ for all i.

The process $i \mapsto X_a(i)$ is non-decreasing in $i \in [1, n-1]$. Moreover, given $\mathcal{F}_{T_a(i)}$, the increment $X_a(i+1) - X_a(i)$ is stochastically dominated by the following binomial random variable

$$X_a(i+1) - X_a(i) \leq_{st} \text{Bin}\left(\widehat{d}_a(i+1), \frac{(S_a(i)-1)^+}{m^{(n)} - 2(X_a(i)+i)}\right). \tag{2.22}$$

Note that if $i > I_a$, then $S_a(i) = 0$ and $X_a(i+1) - X_a(i) = 0$, so that (2.22) is still valid.

For $i < \frac{n}{2}$, we have

$$\begin{aligned}\frac{(S_a(i)-1)^+}{m^{(n)}-2(X_a(i)+i)} &\leq \frac{\widehat{S}_a(i)-2X_a(i)}{m^{(n)}-2(X_a(i)+i)}\\ &\leq \frac{\widehat{S}_a(i)}{m^{(n)}-2i}\\ &\leq \frac{\max_{\ell\leq i}\widehat{S}_a(\ell)}{n-2i}.\end{aligned}$$

Hence, we obtain for $i < n/2$ that

$$X_a(i) \leq_{st} \text{Bin}\left(\max_{\ell\leq i}\widehat{S}_a(\ell)+i, \frac{\max_{\ell\leq i}\widehat{S}_a(\ell)}{n-2i}\right). \tag{2.23}$$

An important ingredient in the proof will be the coupling of the forward-degrees sequence $\{\widehat{d}(i)\}$ to an i.i.d. sequence in the range $i \leq \beta_n$, that we provide in the next subsection.

Recall that we defined α_n and β_n as follows (c.f. Equation (2.14))

$$\alpha_n = \log^3 n, \text{ and } \beta_n = 3\sqrt{\frac{\lambda}{\nu-1}n\log n}\,.$$

2.6.2 Coupling the forward-degrees sequence $\widehat{d}_a(i)$

We now present a coupling of the variables $\{\widehat{d}_a(1),...,\widehat{d}_a(k)\}$ valid for $k \leq \beta_n$, where β_n is defined in Equation (2.14), with an i.i.d. sequence of random variables, that we now define.

Denote the order statistics of the sequence of degrees $(d_i^{(n)})$ by

$$d_{(1)}^{(n)} \leq d_{(2)}^{(n)} \leq \cdots \leq d_{(n)}^{(n)}\,. \tag{2.24}$$

Define $\underline{m}^{(n)} := \sum_{i=1}^{n-\beta_n} d_{(i)}^{(n)}$ and let $\underline{\pi}^{(n)}$ be the size-biased empirical distribution with the β_n highest degrees in (2.24) removed, i.e.,

$$\underline{\pi}_k^{(n)} := \frac{\sum_{i=1}^{n-\beta_n}(k+1)\mathbb{1}_{\left(d_{(i)}^{(n)}=k+1\right)}}{\underline{m}^{(n)}}.$$

Similarly, define $\overline{m}^{(n)} := \sum_{i=(\beta_n+1)\Delta_n}^{n} d_{(i)}^{(n)}$ and let $\overline{\pi}^{(n)}$ be the size-biased empirical distribution with the $(\beta_n+1)\Delta_n$ lowest degrees in (2.24) removed, i.e.,

$$\overline{\pi}_k^{(n)} := \frac{\sum_{i=(\beta_n+1)\Delta_n}^{n}(k+1)\mathbb{1}_{\left(d_{(i)}^{(n)}=k+1\right)}}{\overline{m}^{(n)}}.$$

Note that by Condition 1.17, we have $\beta_n\Delta_n = o(n)$ which implies that both the distributions $\underline{\pi}^{(n)}$ and $\overline{\pi}^{(n)}$ converge to the size-biased distribution q defined in Equation (2.1).

The following lemma shows that the forward-degree of the i-th vertex given the forward-degrees of all the previous vertices is stochastically between two random variables with lower and upper distributions $\overline{\pi}^{(n)}$ and $\underline{\pi}^{(n)}$ defined above, provided that $i \le \beta_n$. More precisely,

Lemma 2.11. *For a uniformly chosen vertex a, we have for all $i \le \beta_n$,*

$$\underline{D}_i^{(n)} \le_{st} \left(\widehat{d}_a(i) \,|\, \widehat{d}_a(1),\dots,\widehat{d}_a(i-1)\right) \le_{st} \overline{D}_i^{(n)}, \tag{2.25}$$

where $\underline{D}_i^{(n)}$ (resp. $\overline{D}_i^{(n)}$) are i.i.d. with distribution $\underline{\pi}^{(n)}$ (resp. $\overline{\pi}^{(n)}$).

In particular, we have for all $i \le \beta_n$,

$$\sum_{k=1}^{i} \underline{D}_k^{(n)} \le_{st} \sum_{k=1}^{i} \widehat{d}_a(k) \le_{st} \sum_{k=1}^{i} \overline{D}_k^{(n)}.$$

Proof. Fix the sequence of degrees $\{d_i^{(n)}\}$ and the initial vertex a. We now prove that conditionally on the values of $(\widehat{d}_a(1),...,\widehat{d}_a(j-1))$, the random variable $\widehat{d}_a(j)$ is stochastically smaller than $\overline{D}_j^{(n)}$ provided that $j \le \beta_n$. This can be seen by a simple coupling argument as follows. In the following we will look at the half-edges adjacent to a node i as balls in a bin labeled with i, so the corresponding bin to the node i has d_i balls. First order the balls from 1 to m consistently with the order statistics. In other words, the values given to balls in each bin form an interval of consecutive numbers, and the values of the balls in the i-th bin is smaller than the values of the balls in the $(i+1)$-th bin for each i.

Given the sequence $(d_a, \widehat{d}_a(1),...,\widehat{d}_a(j-1))$, choose uniformly at random a set of $j-1$ bins containing respectively $\widehat{d}_a(1)+1,\dots,\widehat{d}_a(j-1)+1$ balls and color in red all the balls of these bins. In order to get a sample for $\overline{D}_j^{(n)}$, pick a ball at random among all balls in the last $n-(\beta_n+1)\Delta_n$ bins and set $\overline{D}_j^{(n)}$ to be equal to the size of the selected bin minus one. We now define a random variable $\tilde{d}_j$ such that $\widehat{d}_a(j) \le_{st} \tilde{d}_j \le \overline{D}_j^{(n)}$, obtaining the desired inequality. If the ball picked in defining $\overline{D}_j^{(n)}$ is white, set $\tilde{d}_j = \overline{D}_j^{(n)}$. If there are red balls in the last $n-(\beta_n+1)\Delta_n$ bins, and if the chosen ball is red, suppose that this ball is the ℓ-th ball among all the red balls in the last $n-(\beta_n+1)\Delta_n$ bins for the induced order by the enumeration of the balls, and then define $\tilde{d}_j$ to be the size of the bin containing the ℓ-th white ball minus one. Since $d_a + \widehat{d}_a(1) + \cdots + \widehat{d}_a(j-1) \le \beta_n\Delta_n$, this ball is in one of the first $(\beta_n+1)\Delta_n$ bins. In other

words, in all cases we have $\tilde{d}_j \leq \overline{D}_j^{(n)}$. By the definition of $\tilde{d}_j$ and the definition of stochastically dominance, it is fairly easy to show that , $\widehat{d}_a(j) \leq_{st} \tilde{d}_j$, conditioned on $(\widehat{d}_a(1), ..., \widehat{d}_a(j-1))$. Thus, we obtain one of the inequalities. A similar argument proves the other inequality $\underline{D}_j^{(n)} \leq_{st} \widehat{d}_j$ conditioned on the sequence $(d_a, \widehat{d}_a(1), ..., \widehat{d}_a(j-1))$, and so the first part of the lemma follows.

The second statement follows directly from the following basic result (whose proof is provided for the sake of completeness). □

Lemma 2.12. *Let $X_1, ..., X_t$ be a random process adapted to a filtration $\mathcal{F}_0 = \sigma[\emptyset], \mathcal{F}_1, ..., \mathcal{F}_t$, and let $\mathbf{\Sigma}_t = X_1 + ... + X_t$. Consider a distribution μ such that $(X_{s+1}|\mathcal{F}_s) \geq_{st} \mu$ (resp. $(X_{s+1}|\mathcal{F}_s) \leq_{st} \mu$) for all $0 \leq s \leq t-1$. Then $\mathbf{\Sigma}_t$ is stochastically larger (resp. smaller) than the sum of t i.i.d. μ-distributed random variables.*

Proof. We only consider the case $(X_{s+1} \,|\, \mathcal{F}_s) \leq_{st} \mu$. The other case follows similarly. By induction, it suffices to prove that for two random variables X_1, X_2, and for distributions μ_1, μ_2, if that $X_1 \leq_{st} \mu_1$, and $(X_2 \,|\, X_1) \leq \mu_2$, then

$$X_1 + X_2 \leq Z_{\mu_1} + Z_{\mu_2},$$

where Z_{μ_1} and Z_{μ_2} are independent random variables with distributions μ_1 and μ_2, respectively.

Note that for any x such that $\mathbb{P}(X_1 = x) > 0$, and for any t, we have

$$\mathbb{P}(X_1 + X_2 > t \mid X_1 = x) \leq \mathbb{P}(Z_{\mu_2} > t - x) = 1 - F_{\mu_2}(t - x),$$

where F_{μ_2} denotes the cumulative distribution function for distribution μ_2. By taking expectation over X_1, and since F_{μ_2} is non-decreasing, we obtain

$$\begin{aligned} \mathbb{P}(X_1 + X_2 > t) &\leq \mathbb{E}[1 - F_{\mu_2}(t - X_1)] \\ &\leq \mathbb{E}[1 - F_{\mu_2}(t - Z_{\mu_1})] = \mathbb{P}(Z_{\mu_1} + Z_{\mu_2} > s). \end{aligned}$$

□

2.7 Proof of the Upper Bound

In this section we present the proof of the upper bound for Theorem 2.2. Namely we prove that for any $\epsilon > 0$, with high probability for all vertices u and v which are in the same component

(i.e., such that $\mathrm{dist}_w(u,v) < \infty$), we have

$$\begin{aligned}\mathrm{dist}_w(u,v) &\leq \left(\frac{1}{\nu-1} + \frac{2}{d_{\min}}\mathbb{1}_{(d_{\min}\geq 3)} + \frac{1}{(1-q_1)}\mathbb{1}_{(d_{\min}=2)} + \frac{2}{1-\beta_*}\mathbb{1}_{(d_{\min}=1)}\right)(1+\epsilon)\log n \\ &= (1+\epsilon)(\frac{1}{\nu-1}\log n + s_n).\end{aligned}$$

Recall that we defined $\alpha_n = (\log n)^3$ and $\beta_n = 3\sqrt{\frac{\lambda}{\nu-1} n \log n}$. The proof will be based on the following two technical propositions. For the sake of readability, we postpone the proof of these two propositions to the end of this section.

The first one roughly says that for all u and v in the same component, the growing balls centered at u and v intersect w.h.p. provided that they contain each at least β_n nodes. More precisely,

Proposition 2.13. *We have w.h.p.*

$$\mathrm{dist}_w(u,v) \leq T_u(\beta_n) + T_v(\beta_n), \text{ for all } u \text{ and } v.$$

The above proposition shows that in proving the upper bound, it will be enough to control the random variable $T_u(\beta_n)$ for each node u in V. It turns out that in the range between α_n and β_n, in the three cases $d_{\min} \geq 3$, $d_{\min} = 2$, and $d_{\min} = 1$, $T_u(k)$ have more or less the same behavior, namely, it takes time at most roughly half of the typical (weighted) distance to go from size α_n to β_n. More precisely,

Proposition 2.14. *For a uniformly chosen vertex u and any $\epsilon > 0$, we have*

$$\mathbb{P}\left(T_u(\beta_n) - T_u(\alpha_n) \geq \frac{(1+\epsilon)\log n}{2(\nu-1)} \mid I_u \geq \alpha_n\right) = o(n^{-1}).$$

The conditioning $I_u \geq \alpha_n$ is here to ensure that the connected component which contains u has size at least α_n. In particular, note that one immediate corollary of the two above propositions is that two nodes whose connected components have size at least α_n are in the same component (necessarily the giant component), and that the two balls of size β_n centered at these two vertices intersect w.h.p.

Using the above two propositions, we are only left to understand $T_u(\alpha_n)$, and for this we will need to consider the three cases separately. Before going through the proof of the upper bound in these three cases, we need one more result. Consider the exploration process started

at a vertex a. We will need to find lower bounds for $S_a(k)$ in the range $1 \le k \le \alpha_n$. Recall that we defined $\gamma_a(k)$ as the number of nodes of forward-degree at least two in the growing balls centered at a, c.f. Equation (2.20) for the precise definition. These nodes are roughly all the ones which could contribute to the growth of the random variable $S_a(k)$. Now define the two following events.

$$\begin{aligned} R_a &:= \big\{S_a(k) \ge d_{\min} + \gamma_a(k), \text{ for all } 0 \le k \le \alpha_n - 1\big\}, \\ R'_a &:= \big\{S_a(k) \ge \gamma_a(k), \text{ for all } 0 \le k \le \alpha_n - 1\big\}. \end{aligned}$$

Lemma 2.15. *Assume $d_a \ge 2$ and $\widehat{d}_a(i) \ge 1$ for all $1 \le i \le \alpha_n$. Then we have*

$$\mathbb{P}\left(R_a \mid \widehat{d}_a(1), \dots, \widehat{d}_a(n-1)\right) \ge 1 - o(\log^{10} n/n), \tag{2.26}$$

$$\mathbb{P}\left(R'_a \mid \widehat{d}_a(1), \dots, \widehat{d}_a(n-1)\right) \ge 1 - o(n^{-3/2}). \tag{2.27}$$

In particular, $\mathbb{P}(R_a) \ge 1 - o(\log^{10} n/n)$ and $\mathbb{P}(R'_a) \ge 1 - o(n^{-3/2})$.

Proof. Since $\widehat{d}_a(i) \ge 1$, we have for all $k \le \alpha_n$,

$$d_{\min} + \gamma_a(k) \le d_a + \gamma_a(k) \le \widehat{S}_a(k) \le \alpha_n \Delta_n = o(n). \tag{2.28}$$

Thus, since $d_a \ge 2$ and $S_a(k) = \hat{S}_a(k) - 2X_a(k)$, we have

$$\{X_a(\alpha_n) = 0\} \subset R_a, \quad \{X_a(\alpha_n) \le 1\} \subset R'_a.$$

Note that the inequalities in (2.28) are true for any sequence such that $1 \le \widehat{d}_a(i) \le \Delta_n$. In particular, in the rest of the proof we condition on a realisation of the sequence

$$\mathbf{d} = (d_a, \widehat{d}_a(1), \dots, \widehat{d}_a(n-1)).$$

Note that since $\widehat{d}_a(i) \ge 1$, $\widehat{S}_a(k)$ is non-decreasing in k and so $\max_{k \le \alpha_n} \widehat{S}_a(k) = \widehat{S}_a(\alpha_n)$.

We distinguish two cases depending on whether or not $\widehat{S}_a(\alpha_n)$ is smaller than $3\alpha_n$. Denote this event by $\mathcal{Q}$ (and its complementary by $\mathcal{Q}^c$), i.e.,

$$\mathcal{Q} := \{\, \widehat{S}_a(\alpha_n) < 3\alpha_n \,\}.$$

- **Case 1)** $\widehat{S}_a(\alpha_n) < 3\alpha_n$. Conditioning on $\mathcal{Q}$, by (2.23) we have

$$X_a(\alpha_n) \le_{st} \mathrm{Bin}\left(4\alpha_n, \frac{3\alpha_n}{n - 2\alpha_n}\right).$$

Thus, we have

$$\begin{aligned}\mathbb{P}\left(X_a(\alpha_n) \geq 1 \mid \mathcal{Q}, \mathbf{d}\right) &\leq \mathbb{P}\left(\mathrm{Bin}\left(4\alpha_n, \tfrac{3\alpha_n}{n-2\alpha_n}\right) \geq 1\right) \leq O(\alpha_n^2/n),\\ \mathbb{P}\left(X_a(\alpha_n) \geq 2 \mid \mathcal{Q}, \mathbf{d}\right) &\leq \mathbb{P}\left(\mathrm{Bin}\left(4\alpha_n, \tfrac{3\alpha_n}{n-2\alpha_n}\right) \geq 2\right) \leq O(\alpha_n^4/n^2).\end{aligned}$$

We infer that

$$\begin{aligned}\mathbb{P}\left((R_a)^c \mid \mathcal{Q}, \mathbf{d}\right) &\leq O(\alpha_n^2/n),\\ \mathbb{P}\left((R'_a)^c \mid \mathcal{Q}, \mathbf{d}\right) &\leq O(\alpha_n^4/n^2).\end{aligned}$$

- **Case 2)** $\widehat{S}_a(\alpha_n) \geq 3\alpha_n$. Note that in this case, we still have $\max_{k\leq\alpha_n} \widehat{S}_a(k) = \widehat{S}_a(\alpha_n) \leq \alpha_n\Delta_n = o(n)$. Moreover, there exists $k \leq \alpha_n$ such that for all $\ell \leq k$, $\widehat{S}_a(\ell) < 3\alpha_n$ and $\widehat{S}_a(k+1) \geq 3\alpha_n$. Note that since we have conditioned on the degree sequence $\mathbf{d}$, the value of k is deterministic (k is not a random variable). Conditioning on the event $\mathcal{Q}^c$, we obtain by (2.23)

$$\begin{aligned}X_a(k) &\leq_{st} \mathrm{Bin}\left(4\alpha_n, \frac{3\alpha_n}{n-2\alpha_n}\right), \qquad \text{and}\\ X_a(\alpha_n) &\leq_{st} \mathrm{Bin}\left(\alpha_n(\Delta_n+1), \frac{\alpha_n\Delta_n}{n-2\alpha_n}\right).\end{aligned} \tag{2.29}$$

By Condition 2.1, there exists a $\tau > 0$ such that $\Delta_n := O(n^{1/2-\tau})$. Let $m = \lceil 2\tau^{-1}\rceil$. Combining the last (stochastic) inequality together with the Chernoff inequality applied to the right-hand side binomial random variable, we obtain

$$\begin{aligned}\mathbb{P}\left(X_a(\alpha_n) \geq m \mid \mathcal{Q}^c, \mathbf{d}\right) &\leq \mathbb{P}\left(\mathrm{Bin}\left(\alpha_n(\Delta_n+1), \frac{\alpha_n\Delta_n}{n-2\alpha_n}\right) \geq m\right)\\ &= O\left((\Delta_n^2\alpha_n^2/n)^m\right) = o(n^{-3}).\end{aligned}$$

We notice that for all $\ell > k$, we have $S_a(\ell) \geq 2\alpha_n - 2X_a(\alpha_n)$. Also for n large enough, we have $2\alpha_n - 2m \geq d_{\min} + \gamma_a(\ell)$. Therefore,

$$\begin{aligned}\left\{X_a(k) = 0,\ X_a(\alpha_n) \leq m,\ \mathcal{Q}^c\right\} &\subset R_a \cap \mathcal{Q}^c, \qquad \text{and}\\ \left\{X_a(k) \leq 1,\ X_a(\alpha_n) \leq m,\ \mathcal{Q}^c\right\} &\subset R'_a \cap \mathcal{Q}^c.\end{aligned}$$

This in turn implies that

$$\begin{aligned}\mathbb{P}\left((R_a)^c \mid \mathcal{Q}^c, \mathbf{d}\right) &\leq \mathbb{P}\left(X_a(k) \geq 1 \mid \mathcal{Q}^c\right) + \mathbb{P}\left(X_a(\alpha_n) \geq m \mid \mathcal{Q}^c\right) \leq O(\alpha_n^2/n)\\ \mathbb{P}\left((R'_a)^c \mid \mathcal{Q}^c, \mathbf{d}\right) &\leq \mathbb{P}\left(X_a(k) \geq 2 \mid \mathcal{Q}^c\right) + \mathbb{P}\left(X_a(\alpha_n) \geq m \mid \mathcal{Q}^c\right) \leq O(\alpha_n^4/n^2).\end{aligned}$$

In the above inequalities, we used (stochastic) Inequality (2.29) and Case 1 to bound the terms $\mathbb{P}\left(X_a(k) \geq 1 \mid \mathcal{Q}^c\right)$ and $\mathbb{P}\left(X_a(k) \geq 2 \mid \mathcal{Q}^c\right)$.

The lemma follows by the definition of α_n. □

We are now in position to provide the proof of the upper bound in the three different cases depending on whether $d_{\min} \geq 3$, $d_{\min} = 2$, or $d_{\min} = 1$. Recall, for the ease of reading, the definition of the two events R_a and R'_a that we will use throughout the proof below.

$$\begin{aligned} R_a &= \big\{S_a(k) \geq d_{\min} + \gamma_a(k), \text{ for all } 0 \leq k \leq \alpha_n - 1\big\}, \\ R'_a &= \big\{S_a(k) \geq \gamma_a(k), \text{ for all } 0 \leq k \leq \alpha_n - 1\big\}. \end{aligned}$$

The proof will be based on the analysis of these events. In particular, to justify that we have to consider these three different cases, we note that in the case $d_{\min} \geq 3$, the value of $\gamma_a(k)$ is always k, while in the case $d_{\min} = 2$ we need to control the length of the paths consisting of the vertices of degree two which contribute to the value of $\gamma_a(k)$, and in the case $d_{\min} = 1$ we need to do a kind of similar analysis as in the case $d_{\min} = 2$ but in a modified configuration model which consists of the 2-core of the graph. We also emphasize that one other important difference between the case $d_{\min} \geq 3$ and the two other cases $d_{\min} = 1, 2$ is that in the former case, as we will prove, the graph is connected with high probability, while in the two later cases $d_{\min} = 1, 2$ we also need to consider the small components of the 2-core. In addition, in the case $d_{\min} = 1$ we need to consider the vertices which are connected to the small components of 2-core and also, the tree components.

In the following, we will use the following property of the exponential random variables, without sometimes mentioning. If Y is an exponential random variable of rate μ, then for any $\theta < \mu$, we have $\mathbb{E}\left[e^{\theta Y}\right] = \frac{\mu}{\mu - \theta}$.

Proof of the upper bound in the case $d_{\min} \geq 3$. Let a be a node of the graph. Consider the exploration process defined in Section 2.6.1. First, note that in this case, the conditions $\hat{d}_a(i) \geq 1$ of Lemma 2.15 are automatically verified. Thus, as an immediate corollary, we obtain

Corollary 2.16. *We have $\mathbb{P}(I_a \geq \alpha_n) \geq 1 - o(n^{-3/2})$.*

Proof. Indeed, for $d_{\min} \geq 3$, we have $\gamma_a(k) = k$ so that

$$R'_a \subseteq \{I_a \geq \alpha_n\} = \{S_a(k) \geq 1, \text{ for all } 0 \leq k \leq \alpha_n - 1\}.$$

Now apply Lemma 2.15. □

We will need the following lemma.

Lemma 2.17. *Assume $d_{\min} \geq 3$. For a uniformly chosen vertex a, and any $\epsilon, \ell > 0$, we have*

$$\mathbb{P}\left(T_a(\alpha_n) \geq \epsilon \log n + \ell\right) = o(n^{-1} + e^{-d_{\min}\ell}).$$

Proof. Recall that given the sequence $S_a(k)$, for $k < I_a$, the random variables $T_a(k+1) - T_a(k)$ are i.i.d. exponential random variables with mean $S_a(k)^{-1}$.

Assume first that R'_a holds and consider the following two cases based on whether the event R_a holds or not.

- **Case 1) R_a holds.** By the definition of R_a, we have $S_a(k) \geq d_{\min} + k$ for all $k \leq \alpha_n - 1$. Conditioning on R_a, we have for any $k < \alpha_n$,

$$T_a(k+1) - T_a(k) \leq_{st} Y_k = \text{Exp}\left(d_{\min} + k\right),$$

and the random variables Y_k are all independent. Hence, we have

$$\begin{aligned}
\mathbb{E}\left[e^{d_{\min}(T_a(\alpha_n) - T_a(1))} \mid R_a\right] &\leq \mathbb{E}\left[e^{d_{\min}(\sum_{k=1}^{\alpha_n - 1} Y_k} \mid R_a\right] \\
&\leq \prod_{k=1}^{\alpha_n - 1} \left(1 + \frac{d_{\min}}{k}\right) \\
&\leq \exp\left[d_{\min} \sum_{k=1}^{\alpha_n - 1} \frac{1}{k}\right] \\
&\leq \alpha_n^{d_{\min}} = (\log n)^{3d_{\min}},
\end{aligned}$$

for n large enough. Thus, by Markov's inequality, we have for any $\epsilon > 0$,

$$\begin{aligned}
\mathbb{P}\big(T_a(\alpha_n) - T_a(1) \geq \epsilon \log n + \ell \mid R_a\big) &\leq (\log n)^{3d_{\min}} \exp(-d_{\min}(\epsilon \log n + \ell)) \\
&= \frac{(\log n)^{3d_{\min}}}{n^{\epsilon d_{\min}}} e^{-d_{\min}\ell} = o(e^{-d_{\min}\ell}).
\end{aligned}$$

By assumption (since R_a holds), $S_a(0) \geq d_{\min}$, and so we also have $T_a(1) \leq_{\text{st}} \text{Exp}(d_{\min})$. Therefore,

$$\mathbb{P}\big(T_a(1) \geq \epsilon \log n + \ell \mid R_a\big) \leq \frac{\exp(-\ell d_{\min})}{n^{\epsilon d_{\min}}},$$

and we conclude

$$\mathbb{P}\left(T_a(\alpha_n) \geq \epsilon \log n + \ell \mid R_a\right) = o(e^{-d_{\min}\ell}).$$

- **Case 2) R_a does not hold.** In this case, for any $k < \alpha_n$, by conditioning on $(R_a)^c \cap R'_a$, we have

$$T_a(k+1) - T_a(k) \leq_{st} Y_k = \text{Exp}(k),$$

and all the Y_k's are independent. We have

$$\begin{aligned}\mathbb{E}\left[e^{(T_a(\alpha_n)-T_a(2))} \mid R_a^c \cap R'_a\right] &\leq \prod_{k=2}^{\alpha_n - 1}\left(\frac{k}{k-1}\right)\\ &= \alpha_n - 1\\ &= \log^3 n - 1.\end{aligned}$$

By Markov's inequality, we have

$$\mathbb{P}\big(T_a(\alpha_n) - T_a(2) \geq \epsilon \log n + \ell \mid R_a^c \cap R'_a\big) \leq \log^3 n \exp(-\epsilon \log n - \ell) = o(n^{-\epsilon/2}).$$

We also have $T_a(1) \leq_{\text{st}} \text{Exp}(1)$ and $T_a(2) - T_a(1) \leq_{\text{st}} \text{Exp}(1)$, and we conclude in this case

$$\mathbb{P}\left(T_a(\alpha_n) \geq \epsilon \log n + \ell \mid R_a^c \cap R'_a\right) = o(n^{-\epsilon/2}).$$

Putting all the above inequalities together, we have

$$\begin{aligned}\mathbb{P}\big(T_a(\alpha_n) \geq \epsilon \log n + \ell\big) &\leq 1 - \mathbb{P}(R'_a) + (1 - \mathbb{P}(R_a))n^{-\epsilon/2} + o(e^{-d_{\min}\ell})\\ &\leq o(n^{-1} + e^{-d_{\min}\ell}),\end{aligned}$$

as desired. □

We can now finish the proof of the upper bound. By Proposition 2.13, we have (w.h.p.)

$$\begin{aligned}\text{flood}_w(G(n,(d_i)_1^n)) &= \max\{\text{dist}_w(a,b),\, b \in V,\, \text{dist}_w(a,b) < \infty\} \qquad (2.30)\\ &\leq T_a(\beta_n \wedge I_a) + \max_b T_b(\beta_n \wedge I_b), \quad \text{and}\\ \text{diam}_w(G(n,(d_i)_1^n)) &= \max\{\text{dist}_w(a,b),\, a,b \in V,\, \text{dist}_w(a,b) < \infty\} \qquad (2.31)\\ &\leq 2\max_a T_a(\beta_n \wedge I_a),\end{aligned}$$

where a is chosen uniformly at random in (2.30).

By Proposition 2.14 and Corollary 2.16, and Lemma 2.17 applied to $\ell = \epsilon \log n$, we obtain that for a uniformly chosen vertex a and any $\epsilon > 0$, we have

$$\mathbb{P}\left(T_a(\beta_n) \geq \frac{1}{2(\nu-1)}(1+\epsilon)\log n + 2\epsilon \log n\right) = o(1). \tag{2.32}$$

Indeed the above probability can be bounded above by

$$\mathbb{P}\left(T_a(\alpha_n) \geq 2\epsilon \log n\right) + \mathbb{P}\left(T_a(\beta_n) - T_a(\alpha_n) \geq \frac{1+\epsilon}{2(\nu-1)}\log n \mid I_a \geq \alpha_n\right) + \mathbb{P}(I_a < \alpha_n),$$

and this is $o(1)$ by the above cited results.

Furthermore, by Proposition 2.14 and Corollary 2.16, and Lemma 2.17 applied to $\ell = \frac{\log n}{d_{\min}}$, we obtain that for a uniformly chosen vertex b and any $\epsilon > 0$, we have

$$\mathbb{P}\left(T_b(\beta_n) \geq \left(\frac{1}{2(\nu-1)} + \frac{1}{d_{\min}}\right)(1+\epsilon)\log n\right) = o(n^{-1}). \tag{2.33}$$

Indeed the above probability can be bounded above by

$$\mathbb{P}\left(T_b(\alpha_n) \geq \frac{1+\epsilon}{d_{\min}}\log n\right) + \mathbb{P}\left(T_b(\beta_n) - T_b(\alpha_n) \geq \frac{1+\epsilon}{2(\nu-1)}\log n \mid I_b \geq \alpha_n\right) + \mathbb{P}(I_b < \alpha_n),$$

and this is $o(n^{-1})$ by the above cited results.

Applying Equation (2.33) and a union bound over b, we obtain

$$\mathbb{P}\left(\forall b,\ T_b(\beta_n) \leq \left(\frac{1}{2(\nu-1)} + \frac{1}{d_{\min}}\right)(1+\epsilon)\log n\right) = 1 - o(1). \tag{2.34}$$

We conclude by Equations (2.30), (2.32), and (2.34) that w.h.p. (for any $\epsilon > 0$)

$$\frac{\mathrm{flood}_w(G(n,(d_i)_1^n))}{\log n} \leq (1+\epsilon)\big(\frac{1}{\nu-1} + \frac{1}{d_{\min}}\big).$$

Clearly, the two equations (2.31) and (2.34) imply the bound on the diameter, that w.h.p.

$$\frac{\mathrm{diam}_w(G(n,(d_i)_1^n))}{\log n} \leq (1+\epsilon)\big(\frac{1}{\nu-1} + \frac{2}{d_{\min}}\big).$$

The proof of the upper bound in this case is now complete.

We end this section by the following remark on the connectivity of the random graph. We note that the above arguments show that the graph $G(n,(d_i)_1^n)$, satisfying Conditions 1.17, 1.22,

and 2.1 with $d_{\min} \geq 3$, is connected w.h.p. Indeed by Lemma 2.15, R'_a holds with probability at leat $1-o(n^{-1})$. With probability $1-o(n^{-1})$, for a uniformly chosen vertex a, we have $S_k(a) \geq 1$ for all $1 \leq k \leq \alpha_n$. By a union bound argument, with probability $1-o(1)$ the size of the growing ball centered at a reaches α_n for all nodes a in the graph. Using Lemma 2.22 shows that for all nodes, this cluster also reaches β_n. The connectivity then follows by applying Proposition 2.13.

Proof of the upper bound in the case $d_{\min} = 2$. Consider the exploration process defined in Section 2.6.1 starting from a. Recall the definitions (2.20) and (2.21): $\gamma_a(i)$ is the number of nodes with forward-degree (strictly) larger than one until the i-th exploration step, and $\overline{T}_a(k)$ is the first time that the k-th node with the forward-degree (strictly) larger than one appears in the exploration process started at node a. We also define the sets

$$L_a(k) := \big\{\ell,\, \overline{T}_a(k) \leq T_a(\ell+1) < \overline{T}_a(k+1)\big\},$$

for $k \geq 0$, and let $n_a(k)$ be the smallest ℓ in $L_a(k)$. Clearly, we have $n_a(k) \geq k-1$ and

$$\gamma_a^{-1}(k) = L_a(k) = [n_a(k), n_a(k+1) - 1].$$

For $x, y \in \mathbb{R}$, we denote $x \wedge y = \min(x, y)$. We will need the following lemma, equivalent to Lemma 2.17 (in the case $d_{\min} \geq 3$).

Lemma 2.18. *For a uniformly chosen vertex a, any $x > 0$, and any $\ell = O(\log n)$, we have*

$$\mathbb{P}\Big(T_a(\alpha_n \wedge I_a) \geq x \log n + \ell\Big) \leq o(n^{-1}) + o(e^{-2(1-q_1)\ell}).$$

Proof. Recall that given the sequence $S_a(k)$, for $k < I_a$, the random variables $T_a(k+1) - T_a(k)$ are i.i.d. exponential random variables with mean $S_a(k)^{-1}$. First write

$$\begin{aligned} T_a(\alpha_n) &= \sum_{0 \leq j < \alpha_n} T_a(j+1) - T_a(j) \\ &\leq \sum_{k \leq K_n} \overline{T}_a(k+1) - \overline{T}_a(k), \end{aligned}$$

where K_n is the largest integer such that $n_a(K_n) \leq \alpha_n$.

We now show that for any $x > 0$ and $\ell = O(\log n)$,

$$\mathbb{P}\Big(T_a(\alpha_n) \geq x \log n + \ell, R_a\Big) = o(e^{-2(1-q_1)\ell}). \tag{2.35}$$

Remark that a sum of a geometric (with parameter π) number of independent exponential random variables with parameter μ is distributed as an exponential random variable with parameter $(1-\pi)\mu$. For any $k \leq K_n$, we have:

$$\overline{T}_a(k+1) - \overline{T}_a(k) = \sum_{j \in L_a(k)} T_a(j+1) - T_a(j)$$

Assume R_a holds, then we have $S_a(j) \geq 2+k$ for all $j \in [n_a(k), n_a(k+1)-1] = L_a(k)$. Thus,

$$T_a(j+1) - T_a(j) \leq_{st} Y_{k,i} \sim \mathrm{Exp}(2+k),$$

where $i = j - n_a(k) + 1$, and all the $Y_{k,i}$'s are independent. (for $i = 1, \ldots, |L_a(k)|$, $Y_{k,i}$ are exponential random variables with rate $2+k$.)

For any positive t and θ, we obtain

$$\begin{aligned}
\mathbb{P}\left(T_a(\alpha_n) - \overline{T}_a(1) \geq t, R_a\right) &\leq \mathbb{E}\left[\mathbb{E}\left[\mathbb{1}(R_a) \prod_{1 \leq k \leq K_n} e^{\theta(\overline{T}_a(k+1) - \overline{T}_a(k))} \mid d_a, \ldots, \widehat{d}_a(n-1)\right]\right] e^{-\theta t} \\
&= \mathbb{E}\left[\prod_{1 \leq k \leq K_n} e^{\theta \sum_{i=1}^{|L_a(k)|} Y_{k,i}} \mathbb{P}\left(R_a \mid d_a, \ldots, \widehat{d}_a(n-1)\right)\right] e^{-\theta t} \\
&\leq \prod_{1 \leq k \leq \alpha_n} \left(1 + \frac{\theta}{(2+k)(1-\underline{\pi}_1^{(n)}) - \theta}\right) e^{-\theta t},
\end{aligned}$$

where in the last inequality, we used the fact that the probability for a new node to have forward-degree one is at most $\underline{\pi}_1^{(n)}$, and so the length $|L_a(k)|$ is dominated by a geometric random variable with parameter $\underline{\pi}_1^{(n)}$. Taking $\theta = 2(1-\underline{\pi}_1^{(n)})$ in the above inequality, we get

$$\begin{aligned}
\mathbb{P}\left(T_a(\alpha_n) - \overline{T}_a(1) \geq t, R_a\right) &\leq \prod_{k \leq \alpha_n} \left(1 + \frac{2(1-\underline{\pi}_1^{(n)})}{(1-\underline{\pi}_1^{(n)})i}\right) e^{-2(1-\underline{\pi}_1^{(n)})t} \\
&\leq \exp\left[2 \sum_{i=1}^{\alpha_n - 1} \frac{1}{i}\right] e^{-\theta t} < \alpha_n^3 e^{-2(1-\underline{\pi}_1^{(n)})t}.
\end{aligned}$$

In the same way, we can easily deduce that

$$\left(\overline{T}_a(1) \mid R_a\right) \leq_{\mathrm{st}} \mathrm{Exp}(2(1-\underline{\pi}_1^{(n)})).$$

Let $t = x \log n + \ell$, and note that $\ell \leq C \log n$ for some constant $C > 0$ (by assumption $\ell = O(\log n)$). Take any $0 < \epsilon < x(1-q_1)(C+x)^{-1}$; since for n sufficiently large, we have

$\underline{\pi}_1^{(n)} \le q_1 + \epsilon$, we obtain

$$\mathbb{P}\left(T_a(\alpha_n) \ge x \log n + \ell, R_a \right) \le \frac{\alpha_n^3}{n^{2(x(1-q_1-\epsilon)-\epsilon C)}} e^{-2(1-q_1)\ell},$$

and (2.35) follows. Note that $x(1 - q_1 - \epsilon) - \epsilon C > 0$ by the choice of ϵ.

Assume now that the event $R'_a \cap R_a^c$ holds. Two cases can happen: either $I_a < \alpha_n$ or $I_a \ge \alpha_n$. If $I_a < \alpha_n$, then by the definition of R'_a, $0 = S_a(I_a) \ge \gamma_a(I_a)$, i.e., $\gamma_a(I_a) = 0$. In other words, the component of a is a union of cycles (or loops) having node a as a common node, and with total number of edges less than α_n. Hence, in this case, we have

$$\begin{aligned}
&\mathbb{P}\Big(R'_a,\, R_a^c,\, I_a < \alpha_n, T_a(I_a) \ge x \log n + \ell \Big) \\
&\qquad \le \mathbb{P}\left(R_a^c \mid d_a, \ldots, \widehat{d}_a(n-1) \right) \left(\sum_{0 \le k \le \alpha_n} (\underline{\pi}_1^{(n)})^k \int_{x \log n + \ell}^{\infty} t^k \frac{e^{-t}}{k!} dt \right) \\
&\qquad \le \frac{\log^{10} n}{n} (1 - \underline{\pi}_1^{(n)})^{-1} \exp\Big(-(1 - \underline{\pi}_1^{(n)})(x \log n + \ell) \Big) = o(n^{-1}),
\end{aligned}$$

where the last inequality follows from Inequality (2.26) in Lemma 2.15.

In the second case, when $I_a \ge \alpha_n$, let

$$\mathcal{Q} = R'_a \cap R_a^c \cap \{\, I_a \ge \alpha_n \,\}.$$

If $\mathcal{Q}$ holds, by the definition of R'_a, we have $S_a(j) \ge k$ for all $j \in L_a(k)$. Thus,

$$T_a(j+1) - T_a(j) \le_{st} Y_{k,i} \sim \mathrm{Exp}(k),$$

where $i = j - n_a(k) + 1$, and all the $Y_{k,i}$'s are independent. (for $i = 1, \ldots, |L_a(k)|$, $Y_{k,i}$ are exponential random variables with rate k.) Hence, by the same argument as above, we have

$$\begin{aligned}
\mathbb{P}\left(T_a(\alpha_n) - \overline{T}_a(2) \ge t, \mathcal{Q} \right) &\le \mathbb{E}\left[\mathbb{E}\left[\mathbb{1}(\mathcal{Q}) \prod_{2 \le k \le K_n} e^{\theta(\overline{T}_a(k+1) - \overline{T}_a(k))} \mid d_a, \ldots, \widehat{d}_a(n-1) \right] \right] e^{-\theta t} \\
&\le \mathbb{E}\left[\prod_{2 \le k \le K_n} e^{\theta \sum_{i=1}^{|L_a(k)|} Y_{k,i}} \mathbb{P}\left(R_a^c \mid d_a, \ldots, \widehat{d}_a(n-1) \right) \right] e^{-\theta t} \\
&\le \prod_{2 \le k \le \alpha_n} \left(1 + \frac{\theta}{k(1 - \underline{\pi}_1^{(n)}) - \theta} \right) e^{-\theta t} o\left(\frac{\log^{10} n}{n} \right),
\end{aligned}$$

where the last inequality follows from Inequality (2.26) in Lemma 2.15. Thus, taking $\theta = 1-\underline{\pi}_1^{(n)}$ gives

$$\begin{aligned}\mathbb{P}\left(T_a(\alpha_n) - \overline{T}_a(2) \geq t, \mathcal{Q},\right) &\leq \prod_{2\leq k\leq \alpha_n}\left(1+\frac{1}{k-1}\right)e^{-(1-\underline{\pi}_1^{(n)})t}o\left(\frac{\log^{10} n}{n}\right)\\ &\leq \exp\left(\sum_{k=2}^{\alpha_n}\frac{1}{k-1}\right)e^{-(1-\underline{\pi}_1^{(n)})t}o\left(\frac{\log^{10} n}{n}\right)\\ &\leq e^{-(1-\underline{\pi}_1^{(n)})t}o\left(\frac{\log^{16} n}{n}\right).\end{aligned}$$

Since $d_a \geq 2$, we can easily deduce that

$$\left(\overline{T}_a(2) \mid \mathcal{Q}\right) \leq_{\text{st}} \text{Exp}(2(1-\underline{\pi}_1^{(n)})) + \text{Exp}(1-\underline{\pi}_1^{(n)}),$$

with these two exponential being independent and independent of $\mathcal{Q}$. Hence, we have

$$\mathbb{P}\left(\overline{T}_a(2) \geq t \mid \mathcal{Q}\right) \leq \int_t^\infty 2(1-\underline{\pi}_1^{(n)})\left(e^{-(1-\underline{\pi}_1^{(n)})x} - e^{-2(1-\underline{\pi}_1^{(n)})x}\right) \leq 2e^{-(1-\underline{\pi}_1^{(n)})t}.$$

Thus,

$$\mathbb{P}\left(T_a(\alpha_n) \geq t, \mathcal{Q}\right) \leq e^{-(1-\underline{\pi}_1^{(n)})t}o\left(\frac{\log^{16} n}{n}\right).$$

Similar to the case where R_a holds (by fixing a constant ϵ small enough and using that for n sufficiently large $\underline{\pi}_1^{(n)} \leq q_1 + \epsilon$ for n large enough), we get

$$\mathbb{P}\left(T_a(\alpha_n) \geq x\log n + \ell, \mathcal{Q}\right) \leq o\left(\frac{\log^{16} n}{n^{1+(1-q_1-\epsilon)C}}\right) = o(n^{-1}).$$

Putting all the above arguments together, and considering the three disjoint cases $(R'_a)^c$ holds, R_a holds, and $R'_a \cap R_a^c$ holds (in which case either $I_a < \alpha_n$ or $I_a \geq \alpha_n$), we conclude that

$$\mathbb{P}\left(T_a(\alpha_n \wedge I_a) \geq x\log n + \ell\right) \leq o(e^{-2(1-q_1)\ell}) + o(n^{-1}) + 1 - \mathbb{P}(R'_a).$$

To conclude the proof it suffices to use Lemma 2.15. □

We can now finish the proof of the upper bound in the case $d_{\min} = 2$. By Proposition 2.14, and Lemma 2.18 applied to $\ell = \frac{\log n}{2(1-q_1)}$, we obtain that for a uniformly chosen vertex a and any $\epsilon > 0$, we have

$$\mathbb{P}\left(\infty > T_a(\beta_n \wedge I_a) \geq \left(\frac{1}{2(\nu-1)} + \frac{1}{2(1-q_1)}\right)(1+\epsilon)\log n\right) = o(n^{-1}). \tag{2.36}$$

Indeed the above probability can be bounded above by

$$\mathbb{P}\left(T_a(\alpha_n \wedge I_a) \geq \frac{1+\epsilon}{2(1-q_1)}\log n\right) + \mathbb{P}\left(T_a(\beta_n) - T_a(\alpha_n) \geq \frac{1+\epsilon}{2(\nu-1)}\log n \mid I_a \geq \alpha_n\right),$$

and this is $o(n^{-1})$ by the above cited results.

