Ultra-Fast Rumor Spreading in Social Networks

Nikolaos Fountoulakis* Konstantinos Panagiotou† Thomas Sauerwald‡

Abstract

We analyze the popular push-pull protocol for spreading a rumor in networks. Initially, a single node knows of a rumor. In each succeeding round, every node chooses a random neighbor, and the two nodes share the rumor if one of them is already aware of it. We present the first theoretical analysis of this protocol on random graphs that have a power law degree distribution with an arbitrary exponent $\beta > 2$.

Our main findings reveal a striking dichotomy in the performance of the protocol that depends on the exponent of the power law. More specifically, we show that if $2 < \beta < 3$, then the rumor spreads to almost all nodes in $\Theta(\log \log n)$ rounds with high probability. On the other hand, if $\beta > 3$, then $\Omega(\log n)$ rounds are necessary.

We also investigate the asynchronous version of the push-pull protocol, where the nodes do not operate in rounds, but exchange information according to a Poisson process with rate 1. Surprisingly, we are able to show that, if $2 < \beta < 3$, the rumor spreads even in *constant* time, which is much smaller than the typical distance of two nodes. To the best of our knowledge, this is the first result that establishes a gap between the synchronous and the asynchronous protocol.

1 Introduction

Current estimates [1] reveal that roughly two billion people are using every day the Internet and its numerous services, like E-mail, the World Wide Web, and social networks. Especially social networks provide new and easily accessible ways for interaction and communication among individuals, thus making the Internet an ideal environment for the spread of all kinds of information. The dynamics of such information spreading processes constitute an important topic not only in computer science, but in several other disciplines, like economics and sociology, as well. In the present work, we address the fundamental question about whether and how the structure of certain models for real-world networks impacts the spread of information.

Empirical observations confirm that information disseminates very fast, especially in the occurrence of extraordinary or unexpected events, like earthquakes, plane crashes, or other emergencies originating from natural or human activities. Using extensive phone call records, Bagrow et al. [2] and Candia et al. [5] discovered that such exceptional events trigger enormous communication spikes all over the world, thus enabling a rapid propagation of the news. Modern media like *Facebook* and *Twitter* are further propelling this development.

The focus of this work is to perform a thorough analysis of the popular *phone-call* or *push-pull* protocol of Demers et al. [13] (see also Karp et al. [21]). Suppose that we are given a graph whose nodes represent individual entities, and each edge stands for some kind of interaction between them. Initially, there is a single node that knows of a rumor. The protocol then proceeds in rounds. In each such round, every node chooses a random neighbor and the two nodes share the rumor, if at least one of them is aware of it. We also study a natural asynchronous version of this protocol, see e.g. [4, 23], where each node repeatedly contacts a randomly chosen neighbor following a Poisson Process (i.e., the waiting times between two consecutive contacts are exponentially distributed with mean 1).

Let us now turn our attention to the graph models that we are going to use. There is a considerable amount of experimental research devoted to the study of properties of real-world networks. More than a decade ago, Faloutsos et al. [17] observed that the Internet exhibits a so-called *scale-free* nature: the degree sequence follows a *power law* distribution, which means that the proportion of vertices of degree k scales like $k^{-\beta}$, for all sufficiently large k, and some $\beta > 2$. This result came back then as a surprise to the networking community, and stirred significant interest in exploring the causes of this phenomenon. However, power laws have been observed in several other disciplines as well. Examples include citations in the academic literature, frequencies of words in languages, the degree sequences of several social networks, populations of cities, frequencies of names, ...,

*School of Mathematics, University of Birmingham, Edgbaston B15 2TT, United Kingdom. This author was supported by a Marie Curie Intra-European Research Fellowship PIEF-GA-2009-255115 held at the Max Planck Institute for Informatics.

†Max Planck Institute for Informatics, 66123 Saarbrücken, Germany.

‡Max Planck Institute for Informatics, 66123 Saarbrücken, Germany.

and literally hundreds of other domains; we refer the reader to the excellent surveys by Newman [24] and Mitzenmacher [22], which contain a far more exhaustive list and many references to the relevant literature. Moreover, most of the studied networks, and in particular the social and technological ones, are observed to have $2 < \beta < 3$, implying that the second-order average degree is significantly larger than the average degree.

In this paper we present the first theoretical analysis of the performance of both the synchronous and asynchronous push-pull protocol on a random graph model that has a power law degree distribution with arbitrary exponent $\beta > 2$, and which have been proposed as models for real-world networks. We are interested in the number of rounds/the time that is needed until *almost all* nodes have received the rumor. More specifically, for any $\varepsilon > 0$, we are interested in bounds that hold with high probability, and within which the protocol spreads the information to a $(1-\varepsilon)$-fraction of the nodes. This approach is somehow different from the classical settings, where typically it is required that all nodes become informed. However, this does not limit the applicability of our results: it is well-known that in real-world networks there exist "outliers", i.e., nodes which are connected only by very long paths to the remaining graph. The observed fraction of such outliers is measured to be typically at around 5 - 10% of the total number of nodes, see the recent paper [20] by Kang et al., and references therein. Hence, if we want to inform all nodes, then the time needed will unavoidably be dominated only by a small fraction of them.