Applying Equation (2.36) (and Lemma 2.18) and a union bound over a, we obtain

$$\mathbb{P}\left(\forall a,\ T_a(\beta_n \wedge I_a) \leq \left(\frac{1}{2(\nu-1)} + \frac{1}{2(1-q_1)}\right)(1+\epsilon)\log n\right) = 1 - o(1). \tag{2.37}$$

Hence by Proposition 2.13, we have w.h.p.

$$\frac{\mathrm{diam}_w(G(n,(d_i)_1^n))}{\log n} \leq (1+\epsilon)\big(\frac{1}{\nu-1} + \frac{1}{1-q_1}\big).$$

This proves the bound on the diameter. To obtain the upper bound for the flooding, we use Equation (2.37), and proceed as above by applying Proposition 2.14, and Lemma 2.18 applied to $\ell = \epsilon \log n$, to obtain that for a uniformly chosen vertex b, we have

$$\mathbb{P}\left(T_b(\beta_n \wedge I_b) \leq (\frac{1+\epsilon}{2(\nu-1)} + \epsilon)\log n\right) = 1 - o(1). \tag{2.38}$$

Clearly, Equations (2.37) and (2.38) imply the bound on the flooding.

The proof of the upper bound in this case is now complete.

Proof of the upper bound in the case $d_{\min} = 1$. We denote by $\mathcal{C}_a$ the event that a is connected to the 2-core of G_n. It is well-known (c.f. Section 2.5) that the condition $\nu > 1$ ensures that the 2-core of G_n has size $\Omega(n)$, w.h.p.

We consider the graph $\tilde{G}_n(a)$ obtained from G_n by removing all vertices of degree one except a until no such vertices exist. If the event $\mathcal{C}_a$ holds, $\tilde{G}_n(a)$ consists of the 2-core of G_n and the unique path (empty if a belongs to the 2-core) from a to the 2-core. While, if the event $\mathcal{C}_a^c$ holds, then the graph $\tilde{G}_n(a)$ is the union of the 2-core of G_n and the isolated vertex a.

In order to bound the weighted distance between two vertices a and b, in what follows, we will consider two cases depending on whether both the vertices a and b are connected to the 2-core (i.e., the events $\mathcal{C}_a$ and $\mathcal{C}_b$ both hold), or both the vertices a and b belong to the same tree component of the graph. In the former case, we will show how to adapt the analysis we made in the case $d_{\min} = 2$ to this case. And in the latter case, we directly bound the diameter of all the tree components of the graph.

First note that $\tilde{G}_n(a)$ can be constructed by means of a configuration model with a new degree sequence $\tilde{d}$, c.f. Section 2.5. Consider the exploration process on the graph $\tilde{G}_n(a)$ and denote by $\widetilde{T}_a(i)$ the first time the ball $\widetilde{B}_w(a,t)$ in $\tilde{G}_n(a)$ reaches size $i+1$. Also, $\tilde{I}_a$ is defined similarly to I_a for the graph $\tilde{G}_n(a)$. We need the following lemma.

Lemma 2.19. *For a uniformly chosen vertex a, any $x>0$ and any $\ell = O(\log n)$, we have*

$$\mathbb{P}\left(\widetilde{T}_a(\alpha_n \wedge \tilde{I}_a) \geq x\log n + \ell\right) \leq o(n^{-1}) + o(e^{-(1-\beta_*)\ell}).$$

Proof. First note that if $\mathcal{C}_a$ does not hold, i.e., if a is not connected to the 2-core, we will have $\tilde{I}_a = 0$ (since $\tilde{d}_a = 0$), and there is nothing to prove. Now the proof follows the same lines as in the proof of Lemma 2.18. Note that conditional on $\mathcal{C}_a$, we have $\tilde{d}_a \geq 1$, hence by Lemma 2.15, we have $\mathbb{P}(R_a \mid \mathcal{C}_a, \tilde{\mathbf{d}}) \geq 1 - o(\log^{10} n/n)$, and similarly for R'_a. The only difference we have to highlight here, compared to the proof of Lemma 2.18, is that conditional on $R_a \cap \mathcal{C}_a$, we have $\tilde{S}_a(j) \geq 1+k$ for all $j \in \tilde{L}_a(k)$, where $\tilde{S}_a(j)$ and $\tilde{L}_a(k)$ are defined in the same way as $S_a(j)$ and $L_a(k)$ for the graph $\tilde{G}(a)$. Take now $\theta = 1 - \underline{\tilde{\pi}}_1^{(n)}$ in the Chernoff bound, used in the proof of Lemma 2.18, where $\underline{\tilde{\pi}}^{(n)}$ is defined as $\underline{\pi}^{(n)}$ for the degree sequence $(\tilde{d}_1^{(n)}, ..., \tilde{d}_{\tilde{n}}^{(n)})$. The rest of the proof of Lemma 2.18 can then be easily adapted to obtain the same result provided we replace $2(1-q_1)$ by $(1-\beta_*)$, which is precisely the statement of the current lemma. (Note that $\beta_* = \tilde{q}_1$, c.f. Section 2.5.) □

By Proposition 2.14 applied to the graph $\tilde{G}_n(a)$ (note that $\tilde{\nu} = \nu$, cf. see Section 2.5), and Lemma 2.19 applied to $\ell = \frac{\log n}{1-\beta_*}$, we obtain that for a uniformly chosen vertex a and any $\epsilon > 0$, we have

$$\mathbb{P}\left(\infty > \widetilde{T}_a(\beta_n \wedge \tilde{I}_a) \geq \left(\frac{1}{2(\nu-1)} + \frac{1}{1-\beta_*}\right)(1+\epsilon)\log n\right) = o(n^{-1}). \tag{2.39}$$

Indeed the above probability can be bounded above by

$$\mathbb{P}\left(\widetilde{T}_a(\alpha_n \wedge \tilde{I}_a) \geq \frac{1+\epsilon}{1-\beta_*}\log n\right) + \mathbb{P}\left(\widetilde{T}_a(\beta_n) - \widetilde{T}_a(\alpha_n) \geq \frac{1+\epsilon}{2(\nu-1)}\log n \mid \tilde{I}_a \geq \alpha_n\right),$$

and this is $o(n^{-1})$ by the above cited results.

Applying Equation (2.39) (and Lemma 2.19) and a union bound over a, we obtain

$$\mathbb{P}\left(\forall a,\ \widetilde{T}_a(\beta_n \wedge \tilde{I}_a) \leq \left(\frac{1}{2(\nu-1)} + \frac{1}{1-\beta_*}\right)(1+\epsilon)\log n\right) = 1 - o(1). \tag{2.40}$$

To obtain the upper bound for the flooding, we use Equation (2.40), and proceed as above by using Lemma 2.19 applied to $\ell = \epsilon \log n$, to obtain that for a uniformly chosen vertex b, we have

$$\mathbb{P}\left(\widetilde{T}_b(\beta_n \wedge \tilde{I}_b) \leq (\frac{1+\epsilon}{2(\nu-1)} + \epsilon)\log n\right) = 1 - o(1). \tag{2.41}$$

Clearly, the two equations (2.40) and (2.41) together with Proposition 2.13 (since $\widetilde{T}_a(k) \geq T_a(k)$ for all k), imply the desired upper bound on the (weighted) flooding and (weighted) diameter on the giant component of G_n and also on every components containing a cycle, i.e, connected to 2-core.

At this point, we are only left to bound the (weighted) diameter and the (weighted) flooding of the tree components. In particular, the following lemma concludes the proof.

Lemma 2.20. *For two uniformly chosen vertices a, b, and any $\epsilon > 0$, we have*

$$\mathbb{P}\left(\frac{1+\epsilon}{1-\beta_*}\log n < \mathrm{dist}_w(a,b) < \infty,\ \mathcal{C}_a^c,\ \mathcal{C}_b^c\right) = o(n^{-2}).$$

Proof. We consider the graph $\tilde{G}_n(a,b)$ obtained from G_n by removing vertices of degree less than two except a and b until no such vertices exist. As shown in Section 2.5, the random graph $\tilde{G}_n(a,b)$ can be still obtained by a configuration model, and has the same asymptotic parameters as the random graph $\tilde{G}_n(a)$ in the proof of the previous lemma. We denote again by $\tilde{d}$, the degree sequence of the random graph $\tilde{G}_n(a,b)$. Also, $\tilde{T}_a$ and $\tilde{I}_a$ are defined similarly for the graph $\tilde{G}_n(a,b)$.

Trivially, we can assume $\tilde{d}_a = 1$ and $\tilde{d}_b = 1$, otherwise, either they are not in the same component and so $\mathrm{dist}_w(a,b) = \infty$, or one of them is in the 2-core, i.e., one of the two events $\mathcal{C}_a$ or $\mathcal{C}_b$ holds. Consider now the exploration process started at a until time k^* which is the first time either a node with forward-degree (strictly) larger than one appears or the time that the unique half-edge adjacent to b is chosen by the process. Let v^* be the node chosen at k^*. Note

that $\tilde{d}_{v^*} = 1$ if and only if the half-edge incident to b is chosen at k^*. We have

$$\begin{aligned}
\mathbb{P}\left(\frac{1+\epsilon}{1-\beta_*}\log n < \mathrm{dist}_w(a,b) < \infty, , \ \mathcal{C}_a^c,\ \mathcal{C}_b^c\right) &= \mathbb{P}\left(\widetilde{T}_a(k^*) > \frac{1+\epsilon}{1-\beta_*}\log n, v^* = b, \tilde{d}_a = \tilde{d}_b = 1\right) \\
&\leq \mathbb{P}\left(\widetilde{T}_a(k^*) > \frac{1+\epsilon}{1-\beta_*}\log n, v^* = b \mid \tilde{d}_a = \tilde{d}_b = 1\right) \\
&= \mathbb{P}\left(\widetilde{T}_a(k^*) > \frac{1+\epsilon}{1-\beta_*}\log n \mid \tilde{d}_a = \tilde{d}_b = 1\right) \times \\
&\qquad \mathbb{P}\left(\tilde{d}_{v^*} = 1 \mid \tilde{d}_{v^*} \neq 2, \tilde{d}_a = \tilde{d}_b = 1\right) \\
&= o(n^{-2}).
\end{aligned}$$

To prove the last equality above, first note $\mathbb{P}(\tilde{d}_{v^*} = 1 \mid \tilde{d}_{v^*} \neq 2, \tilde{d}_a = \tilde{d}_b = 1) = O(\frac{1}{n})$, this holds since $\nu = \tilde{\nu} > 1$ and v^* will be chosen before $o(n)$ steps, i.e., $k^* = o(n)$ (we will indeed prove something much stronger, that $k^* = O(\log n)$, c.f. Lemma 2.25 in the next section). And second, $\mathbb{P}\left(\widetilde{T}_a(k^*) > \frac{1+\epsilon}{1-\beta_*}\log n \mid \tilde{d}_a = \tilde{d}_b = 1\right) = o(1/n)$, this follows by the same argument as in the proof of Lemma 2.19 applied to $\tilde{G}_n(a,b)$, and by setting $\ell = \frac{(1+\epsilon)\log n}{1-\beta_*}$. □

The proof of the upper bound in this case is now completed by taking a union bound over all a and b. We end this section by presenting the proof of Proposition 2.13 and Proposition 2.14 in the next subsection.

2.7.1 Proof of Proposition 2.13 and Proposition 2.14

We start this section by giving some preliminary results that we will need in the proof of Proposition 2.13 and Proposition 2.14.

Lemma 2.21. *Let $\underline{D}_i^{(n)}$ be i.i.d. with distribution $\underline{\pi}^{(n)}$. For any $\eta < \nu$, there is a constant $\gamma > 0$ such that for n large enough we have*

$$\mathbb{P}\left(\underline{D}_1^{(n)} + \cdots + \underline{D}_k^{(n)} \leq k\eta\right) \leq e^{-\gamma k}. \tag{2.42}$$

Proof. Let D^* be a random variable with distribution $\mathbb{P}(D^* = k) = q_k$ given in Equation (2.1) so that $\mathbb{E}[D^*] = \nu$. Let $\phi(\theta) = \mathbb{E}[e^{-\theta D^*}]$. For any $\epsilon > 0$, there exists $\theta_0 > 0$ such that for any $\theta \in (0, \theta_0)$, we have

$$\log \phi(\theta) < (-\nu + \epsilon)\theta.$$

By Condition 1.17 and the fact that $\beta_n \Delta_n = o(n)$, i.e., $\sum_{i=n-\beta_n+1}^{n} d_{(i)}^{(n)} = o(n)$, we have for any $\theta > 0$,

$$\lim_{n\to\infty} \underline{\phi}^{(n)}(\theta) = \phi(\theta),$$

where $\underline{\phi}^{(n)}(\theta) = \mathbb{E}[e^{-\theta \underline{D}_1^{(n)}}]$. Also, for $\theta > 0$,

$$\mathbb{P}\left(\underline{D}_1^{(n)} + \cdots + \underline{D}_k^{(n)} \le \eta k\right) \le \exp\left(k\left(\theta\eta + \log \underline{\phi}^{(n)}(\theta)\right)\right).$$

Fix $\theta < \theta_0$ and let n be sufficiently large so that $\log \underline{\phi}^{(n)}(\theta) \le \log \phi(\theta) + \epsilon$. This yields

$$\begin{aligned}\mathbb{P}\left(\underline{D}_1^{(n)} + \cdots + \underline{D}_k^{(n)} \le \eta k\right) &\le \exp\left(k\left(\theta\eta + \log\phi(\theta) + \epsilon\theta\right)\right) \\ &\le \exp\left(k\theta\left(\eta - \nu + 2\epsilon\right)\right),\end{aligned}$$

which concludes the proof. □

The following lemma is the main step in the proof of both propositions.

Lemma 2.22. *For any $\epsilon > 0$, define the event*

$$R_a'' := \left\{ S_a(k) \ge \frac{\nu - 1}{1+\epsilon} k, \text{ for all } \alpha_n \le k \le \beta_n \right\}.$$

For a uniformly chosen vertex a, we have $\mathbb{P}\left(R_a'' \mid I_a \ge \alpha_n\right) \ge 1 - o(n^{-5})$.

Before giving the proof of this lemma, we recall the following basic result and one immediate corollary, for the proof see for example [120, Theorem 1].

Lemma 2.23. *Let $n_1, n_2 \in \mathbb{N}$ and $p_1, p_2 \in (0,1)$. We have $\mathrm{Bin}(n_1, p_1) \le_{st} \mathrm{Bin}(n_2, p_2)$ if and only if the following conditions hold*

(i) $n_1 \le n_2$,

(ii) $(1-p_1)^{n_1} \ge (1-p_2)^{n_2}$.

In particular, we have

Corollary 2.24. *If $x \le y = o(n)$, we have (for n large enough)*

$$x - \mathrm{Bin}(x, \sqrt{x/n}) \le_{st} y - \mathrm{Bin}(y, \sqrt{y/n}).$$

Proof. By the above lemma, it is sufficient to show

$$(x/n)^{x/2} \geq (y/n)^{y/2},$$

and this is true because s^s is decreasing near zero (for $s < e^{-1}$). □

Now we go back to the proof of Lemma 2.22.

Proof of Lemma 2.22. By Lemmas 2.11 and 2.21, for any $\epsilon > 0$, $k \geq \alpha_n$ and n large enough, we have

$$\mathbb{P}\left(\widehat{d}_a(1) + ... + \widehat{d}_a(k) \leq \frac{\nu}{1+\epsilon/2}k\right) \leq e^{-\gamma k} = o(n^{-6}).$$

We infer that with probability at least $1 - o(n^{-6})$, for any $k \leq \beta_n$,

$$\frac{\nu-1}{1+\epsilon/2}k < d_a + \widehat{d}_a(1) + ... + \widehat{d}_a(k) - k < (k+1)\Delta_n = o(n).$$

By the union bound over k, we have with probability at least $1-o(n^{-5})$ that for all $\alpha_n \leq k \leq \beta_n$,

$$\frac{\nu-1}{1+\epsilon/2}k < \widehat{S}_a(k) < (k+1)\Delta_n = o(n). \tag{2.43}$$

Hence in the remaining of the proof we can assume that the above condition is satisfied.

By Lemma 2.9, Corollary 2.24 and Inequality (2.43), conditioning on $\widehat{S}_a(k)$ and $\{I_a \geq k\}$, we have

$$\begin{aligned}\left(S_a(k) \mid \{I_a \geq k\}\right) &\geq_{st} \frac{\nu-1}{1+\epsilon/2}k - \mathrm{Bin}\left(\frac{\nu-1}{1+\epsilon/2}k, \sqrt{\left(\frac{\nu-1}{1+\epsilon/2}k\right)/n}\right) \\ &\geq_{st} \frac{\nu-1}{1+\epsilon/2}k - \mathrm{Bin}\left(\nu k, \sqrt{\nu k/n}\right).\end{aligned}$$

By Chernoff's inequality, since $k\sqrt{k/n} = o(k/\sqrt{\alpha_n})$, we have

$$\mathbb{P}\left(\mathrm{Bin}(\nu k, \sqrt{\nu k/n}) \geq k/\sqrt{\alpha_n}\right) \leq \exp\left(-\frac{1}{3}k/\sqrt{\alpha_n}\right) = o(n^{-6}).$$

Moreover, conditioned on $\{I_a \geq k\}$, we have with probability at least $1 - o(n^{-6})$,

$$S_a(k) \geq \frac{\nu-1}{1+\epsilon/2}k - \frac{k}{\sqrt{\alpha_n}} \geq \frac{\nu-1}{1+\epsilon}k,$$

for n large enough. Defining

$$R''_a(k) := \left\{ S_a(k) \geq \frac{\nu - 1}{1+\epsilon} k \right\} \text{ for } \alpha_n \leq k \leq \beta_n,$$

so that $R''_a = \bigcap_{k=\alpha_n}^{\beta_n} R''_a(k)$, we have

$$\mathbb{P}\left(R''_a(k) \mid I_a \geq k \right) \geq 1 - o(n^{-6}). \tag{2.44}$$

Thus, by using the fact that $R''_a(k-1) \subset \left\{ I_a \geq k \right\}$, we get

$$\begin{aligned}
\mathbb{P}\left(R''_a \mid I_a \geq \alpha_n \right) &= 1 - \mathbb{P}\left(\bigcup_{k=\alpha_n}^{\beta_n} R''_a(k)^c \mid I_a \geq \alpha_n \right) \\
&= 1 - \mathbb{P}\left(R''_a(\alpha_n)^c \cup \bigcup_{k=\alpha_n+1}^{\beta_n} \left(R''_a(k)^c \cap R''_a(k-1) \right) \mid I_a \geq \alpha_n \right) \\
&\geq 1 - \mathbb{P}\left(R''_a(\alpha_n)^c \cup \bigcup_{k=\alpha_n+1}^{\beta_n} \left(R''_a(k)^c \cap \left\{ I_a \geq k \right\} \right) \mid I_a \geq \alpha_n \right) \\
&\geq 1 - \sum_{k=\alpha_n}^{\beta_n} \mathbb{P}\left(R''_a(k)^c \mid I_a \geq k \right) \\
&\geq 1 - o(n^{-5}),
\end{aligned}$$

which concludes the proof. □

We are now in position to provide the proof of both the propositions.

Proof of Proposition 2.13. Fix two vertices u and v. We can assume that $T_u(\beta_n), T_v(\beta_n) < \infty$, i.e., $I_u, I_v \geq \beta_n$, otherwise the statement of the proposition holds trivially for u and v. Note that $\mathrm{dist}_w(u,v) \leq T_u(\beta_n) + T_v(\beta_n)$ is equivalent to

$$B_w(u, T_u(\beta_n)) \cap B_w(v, T_v(\beta_n)) \neq \emptyset.$$

Hence, to prove the proposition we need to bound the probability that $B_w(v, T_v(\beta_n))$ does not intersect $B_w(u, T_u(\beta_n))$.

First consider the exploration process for $B_w(u,t)$ until reaching $t = T_u(\beta_n)$. We know by Lemma 2.22 that with probability at least $1 - o(n^{-5})$,

$$S_u(\beta_n) \geq (\nu - 1 - o(1))\beta_n.$$

(In other words, there are at least $(\nu - 1 - o(1))\beta_n$ half-edges in $B_w(u, T_u(\beta_n))$.)

Next, begin exposing $B_w(v,t)$. Each matching adds a uniform half-edge to the neighborhood of v. Therefore, the probability that $B_w(v, T_v(\beta_n))$ does not intersect $B_w(u, T_u(\beta_n))$ is at most

$$\left(1 - \frac{(\nu - 1 - o(1))\beta_n}{m^{(n)}}\right)^{\beta_n} \leq \exp[-(9 - o(1)) \log n] < n^{-4}$$

for large n (recall that $\beta_n^2 = \frac{9\lambda n \log n}{\nu - 1}$). The union bound over u and v completes the proof. □

Proof of Proposition 2.14. Conditioning on the event R''_a defined in Lemma 2.22, we have for any $\alpha_n \leq k \leq \beta_n$,

$$T_a(k+1) - T_a(k) \leq_{st} Y_k \sim \text{Exp}(S_a(k)) \leq_{st} \text{Exp}\left(\frac{\nu - 1}{1 + \epsilon} k\right),$$

and all the Y_k's are independent.

Letting $s = \sqrt{\alpha_n}$, for n large enough we obtain that

$$\begin{aligned}
\mathbb{E}\left[e^{s(T_a(\beta_n) - T_a(\alpha_n))} \mid R''_a\right] \leq \prod_{k=\alpha_n}^{\beta_n - 1} \left(1 + \frac{s}{\frac{(\nu-1)k}{1+\epsilon} - s}\right) &\leq \prod_{k=\alpha_n}^{\beta_n - 1} \left(1 + \frac{s(1+2\epsilon)}{(\nu-1)k}\right) \\
&\leq \exp\left[\frac{s(1+2\epsilon)}{\nu - 1} \sum_{k=\alpha_n}^{\beta_n - 1} \frac{1}{k}\right] \\
&\leq \exp\left[\frac{s(1+3\epsilon)\log n}{2(\nu - 1)}\right].
\end{aligned}$$

By Markov's inequality,

$$\begin{aligned}
\mathbb{P}\left(T_a(\beta_n) - T_a(\alpha_n) \geq \frac{(1+4\epsilon)\log n}{2(\nu-1)} \mid I_a \geq \alpha_n\right) &\leq 1 - \mathbb{P}(R''_a) + \mathbb{E}\left[e^{s(T_a(\beta_n) - T_a(\alpha_n))} \mid R''_a\right] \\
&\quad . \exp\left(-\frac{s(1+4\epsilon)\log n}{2(\nu-1)}\right) \\
&\leq \exp\left(-\frac{s\epsilon \log n}{2(\nu-1)}\right) + o(n^{-5}) = o(n^{-1}),
\end{aligned}$$

which concludes the proof. □

2.8 Proof of the Lower Bound

In this section we present the proof of the lower bound for Theorem 2.2. To prove the lower bound, it suffices to show that for any $\epsilon > 0$, there exist w.h.p. two vertices u and v such that

$$\begin{aligned}\mathrm{dist}_w(u,v) >& \left(\frac{1}{\nu-1} + \frac{2}{d_{\min}}\mathbb{1}_{(d_{\min}\geq 3)} + \frac{1}{(1-q_1)}\mathbb{1}_{(d_{\min}=2)} + \frac{2}{1-\beta_*}\mathbb{1}_{(d_{\min}=1)}\right)(1-\epsilon)\log n \\ =&(1-\epsilon)(\frac{1}{\nu-1}\log n + s_n).\end{aligned}$$

As in the proof of the upper bound, the proof will be different depending whether $d_{\min} = 1, 2$, or ≥ 3. So we start this section by proving some preliminary results, including some new notations and definitions, that we will need in the proof for these three cases, and then divide the end of the proof into three cases.

Fix a vertex a in $G_n \sim G(n,(d_i)_1^n)$, and consider the exploration process, defined in Section 2.6.1. Recall that $\overline{T}_a(1)$ is the first time when the ball centered at a contains a vertex of forward-degree at least two (i.e., degree at least 3), c.f. Equation-Definition (2.21). To simplify the notation, we denote by C_a the ball centered at a containing exactly one node (possibly in addition to a) of degree at least 3:

$$C_a := B_w(a, \overline{T}_a(1)). \tag{2.45}$$

Note that there is a vertex u (of degree $d_u \geq 3$) in C_a which is not in any ball $B_w(a,t)$ for $t < \overline{T}_a(1)$ and we have $\max_{v\in C_a} \mathrm{dist}_w(a,v) = \mathrm{dist}_w(a,u)$. We define the degree of C_a as

$$\deg(C_a) = d_a + d_u - 2. \tag{2.46}$$

Remark that at time $\overline{T}_a(1)$ of the exploration process defined in Section 2.6.1 starting from a, we have at most $\deg(C_a)$ free half-edges, i.e., the list L contains at most $\deg(C_a)$ half-edges. (we have the equality if the tree excess until time $\overline{T}_a(1)$ is zero.) The following lemma shows that the size of C_a is relatively small.

Lemma 2.25. *Consider a random graph $G(n,(d_i)_1^n)$ where the degrees d_i satisfy Conditions 1.17 and 1.22. There exists a constant $M > 0$, independent of n, such that w.h.p. for all the nodes a of the graph, we have $|C_a| \leq M \log n$.*

Proof. We consider the exploration process, defined in Section 2.6.1, starting from a uniformly chosen vertex a, and use the coupling of the forward-degrees we described in subsection 2.6.2.

Recall in particular that each forward-degree $\widehat{d}(i)$ conditioned on the previous forward-degrees is stochastically larger than a random variable with distribution $\underline{\pi}^{(n)}$. This shows that, at each step of the exploration process, the probability of choosing a node of degree at most two (forward-degree one or zero) will be at most $\underline{\pi}_0^{(n)} + \underline{\pi}_1^{(n)} < 1 - \epsilon$, for some $\epsilon > 0$ (note that the asymptotic mean of $\underline{\pi}^{(n)}$ is ν, and by assumption $\nu > 1$). We conclude that there exists a constant $M > 0$ such that for all large n,

$$\mathbb{P}(|C_a| > M \log n) = o(n^{-1}).$$

The union bound over a completes the proof. □

For two subsets of vertices $U, W \subset V$, the (weighted) distance between U and W is defined as usual,

$$\mathrm{dist}_w(U, W) := \min\{\ \mathrm{dist}_w(u, w) \mid u \in U,\ w \in W\ \}.$$

For two nodes a, b, define the event $\mathcal{H}_{a,b}$ as

$$\mathcal{H}_{a,b} := \left\{ \frac{1-\epsilon}{\nu - 1} \log n < \mathrm{dist}_w(C_a, C_b) < \infty \right\}. \tag{2.47}$$

Note that $\frac{\log n}{\nu - 1}$ is the typical distance, so the left inequality in the definition of the above event means that C_a and C_b have the right typical distance in the graph (modulo a factor $(1-\epsilon)$). The right inequality simply means that a and b belong to the same connected component. The following proposition is the crucial step in the proof of the lower bound, the proof of which is postponed to the end of this section.

Proposition 2.26. *Consider a random graph $G(n, (d_i)_1^n)$ with i.i.d. rate one exponential weights on its edges. Suppose that the degree sequence $(d_i)_1^n$ satisfies Conditions 1.17 and 1.22. Assume that the number of nodes with degree one satisfy $u_1^{(n)} = o(n)$, and let a and b be two distinct vertices such that $\mathrm{deg}(C_a) = O(1)$, and $\mathrm{deg}(C_b) = O(1)$. Then for all $\epsilon > 0$,*

$$\mathbb{P}\left(\mathcal{H}_{a,b}\right) = 1 - o(1).$$

Furthermore, the same result holds without the condition $\mathrm{deg}(C_a) = O(1)$ if the node a is chosen uniformly at random and $\mathrm{deg}(C_b) = O(1)$.

Assuming the above proposition, we now show that

(*i*) If the minimum degree $d_{\min} \geq 3$, then there are pairs of nodes a and b of degree $d_{\min}$ such that the event $\mathcal{H}_{a,b}$ holds and in addition all the weights on the edges adjacent to a or b are at least $(1-\epsilon)\log n/d_{\min}$ w.h.p., for all $\epsilon > 0$.

(*ii*) If the minimum degree $d_{\min} = 2$, then there are pairs of nodes a and b of degree two such that $\mathcal{H}_{a,b}$ holds and in addition, the closest nodes to each with forward-degree at least two is at distance at least $(1-\epsilon)\log n/(2(1-q_1))$ w.h.p., for all $\epsilon > 0$.

(*iii*) If the minimum degree $d_{\min} = 1$, then there are pairs of nodes of degree one such that $\mathcal{H}_{a,b}$ holds and in addition, the closest node to each which belongs to the 2-core is at least $(1-\epsilon)\log n/(1-\beta_*)$ away w.h.p., for all $\epsilon > 0$.

This will finish the proof of the claimed lower bound.

Proof of the lower bound in the case $d_{\min} \geq 3$. Let V^* be the set of all vertices of degree $d_{\min}$. We call a vertex u in V^* *good* if $\overline{T}_u(1)$ is at least $(1-\epsilon)s_n$, i.e.,

$$\overline{T}_u(1) \geq (1-\epsilon)s_n,$$

and if in addition, $\deg(C_u) \leq K$ for a constant K chosen as follows. Let $\widehat{D}$ be a random variable with the size-biased distribution, i.e., $\mathbb{P}(\widehat{D} = k) = q_k$. The constant K is chosen in order to have with positive probability $\widehat{D} \leq d_{\min} - 1 + K$, i.e.,

$$y = y_K := \mathbb{P}\left(\widehat{D} \leq d_{\min} - 1 + K\right) > 0. \tag{2.48}$$

It will be convenient to consider the two events in the definition of good vertices separately, namely, for a vertex $u \in V^*$, define

$$\mathcal{E}_u := \left\{\overline{T}_u(1) \geq (1-\epsilon)s_n\right\}, \text{ and} \tag{2.49}$$

$$\mathcal{E}'_u := \left\{\deg(C_u) \leq K\right\}. \tag{2.50}$$

We note that in the case $d_{\min} \geq 3$, the event $\mathcal{E}_u$ for $u \in V^*$ is equivalent to having a weight greater than $(1-\epsilon)s_n$ on all the $d_{\min}$ edges connected to u, and clearly, the two above events $\mathcal{E}'_u$ and $\mathcal{E}_u$ are independent (conditionally on $u \in V^*$, i.e., $d_u = d_{\min}$).

For $u \in V^*$, let A_u be the event that u is good, $A_u := \mathcal{E}_u \cap \mathcal{E}'_u$, and let Y be the total number of good vertices, $Y := \sum_u \mathbb{1}_{A_u}$. In the following, we first obtain a bound for the expected value

of Y, and then use the second moment inequality to show that w.h.p. $Y = \Omega(n^\epsilon)$.

$$\begin{aligned}
\mathbb{E}[Y] &= \sum_{u \in V^*} \mathbb{P}(A_u) = \sum_{u \in V^*} \mathbb{P}(\mathcal{E}_u) \cdot \mathbb{P}(\mathcal{E}'_u) && \text{(By the independence of } \mathcal{E}_u \text{ and } \mathcal{E}'_u) \\
&= \sum_{u \in V^*} n^{\epsilon-1} \mathbb{P}(\mathcal{E}'_u) && \text{(Since } \mathbb{P}(\mathcal{E}_u) = \mathbb{P}\left(\mathrm{Exp}(d_{\min}) \geq (1-\epsilon)s_n\right)) \\
&= \sum_{u \in V^*} n^{\epsilon-1}(1 \pm o(1))y \\
&= (1 \pm o(1)) p_{d_{\min}} n^\epsilon y.
\end{aligned}$$

In the equality before last one, we used the coupling argument described in Section 2.6.2 to bound the forward-degrees from above (and below) by i.i.d. random variables having distributions $\overline{\pi}^{(n)}$ (and $\underline{\pi}^{(n)}$), and then used the fact that the asymptotic distributions of both $\overline{\pi}^{(n)}$ and $\underline{\pi}^{(n)}$ coincides with the size biased distribution $\{q_k\}$. This is where the factor $(1 \pm o(1))y$ in the above equality comes from. The last equality is simply obtained from Condition 1.17, which implies that $|V^*| = (1 \pm o(1)) p_{d_{\min}} n$.

We now show that $\mathrm{Var}(Y) = o(\mathbb{E}[Y]^2)$. Applying the Chebyshev inequality, this will show that $Y \geq \frac{2}{3} p_{d_{\min}} n^\epsilon$ with high probability.

We have

$$\begin{aligned}
\mathbb{E}[Y^2] &= \mathbb{E}\left[\left(\sum_{u \in V^*} \mathbb{1}_{A_u} \right)^2 \right] = \mathbb{E}\left[\sum_{u,v \in V^*} \mathbb{1}_{A_u} \mathbb{1}_{A_v} \right] \\
&= \mathbb{E}\left[\sum_{u,v \in V^* : C_u \cap C_v \neq \emptyset} \mathbb{1}_{A_u} \mathbb{1}_{A_v} + \sum_{u,v \in V^* : C_u \cap C_v = \emptyset} \mathbb{1}_{A_u} \mathbb{1}_{A_v} \right] \\
&= \mathbb{E}\left[\sum_{u \in V^*} \mathbb{1}_{A_u} \sum_{v \in V^* : C_u \cap C_v \neq \emptyset} \mathbb{1}_{A_v} + \sum_{u,v \in V^* : C_u \cap C_v = \emptyset} \mathbb{1}_{A_u} \mathbb{1}_{A_v} \right] \\
&\leq \mathbb{E}\left[\sum_{u \in V^*} \mathbb{1}_{A_u} (K+2) + \sum_{u,v \in V^* : C_u \cap C_v = \emptyset} \mathbb{1}_{A_u} \mathbb{1}_{A_v} \right] \\
&= (K+2)\mathbb{E}[Y] + \mathbb{E}\left[\sum_{u,v \in V^* : C_u \cap C_v = \emptyset} \mathbb{1}_{A_u} \mathbb{1}_{A_v} \right],
\end{aligned}$$

where to obtain the inequality above we used the fact that for each node u, when the event A_u holds, the degree of C_u is at most K and so $C_u \cap C_v \neq \emptyset$ happens at most for $(K+2)$ nodes v.

Indeed, each C_v consists of v and another node of the graph. A simple analysis shows that such a v either belongs to C_u (two possibilities) or is among the neighbors of C_u (at most K choices).

Moreover, for any pair of vertices $u, v \in V^*$, such that $C_u \cap C_v = \emptyset$, conditioning on A_u does not have much effect on the asymptotic of the degree distribution. Indeed, by the coupling argument of Section 2.6.2, we have for $u, v \in V^*$ such that $C_u \cap C_v = \emptyset$,

$$\begin{aligned}\mathbb{P}(A_v \mid A_u) &= \mathbb{P}(\mathcal{E}_v \mid A_u)\mathbb{P}(\mathcal{E}'_v \mid A_u) = \mathbb{P}(\mathcal{E}_v)\mathbb{P}(\mathcal{E}'_v \mid A_u) \\ &\geq \mathbb{P}\big(\mathrm{Exp}(d_{\min}) \geq s_n\big)\,\mathbb{P}\big(\,\overline{D}^{(n)} \leq d_{\min} - 1 + K\,\big)\end{aligned}$$

where $\overline{D}^{(n)}$ is a random variable with distribution $\overline{\pi}^{(n)}$. Similarly, we have

$$\mathbb{P}(A_v \mid A_u) \leq \mathbb{P}(\mathrm{Exp}(d_{\min}) \geq s_n)\mathbb{P}(\underline{D}^{(n)} \leq d_{\min} - 1 + K),$$

where $\underline{D}^{(n)}$ is a random variable with distribution $\underline{\pi}^{(n)}$. We infer that

$$\mathbb{P}(A_v \mid A_u) = (1 \pm o(1))y\; n^{-1+\epsilon},$$

and then using the estimate $\mathbb{P}(A_v) = (1 \pm o(1))yn^{\epsilon-1}$ we obtained above, we obtain (conditioned on $C_u \cap C_v = \emptyset$)

$$\mathbb{P}(A_v \cap A_u) = (1 \pm o(1))\mathbb{P}(A_u)\mathbb{P}(A_v).$$

This shows that

$$\begin{aligned}\mathbb{E}\left[\sum_{u,v \in V^*:\, C_u \cap C_v = \emptyset} \mathbb{1}_{A_u}\mathbb{1}_{A_v}\right] &\leq (1 + o(1)) \sum_{u,v \in V^*} \mathbb{P}(A_u)\mathbb{P}(A_v) \\ &= (1 + o(1))\mathbb{E}[Y]^2.\end{aligned}$$

Hence, we have

$$\mathrm{Var}[Y] = \mathbb{E}[Y^2] - \mathbb{E}[Y]^2 \leq (K+2)\mathbb{E}[Y] + o(\mathbb{E}[Y]^2) = o(\mathbb{E}[Y]^2).$$

This finishes the proof of the fact that $Y \geq \frac{2}{3}p_{d_{\min}}n^{\epsilon}$ with high probability.

We consider first the flooding, and obtain the corresponding lower bound. Let Y' denote the number of good vertices that are at distance at most $(1-\epsilon)s_n + \frac{(1-\epsilon)}{\nu - 1}\log n$ from a vertex a (chosen uniformly at random). It is clear that the lower bound follows by showing that $Y' < Y$ with high probability, i.e., $Y - Y' > 0$ w.h.p. To show this, we will bound the expected value of Y' and use Markov's inequality.

Since by Condition 1.17, V^* has size linear in n by applying Proposition 2.26, we obtain that for a uniformly chosen vertex $u \in V^*$, conditioning on A_u, we have $\mathbb{P}(\mathcal{H}_{a,u}) = 1 - o(1)$. Indeed, the two events $\mathcal{H}_{a,u}$ and $\mathcal{E}_u$ are independent, and conditioning on $\mathcal{E}'_u$ is the same as conditioning on $\deg(C_u) \leq K = O(1)$. Therefore, for a uniformly chosen vertex u in V^*, we have

$$\mathbb{P}\left(A_u \cap \mathcal{H}^c_{a,u}\right) = o(\mathbb{P}(A_u)),$$

where $\mathcal{H}^c_{a,i}$ denotes the complementary event of $\mathcal{H}_{a,i}$, i.e., the event that $\mathcal{H}_{a,i}$ does not occur. Thus, a straightforward calculation shows that

$$\mathbb{E}[Y'] = o(\mathbb{E}[Y]) = o(n^\epsilon).$$

By Markov's inequality, we conclude that $Y' \leq \frac{1}{3} p_{d_{\min}} n^\epsilon$ w.h.p., and hence $Y - Y'$ is w.h.p. positive. This implies the existence of a vertex u whose distance from a is at least $\left(\frac{1}{\nu-1} + \frac{1}{d_{\min}}\right)(1-\epsilon)\log n$. Hence for any $\epsilon > 0$ we have w.h.p.

$$\text{flood}_w(G_n) \geq \max_{u \in V^*} \text{dist}_w(a,u) \geq \left(\frac{1}{\nu-1} + \frac{1}{d_{\min}}\right)(1-\epsilon)\log n.$$

We now turn to the proof of the lower bound for the (weighted) diameter of the graph. The proof will follows the same strategy as for the flooding, but this time we need to consider the pairs of good vertices. Let R denote the number of pairs of distinct good vertices. Recall we proved above that w.h.p. $Y \geq \frac{2}{3}\mathbb{E}[Y]$. Thus,

$$R = Y(Y-1) \geq \frac{2\mathbb{E}[Y]}{3}\left(\frac{2\mathbb{E}[Y]}{3} - 1\right) > \frac{1}{4}\mathbb{E}[Y]^2.$$

The probabilities that u and v are both good and $\mathcal{H}_{u,v}$ does not happen can be bounded as follows.

$$\begin{aligned}
\mathbb{P}\left(A_u \cap A_v \cap \mathcal{H}^c_{u,v}\right) &= \mathbb{P}(A_u \cap A_v)\mathbb{P}(\mathcal{H}^c_{u,v} \mid A_u, A_v) \\
&= \mathbb{P}(A_u \cap A_v)\mathbb{P}\left(\mathcal{H}^c_{u,v} \mid \deg(C_u) \leq K, \deg(C_v) \leq K\right) \\
&\qquad \text{(We used the independence of } \mathcal{H}_{u,v} \text{ and } \mathcal{E}_u \text{ and } \mathcal{E}_v\text{)} \\
&= o(\mathbb{P}(A_u \cap A_v)). \qquad (2.51)
\end{aligned}$$

The last equality follows from Proposition 2.26, since C_u and C_v are of degree $O(1)$.