Our Results The family of random graphs that we consider is asymptotically equivalent to a model described by Chung and Lu [11], who introduced it as a general purpose model for generating graphs with a power law degree sequence. Consider the vertex set $[n] := \{1, \ldots, n\}$. Every vertex $v \in [n]$ is assigned a positive weight w_v, and the edge uv is included independently in the graph with probability proportional to $w_u w_v$. Note that the expected degree of v is close to w_v. With high probability the degree sequence of the resulting graph follows a power law, provided that the sequence of weights follows a power law (see [29] for a detailed discussion). When the resulting random graph has a power law degree sequence with exponent β, we say that the sequence of weights $(w_1, \ldots, w_n)$ is *power law type* β (see Definition 2.2).

Chung and Lu [10] proved that with high probability such a graph has a *giant connected component* that contains a linear fraction of the nodes, whereas every other component only contains $O(\log n)$ nodes. Last but not least, such graphs are typically *ultra-small worlds*, i.e., the average distance of any two nodes is small, namely of order $O(\log \log n)$ [11, 14]. Let $CL(\mathbf{w}(n)) = CL(w_1, \ldots, w_n)$ be a graph drawn from this distribution. We show the following result, which establishes that the synchronous push-pull algorithm is extremely fast in the range $2 < \beta < 3$.

THEOREM 1.1. *Let $\varepsilon > 0$. Let $G = CL(w_1, \ldots, w_n)$, where the w_i's follow a power law distribution with exponent β. Assume that initially the rumor is located on a vertex of the giant component of G that is chosen uniformly at random. There is a positive integer $n_0 = n_0(\varepsilon, \beta)$ such that for any $n > n_0$, with probability at least $1-\varepsilon$, the following hold.*

(i) If $2 < \beta < 3$, then there is a $c = c(\beta) > 0$ such that after at most $c \log \log n$ rounds of the synchronous protocol all except of at most εn nodes in the largest component of G have received the rumor;

(ii) If $\beta > 3$, then the synchronous protocol needs $\Omega(\log n)$ rounds to spread the rumor to more than εn nodes.

We want to remark that the constant c depends only on β and not on ε. The above theorem is best possible for $2 < \beta < 3$. Indeed, the distance of two randomly selected vertices is with high probability at least $\frac{2}{|\log(\beta-2)|} \log \log n$; see [14]. So, a successful spread of the rumor to almost all vertices cannot be achieved in a smaller number of rounds. Moreover, a spread to all vertices is lower bounded by $\Omega(\log n)$, which is the diameter of a Chung-Lu graph for *any* $\beta > 2$; see [11]. We also show the following result for the asynchronous protocol, which says that the rumor spreads in constant time.

THEOREM 1.2. *Let $G = CL(w_1, \ldots, w_n)$, where the w_i's follow a power law distribution with exponent β, where $2 < \beta < 3$, and let $\varepsilon > 0$. There is a $T = T(\varepsilon, \beta) > 0$ and a positive integer $n_0 = n_0(\varepsilon, \beta)$ such that for any $n > n_0$, with probability at least $1-\varepsilon$, after T rounds of the asynchronous protocol all but εn nodes of the largest component of G will have received the rumor.*

Related work There is a huge amount of literature devoted to the study of several variations of the push-pull protocol on many classes of (random) graphs, and in particular, on models of social networks. We focus here only on the work that relates directly to our results. Giakkoupis [18], improving upon previous work by Chierichetti, Lattanzi, and Panconesi [8, 9], showed that the algorithm distributes a rumor to all nodes of a connected graph G in $O(\phi^{-1} \log n)$ rounds with high probability, where n is the number of nodes, and ϕ is the *conductance* of G. The conductance is a standard measure of the expansion properties of a given graph. There are also further studies that connect related expansion parameters to the runtime of rumor spreading, e.g., [4, 6, 23], however, all these upper bounds are at least logarithmic in n.

Regarding the performance of the push-pull protocol on the preferential attachment model, where $\beta = 3$, Doerr, Fouz, and Friedrich [15] showed that it disseminates the information to all nodes of the classical preferential attachment random graph in $\Theta(\log n)$ rounds. The best previously known bound was obtained by Chierichetti et al. [7], and was of order $\log^2 n$. Moreover, by considering a push-pull strategy with memory, where each node never contacts a neighbor twice in a row, the authors of [15] showed that the rumor propagates to all nodes in time $\Theta(\frac{\log n}{\log \log n})$. This was the first time that a sublogaritmic bound for the performance of a variation of the push-pull protocol was shown. Moreover, Elsässer [16] studied a variant of the push-pull protocol on Chung-Lu random graphs with a power law degree distribution and at least logarithmic minimum degree. The author shows that the algorithm completes the broadcast in $O(\log n)$ rounds and uses only $o(n \log n)$ messages.

Apart from the classical synchronous protocols, their asynchronous counterparts have also been studied [4, 19, 23]. For regular graphs, it was shown in [26] that several synchronous and asynchronous models have asymptotically the same running time. In particular, it follows from these result that for any regular graph, the running time of the synchronous push-pull protocol is asymptotically bounded from above by the running time of its asynchronous version. Our analysis demonstrates that this is not the case for the class of (non-regular) random graphs we consider.

Proof outline A central tool in our proof is the so-called *efficient connector*. An efficient connector is a vertex of bounded degree such that, once one of its neighbors has received the rumor, it pulls the rumor and pushes to all other neighbors in a very short time. This notion gives rise to an auxiliary graph H containing an edge joining two vertices x and y, if there is an efficient connector which is incident to both of them. If H has diameter D, then any rumor in H will spread to all vertices in $O(D)$ rounds. The key point in our analysis is that this graph turns out to be essentially distributed as a (smaller) Chung-Lu random graph. This allows us to use known results about typical distances in such random graphs in the proof regarding the performance of the synchronous protocol.