To conclude, consider R' the number of pairs of good vertices that are at distance at most $(1-\epsilon)(2s_n + \frac{\log n}{\nu-1})$. By using Equation (2.51), we have $\mathbb{E}R' = o(\mathbb{E}[Y]^2)$. Applying Markov's

inequality, we obtain that w.h.p. $R' \leq \frac{1}{6}(\mathbb{E}[Y])^2$, and thus, $R - R'$ is w.h.p positive. This implies that for any $\epsilon > 0$, we have w.h.p.

$$\mathrm{diam}_w(G_n) \geq \max_{u,v \in V^*} \mathrm{dist}_w(u,v) \geq \left(\frac{1}{\nu - 1} + \frac{2}{d_{\min}}\right)(1-\epsilon)\log n.$$

Proof of the lower bound in the case $d_{\min} = 2$. Let V^* be the set of vertices with degree two. Again we call a vertex u in V^* *good* if both the events $\mathcal{E}_u$ and $\mathcal{E}'_u$ hold. Recall the definition of the two events

$$\mathcal{E}_u := \left\{\overline{T}_u(1) \geq (1-\epsilon)s_n\right\}, \text{ and} \tag{2.52}$$

$$\mathcal{E}'_u := \left\{\deg(C_u) \leq K\right\}. \tag{2.53}$$

where in this case we choose K in such a way that

$$y := \mathbb{P}\left(\widehat{D} \leq 1 + K \mid \widehat{D} \geq 2\right) > 0. \tag{2.54}$$

The random variable $\widehat{D}$ has the size-biased distribution $\{q_k\}$. For a vertex u of degree two, we denote as in the previous case by A_u the event that u is good, and define $Y := \sum_u \mathbb{1}_{A_u}$ to be the number of good vertices. The strategy of the proof will be the same as in the previous case, the details change. So we will obtains bounds for the expected value and the variance of Y, and then use the second moment method to show that w.h.p $Y = \Omega(n^\epsilon)$. The rest of the proof will be similar to the case $d_{\min} \geq 3$ (based as before on the use of Proposition 2.26).

Consider the exploration process defined in Section 2.6.1, starting from a node $u \in V^*$. At the beginning, each step of the exploration process is an exponential with parameter two (since there are two yet-unmatched half-edges adjacent to the explored vertices). In each step, the probability that the new half-edge of the list L does not match to the other half-edge of L (which of course corresponds to the case that u is not in the giant component) is at least $1 - 1/n$. This follows by observing that there are at least n yet-unmatched half-edges (by $\nu > 1$), and by using Lemma 2.25 (which says that before $M \log n$ steps the exploration process meets a vertex of forward-degree at least two). By the forward-degree coupling arguments of Section 2.6.2, the probability that a new matched node be of forward-degree one is at least $\overline{\pi}_1^{(n)}$. This shows that, with probability at least $(1-1/n)\overline{\pi}_1^{(n)}$ the exploration process adds a new node of forward-degree one. This shows that the first step in the exploration process a vertex of forward-degree at least

two is added will be stochastically bounded below by a geometric random variable of parameter $(1-1/n)\overline{\pi}_1^{(n)}$. Each step takes rate two exponential time. Therefore,

$$\mathbb{P}(\mathcal{E}_u) = \mathbb{P}(\overline{T}_u(1) \geq (1-\epsilon)s_n) \geq \mathbb{P}\left(\mathrm{Exp}\left(2\left(1-(1-1/n)\overline{\pi}_1^{(n)}\right)\right) \geq (1-\epsilon)s_n\right).$$

In the last inequality we used the fact that a sum of a geometric (with parameter π) number of independent exponential random variables of rate μ is distributed as an exponential random variable of rate $(1-\pi)\mu$. Note that this in particular shows that

$$\mathbb{P}(\mathcal{E}_u) \geq (1-o(1))\exp(-2(1-q_1)(1-\epsilon)s_n) = (1-o(1))n^{\epsilon-1}.$$

By the coupling arguments of Section 2.6.2 (and by using Lemma 2.25), similar to the previous case $d_{\min} \geq 3$, we have

$$\mathbb{P}(\mathcal{E}'_u \mid \mathcal{E}_u) = (1 \pm o(1))y.$$

We conclude that

$$\mathbb{P}(A_u) = \mathbb{P}(\mathcal{E}'_u \mid \mathcal{E}_u)\mathbb{P}(\mathcal{E}_u) \;\geq\; (1 \pm o(1))y\, n^{\epsilon-1}.$$

This shows, as before, that

$$\mathbb{E}[Y] = \sum_{u \in V^*} \mathbb{P}(A_u) \;\geq\; (1 \pm o(1))\, y\, p_2\, n^{\epsilon}.$$

Moreover, for any pair of vertices $u, v \in V^*$ such that $C_u \cap C_v = \emptyset$, conditioning on A_u does not have much effect on the asymptotic of the degree distribution (by Lemma 2.25 the size of each component C_u is at most $M \log n$), and hence, we deduce again, by the coupling argument of Section 2.6.2, that for u and v such that $C_u \cap C_v = \emptyset$,

$$\mathbb{P}(A_v \cap A_u) = (1 \pm o(1))\mathbb{P}(A_u)\mathbb{P}(A_v).$$

We infer as before,

$$\begin{aligned}
\mathrm{Var}(Y) &= \mathbb{E}[Y^2] - \mathbb{E}[Y]^2 = \mathbb{E}\left[\sum_{u,v\in V^*} \mathbb{1}_{A_u}\mathbb{1}_{A_v}\right] - \mathbb{E}[Y]^2 \\
&= \mathbb{E}\left[\sum_{u,v\in V^*:\, C_u\cap C_v\neq\emptyset} \mathbb{1}_{A_u}\mathbb{1}_{A_v} + \sum_{u,v\in V^*:\, C_u\cap C_v=\emptyset} \mathbb{1}_{A_u}\mathbb{1}_{A_v}\right] - \mathbb{E}[Y]^2 \\
&= \mathbb{E}\left[\sum_{u\in V^*} \mathbb{1}_{A_u} \sum_{v\in V^*:\, C_u\cap C_v\neq\emptyset} \mathbb{1}_{A_v} + \sum_{u,v\in V^*:\, C_u\cap C_v=\emptyset} \mathbb{1}_{A_u}\mathbb{1}_{A_v}\right] - \mathbb{E}[Y]^2 \\
&\leq (K+1)(M\log n)\ \mathbb{E}[Y] + \mathbb{E}\left[\sum_{u,v\in V^*:\, C_u\cap C_v=\emptyset} \mathbb{1}_{A_u}\mathbb{1}_{A_v}\right] - \mathbb{E}[Y]^2 \qquad \text{(By Lemma 2.25)} \\
&= o(\mathbb{E}[Y]^2).
\end{aligned}$$

In the inequality above, we used Lemma 2.25 to bound w.h.p. the size of all C_w by $M\log n$ (for some large enough M) for any node w in the graph, and used the fact that if the event A_u holds, then there are at most K edges out-going from C_u. Each of the vertices v with the property that $C_u \cap C_v \neq \emptyset$ should be either already on C_u or connected with a path consisting only of vertices of degree two to C_u (in which case, this path should belong to C_v). A simple analysis then shows that the number of nodes v with the property that $C_u \cap C_v \neq \emptyset$ is bounded by $(K+1)M\log n$, and the inequality follows. The rest of the proof follows by using Proposition 2.26, similar to the case $d_{\min} \geq 3$.

Proof of the lower bound in the case $d_{\min} = 1$. Consider the 2-core algorithm, and stop the process the first time the number of nodes of degree one drops below $n^{1-\epsilon/2}$. Let V^* be the set of all nodes of degree one at this time. We denote by $\tilde{G}_n(V^*)$ the graph constructed by configuration model on the set of remaining nodes (this is indeed the V^*-augmented 2-core). Observe that proving the lower bound on the graph $\tilde{G}_n(V^*)$ gives us the lower bound on G_n.

Since $|V^*| = o(n/\log n)$, and the 2-core has linear size in n, w.h.p. the degree sequence of $\tilde{G}_n(V^*)$ has the same asymptotic as the degree sequence in the 2-core of G_n (see Section 2.5, Lemma 2.8 for more details). In particular, we showed in Section 2.5 that for the size-biased degree sequence of the 2-core's degree distribution, we have $\tilde{q}_1 = \beta_*$, and for its mean, we have $\tilde{\nu} = \nu$.

Repeating the coupling arguments of Section 2.6.2 and defining $\tilde{\tilde{\pi}}^{(n)}$ (similar to the definition

of $\overline{\pi}^{(n)}$) for the degree sequence of $\tilde{G}_n(V^*)$, we infer that $\tilde{\overline{\pi}}_1^{(n)} \to \beta_*$.

Similar as before, call a vertex u in V^* *good* if both the events $\mathcal{E}_u$ and $\mathcal{E}'_u$ hold. Recall the definition of the two events

$$\mathcal{E}_u := \left\{\overline{T}_u(1) \geq (1-\epsilon)s_n\right\}, \text{ and} \tag{2.55}$$

$$\mathcal{E}'_u := \{\deg(C_u) \leq K\}. \tag{2.56}$$

here the constant $K \geq 2$ is chosen with the property that $\tilde{q}_K > 0$ ($\tilde{q}$ is the size-biased probability mass function corresponding to the 2-core, c.f. Section 2.5).

Note that in this case $s_n = \frac{\log n}{(1-\beta_*)}$. Consider the exploration process starting from a node $u \in V^*$. At the beginning, each step of the exploration process is an exponential of rate one, and the probability that each new matched node be of forward-degree exactly one is at least $\tilde{\overline{\pi}}_1^{(n)}$. Similar to the case of $d_{\min} = 2$, we obtain

$$\begin{aligned}
\mathbb{P}(A_u) &\geq (1 \pm o(1))\tilde{q}_K \, \mathbb{P}\left(\mathrm{Exp}\left(1 - \tilde{\overline{\pi}}_1^{(n)}\right) \geq (1-\epsilon)s_n\right) \\
&= (1 \pm o(1))\tilde{q}_K \exp(-(1-\beta_*)(1-\epsilon)s_n) \\
&= (1 \pm o(1))\tilde{q}_K \, n^{-1+\epsilon}.
\end{aligned}$$

This shows that

$$\mathbb{E}[Y] = \sum_{u \in V^*} \mathbb{P}(A_u) \geq n^{1-\epsilon/2}(1 \pm o(1))\tilde{q}_K \, n^{-1+\epsilon} = (1 \pm o(1))\tilde{q}_K n^{\epsilon/2},$$

Similarly, we obtain that $\mathrm{Var}(Y) = o(\mathbb{E}[Y]^2)$, and the rest of the proof follows similarly to the precedent cases by using Proposition 2.26 for $\tilde{G}_n(V^*)$. Note that in $\tilde{G}_n(V^*)$, the number of vertices of degree one is $o(n) = o(|\tilde{G}_n(V^*)|)$ and thus, Proposition 2.26 can be applied.

At the present we are only left to prove Proposition 2.26.

2.8.1 Proof of Proposition 2.26

In this section we present the proof of Proposition 2.26. It is shown in [110, 129] (see Theorem 1.24 in Chapter 1) that the giant component of a random graph $G(n, (d_i)_1^n)$ for $(d_i)_1^n$ satisfying Condition 1.17 contains w.h.p. all but $o(n)$ vertices (since $\nu > 1$ and $u_0^{(n)} + u_1^{(n)} = o(n)$). This immediately shows that $\mathbb{P}(\mathrm{dist}_w(C_a, C_b) < \infty) = 1 - o(1)$. Define $t_n := \frac{1-\epsilon}{2(\nu-1)} \log n$. So to prove the proposition, we need to prove that $\mathrm{dist}_w(C_a, C_b)$ is lower bounded by t_n w.h.p. in the case where $\deg(C_b) = O(1)$ and either $\deg(C_a) = O(1)$ or a is chosen uniformly at random.

In the case where a is chosen uniformly at random, it is easy to deduce, by using Markov's inequality, that we have w.h.p. $\deg(C_a) \leq \log n$. Indeed, this is true since $\deg(C_a)$ is asymptotically distributed as $\left(D + \widehat{D} - 1 \mid \widehat{D} \geq 2\right)$, where $\widehat{D}$ is a random variable with the size-biased distribution, and D is independent of $\widehat{D}$ with the degree distribution $\{p_k\}$. (To show this, one can use the coupling argument of Section 2.6.2 to bound $\deg(C_a)$ stochastically from above.) And, since this latter random variable has finite moment (by Condition 1.17), by applying Markov's inequality, we obtain w.h.p. $\deg(C_a) \leq \log n$. This shows that in both cases stated in the proposition, we can assume that $\deg(C_b) = O(1)$ and $\deg(C_a) \leq \log n$.

We now consider the exploration process defined in Section 2.6.1 starting from C_a, i.e., we start the exploration process with $B = C_a$, and apply the steps one and two of the process. In a similar way we defined $T_a(i)$, we define $T_{C_a}(i)$ to be the time of the i-th step in this continuous-time exploration process. Similarly, let $\widehat{d}_{C_a}(i)$ be the forward-degree of the vertex added at i-th exploration step for all $i \geq 1$, and define

$$\widehat{S}_{C_a}(i) := \deg(C_a) + \widehat{d}_{C_a}(1) + ... + \widehat{d}_{C_a}(i) - i, \tag{2.57}$$

and define $S_{C_a}(i)$ similarly, so that we have $S_{C_a}(i) \leq \widehat{S}_{C_a}(i)$. Note that $T_{C_a}(i)$ obviously satisfies

$$T_{C_a}(i+1) - T_{C_a}(i) = \text{Exp}\left(S_{C_a}(i)\right) \geq_{st} Y_i \sim \text{Exp}\left(\widehat{S}_{C_a}(i)\right),$$

where the random variables Y_i are all independent.

Also, we infer (by Lemma 2.11) that

$$\widehat{S}_{C_a}(i) \leq_{st} \log n + \sum_{j=1}^{i} \overline{D}_j^{(n)} - i, \tag{2.58}$$

where $\overline{D}_j^{(n)}$ are i.i.d with distribution $\overline{\pi}^{(n)}$.

Let $\overline{\nu}^{(n)}$ be the expected value of $\overline{D}_1^{(n)}$ which is

$$\overline{\nu}^{(n)} := \sum_k k \overline{\pi}_k^{(n)},$$

and define $z_n = \sqrt{n/\log n}$. We will show later that the two growing balls in the exploration processes started from C_a and C_b, for a and b as in the proposition, will not intersect w.h.p. provided that they are of size less than z_n. We now prove that $T_{C_a}(z_n) \geq t_n$ with high probability.

For this, let us define

$$T'(k) \sim \sum_{i=1}^{k} \mathrm{Exp}\left(\log n + \sum_{j=1}^{i} \overline{D}_j^{(n)} - i\right),$$

where all the exponential variables in the above sum are independent, such that by the above arguments, we have

$$T_{C_a}(z_n) \geq_{st} T'(z_n).$$

We need the following lemma. (We define $\mathrm{Exp}(s) := +\infty$ for $s \leq 0$.)

Lemma 2.27. *Let $X_1, ..., X_t$ be a random process adapted to a filtration $\mathcal{F}_0 = \sigma[\emptyset], \mathcal{F}_1, ..., \mathcal{F}_t$, and let $\mu_i = \mathbb{E}X_i$, $\mathbf{\Sigma}_i = X_1 + ... + X_i$, $\Lambda_i = \mu_1 + ... + \mu_i$. Let $Y_i \sim \mathrm{Exp}(\Sigma_i)$, and $Z_i \sim \mathrm{Exp}(\Lambda_i)$, where all exponential variables are independent. Then we have*

$$Y_1 + ... + Y_t \geq_{st} Z_1 + ... + Z_t.$$

Proof. By Jensen's inequality, it is easy to see that for positive random variable X, we have

$$\mathrm{Exp}(X) \geq_{st} \mathrm{Exp}(\mathbb{E}X).$$

Then by induction, it suffices to prove that for a pair of random variables X_1, X_2 we have $Y_1 + Y_2 \geq_{st} Z_1 + Z_2$. We have

$$\begin{aligned} \mathbb{P}(Y_1 + Y_2 > s) &= \mathbb{E}_{X_1}[\mathbb{P}(Y_1 + Y_2 > s | X_1)] \\ &\geq \mathbb{E}_{X_1}[\mathbb{P}(\mathrm{Exp}(X_1) + \mathrm{Exp}(X_1 + \mu_2) > s)] \\ &\geq \mathbb{P}(Z_1 + Z_2 > s). \end{aligned}$$

□

We infer by Lemma 2.27,

$$T'(z_n) \geq_{st} \sum_{i=0}^{z_n} \mathrm{Exp}\left(\log n + (\overline{\nu}^{(n)} - 1)i\right) =: T^*(z_n),$$

where all exponential variables are independent.

We now let $b_n := \log n - (\overline{\nu}^{(n)} - 1)$, so that we have

$$\begin{aligned}\mathbb{P}(T^*(z_n) \leq t) &\leq \int_{\sum x_i \leq t} e^{-\sum_{i=1}^{z_n}((\overline{\nu}^{(n)}-1)i+b_n)x_i} dx_1 ... dx_{z_n} \prod_{i=1}^{z_n}((\overline{\nu}^{(n)}-1)i+b_n) \\ &= \int_{0\leq y_1 \leq \cdots \leq y_{z_n} \leq t} e^{-(\overline{\nu}^{(n)}-1)\sum_{i=1}^{z_n} y_i} e^{-b_n y_{z_n}} dy_1 ... dy_{z_n} \prod_{i=1}^{z_n}((\overline{\nu}^{(n)}-1)i+b_n),\end{aligned}$$

where $y_k = \sum_{i=0}^{k-1} x_{z_n - i}$. Letting y play the role of y_{z_n}, and accounting for all permutations over $y_1, ..., y_{z_n - 1}$ (giving each such variable the range $[0, y]$), we obtain

$$\begin{aligned}\mathbb{P}(T^*(z_n) \leq t) &\leq (\overline{\nu}^{(n)}-1)^{z_n} \frac{\prod_{i=1}^{z_n}(i + \frac{b_n}{\overline{\nu}^{(n)}-1})}{(z_n-1)!} \\ &\quad \cdot \int_0^t e^{-(\overline{\nu}^{(n)}-1+b_n)y} \left(\int_{[0,y]^{z_n-1}} e^{-(\overline{\nu}^{(n)}-1)\sum_{i=1}^{z_n-1} y_i} dy_1 ... dy_{z_n-1} \right) dy \\ &= z_n \frac{\prod_{i=1}^{z_n}(i + \frac{b_n}{\overline{\nu}^{(n)}-1})}{z_n!} (\overline{\nu}^{(n)}-1) \\ &\quad \cdot \int_0^t e^{-(\overline{\nu}^{(n)}-1+b_n)y} \left(\prod_{i=1}^{z_n-1} \int_0^y (\overline{\nu}^{(n)}-1) e^{-(\overline{\nu}^{(n)}-1)y_i} dy_i \right) dy \\ &= z_n \prod_{i=1}^{z_n}(1 + \frac{b_n}{(\overline{\nu}^{(n)}-1)i})(\overline{\nu}^{(n)}-1) \\ &\quad \cdot \int_0^t e^{-(\overline{\nu}^{(n)}-1+b_n)y} \left(1 - e^{-(\overline{\nu}^{(n)}-1)y}\right)^{z_n-1} dy \\ &\leq c z_n^{\frac{b_n}{\overline{\nu}^{(n)}-1}+1} (\overline{\nu}^{(n)}-1) \int_0^t e^{-(\overline{\nu}^{(n)}-1+b_n)y} \left(1 - e^{-(\overline{\nu}^{(n)}-1)y}\right)^{z_n-1} dy,\end{aligned}$$

where $c > 0$ is an absolute constant. Recall that $t_n = \frac{1-\epsilon}{2(\nu-1)} \log n$, and $z_n = \sqrt{n/\log n}$. Now we use the fact that $\left(1 - e^{-(\overline{\nu}^{(n)}-1)y}\right)^{z_n-1} \leq e^{-n^\alpha}$, for some $\alpha > 0$ and for all $0 \leq y \leq t_n$. We infer

$$\mathbb{P}(T^*(z_n) \leq t_n) \leq c(\overline{\nu}^{(n)}-1) z_n^{\frac{b_n}{\overline{\nu}^{(n)}-1}+1} \int_0^{t_n} e^{-n^\alpha} dy = o(n^{-4}),$$

since $b_n = O(\log n)$. Hence, we have w.h.p.

$$|B_w(C_a, t_n)| \leq z_n.$$

(Here naturally, for $W \subseteq V$, we let $B_w(W, t) = \{b, \text{ such that } \text{dist}_w(W, b) \leq t\}$.)

Similarly for b, and exposing $B_w(C_b, t_n)$, again w.h.p we obtain a set of size at most z_n. Now remark that, because each matching is uniform among the remaining half-edges, the probability of hitting $B_w(C_a, t_n)$ is at most $\widehat{S}_{C_a}(z_n)/n$.

Let $\epsilon_n := \log\log n$. By Markov's inequality we have

$$\begin{aligned}\mathbb{P}\left(\widehat{S}_{C_a}(z_n) \geq z_n\epsilon_n\right) &\leq \mathbb{E}\widehat{S}_{C_a}(z_n)/z_n\epsilon_n \\ &= \frac{K + (\overline{\nu}^{(n)} - 1)(z_n + \lambda_n)}{z_n\epsilon_n} = o(1).\end{aligned}$$

We conclude

$$\begin{aligned}\mathbb{P}\left(B_w(C_a, t_n) \cap B_w(C_b, t_n) \neq \emptyset\right) &\leq \mathbb{P}\left(|B_w(C_a, t_n)| > z_n\right) + \mathbb{P}\left(|B_w(C_b, t_n)| > z_n\right) \\ &\quad + \mathbb{P}\left(\widehat{S}_{C_a}(z_n) \geq z_n\epsilon_n\right) + \epsilon_n z_n^2/n \\ &= o(1).\end{aligned}$$

This completes the proof of Proposition 2.26.

2.9 Proof of Corollary 2.6

First we prove that for an r-regular graph $G = (V, E)$, the dynamic evolution of informed nodes in continuous-time broadcast when each node is endowed with a Poisson process with rate one corresponds exactly to the flooding time with exponential random weights on edges with mean r. For a node $a \in V$, let $\mathcal{I}(a, t)$ denote the set of informed nodes at time t when the broadcast process is started from a. Indeed, using a coupling argument, we will show that the random map $\mathcal{I}(a, .)$ from $[0, \infty)$ to subsets of V has the same law as $B_w(a, .)$ when the weights are exponentials with mean r. To this aim, from the asynchronous broadcasting model, we construct weights on the edges of the graph and show that these weights are independent exponential with mean r.

For $v \in V$, let $T(v)$ denote the time at which the node v becomes informed in the asynchronous broadcast model. For an edge $\{u, v\} \in E$, let $\tau_i(u, v)$ be the i-th time the node u contacts the node v. The weight of the edge $e = \{u, v\}$ is defined as follows. Without loss of generality assume that $T(u) < T(v)$, i.e., if u is informed before v, then

$$w_e := \min_i\{\ \tau_i(u, v) - T(u) \mid \tau_i(u, v) > T(u)\ \}.$$

Thanks to the memoryless property of the Poisson process, $\{w_e, e \in E\}$ are independent exponential random variables with mean r. Indeed since u has r neighbors, the waiting time (after $T(u)$) until the node u contacts v is distributed as a sum of a geometric (with parameter $1/r$) number of independent exponential random variables with parameter one. Now recall that a sum of a geometric (with parameter s) number of independent exponential random variables with parameter μ is distributed as an exponential random variable with parameter $(1-s)\mu$ (i.e., with mean $1/(1-s)\mu$).

In addition, it is easy to see that we have $\mathcal{I}(a,t) = B_w(a,t)$ for all $t \geq 0$. Hence the asynchronous broadcast time corresponds to the flooding time with exponential weights with mean r. We conclude the proof by using Corollary 2.4,

$$\begin{aligned} \mathrm{ABT}(G) &= r\left(\frac{1}{r-2}+\frac{1}{r}\right)\log n + o_p(\log n) \\ &= 2\left(\frac{r-1}{r-2}\right)\log n + o_p(\log n). \end{aligned}$$

Discussion

In this chapter, we have studied the impact of the i.i.d. exponential random edge weights on distances in configuration model. It would be of interest to study the effect of weights even further, for example, by studying the case where the weights are i.i.d. random variables with distribution equal to E^s, where E is an exponential random variable with mean 1 and $s \in [0;1)$. The case $s = 0$ corresponds to the graph distance as studied in [94, 69], while the case $s = 1$ corresponds to the case with i.i.d. exponential weights as studied here, and in [24]. (e.g., see [22] for the similar results in the case of complete graphs). These and some other related issues are subject of a future work we are currently investigating.

Chapter 3

Bootstrap Percolation, Diffusion, and Cascades

Abstract. *In this chapter, we consider diffusion in random graphs with given vertex degrees. Our diffusion model can be viewed as a variant of a cellular automaton growth process: assume that each node can be in one of the two possible states, inactive or active. The parameters of the model are two given functions* $\theta : \mathbb{N} \to \mathbb{N}$ *and* $\alpha : \mathbb{N} \to [0, 1]$*. At the beginning of the process, each node* v *of degree* d_v *becomes active with probability* $\alpha(d_v)$ *independently of the other vertices. Presence of the active vertices triggers a percolation process: if a node* v *is active, it remains active forever. And if it is inactive, it will become active when at least* $\theta(d_v)$ *of its neighbors are active. In the case where* $\alpha(d) = \alpha$ *and* $\theta(d) = \theta$*, for each* $d \in \mathbb{N}$*, our diffusion model is equivalent to what is called bootstrap percolation. Our main result is a theorem which enables us to find the final proportion of the active vertices in the asymptotic case, i.e., when* $n \to \infty$*. This is done via analysis of the process on the multigraph counterpart of the graph model.*

Contents

3.1 Introduction

There is a vast literature on epidemics on complex networks (see for example [54, 55, 134, 121, 122]). The diffusion model we consider in this chapter is a generalization of bootstrap percolation in an arbitrary graph (modeling a given network). Let $G = (V, E)$ be a connected graph. Given two vertices i and j, we write $i \sim j$ if $\{i, j\} \in E$. The threshold associated to a node i is $\theta(d_i)$, where d_i is the degree of i, and $\theta : \mathbb{N} \to \mathbb{N}$ is given fixed function. Assume that each node can be in one of the two possible states: inactive or active. Let $\alpha : \mathbb{N} \to [0, 1]$ be a fixed given function. At time 0, each node i becomes active with probability $\alpha(d_i)$ independently of all the other vertices. At time $t \in \mathbb{N}$, the state of each node i will be updated according to a deterministic process: if a node i was active at time $t-1$, it will remains active at time t. Otherwise, i will become active if at least $\theta(d_i)$ of its neighbors were active at time $t-1$.

In the case where $\alpha(d) = \alpha$ and $\theta(d) = \theta$, for each $d \in \mathbb{N}$, our diffusion model is equivalent to what is called *bootstrap percolation*. This model has a rich history in statistical physics, mostly on $G = \mathbb{Z}^d$ and finite boxes. Bootstrap percolation serves as a useful model to describe a growing list of complex phenomena, including neuronal activity [35, 43, 147, 80], jamming transitions and glassy dynamics [81, 143, 148], and magnetic systems [142]. Bootstrap percolation was first mentioned and studied in the statistical physics literature by Chalupa et al. in [40]. The problem of complete occupation on $\mathbb{Z}^2$ was solved by van Enter in [61]. A short physics survey is [4]. Bootstrap percolation also has connections to the dynamics of the Ising model at zero temperature [71]. Some further references for bootstrap percolation on various graphs are Cerf and Manzo [39], Holroyd [97] (grids); Balogh and Bollobás [14] (hypercube); Balogh, Peres and Pete [15], Fontes and Schonmann [70] (infinite trees); Balogh and Pittel [16] (random regular graphs); and recently Janson [106] (Erdős-Rényi random graphs).

Let G be a graph with n nodes, i.e., $|V| = n$. Let A denote the adjacency matrix of G, with $A_{ij} = 1$ if $i \sim j$, and $A_{ij} = 0$ otherwise. The state of the network at time t can be described by the vector $(X_t(i))_{i=1}^n$: $X_t(i) = 1$ if the node i is active at time t, and $X_t(i) = 0$ otherwise. Remark that $X_0(i)$ is a Bernoulli random variable with parameter $\alpha(d_i)$. The evolution of this vector at time $t+1$ follows the following functional equation, i.e., at each time step $t+1$, each node v applies:

$$X_{t+1}(i) \quad = \quad X_t(i) + (1 - X_t(i))\mathbb{1}\left(\sum_j A_{ij} X_t(j) \geq \theta(d_i)\right), \tag{3.1}$$

where $\mathbb{1}(\Xi)$ denotes the indicator of an event Ξ; this is 1 if Ξ holds and 0 otherwise.

From the definition, $X_t(i)$ is non-decreasing. Indeed, the equation (3.1) implies again that $X_t(i+1) \geq X_t(i)$. Define $\Phi^{(n)}(\alpha, \theta, t)$ as

$$\Phi^{(n)}(\alpha, \theta, t) := n^{-1} \sum_{j=1}^{n} \mathbb{E}[X_t(j)].$$

We are interested in finding the asymptotic value when $n \to \infty$, of

$$\Phi^{(n)}(\alpha, \theta) := \lim_{t \to \infty} \Phi^{(n)}(\alpha, \theta, t)$$

in the case of random graphs with given vertex degrees (e.g., see Section 1.4).

Remark 3.1. We have to emphasize that the results of this chapter can be as well applied to some other random graphs models by conditioning on the vertex degrees. For example, for the Erdős-Rényi random graph $\mathrm{ER}(n, p)$, where every edge is present with probability p, with $np \to \lambda \in (0, \infty)$, the assumptions hold with (p_k) a Poisson distribution with mean λ:

$$p_k = e^{-\lambda} \frac{\lambda^k}{k!}.$$

3.1.1 Main results

We now present the main results of this chapter. Consider a random graph $G(n, (d_i)_1^n)$, where the degree sequence $(d_i)_1^n$ satisfies Condition 1.17. Let D be a random variable with integer values, and with distribution $\mathbb{P}(D = r) = p_r$, $r \in \mathbb{N}$. The two functions $\alpha : \mathbb{N} \to [0, 1]$ and $\theta : \mathbb{N} \to \mathbb{N}$ are given as before. We define the function $f_{\alpha,\theta} : [0, 1] \to \mathbb{R}$ as follows

$$f_{\alpha,\theta}(y) := \lambda y^2 - y\, \mathbb{E}\left[\left(1 - \alpha(D)\right) D\, \mathbb{1}\left(\mathrm{Bin}(D - 1, 1 - y) < \theta(D)\right) \right]. \tag{3.2}$$

Let $y^* = y^*_{\alpha,\theta}$ be the largest solution to $f_{\alpha,\theta}(y) = 0$, i.e.,

$$y^* := \sup \{ y \in [0,1] \mid f_{\alpha,\theta}(y) = 0 \}.$$

Remark that such y^* exists because $y = 0$ is a solution, and $f_{\alpha,\theta}$ is continuous.

The main result of this chapter is the following.

Theorem 3.2. *Consider a random graph $G(n,(d_i)_1^n)$ where the degree sequence $(d_i)_1^n$ satisfies Condition 1.17. Then we have:*

1. *If $y^* = 0$, i.e., if $f_{\alpha,\theta}(y) > 0$ for all $y \in (0,1]$, then $\Phi^{(n)}(\alpha,\theta) = 1 - o_p(1)$.*
2. *If $y^* > 0$ and furthermore y^* is not a local minimum point of $f_{\alpha,\theta}(y)$, then*

$$\Phi^{(n)}(\alpha,\theta) = 1 - \mathbb{E}\,[\,(1-\alpha(D))\,\mathbb{1}\,(\,\mathrm{Bin}(D,1-y^*) \,<\, \theta(D))\,] + o_p(1).$$

It is easy to see that in the case $y^* = 0$, we necessarily have $\theta(d) < d$ for all $d \in \mathbb{N}$ with $p_d > 0$ and $\alpha(d) < 1$.

We now look at the diffusion with one initial active node. Let us call the following condition the cascade condition:

$$\mathbb{E}\,[\,D\,] \,<\, \mathbb{E}\,\big[\,D(D-1)\mathbb{1}_{(\theta(D)=1)}\,\big]. \tag{3.3}$$

The second theorem of this chapter is the following:

Theorem 3.3 (The cascade condition). *Consider a random graph $G(n,(d_i)_1^n)$ where the degree sequence $(d_i)_1^n$ satisfies Condition 1.17.*

- *If the cascade condition (3.3) is satisfied, then there exists w.h.p. a single node v which can trigger a global cascade, i.e., v can activate a strictly positive fraction of the total population w.h.p.*
- *If $\mathbb{E}\,[\,D\,] \,>\, \mathbb{E}\,\big[\,D(D-1)\mathbb{1}_{(\theta(D)=1)}\,\big]$, then for any $k = o(n)$, we have*

$$|\mathcal{C}(1,2,...,k)| = o_p(n),$$

where for $W \subseteq V$, $\mathcal{C}(W)$ denote the final active nodes when we start the diffusion with initial active nodes W.

Remark 3.4. We note that in the case where $\theta(d) = \theta d$, Watts [151] obtained the same condition by a heuristic argument validated through simulations. Our theorem provides as a very special case a mathematical proof of his heuristic results.

In the rest of this section, we provide some of the applications of our main theorems above. But let us first briefly explain the methods used to derive Theorems 3.2 and 3.3. The base of our approach is some standard techniques similar to those used by Balogh and Pittel [16] for the special d-regular case problem, Cain and Wormald [38] for the k-core problem and Molloy and Reed [129] for the giant component problem. This means we consider the diffusion process on the random configuration model and describe the dynamics of the diffusion by a Markov chain. The proof of Theorem 3.2 is mainly based on a method introduced by Wormald in [154] for the analysis of a discrete random process by using differential equations (see Appendix A.4 for more details). However, our model is more general and new difficulties arise in treating the Markov chain and proving the convergence results. One special difficulty is that, contrary to [16], here the number of variables is a function of n (and so is not constant). The proof of Theorem 3.3 is based on Theorem 3.2 and a theorem of Janson [103] (Theorem 1.30 in Chapter 1) for the study of percolation in a random graph with given vertex degrees. We refer to Section 3.4.3 for more details.

3.1.2 The k-core in random graphs with given degree sequence

As discussed in Section 1.4.2, the *k-core* of a given graph G, denoted by $\mathrm{Core}_k(G)$, is the largest induced subgraph of G with minimum vertex degree at least k. Let $\mathrm{Core}_k^{(n)}$ be the k-core of the graph $G(n,(d_i)_1^n)$.

The existence of a large k-core in a random graph with a given degree sequence has been studied by several authors (see Section 1.4.2). Theorem 3.2 allows us to unify all these results into a single theorem. In fact by assuming the functions α and θ to be equal to $\hat{\alpha}(d) = \mathbb{1}(d < k)$ and $\hat{\theta}(d) = (d-k+1)_+ = (d-k+1)\mathbb{1}(d \geq k)$ respectively, we obtain

$$\frac{\mathrm{Core}_k^{(n)}}{n} = 1 - \Phi^{(n)}(\hat{\alpha}, \hat{\theta}).$$

Let $\hat{y} = y^*_{\hat{\alpha},\hat{\theta}}$ be the largest solution to $f_{\hat{\alpha},\hat{\theta}}(y) = 0$.

Corollary 3.5. *Consider a random graph $G(n,(d_i)_1^n)$ where the degree sequence $(d_i)_1^n$ satisfies Condition 1.17. Let $k \geq 2$ be fixed, and let $\mathrm{Core}_k^{(n)}$ be the k-core of $G(n,(d_i)_1^n)$. Then we have:*

1. *If $\hat{y} = 0$, i.e., if $f_{\hat{\alpha},\hat{\theta}}(y) > 0$ for all $y \in (0,1]$, then w.h.p.* $\mathrm{Core}_k^{(n)} = o(n)$.

2. *If $\hat{y} > 0$ and furthermore $\hat{y}$ is not a local minimum point of $f_{\hat{\alpha},\hat{\theta}}(y)$, then w.h.p.*

$$\mathrm{Core}_k^{(n)} = n\,\mathbb{P}\left(\,\mathrm{Bin}(D,\hat{y}) \geq k\,\right) + o(n).$$

3.1.3 Bootstrap percolation on random regular graphs

In the case of random regular graphs, i.e., in the case $d_i = d$ for all i, our diffusion model is equivalent to *bootstrap percolation*. Bootstrap percolation on the random regular graph $G(n,d)$ with fixed vertex degree d was studied by Balogh and Pittel in [16]. By Theorem 3.2 we can recover a large part of their results. Let $A_f^{(n)}$ be the final set of active vertices. We find that

Corollary 3.6 (Balogh-Pittel [16])**.** *Let the three parameters $\alpha, \theta \in [0,1]$ and $d \geq 1$ be given with $1 \leq \theta \leq d-1$. Consider bootstrap percolation on the random d-regular graph $G(n,d)$ in which each vertex is initially active independently at random with probability α and the threshold is θ. Let α_c be defined by*

$$\alpha_c := 1 - \inf_{0<y\leq 1} \frac{y}{\mathbb{P}\big(\mathrm{Bin}(d-1,1-y) \leq \theta - 1\big)}\,.$$

We have

(i) If $\alpha > \alpha_c$, then $|A_f^{(n)}| = n - o_p(n)$.

(ii) If $\alpha < \alpha_c$, then w.h.p. a positive proportion of the vertices remain inactive. More precisely, if $y^* = y^*(\alpha)$ is the largest $y \leq 1$ such that $\mathbb{P}\left(\mathrm{Bin}(d-1,1-y) \leq \theta - 1\right)/y = (1-\alpha)^{-1}$, then

$$\frac{|A_f^{(n)}|}{n} \xrightarrow{p} 1 - (1-\alpha)\mathbb{P}\big(\,\mathrm{Bin}(d,1-y^*) \leq \theta - 1\,\big) < 1.$$

Proof. It remains only to show that in case (ii), y^* is not a local minimum point of

$$f_{\alpha,\theta}(y) = dy^2\big(1 - (1-\alpha)\frac{\mathbb{P}\big(\mathrm{Bin}(d-1,1-y) \leq \theta - 1\big)}{y}\big).$$

In fact, $\mathbb{P}\big(\,\mathrm{Bin}(d-1,1-y) \leq \theta - 1\big)/y$ is decreasing in y when $\theta = d-1$ and has only one minimum point when $\theta < d-1$ (see [16] for details). Thus for $\theta < d-1$, the only local minimum point is the global minimum point $\hat{y}$ with $\mathbb{P}\big(\mathrm{Bin}(d-1,1-\hat{y}) \leq \theta - 1\big)/\hat{y} = (1-\alpha_c)^{-1}$ and otherwise, when $\theta = d-1$, there is no local minimum point. □

In this case, Balogh and Pittel [16] have also studied the threshold in greater detail by allowing α to depend on n; we have

- if $n^{1/2}(\alpha(n) - \alpha_c) \to \infty$, then w.h.p. $|A_f^{(n)}| = n$;
- if $n^{1/2}(\alpha_c - \alpha(n)) \to \infty$, then w.h.p. $|A_f^{(n)}| < n$ and furthermore

$$|A_f^{(n)}| = n\left(1 - (1-\alpha(n))\mathbb{P}\big(\,\mathrm{Bin}(d, 1-y^*) \le \theta - 1\,\big)\right) + O_p(n^{1/2}(\alpha_c - \alpha(n))^{-1/2}).$$

It would be interesting to generalize these results to our case. Note that Balogh and Pittel [16] do not use Wormald's theorem. Indeed they analyze directly the system of differential equations via exponential supermartingales by using its integrals to show that the percolation process undergoes relatively small fluctuations around the deterministic trajectory.

3.1.4 Contagion threshold for random networks

A famous example of diffusion is given by *network coordination game* proposed by Morris [130], that we briefly explain below. Other references in this area are [115, 119, 100].

Consider the following game-theoretic diffusion model: At time t, each node plays a 2×2 game with each neighbor and chooses an action from the space $S = \{A, B\}$. Payoffs from each interaction in each period are given by a function $\pi(s, s')$ where $s, s' \in S$, and they are summarized in the following symmetric matrix: $\begin{pmatrix} x & w \\ z & y \end{pmatrix}$. In other words, $\pi(A, A) = x$, $\pi(A, B) = w$, $\pi(B, A) = z$ and $\pi(B, B) = y$. Assume that $y > w$ and $x > z$, which basically means that the game is a coordination game (whose strict Nash equilibrium are (A, A) and (B, B)). An individual's payoff is simply the sum of the payoffs obtained across all the bilateral games in which (s)he is involved.

Then if the degree of node i is d_i, and n_i^B is her number of neighbors playing B, then the payoff to i from choosing A is

$$\Pi_i(A) = wn_i^B + x(d_i - n_i^B),$$

while the payoff from choosing B is

$$\Pi_i(B) = yn_i^B + z(d_i - n_i^B).$$

This implies that, if the proportion of neighbors choosing B is higher than

$$q := \frac{x-z}{x-z+y-w}$$

then i's best response is to choose B. Otherwise i chooses A.

A number of qualitative insights can be derived from this simple diffusion model. Clearly, a network where all nodes play A is an equilibrium of the game as in the state where all nodes play B. Consider a network where all nodes initially play A. If a small number of nodes are forced to adopt strategy B and we apply best-response updates to other nodes in the network, then these nodes will be repeatedly applying the following rule: switch to B if enough of your neighbors have already adopted B. There can be a cascading sequence of nodes switching to B such that a network-wide equilibrium is reached in the limit. This equilibrium may involve uniformity with all nodes adopting B or it may involve coexistence, with the nodes partitioned into a set adopting B and a set sticking to A. In [130], Morris considers the case of an infinite graph G and provides graph-theoretic characterizations of the different types of equilibrium that can arise.

We now compute the contagion threshold for random networks. We define the contagion threshold of the graph to be the maximum q for which a single individual (who adopts B) can trigger a global cascade, i.e. activate a strictly positive fraction of the total population, with positive probability. This notion is the natural extension of the contagion threshold defined in [130] for regular graphs.

Then as a corollary of Theorem 3.3 (by setting $\theta(d) = \lceil qd \rceil$) we have:

Corollary 3.7. *Consider a random graph $G(n,(d_i)_1^n)$ where the degree sequence $(d_i)_1^n$ satisfies Condition 1.17. Then the contagion threshold q_c is given by*

$$q_c = \sup\left\{ q : \sum_{2 \le r < q^{-1}} r(r-1)p_r > \lambda \right\}. \tag{3.4}$$

We note that in this case, Watts [151] obtained the same condition by a heuristic argument validated through simulations. Recently in [123], Lelarge analyzes the possible equilibria of this game and identifies conditions for the coexistence of both strategies. We refer to [123] for more on this.

3.1.5 Simulation

We consider the bootstrap percolation model, i.e., where $\theta(d) = \theta$, and $\alpha(d) = \alpha$ over all nodes. We assume a Gaussian distribution for degrees, i.e., we let for $k \in \mathbb{N}$;

$$\mathbb{P}(D = k) = exp\left(\frac{-(k-\overline{k})^2}{2\sigma^2}\right)/Z(\overline{k}, \sigma),$$

where $Z(\overline{k}, \sigma)$ is a normalizing constant. We assume $\overline{k} = 50$ and $\sigma = 15$ based on the experimental results on neural networks [35, 43, 145].

Figure 3.1 shows the three dimensional representation of the final activated nodes

$$\Phi(\alpha, \theta) := \lim_{n\to\infty} \Phi^{(n)}(\alpha, \theta)$$

as a function of α and θ. This shows that both parameter θ and α have transition values, α_c and θ_c, where the solution changes qualitatively.

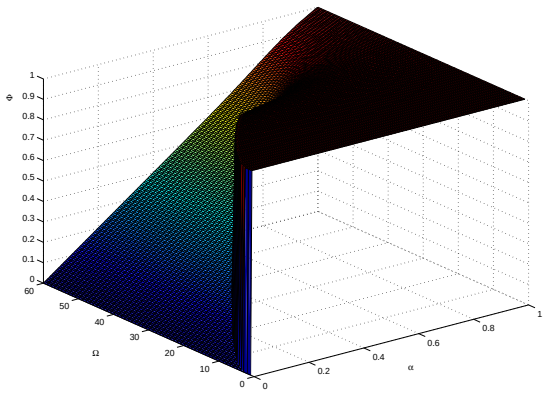

Figure 3.1: The final fraction of activated nodes as a function of α and θ. Here $\mathbb{P}(D = k) \sim exp\left(\frac{-(k-\overline{k})^2}{2\sigma^2}\right)$ with $\overline{k} = 50$ and $\sigma = 15$.

Let us assume that θ is fixed. Then there exists a critical value for the fraction of initially activated nodes (i.e., α) beyond which the global activity jumps to an almost complete activation of the network while below this critical value the diffusion essentially does not spread. Indeed as we can see in Figure 3.2, the map $\alpha \to \Phi(\alpha) := \lim_{n\to\infty} \Phi^{(n)}(\alpha, \theta)$ exhibits a point of discontinuity.

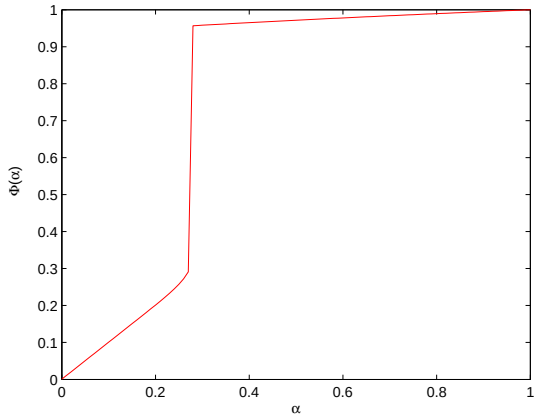

Figure 3.2: The final fraction of activated nodes as a function of α in bootstrap percolation model. Here $\mathbb{P}(D = k) \sim exp\left(\frac{-(k-\overline{k})^2}{2\sigma^2}\right)$ with $\overline{k} = 50$, $\sigma = 15$ and $\theta = 25$.

In the next section we consider the continuous-time version of the bootstrap percolation model, and we derive a general bound for the fraction of active nodes in terms global graph properties. Then in Section 3.3 we describe an approximation to the local structure of the graph by an appropriate branching process, and give a heuristic argument which leads quickly to derive Theorem 3.2. Then the diffusion process on $G^*(n, (d_i)_1^n)$ is studied in detail. The proof of our results are based on the use of differential equations for solving discrete random processes, and this is due to Wormald [154] (This is also discussed in Section A.4). The proofs of our main results, Theorem 3.2 and Theorem 3.3, are given in Section 3.4.

3.2 Continuous-time Dynamic

We first consider the following continuous-time version of the bootstrap percolation model. Assume that each non infected node i updates its state at rate one, and it becomes active if $\sum_{j\sim i} X_t(j) > \theta$. The state at time t is represented by a vector X_t. Denote by A the adjacency matrix of the graph G, and let $\lambda_1(A)$ be the spectral radius of A, namely, its largest eigenvalue. In addition we will assume that the graph is connected, and $\lambda_1(A)$ has multiplicity one. Therefore

we have, $X_0(i) = \text{Ber}(\alpha)$, for all $i \in V$, and

$$X(i) : 0 \to 1 \quad \text{at rate } \mathbb{1}\left(\sum_j A_{ij} X_t(j) > \theta\right).$$

Note that $\mathbb{1}\left(\sum_j A_{ij} X_t(j) > \theta\right) \leq \frac{\sum_j A_{ij} X_t(j)}{\theta}$. We now consider the continuous time Markov process $Z_t = (Z_t(i))_{i \in V}$, with $Z_0 = X_0$, and transition rate:

$$Z(i) : k \to k+1 \quad \text{at rate } \frac{\sum_{j=1}^n A_{ij} Z_t(j)}{\theta},$$

standard coupling arguments yield $X_t \leq_{st} Z_t$ for all $t \geq 0$, where $X \leq_{st} Z$ denotes that Z stochastically dominates X (see Section A.2). This implies that $\sum_{i=1}^n \mathbb{E}(X_t(i)) \leq \sum_{i=1}^n \mathbb{E}(Z_t(i))$. Moreover, the transition rates of the process Z are such that

$$\frac{d\mathbb{E}[Z_t]}{dt} = \frac{A}{\theta}\mathbb{E}[Z_t].$$

Hence we obtain

$$\mathbb{E}[Z_t] = e^{\frac{t}{\theta}A}\, \mathbb{E}[Z_0]. \tag{3.5}$$

Using Cauchy-Schwartz inequality, we obtain that $\sum_{i=1}^n \mathbb{E}(Z_t(i)) \leq ||\mathbb{E}(Z_t)||_2 ||1||_2$. Combining this with (3.5), we have that

Proposition 3.8. *Let β_t be the proportion of nodes that are active by time t. Then*

$$\beta_t := \frac{\sum_{i=1}^n \mathbb{E}(X_t(i))}{n} \leq \alpha e^{\frac{\lambda_1(A)}{\theta} t}.$$

Moreover if the G is a regular graph with degree d, then, using the spectral decomposition of the matrix $e^{\frac{t}{\theta}A}$, we have that

$$\beta_t \leq \frac{\alpha}{d} e^{\frac{d}{\theta} t}.$$

The above result states that the number of active nodes increases at most exponentially in time and that the speed is given by $\frac{\lambda_1(A)}{\theta}$. Similar results have been found in [78] in the case of the Susceptible-Infected-Susceptible (SIS) epidemic.

3.3 Branching Process Approximation

In this section we give a heuristic argument which leads quickly to derive our main result, Theorem 3.2. Intuitively, it is useful to note that a random graph $G^*(n, (d_i)_1^n)$ under Condition 1.17

locally behaves like a branching process. Hence we start our analysis by looking at the branching process. If we start with a given vertex x, then the number of neighbors (the first generation in the branching process) has distribution p_r. This is not true for the second generation. 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 is for $k \geq 1$: $q_{k-1} = \frac{kp_k}{\lambda}$. Recall that the $k-1$ on the left-hand side comes from the fact that we used up one edge connecting to the vertex. Note that q has finite mean (by Condition 1.17)

$$\nu := \sum_j jq_j = \frac{1}{\lambda} \sum_{j=1}^{\infty} j(j-1)p_j. \tag{3.6}$$

For any $d > 0$, we let T_d denote the random rooted tree with a fixed root $\emptyset$ generated as follows. First draw an integer k with distribution p_k, and connect the root $\emptyset$ to k offspring. Then recursively, for each node in the last generation, generate an integer k independently with distribution q_k, and connect the node to k new nodes. This is repeated until the tree has d generations. Notice that the random infinite tree T_∞ is well defined. Now we look at the contagion process on the infinite tree T_∞.

For a node i, we denote by $\mathrm{gen}(i) \in \mathbb{N}$, the generation of i, i.e., the length of the minimal path from $\emptyset$ to i. Also we denote $i \to j$ if i belongs to the children of j, i.e., $\mathrm{gen}(i) = \mathrm{gen}(j) + 1$ and j is on the minimal path from $\emptyset$ to i. For an edge (i, j) with $i \to j$, we denote by $T_{i \to j}$ the sub-tree of T_∞ with root i obtained by the deletion of edge (i, j) from T_∞. Now consider the contagion process in T_∞. We encode the initial active population by a vector χ, where $\chi_i = 1$ if the node i is active and $\chi_i = 0$ otherwise. Then χ_i is a Bernoulli random variable with parameter $\alpha(d_i)$ independent of everything else. For a given vector χ, we say that node $i \neq \emptyset$ is infected from $T_{i \to j}$ if the node i is infected in $T_{i \to j}$ with the same threshold and vector χ for $T_{i \to j}$ as T_∞. We denote by $Y_t(i)$ the corresponding indicator function with value 1 if i is infected from $T_{i \to j}$ at time t and 0 otherwise.

Lemma 3.9. *We have*

$$X_{t+1}(\emptyset) = 1 - (1 - \chi_\emptyset) \mathbb{1}\left(\sum_{i \sim \emptyset} Y_t(i) < \theta(d_\emptyset) \right). \tag{3.7}$$

Proof. We have

$$X_{t+1}(\emptyset) = 1 - (1 - \chi_\emptyset) \mathbb{1}\left(\sum_{i \sim \emptyset} X_t(i) < \theta(d_\emptyset) \right),$$

and let

$$Z_{t+1}(\emptyset) = 1 - (1 - \chi_\emptyset)\mathbb{1}\left(\sum_{i\sim\emptyset} Y_t(i) < \theta(d_\emptyset)\right).$$

Now we show $\forall t \geq 0 :\ X_t(\emptyset) = Z_t(\emptyset)$ which is clear if $\chi_\emptyset = 1$. Suppose now that $\chi_\emptyset = 0$, hence $X_0(\emptyset) = Z_0(\emptyset) = 0$. By definition of $Y(i)$ we have $\forall i \sim \emptyset :\ Y_t(i) \leq X_t(i)$ and then $Z_{t+1}(\emptyset) \leq X_{t+1}(\emptyset)$. Hence we need to show $Z_{t+1}(\emptyset) \geq X_{t+1}(\emptyset)$. Suppose it is false and consider the first time s that the inequality is violated, i.e. $X_s(\emptyset) = 1,\ Z_s(\emptyset) = 0$. Since s is the first time that it happens we have $X_{s-2}(\emptyset) = 0$ and then by definition of $Y(i)$ we will have $\forall i \sim \emptyset :\ Y_{s-1}(i) = X_{s-1}(i)$ which implies $X_s(\emptyset) = Z_s(\emptyset)$. □

The representation (3.7) is crucial to our analysis because, thanks to the tree structure, the random variables $(Y_t(i), i \sim \emptyset)$ are independent of each other and identically distributed. More precisely, a simple induction shows that for $i \neq \emptyset$:

$$Y_{t+1}(i) \quad = \quad 1 - (1 - \chi_i)\mathbb{1}\left(\sum_{j\to i} Y_t(j) < \theta(d_i)\right). \tag{3.8}$$

Note that (3.8) allows to compute all the $Y_t(i)$ recursively, starting with $Y_0(i) = \chi_i$. Hence a simple induction on t shows that the random variables $(Y_t(i), i \sim \emptyset)$ are independent of each other. Let D and $\widehat{D}$ be random variables with the distribution $\mathbb{P}(D = r) = p_r$, and $\mathbb{P}(\widehat{D} = r) = q_r$. In view of (3.8), it is natural to introduce the following Recursive Distributional Equation (RDE):

$$Y \stackrel{d}{=} 1 - (1 - \chi(\widehat{D}+1))\mathbb{1}\left(\sum_{l=1}^{\widehat{D}} Y_l < \theta(\widehat{D}+1)\right), \tag{3.9}$$

where $\chi(d)$ is a Bernoulli random variable with parameter $\alpha(d)$, Y and Y_l are i.i.d. copies and all random variables are independent of each others. RDE for the process Y plays a similar role as the equation $\mu = K\mu$ for the stationary distribution of a Markov chain with kernel K, see [9].

Proposition 3.10. *Let $y = \mathbb{P}[Y = 0]$, where the distribution of Y solves the RDE (3.9). We have $f_{\alpha,\theta}(y) = 0$ where $f_{\alpha,\theta}(y)$ is defined by (3.2). Furthermore the state of the root $X_\emptyset$ is a Bernoulli random variable with parameter $1 - \mathbb{E}\left[(1 - \alpha(D))\mathbb{1}(\mathrm{Bin}(D, 1-y) < \theta(D))\right]$.*

Proof. By taking expectation in (3.9), we get

$$\begin{aligned} 1-y &= 1-\sum_{d\geq 0} q_d(1-\alpha(d+1))\mathbb{P}(\mathrm{Bin}(d,1-y)<\theta(d+1)) \\ &= 1-\sum_{d\geq 1}\frac{dp_d}{\lambda}(1-\alpha(d))\mathbb{P}(\mathrm{Bin}(d-1,1-y)<\theta(d)) \\ &= 1-\frac{\mathbb{E}\left[(1-\alpha(D))D\mathbb{1}(\mathrm{Bin}(D-1,1-y)<\theta(D))\right]}{\lambda}. \end{aligned}$$

Then multiplying λy in two sides of the above equation gives $f_{\alpha,\theta}(y)=0$.
By (3.7), the state of the root follows:

$$X_\emptyset = 1-(1-\chi(d_\emptyset))\mathbb{1}\left(\sum_{i\sim\emptyset} Y_i < \theta(d_\emptyset)\right).$$

Then taking expectation gives:

$$\mathbb{E}[X_\emptyset] = 1-\mathbb{E}\left[(1-\alpha(D))\mathbb{1}(\mathrm{Bin}(D,1-y)<\theta(D))\right],$$

and the second part of the proposition follows. □

3.4 Diffusion Process in $G^*(n,(d_i)_1^n)$

In this section we present the proofs of Theorem 3.2 and Theorem 3.3.

3.4.1 The Markov chain

The aim of this section is to describe the dynamics of the diffusion process as a Markov chain, which is perfectly tailored for the asymptotic study. We consider diffusion process on $G^*(n,(d_i)_1^n)$ where the sequence $(d_i)_1^n$, satisfies Condition 1.17. Let $m^{(n)} := \sum_{i=1}^n d_i$ denote the number of half-edges in the configuration model.

We introduce the sets $S_1,...,S_n, |S_i| = d_i$, representing the vertices $1,\dots,n$, respectively. Let M_n be a uniform random matching on $S=\cup_i S_i$ which gives us $G^*(n,(d_i)_1^n)$. Let $\mathbf{A}(0)$ and $\mathbf{I}(0)$ be the initial sets of active and inactive vertices, respectively. In particular we have

$$V = \mathbf{A}(0) \biguplus \mathbf{I}(0).$$

Let $S_i(0) := S_i$ denote the initial set of half-edges hosted by the vertex i. We call the half-edges of a subset $S_i(t)$ active (resp. inactive) if $i \in \mathbf{A}(t)$ (resp. $i \in \mathbf{I}(t)$). We define the following process: in step 0, we pick a pair (a,b), with $a \in S_i$ and $b \in S_j$ such that $i \in \mathbf{A}(0)$, and then delete both a and b from S_i and S_j respectively. Recursively, after t steps, we have the set of (currently) active vertices at step t, $\mathbf{A}(t)$, and the set of (currently) inactive vertices at step t, $\mathbf{I}(t)$. We also denote by $S_i(t)$ the state of set S_i at step t. At step $t+1$, we do the following

- We pick an active half-edge $a \in S_i(t)$ for $i \in \mathbf{A}(t)$;
- We identify its partner $b : (a,b) \in M_n$;
- And we delete both a and b from the sets $S_i(t)$ and $S_j(t)$;
- If j is currently inactive, and b is the $\theta(d_j)$-th half-edge deleted from the initial set S_j, then j becomes active from this moment on.

The system is described in terms of

- $A^{(n)}(t)$: the number of half-edges belonging to active vertices at time t;
- $I_{d,j}^{(n)}(t)$, $0 \le j < \theta(d)$, the number of inactive nodes with degree d, and j deleted half-edges, i.e., j active neighbors at time t;
- $I^{(n)}(t)$ the number of inactive nodes at time t.

It is easy to see that the following identities hold:

$$\begin{aligned} A^{(n)}(t) &= \sum_{i \in \mathbf{A}(t)} |S_i(t)|, \\ I_{d,j}^{(n)}(t) &= |\left\{ i \in \mathbf{I}(t) : d_i = d,\ |S_i(t)| = d - j \right\}|, \quad 0 \le j < \theta(d), \\ I^{(n)}(t) &= \sum_{d} \sum_{j=0}^{\theta(d)-1} I_{d,j}^{(n)}(t). \end{aligned} \tag{3.10}$$

Because at each step we delete two half-edges and the number of half-edges at time 0 is $m^{(n)}$, the number of existing half-edges at time t will be $m^{(n)} - 2t$ and we have

$$A^{(n)}(t) = m^{(n)} - 2t - \sum_{d} \sum_{j < \theta(d)} (d-j) I_{d,j}^{(n)}(t). \tag{3.11}$$

The process will finish at the stopping time $T_f^{(n)}$ which is the first time $t \in \mathbb{N}$ where $A^{(n)}(t) = 0$. The final number of active vertices will be $|A_f^{(n)}| = n - I^{(n)}(T_f^{(n)})$. By the definition of our process

$$\left\{A^{(n)}(t), \{I_{d,j}^{(n)}(t)\}_{d,j<\theta(d)}\right\}_{t\geq 0}$$

is Markov. We write the transition probabilities of the Markov chain. There are three possibilities for B, the partner of a half-edge e of an active node A at time $t+1$.

1. B is active. The probability of this event is $\frac{A^{(n)}(t)}{m^{(n)}-2t-1}$, and we have

$$\begin{aligned} A^{(n)}(t+1) &= A^{(n)}(t) - 2, \\ I_{d,j}^{(n)}(t+1) &= I_{d,j}^{(n)}(t), \quad (0 \leq j < \theta(d)). \end{aligned}$$

2. B is inactive of degree d and the half-edge e is the $(k+1)$-th deleted half-edge, and $k+1 < \theta(d)$. The probability of this event is $\frac{(d-k)I_{d,k}^{(n)}(t)}{m^{(n)}-2t-1}$, and we have

$$\begin{aligned} &A^{(n)}(t+1) = A^{(n)}(t) - 1, \\ &I_{d,k}^{(n)}(t+1) = I_{d,k}^{(n)}(t) - 1, \\ &I_{d,k+1}^{(n)}(t+1) = I_{d,k+1}^{(n)}(t) + 1, \\ &I_{d,j}^{(n)}(t+1) = I_{d,j}^{(n)}(t), \text{ for } 0 \leq j < \theta(d),\ j \neq k, k+1. \end{aligned}$$

3. B is inactive of degree d and e is the $\theta(d)$-th deleted half-edge of B. The probability of this event is $\frac{(d-\theta(d)+1)I_{d,\theta(d)-1}^{(n)}}{m^{(n)}-2t-1}$. The next state is

$$\begin{aligned} &A^{(n)}(t+1) = A^{(n)}(t) + d - \theta(d) - 1, \\ &I_{d,j}^{(n)}(t+1) = I_{d,j}^{(n)}(t), \quad (0 \leq j < \theta(d) - 1), \\ &I_{d,\theta(d)-1}^{(n)}(t+1) = I_{d,\theta(d)-1}^{(n)}(t) - 1. \end{aligned}$$

Let F_t denote the pairing generated by time t, i.e., $F_t = \{e_1, e_2\}$ is the set of half-edges picked at time t. We obtain the following equations for the expectations of $A^{(n)}(t+1), \{I_{d,j}^{(n)}(t+1)\}_{d,j<\theta(d)}$

conditioned on $A^{(n)}(t), \{I_{d,j}^{(n)}(t)\}_{d,j<\theta(d)}$:

$$\begin{aligned}
\mathbb{E}\big[A^{(n)}(t+1) - A^{(n)}(t) \mid F_t\big] &= -1 + \frac{-A^{(n)}(t) + \sum_d (d-\theta(d)+1)(d-\theta(d)) I_{d,\theta(d)-1}^{(n)}(t)}{m^{(n)} - 2t - 1}, \\
\mathbb{E}\big[I_{d,0}^{(n)}(t+1) - I_{d,0}^{(n)}(t) \mid F_t\big] &= -\frac{d I_{d,0}^{(n)}(t)}{m^{(n)} - 2t - 1}, \\
\mathbb{E}\big[I_{d,j}^{(n)}(t+1) - I_{d,j}^{(n)}(t) \mid F_t\big] &= \frac{(d-j+1) I_{d,j-1}^{(n)}(t) - (d-j) I_{d,j}^{(n)}(t)}{m^{(n)} - 2t - 1}.
\end{aligned}$$

We will show in the next section, that the trajectory of these variables throughout the algorithm is a.a.s. (asymptotically almost surely, as $n \to \infty$) close to the solution of the deterministic differential equations suggested by these equations.

3.4.2 Proof of Theorem 3.2

The proof of Theorem 3.2 is mainly based on Theorem A.10. Indeed we will apply this theorem to show that the trajectory of $I_{d,j}^{(n)}$ throughout the algorithm is a.a.s. close to the solution of the deterministic differential equations suggested by these equations.

Let (DE) be the following system of differential equations:

$$\begin{aligned}
i'_{d,0}(\tau) &= \frac{-d i_{d,0}(\tau)}{\lambda - 2\tau}, \\
i'_{d,j}(\tau) &= \frac{(d-j+1) i_{d,j-1}(\tau) - (d-j) i_{d,j}(\tau)}{\lambda - 2\tau} \quad (0 < j < \theta(d)),
\end{aligned}$$

with $\tau \in [0, \lambda/2)$, and initial conditions

$$i_{d,0}(0) = p_d(1-\alpha(d)) \qquad , \qquad i_{d,j}(0) = 0 \quad \text{for } 0 < j < \theta(d).$$

Lemma 3.11. *The solution of the system of differential equations* (DE) *is*

$$i_{d,j}(\tau) = p_d(1-\alpha(d)) \binom{d}{j} y^{d-j} (1-y)^j ,$$

where $y = (1 - 2\tau/\lambda)^{1/2}$.

Proof. Let $u = u(\tau) = -\frac{1}{2}\ln(\lambda - 2\tau)$. Then $u(0) = -\frac{1}{2}\ln(\lambda)$, u is strictly monotone and so is the inverse function $\tau = \tau(u)$. Let $f_{d,j}(u) = i_{d,j}(\tau(u))$. We write the system of differential

equations above with respect to u:

$$\begin{aligned} f'_{d,0}(u) &= -df_{d,0}(u), \\ f'_{d,j}(u) &= (d-j+1)f_{d,j-1}(u) - (d-j)f_{d,j}(u). \end{aligned}$$

We now show by induction that the solution is

$$f_{d,j}(u) = e^{-(d-j)(u-u(0))} \sum_{r=0}^{j} \binom{d-r}{j-r} \left(1 - e^{-(u-u(0))}\right)^{j-r} f_{d,r}(u(0)),$$

$0 \le j \le \theta(d) - 1$. This formula is obviously true for $j = 0$. Suppose it holds for some $j \ge 0$. Then using

$$\frac{d}{du}(f_{d,j}(u)e^{(d-j-1)(u-u(0))}) = e^{(d-j-1)(u-u(0))}(d-j)f_{d,j}(u),$$

and by inductive hypothesis, we obtain

$$\begin{aligned} f_{d,j+1}(u)e^{(d-j-1)(u-u(0))} &- f_{d,j+1}(u(0)) \\ = (d-j) &\sum_{r=0}^{j} f_{d,r}(u(0)) \binom{d-r}{j-r} \int_{u(0)}^{u} \left(1 - e^{-(v-u(0))}\right)^{j-r} e^{-(v-u(0))} dv \\ = &\sum_{r=0}^{j} f_{d,r}(u(0)) \frac{d-j}{j-r+1} \binom{d-r}{j-r} \left(1 - e^{-(u-u(0))}\right)^{j-r+1} \\ = &\sum_{r=0}^{j} f_{d,r}(u(0)) \binom{d-r}{j+1-r} \left(1 - e^{-(u-u(0))}\right)^{j-r+1}. \end{aligned}$$

Hence we have

$$f_{d,j}(u) = e^{-(d-j)(u-u(0))} \sum_{r=0}^{j} \binom{d-r}{j-r} \left(1 - e^{-(u-u(0))}\right)^{j-r} f_{d,r}(u(0)),$$

for $0 \le j \le \theta(d) - 1$. By going back to τ, we obtain

$$i_{d,j}(\tau) = y^{d-j} \sum_{r=0}^{j} i_{d,j}(0) \binom{d-r}{j-r} (1-y)^{j-r}, \; y = (1 - 2\tau/\lambda)^{1/2}.$$

It is then easy to finish the proof. □

By Condition 1.17, we know

$$\lambda = \sum_k k p_k \in (0, \infty).$$

Then, for all $\epsilon > 0$ there exist a constant K_ϵ, such that

$$\sum_{d \geq K_\epsilon} dp_d < \epsilon.$$

By Lemma 3.11

$$i_{d,j}(\tau) = p_d(1-\alpha(d))\binom{d}{j} y^{d-j}(1-y)^j \leq p_d.$$

Let $N^{(n)}(d)$ denote the number of vertices with degree d. Again, by Condition 1.17,

$$\sum_d kN^{(n)}(d)/n \to \lambda \in (0,\infty).$$

Therefore, for n large enough,

$$\sum_{d \geq K_\epsilon} dN^{(n)}(d)/n < \epsilon.$$

Hence we obtain

$$\begin{aligned} \sum_{d \geq K_\epsilon,\, j < \theta(d)} d \,|\, I_{d,j}^{(n)}(t)/n - i_{d,j}(t/n) \,| &\leq \sum_{d \geq K_\epsilon,\, j < \theta(d)} d \left(I_{d,j}^{(n)}(t)/n + i_{d,j}(t/n) \right) \\ &\leq \sum_{d \geq K_\epsilon,\, j < \theta(d)} d \left(N^{(n)}(d)/n + p_d \right) < 2\epsilon. \end{aligned} \tag{3.12}$$

Let us define

$$a(\tau) := \lambda - 2\tau - \sum_{d,j<\theta(d)} (d-j) i_{d,j}(\tau), \quad \text{and} \tag{3.13}$$

$$i(\tau) := \sum_{d,j<\theta(d)} i_{d,j}(\tau). \tag{3.14}$$

Then by Lemma 3.11, we have

$$\begin{aligned} a(\tau) &= \lambda - 2\tau - \sum_d \sum_{j<\theta(d)} (d-j) i_{d,j}(\tau) \\ &= \lambda y^2 - y\,\mathbb{E}\big[(1-\alpha(D))D\mathbb{1}\big(\mathrm{Bin}(D-1,1-y) < \theta(D)\big)\big] \\ &= f_{\alpha,\theta}(y), \quad \text{and} \end{aligned}$$

$$\begin{aligned} i(\tau) &= \sum_d \sum_{j<\theta(d)} p_d(1-\alpha(d))\binom{d}{j} y^{d-j}(1-y)^j \\ &= \mathbb{E}\left[(1-\alpha(D))\mathbb{1}\left(\mathrm{Bin}(D,1-y) < \theta(D)\right)\right], \end{aligned}$$

where $y = (1 - 2\tau/\lambda)^{1/2}$, and D is a random variable with distribution $\mathbb{P}(D = r) = p_r$.

For $\epsilon > 0$, we let K_ϵ be defined as above, and we let $b(\epsilon) := \sum_{d<K_\epsilon} \theta(d)$. We now define the domains $\Omega(\epsilon)$ as

$$\Omega(\epsilon) := \Big\{ \big(\tau, \{i_{d,j}\}_{d<K_\epsilon, j<\theta(d)}\big) \in \mathbb{R}^{b(\epsilon)+1} \ : \ -\epsilon < i_{d,j} < 1, -\epsilon < \tau < \frac{\lambda(1-\epsilon^2)}{2}, \\ \lambda - 2\tau - \sum_{d<K_\epsilon} \sum_{j<\theta(d)} (d-j) i_{d,j} > \epsilon \ \Big\}.$$

Let $T_\Omega^{(n)}$ be the stopping time for Ω which is the first time t when

$$\Big(t/n, \{I_{d,j}^{(n)}(t/n)\}\Big) \notin \Omega.$$

We will use Theorem A.10. The domain $\Omega(\epsilon)$ is a bounded open set which contains all initial values of variables which may happen with positive probability. Each variable is bounded by a constant times n. By the definition of our process, the Boundedness Hypothesis is satisfied with $\beta(n) = 1$. Trend Hypothesis is satisfied by some $\lambda_1(n) = O(1/n)$. Finally the third condition (Lipschitz Hypothesis) of the theorem is also satisfied since $\lambda - 2\tau$ is bounded away from zero. Note that for $0 < j < \theta(d)$, we have $I_{d,j}(0) = 0$, and by Condition 1.17 and by definition, $I_{d,0}(0)/n \xrightarrow{p} p_d(1 - \alpha(d))$. Then we set $\lambda = O(n^{-1/4}) > \lambda_1$. The conclusion of Theorem A.10 now gives for all $d \le K_\epsilon$

$$I_{d,j}^{(n)}(t) = n i_{d,j}(t/n) + O(n^{3/4}) \tag{3.15}$$

with probability $1 - O(n^{7/4} \exp(-n^{1/4}))$ uniformly for all $t \le n\sigma$, where $\sigma = \sigma(n)$ is the supremum of those τ for which the solution of the differential equations (DE) can be extended before reaching within l^∞-distance $Cn^{-1/4}$ of the boundary of $\Omega(\epsilon)$.

Then we have by (3.15)

$$\sup_{t \le n\sigma} \sum_{d<K_\epsilon} \sum_{j<\theta(d)} d \left| I_{d,j}^{(n)}(t)/n - i_{d,j}(t/n) \right| = o_p(1). \tag{3.16}$$

Hence by (3.12) we obtain

$$\begin{aligned}\sup_{t\le n\sigma}\left|A^{(n)}(t)/n-a(t/n)\right| &= \sup_{t\le n\sigma}\left|m^{(n)}/n-\lambda-\sum_d\sum_{j<\theta(d)}(d-j)\left(I_{d,j}^{(n)}(t)/n-i_{d,j}(t/n)\right)\right| \\ &\le \left|m^{(n)}/n-\lambda\right|+\sup_{t\le n\sigma}\sum_d\sum_{j<\theta(d)} d\left|I_{d,j}^{(n)}(t)/n-i_{d,j}(t/n)\right| \\ &\le o_p(1)+\sup_{t\le n\sigma}\sum_{d\ge K_\epsilon}\sum_{j<\theta(d)} d\left|I_{d,j}^{(n)}(t)/n-i_{d,j}(t/n)\right| \\ &\le 2\epsilon+o_p(1), \end{aligned} \tag{3.17}$$

and by the same argument

$$\begin{aligned}\sup_{t\le n\sigma}\left|I^{(n)}(t)/n-i(t/n)\right| &= \sup_{t\le n\sigma}\left|\sum_d\sum_{j<\theta(d)}\left(I_{d,j}^{(n)}(t)/n-i_{d,j}(t/n)\right)\right| \\ &\le o_p(1)+\sup_{t\le n\sigma}\sum_{d\ge K_\epsilon}\sum_{j<\theta(d)} d\left|I_{d,j}^{(n)}(t)/n-i_{d,j}(t/n)\right| \\ &\le 2\epsilon+o_p(1). \end{aligned} \tag{3.18}$$

To analyze σ, we need to determine which constraint is violated when the solution reaches the boundary of $\Omega(\epsilon)$. It cannot be the first constraint, because (3.15) must give asymptotically feasible values of $I_{d,j}^{(n)}$ until the boundary is approached. It remains to determine which of the last two constraints is violated when $\tau=\sigma$.

First assume $f_{\alpha,\theta}(y)>0$ for all $y\in(0,1]$, i.e., $y^*=0$. Then we have $a(\tau)>0$ for all $\tau\in[0,\lambda)$, which is

$$\lambda-2\tau-\sum_d\sum_{j<\theta(d)}(d-j)i_{d,j}>0.$$

Now note that $\lambda-2\tau-\sum_{d<K_\epsilon}\sum_{j<\theta(d)}(d-j)i_{d,j}\le\epsilon$ implies $f_{\alpha,\theta}(y)<\epsilon$. Then by continuity of the function $f_{\alpha,\theta}$, we conclude that in this case we can choose ϵ small enough, such that for any $\epsilon_0>0$, and for n large enough, we will have w.h.p. $\sigma>\lambda-\epsilon_0$. Then by (3.17),

$$T_f^{(n)}>n(\lambda-\epsilon_0),$$

and (3.18) implies $\Phi^{(n)}(\alpha,\theta)=1-o_p(1)$.

Consider now $y^*>0$, and suppose further that y^* is not a local minimum point of $f_{\alpha,\theta}(y)$. This means $f_{\alpha,\theta}(y)<0$ for some interval (y^*-a,y^*). We infer that the last constraint is violated

at time $\hat{\tau} \sim \lambda(1-(y^*)^2)/2$. We apply Corollary A.14 with $\hat{D}$ the domain $\Omega(\epsilon)$ defined above, and the domain D replaced by $\Omega'(\epsilon)$, which is the same as $\Omega(\epsilon)$ except that the last constraint is omitted:

$$\Omega'(\epsilon) := \left\{ \left(\tau, \{i_{d,j}\}_{d<K_\epsilon, j<\theta(d)}\right) \in \mathbb{R}^{b(\epsilon)+1} \ : \ -\epsilon < i_{d,j} < 1, -\epsilon < \tau < \frac{\lambda(1-\epsilon^2)}{2} \right\}.$$

This gives us the convergence upto the point where the solution leaves $\Omega'(\epsilon)$ or when

$$\lambda - 2t/n - \sum_{d<K_\epsilon} \sum_{j<\theta(d)} (d-j) I^{(n)}_{d,j}(t)/n > \epsilon$$

is violated. Since $a(\tau)$ begins to go negative after $\hat{\tau}$, and by (3.17) it follows that

$$\lambda - 2t/n - \sum_{d<K_\epsilon} \sum_{j<\theta(d)} (d-j) I^{(n)}_{d,j}(t)/n > \epsilon$$

must be violated almost asymptotic surely. Then it is clear (by choosing ϵ small enough) that in this case for any $\epsilon' > 0$, and for n large enough, we will have w.h.p. $T^{(n)}_f/n \in (\hat{\tau}-\epsilon', \hat{\tau}+\epsilon')$, which gives $T^{(n)}_f/n \xrightarrow{p} \hat{\tau}$. We conclude by (3.18)

$$|A^{(n)}_f| = n - I^{(n)}(T^{(n)}_f) = n - n\mathbb{E}\left[(1-\alpha(D))\mathbb{1}\left(\mathrm{Bin}(D, 1-y^*) < \theta(D)\right)\right] + o_p(n),$$

which completes the proof for $G^*(n, (d_i)_1^n)$.

Now it suffices to use Corollary 1.20, to transfer the result from $G^*(n, (d_i)_1^n)$ to $G(n, (d_i)_1^n)$.

3.4.3 Proof of Theorem 3.3

For each node i, let $\mathcal{C}(i)$ denote the final set of active nodes when in the starting state of the procedure the node i is the only active node. Clearly if $j \in \mathcal{C}(i)$, then $\mathcal{C}(j) \subseteq \mathcal{C}(i)$. Let $\alpha(d) = \alpha$ for each $d \in \mathbb{N}$. We define $\gamma_\theta(D) := \frac{\mathbb{E}[D(D-1)\mathbf{1}(\theta(D)=1)]}{\mathbb{E}[D]}$.

We first prove that if $\gamma_\theta(D) > 1$, then there exists a single node which can activate a positive fraction of the population. To do this, we use Theorem 1.30 about the existence of a giant component in the percolated graph.

We consider $\pi_\theta(d) = \mathbb{1}(\theta(d) = 1)$. Let $G^*(n, (d_i)_1^n)_{\pi_\theta}$ be the random graph obtained by deleting all the nodes of $G^*(n, (d_i)_1^n)$ for which the threshold is greater than 1. Hence $G^*(n, (d_i)_1^n)_{\pi_\theta}$ is a subgraph of $G^*(n, (d_i)_1^n)$ and we have

$$v \in G^*(n, (d_i)_1^n)_{\pi_\theta} \ \text{ if and only if } \ v \in G^*(n, (d_i)_1^n) \ \&\ \theta(d_v) = 1.$$

It is clear that to prove the existence of a node v which can trigger a global cascade in $G^*(n,(d_i)_1^n)$ w.h.p., it suffices to prove that w.h.p. there is a giant component in the random percolated graph $G^*(n,(d_i)_1^n)_{\pi_\theta}$. Indeed the threshold of every node in the giant component of $G^*(n,(d_i)_1^n)_{\pi_\theta}$ is equal to one and then each node in the giant component can activate the whole component. By Theorem 1.30, there is w.h.p. a giant component in $G^*(n,(d_i)_1^n)_{\pi_\theta}$ if and only if

$$\begin{aligned}\lambda &< \sum_{d=0}^{\infty} d(d-1)\pi_\theta(d)p_d \\ &= \sum_{d=0}^{\infty} d(d-1)\mathbb{1}\left(\theta(d)=1\right)p_d \\ &= \mathbb{E}\left[D(D-1)\mathbb{1}\left(\theta(D)=1\right)\right].\end{aligned}$$

We now prove that if $\gamma_\theta(D)<1$, then for any $k=o(n)$, we have $|\mathcal{C}(1,2,...,k)|=o_p(n)$. We will actually prove that if $\gamma_\theta(D)<1$, w.h.p. we will have $\lim_{\alpha\to 0}\Phi(\alpha,\theta)=0$ which implies the claim. Define

$$\begin{aligned}f_\theta(y) &:= \lim_{\alpha\to 0} f_{\alpha,\theta}(y) \\ &= \lambda y^2 - y\,\mathbb{E}\big[D\mathbb{1}\big(\mathrm{Bin}(D-1,1-y)<\theta(D)\big)\big].\end{aligned}$$

Clearly we have $f_\theta(1)\geq 0$. We claim that if $\gamma_\theta(D)<1$, then $f_\theta(1-\epsilon)<0$ for sufficiently small $\epsilon>0$. Indeed, we have

$$\begin{aligned}f_\theta(1-\epsilon) &= \lambda(1-\epsilon)^2-(1-\epsilon)\mathbb{E}\left[D\mathbb{1}\big(\mathrm{Bin}(D-1,\epsilon)<\theta(D)\big)\right] \\ &= \lambda(1-\epsilon)^2-(1-\epsilon)\big(\lambda-\mathbb{E}\left[D\mathbb{1}\big(\mathrm{Bin}(D-1,\epsilon)\geq\theta(D)\big)\right]\big) \\ &= \lambda(1-2\epsilon)-(1-\epsilon)(\lambda-\mathbb{E}\left[D(D-1)\mathbb{1}\big(\theta(D)=1\big)\right]\epsilon)+o(\epsilon) \\ &= \big(-\lambda+\mathbb{E}\left[D(D-1)\mathbb{1}\big(\theta(D)=1\big)\right]\big)\epsilon+o(\epsilon),\end{aligned}$$

which is negative for $\gamma_\theta(D)<1$. We infer that w.h.p. $\lim_{\alpha\to 0}y^*=1$. And this in turn implies, by Theorem 3.2, that w.h.p. $\lim_{\alpha\to 0}\Phi(\alpha,\theta)=0$. This completes the proof.