More importantly, if we restrict H to vertices with weight larger than w, then its density increases polynomially in w. This allows us to be much more demanding regarding the "efficiency" of a connector. In particular, when considering the asynchronous protocol, we may require that the rumor is spread to all neighbors in time $O(1/D)$. This implies that the rumor is spread over H in constant time. To pass the rumor efficiently to H, we use the property that when $2 < \beta < 3$ the neighborhood of most vertices grows superexponentially – this is not the case when $\beta > 3$. This allows us to adjust the efficiency of the connectors in such a way that the rumor reaches H from most vertices in V in constant time. An illustration of our proof is given in Figure 1.

We will begin our analysis with the asynchronous protocol. The reason is that the result for the synchronous protocol will follow from this by, roughly speaking, rounding up the "efficiency" of the connectors to an integer by means of a proper coupling.

2 Random graph models and notation

The model of the random graphs, which will serve as the underlying graph over which both protocols run, is asymptotically equivalent to a model considered by Chung and Lu [11], and is a special case of the so-called *inhomogeneous random graph*, which was introduced Söderberg [27] and was studied in great detail by Bollobás, Janson and Riordan in [3].

We begin this section with a precise definition of the considered models. In Section 2.2 we then collect the relevant results about the degree sequence and the average distances, that will be heavily used in the subsequent sections.

2.1 Chung-Lu random graphs In order to define the model we consider for any $n \in \mathbb{N}$ the vertex set $[n] = \{1, \dots, n\}$. Each vertex i is assigned a positive weight $w_i(n)$, and we will write $\mathbf{w}(n) = (w_1(n), \dots, w_n(n))$. We assume in the remainder that the weights are deterministic, and we will suppress a possible dependence on n, whenever it is obvious from the context. For any $S \subseteq [n]$, set

$$W_S(\mathbf{w}) := \sum_{i \in S} w_i.$$

In our random graph model, the event of including the edge $\{i, j\}$ in the resulting graph is independent of the events of including all other edges, and equals

$$p_{ij}(\mathbf{w}) = \min\left\{\frac{w_i w_j}{W_{[n]}(\mathbf{w})}, 1\right\}. \tag{2.1}$$

This model was considered by Chung et al., who studied in a series of papers [10–12], for fairly general choices of $\mathbf{w}$, several properties of the resulting graphs, such as the average path length or the distribution of the component sizes. We will refer to this model as the *Chung-Lu* model, and we shall write $CL(\mathbf{w})$ for a random graph in which each possible edge $\{i, j\}$ is included independently with probability as in (2.1). Moreover, we will suppress the dependence on $\mathbf{w}$, if it is clear from the context to which sequence of weights we refer to.

Note that in a Chung-Lu random graph, the weights essentially control the *expected* degrees of the vertices. Indeed, if we ignore the minimization in (2.1), and also allow a loop at vertex i, then the expected degree of that vertex is $\sum_{j=1}^{n} w_i w_j / W_{[n]} = w_i$. In the general case, a similar asymptotic statement is true, unless the weights fluctuate too much. Consequently, the choice of $\mathbf{w}$ has a significant effect on the degree sequence of the resulting graph. For example, the authors of [11] choose $w_i = d\left(\frac{\beta-2}{\beta-1}\right)(\frac{n}{i+i_0})^{1/(\beta-1)}$, which typically results in a graph with a power law degree sequence with exponent β, average degree d, and maximum degree proportional to $(n/i_0)^{1/(\beta-1)}$, where i_0 was chosen such that this expression is $O(n^{1/2})$. Our results will hold in a more general setting, where larger fluctuations around a "strict" power law are allowed, and also larger maximum degrees are possible, thus allowing a greater flexibility in the choice of the parameters.

In our analysis it will be very useful to consider a slight variation of the above model. Suppose that instead of (2.1) we create a *multi*graph by including X_{ij} edges that join i to j, where X_{ij} is a Poisson random variable with mean $w_i w_j / W_{[n]}$. In other words, the probability p'_{ij} that i and j are connected by (at least) one edge is given by the relation

$$1 - p'_{ij} = e^{-w_i w_j / W_{[n]}}. \tag{2.2}$$

This model was studied in [25] by Norros and Reittu, who showed that it behaves similar to the Chung-Lu model in terms of the distance between two randomly chosen vertices. We shall denote a random graph drawn from this distribution by $NR(\mathbf{w})$.

2.2 Power law weight distributions Following van der Hofstad [29], let us write for any $n \in \mathbb{N}$ and any sequence $\mathbf{w} = (w_1(n), \ldots, w_n(n))$ of weights

$$F_n(x) = n^{-1} \sum_{i=1}^{n} \mathbf{1}[w_i(n) \leqslant x],$$

where $\mathbf{1}[w_i(n) \leqslant x]$ is the indicator function which is equal to 1 if and only if $w_i(n) \leqslant x$. This is the empirical distribution function of the weight of a randomly chosen vertex. We will assume that F_n satisfies the following two conditions.

DEFINITION 2.1. *We say that $(F_n)_{n\geqslant 1}$ is* regular*, if it has the following two properties.*

- **[Weak convergence of weight]** *There is a distribution function $F : [0, \infty) \to [0, 1]$ such that for all x at which F is continuous $\lim_{n\to\infty} F_n(x) = F(x)$.*
- **[Convergence of average weight]** *Let W_n be a random variable with distribution function F_n, and let W_F be a random variable with distribution function F. Then $\lim_{n\to\infty} \mathbb{E}[W_n] = \mathbb{E}[W_F]$.*

The regularity of $(F_n)_{n\geqslant 1}$ guarantees two important properties. First, the weight of a random vertex is close to the value of a given random variable. Moreover, this variable has finite mean. Thus, as will be made more precise in a few paragraphs, the resulting graph has bounded average degree. Apart from regularity, our focus will be on weight sequences that give rise to power-law degree distributions.

DEFINITION 2.2. *We say that a regular sequence $(F_n)_{n\geqslant 1}$ is* of power law type β, *if there are constants $0 < \gamma_1 < \gamma_2$ and $x_0 > 0$ as well as a function $(\log\log n)^{-1} \ll \alpha(n) \leqslant \frac{1}{\beta-1}$ such that for all $x_0 \leqslant x \leqslant n^{\alpha(n)}$*

$$\gamma_1 x^{-\beta+1} \leqslant 1 - F_n(x) \leqslant \gamma_2 x^{-\beta+1},$$

and $F_n(x) = 1$, for $x \leqslant x_0$, whereas $F_n(x) = 0$, for $x > n^{\alpha(n)}$. (Here and elsewhere the notation $\ll$ means "asymptotically smaller".)

3 Analysis of the protocols

In this section we perform our analysis for Chung-Lu random graphs $CL(\mathbf{w})$. Recall that an illustration of this analysis is given in Figure 1. We will assume in the remainder that the empirical distribution $(F_n)_{n\geqslant 1}$ of $(\mathbf{w}(n))_{n\geqslant 1}$ is of power law type β, cf. Definition 2.2, where $2 < \beta < 3$. We start with the analysis of the asynchronous protocol. All lemmas whose proof is omitted here are proved in Section 6.

3.1 The asynchronous protocol – Proof of Theorem 1.2 To avoid ambiguities, we denote by Ω_G the probability space for the graph and by Ω_{RS} the probability space for the rumor spreading protocol. For the analysis, we will use the following equivalent description of the asynchronous protocol. For every vertex $u \in V$, we have an independent Poisson Processes $\mathcal{PP}(u)$ over the non-negative reals [28], and an infinite list $(L_t(u))_{t\in\mathbb{N}}$ of randomly chosen neighbors. If the vertex u or a neighbor of u is informed at time 0,

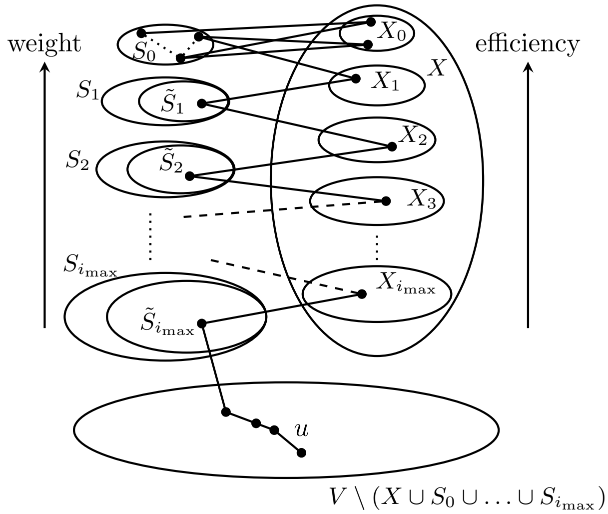

Figure 1: Illustration of the way a rumor initiated at u is conveyed to the kernel S_0. The weights of the sets $S_{i_{\max}}, S_{i_{\max}-1}, \ldots, S_0$ are super-exponentially increasing, starting with a (large) constant weight in $S_{i_{\max}}$. The sets $X_{i_{\max}}, X_{i_{\max}-1}, \ldots, X_0 \subseteq X$ are all vertices of constant weight and degree, but their efficiency increases exponentially (starting with a constant efficiency in $X_{i_{\max}}$). Every vertex in $\tilde{S}_i$ is connected to a vertex in $\tilde{S}_{i-1}$ via a vertex in X_i. The dotted edges within S_0 represent the edges of the graph H.

then u follows $\mathcal{PP}(u)$, i.e., when the process $\mathcal{PP}(u)$ ticks for the k-th time, it exchanges a rumor with neighbor $L_k(u)$ (if any of the two nodes knows the rumor). However, if u and its neighbors are uninformed, then we may assume that u does nothing (note that in the original definition of the process, u would communicate with randomly chosen neighbors; this has no effect, since the waiting times are exponentially distributed and thus memoryless). Now suppose that at time $\tau(u) \in \mathbb{R}$, the first neighbor of u becomes informed. Then u starts to follow the Poisson process $\mathcal{PP}(u)$, i.e., when the process $\mathcal{PP}(u)$ ticks for the k-th time at time t_k, u communicates with neighbor $L_k(u)$ at time $\tau(u) + t_k$. Here we implicitly assume that there is some oracle that tells u at the right time to start using the Poisson process. In other words, we make a "thought experiment" about the behavior of the vertices, which, however, does not affect the execution of the protocol at all.