Conclusion and Future Work

In this chapter, we have studied diffusion and bootstrap percolation in a random graph with a given degree sequence. Our main result is a theorem which enables to find the final proportion of

the active vertices in the asymptotic case, i.e., when $n \to \infty$. It would be interesting to obtain quantitative versions of our results, such as large deviation estimates and central limit theorems. But this seems to be more involved due to the generality of our model (see for example [109] for some related work on the particular problem of k-core).

Coexistence in bootstrap percolation. Consider the threshold θ bootstrap percolation model on random regular graphs with degree d, and initial density α. It is known (c.f., Corollary 3.6) that there exists a nontrivial critical value for α, which we call α_f , such that for $\alpha > \alpha_f$, the final bootstrapped configuration is fully occupied for almost every initial configuration, and for $\alpha < \alpha_f$ the final bootstrapped configuration has density of occupied vertices less than 1. In our future work, we establish the existence of a distinct critical value for α, α_c, such that $0 < \alpha_c < \alpha_f$, with the following properties: if $\alpha < \alpha_c$, then for almost every initial configuration, there is no giant cluster of occupied vertices (a component containing at least a positive fraction of occupied vertices) in the final bootstrapped configuration; and if $\alpha > \alpha_c$, then for almost every initial configuration, there is a giant cluster of occupied vertices. It is interesting to note that when $\alpha < \alpha_f$, we have a giant cluster of vacant sites. Therefore, in the intermediate phase between α_c and α_f giant clusters of vacant and of occupied sites coexist. We refer to [70] for the similar results in the case of infinite regular trees.

Chapter 4

Contagion in Financial Networks

Cascades of insolvency across financial institutions may be modeled as a contagion process on a network representing their mutual exposures. We derive rigorous asymptotic results for the magnitude of contagion in a large financial network and give an analytical expression for the asymptotic fraction of defaults, in terms of network characteristics. These results extend previous studies on contagion in (unweighted) random graphs to inhomogeneous directed graphs with a given degree sequence and arbitrary distribution of weights. We use our result to obtain a criterion for the resilience of a large financial network to the default of a small group of financial institutions. Our results emphasize the role played by "contagious exposures" and show that institutions that are both highly connected and over-exposed are those which contribute most to network instability in case of default. Our asymptotic results are shown to be in good agreement with simulations for networks whose sizes are realistic, showing the relevance of the large network limit for macro-prudential regulation.

Contents

4.1 Introduction

The recent financial crisis has highlighted the interconnectedness of financial institutions worldwide and led to an increased awareness of the impact of network externalities when considering financial stability [44].

The interrelations among financial institutions may be modeled in terms of a network whose characteristics turn out to be heterogeneous and complex in nature [34]. Many of the existent frameworks [150, 59, 60] to assess systemic risk are based on a paper by Eisenberg and Noe [57]. In their model, when a firm A cannot meet all its financial obligations, it defaults and the amount recovered from A's debtors is redistributed to its creditors proportionally to their outstanding credit. It is shown that a clearing payment vector always exists, giving for each node the value of its total payable debt. Elsinger et al. [59, 60] incorporate a stochastic model for the external assets, and thus a probability that any bank defaults due to external shocks, and then use the clearing mechanism affecting the whole system to find the equilibrium payment vector. Upper's analysis [150] is based on the reconstruction of the interbank liabilities matrix by maximization of entropy, and the Eisenberg and Noe [57] clearing mechanism, to find the

size of the contagion triggered by a single firm. One observation is that, the model in [57] was conceived for payment systems, where all obligations must be paid in the short term, and this situation can be approximated by instantaneous clearing. In banking networks however, financial obligations do not necessarily represent cash flows, and consequently a mechanism that instantaneously clears all obligations does not exist. Indeed, the exposures may represent debt with different maturities, market valuations of any kind of financial contracts that are due at different time horizons. More importantly, under this model, the total loss that propagates through the financial network due to the default of a firm is limited to the total debt of the initial defaulting firms, so one does not see an effect of loss amplification. Secondly, as [149, 127, 49] point out, using maximization of entropy to reconstruct balance sheets may bias results: the threshold for a shock giving place to contagion may be larger, but conditionally on cascades occurring, the effects may be more widespread.

Studies like Gai and Kapadia [77], and Nier et al. [133] analyze contagion on random graphs. The former give an asymptotic formula for the size of contagion in a model with arbitrary degree distribution and equal weights, found under the mean field approximation. The latter considers an Erdős-Rényi graph, and computes the default cluster for each bank by simulation with an exogenous recovery rate.

Another branch of contagion studies, from the seminal paper of Allen and Gale [11] to the recent work of Battiston et al. [18], is mostly concerned by the impact of connectivity on systemic risk in regular unweighted networks, where all the exposures to counterparties are equal. While the former paper concludes on the benefits of an increase in connectivity thanks to the risk sharing effect, the latter shows that systemic risk is generally not monotonically decreasing with the diversification since their model incorporates a negative trend reinforcement that is stronger with connectivity. Battiston et al. [18] analyze several channels of contagion on regular graphs. In their model the robustness, that can be seen as a capital ratio, of all nodes evolve according to some linearly coupled SDEs before any defaults occur. The distribution of the robustness can be found in closed form. The final fraction of defaults is given then by the mean field approximation.

In this context, we would like to understand how the network topology influences the contagion magnitude, and what is the amplification of the number of defaults due to domino effects. Moreover, since the exposures are deeply heterogeneous, we need to understand the role of different nodes in the default propagation not only in terms of their connectivity, but also in

terms of how their assets are distributed among their counterparties. We introduce a notion of 'contagiousness' of links, in the sense that a link is contagious if the corresponding exposure is larger than the capital of the node to which it belongs. We study the skeleton of such contagious links, and find conditions such that there is no strongly giant connected component.

Our work extends previous results known from the random graph literature, to allow for a prescribed sequence of in/out degrees and exposures. The related and well studied problems are the existence of giant components (e.g., see [47], and Section 1.4.5), k-core problem (e.g., see Section 1.4.2), and bootstrap percolation presented in the last chapter. We use the differential equation method for random graph processes, to prove rigorously our main result, Theorem 4.8, representing a convergence in probability as the number of nodes $n \to \infty$, of the cascade size to a limit known in closed form.

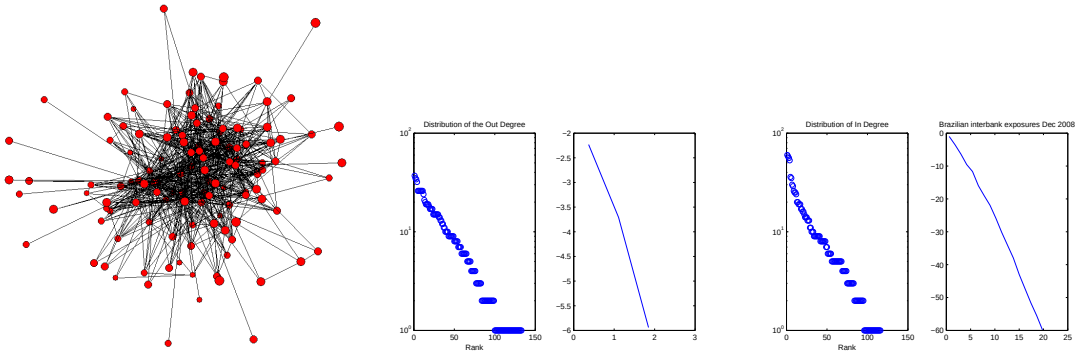

Figure 4.1: (a) The Brazilian interbank network, (b) The out-degree (number of debtors) has a Pareto tail distribution with exponent ≈ 1.7, (c) The in-degree (number of creditors) has a Pareto tail distribution with exponent ≈ 3 . Source: Cont et al. [45].

4.1.1 Summary

In this chapter, we propose a probabilistic approach to contagion modeling under incomplete information, which consists of embedding the financial network into the probability space of random networks with prescribed balance sheet data. We study contagion on a sequence of such networks of size tending to infinity and give conditions on the degree and exposure sequence such that the relevant quantities converge (i.e. the final fraction of defaults). Based on local parameters evaluated for each bank like its number of counterparties and 'contagious' links,

we find in closed form the limit in probability of the size of a default cascade generated by a random initial shock as the number of nodes $n \to \infty$. We introduce a global criterion of resilience of the financial network to small initial shocks, in which the contribution to systemic risk of every node becomes apparent. The influence of the network topology on the cascade size is now quantifiable since the expression for the asymptotic size of cascade involves the empirical distribution of degrees.

This approach allows us to obtain several new results:

- We obtain an asymptotic expression for the size of a default cascade in a large network, in terms of the characteristics of the network, extending previous results for homogeneous undirected random graphs to heterogeneous, weighted networks. These asymptotic results are shown to be in good agreement with simulations for networks with large but realistic sizes.
- We obtain an analytical criterion for the resilience of a large financial network to the default of one or several institutions, in terms of the characteristics of the network.
- The analytical nature of these results allows to analyze the influence of network characteristics, in a general setting, more explicitly than in previous studies. In particular, our results underline the role played by *contagious exposures* and show that institutions which are both highly connected and overexposed with respect to their capital may act as potential hubs for default contagion.
- Our results show the importance of taking into account the *heterogeneity* of financial networks when discussing issues of financial stability and contagion. In particular we show that, contrarily to the intuition conveyed by examples based on homogeneous networks, in presence of heterogeneity the relation between (average) connectivity of a network and its resilience to contagion is not monotonous.

4.1.2 Outline

The chapter is organized as follows. In Section 4.2 we give a model for the financial network on which we define the default dynamics, and describe the probabilistic setting we work with throughout the chapter. Our main result concerning the asymptotic of the final fraction of defaulted banks is given in Section 4.3. The examples in Section 4.4 show how the final fraction

depends on the magnitude of shocks, and point out the possible appearance of phase transitions. We compare different networks topologies, by the minimal capital ratio for which the network is resilient to contagion. Networks with the same average connectivity have different amplifications of initial shocks, and their degree of resilience to contagion is significantly changed when heterogeneity is introduced. Technical proofs are given in Section 4.5.

4.2 Models

In this section, we first introduce a model of a financial network, then describe cascade dynamics on this network, and finally the probabilistic setting we work with throughout the chapter.

4.2.1 The financial network

Interlinkages across balance sheets of financial institutions may be modeled as a network in which the vertices represent financial institutions (banks, companies, hedge funds, etc.), and links represent the interbank exposures between these institutions.

Let $\mathbf{g} = (\mathbf{v}, \mathbf{e})$ be a weighted directed graph on the vertex set $\mathbf{v} = [1, \dots, n]$. There is a weighted directed link $e(i, j)$, between the node i and j, if and only if j has a financial obligation towards i; $e(i, j)$ representing in this case the dollar amount of this exposure. Otherwise $e(i, j) = 0$.

Let L_i denote the *total interbank debt* of a node i, which is equal to the sum of all other nodes' exposures to i:

$$L_i := \sum_{j \in \mathbf{v}} e(j, i).$$

The *total interbank assets* A_i of a node i is the sum of all i's exposures

$$A_i := \sum_{j \in \mathbf{v}} e(i, j).$$

Table 4.1 shows a stylized balance sheet of a bank.

A bank faces unexpected losses, coming either from inside the financial network in the form of unpaid obligations by the debtors, or from outside the financial network in the form of external shocks emanating from the aggregate economy. The net worth of the bank, given by its **capital**

Assets	Liabilities
Interbank assets $A_i = \sum_j e(i,j)$	Interbank liabilities $L_i = \sum_j e(j,i)$ Deposits D_i
Other assets x_i	Net worth $c_i = \gamma_i A_i$

Table 4.1: Stylized balance sheet of a bank.

c_i, represents its capacity for absorbing losses before it becomes insolvent. We define the ratio γ_i as

$$\gamma_i := \frac{c_i}{A_i}.$$

We will refer to γ_i as "capital ratio" although technically it is the ratio of capital to interbank assets and not total assets. An institution is *insolvent* if its net worth is negative or zero, in which case we set $\gamma_i = 0$.

Definition 4.1. *A financial network* $(\mathbf{e}, \gamma)$ *is defined by*

- *a matrix of exposures* $(e(i,j))_{1\leq i,j\leq n}$*, and*
- *a set of capital ratios* $(\gamma_i)_{1\leq i\leq n}$*.*

The in-degree of a bank i, denoted by $d^-(i)$, is the number of its creditors

$$d^-(i) = \#\{j \in \mathbf{v} \mid e(j,i) > 0\},$$

while its out-degree, denoted by $d^+(i)$, is the number of its debtors

$$d^+(i) = \#\{j \in \mathbf{v} \mid e(i,j) > 0\}.$$

If a node has out-degree j and in-degree k, we will say that it has degree (j,k).

In a financial network $(\mathbf{e}, \gamma)$, the set of initially insolvent institutions is represented by

$$\mathbb{D}_0(e,\gamma) = \{i \in \mathbf{v} \mid \gamma_i = 0\}.$$

Empirical studies on interbank exposures [34, 45] show these networks to have complex, heterogeneous structure: Figure 4.1 displays the Brazilian interbank network, studied in [45]. In particular, one observes a heavy-tailed cross sectional distribution of degrees and exposures.

4.2.2 Default cascades

We describe now the default cascade on the financial network $(\mathbf{e}, \gamma)$. A node defaults when its capital is wiped out, i.e., $\gamma_i = 0$. We first suppose that there exists a single bank i such that its capital is wiped out by an initial shock. Then its creditors incur a loss proportional to their outstanding exposures

$$l_{j,i} := (1 - R_i)e(j, i), \text{ for all } j \in \mathbf{v},$$

where we denoted by R_i the recovery rate for the interbank liability of bank i. The loss $l_{j,i}$ will in turn erode j's capital to a level

$$(\gamma_j A_j - l_{j,i})_+,$$

where for $a \in \mathbb{R}$; $(a)_+ := \max(a, 0)$.

If the loss is greater than j's capital, then j defaults. Bank j's default may generate further rounds of defaults, with losses accumulating. The recovery rate is a key element, depending on how much of i's assets, A_i, are recovered from i's own debtors. In the most optimistic scenario all assets recovered by a defaulted bank are redistributed to creditors as for example in the paper by Eisenberg and Noe [57]. In practice however and especially in the short term, the recovery rates are very low. In the remainder of the chapter we will use an exogenous parameter, R representing the recovery rate of a defaulted bank, constant over all banks, i.e., $R_i = R$, for all $i \in \mathbf{v}$.

The previous dynamics can be extended to the case where before the cascade several banks default. We denote the initial defaults, i.e., $\{i \in \mathbf{v} \mid \gamma_i = 0\}$, *fundamental defaults*. The following rounds of default result from nonpayment of liabilities by the defaulted banks. We denote such subsequent defaults *contagious defaults* as in [59].

One can write the deterministic process of contagion in a technical manner, introducing the sequence

$$\mathbb{D}^{\mathbf{e},\gamma} = (\mathbb{D}_r^{\mathbf{e},\gamma})_{r \geq 0},$$

with $\mathbb{D}_r^{\mathbf{e},\gamma}$ the set of defaults up to round r.

Definition 4.2 (Default contagion in the financial network $(\mathbf{e}, \gamma)$)**.**

(i) Set $\mathbb{D}_0^{\mathbf{e},\gamma} := \{i \in \mathbf{v} \mid \ \gamma_i = 0\}$.

(ii) At round $r \geq 1$, set

$$\mathbb{D}_r^{\mathbf{e},\gamma} := \{i \in \mathbf{v} \mid \ \gamma_i A_i < (1-R) \sum_{j \in \mathbb{D}_{r-1}^{\mathbf{e},\gamma}} e(i,j)\}. \tag{4.1}$$

(iii) The cascade stops when no more defaults occur.

The sequence $\mathbb{D}^{\mathbf{e},\gamma}$ is non-decreasing (in each step the new nodes are added to the set of defaults), and we have

$$\mathbb{D}_0^{\mathbf{e},\gamma} \subseteq \mathbb{D}_1^{\mathbf{e},\gamma} \subseteq \cdots \subseteq \mathbb{D}_{n-1}^{\mathbf{e},\gamma}.$$

When the network is of size n, it is easy to see that the process stops at most after $n-1$ time steps. We are interested in the fraction of defaults generated by the cascade.

Definition 4.3. The final fraction of defaults, denoted by $\alpha_n(\mathbf{e}, \gamma)$, in the financial network $(\mathbf{e}, \gamma)$ is given by

$$\alpha_n(\mathbf{e}, \gamma) = \frac{|\mathbb{D}_{n-1}^{\mathbf{e},\gamma}|}{n}. \tag{4.2}$$

4.2.3 A random network model

In the previous sections, we modeled the financial network by a weighted graph $\mathbf{e}$ with the vertex set $[1, \ldots, n]$, and the corresponding sequence of capital ratios $\gamma = (\gamma_i)_{i=1}^n$. The set of exposures of node i to its counterparties is given by $\{e(i,j) > 0\}$.

To study large network asymptotics, we introduce a random network model as an ensemble of which our network is a typical sample. We embed the network $\mathbf{e}$ in a sequence of networks $(\mathbf{e_n}, \gamma_\mathbf{n})_{n \geq 1}$. The sequences of in- and out- degrees in these networks, also indexed by n, are denoted $\mathbf{d_n^+} = \{d_n^+(i)\}_{i=1}^n$, and $\mathbf{d_n^-} = \{d_n^-(i)\}_{i=1}^n$, respectively. Each network $\mathbf{e_n}$ is seen as a realization of a random weighted graph $\mathbf{E_n}$.

Definition 4.4. Let $\mathcal{G}_n(\mathbf{e_n}, \mathbf{d_n^+}, \mathbf{d_n^-})$ be the set of all weighted directed graphs with degree sequence $\mathbf{d_n^+}, \mathbf{d_n^-}$ such that, for any node i, the set of exposures is given by the non-zero elements

of line i in the exposure matrix $\mathbf{e_n}$. On a probability space $(\Omega, \mathcal{A}, \mathbb{P})$, we define $\mathbf{E_n}$ as a random network uniformly distributed on $\mathcal{G}_n(\mathbf{e_n}, \mathbf{d_n^+}, \mathbf{d_n^-})$. We endow the nodes in $\mathbf{E_n}$ with the capital ratios $\gamma_\mathbf{n}$.

As given by the above definition, the financial network is modeled under incomplete information. Disclosure of counterparty identity is not required in our framework, and thus allows for valuable confidentiality. On the other hand, we require important information, as the exact composition of balance sheets: the size of all exposures and the connectivity of each node which determine the crucial characteristics of the network and its response to external shocks.

Remark 4.5. The random financial network $\mathbf{E_n}$, is a random matrix taken uniformly over all $n \times n$ matrices having the following properties:

- for every i, $1 \leq i \leq n$, line i in $\mathbf{E_n}$ is a permutation of line i in $\mathbf{e_n}$, with the constraint that $E_n(i,i) = 0$,
- on every column $1 \leq j \leq n$, the number of non zero elements in $\mathbf{E_n}$ is the same as in $\mathbf{e_n}$,

and endowed with the capital ratios $\gamma_\mathbf{n}$. So we have the nodes in $\mathbf{E_n}$ with the capital ratios $\gamma_\mathbf{n}$, such that

$$\forall i = 1 \dots n, \quad \{\mathbf{E_n}(i,j), \quad \mathbf{E_n}(i,j) > 0\} = \{\mathbf{e_n}(i,j), \quad \mathbf{e_n}(i,j) > 0\},$$
$$\#\{j \in \mathbf{v},\ \mathbf{E_n}(j,i) > 0\} = d_n^+(j), \quad \text{and} \qquad \#\{j \in \mathbf{v},\ \mathbf{E_n}(i,j) > 0\} = d_n^-(i).$$

Definition 4.4, is equivalent to the representation of the financial system by an unweighted graph chosen uniformly among all graphs with the degree sequence $(\mathbf{d_n^+}, \mathbf{d_n^-})$, in which we assign to node i's out-going edges the set of weights $W_n(i) := \{e_n(i,j) > 0\}$. A standard method for studying random graphs with prescribed degree sequence is to consider the related (random) multigraph with the same degree sequence, constructed by configuration model and denoted by $G_n^*(\mathbf{d_n^+}, \mathbf{d_n^-}, \mathbf{e_n})$, and then condition on this multigraph being simple. We refer to Section 4.5, where we present the proof of our main results.

4.3 Results

We consider a sequence of random financial networks as introduced above. Our goal is to study the behavior of $\alpha_n(\mathbf{E_n}, \gamma_\mathbf{n})$ which represents the size of the cascade generated by the default of initially insolvent institutions $\mathbb{D}_0(\mathbf{E_n}, \gamma_\mathbf{n}) = \{i \mid \gamma_n(i) = 0\}$.

4.3.1 Assumptions

We denote by m_n the total number of links in the network $\mathbf{e_n}$

$$m_n := \sum_{i=1}^{n} d_n^+(i) = \sum_{i=1}^{n} d_n^-(i),$$

and we introduce the *empirical distribution of the degrees* as

$$\mu_n(j,k) := \frac{1}{n}\#\{i : d_n^+(i) = j, d_n^-(i) = k\}.$$

From now on, we assume that the degree sequence $\{\mathbf{d_n^+}\}$ and $\{\mathbf{d_n^-}\}$ satisfy the following regularity conditions analogous to Condition 1.17.

Assumption 4.6. For each $n \in \mathbb{N}$, $\mathbf{d_n^+} = \{(d_n^+(i))_{i=1}^n\}$ and $\mathbf{d_n^-} = \{(d_n^-(i))_{i=1}^n\}$ are sequences of nonnegative integers with $\sum_{i=1}^n d_n^+(i) = \sum_{i=1}^n d_n^-(i)$, and such that, for some probability distribution $\mu(j,k)$, independent of n:

1. The degree density condition: the empirical proportion $\mu_n(j,k)$ of vertices of degree (j,k) tends to $\mu(j,k)$: $\mu_n(j,k) \to \mu(j,k)$ as $n \to \infty$;
2. The finite expectation property: $\sum_{j,k} j\mu(j,k) = \sum_{j,k} k\mu(j,k) =: \lambda \in (0,\infty)$;
3. The second moment property: $\sum_{i=1}^n (d_n^+(i))^2 + (d_n^-(i))^2 = O(n)$.

In particular, the above assumption implies that $m_n/n \to \lambda$, as $n \to \infty$ (c.f. Remark 1.18, and [84, Theorems 5.4.2]).

We denote by $\Sigma_i^{\mathbf{e}}$ the set of permutations of the counterparties of i in a network $\mathbf{e}$. We now map the sequences of continuous exposures and capital ratios into discrete sequences, representing the default threshold for each node.

For each node i and permutation $\tau \in \Sigma_i^{\mathbf{e}}$, we define

$$\Theta(i,\tau,\mathbf{e}) := \min\left\{ k \geq 0 \mid \gamma_i \sum_{j=1}^{d^+(i)} e(i,j) < \sum_{j=1}^{k} (1-R)e(i,\tau(j)) \right\}, \tag{4.3}$$

which represents the threshold function; conditional on the order τ in which i's counterparties may default, this function determines how many defaults i's capital buffer can withstand before i defaults.

Let us define

$$p_n(j,k,\theta) := \frac{\#\{(i,\tau) \mid 1 \leq i \leq n,\ \tau \in \Sigma_i^{\mathbf{e_n}},\ d_n^+(i) = j,\ d_n^-(i) = k,\ \Theta(i,\tau,\mathbf{e_n}) = \theta\}}{n\mu_n(j,k)j!}.$$

For example, when we take $\theta = 1$, then $n\mu_n(j,k)j!p_n(j,k,1)$ is simply the total number of 'contagious links' that enter a node with degree (j,k). We say that a link is *contagious* if it represents an exposure of a node larger than its capital.

From now on, we assume that the limit of $p_n(j,k,\theta)$ exists when $n \to \infty$. Under this assumption and Assumption 4.6, we show in Section 4.5, that the fraction of nodes with degree (j,k), which default after having θ defaulted out-going neighbors, in the random financial network $\mathbf{E}_n$, converges in probability to this same limit.

Assumption 4.7. There exists a function $p : \mathbb{N}^3 \to [0,1]$, such that for all $j,k,\theta \in \mathbb{N}$ $(\theta \leq j)$

$$p_n(j,k,\theta) \longrightarrow p(j,k,\theta), \quad \text{as } n \to \infty.$$

Under this assumption, we will see in Section 4.5 that $p(j,k,\theta)$ is also the limit, in probability, of the fraction of nodes with degree (j,k) which become insolvent after θ of their counterparties default:

- $p(j,k,0)$ represents the proportion of initially insolvent nodes with degree (j,k);
- $p(j,k,1)$ represents the proportion of nodes with degree (j,k) which are 'vulnerable' i.e. may become insolvent due to the default of a single counterparty.

For example assume that for all n, the exposures of all nodes are independent, and identically distributed for all nodes with the same degree (j,k), with a law depending on j and k, but not on n, denoted by $F_X(j,k)$. Assume also that the same holds for the sequence of capital ratios, i.e., independent variables with a law depending on j and k, but not on n, denoted by $F_\gamma(j,k)$. Then it is easy to see that, by the law of large numbers, Assumption 4.7 holds, and the limit $p(j,k,\theta)$ is given by

$$p(j,k,\theta) = \mathbb{P}(X(\theta) > \gamma \sum_{l=1}^{j} X(l) - \sum_{l=1}^{\theta-1}(1-R)X(l) \geq 0),$$

with $(X(l))_{l=1}^{j}$ are i.i.d. random variables with distribution $F_X(j,k)$, and γ an independent random variable of law $F_\gamma(j,k)$.

4.3.2 The asymptotic size of contagion

We consider the representation of the financial network by a random graph as described in Section 4.2.3. The quantity $\alpha_n(\mathbf{E_n}, \gamma_\mathbf{n})$ represents the final fraction of defaults in the random network of size n.

Let us define

$$\beta(j, \pi, \theta) := \mathbb{P}(\text{Bin}(j, \pi) \geq \theta) = \sum_{l \geq \theta}^{j} \binom{j}{l} \pi^l (1-\pi)^{j-l}.$$

We let $p(j, k, \theta)$ be the limit in Assumption 4.7. We show in Lemma 4.17 that, this quantity also represents the asymptotic limit in probability of the fraction of nodes with out-degree j and in-degree k, that will default when θ of their debtors default. We define the function $I : [0, 1] \to [0, 1]$ as

$$I(\pi) := \sum_{j,k} \frac{\mu(j, k)k}{\lambda} \sum_{\theta=0}^{j} p(j, k, \theta)\beta(j, \pi, \theta). \tag{4.4}$$

$I(\pi)$ has the following interpretation: if the out-going neighbor of a randomly chosen node defaults with probability π, then $I(\pi)$ is the expected fraction of defaulted out-going neighbors after one iteration of the cascade.

Let π^* be the smallest fixed point of I in $[0, 1]$, i.e.,

$$\pi^* = \inf\{\pi \in [0, 1] \mid I(\pi) = \pi\}.$$

Note that I admits at least one fixed point. Indeed, I is a continuous increasing function, and we have

$$I(1) = \sum_{j,k} \frac{\mu(j, k)k}{\lambda} \sum_{\theta=0}^{j} p(j, k, \theta) \leq 1,$$

since $\sum_{\theta=0}^{j} p(j, k, \theta) \leq 1$. Moreover

$$I(0) = \sum_{j,k} \frac{\mu(j, k)k}{\lambda} p(j, k, 0) \geq 0.$$

The main result of this chapter is the following.

Theorem 4.8. *Consider a sequence of exposure matrices and capital ratios*

$$\{(\mathbf{e_n})_{\mathbf{n}\geq\mathbf{1}}, (\gamma_\mathbf{n})_{\mathbf{n}\geq\mathbf{1}}\},$$

satisfying Assumptions 4.6 and 4.7, and the corresponding sequence of random matrices $\{(\mathbf{E_n})_{n\geq 1}\}$.

Let π^* *be the smallest fixed point of* I *in* $[0,1]$, *i.e.,*

$$\pi^* = \inf\{\pi \in [0,1] \mid I(\pi) = \pi\}.$$

We have

1. *If* $\pi^* = 1$, *i.e. if* $I(\pi) > \pi$ *for all* $\pi \in [0,1)$, *then asymptotically all nodes default during the cascades*

$$\alpha_n(\mathbf{E_n}, \gamma_\mathbf{n}) \xrightarrow{p} 1.$$

2. *If* $\pi^* < 1$, *and furthermore* π^* *is a stable fixed point of* I $(I'(\pi^*) < 1)$, *then the asymptotic fraction of defaults satisfy*

$$\alpha_n(\mathbf{E_n}, \gamma_\mathbf{n}) \xrightarrow{p} \sum_{j,k} \mu(j,k) \sum_{\theta=0}^{j} p(j,k,\theta)\beta(j,\pi^*,\theta).$$

A proof of this theorem is given in Section 4.5.3. We give now a heuristic argument based on the branching process approximation.

Remark 4.9 (Out-going branching process approximation)**.** The value π^* is the probability that an edge taken at random ends in a defaulted node. The intuition behind the function I becomes apparent by an heuristic argument. We can approximate the local structure of a randomly chosen vertex by the following branching process (see [10, 20], and Section 3.3): the root ϕ with probability $\mu(j_r, k_r)$ has degree (j_r, k_r). In this branching process the children of a node are its out-going neighbors. Each child has degree (j,k) with probability $\frac{\mu(j,k)k}{\lambda}$, and with probability equal to $p(j,k,\theta)$ defaults when θ of their own children default. Thanks to the infinite tree structure, the state of default of the root's children are Bernoulli and i.i.d. variables with the same distribution as the state of default of their own children (Bernoulli of parameter π). So the probability of any of the root's children to have defaulted solves the equation

$$\pi = \sum_{j,k} \frac{\mu(j,k)k}{\lambda} \sum_{\theta=0}^{j_r} p(j,k,\theta)\beta(j,\pi,\theta) =: I(\pi).$$

If the solution of this equation is π^*, then the default probability of the root is given by

$$\sum_{j_r,k_r} \mu(j_r,k_r) \sum_{\theta=0}^{j_r} p(j_r,k_r,\theta)\beta(j_r,\pi^*,\theta).$$

4.3.3 Resilience to contagion

In this subsection we introduce a notion of resilience to contagion of the financial network. Some empirical studies show that the default of a single bank can trigger the default of up to 15% of the network [150, 45]. This means that for some banks the default cluster size is a significant fraction of the whole system. We consider the case where the initial defaults represent a small number of the nodes.

Let us call the following condition *the resilience condition*:

$$\sum_{j,k} \frac{\mu(j,k)jk}{\lambda} p(j,k,1) < 1. \tag{4.5}$$

As a corollary of Theorem 4.8 we have:

Corollary 4.10. *Consider the sequence of random financial networks* $(\mathbf{E_n}, \gamma_\mathbf{n})$ *satisfying Assumption 4.6 and 4.7. If the resilience condition (4.5) is satisfied, then for any* $k = o(n)$*, we have*

$$|\mathcal{D}(1,2,...,k)| = o_p(n),$$

where for $W \subseteq V$*,* $\mathcal{D}(W)$ *denote the final defaulted nodes when we start the diffusion with initial defaulted nodes* W*.*

Proof. We let a fraction ϵ of all nodes represent fundamental defaults. More precisely for each j, k, let

$$p(j,k,0) = \epsilon.$$

Then we have

$$I(\alpha) = \sum_{j,k} \frac{\mu(j,k)k}{\lambda} \sum_{\theta=0}^{j} p(j,k,\theta)\beta(j,\alpha,\theta).$$

Using a first order expansion of $\beta(j,\alpha,\theta)$ in α at 0:

$$\beta(j,\alpha,\theta) = 1_{\{\theta=0\}} + \alpha j 1_{\{\theta=1\}} + o(\alpha).$$

Then,

$$I(\alpha) = \sum_{j,k} \frac{\mu(j,k)k}{\lambda} (\epsilon + \alpha j p(j,k,1)) + o(\alpha).$$

Let π_ϵ be the smallest fixed point of I. Given Condition 4.5, for $\alpha > 0$ and small enough,

$$\lim_{\epsilon \to 0} I(\alpha) = \alpha \sum_{j,k} \frac{\mu(j,k)jk}{\lambda} p(j,k,1) + o(\alpha) < \alpha.$$

On the other hand we have seen that $I(0) \geq 0$. Thus $\lim_{\epsilon \to 0} \pi^*_\epsilon = 0$, and so by Theorem 4.8 the final fraction is $o(n)$ with high probability. Hence the Corollary follows. □

We show that the converse also holds, that if

$$\sum_{j,k} \frac{\mu(j,k)jk}{\lambda} p(j,k,1) > 1,$$

then the network is not resilient. Indeed the proof follows by the theorem below which states a condition under which with high probability there exists a giant component, strongly interlinked by contagious links (i.e., there is a directed path of contagious links from any node to another in the component). The proof is based on Theorem 1.34 (of Chapter 1), and given in Section 4.5.4.

Theorem 4.11. *Consider the sequence of random financial networks* $(\mathbf{E_n}, \gamma_\mathbf{n})$ *satisfying Assumption 4.6 and 4.7. If*

$$\sum_{j,k} \frac{\mu(j,k)jk}{\lambda} p(j,k,1) > 1, \tag{4.6}$$

then with high probability, there exists set of nodes representing a positive fraction of the financial system, strongly interlinked such that any node belonging to this set can trigger the default of all nodes in the set.

Note that, given the topology of the network, Condition 4.5 sets limits on the asymptotic fraction of contagious links $p_n(j,k,1)$.

For financial networks, Gai and Kapadia [77] give the following cascade condition by a heuristic argument, that we rewrite in our notation

$$1 - \sum_{j,k} \frac{jk}{\lambda} \mu(j,k) v_j < 0, \tag{4.7}$$

with v_j the probability that a bank with out-degree j is vulnerable, i.e. exposed to the default of a single neighbor. This condition can be seen as a particular case of Condition 4.6, in which the assets and capital buffers are i.i.d. variables drawn from an appropriate distribution. In such a

case the convergence Assumption 4.7 is satisfied by the law of large numbers. The model of [77] is an extension to the directed graph case of the generic model of global cascades proposed in a seminal paper by Watts [151], What is actually stated in [77] is that if the cascade condition is satisfied, then the *expected* size of a cascade starting from a randomly chosen node diverges, with the expectation taken over the law of the random graph with the given degree distribution and the uniform law of the initial default. The theorems we reproduced above show stronger results of convergence in probability as the size of the graph tends to infinity. More importantly, there is no reliance on a probabilistic model for the degree sequence or the balance sheet data: real balance sheet data can be used (which must verify the mild Assumptions 4.6 and 4.7). As such, what appears in other models since the probability of a node to be vulnerable represents here the limit when $n \to \infty$ of the fraction of contagious links, a directly measurable quantity. Moreover as the capital ratios can be prescribed, it suffices to set to zero the capital ratio of certain banks in order to have prescribed defaults. All this is crucial if one wants to use the resilience measure in a stress testing framework (see Section 4.4.3).

Remark 4.12 (Too interconnected to fail?)**.** We suppose that the resilience condition 4.5 is satisfied. Let π^*_ϵ be the smallest fixed point of I in $[0,1]$, when a fraction ϵ of all nodes represent fundamental defaults, i.e., $p(j,k,0) = \epsilon$ for all j,k. We obtain then, by a first order approximation of the function I, around ϵ

$$\pi^*_\epsilon - \frac{\epsilon}{1 - \sum_{j,k} \frac{\mu(j,k)jk}{\lambda} p(j,k,1)} = o(\epsilon).$$

Also, by a first order approximation of the function

$$\pi \to \sum_{j,k} \mu(j,k) \sum_{\theta=0}^{j} p(j,k,\theta)\beta(j,\pi,\theta),$$

giving the asymptotic fraction of defaults in Theorem 4.8, we have

$$\alpha_n(\mathbf{E_n}, \gamma_\mathbf{n}) - \epsilon(1 + \frac{\sum_{j,k} j\mu(j,k)p(j,k,1)}{1 - \sum_{j,k} \frac{\mu(j,k)jk}{\lambda} p(j,k,1)}) = o_p(\epsilon). \tag{4.8}$$

Suppose now that the defaulting initial fraction involves only nodes with degree (d^+, d^-), and we denote $\pi^*_\epsilon(d^+, d^-)$, the smallest fixed point of I in $[0,1]$, in the case where $p(d^+, d^-, 0) = \epsilon$, and $p(j,k,0) = 0$ for all $(j,k) \neq (d^+, d^-)$. Then we obtain in this case

$$\alpha_n(\mathbf{E_n}, \gamma_\mathbf{n}) - \epsilon\mu(d^+, d^-)(1 + \frac{d^-}{\lambda} \frac{\sum_{j,k} \frac{\mu(j,k)jk}{\lambda} p(j,k,1)}{1 - \sum_{j,k} \frac{\mu(j,k)jk}{\lambda} p(j,k,1)}) = o_p(\epsilon).$$

This simple formula shows that there are basically two factors that determine how small initial shocks are amplified by the financial network: *the weakness* of a node represented by its in-degree d^-, and the average weakness in the network, the weakness of a node being its average number of out-going contagious links, i.e., $jp(j,k,1)$.

4.4 Numerical Results

In this section, we apply the previous results to networks whose size are realistic, and asses the effect of heterogeneity in the empirical distributions as well as the relation between resilience and connectivity.

4.4.1 Relevance of asymptotics for contagion study

The results in Section 4.3 hold in the asymptotic limit when $n \to \infty$. The real financial networks may have several thousands of nodes. Indeed, the Federal Deposit Insurance Corporation reports insuring a number of 7969 institutions as of 3/18/2010 while the European Central Bank advances a figure of 8350 monetary financial institutions in the euro area (80% credit institutions and 20% money market funds).

In the purpose of studying the relevance of asymptotic formula, we generate by Blanchard's algorithm [27] a scale free network of 10000 nodes with Pareto distributed exposures. This model can be seen as a static version of the preferential attachment model. Conditional on the sequence of out-degrees, an arbitrary out-going edge will be assigned to an end-node with probability proportional to the power α of the node's out-degree (for some real parameter α), independently from all other edges. The empirical distribution of the out-degree is assumed to converge to a power law with tail coefficient γ^+

$$\mu_n^+(j) := \#\{i \mid d_n^+(i) = j\} \overset{n\to\infty}{\to} \mu^+(j) \sim j^{\gamma^+ + 1}.$$

From the graph's construction, it is easy to see that the conditional law of the in-degree is a Poisson distribution

$$P(D^- = k \mid D^+ = j) = e^{-\lambda(j)} \frac{\lambda(j)^k}{k!},$$

with $\lambda(j) = \frac{j^\alpha \mathbb{E}[D^+]}{\mathbb{E}[(D^-)^\alpha]}$, . The main theorem in [27] states that the marginal distribution of the in-degree has a Pareto tail with exponent $\gamma^- = \frac{\gamma^+}{\alpha}$, provided $1 \leq \alpha < \gamma^+$. For $\alpha > 0$, one

obtains positive correlation between in- and out-degrees.