In the following, we now describe a way how every node u generates the list $(L_k(u))_{k \in \mathbb{N}}$. For the k-th element, node u generates a uniformly random number in $U_k(u) \in [0, 1]$. Hence, each node u generates an infinite list of uniform random numbers $(U_k(u))_{k \in \mathbb{N}}$ in $[0, 1]$ and we may assume that the node does so before the graph is exposed. Now if we expose the neighborhood of u with neighbors labeled and ordered from 1 to $\deg(u)$, then u generates the list $(L_k(u))_{k \in \mathbb{N}}$ where for each $k \in \mathbb{N}$

$$L_k(u) \text{ is set to be that } j \in [\deg(u)] \text{ such that } U_k(u) \in \left[\frac{j-1}{\deg(u)}, \frac{j}{\deg(u)}\right).$$

This two-stage procedure for generating the lists turns out to be handy, as we may say something about how fast the rumor will be spread by u to its neighbors just by looking at the sequence $(U_k(u))_{k \geqslant 1}$ without exposing any edge in G.

Definition 3.1. *Let d be any integer and $\rho \in \mathbb{R}$. Then we call a node $u \in V$ a (ρ, d)-efficient asynchronous connector, if the following two conditions hold:*

- *For all integers $1 \leqslant k \leqslant 2d$, $U_k(u) \in \left[\frac{k-1}{2d}, \frac{k}{2d}\right)$ and for all integers $2d < k \leqslant 4d$, $U_k(u) \in \left[\frac{k-2d-1}{2d}, \frac{k-2d}{2d}\right)$.*
- *The clock of $\mathcal{PP}(u)$ ticks at least $4d$ times within the interval $[0, \rho)$.*

Let us now explain this definition. Suppose that u is a (ρ, d)-efficient asynchronous connector, and additionally, the degree of u is not larger than d. Then if at time $\tau(u)$ the first neighbor of u becomes informed, then all neighbors of u and u itself are informed by time $\tau(u) + \rho$. Note that Definition 3.1 uses only the probability space Ω_{RS}. We use the following simple lower bounds on the probability that a vertex u is an efficient connector.

Lemma 3.1. *Let $\rho \in \mathbb{R}_+$ and $d \in \mathbb{N}$. Then,*

$$\Pr[\, u \text{ is a } (\rho, d)\text{-efficient asynchronous connector}] \geqslant \left(\frac{\rho}{16d^2}\right)^{4d}.$$

Having defined the notion of an efficient connector, we are now ready to formulate the general proof strategy. Additionally, with some small modifications, we can also obtain the result for the synchronous algorithm (see Section 3.2).

First, we define a *kernel* which contains all nodes with sufficiently large weight. More precisely, we set $T_0 := (\log n)^A$, where $A = 3(3-\beta)^{-2}$. Then the *kernel* is the set of vertices with weight at least T_0, i.e.,

$$S_0 := \{v \in V : w_v \geqslant T_0\}.$$

Also, let X be the set of εn nodes with smallest weight, where we assume that $\varepsilon \in (0, \frac{1}{2})$. Note that for any x in X

$$\gamma_1^{1/(\beta-1)} \leqslant w_x \leqslant (2\gamma_2)^{1/(\beta-1)}. \tag{3.3}$$

Indeed, the assumption that F_n is of power law type β (cf. Definition 2.2) implies that $F_n(\gamma_1^{1/(\beta-1)}) \leqslant 1 - \gamma_1(\gamma_1^{1/(\beta-1)})^{-\beta+1} = 0$. This shows the lower bound on w_x. To see the upper bound, a similar calculation shows that $F_n\left((2\gamma_2)^{1/(\beta-1)}\right) \geqslant \frac{1}{2} \geqslant \varepsilon$. We shall also write $w_{\min} = \gamma_1^{1/(\beta-1)}$. Our first goal is to show that there are many efficient connectors that have two neighbors in S_0. To this end, we prove the following lemma.

Lemma 3.2. *Let $\varepsilon \in (0, \frac{1}{2})$. Set $\alpha = 2(\beta-2)/(\beta-1) \in (0,1)$ and $\lambda = \lceil \log_{1/\alpha} \log(T_0) \rceil$. Then there is a $c = c(\alpha, \varepsilon) > 0$ so that with probability $1 - o(1)$, there exist disjoint sets $X_0, X_1, \ldots, X_\lambda \subseteq X$ with the following properties:*

- *each node in X_i, $0 \leqslant i \leqslant \lambda$, is not adjacent to any other node in X,*
- *for each $0 \leqslant i \leqslant \lambda$, $|X_i| \geqslant c \cdot \alpha^{8(\lambda-i)} \cdot n$,*
- *for each $u \in X_i$, $0 \leqslant i \leqslant \lambda$, u is an $(\alpha^{\lambda-i}, 2)$-efficient asynchronous connector.*

For the proof that the rumor is quickly disseminated within the kernel, we only use the connectors in X_0 that have degree 2 (we shall use the other connectors in $X_1, \ldots, X_\lambda$ later in the analysis). We are going to prove that there are sufficiently many efficient connectors in the set X_0 that connect two vertices in S_0. To this end, we define an auxiliary graph $H = (V_H, E_H)$ with $V_H = S_0$ and

$$\begin{aligned} E_H := \{\{x, y\} \mid\ & x, y \in V_H\ \exists v \in X_0 \colon \\ & v \text{ is a } ((\log T_0)^{-1}, 2)\text{-e. c. with } N(v) = \{x, y\}\}, \end{aligned}$$

where *e. c.* stands for efficient asynchronous connector. The next lemma proves that the diameter of H is $O(\log\log n)$. Thus, because of the fact that $(\log T_0)^{-1} = \Theta((\log\log n)^{-1})$, as soon as any one of the nodes in V_H learns the rumor, then after an additional time of $O(1)$ all other nodes will know it as well.

Lemma 3.3. *With probability at least $1 - o(1)$, all pairs of vertices in H are connected by a path of length at most $\frac{2+o(1)}{|\log(\beta-2)|} \log\log n$.*

It remains to prove the existence of efficient paths from most vertices $V \setminus S_0$ to the set S_0. To this end, we partition $V(G)$ according to the weight of the vertices. First, we define a sequence of thresholds $(T_i)_{i \geqslant 1}$. Recall the definition of S_0, where we set $T_0 = (\log n)^A$. Let $T_i := T_0^{\alpha^i}$, where $\alpha = 2(\beta-2)/(\beta-1) \in (0,1)$. Moreover, for any integer i with $1 \leqslant i \leqslant \lceil \log_{1/\alpha} \log T_0 \rceil =: \lambda$, we set

$$S_i = \{v \in V \colon T_i \leqslant w_v < T_{i-1}\},$$

where $T_{-1} := \infty$. The task of connecting most vertices in V to S_0 is performed by the following two lemmas.

Lemma 3.4. *Let $\varepsilon \in (0, \frac{1}{2})$ and set $\alpha = 2(\beta-2)/(\beta-1) \in (0,1)$. Then there exists $\kappa = \kappa(\varepsilon) \geqslant 0$ so that with probability at least $1 - \varepsilon$, there exists a sequence of subsets $\tilde{S}_i \subseteq S_i$, $1 \leqslant i \leqslant i_{\max} = \lambda - \kappa$, that satisfies the following properties:*

- *$|\tilde{S}_i| \geqslant |S_i|/2$,*
- *each vertex $u \in \tilde{S}_i$ is connected to $\tilde{S}_{i-1}$ by a vertex $x \in X_i$, which is an $(\alpha^{\lambda-i}, 2)$-efficient asynchronous connector and satisfies $\deg(x) = 2$, where $\tilde{S}_0 = S_0 = V_H$.*

The final step is to prove that there is an efficient path from most vertices in the giant component to the set $\cup_{i=0}^{i_{\max}} \tilde{S}_i$ from Lemma 3.4, which is the subject of the next lemma.

Lemma 3.5. *Let $\varepsilon \in (0, \frac{1}{2})$ and let $\tilde{S} = \cup_{i=0}^{i_{\max}} \tilde{S}_i$ be sequence of sets from Lemma 3.4. Then there are positive integers $r = r(\varepsilon)$ and $\Delta = \Delta(\varepsilon)$ so that with probability at least $1 - \varepsilon$, all but an ε fraction of the vertices of the giant component of $CL(\mathbf{w})$ are such that for each of these vertices u there is a path $P = (u_1 = u, u_2, \ldots, u_r)$ with $u_r \in \tilde{S}$ which does not use vertices in X, and the degrees of all vertices on that path except for u_r are upper bounded by Δ.*

Proof of Theorem 1.2 Let us now prove the upper bound of $O(1)$ on the rumor spreading time. By combining Lemma 3.3, Lemma 3.4 and Lemma 3.5, we conclude that there is, for all pairs u, v in the giant component except for at most εn^2 pairs, a path $P_{u,v} = (u_1 = u, u_2, \ldots, u_\ell = v)$ with the following properties:

- $\ell = \tilde{c} \log\log n$, where $\tilde{c} > 0$ is some constant;
- there are constants $c_1, c_2 \leqslant r$, such that all the degrees of $u_1, \ldots, u_{c_1}$ and the degrees of $u_{\ell-c_2}, \ldots, u_\ell$ are bounded by some constant $\Delta > 0$;
- for the subpath $\tilde{P}_{u,v} = (u_{c_1+1}, \ldots, u_{\ell-c_2-1})$, there is an index set $\mathcal{I} \subseteq \{c_1+1, \ldots, \ell-c_2-1\}$ associated with values Δ_i for each $i \in \mathcal{I}$, so that
  - $\mathcal{I}$ contains at least every second vertex on $\tilde{P}_{u,v}$;
  - for every $i \in \mathcal{I}$, u_i is a $(\Delta_i, 2)$-efficient asynchronous connector and $\deg(u_i) = 2$;
  - $\sum_{i \in \mathcal{I}} \Delta_i = O(1)$.

Note that once the rumor reaches the vertex u_{c_1+1}, we know that it will reach the vertex $u_{\ell-c_2-1}$ in time

$\sum_{i\in\mathcal{I}}\Delta_i = O(1)$, using the definition of $(\Delta_i, 2)$-efficient asynchronous connector along with $\deg(u_i) = 2$ for $i \in \mathcal{I}$. Hence, it only remains to bound the time for the rumor to go from $u_1 = u$ to u_{c_1}, and from $u_{\ell-c_2}$ to $u_\ell = v$. The expected time for w to transmit the rumor to its neighbor on $P_{u,v}$ can upper bounded by Δ. Hence, the expected time for the rumor to traverse the whole path $P_{u,v}$ is at most

$$\Delta \cdot (c_1 + c_2) + \sum_{i\in\mathcal{I}} \Delta_i,$$

which is a constant. Using Markov's inequality completes the proof.

3.2 The synchronous protocol – Proof of Theorem 1.1(i)

We first adjust the definition of efficient connector to the synchronous model.