The distribution of this simulated network's degrees and exposures is given in Figure 4.9, and is based on the empirical analysis of the Brazilian network, June 2007 [45]. On one hand we

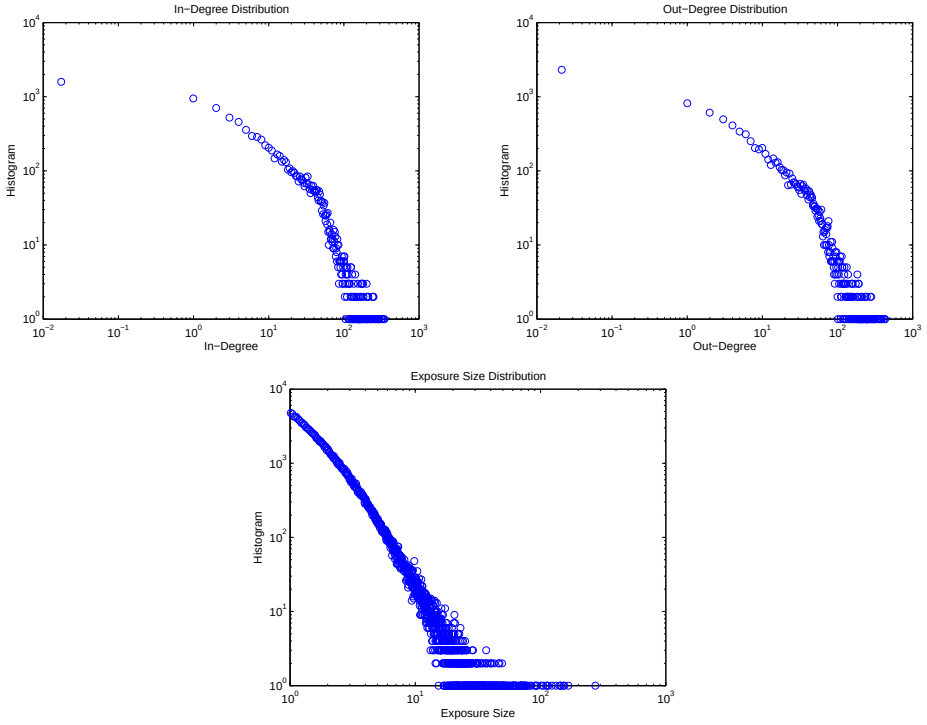

Figure 4.2: (a) The distribution of out-degree has a Pareto tail with exponent 2.19, (b) The distribution of the in-degree has a Pareto tail with exponent 1.98, (c) The distribution of the exposures (tail-exponent 2.61).

make a simulation of the default contagion starting with a random set of defaults representing 0.1% of all nodes (chosen uniformly among all nodes). On the other hand we plug the empirical distribution of the degrees, and the fraction of contagious links into Formula (4.8) for the amplification of a very small number of initial defaults. Figure 4.3 plots these values for varying values of the minimal capital ratios. We find a good similarity between the theoretical and the empirical amplification of the default number. We can clearly see that for minimal capital ratios γ_{min} less than the respective critical value γ^*_{min}, the simulated amplification explodes. Figure 4.4 plots the simulated final fraction of defaults starting from one fundamental default in a simulated scale free network as a function of the in-degree, versus the theoretical slope given

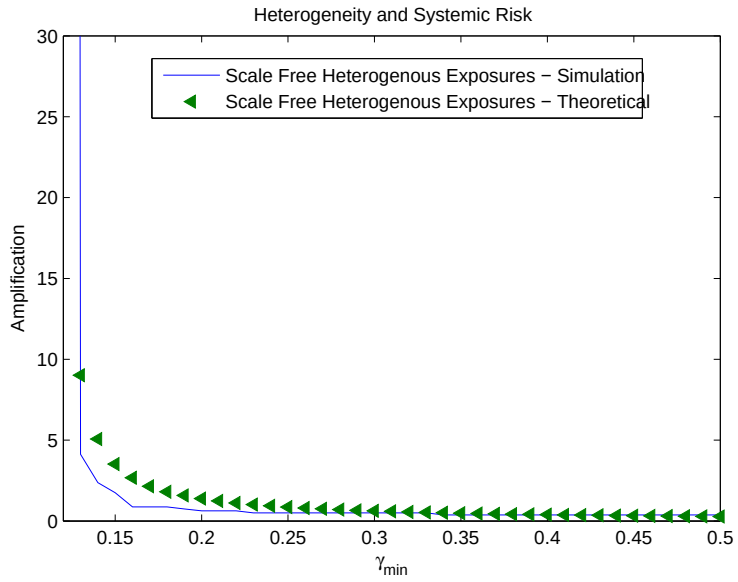

Figure 4.3: Amplification of the default number in a Scale-Free Network. The in- and out-degree of the scale-free network are Pareto distributed with tail coefficients 2.19 and 1.98 respectively, the exposures are Pareto distributed with tail coefficient 2.61, $n = 10000$.

above.

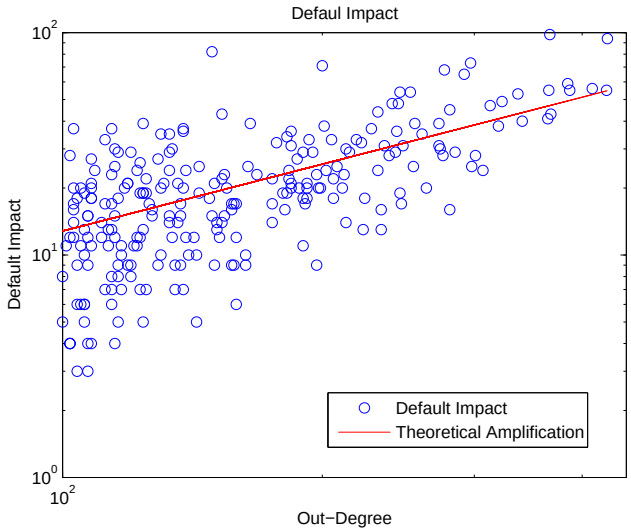

Figure 4.4: Number of defaulted nodes

4.4.2 The impact of heterogeneity

In the examples of the previous section, we can compute the minimal capital ratio γ^*_{min}, such that the network is resilient under the criterion (4.5). Two factors contribute to the sum in Condition 4.5, the connectivity of the node and its 'weakness'. We plot in Figure 4.5 three cases: a scale free network with heterogeneous weights, a scale free network with equal weights, and an homogeneous degree network (Erdős-Rényi) with equal weights, all having the same average degree.

4.4.3 Stress testing

In the Supervisory Capital Assessment Program, implemented by the Board of Governors of the Federal Reserve System in 2009, the top 19 banks in the US were asked to project their losses and resources under a macroeconomic shock scenario[1]. The program determined which of the large banks needed to augment its capital base in order to withstand the projected losses. So

[1] Available at http://www.federalreserve.gov/bankinforeg/bcreg20090424a1.pdf

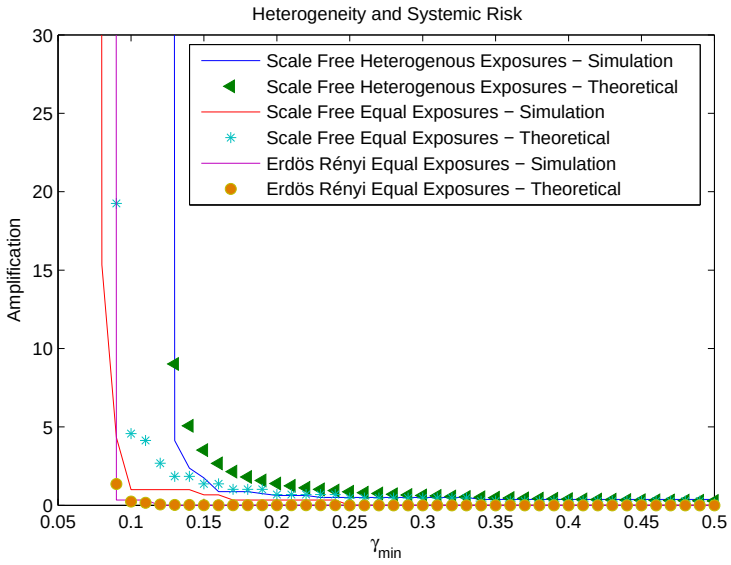

Figure 4.5: Amplification of the number of defaults in a Scale-Free Network (in and in-degree of the scale-free network are Pareto distributed with tail coefficients 2.19 and 1.98 respectively, the exposures are Pareto distributed with tail coefficient 2.61), the same network with equal weights and an Erdős-Rényi Network with equal exposures $n = 10000$.

far, the network effects such as the cascade behavior described in this chapter have not been assessed by the regulator in such stress tests.

This section presents a simple example of a stress test in which the network amplification of initial defaults becomes the major source of systemic risk. The sample network mimics the stylized properties of real financial networks with deeply heterogeneous degrees and exposures [34, 45].

Stress testing resilience to macroeconomic shocks.

Consider a banking system in which the ratio γ_i of each bank's capital to its total assets is restricted to be greater than a minimal capital ratio: $\gamma_i \geq \gamma_{\min}$. If the ratio of institution i's interbank assets to its total assets is denoted by LR_i, then

$$c_i = \gamma_i A_i \frac{1}{LR_i} > 0. \tag{4.9}$$

In a stress testing framework, we consider scenarios in which a given shock is applied to balance sheets of banks, resulting in the loss of a fraction $0 \leq S \leq 1$ of their external assets. To assess how such a stress scenario affects the resilience of the network to contagion, we evaluate the impact on the network of the default of a (small) fraction ϵ of nodes under stress scenarios of variable severity.

Using the notations in Table 4.1, the remaining capital of bank i is then given by

$$c_i(S) = (A_i + x_i \cdot (1-S) - L_i) \cdot \epsilon_i = (A_i + A_i(\frac{1}{LR_i} - 1) \cdot (1-S) - \frac{A_i}{LR_i}(1-\gamma_i))\epsilon_i,$$

where ϵ_i are independent variables with

$$\mathbb{P}(\epsilon_i = 1) = \epsilon = 1 - \mathbb{P}(\epsilon_i = 0),$$

$\epsilon_i = 1$ indicating whether i is in default in the stress scenario under consideration.

This can be re-written so as to underline the effect of the shock S on the capital

$$c_i(S) = \gamma_i A_i \frac{1}{LR_i}(1 - \frac{S}{\gamma_i}(1 - LR_i))\epsilon_i,$$

which means that a loss equal to a fraction S of the external assets translates into a loss equal to a fraction $Z_i := \frac{S}{\gamma_i}(1 - LR_i)$ of the capital buffer. Thus, in the stress scenario characterized by a macroeconomic shock (S, ϵ), the ratio of capital to interbank assets is given by

$$\gamma_i(S, \epsilon) = \gamma_i(1 - \frac{S}{\gamma_i}(1 - LR_i))\epsilon_i. \tag{4.10}$$

Starting from this expression, one can use the results of Section 4.3 to evaluate the resilience of the network and the fraction of final defaults as a function of the size of the macroeconomic shock S, without resorting to large scale simulations. In particular, given that the Condition (4.5) will depend on the shock size S, we will see that there is a *threshold* for the magnitude of S above which it destabilizes the network and makes it vulnerable to contagion. This 'phase transition' indicate that a given network has a maximal tolerance for stress; we will see in fact that this threshold may be easily computed from the characteristics of the network.

This approach is applicable to any large network, with an arbitrary distribution of exposures and degrees. To provide some analytical insight into the impact of macroeconomic shocks on the resilience to contagion, we will consider in the next two examples the case where both LR_i and γ_i are constant and equal to LR and γ_{min} respectively. Figures for the lending ratio LR have been given by [77, 150]. We will take $LR = 20\%$ and $\gamma_{min} = 10\%$.

Then the fraction of capital lost in the stress scenario is given by

$$Z = \frac{S}{\gamma_{min}}(1 - LR),$$

so we have

$$\gamma_i(Z) = \gamma_{min}(1 - Z)\epsilon_i.$$

One can observe that in this model, if $Z = 1$, a trivial global cascade ensues, in which all nodes are fundamental defaults: $\forall i,\ \gamma_i(Z) = 0$. However, as we shall see in the examples in the next sections, a sharp transition in the magnitude of the cascade occur for a threshold value of Z well below 1, which depends on the network characteristics.

An example of infinite network.

We first apply the results to an infinite random scale-free network. Such a network may be obtained as the limit when $n \to \infty$ in Blanchard's random graph model [27] (e.g., see Section 4.4.1).

The exposures of each bank with out-degree j are assumed to be independent, and follow a Pareto law. The average exposure is an increasing deterministic function of j. We denote this law F_j.

Note that in this case the limit function $p(j, k, \theta)$ does not depend on the in-degree k (we denote this simply by $p(j, \theta)$), and the function I, whose smallest zero determines the final

fraction of defaults (see Theorem 4.8), simplifies to

$$\begin{aligned} I(\pi) &= \sum_j \mu^+(j) \frac{\lambda(j)}{\lambda} \sum_{\theta=0}^{j} p(j,\theta)\beta(j,\pi,\theta) \\ &= \sum_j \mu^+(j) \frac{j^\alpha}{\mathbb{E}[(D^+)^\alpha]} \sum_{\theta=0}^{j} p(j,\theta)\beta(j,\pi,\theta) \\ &= \sum_j \hat{\mu}^+(j) \sum_{\theta=0}^{j} p(j,\theta)\beta(j,\pi,\theta), \end{aligned} \tag{4.11}$$

with $\alpha = \gamma^+/\gamma^-$, and $\hat{\mu}^+$ the size-weighted out-degree distribution given by

$$\hat{\mu}^+(j) = \mu^+(j) \frac{j^\alpha}{\mathbb{E}[(D^+)^\alpha]},$$

which is the probability that the end node of a randomly chosen edge has an out-degree equal to j. Since the out-degree distribution is a Pareto distribution, the size biased out-degree distribution is also Pareto, but with a heavier tail with exponent $\gamma^+ - \alpha$. The resilience condition 4.5 then simplifies to

$$\sum_j \hat{\mu}^+(j) j p(j,1) < 1. \tag{4.12}$$

Under the macroeconomic shock Z, the function $p(j,\theta)$ is given by

$$p(j,\theta) = \mathbb{P}(X(\theta) > \gamma(Z) \sum_{l=1}^{j} X(l) - \sum_{l=1}^{\theta-1} (1-R)X(l) \geq 0),$$

where $(X(l))_{l=1}^{j}$ are i.i.d. random variables with law F_j under $\mathbb{P}$ and $\gamma(Z)$ is given by (4.10). The function $p(j,\theta)$ is plotted in Figure 4.4.3 for a given value of the macroeconomic shock Z. The steep increase with the number of counterparty defaults θ shows how much the system is prone to contagion, especially for the institutions whose assets are concentrated across a small number of counterparties (i.e nodes with small out-degrees).

We consider that a node defaults in the first round with probability ϵ, such that $p(j,0) = \epsilon$, for all j. We plot the function I given by (4.11) for several values of the macroeconomic shock Z in Figure 4.7. We notice that the function I has three zeros for smaller values of Z, the smallest being close to zero, and as Z reaches a threshold value Z_c (in this case 42%) its only zero is close to one.

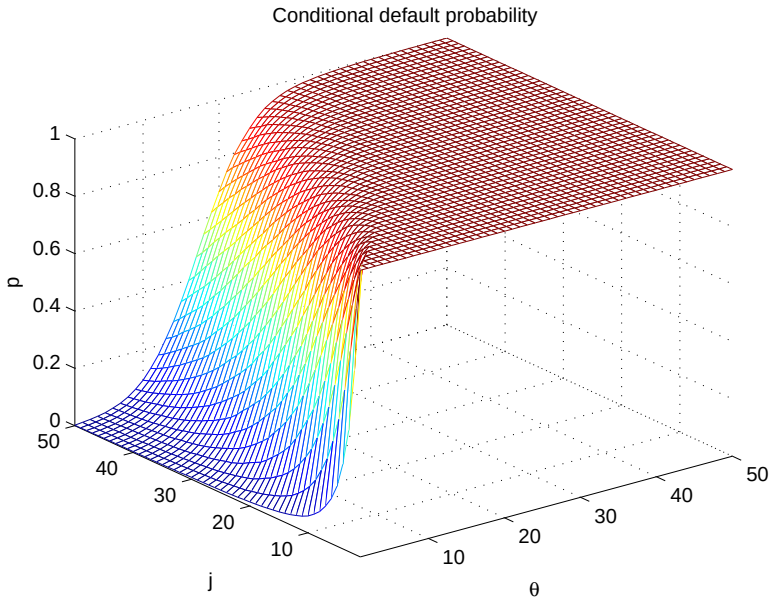

Figure 4.6: The conditional probability of default, Minimal capital ratio = 8%, Macroeconomic shock = 20%, Recovery rate = 0.

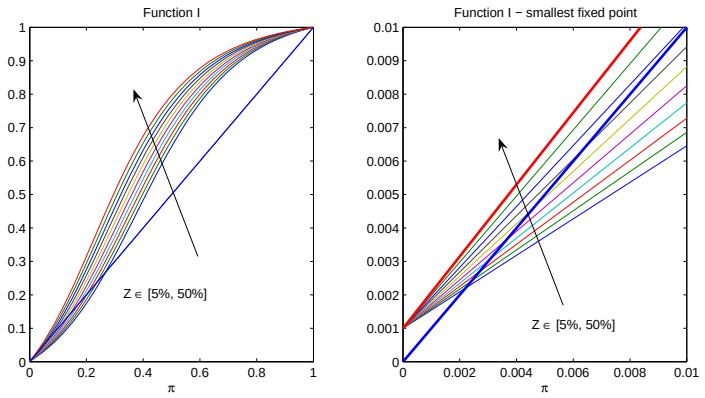

Figure 4.7: Function I for varying size of macroeconomic shock. Fraction of initial defaults = 0.1 %

As stated in Theorem 4.8, if the resilience measure is positive, then with high probability, as the initial fraction of defaults tends to 0, no global cascades appear. On the other hand, if the resilience measure is negative, the skeleton of 'contagious' links percolates, i.e. represents a positive fraction of the whole system, and we observe global cascades for any arbitrarily small fraction $\epsilon > 0$ of initial defaults chosen uniformly among all nodes. The verification of Theorem 4.8 is shown in Figure 4.8. In the non-resilient regime global cascades may occur no matter how

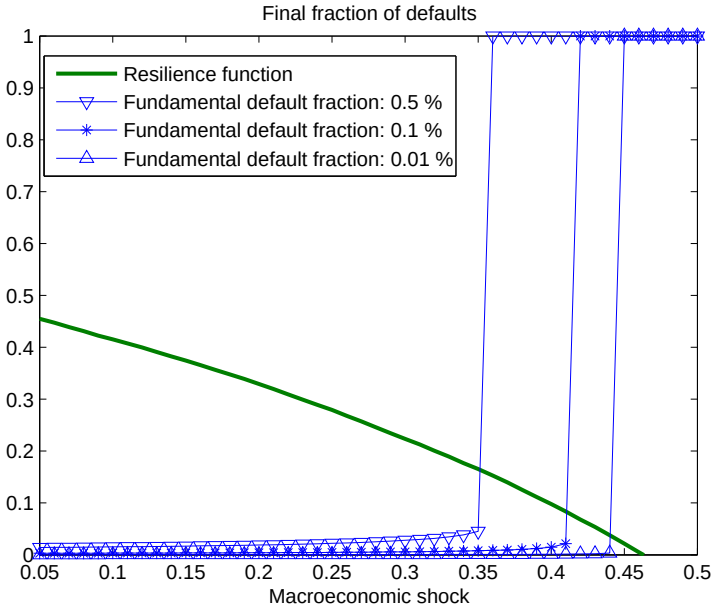

Figure 4.8: Final fraction of defaults: infinite network

small the initial fraction of defaults is. On the contrary, in the resilient regime of the infinite network, if the initial fraction of defaults is small enough, global cascades are not possible. Therefore, the condition of positivity of the resilience measure is a necessary, but not sufficient condition for non occurrence of global cascades.

A finite scale-free network.

We apply the results to a sample scale free network of 2000 nodes with heterogeneous degrees and exposures, generated from Blanchard's random graph model [27]. The empirical distribution of the sample network's degrees and exposures is shown in Figure 4.9, and its parameters were

based on the analysis of the Brazilian [45] and Austrian [34] networks.

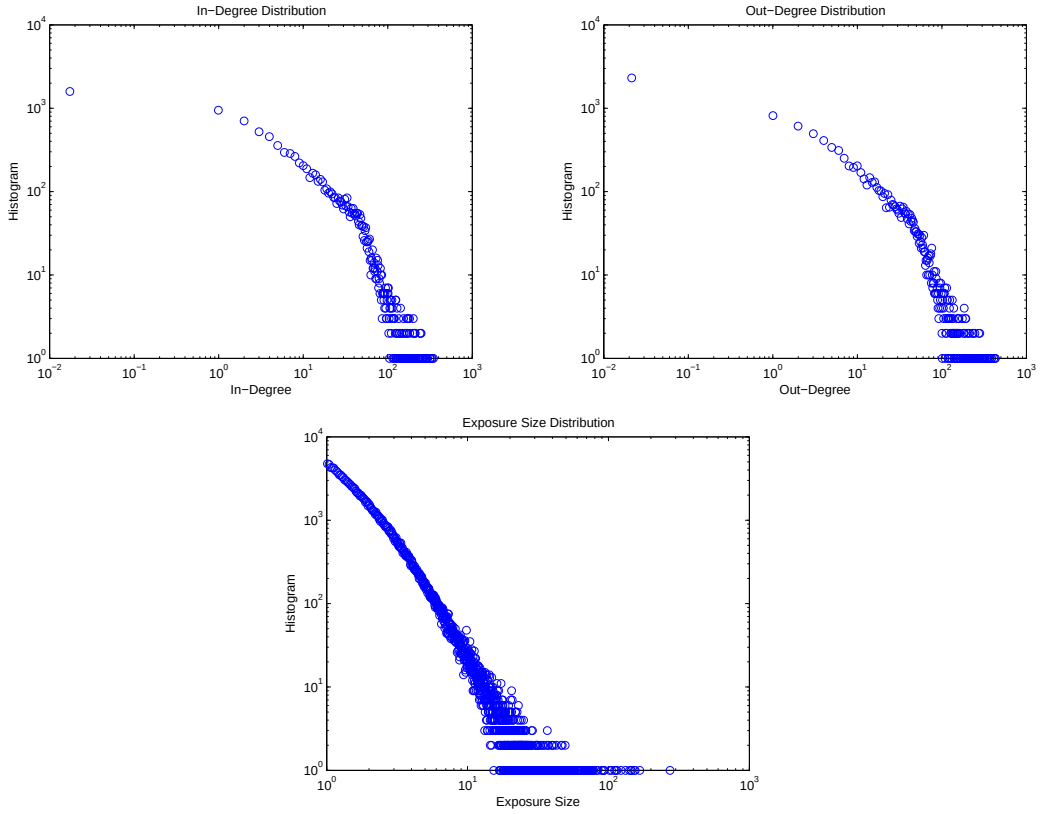

Figure 4.9: (a) The distribution of out-degree has a Pareto tail with exponent 3.5, (b) The distribution of the in-degree has a Pareto tail with exponent 2.5, (c) The distribution of the exposures has a Pareto tail with exponent 2.1.

As Figure 4.10 shows, we obtain highly correlated asset and liabilities sizes and the average exposure is increasing with the number of debtors for the more connected nodes. These properties are both observed in the empirical data.

In the finite sample, condition 4.12 translates to a condition on the average over all nodes of their number of 'contagious' links with a weight proportional to the out-degree to the power α:

$$\frac{1}{n}\sum_i w_i q_i < 1 \tag{4.13}$$

with $q_i := \#\{j \in v \mid e(i,j) > c_i\}$ and $w_i := \frac{(d^+(i))^\alpha}{\sum_l (d^+(l))^\alpha}$.

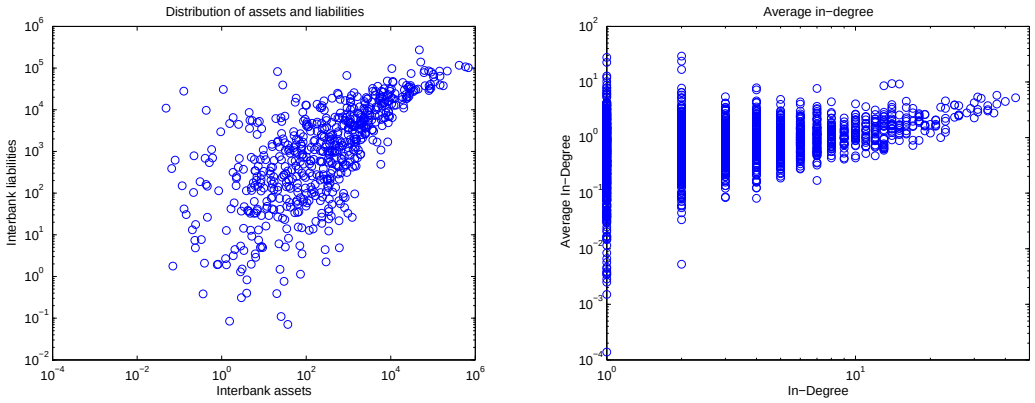

Figure 4.10: (a) Assets and liabilities, (b) Average exposures and connectivity

If α is positive, so the more correlated the in-degree and the out-degree are, the more weight is given to the most interconnected nodes. This confirms the intuition that the nodes posing the highest systemic risk are those both overexposed and interconnected, but not necessarily the largest in terms of balance sheet size.

The value $p(j, 1)$ represents the limit fraction of contagious links entering nodes with out-degree j in the limit network. Figure 4.11 shows the good accordance between the theoretical values and the values computed in the sample network. This suggests that in practice, there is

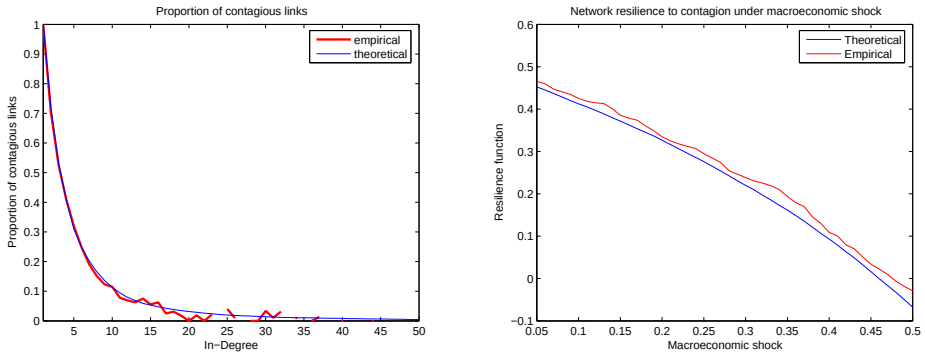

Figure 4.11: (a)Proportion of contagious links. (b)Resilience measure for varying size of macroeconomic shock in the sample and limit random network

no need to estimate the parameters of the limit distribution, but instead work directly with the empirical data.

Definition 4.13 (Empirical resilience measure)**.** *In a network* $(\mathbf{e}, \gamma)$ *of size* n*, we define the empirical resilience measure*

$$1 - \frac{1}{m_n} \sum_i d^-(i) q_i, \tag{4.14}$$

where m_n *is the total number of links in the network.*

We conduct the following simulation on the sample network: two nodes, uniformly selected among all nodes of the network initially default. Then for each value of the macroeconomic shock Z and the corresponding sizes of the capital buffers, we compute the final fraction of defaults. In light of Figure 4.8, in the infinite network, for an initial fraction of defaults representing 0.1% of the network, the positivity of the resilience measure is also sufficient for global cascades not to occur.

The results are plotted in Figure 4.12 along with the 'empirical' resilience measure. We observe that for a given network and set of initial defaults, there exists a threshold value of the macroeconomic shock, beyond which the contagion spreads to essentially the whole network. If the initial fraction of defaults is small enough, the threshold value is given by the value of Z for which the empirical resilience measure becomes zero. This suggests the existence of a first order phase transition marked by the point where the resilience measure becomes negative. We thus verify Theorem 4.11 on the emergence of the giant vulnerable component, i.e. strongly connected skeleton of contagious links, when the resilience function becomes negative.

By virtue of Theorem 4.8, the threshold value is smaller than the value of Z, for which the empirical resilience measure becomes zero. Also, the proximity of those two values depends strongly on the initial set of defaults. In the considered example, there is a good agreement between the two values, as the initial fraction of defaults is very small.

4.5 Proofs

We denote by $N_n^+(i)$, the set of bank i's *debtors*

$$N_n^+(i) = \{j \in \mathbf{v} \mid e_n(i,j) > 0\},$$

such that the out-degree of node i is $d_n^+(i) = |N_n^+(i)|$, while the set of its *creditors* is

$$N_n^-(i) = \{j \in \mathbf{v} \mid e_n(j,i) > 0\},$$

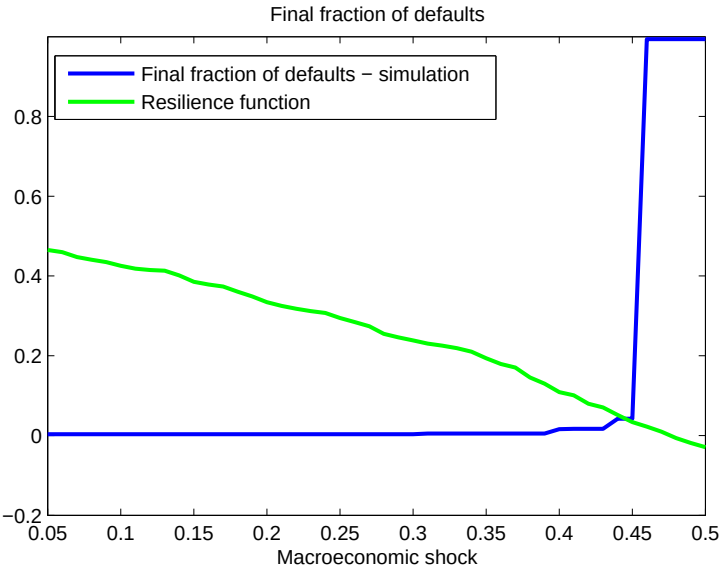

Figure 4.12: Final fraction of defaults triggered by an initial fraction of defaults representing 0.1% of the total network

and their number is the in-degree $d_n^-(i) := |N_n^-(i)|$.

The representation of the financial system by a weighted random matrix of size $n \times n$, as in Definition 4.4, is equivalent to the representation by an unweighted graph chosen uniformly among all graphs with the degree sequence $(\mathbf{d_n^+}, \mathbf{d_n^-})$, in which we assign to node i's out-going edges the weights $(e_n(i,j))_{j \in N_n^+(i)}$. We denote this random network by $G_n(\mathbf{d_n^+}, \mathbf{d_n^-}, \mathbf{e_n})$. A standard method for studying random graphs with prescribed degree distributions, as we saw in the last chapters, is to consider the related multigraph with the same degree sequence, constructed by the configuration model (denoted by $G_n^*(\mathbf{d_n^+}, \mathbf{d_n^-}, \mathbf{e_n})$), and then condition on this multigraph being simple, e.g., see Section 1.4.5.

It is quite easy to see that, conditional on the resulting multigraph being a simple graph, we obtain a uniformly distributed random digraph with the given degree sequence, i.e., $G_n(\mathbf{d_n^+}, \mathbf{d_n^-}, \mathbf{e_n})$.

In particular any property that holds w.h.p. on the random multigraph $G_n^*(\mathbf{d_n^+}, \mathbf{d_n^-}, \mathbf{e_n})$, it holds w.h.p. on the random graph $G_n(\mathbf{d_n^+}, \mathbf{d_n^-}, \mathbf{e_n})$ provided

$$\liminf_{n\to\infty} \mathbb{P}_n(G_n^*(\mathbf{d_n^+}, \mathbf{d_n^-}, \mathbf{e_n}) \text{ is simple}) > 0. \tag{4.15}$$

In recent work (c.f. see Theorem 1.19 in Chapter 1), Janson [104] has studied the probability of the random multigraph to be simple. One can translate the same proof to the directed case, and shows that the condition $\sum_{i=1}^{n}(d_n^+(i))^2 + (d_n^-(i))^2 = O(n)$ implies (4.15). Indeed in the non-directed case, Janson [104] shows that when $m^{(n)} := \sum_{i=1}^{n} d_i^{(n)} \to \infty$, one has

$$\mathbb{P}(G^*(n,(d_i)_1^n) \text{ is simple}) = \exp\left(-\frac{1}{2}\sum_i \lambda_{ii} - \sum_{i<j}(\lambda_{ij} - \log(1+\lambda_{ij}))\right) + o(1),$$

where for $1 \le i,j \le n$; $\lambda_{ij} := \frac{\sqrt{d_i(d_i-1)d_j(d_j-1)}}{m^{(n)}}$; in particular $\lambda_{ii} = \frac{d_i(d_i-1)}{m^{(n)}}$. The proof of these results is based on counting vertices with at least one loop and pairs of vertices with at least two edges between them, disregarding the number of parallel loops or edges. The same argument applies to the directed case, and one can show that when $m_n := \sum_{i=1}^{n} d_n^+(i) = \sum_{i=1}^{n} d_n^-(i) \to \infty$, then

$$\mathbb{P}(G_n^*(\mathbf{d_n^+}, \mathbf{d_n^-}, \mathbf{e_n}) \text{ is simple}) = \exp\left(-\frac{1}{2}\sum_i \lambda_{ii} - \sum_{i<j}(\lambda_{ij} - \log(1+\lambda_{ij}))\right) + o(1),$$

where for $1 \le i,j \le n$; $\lambda_{ij} = \frac{\sqrt{d_n^+(i)d_n^-(i)d_n^+(j)d_n^-(j)}}{m_n}$; in particular $\lambda_{ii} = \frac{d_n^+(i)d_n^-(i)}{m_n}$.

Then our strategy is to study contagion on $G_n^*(\mathbf{d_n^+}, \mathbf{d_n^-}, \mathbf{e_n})$ endowed with the sequence of capital ratios $\gamma_{\mathbf{n}}$, verifying conditions on the degree sequence for the assumption above (4.15) to hold, and then translate all results holding w.h.p. to the random financial system as defined in Definition 4.4.

We identify from now on the notion of a graph and that of a matching of the set of all in-coming half-edges and the set of all out-going half-edges.

Remark 4.14. One can observe that a uniform matching on a set can be obtained sequentially: choose an in-coming half-edge according to any rule (random or deterministic), and then choose the corresponding out-going half-edge uniformly over the unmatched out-going half-edges. The configuration model is particularly appropriate for the study of epidemics, as we saw in the last chapter, since the in-coming half-edges can be chosen sequentially, when epidemics spreads to their corresponding node.

4.5.1 Coupling

We are given the set of nodes $[1,\dots,n]$, and their sequence of degrees $(\mathbf{d_n^+},\mathbf{d_n^-})$. For each node i, we fix an indexing of its out-going and in-coming half-edges, ranging in $[1,\dots,d_n^+(i)]$, and $[1,\dots,d_n^-(i)]$ respectively. Furthermore, all out-going half-edges are given a global label in the range $[1,\dots,m_n]$, with m_n the total number of out-going (in-coming) half-edges. Similarly, all in-coming half-edges are given a global label in the range $[1,\dots,m_n]$.

For a set A, we denote by Σ_A the set of permutations of A. For the sequence of edge weights and capital ratios, $(\mathbf{e_n},\gamma_\mathbf{n})$, we generate the random graph $\tilde{G}_n(\mathbf{d_n^+},\mathbf{d_n^-},\mathbf{e_n},\gamma_\mathbf{n})$, by the following algorithm:

1. For each node i, choose a permutation $\tau_n(i)\in\Sigma_{N_n^+(i)}$ uniformly at random.

2. Color all in-coming and out-going half-edges in black. Define the set of initially defaulted nodes
$$\mathcal{D}_0 := \bigcup_{i,\gamma_n(i)=0}\{i\}.$$
Set for all nodes in $[1,\dots,n]\backslash\mathcal{D}_0$, $c_i=\gamma_n(i)\sum_{l\in N_n^+(i)}e_n(i,l)$.

3. At step $k\geq 1$, if the set of in-coming black half-edges belonging to nodes in $\mathcal{D}_{k-1}$ is empty, denote $\mathcal{D}_f$ the set $\mathcal{D}_{k-1}$. Otherwise:

   (a) Choose among all in-coming black half-edges of the nodes in $\mathcal{D}_{k-1}$, the in-coming half-edge with the lowest global label and color it in red.

   (b) Choose a node i with probability proportional to its number of black out-going half-edges and set $\pi_n(k)=i$. Let i have $l-1$ out-going half-edges colored in red. Choose its $\tau_i(l)$-th out-going half-edge and color it in red. If the node $i\notin\mathcal{D}_{k-1}$ and the weight $(1-R)e_n(i,\tau_i(l))$ is larger than i's remaining capital then $\mathcal{D}_k=\mathcal{D}_{k-1}\bigcup\{i\}$. Otherwise, the capital of node i becomes $c_i-(1-R)e_n(i,\tau_i(l))$.

   (c) Match node i's $\tau_i(l)$-th out-going half-edge to the in-coming half-edge selected at step (3a) to form an edge.

4. Choose a random uniform matching of the remaining out-going half-edges and match them to the remaining in-coming half-edges in increasing order and color them all in red.

Lemma 4.15. *The random graph $\tilde{G}_n(\mathbf{d}_\mathbf{n}^+, \mathbf{d}_\mathbf{n}^-, \mathbf{e_n}, \gamma_\mathbf{n})$ has the same distribution as $G_n^*(\mathbf{d}_\mathbf{n}^+, \mathbf{d}_\mathbf{n}^-, \mathbf{e_n})$. Furthermore the set $\mathcal{D}_f$ at the end of the above algorithm is the final set of defaulted nodes in the graph $\tilde{G}_n(\mathbf{d}_\mathbf{n}^+, \mathbf{d}_\mathbf{n}^-, \mathbf{e_n}, \gamma_\mathbf{n})$, endowed with the exposures $\mathbf{e_n}$ and capital ratios $\gamma_\mathbf{n}$.*

Proof. The second claim is trivial. Let us prove the first claim. We denote by σ_n^+ and σ_n^- the random permutations in $\Sigma_{[1,\dots,m_n]}$, representing the order in which the above algorithm selects the in-coming / out-going edges. At step k of the above construction, in-coming half-edge with global label $\sigma_n^-(k)$ is matched to out-going half-edge with global label $\sigma_n^+(k)$ to form an edge. The permutation σ_n^+ is determined by the set of permutations $(\tau_n(i))_{i=1,\dots,n}$ and the sequence π_n of size m_n, representing the sequence of nodes selected at Step k-(3b) of the algorithm (each node i appears in sequence π_n exactly $d_n^+(i)$ times). It is easy to see that σ_n^+ is a uniform permutation among all permutations in $\Sigma_{[1,\dots,m_n]}$, since $(\tau_n(i))_{i=1,\dots,n}$ are uniformly distributed and at each step of the algorithm we choose a node with probability proportional to its black out-going half-edges. On the other hand, the value of $\sigma_n^-(k)$ depends in a deterministic manner on

$$(\mathbf{e_n}, \gamma_\mathbf{n}, \sigma_n^+(1), \dots, \sigma_n^+(k-1)).$$

The out-going half-edge with global label j is matched with the in-coming half-edge with global label $(\sigma_n^- \circ (\sigma_n^+)^{-1})(j)$. In order to prove our claim it is enough to prove that the permutation $(\sigma_n^- \circ (\sigma_n^+)^{-1})$ is uniformly distributed among all permutations of m_n. Indeed, for an arbitrary permutation ξ belonging to the set $\Sigma_{[1,\dots,m_n]}$ we have that

$$\mathbb{P}\left(\sigma_n^+(j) = \xi^{-1}(\sigma_n^-(j)) \mid \sigma_n^+(1), \dots, \sigma_n^+(j-1)\ ,\ \sigma_n^+(k) = \xi^{-1}(\sigma_n^-(k)) \text{ for all } k < j\right) = \frac{1}{m_n - j + 1}.$$

Conditional on the knowledge of $(\sigma_n^+(1), \dots, \sigma_n^+(j-1))$, $\sigma_n^-(j)$ is deterministic. Also, by conditioning on : $\forall k < j,\ \sigma_n^+(k) = \xi^{-1}(\sigma_n^-(k))$, then

$$\xi^{-1}(\sigma_n^-(j)) \in \mathcal{T} := [1, \dots, m] \backslash \{\sigma_n^+(1), \dots, \sigma_n^+(j-1)\},$$

of cardinal $m_n - j + 1$. In the above algorithm, $\sigma_n^+(j)$ has uniform law over $\mathcal{T}$. Then the probability to chose $\xi^{-1}(\sigma_n^-(j))$ is $\frac{1}{m_n - j + 1}$.

By the law of iterated expectations, we obtain that

$$\mathbb{P}(\sigma_n^- \circ (\sigma_n^+)^{-1} = \xi) = \mathbb{P}(\sigma_n^+ = \xi^{-1} \circ \sigma_n^-) = \frac{1}{m_n!}.$$

This and the fact that the last step of the algorithm is a conditionally uniform match conclude the proof.

□

Corollary 4.16. *We can find the final set of defaulted nodes $\mathcal{D}_f$ of the above algorithm in the following manner: once the permutation $\tau_n(i)$ is chosen, assign to each node its corresponding threshold $\theta_n(i) = \Theta(i, \tau_n(i), \mathbf{e_n}, \gamma_\mathbf{n})$ as in Definition 4.3, and forget everything about $(\mathbf{e_n}, \gamma_\mathbf{n})$. Now replace Step (3b) of the algorithm by the fact that node i defaults the first time it has $\theta_n(i)$ out-going half-edges colored in red, i.e., at step*

$$\inf\{k \geq 1,\ s.t.\ \theta_n(i) = \#\{1 \leq l \leq k,\ \pi_n(l) = i\}\}.$$

We denote by $\tilde{G}_n(\mathbf{d}_\mathbf{n}^+, \mathbf{d}_\mathbf{n}^-, \theta_\mathbf{n})$ the random graph resulting from this modified algorithm. Let $N_n(j, k, \theta)$ denote the number of nodes with degree (j, k), and threshold θ after choosing uniformly the random permutations $\tau_\mathbf{n}$ in the above algorithm.

Lemma 4.17. *We have*

$$\frac{N_n(j,k,\theta)}{n} \xrightarrow{p} \mu(j,k)p(j,k,\theta),$$

as n tends to infinity.

Proof. For any node i with degree (j, k), the probability that its random threshold $\Theta(i, \tau_n(i), \mathbf{e_n}, \gamma_\mathbf{n})$, is equal to θ is

$$\nu_n(i,\theta) := \frac{\#\{\tau \in \Sigma_{N_n^+(i)} \mid \Theta(i,\tau,\mathbf{e_n},\gamma_\mathbf{n}) = \theta\}}{j!}.$$

Then we have

$$N_n(j,k,\theta) = \sum_{i,\ d_n^+(i)=j,\ d_n^-(i)=k} \mathrm{Ber}(\nu_n(i,\theta)).$$

By Assumption 4.7 we have

$$\mathbb{E}[N_n(j,k,\theta)/n] = \mu_n(j,k)p_n(j,k,\theta) \overset{n\to\infty}{\to} \mu(j,k)p(j,k,\theta),$$

and

$$\mathrm{Var}[N_n(j,k,\theta)/n] = \frac{\sum_{i,\ d_n^+(i)=j,\ d_n^-(i)=k} \nu_n(i,\theta)(1-\nu_n(i,\theta))}{n^2} \overset{n\to\infty}{\to} 0.$$

Now it is easy to conclude the proof by Chebysev's inequality.

□

4.5.2 Markov chain for contagion dynamics

In the previous section, we have replaced the description based on default rounds, given in section (4.2.2), by an equivalent one based on successive bilateral interactions. By *interaction* we mean coupling an in-coming edge with an out-going edge. At each step of this algorithm, we have one interaction only between two banks, yielding at most one default. This allows for a simpler Markov chain description, while having the same final set of defaults.

We describe in what follows the contagion process on the unweighted graph $\tilde{G}_n(\mathbf{d}_\mathbf{n}^+, \mathbf{d}_\mathbf{n}^-, \theta_\mathbf{n})$, with the sequence of random thresholds $(\theta_n(i) = \Theta(i, \tau_n(i), \mathbf{e}_\mathbf{n}, \gamma_\mathbf{n}))_{1\leq i\leq n}$ in terms of the dynamics of a Markov chain.

At any given time t banks are partitioned into two sets, *defaulted* $\mathcal{D}_n(t)$ and *solvent* $\mathcal{S}_n(t)$. We describe our Markovian system in terms of:

- $D_n^{j,k,\theta}(t)$, the number of defaulted banks at time t with degree (j,k) and default threshold θ,
- $S_n^{j,k,\theta,l}(t)$, $l < \theta \leq j$, the number of solvent banks with degree (j,k), default threshold θ and l defaulted debtors before time t.

We can introduce further variables

- $D_n(t)$: the number of defaulted banks at time t,
- $D_n^+(t)$: the number of black out-going edges belonging to defaulted banks,
- $D_n^-(t)$: the number of black in-coming edges belonging to defaulted banks,
- $S_n^+(t)$: the number of black out-going edges belonging to solvent banks at time t,

for which it is easy to see that the following identities hold:

$$\begin{aligned} S_n^+(t) &= \sum_{j,k} \sum_{0\leq l<\theta\leq j} (j-l) S_n^{j,k,\theta,l}(t), \\ D_n^-(t) &= \sum_{j,k,0\leq\theta\leq j} k D_n^{j,k,\theta}(t) - t, \\ D_n(t) &= \sum_{j,k,0\leq\theta\leq j} D_n^{j,k,\theta}(t). \end{aligned}$$

Because at each step we color in red one out-going edge and the number of black out-going edges at time 0 is m_n, the number of black out-going edges at time t will be $m_n - t$ and we have

$$D_n^+(t) + S_n^+(t) = m_n - t.$$

By definition, $\mathbf{Y_n}(t) = \left(D_n^{j,k,\theta}(t), S_n^{j,k,\theta,l}(t)\right)_{j,k,0\leq l<\theta\leq j}$ represents a Markov chain. Let $(\mathcal{F}_{n,t})_{t\geq 0}$ be its natural filtration. Let us define the operator $\vee$ as

$$x \vee y = \max(x, y).$$

We define the stopping time T_n, which represents the time the default cascade ends

$$T_n = inf\{0 \leq t \leq m_n,\ D_n^-(t) = 0\}. \tag{4.16}$$

The final number of defaulted banks will be $\mathcal{D}_f := D_n(T_n)$.

We write the transition probabilities of the Markov chain. For $t < T_n$, there are three possibilities for the partner B of an in-coming edge of a defaulted node A at time $t+1$:

1. B is in default, the next state is $\mathbf{Y_n}(t+1) = \mathbf{Y_n}(t)$.

2. B is solvent, has degree (j,k) and default threshold θ, and this is the $(l+1)$-th deleted out-going edge and $l+1 < \theta$. The probability of this event is $\frac{(j-l)S_n^{j,k,\theta,l}(t)}{m_n - t}$. The changes for the next state will be

$$S_n^{j,k,\theta,l}(t+1) = S_n^{j,k,\theta,l}(t) - 1,$$
$$S_n^{j,k,\theta,l+1}(t+1) = S_n^{j,k,\theta,l+1}(t) + 1.$$

3. B is solvent, has degree (j,k) and default threshold θ and this is the θ-th deleted out-going edge. Then with probability $\frac{(j-\theta+1)S_n^{j,k,\theta,\theta-1}(t)}{m_n - t}$ we have

$$D_n^{j,k,\theta}(t+1) = D_n^{j,k,\theta}(t) + 1,$$
$$S_n^{j,k,\theta,\theta-1}(t+1) = S_n^{j,k,\theta,\theta-1}(t) - 1.$$

Let Δ_t be the difference operator:

$$\Delta_t Y := Y(t+1) - Y(t).$$

We obtain the following equations for the expectation of $\mathbf{Y_n}(t+1)$, conditional on $\mathcal{F}_{n,t}$, by averaging over the possible transitions:

$$\begin{aligned}
\mathbb{E}\left[\Delta_t S_n^{j,k,\theta,0}|\mathcal{F}_{n,t}\right] &= -\frac{jS_n^{j,k,\theta,0}(t)}{m_n - t}, \\
\mathbb{E}\left[\Delta_t S_n^{j,k,\theta,l}|\mathcal{F}_{n,t}\right] &= \frac{(j-l+1)S_n^{j,k,\theta,l-1}(t)}{m_n - t} - \frac{(j-l)S_n^{j,k,\theta,l}(t)}{m_n - t}, \\
\mathbb{E}\left[\Delta_t D_n^{j,k,\theta}|\mathcal{F}_{n,t}\right] &= \frac{(j-\theta+1)S_n^{j,k,\theta,\theta-1}(t)}{m_n - t}.
\end{aligned} \tag{4.17}$$

The initial condition is

$$\begin{aligned}
S_n^{j,k,\theta,l}(0) &= N_n(j,k,\theta)\mathbb{1}(l=0)\mathbb{1}(\theta>0), \\
D_n^{j,k,\theta}(0) &= N_n(j,k,0)\mathbb{1}(\theta=0).
\end{aligned}$$

We will show in the next section that the trajectory of these variables for $t \leq T_n$ is close to the solution of the deterministic differential equations suggested by equations (4.17) with high probability.

4.5.3 Proof of Theorem 4.8

The proof of Theorem 3.2 is mainly based on Theorem A.10.

First we define the following set of differential equations denoted by (DE):

$$\begin{aligned}
(s^{j,k,\theta,0})'(\tau) &= -\frac{js^{j,k,\theta,0}(\tau)}{\lambda - \tau}, \\
(s^{j,k,\theta,l})'(\tau) &= \frac{(j-l+1)s^{j,k,\theta,l-1}(\tau)}{\lambda - \tau} - \frac{(j-l)s^{j,k,\theta,l}(\tau)}{\lambda - \tau}, \\
(\delta^{j,k,\theta})'(\tau) &= \frac{(j-\theta+1)s_{\theta-1}^{j,k,\theta}(\tau)}{\lambda - \tau},
\end{aligned}$$

with initial conditions

$$\begin{aligned}
s^{j,k,\theta,l}(0) &= \mu(j,k)p(j,k,\theta)\mathbb{1}(l=0)\mathbb{1}(\theta>0), \\
\delta^{j,k,\theta}(0) &= \mu(j,k)p(j,k,0)\mathbb{1}(\theta=0).
\end{aligned}$$

Lemma 4.18. *The system of differential equations (*DE*) admits the unique solution*

$$y(\tau) := \left(\delta^{j,k,\theta}(\tau), s^{j,k,\theta,l}(\tau)\right)_{j,k,0\leq l<\theta\leq j},$$

in the interval $0 \leq \tau < \lambda$, *with*

$$\begin{aligned} s^{j,k,\theta,l}(\tau) &:= \mu(j,k)p(j,k,\theta)\binom{j}{l}(1-\frac{\tau}{\lambda})^{j-l}(\frac{\tau}{\lambda})^{l}\mathbb{1}_{\{\theta>0\}}, \\ \delta^{j,k,\theta}(\tau) &:= \mu(j,k)p(j,k,\theta)\mathbb{P}\left(\mathrm{Bin}(j,\frac{\tau}{\lambda}) \geq \theta\right). \end{aligned} \tag{4.18}$$

Proof. We denote by DE^K, the set of differential equations defined above, restricted to $j \vee k < K$, and by $b(K)$ the dimension of the restricted system. Since the derivatives of the functions $\left(\delta^{j,k,\theta}(\tau), s^{j,k,\theta,l}(\tau)\right)_{j\vee k<K, 0\leq l<\theta\leq j}$ depend only on τ and the same functions, by a standard result in the theory of ordinary differential equations [99, Chapter 2, Theorem 11], there is an unique solution of DE^K in any domain of the type $(-\epsilon, \lambda) \times R$, with R a bounded subdomain of $\mathbb{R}^{b(K)}$ and $\epsilon > 0$. The solution of (DE) is defined to be the set of functions solving all the finite systems $(\mathrm{DE}^K)_{K\geq 1}$.

We solve now the system (DE). Let $u = u(\tau) = -ln(\lambda - \tau)$. Then $u(0) = -ln(\lambda)$, u is strictly monotone and so is the inverse function $\tau = \tau(u)$. We write the system of differential equations (DE) with respect to u:

$$\begin{aligned} (s^{j,k,\theta,0})'(u) &= -js^{j,k,\theta,0}(u), \\ (s^{j,k,\theta,l})'(u) &= (j-l+1)s^{j,k,\theta,l-1}(u) - (j-l)s^{j,k,\theta,l}(u), \\ (\delta^{j,k,\theta})'(u) &= (j-\theta+1)s^{j,k,\theta}_{\theta-1}(u). \end{aligned}$$

Then we have

$$\frac{d}{du}(s^{j,k,\theta,l+1}e^{(j-l-1)(u-u(0))}) = (j-l)s^{j,k,\theta,l}(u)e^{(j-l-1)(u-u(0))},$$

and by induction, we find

$$s^{j,k,\theta,l}(u) = e^{-(j-l)(u-u(0))}\sum_{r=0}^{l}\binom{j-r}{l-r}\left(1-e^{-(u-u(0))}\right)^{l-r}s^{j,k,\theta}_{r}(u(0)).$$

By going back to τ, we have

$$s^{j,k,\theta,l}(\tau) = (1-\frac{\tau}{\lambda})^{j-l}\sum_{r=0}^{l}s^{j,k,\theta}_{r}(0)\binom{j-r}{l-r}(\frac{\tau}{\lambda})^{l-r}.$$

Then, by using the initial conditions, we find

$$s^{j,k,\theta,l}(\tau) = \mu(j,k)p(j,k,\theta)\binom{j}{l}(1-\frac{\tau}{\lambda})^{j-l}(\frac{\tau}{\lambda})^{l}\mathbb{1}_{\{\theta>0\}}.$$

We have

$$\begin{aligned}(\delta^{j,k,\theta})'(\tau) &= \frac{(j-\theta+1)s_{\theta-1}^{j,k,\theta}}{\lambda(1-\frac{\tau}{\lambda})} \\ &= \frac{1}{\lambda}(j-\theta+1)\mu(j,k)p(j,k,\theta)\binom{j}{\theta-1}(1-\frac{\tau}{\lambda})^{j-\theta}(\frac{\tau}{\lambda})^{\theta-1} \\ &= \frac{j\mu(j,k)}{\lambda}p(j,k,\theta)\mathbb{P}(\mathrm{Bin}(j-1,\frac{\tau}{\lambda})=\theta-1).\end{aligned}$$

Finally, by using the fact that

$$\frac{\partial}{\partial p}\mathbb{P}(\mathrm{Bin}(N,p)>K) = N\mathbb{P}(\mathrm{Bin}(N-1,p)=K),$$

and by the initial conditions, we find that

$$\delta^{j,k,\theta}(\tau) = \mu(j,k)p(j,k,\theta)\mathbb{P}\left(\mathrm{Bin}(j,\frac{\tau}{\lambda})\geq\theta\right).$$

□

Let us define, for $0 \leq \tau \leq \lambda$;

$$\begin{aligned}\delta^{-}(\tau) &:= \sum_{j,k,\theta} k\delta^{j,k,\theta}(\tau) - \tau, \quad \text{and} \\ \delta(\tau) &:= \sum_{j,k,\theta}\delta^{j,k,\theta}(\tau),\end{aligned}$$

with $\delta^{j,k,\theta}$ given in Lemma 4.18. We have

$$\begin{aligned}\delta^{-}(\tau) &= \sum_{j,k,\theta} k\delta^{j,k,\theta}(\tau) - \tau \\ &= \sum_{j,k,\theta\leq j} k\mu(j,k)p(j,k,\theta)\mathbb{P}\left(\mathrm{Bin}(j,\frac{\tau}{\lambda})\geq\theta\right) - \tau \qquad (4.19) \\ &= \lambda(I(\frac{\tau}{\lambda}) - \frac{\tau}{\lambda}),\end{aligned}$$

and

$$\delta(\tau) := \sum_{j,k,\theta}\mu(j,k)p(j,k,\theta)\mathbb{P}\left(\mathrm{Bin}(j,\frac{\tau}{\lambda})\geq\theta\right). \qquad (4.20)$$

We now proceed to the proof of Theorem 4.8, whose aim is to approximate the value $D_n(T_n)/n$, as $n \to \infty$. We base the proof on Theorem A.10. We first need to bound the contribution of higher order terms in the infinite sums (4.19) and (4.20).

Let us fix an arbitrary constant $\epsilon > 0$. By Condition 4.6, we know

$$\lambda = \sum_{j,k} k\mu(j,k) = \sum_{j,k} j\mu(j,k) \in (0,\infty).$$

Then, there exists an integer K_ϵ, such that

$$\sum_{k\geq K_\epsilon} \sum_{j} k\mu(j,k) + \sum_{j\geq K_\epsilon} \sum_{k} j\mu(j,k) < \epsilon,$$

which implies that

$$\sum_{j\vee k\geq K_\epsilon} k\mu(j,k) < \epsilon.$$

It follows that for all $0 \leq \tau \leq \lambda$:

$$\sum_{j\vee k\geq K_\epsilon, 0\leq\theta\leq j} k\mu(j,k)p(j,k,\theta)\mathbb{P}\left(\mathrm{Bin}(j,\frac{\tau}{\lambda}) \geq \theta\right) < \epsilon. \tag{4.21}$$

The number of vertices with degree (j,k) is $n\mu_n(j,k)$. Again, by Condition 4.6,

$$\sum_{j,k} k\mu_n(j,k) = \sum_{j,k} j\mu_n(j,k) \to \lambda \in (0,\infty).$$

Therefore, for n large enough, $\sum_{j\vee k\geq K_\epsilon} k\mu_n(j,k) < \epsilon$, and for all $0 \leq t \leq m_n$:

$$\sum_{j\vee k\geq K_\epsilon, 0\leq\theta\leq j} kD_n^{j,k,\theta}(t)/n < \epsilon. \tag{4.22}$$

For $K \geq 1$, we denote

$$\begin{aligned}
\mathbf{y}^K &:= \left(\delta^{j,k,\theta}(\tau), s^{j,k,\theta,l}(\tau)\right)_{j\vee k<K,\ 0\leq l<\theta\leq j}, \quad \text{and} \\
Y_n^K &:= \left(D_n^{j,k,\theta}(\tau), S_n^{j,k,\theta,l}(\tau)\right)_{j\vee k<K,\ 0\leq l<\theta\leq j},
\end{aligned}$$

both of dimension $b(K)$, where $\delta^{j,k,\theta}(\tau)$ and $s^{j,k,\theta,l}(\tau)$ are solutions to the system (DE).

We now define the domain $\Omega(\epsilon)$ as

$$\begin{aligned}
\Omega(\epsilon) = \{\left(\tau, y^{K_\epsilon}\right) \in \mathbb{R}^{b(K_\epsilon)+1} \ : \ & -\epsilon < \delta^{j,k,\theta} < 1\ ,\ -\epsilon < s^{j,k,\theta,l} < 1, \\
& -\epsilon < \tau < \lambda - \epsilon\ ,\ \sum_{j\vee k<K_\epsilon} \sum_{\theta} k\delta^{j,k,\theta} - \tau > \epsilon\}.
\end{aligned}$$

Let $T_\Omega^{(n)}$ be the stopping time for Ω which is the first time t when

$$(t/n, Y_n) \notin \Omega.$$

We will use Theorem A.10. The domain $\Omega(\epsilon)$ is a bounded open set which contains all initial values of variables which may happen with positive probability. Each variable is bounded by a constant times n ($C_0 = 1$). By the definition of our process, the Boundedness Hypothesis is satisfied with $\beta(n) = 1$. Trend Hypothesis is satisfied by some $\lambda_1(n) = O(1/n)$. Finally the third condition (Lipschitz Hypothesis) of the theorem is also satisfied since $\lambda - \tau$ is bounded away from zero. Then by Theorem A.10, and by using Lemma 4.17 for convergence of initial values, we have

$$\mathbf{Y}_n^{K_\epsilon}(t) = n\mathbf{y}^{K_\epsilon}(t/n) + O(n^{3/4}), \tag{4.23}$$

with probability $1 - O(n^{7/4}exp(-n^{-1/4}))$ uniformly for all $t \leq n\sigma(n)$ where $\sigma = \sigma(n)$ is the supremum of those τ for which the solution of the differential equations (DE) can be extended before reaching within l^∞-distance $Cn^{-1/4}$ of the boundary of $\Omega(\epsilon)$.

Then using (4.21) and (4.22), we have

$$\begin{aligned}
\sup_{t\leq n\sigma} \left|D_n^-(t)/n - \delta^-(t/n)\right| &= \sup_{t\leq n\sigma} |\sum_{j,k}\sum_{\theta\leq j} k(D_n^{j,k,\theta}(t)/n - \delta^{j,k,\theta}(t/n))| \\
&\leq \sup_{t\leq n\sigma} \sum_{j,k}\sum_{\theta\leq j} k\left|D_n^{j,k,\theta}(t)/n - \delta^{j,k,\theta}(t/n)\right| \\
&\leq \sup_{t\leq n\sigma} \sum_{j\vee k\leq K_\epsilon}\sum_{\theta\leq j} k\left|D_n^{j,k,\theta}(t)/n - \delta^{j,k,\theta}(t/n)\right| + 2\epsilon \\
&\leq o_p(1) + 2\epsilon,
\end{aligned} \tag{4.24}$$

and similarly,

$$\begin{aligned}
\sup_{t\leq n\sigma} |D_n(t)/n - \delta(t/n)| &\leq \sup_{t\leq n\sigma} \sum_{j\vee k\leq K_\epsilon}\sum_{\theta\leq j} \left|D_n^{j,k,\theta}(t)/n - \delta^{j,k,\theta}(t/n)\right| + 2\epsilon \\
&\leq o_p(1) + 2\epsilon.
\end{aligned} \tag{4.25}$$

To analyze σ, we need to determine which constraint is violated when the solution reaches the boundary of $\Omega(\epsilon)$. It cannot be the first two constraints, because (4.23) must give asymptotically feasible values of Y_n until the boundary is approached. It remains to determine which of the last two constraints is violated when $\hat{\tau} = \sigma$. There are two cases:

First assume $I(\pi) > \pi$ for all $\pi \in [0, 1)$, i.e., $\pi^* = 1$. Then we have $\delta^-(\tau) > 0$ for all $\tau \in [0, \lambda)$, which is

$$\sum_{j,k,\theta} k\delta^{j,k,\theta}(\tau) - \tau > 0.$$

We now note that $\sum_{j \vee k < K_\epsilon} \sum_\theta k\delta^{j,k,\theta}(\tau) - \tau \leq \epsilon$, implies

$$I(\tau/\lambda) - \tau/\lambda \leq \epsilon/\lambda.$$

Then by continuity of the function I we conclude that in this case we can choose ϵ small enough, such that for any ϵ_0, and for n large enough, we will have w.h.p. $\sigma > \lambda - \epsilon_0$. Then we conclude $T_n > n(\lambda - \epsilon_0)$, which gives us

$$D_n(T_n) = n - o_p(n).$$

Now consider the case $\pi^* < 1$, and furthermore π^* is a stable fixed point of $I(\pi)$. Then by definition of π^* and by using the fact that $I(1) \leq 1$, we have $I(\pi) < \pi$ for some interval $(\pi^*, \pi^* + \tilde{\pi})$. Then $\delta^-(\tau)$ is negative in an interval $(\hat{\tau}, \hat{\tau} + \tilde{\tau})$, with $\hat{\tau} = \lambda\pi^*$.

We apply Corollary A.14 with $\hat{D}$ the domain $\Omega(\epsilon)$ defined above, and the domain D replaced by $\Omega'(\epsilon)$, which is the same as $\Omega(\epsilon)$ except that the last constraint is omitted:

$$\Omega'(\epsilon) = \{\left(\tau, y^{K_\epsilon}\right) \in \mathbb{R}^{b(K_\epsilon)+1} \ : \ -\epsilon < \delta^{j,k,\theta} < 1 \ , \ -\epsilon < s^{j,k,\theta,l} < 1, \ -\epsilon < \tau < \lambda - \epsilon\}.$$

This gives us the convergence upto the point where the solution leaves $\Omega'(\epsilon)$ or when

$$\sum_{j \vee k < K_\epsilon} \sum_\theta k\delta^{j,k,\theta}(\tau) - \tau > \epsilon$$

is violated. Since $\delta^-(\tau)$ begins to go negative after $\hat{\tau}$, and by (4.24) it follows that

$$\sum_{j \vee k < K_\epsilon} \sum_\theta k\delta^{j,k,\theta}(\tau) - \tau > \epsilon$$

must be violated almost asymptotic surely. Then it is easy to conclude (by choosing ϵ small enough) that in this case for any $\epsilon' > 0$, and for n large enough, we will have w.h.p. $T_n/n \in (\hat{\tau} - \epsilon', \hat{\tau} + \epsilon')$, which gives $T_n/n \xrightarrow{p} \hat{\tau}$. We conclude by (4.25), $D_n(T_n)/n \xrightarrow{p} \delta(\hat{\tau})$. It is now clear that (4.20) implies the desired result.

4.5.4 Proof of Theorem 4.11

The strong connectivity for spaces of sparse random directed graphs with prescribed degree sequence has been studied by Cooper and Frieze in [47], e.g., see Theorem 1.34 in Chapter 1.

Let λ_n represent the average degree (then by Assumption 4.6, $\lambda_n \to \lambda$ as $n \to \infty$), and $\mu_n(j,k)$ represent the empirical distribution of the degrees, which is furthermore assumed to be proper (satisfy Condition 1.33), then Theorem 1.34 states that if the average (size-biased) directed degree ν in the graph is greater than 1, i.e.,

$$\nu := \sum_{j,k} jk \frac{\mu(j,k)}{\lambda} > 1, \tag{4.26}$$

then the random directed graph contains w.h.p. a strongly connected giant component.

We remark that Theorem 1.34 is given under stronger assumptions on the degree sequence, adding to Assumption 4.6 the following three conditions, in which Δ_n denotes the maximum degree:

- $\rho_n := \max\left(\sum_{i,j} \frac{i^2 j \mu_n(i,j)}{\lambda_n}, \sum_{i,j} \frac{j^2 i \mu_n(i,j)}{\lambda_n}\right) = o(\Delta_n)$, if $\Delta_n \to \infty$;
- $\Delta_n \leq \frac{n^{1/12}}{\log n}$;
- As $n \to \infty$, $\nu_n := \sum_{j,k} jk \frac{\mu_n(j,k)}{\lambda_n} \to \nu \in (0, \infty)$.

A first reason for adding these conditions in [47] is to ensure that Equation (4.15) holds. However, following Janson [104], the restricted set of conditions 4.6 is sufficient. The second reason is that [47] gives a more precise results on the structure of the giant component, e.g., see Section 1.4.5. For our purpose, to find the sufficient condition for the existence of a strongly connected giant component, we show that these supplementary conditions may be dropped.

It is easy to see that a bounded degree sequence (i.e., $\Delta_n = O(1)$) which satisfies Assumption 4.6 is proper. We use this fact in the following.

Lemma 4.19. *Consider the random directed graph $G(n, \mathbf{d_n^+}, \mathbf{d_n^-})$, where the degree sequence satisfies Assumption 4.6. If*

$$\sum_{j,k} jk \frac{\mu(j,k)}{\lambda} > 1, \tag{4.27}$$

then the graph contains w.h.p. a strongly connected giant component.

Proof. First note that by the second moment property (in Assumption 4.6) and Fatou's lemma, there exists a constant C such that

$$\begin{aligned}\sum_{j,k} jk\mu(j,k) &\leq \sum_{j,k}(j^2+k^2)\mu(j,k) \\ &\leq \liminf_{n\to\infty}\sum_{j,k}(j^2+k^2)\mu_n(j,k) \leq C.\end{aligned}$$

It follows that for arbitrary $\epsilon > 0$, there exists a constant Δ_ϵ such that

$$\sum_{j\wedge k>\Delta_\epsilon} jk\mu(j,k) \leq \epsilon.$$

Thus (assuming (4.27) holds), by choosing ϵ small enough, there exists a constant Δ_ϵ such that

$$\sum_{j\wedge k\leq\Delta_\epsilon} jk\frac{\mu(j,k)}{\lambda} > 1.$$

We now modify the random graph such that the maximum degree is equal to Δ_ϵ. For every node i such that $d_n^+(i)\wedge d_n^-(i) > \Delta_\epsilon$, all its in-coming (resp. out-going) half-edges are transferred to new nodes with degree $(0,1)$ (resp. with degree $(1,0)$).

Since these newly created nodes cannot be part of any strongly connected component, it follows that, if the modified graph contains such a component, then necessarily the initial graph also does. It is then enough to evaluate Equation (4.26) for this modified graph, which by construction verifies the Assumption 4.6, for the new empirical distribution $\tilde{\mu}$ with the average degree $\tilde{\lambda}$. Also, since the degrees of the modified graph are bounded, the supplementary conditions above also hold, i.e., the degree sequence is proper, and we can apply Cooper and Frieze's result.

We are only left to show that $\sum_{j,k} jk\frac{\tilde{\mu}(j,k)}{\tilde{\lambda}} > 1$. Indeed, we have that

$$\begin{aligned}\sum_{j,k} jk\frac{\tilde{\mu}(j,k)}{\tilde{\lambda}} &= \sum_{j\wedge k\leq\Delta_\epsilon} jk\frac{\tilde{\mu}(j,k)}{\tilde{\lambda}} \\ &= \sum_{0<j,k\leq\Delta_\epsilon} jk\frac{\tilde{\mu}(j,k)}{\tilde{\lambda}} \\ &= \sum_{0<j,k\leq\Delta_\epsilon} jk\frac{\mu(j,k)}{\lambda} > 1.\end{aligned}$$

The last equality follows from the fact that for $0 < j, k \leq \Delta_\epsilon$, we have

$$\frac{\tilde{\mu}(j,k)}{\tilde{\lambda}} = \frac{\mu(j,k)}{\lambda}.$$

This is true since the total number of edges, and the number of nodes with degree j, k for $0 < j, k \leq \Delta_\epsilon$, stays unmodified. □

We now proceed to the proof of Theorem 4.11. Our proof is based on ideas applied in [72, 103] for site and bond percolation in configuration model (c.f. see Section 1.4.4).

Our aim is to show that the skeleton of contagious links in the random financial network is still described by configuration model, with a degree sequence verifying Assumption 4.6, and then apply Lemma 4.19.

For each node i, the set of contagious out-going edges is denoted by $C_n(i)$, which is

$$C_n(i) := \{1 \leq l \leq d_n^+(i) : (1-R)e_n(i,l) > \gamma_n(i)\}.$$

Let $c_n^+(i)$ be their number, i.e., $c_n^+(i) := \#C_n(i)$.

We denote by G_n^c the unweighted skeleton of contagious links in the random financial network $G_n(\mathbf{d_n^+}, \mathbf{d_n^-}, \mathbf{e_n})$, endowed with the capital ratios $\gamma_\mathbf{n}$.

In order to characterize the law of G_n^c, we adapt Janson's method (see Section 1.4.4) for the directed case.

Lemma 4.20. *The unweighted skeleton of contagious links G_n^c has the same law as the random graph constructed by the following algorithm:*

1. *Let $\tilde{n} := n + m_n - \sum_{i=1}^n c_n^+(i)$. Replace the degree sequence (d_n^+, d_n^-) of size n by the degree sequence $(\tilde{\mathbf{d}}_{\tilde{\mathbf{n}}}^+, \tilde{\mathbf{d}}_{\tilde{\mathbf{n}}}^-)$ of size $\tilde{n}$, with*

$$\forall\ 1 \leq i \leq n,\ \tilde{d}_{\tilde{n}}^+(i) = c_n^+(i),\ \tilde{d}_{\tilde{n}}^-(i) = d_n^-(i),$$
$$\forall\ n+1 \leq i \leq \tilde{n},\ \tilde{d}_{\tilde{n}}^+(i) = 1,\ \tilde{d}_{\tilde{n}}^+(i) = 0.$$

2. *Construct the random unweighted graph $G_{\tilde{n}}^*(\tilde{\mathbf{d}}_{\tilde{\mathbf{n}}}^+, \tilde{\mathbf{d}}_{\tilde{\mathbf{n}}}^-)$ with $\tilde{n}$ nodes, and the degree sequence $(\tilde{\mathbf{d}}_{\tilde{\mathbf{n}}}^+, \tilde{\mathbf{d}}_{\tilde{\mathbf{n}}}^-)$ by configuration model.*

3. *Delete $n^+ = \tilde{n} - n$ randomly chosen nodes with out-degree 1 and in-degree 0.*

Proof. The skeleton G_n^c can be obtained in a two-step "explosion" procedure, rather than directly removing all non-contagious links: first all non-contagious links in $G_n^*(\mathbf{d_n^+}, \mathbf{d_n^-}, \mathbf{e_n})$ are disconnected form their end nodes and transferred to newly created nodes of degree $(1,0)$. Then delete all new nodes and their incident edges. Looking at graphs as configurations, and since the first step changes the total number of nodes but not the number of half-edges, it is easy to see that there is a one to one correspondence between the configurations before and after the explosions. Then the graph after explosions is still described by the configuration model, and thus has the same law as $G_{\tilde{n}}^*(\tilde{\mathbf{d}}_{\tilde{\mathbf{n}}}^+, \tilde{\mathbf{d}}_{\tilde{\mathbf{n}}}^-)$. Finally, by symmetry the nodes with out-degree 1 and in-degree 0 are equivalent, so one may remove randomly the appropriate number of them. □

Note that since the degree sequence before the explosions verifies Assumption 4.6, so does the degree sequence after explosions. Moreover, since we are interested in its strongly connected component, and nodes of degrees $(1,0)$ will not be included, we can actually apply Lemma 4.19 to the random graph resulting by the above algorithm. Hence we study the strongly connected component in the intermediate graph $G_{\tilde{n}}^*(\tilde{\mathbf{d}}_{\tilde{\mathbf{n}}}^+, \tilde{\mathbf{d}}_{\tilde{\mathbf{n}}}^-)$.

Let us denote by $l_{\tilde{n}}(j,k)$, the number of nodes with out-degree j and in-degree k in the graph $G_{\tilde{n}}^*(\tilde{\mathbf{d}}_{\tilde{\mathbf{n}}}^+, \tilde{\mathbf{d}}_{\tilde{\mathbf{n}}}^-)$, and by $\tilde{\lambda}_{\tilde{n}}$, the average degree. Then the average (size-biased) directed degree in this random graph is given by $\tilde{\nu}_n := \sum_{j,k} jkl_{\tilde{n}}(j,k)/(\tilde{\lambda}_{\tilde{n}}\tilde{n})$.

We first observe that $\tilde{\lambda}_{\tilde{n}}\tilde{n} = m_n$, since the number of edges is unchanged after explosions. For every $k > 0$, the quantity $\sum_j jl_{\tilde{n}}(j,k)$ represents the number of out-going edges belonging to nodes with in-degree k in the graph after explosions, which in turn represents the number of contagious out-going edges belonging to nodes with in-degree k in the graph before explosions. But so does $\sum_j jp_n(j,k,1)n\mu_n(j,k)$.

So we have for all k

$$\begin{aligned}\sum_j j\frac{l_{\tilde{n}}(j,k)}{\lambda_{\tilde{n}}\tilde{n}} &= \frac{1}{\lambda_{\tilde{n}}\tilde{n}}\sum_j jp_n(j,k,1)n\mu_n(j,k)\\ &= \sum_j jp_n(j,k,1)\frac{\mu_n(j,k)}{\lambda_n}\\ &\overset{n\to\infty}{\to} \sum_j jp(j,k,1)\frac{\mu(j,k)}{\lambda},\end{aligned}$$

where the convergence holds by the second moment property in Assumption 4.6.

Applying Lemma 4.19 to the sequence of degrees in the graph after explosions shows that

when

$$\sum_k k \lim_n \sum_j j \frac{l_{\tilde{n}}(j,k)}{\lambda_{\tilde{n}} \tilde{n}} = \sum_k k \sum_j j p(j,k,1) \frac{\mu(j,k)}{\lambda} > 1,$$

then with high probability there exists a strongly connected giant component in the skeleton of contagious links. This completes the proof.

4.5.5 Appendix: Size of default cascade

In this section, we consider the structure of the skeleton of contagious links more in detail. Let us define the *susceptibility* of a random financial network

$$\chi(\mathbf{E_n}, \gamma_\mathbf{n}) := \frac{1}{n} \sum_{v \in [1,\ldots,n]} |C(v)|, \tag{4.28}$$

with $C(v)$ the default cluster of v containing all nodes from which v is reachable by a directed path of contagious links.

The skeleton of contagious links is the subgraph obtained by retaining only the contagious links in the initial network. Thus, if we consider the new degree sequence for this subgraph, it is still a random graph chosen uniformly from all graphs with this degree sequence (e.g., see Section 4.5.4), so we can still apply asymptotic results for the random configuration model. In particular, Janson [105] shows that the susceptibility of the random graph with given vertex degrees converges under mild conditions to the expected cluster size in the corresponding branching process, which may be defined as a Galton-Watson branching process with initial offspring ξ_0 and general offspring ξ. We define

$$\tilde{\lambda} := \sum_{j,k} j \mu(j,k) p(j,k,1),$$

the average number of contagious links and note that the fraction of contagious links is $T := \frac{\tilde{\lambda}}{\lambda}$. The generating function of the initial offspring ξ_0 is

$$G_0(y) = \sum_{k_0, j, k \geq k_0} \mu(j,k) \binom{k}{k_0} (1-T)^{k-k_0} T^{k_0} y^{k_0} = \sum_{j,k} \mu(j,k)(1-T+Ty)^k,$$

while the generating function of the general offspring is

$$G(y) = \sum_{j,k} \frac{j \mu(j,k) p(j,k,1)}{\tilde{\lambda}} (1-T+Ty)^k.$$

It is easy to see that G_0 represents the generating function of the number of links pointing into a randomly chosen node after bond percolation with probability T (each incoming edge is removed with probability $1-T$ independently of all other incoming edges). In terms of our network model, G represents the generating function of the number of contagious links ending in a node which is start of a randomly chosen contagious link. The probability that such a node has degree (j,k) is given by a weighted version of μ: $\frac{j\mu(j,k)p(j,k,1)}{\tilde{\lambda}}$. We have that

$$\mathbb{E}(\xi) = G'(1) = \sum_{j,k} \frac{j\mu(j,k)p(j,k,1)}{\tilde{\lambda}} kT = \sum_{j,k} \frac{jk\mu(j,k)}{\lambda} p(j,k,1),$$

and

$$\mathbb{E}(\xi_0) = G_0'(1) = \sum_{j,k} k\mu(j,k)T = \tilde{\lambda}.$$

For a branching process with initial offspring ξ_0 and general offspring ξ, its susceptibility is given by $1 + \frac{\mathbb{E}\xi_0}{(1-\mathbb{E}\xi)_+}$ (see [105, Theorem 3.1]). By virtue of [105, Theorem 3.3] applied to the skeleton of contagious links, under Conditions 4.6 and 4.7, the average cascade size converges in probability (and in fact in L^1, in the subcritical case when $\mathbb{E}(\xi) < 1$) to the susceptibility of the corresponding branching process. We have

- If the resilience measure is strictly positive,

$$\chi(\mathbf{E_n}, \gamma_\mathbf{n}) \xrightarrow{L^1} \chi_\infty := 1 + \frac{\sum_{j,k} j\mu(j,k)p(j,k,1)}{1 - \sum_{j,k} \frac{jk}{\lambda}\mu(j,k)p(j,k,1)}.$$

- If the resilience measure is zero or negative,

$$\chi(\mathbf{E_n}, \gamma_\mathbf{n}) \xrightarrow{p} \infty.$$

We show thus by a different method that the positivity of the resilience measure is a necessary condition for the non-occurrence of global cascades: this condition is equivalent to the non-explosion of the branching process associated to the skeleton of contagious links

$$\mathbb{E}(\xi) < 1.$$

The full distribution of the size of the default cluster can be computed once the generating functions G_0 and G are known (see Bertoin and Sidoravicius [20, Theorem 1] which connects the structure of clusters in random graphs with prescribed degree distributions to branching

processes and Newman et al. [131] for the derivation in case of branching processes). We define the generating function H of the size of the default cluster generated by a randomly chosen contagious edge, which verifies the condition $H(y) = yG(H(y))$. The generating function H_0 of the size of a default cluster is then given by $H_0(y) = yG_0(H(y))$. If the resilience measure is negative, then the probability of a large scale epidemic triggered by a single node is equal to the explosion probability of the branching process. If we let y^* be the smallest solution of

$$y = \sum_{j,k} \frac{j\mu(j,k)p(j,k,1)}{\tilde{\lambda}}(1 - T + Ty)^k,$$

then the probability of a global cascade is given by

$$1 - \sum_{j,k} \mu(j,k)(1 - T + Ty^*)^k.$$

This last formula confirms the observations in Gleeson [79] that the probability of occurrence of a global cascade strongly depends on the out-degree distribution even when the average cascade size does not, such as in cases where the degree distribution factorizes and the fraction of contagious links does not depend on the out-degrees.

Discussion

We have proposed a framework for evaluating the impact of a macroeconomic shock on the resilience of a banking network to contagion effects. Our approach complements existing stress tests used by regulators and suggests to monitor the capital adequacy of each institution with regard to its *largest exposures*. In practice, such a stress tests may be implemented in a decentralized fashion by requesting banks to project the effect of a macroeconomic stress scenario on their balance sheets, and report the quantities of interest -mainly the number of exposures exceeding capital in the stress scenario- regulator, who can then assess the resilience of the network using our proposed resilience measure. Our results also suggest that one need not monitor/know the *entire* network of counterparty exposures, but simply the subgraph of "contagious" links, which is much smaller. Moreover, the regulator can efficiently contain contagion by focusing on fragile nodes, especially those with high connectivity, and their counterparties. Higher capital requirements could be imposed on them to reduce their number of contagious links, and insure that the danger of phase transitions as described above is avoided.

Bibliography

[1] L. A. Adamic. The small world web. In *Lecture Notes in Computer Science*, volume 1696, pages 443–452. Springer, 1999.

[2] L. A. Adamic, B. A. Huberman, A. L. Barabási, R. Albert, H. Jeong, and G. Bianconi. Power-law distribution of the world wide web. *Science*, 287(5461):2115, 2000.

[3] L. Addario-Berry, N. Broutin, and C. Goldschmidt. The continuum limit of critical random graphs. Available at http://arxiv.org/abs/0903.4730.

[4] J. Adler and U. Lev. Bootstrap percolation: visualizations and applications. *Brazilian Journal of Physics*, 33(3):641–644, 2003.

[5] R. Albert and A.-L. Barabási. Statistical mechanics of complex networks. *Reviews of Modern Physics*, 74:47–97, 2002.

[6] R. Albert, H. Jeong, and A. L. Barabasi. The diameter of the world wide web. *Nature*, 401:130–131, 1999.

[7] R. Albert, H. Jeong, and A.-L. Barabasi. Error and attack tolerance of complex networks. *Nature*, 406(6794):378–382, 2000.

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

[9] D. Aldous and A. Bandyopadhyay. A survey of max-type recursive distributional equations. *Annals of Applied Probability*, 15(2):1047–1110, 2005.

[10] D. Aldous and R. Lyons. Processes on unimodular random networks. *Electronic Journal of Probability*, 12:1454–1508, 2007.

[11] F. Allen and D. Gale. Financial contagion. *Journal of Political Economy*, 108(1):1–33, 2000.

[12] N. Alon and J. Spencer. *The probabilistic method*. Wiley-Interscience Series in Discrete Mathematics and Optimization. John Wiley & Sons, New York, second edition, 2000.

[13] K. B. Athreya and P. E. Ney. *Branching processes*. Dover Publications, 2004.

[14] J. Balogh and B. Bollobás. Bootstrap percolation on the hypercube. *Probability Theory and Related Fields*, 134(4):624–648, 2006.

[15] J. Balogh, Y. Peres, and G. Pete. Bootstrap percolation on infinite trees and non-amenable groups. *Combinatorics, Probability and Computing*, 15(5):715–730, 2006.

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

[17] A.-L. Barabasi and R. Albert. Emergence of scaling in random networks. *Science*, 286:509–512, 1999.

[18] S. Battiston, D. D. Gatti, M. Gallegati, B. Greenwald, and J. E. Stiglitz. Liaisons dangereuses: increasing connectivity, risk sharing, and systemic risk. Available at http://ssrn.com/paper=1532069, 2009.

[19] E. A. Bender and E. R. Canfield. The asymptotic number of labeled graphs with given degree sequences. *Journal of Combinatorial Theory, Series A*, 24:296–307, 1978.

[20] J. Bertoin and V. Sidoravicius. The structure of typical clusters in large sparse random configurations. *Journal of Statistical Physics*, 135(1):87–105, 2009.

[21] S. Bhamidi. First passage percolation on locally tree-like networks. I. dense random graphs. *Journal of Mathematical Physics*, 49(12):125218, 2008.

[22] S. Bhamidi and R. van der Hofstad. Weak disorder in the stochastic mean field model of distance. Available at http://arxiv.org/abs/1002.4362, 2010.

[23] S. Bhamidi, R. van der Hofstad, and G. Hooghiemstra. Extreme value theory, Poisson-Dirichlet distributions and first passage percolation on random networks. *Advances in applied probability*, 42(3):706–738, 2010.