Definition 3.2. *Let d and $\rho \geqslant 4d$ be integers. We call a node $u \in V$ a (ρ, d)-efficient synchronous connector if the following condition holds:*

- *for each $1 \leqslant j \leqslant d$, we have*

$$\bigcup_{t=1}^{\lfloor\rho/2\rfloor} U_t(u) \cap \left[\frac{j-1}{2d}, \frac{j}{2d}\right) \neq \emptyset$$
$$\text{and} \bigcup_{t=\lfloor\rho/2\rfloor+1}^{\rho} U_t(u) \cap \left[\frac{j-1}{2d}, \frac{j}{2d}\right) \neq \emptyset.$$

Hence, if a node u is a (ρ, d)-efficient connector and its degree is at most d, then once the first neighbor of u becomes informed at step $\tau(u)$, then all other neighbors of u (and u itself) will have been informed by step $\tau(u) + \rho$. Let us also observe that, for any d, we can make the probability for a node u to be a (ρ, d)-efficient synchronous connector arbitrarily close to 1 by choosing a sufficiently large $\rho = \rho(d)$.

Recall that all vertices in X that are used in Lemma 3.3 and Lemma 3.4 as connectors in $\tilde{S}$ and H, are $(1, d)$-efficient asynchronous connectors. Moreover, the probability for being a $(1, d)$-efficient asynchronous connector is smaller than the probability for being a (ρ, d)-efficient synchronous connector, if we choose $\rho > 1$ to be a sufficiently large constant. Hence we can set up a coupling between the asynchronous and the synchronous model such that whenever a vertex $u \in V$ is a $(1, d)$-efficient asynchronous connector, then it will be also a (ρ, d)-efficient synchronous connector. Hence, corresponding to Section 3.1, we have the following. There is, for all pairs u, v in the giant component except for at most εn^2 pairs, a path $P_{u,v} = (u_1 = u, u_2, \ldots, u_\ell)$ with the following properties:

- $\ell = \tilde{c} \log\log n + 2f(\varepsilon)$, where $\tilde{c} > 0$ is a constant and $f(\varepsilon)$ is a function that only depends on ε,
- there are $c_1, c_2 \leqslant f(\varepsilon)$, such that all the degrees of $u_1, \ldots, u_{c_1}$ and the degrees of $u_{\ell-c_2}, \ldots, u_\ell$ are bounded by some constant $C = C(\varepsilon) > 0$,
- for the subpath $\tilde{P}_{u,v} = (u_{c_1+1}, \ldots, u_{\ell-c_2-1})$, there is an index set $\mathcal{I} \subseteq \{c_1 + 1, \ldots, \ell - c_2 - 1\}$ associated with values Δ_i for each $i \in \mathcal{I}$, so that
  - $\mathcal{I}$ contains at least every second vertex on $\tilde{P}_{u,v}$,
  - for every $i \in \mathcal{I}$, u_i is a $(\rho, 2)$-efficient synchronous connector and $\deg(u_i) = 2$,

Using exactly the same arguments as in Section 3.1, we obtain that the expected time for the rumor to reach v from u is at most

$$(\tilde{c} \log\log n + 2f(\varepsilon)) \cdot \rho + 2f(\varepsilon) \cdot C.$$

This completes the proof of Theorem 1.1(i).

4 Typical properties of Chung-Lu random graphs

The next paragraphs discuss the most relevant properties of the Chung-Lu model, namely its typical degree sequence and the average distance of two randomly chosen vertices.

The degree sequence Let us first turn our attention to the degree sequence of a Chung-Lu random graph. Note that the weight of a random vertex is given by W_n, which is a random variable with distribution F_n. Moreover, a simple calculation shows that the degree of the vertex with weight w_i follows approximately a Poisson distribution with mean w_i. Consequently, we expect that the probability that a random vertex has degree k is close to $e^{-W_n} W_n^k / k!$. The next theorem, which is taken from [3, 29], confirms this intuition and characterizes the degree distribution of a random graph $CL(\mathbf{w}(n))$ for large n.

Theorem 4.1. *Suppose that $(F_n)_{n\geqslant 1}$ is of power law type β, for some $\beta > 2$. Then, for any $k \geqslant 0$, with high probability,*

$$N_k = (1 + o(1)) p_k n, \quad \text{where} \quad p_k = \mathbb{E}\left[e^{-W_F} \frac{W_F^k}{k!}\right].$$

Moreover, there are constants $0 < \gamma_1' < \gamma_2'$ such that $\gamma_1' k^{-\beta+1} \leqslant \sum_{i\geqslant k} p_i \leqslant \gamma_2' k^{-\beta+1}$.

In other words, if $(F_n)_{n\geqslant 1}$ is of power law type β, then also the resulting random graph has a power law degree distribution with exponent β. We want to remark that a similar statement is true for the Norros-Reittu $NR(\mathbf{w})$ model; we omit the details.

Typical distances Perhaps the most crucial parameter in our proofs about the Chung-Lu model is the typical distance of two random nodes. The following theorem follows straightforwardly from the results in [11, 14, 29], and covers the case $2 < \beta < 3$.

THEOREM 4.2. *Suppose that $(F_n)_{n\geqslant 1}$ is of power law type β, for some $2 < \beta < 3$. Let D_n be the distance of two randomly chosen vertices in $CL(\mathbf{w}(n))$. Then, with high probability, if $D_n < \infty$,*

$$D_n = \left(\frac{2}{|\log(\beta-2)|} + o(1)\right) \log\log n.$$

In other words, if the parameters are chosen such that the resulting graph has a power law degree distribution with exponent in $(2,3)$, then the resulting graphs are *ultra-small*, in the sense that most pairs of nodes are at distance $O(\log\log n)$. The next result, also taken from [11, 29], states that the situation is dramatically different if the exponent is larger.