[24] S. Bhamidi, R. van der Hofstad, and G. Hooghiemstra. First passage percolation on random graphs with finite mean degrees. *Annals of Applied probability*, 20(5):1907–1965, 2010.

[25] S. Bhamidi, R. van der Hofstad, and G. Hooghiemstra. First passage percolation on the Erdös-Rényi random graph. Available at http://arxiv.org/abs/1005.4104, 2010.

[26] P. Billingsley. *Probability and measure*. Wiley-Interscience Series in Probability and Mathematical Statistics. John Wiley & Sons, New York, third edition, 1995.

[27] P. Blanchard, C.-H. Chang, and T. Krüger. Epidemic thresholds on scale-free graphs: the interplay between exponent and preferential choice. *Annales Henri Poincaré*, 4(suppl. 2):S957–S970, 2003.

[28] S. Boccaletti, V. Latora, Y. Moreno, M. Chavez, and D. Hwang. Complex networks: structure and dynamics. *Physics Reports*, 424(4-5):175–308, 2006.

[29] B. Bollobás. The evolution of random graphs. *Transactions of the American Mathematical Society*, 286(1):257–274, 1984.

[30] B. Bollobás. *Random graphs*. Cambridge University Press, 2001.

[31] B. Bollobás and W. F. de la Vega. The diameter of random regular graphs. *Combinatorica*, 2(2):125–134, 1982.

[32] B. Bollobás, S. Janson, and O. Riordan. The phase transition in inhomogeneous random graphs. *Random Structures & Algorithms*, 31:3–122, 2007.

[33] B. Bollobás and O. Riordan. *Percolation*. Cambridge University Press, 2006.

[34] M. Boss, H. Elsinger, M. Summer, and S. Thurner. The network topology of the interbank market. *Quantitative Finance*, (4):677–684, 2004.

[35] I. Breskin, J. Soriano, E. Moses, and T. Tlusty. Percolation in living neural networks. *Physical Review Letters*, 97(18), 2006.

[36] T. Britton, S. Janson, and A. Martin-Löf. Graphs with specified degree distributions, simple epidemics and local vaccination strategies. *Advances in applied probability*, 39(4):922–948, 2007.

[37] A. Broder, R. Kumar, F. Maghoul, P. Raghavan, S. Rajagopalan, R. Stata, A. Tomkins, and J. Wiener. Graph structure in the web. *Computer Networks*, 33:309–320, 2000.

[38] J. Cain and N. Wormald. Encores on cores. *Electronic Journal of Combinatorics 13 (2006), RP 81.*

[39] R. Cerf and F. Manzo. The threshold regime of finite volume bootstrap percolation. *Stochastic Processes and their Applications*, 101(1):69–82, 2002.

[40] J. Chalupa, G. R. Reich, and P. L. Leath. Bootstrap percolation on a bethe lattice. *Journal of Physics C*, 12:L31–L35, 1979.

[41] F. Chung and L. Lu. The diameter of sparse random graphs. *Advances in Applied Mathematics*, 26:257–279, 2001.

[42] F. Chung and L. Lu. *Complex graphs and networks.* American Mathematical Society, 2006.

[43] O. Cohen, A. Kesselman, M. R. Martinez, J. Soriano, E. Moses, and T. Tlusty. Quorum percolation: more is different in living neural networks. *Europhysics Letters*, 89(1).

[44] R. Cont. Measuring systemic risk. *Working Paper*, 2010.

[45] R. Cont, A. Moussa, and E. B. e Santos. The brazilian financial system: network structure and systemic risk analysis. Working paper, 2010.

[46] C. Cooper. The cores of random hypergraphs with a given degree sequence. *Random Structures & Algorithms*, 25(4):353–375, 2004.

[47] C. Cooper and A. M. Frieze. The size of the largest strongly connected component of a random digraph with a given degree sequence. *Combinatorics, Probability & Computing*, 13(3):319–337, 2004.

[48] R. Darling and J. Norris. Differential equation approximations for markov chains. *Probability Surveys*, 5:37–79, 2008.

[49] H. Degryse and G. Nguyen. Interbank exposures: an empirical examination of systemic risk in the belgian banking system. *International Journal of Central Banking*, (3(2)):123–171, 2007.

[50] A. Dembo and O. Zeitouni. *Large deviation techniques and applications*. 2nd ed., Springer-Verlag, 1998.

[51] J. Ding, J. H. Kim, E. Lubetzky, and Y. Peres. Anatomy of a young giant component in the random graph. Available at http://arxiv.org/abs/0906.1839.

[52] J. Ding, J. H. Kim, E. Lubetzky, and Y. Peres. Diameters in supercritical random graphs via first passage percolation. Available at http://arxiv.org/abs/0906.1840, 2009.

[53] S. N. Dorogovtsev, A. V. Goltsev, and J. F. F. Mendes. Critical phenomena in complex networks. *Reviews of Modern Physics*, 80(4):1275–1335, 2008.

[54] M. Draief, A. J. Ganesh, and L. Massoulié. Thresholds for virus spread on networks. *Annals of Applied Probability*, 18(2):359–378, 2008.

[55] M. Draief and L. Massoulié. *Epidemics and rumours in complex networks*. Cambridge University Press, New York, NY, USA, 1st edition, 2010.

[56] R. Durrett. *Random graph dynamics*. Cambridge University Press, New York, NY, USA, 2006.

[57] L. Eisenberg and T. H. Noe. Systemic risk in financial systems. *Management Science*, 47(2):236–249, 2001.

[58] R. Elsässer and T. Sauerwald. On the runtime and robustness of randomized broadcasting. *Theoretical Computer Science*, 410(36):3414–3427, 2009.

[59] H. Elsinger, A. Lehar, and M. Summer. Risk assessment for banking systems. *Management Science*, 52(9):1301–1314, 2006.

[60] H. Elsinger, A. Lehar, and M. Summer. Using market information for banking system risk assessment. *International Journal of Central Banking*, pages 137–165, 2006.

[61] A. C. D. van Enter. Proof of Straley's argument for bootstrap percolation. *Journal of Statistical Physics*, 48(3-4):943–945, 1987.

[62] P. Erdős and A. Rényi. On random graphs. I. *Publicationes Mathematicae*, 6:290–297, 1959.

[63] P. Erdös and A. Rényi. On the evolution of random graphs. *Publication of the Mathematical Institute of the Hungarian Academy of Sciences*, 5:17–61, 1960.

[64] P. Erdös and A. Rényi. On the evolution of random graphs. In *Publication of the Mathematical Institute of the Hungarian Academy of Sciences*, pages 17–61, 1960.

[65] H. van den Esker, R. van der Hofstad, and G. Hooghiemstra. Universality for the distance in finite variance random graphs. *Journal of Statistical Physics*, 133(1):169–202, 2008.

[66] H. van den Esker, R. van der Hofstad, G. Hooghiemstra, and D. Znamenski. Distances in random graphs with infinite mean degrees. *Extremes*, 8:111–141, 2006.

[67] W. Feller. *An introduction to probability theory and its applications, 2nd Edition. Vol. 2.* Wiley, 1991.

[68] D. Fernholz and V. Ramachandran. Cores and connectivity in sparse random graphs. *Technical Report TR-04-13, The University of Texas at Austin, Department of Computer Sciences*, 2004.

[69] D. Fernholz and V. Ramachandran. The diameter of sparse random graphs. *Random Structures & Algorithms*, 31(4):482–516, 2007.

[70] L. Fontes and R. Schonmann. Bootstrap percolation on homogeneous trees has 2 phase transitions. *Journal of Statistical Physics*, 132:839–861, 2008.

[71] L. R. Fontes, R. H. Schonmann, and V. Sidoravicius. Stretched exponential fixation in stochastic Ising models at zero temperature. *Communications in Mathematical Physics*, 228:495–518, 2002.

[72] N. Fountoulakis. Percolation on sparse random graphs with given degree sequence. *Internet Mathematics*, 4(4):329–356, 2007.

[73] N. Fountoulakis, A. Huber, and K. Panagiotou. Reliable broadcasting in random networks and the effect of density. *In Proceedings of INFOCOM*, 2010.

[74] N. Fountoulakis and K. Panagiotou. Rumor spreading on random regular graphs and expanders. Available at www.mpi-inf.mpg.de/~fountoul/RRRS.pdf, 2010.

[75] N. Fountoulakis and B. A. Reed. 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, 2009.

[76] A. Frieze and G. Grimmett. The shortest path problem for graphs with random arc-lengths. *Discrete Applied Mathematics*, 10:57–77, 1985.

[77] P. Gai and S. Kapadia. Contagion in financial networks. Available at SSRN:http://ssrn.com/abstract=1577043, 2010.

[78] A. Ganesh, L. Massoulié, and D. Towsley. The effect of network topology on the spread of epidemics. In *IEEE INFOCOM*, pages 1455–1466, 2005.

[79] J. P. Gleeson. Mean size of avalanches on directed random networks with arbitrary degree distributions. *Physical Review E*, 77(5):057101, 2008.

[80] A. V. Goltsev, F. V. de Abreu, S. N. Dorogovtsev, and J. F. F. Mendes. Stochastic cellular automata model of neural networks. Available at http://arxiv.org/abs/0904.2189, 2009.

[81] P. D. Gregorio, A. Lawlor, P. Bradley, and K. A. Dawson. Exact solution of a jamming transition: closed equations for a bootstrap percolation problem. *Proceedings of the National Academy of Sciences*, 102:5669–5673, 2005.

[82] G. Grimmett. *Percolation*. Springer, Berlin, 2nd edition, 1999.

[83] G. Grimmett and H. Kesten. First-passage percolation, network flows and electrical resistances. *Probability Theory and Related Fields*, 66:335–366, 1984. 10.1007/BF00533701.

[84] A. Gut. *Probability: a graduate course*. Springer, New York, 2005.

[85] O. Häggström and R. Pemantle. First passage percolation and a model for competing spatial growth. *Journal of Applied Probability*, 35(3):683–692, 1998.

[86] T. E. Harris. Some mathematical models for branching processes. *Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability*, pages 305–328.

[87] T. E. Harris. *The theory of branching processes*. Springer, Berlin, 1963.

[88] H. Hatami and M. Molloy. The scaling window for a random graph with a given degree sequence. Available at http://arxiv.org/abs/0907.4211.

[89] B. Hayes. Graph theory in practice: Part I. *American Scientist*, 88(1):9–13, 2000.

[90] R. van der Hofstad. Random graphs and complex networks. Available at http://www.win.tue.nl/~rhofstad/NotesRGCN.pdf , 2009.

[91] R. van der Hofstad. *Percolation and random graphs*. Survey on Percolation and Random Graphs, Chapter 6 of New Perspectives on Stochastic Geometry edited by Ilya Molchanov and Wilfrid Kendall, 2010.

[92] R. van der Hofstad, G. Hooghiemstra, and P. Van Mieghem. First-passage percolation on the random graph. *Probability in the Engineering and Informational Sciences*, 15:225–237, 2001.

[93] R. van der Hofstad, G. Hooghiemstra, and P. Van Mieghem. The flooding time in the random graph. *Extremes*, 5(2):111–129, 2002.

[94] R. van der Hofstad, G. Hooghiemstra, and P. Van Mieghem. Distances in random graphs with finite variance degrees. *Random Structures & Algorithms*, 27(1):76–123, 2005.

[95] R. van der Hofstad, G. Hooghiemstra, and D. Znamenski. Distances in random graphs with finite mean and infinite variance degrees. *Electronic Journal of Probability*, 12:703–766, 2007.

[96] F. den Hollander. *Large deviations*. volume 14 of Fields Institute Monographs. American Mathematical Society, Providence, RI, 2000.

[97] A. E. Holroyd. Sharp metastability threshold for two-dimensional bootstrap percolation. *Probability Theory and Related Fields*, 125(2):195–224, 2003.

[98] G. Hooghiemstra and P. Van Mieghem. The weight and hopcount of the shortest path in the complete graph with exponential weights. *Combinatorics, Probability and Computing*, 17(4):537–548, 2008.

[99] W. Hurewicz. *Lectures on ordinary differential equations*. M.I.T. Press, 1958.

[100] M. O. Jackson and L. Yariv. Diffusion on social networks. *Economie publique*, 16:69–82, 2005.

[101] P. Jagers. *Branching processes with biological applications*. Wiley, New York, 1975.

[102] S. Janson. One, two and three times log n/n for paths in a complete graph with random weights. *Combinatorics, Probability and Computing*, 8(4):347–361, 1999.

[103] S. Janson. On percolation in random graphs with given vertex degrees. *Electronic Journal of Probability*, 14:86–118, 2009.

[104] S. Janson. The probability that a random multigraph is simple. *Combinatorics, Probability and Computing*, 18(1-2):205–225, 2009.

[105] S. Janson. Susceptibility of random graphs with given vertex degrees. Available at http://arxiv.org/abs/0911.2636v1, 2009.

[106] S. Janson. Bootstrap percolation on the random graph $G_{n,p}$. Available at http://arxiv.org/abs/1012.3535, 2010.

[107] S. Janson, D. E. Knuth, T. Łuczak, and B. Pittel. The birth of the giant component. *Random Structures & Algorithms*, 4(3):233–359, 1993.

[108] S. Janson and M. Łuczak. A simple solution to the k-core problem. *Random Structures & Algorithms*, 30:50–62, 2007.

[109] S. Janson and M. J. Łuczak. Asymptotic normality of the k-core in random graphs. *Annals of Applied Probability*, 18:1085, 2008.

[110] S. Janson and M. J. Łuczak. A new approach to the giant component problem. *Random Structures & Algorithms*, 34(2):197–216, 2009.

[111] S. Janson, T. Łuczak, and A. Rucinski. *Random graphs*. Wiley-Interscience Series in Discrete Mathematics and Optimization. Wiley-Interscience, New York, 2000.

[112] S. Janson and J. Spencer. A point process describing the component sizes in the critical window of the random graph evolution. *Combinatorics, Probability and Computing*, 16(4):631–658, 2007.

[113] O. Kallenberg. *Foundations of modern probability*. 2nd ed., Springer-Verlag, New York, 2002.

[114] M. Kang and T. g. Seierstad. The critical phase for random graphs with a given degree sequence. *Combinatorics, Probability and Computing*, 17:67–86, 2008.

[115] D. Kempe, J. Kleinberg, and E. Tardos. Maximizing the spread of influence through a social network. In *KDD '03: Proceedings of the ninth ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 137–146, New York, NY, USA, 2003. ACM.

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

[117] H. Kesten. Aspects of first passage percolation. In *École d'été de probabilités de Saint-Flour, XIV—1984*, volume 1180 of *Lecture Notes in Math.*, pages 125–264. Springer, Berlin, 1986.

[118] H. Kesten and B. P. Stigum. A limit theorem for multidimensional Galton-Watson processes. *The Annals of Mathematical Statistics*, 37(5):1211–1223, 1966.

[119] J. Kleinberg. Cascading behavior in networks: Algorithmic and economic issues. In *Algorithmic Game Theory*. Cambridge University Press, 2007.

[120] A. Klenke and L. Mattner. Stochastic ordering of classical discrete distributions. *To appear in Advances in Applied probability*, 2010.

[121] M. Lelarge. Diffusion of innovations on random networks: understanding the chasm. In *Proceedings of WINE 2008*, pages 178–185, 2008.

[122] M. Lelarge. Efficient control of epidemics over random networks. In *Proceedings of SIGMETRICS*, pages 1–12, New York, NY, USA, 2009. ACM.

[123] M. Lelarge. Diffusion and cascading behavior in random networks. Available at `http://arxiv.org/abs/1012.2062`, 2010.

[124] T. Łuczak. Component behavior near the critical point of the random graph process. *Random Structures & Algorithms*, 1:287–310, 1990.

[125] T. Łuczak, B. Pittel, and J. C. Wierman. The structure of a random graph at the point of the phase transition. *Transactions of the American Mathematical Society*, 341(2):721–748, 1994.

[126] P. Van Mieghem. *Performance analysis of communications networks and systems*. Cambridge University Press, 2006.

[127] P. Mistrulli. Assessing financial contagion in the interbank market: a comparison between estimated and observed bilateral exposures. Working paper, 2006.

[128] M. Molloy and B. Reed. A critical point for random graphs with a given degree sequence. *Random Structures & Algorithms*, 6:161–179, 1995.

[129] M. Molloy and B. Reed. The size of the giant component of a random graph with a given degree sequence. *Combinatorics, Probability and Computing*, 7:295–305, 1998.

[130] S. Morris. Contagion. *Review of Economic Studies*, 67(1):57–78, 2000.

[131] M. E. J. Newman, S. H. Strogatz, and D. J. Watts. Random graphs with arbitrary degree distributions and their applications. *Physical Review E*, 64:026118, 2001.

[132] M. E. J. Newman, D. J. Watts, and S. H. Strogatz. Random graph models of social networks. *Proceedings of the National Academy of Sciences of the United States of America*, 99(Suppl 1):2566–2572, 2002.

[133] E. Nier, J. Yang, T. Yorulmazer, and A. Alentorn. Network models and finacial stability. *Journal of Economic Dynamics and Control*, (31):2033–60, 2007.

[134] R. Pastor-Satorras and A. Vespignani. Epidemic spreading in scale-free networks. *Physical Review Letters*, 86(14):3200–3203, 2001.

[135] D. Peleg. Distributed computing: a locality-sensitive approach. *SIAM Monographs on Discrete Mathematics and Applications*, 5, 2000.

[136] B. Pittel. On spreading a rumor. *SIAM Journal on Applied Mathematics*, 1:213–223, 1987.

[137] B. Pittel. On tree census and the giant component in sparse random graphs. *Random Structures & Algorithms*, 1(3):311–342, 1990.

[138] B. Pittel. On the largest component of the random graph at a nearcritical stage. *Journal Combinatorial Theory Series B*, 82(2):237–269, 2001.

[139] B. Pittel, J. Spencer, and N. Wormald. Sudden emergence of a giant k-core in a random graph. *Journal Combinatorial Theory Series B*, 67(1):111–151, 1996.

[140] H. Reittu and I. Norros. On the power-law random graph model of massive data networks. *Performance Evaluation*, 55:3–23, 2004.

[141] O. Riordan and N. Wormald. The diameter of sparse random graphs. Available at `http://arxiv.org/abs/0808.4067`, 2010.

[142] S. Sabhapandit, D. Dhar, and P. Shukla. Hysteresis in the random-field Ising model and bootstrap percolation. *Physical Review Letters*, 88(19):197202, 2002.

[143] M. Sellitto, G. Biroli, and C. Toninelli. Facilitated spin models on bethe lattice: bootstrap percolation, mode-coupling transition and glassy dynamics. *Europhysics Letters*, 69:496–502, 2005.

[144] B. A. Sevastyanov. The theory of branching random processes(russ). *Usephi Mat. Nauk.*, 6:47–99, 1951.

[145] J. Soriano, M. Rodríguez Martínez, T. Tlusty, and E. Moses. Development of input connections in neural cultures. *Proceedings of the National Academy of Sciences of the United States of America*, 2008.

[146] D. Stauffer and A. Aharony. *Introduction to percolation theory*. Taylor and Francis Ed. London, 1985.

[147] T. Tlusty and J. P. Eckmann. Remarks on bootstrap percolation in metric networks. *Journal of Physics A: Mathematical and Theoretical*, 42:205004, 2009.

[148] C. Toninelli, G. Biroli, and D. S. Fisher. Jamming percolation and glass transitions in lattice models. *Physical Review Letters*, 96(3):035702, 2006.

[149] C. Upper. Using counterfactual simulations to assess the danger of contagion in interbank markets. Available at `http://ssrn.com/paper=1014320`, 2007.

[150] C. Upper and A. Worms. Estimating bilateral exposures in the german interbank market: is there a danger of contagion? *Deutsche Bundesbank Discussion Paper No 09/02*, 2002.

[151] D. J. Watts. A simple model of global cascades on random networks. *Proceedings of the National Academy of Sciences*, 99(9):5766–5771, 2002.

[152] D. J. Watts and S. H. Strogatz. Collective dynamics of 'small-world' networks. *Nature*, 393(6684):440–442, 1998.

[153] D. Williams. *Probability with martingales*. Cambridge Mathematical Textbooks. Cambridge University Press, Cambridge, 1991.

[154] N. Wormald. Differential equations for random processes and random graphs. *Annals of Applied Probability*, 5(4):1217–1235, 1995.

[155] N. Wormald. The differential equation method for random graph processes and greedy algorithms. In *Lectures on Approximation and Randomized Algorithms*, 1999.

[156] N. Wormald. Models of random regular graphs. *Surveys in combinatorics*, 267:239–298, 1999.

Appendix A

Probabilistic Methods

In this appendix, we provide some classical probabilistic tools we need through this thesis. For details and proofs of these results, we refer to [26, 12, 50, 90, 155].

A.1 Coupling

The coupling method is a fruitful probabilistic tool. Its range of applications is large: Markov chain, renewal theory, perfect simulation or the Stein-Chen method. In this section, we will record a few simple results on this method.

In general, two random variables X and Y are coupled when they are defined on the same probability space. This means that there is one probability law $\mathbb{P}$ such that $\mathbb{P}(X \in E, Y \in F)$ are defined for all events E and F. This is formalized in the following definition:

Definition A.1 (Coupling of random variables)**.** The random variables $(\hat{X}_1, ..., \hat{X}_n)$ are a coupling of the random variables $X_1, ..., X_n$ when $(\hat{X}_1, ..., \hat{X}_n)$ are defined on the same probability space, and are such that the marginal distribution of $\hat{X}_i$ is the same as the distribution of X_i for all $i = 1, ..., n$, i.e., for all measurable subsets E of $\mathbb{R}$, we have

$$\mathbb{P}(\hat{X}_i \in E) = \mathbb{P}(X_i \in E). \tag{A.1}$$

Couplings are very useful in proving that random variables are related. The key point of the above definition is that while the random variables $X_1, ..., X_n$ may be defined on different

probability spaces, the coupled random variables $(\hat{X}_1, ..., \hat{X}_n)$ are defined on the same probability space. The coupled random variables $(\hat{X}_1, ..., \hat{X}_n)$ are related to the original random variables $X_1, ..., X_n$ by the fact that the marginal distributions of $(\hat{X}_1, ..., \hat{X}_n)$ are equal to the random variables $X_1, ..., X_n$.

Theorem A.2 (Poisson limit for binomial random variables)**.** *Let $\{I_i\}_{i=1}^n$ be independent with $I_i \sim \mathrm{Ber}(p_i)$, and let $\lambda = \sum_{i=1}^n p_i$. Let $X = \sum_{i=1}^n I_i$, and let Y be a Poisson random variable with parameter λ. Then, there exists a coupling $(\hat{X}, \hat{Y})$ of (X, Y) such that*

$$\mathbb{P}(\hat{X} \neq \hat{Y}) \leq \sum_{i=1}^n p_i^2. \tag{A.2}$$

Consequently, for every $\lambda \geq 0$ and $n \in \mathbb{N}$, there exists a coupling $(\hat{X}, \hat{Y})$, where $\hat{X} \sim \mathrm{Bin}(n, \lambda/n)$ and $\hat{Y} \sim \mathrm{Poi}(\lambda)$ such that

$$\mathbb{P}(\hat{X} \neq \hat{Y}) \leq \frac{\lambda^2}{n}. \tag{A.3}$$

A.2 Stochastic Domination

To compare random variables, we will rely on the notion of stochastic ordering, which is defined as follows.

Definition A.3 (Stochastic domination)**.** Let X and Y be two random variables, not necessarily living on the same probability space. The random variable X is stochastically smaller than the random variable Y when the inequality

$$\mathbb{P}(X \leq x) \geq \mathbb{P}(Y \leq x) \tag{A.4}$$

holds for all $x \in \mathbb{R}$. We denote this by $X \leq_{st} Y$.

Strassen's Theorem shows that the stochastic ordering $\leq_{st}$ corresponds to the ordering $\leq$ through a coupling.

Theorem A.4. *The random variable X is stochastically smaller than the random variable Y if and only if there exists a coupling $(\hat{X}, \hat{Y})$ of X, Y such that*

$$\mathbb{P}(\hat{X} \leq \hat{Y}) = 1. \tag{A.5}$$

We now state two consequences of stochastic domination, that we will use later.

Theorem A.5. *Suppose $X \leq_{st} Y$. Then*

(i) we have $\mathbb{E}X \leq \mathbb{E}Y$.

(ii) if $g : \mathbb{R} \to \mathbb{R}$ is non-decreasing, then $g(X) \leq_{st} g(Y)$.

A.3 Probabilistic Bounds

We will often make use of a number of probabilistic bounds, which we will summarize in this section.

Theorem A.6 (Markov inequality)**.** *Let X be a non-negative random variable with $\mathbb{E}[X] < \infty$. Then*

$$\mathbb{P}(X \geq a) \leq \frac{\mathbb{E}[X]}{a}. \tag{A.6}$$

Theorem A.7 (Chebysev inequality)**.** *Let X be a random variable with finite variance σ^2. Then for any real number $\alpha > 0$,*

$$\mathbb{P}\left(|X - \mathbb{E}X| \geq \alpha\right) \leq \frac{\sigma^2}{\alpha^2}. \tag{A.7}$$

We will often rely on bounds on the probability that a sum of independent random variables is larger than its expectation. We consider a sequence of i.i.d. random variables $\{X_i, i \geq 1\}$ distributed like X, taking values in $\mathbb{R}$. Let $S_n = \sum_{i=1}^n X_i$. We are interested in proving exponential bounds for the right tail probability $\mathbb{P}(S_n \geq tn)$ when $t \geq \mathbb{E}X$, as $n \to \infty$. Similar results are easily derived for left tails by considering $\{-X_i, i \geq 1\}$. Let $\alpha > 0$, then

$$\mathbb{P}(S_n \geq t_n) = \mathbb{P}(e^{\alpha S_n} \geq e^{\alpha t n}).$$

It follows using Markov's inequality that

$$\mathbb{P}(S_n \geq tn) \leq e^{-\alpha t n}\mathbb{E}e^{\alpha S_n} = e^{-\alpha t n}\prod_{i=1}^n \mathbb{E}e^{\alpha X_i} \tag{A.8}$$

since the variables X_i are independent. We infer that

$$\mathbb{P}(S_n \geq tn) \leq e^{-\alpha t n + \Lambda(\alpha) n},$$

where $\Lambda(\alpha) := \log \mathbb{E}e^{\alpha X}$ is the *cumulant generating function*. So optimizing the choice of α, we have

$$\mathbb{P}(S_n \geq tn) \leq e^{-n\Lambda^*(t)}, \tag{A.9}$$

where $\Lambda^*(t) := \sup_{\alpha>0} \alpha t - \Lambda(\alpha)$ is the Fenchel-Legendre (convex) dual of λ. The following theorem states that the above upper bound is tight. The proof can be found in [50].

Theorem A.8 (Cramér). *Assume that $\Lambda(\alpha) < \infty$ for some $\alpha > 0$. Let $t \geq \mathbb{E}X$. Then, as $n \to \infty$,*

$$\mathbb{P}(S_n \geq tn) = e^{-n\Lambda^*(t)+o(n)}. \tag{A.10}$$

For more background on large deviations, we refer to [50, 96].

We now state Azuma-Hoeffding inequality concerning martingales. For more details on martingales, we refer the reader to [153].

Theorem A.9 (Azuma-Hoeffding inequality). *Let $X_0, X_1, ..., X_t$ be a martingale such that $X_i - X_{i-1} \leq c_i$, for $1 \leq i \leq t$, for constants c_i. Then for any $\alpha > 0$,*

$$\mathbb{P}(|X_t - X_0| \geq \alpha) \leq 2\exp\left(-\frac{\alpha^2}{2\sum_{i=1}^{t} c_i^2}\right). \tag{A.11}$$

A.4 The Differential Equation Method

In this section we briefly present a method introduced by Wormald in [154] for the analysis of a discrete random process by using differential equations. In particular we recall a general purpose theorem for the use of this method. This method has been used to analyze several kinds of algorithms on random graphs and random regular graphs (see e.g., [38, 48, 129, 155]).

Recall that a function $f(u_1, ..., u_j)$ satisfies a Lipschitz condition on $D \in \mathbb{R}^j$ if a constant $L > 0$ exists with the property that

$$|f(u_1, ..., u_j) - f(v_1, ..., v_j)| \leq L \max_{1\leq i\leq j} |u_i - v_i|$$

for all $(u_1, ..., u_j)$ and $(v_1, ..., v_j)$ in D. For variables $Y_1, ..., Y_b$ and for $D \in \mathbb{R}^{b+1}$, the stopping time $T_D(Y_1, ..., Y_b)$ is defined to be the minimum t such that

$$(t/n; Y_1(t)/n, ..., Y_b(t)/n) \notin D.$$

This is written as T_D when $Y_1, ..., Y_b$ are understood from the context.

The following theorem is the Theorem 5.1 of [155]. In it, "uniformly" refers to the convergence implicit in the $o()$ terms. Hypothesis (1) ensures that Y_t does not change too quickly throughout the process. Hypothesis (2) tells us what we expect for the rate of change to be, and property (3) ensures that this rate does not change too quickly.

Theorem A.10 (Wormald [155])**.** *Let b be given (b is the number of variables). For $1 \leq l \leq b$, suppose $Y_l^{(n)}(t)$ is a sequence of real-valued random variables, such that $0 \leq Y_l^{(n)}(t) \leq C_0 n$ for some constant C_0, and H_t be the history of the sequence, i.e., the sequence*

$$\left\{ Y_j^{(n)}(k),\ 0 \leq j \leq b,\ 0 \leq k \leq t \right\}.$$

Suppose also that for some bounded connected open set $D \subseteq \mathbb{R}^{b+1}$ containing the intersection of $\{(t, z_1, ..., z_b) : t \geq 0\}$ with some neighborhood of

$$\{(0, z_1, ..., z_b) : \mathbb{P}(Y_l^{(n)}(0) = z_l n, 1 \leq l \leq b) \neq 0 \text{ for some } n\},$$

the following three conditions are verified:

1. (Boundedness). *For some function $\beta = \beta(n) \geq 1$ and for all $t < T_D$*

$$\max_{1 \leq l \leq b} |Y_l^{(n)}(t+1) - Y_l^{(n)}(t)| \leq \beta;$$

2. (Trend). *For some function $\lambda = \lambda_1(n) = o(1)$ and for all $l \leq b$ and $t < T_D$*

$$|\mathbb{E}[Y_l^{(n)}(t+1) - Y_l^{(n)}(t) | H_t] - f_l(t/n, Y_1^{(n)}(t)/n, ..., Y_b^{(n)}(t)/n)| \leq \lambda_1;$$

3. (Lipschitz). *For each l the function f_l is continuous and satisfies a Lipschitz condition on D with all Lipschitz constants uniformly bounded.*

Then the following holds

(a) *For $(0, \hat{z}_1, ..., \hat{z}_b) \in D$, the system of differential equations*

$$\frac{dz_l}{ds} = f_l(s, z_1, ..., z_b), \quad l = 1, ..., b,$$

has a unique solution in D, $z_l : \mathbb{R} \to \mathbb{R}$ for $l = 1, \ldots, b$, which passes through $z_l(0) = \hat{z}_l$, $l = 1, \ldots, b,$, and which extends to points arbitrarily close to the boundary of D.

(b) *Let $\lambda > \lambda_1$ with $\lambda = o(1)$. For a sufficiently large constant C, with probability $1 - O\left(\frac{b\beta}{\lambda}\exp\left(-\frac{n\lambda^3}{\beta^3}\right)\right)$, we have*

$$Y_l^{(n)}(t) = nz_l(t/n) + O(\lambda n)$$

uniformly for $0 \le t \le \sigma n$ and for each l. Here $z_l(t)$ is the solution in (a) with $\hat{z}_l = Y_l^{(n)}(0)/n$, and $\sigma = \sigma(n)$ is the supremum of those s to which the solution can be extended before reaching within l^∞-distance $C\lambda$ of the boundary of D.

Proof. The solution is unique from a standard result in the theory of first order differential equations (see Hurewicz[99], Chapter 2, Theorem 11). We now present the proof of part (b). We will use the following supermartingale inequality. The proof follows from exactly the same proof as Azuma's inequality (see [155], Lemma 4.2).

Lemma A.11. *Let $\{X_i\}_{i=0}^t$ be a supermartingale with $X_0 = 0$ and $X_i - X_{i-1} \le c_i$ for $i \ge 1$ and some constants c_i. Then for all $\alpha > 0$,*

$$\mathbb{P}(X_t \ge \alpha) \le \exp\left(-\frac{\alpha^2}{2\sum c_i^2}\right).$$

Let us define $\omega = \lceil n\lambda/\beta \rceil$, $\alpha = n\lambda^3/\beta^3$ and let $0 \le t \le \sigma n$. If $\omega < n^{2/3}$ then $\beta/\lambda > n^{1/3}$ and then the probability in the conclusion is not restricted and there is nothing to prove.

Lemma A.12. *For some constant B with probability $1 - O(e^{-\alpha})$, we have*

$$|Y_l(t+\omega) - Y_l(t) - \omega f_l(t/n, I_1(t)/n, ..., I_b(t)/n)| < B\omega\lambda.$$

Proof. For $0 \le k < \omega$, we have $k\beta/n = O(\lambda)$ and by the Trend and Lipschitz hypotheses

$$\begin{aligned}\mathbb{E}\left[Y_l(t+k+1) - Y_l(t+k)|H_{t+k}\right] &= f_l(\frac{t+k}{n}, \frac{Y_1(t+k)}{n}, ..., \frac{Y_b(t+k)}{n}) + O(\lambda) \\ &= f_l(t/n, Y_1(t)/n, ..., Y_b(t)/n) + O(\lambda).\end{aligned}$$

Hence there exists a function $g(n) = O(\lambda)$ such that conditional on F_t,

$$f(k) := Y_l(t+k) - Y_l(t) - kf_l(t/n, Y_1(t)/n, ..., Y_b(t)/n) - kg(n)$$

is a supermartingale with respect to the sequence σ-fields generated by $F_t, ..., F_{t+\omega}$. By the boundedness hypothesis

$$|f(k+1) - f(k)| \le \beta + O(1) \le \kappa\beta$$

for some constant $\kappa > 0$. Therefore, by Lemma A.11,

$$\mathbb{P}\left(|Y_l(t+\omega) - Y_l(t) - \omega f_l(t/n, Y_1(t)/n, ..., Y_b(t)/n)| \geq \omega g(n) + \kappa\beta\sqrt{\omega\alpha}|F_t\right) \leq 2e^{-\alpha},$$

and so the lemma follows. □

Let $i = \lfloor n\sigma/\omega \rfloor$, and let $h_l(k) = |Y_l(k\omega) - nz_l(k\omega/n)|$ for $0 \leq k \leq i$. We have

$$h_l(k+1) \leq h_l(k) + |A_1| + |A_2| + |A_3|$$

where

$$\begin{aligned} A_1 &= Y_l((k+1)\omega) - Y_l(k\omega) - \omega f_l(k\omega/n, Y_1(k\omega)/n, ..., Y_b(k\omega)/n), \\ A_2 &= \omega z_l'(k\omega/n) + z_l(k\omega/n)n - z_l((k+1)\omega/n)n, \\ A_3 &= \omega f_l(k\omega/n, Y_1(k\omega)/n, ..., Y_b(k\omega)/n) - \omega z_l'(k\omega/n). \end{aligned}$$

By Lemma A.12, we have for a suitable universal constant B', $|A_1| < B'\omega\lambda$ with probability $1 - O(e^{-\alpha})$ (This is the point where the assumption, the scaled variables not approaching within distance $C\lambda$ of the boundary of Ω, is justified). Since f_l satisfies the Lipschitz hypothesis, we have

$$|A_2| = O(n\left(\frac{\omega}{n}\right)^2) < B''\omega^2/n$$

for a suitable constant B''. Finally using the same arguments as above we obtain

$$|A_3| < \frac{B''\omega}{n} h_l(k).$$

Set $B = \max\{B', B''\}$. By induction on k, we infer that

$$\mathbb{P}\left(h_l(k) \geq B_k \text{ for some } k \leq i, 1 \leq l \leq b\right) = O(bie^{-\alpha}), \tag{A.12}$$

where

$$B_k = B\omega\left(\lambda + \omega/n\right)\left((1 + B\omega/n)^k - 1\right)\frac{n}{B\omega}.$$

We have $B_k = O(n\lambda + \omega) = O(n\lambda)$ since β is bounded below. This proves the theorem in the case $t = k\omega$. Assume $t \leq n\sigma$. From time $\lfloor t/\omega \rfloor\omega$ to t the change in Y and z is at most $\omega\beta = O(n\lambda)$ and the theorem follows.

□

Remark A.13. *([155]) A version of the theorem also holds with the number a of variables b function of n, the domain D a function of n but with all Lipschitz constants uniformly bounded, and each function f_l depends only on s and $z_1, ..., z_l$.*

We will also use the following corollary of the above theorem, which is namely Theorem 6.1 of [155]. This theorem states that, as long as Condition 3 holds in D, the solution of the system of equations above can be extended beyond the boundary of $\hat{D}$, into D.

Corollary A.14. *For any set $\hat{D} \subseteq \mathbb{R}^{b+1}$, let $T_{\hat{D}} = T_{\hat{D}}(Y_1^{(n)}, ..., Y_b^{(n)})$ be the minimum t such that $(\frac{t}{n}, \frac{Y_1^{(n)}(t)}{n}, \ldots, \frac{Y_b^{(n)}(t)}{n}) \notin \hat{D}$ (the stopping time). Assume in addition that the first two hypotheses of Theorem A.10 are verified but only within the restricted range $t < T_{\hat{D}}$ of t. Then the conclusions of the theorem hold as before, after replacing $0 \leq t \leq \sigma n$ by $0 \leq t \leq \min\{\sigma n, T_{\hat{D}}\}$.*

Proof. For $1 \leq j \leq b$, define random variables $\hat{Y}_j^{(n)}$ by

$$\hat{Y}_j^{(n)}(t+1) = \begin{cases} Y_j^{(n)}(t+1) & \text{if } t < T_{\hat{D}} \\ Y_j^{(n)}(t) + f_j(t/n, Y_1^{(n)}(t)/n, ..., Y_b^{(n)}(t)/n) & \text{otherwise} \end{cases}$$

for all $t \geq 0$. Then the $\hat{Y}_j^{(n)}$ satisfy the hypotheses of Theorem A.10, and so the corollary follows as $\hat{Y}_j^{(n)}(t) = Y_j^{(n)}(t)$ for $0 \leq t < T_{\hat{D}}$. □

Résumé de thèse.

Le but de cette thèse est d'étudier la diffusion et les épidémies dans un graphe aléatoire avec des degrés donnés. ,Tout d'abord, nous considérons l'impact des poids sur les distances dans les graphes aléatoires dilués. Nous interprétons ces poids comme des retards, et les prenons comme des variables aléatoires exponentielles i.i.d.. Nous analysons le temps d'inondation défini comme le temps minimum nécessaire pour atteindre tous les noeuds à partir d'un noeud choisi d'une manière uniforme, et le diamètre correspondant au pire cas pour le temps d'inondation. Sous certaines conditions de régularité sur les suites de degrés du graphe aléatoire, nous montrons que ces quantités croissent comme le logarithme de n, lorsque la taille du graphe n tend vers l'infini. Nous trouvons également la valeur exacte des préfacteurs. Ce resultat nous permet d'analyser un algorithme de transmission asynchrone dans les graphes aléatoires réguliers. Nous montrons que la version asynchrone de l'algorithme est plus performante que sa version synchronisée : quand la taille du graphe est suffisament grande, il atteindra l'ensemble du réseau plus rapidement, même si le dynamique local est similaire en moyenne.

Deuxièmement, nous étudions la diffusion et les cascades dans les graphes aléatoires. Notre modèle de diffusion peut être considéré comme une variante d'un processus de croissance d'un automate cellulaire : supposons que chaque site puissent être dans l'un des deux états possibles, inactif ou actif. Les paramètres du modèle sont deux fonctions données $\theta : \mathbb{N} \to \mathbb{N}$ et $\alpha : \mathbb{N} \to [0, 1]$. Au début du processus, chaque noeud v de degré d_v devient actif avec une probabilité $\alpha(d_v)$ indépendamment des autres sommets. La présence de sommets actifs déclenche un processus de percolation : si un noeud v est actif, il reste actif pour toujours. Et s'il est inactif, il deviendra actif si au moins $\theta(d_v)$ de ses voisins sont actifs. Dans le cas où $\alpha(d) = \alpha$ et $\theta(d) = \theta$, pour chaque $d \in \mathbb{N}$, notre modèle de diffusion est équivalent à ce qui est appelé "percolation bootstrap'. Notre résultat principal est un théorème qui nous permet de trouver la proportion finale des sommets actifs dans le cas asymptotique, c'est-à-dire lorsque $n \to \infty$. Cela se fait par une analyse du processus sur le multigraphe associé au graphe aléatoire.

Enfin, nous réalisons une analyse asymptotique des cascades de défaut dans les réseaux financiers. En utilisant des méthodes analytiques, nous obtenons une expression pour la taille asymptotique d'une cascade de défaut en fonction des caractéristiques du réseau. Ce résultat est utilisé pour obtenir un critère pour la résilience d'un réseau financier aux chocs de petite taille. Nos résultats soulignent le rôle joué par les expositions contagieuses, et montrent notamment que les noeuds qui sont étroitement liés et surexposés sont ceux qui contribuent le plus à la fragilité du réseau. Ces résultats asymptotiques concordent avec des simulations faites pour les réseaux dont les tailles sont réalistes, montrant la pertinence de l'étude des réseaux de taille infinie pour la régulation macro-prudentielle.

MOTS-CLÉS : graphes aléatoires, percolation de premier passage, diamètre pondéré, transmission asynchrone, percolation bootstrap, cascades, contagion financière, risque systémique.

AMS SUBJECT CLASSIFICATIONS: 05C80, 60K35, 82B43, 60C05, 60E15, 68R10.