THEOREM 4.3. *Suppose that $(F_n)_{n\geqslant 1}$ is of power law type β, for some $\beta > 3$. Let D_n be the distance of two randomly chosen modes in $CL(\mathbf{w}(n))$. Then, for sufficiently large n, with probability at least $1 - n^{-1/2}$, if $D_n < \infty$,*

$$D_n \geqslant \frac{1}{3}\log_\nu n, \quad \textit{where} \quad \nu = \frac{\mathbb{E}\left[W_F^2\right]}{\mathbb{E}\left[W_F\right]^2} > 1.$$

The same results are again also true for the Norros-Reittu $NR(\mathbf{w})$ model.

The following lemma bounds the total weight of certain subsets of vertices, in the case where the weights follow a power law of type $\beta \in (2,3)$ – it will be used many times in our proofs.

LEMMA 4.1. *Let $(F_n)_{n\geqslant 1}$ be of power law type $2 < \beta < 3$. Then, uniformly for $0 < x \leqslant \frac{1}{2}n^{\alpha(n)}$*

$$\sum_{v\,:\,w_v\leqslant x} w_v = W_{[n]} - \Theta(nx^{-\beta+2})$$

$$\textit{and for } x \leqslant n^{\alpha(n)} \qquad \sum_{v\,:\,w_v\leqslant x} w_v^2 = \Theta(nx^{3-\beta}).$$

Proof. We begin with the basic relation

$$(4.4) \qquad \begin{aligned} \frac{1}{n}\sum_{v\,:\,w_v\leqslant x} w_v &= \int_0^x \Pr[y \leqslant W_n \leqslant x]dy \\ &= \int_0^x (1-F_n(y)) - (1-F_n(x))dy. \end{aligned}$$

To estimate the above expression, note that the assumption that $(F_n)_{n\geqslant 1}$ is of power law type $\beta > 2$ implies

$$\int_x^{n^{\alpha(n)}} (1-F_n(y))dy = \Theta(1)\cdot\int_x^\infty y^{-\beta+1}dy = \Theta(x^{-\beta+2}).$$

By putting (4.4) and the above estimate together we obtain that

$$\begin{aligned} \sum_{v\,:\,w_v\leqslant x} w_v &= n\mathbb{E}\left[W_n\right] - n\Theta(x^{-\beta+2}) - x(1-F_n(x)) \\ &= n\mathbb{E}\left[W_n\right] - n\Theta(x^{-\beta+2}). \end{aligned}$$

This proves the first claim. To see the second claim, observe that

$$\sum_{v\,:\,w_v\leqslant x} w_v^2 = 2n\int_0^x y\Pr[y \leqslant W_n \leqslant x]dy.$$

A similar computation as above gives then the desired statement – the details are omitted.

5 Lower bounds – Proof of Theorem 1.1(ii)

LEMMA 5.1. *Consider the Chung-Lu random graph $CL(\mathbf{w}(n))$, where the empirical distribution function $F_n(x)$ is of power law type $\beta > 3$. Then there is a constant $\gamma > 0$ such that after $\gamma\log n$ rounds the number of nodes informed by the synchronous push-pull algorithm is $o(n)$ with probability that is asymptotically bounded away from 0.*

Proof. The lemma follows from Theorem 4.3. To see this, note that the conclusion of Theorem 4.3 implies that there are at most $n^{3/2}$ pairs of nodes in $CL(\mathbf{w})$ with distance at most $\frac{1}{3}\log_\nu n$. This however, guarantees that there are at most, say, $n^{3/4}$ nodes such that the number of nodes within distance $\frac{1}{3}\log_\nu n$ is at least $n^{3/4}$. This shows the claimed statement.

6 Completing the proof of Theorems 1.1 and 1.2

In this final section we collect all proofs omitted in Section 3.

6.1 Proof of Lemma 3.1

Recall that the definition of efficient asynchronous connector consists of two conditions, which are by construction independent events. Let us start with a lower bound for the probability for the first condition. Since for the kth clock tick, node u generates a uniform random variable $U_k(u) \in [0,1)$, we find that the probability that u satisfies the first condition is at least $(2d)^{-4d}$. For the second condition, let us estimate the probability that within the time-interval $[0, 0+\rho)$, the Poisson clock ticks at least $4d$ times. Since the waiting time between two ticks is exponentially distributed with mean 1, we can lower bound this probability by

$$(6.5) \qquad \left(1 - e^{-\tau/(4d)}\right)^{4d} \geqslant \left(\frac{\tau}{8d}\right)^{4d},$$

since $1-e^{-x} \geqslant x/2$ for $x \in (0,1)$. Combining the two lower bounds, we conclude that the probability for u being an efficient (ρ, d) asynchronous connector is at least

$$\left(\frac{\tau}{8d}\right)^{4d} \cdot \left(\frac{1}{2d}\right)^{4d} = \left(\frac{\tau}{16d^2}\right)^{4d}.$$

6.2 Proof of Lemma 3.2 We start with the first condition which says that there may not be any internal edge in $\cup_{i=1}^{\lambda} X_i$. To this end, consider all nodes in X. For a node $x \in X$, let $Z_x = 1$ if x has no edge in X, and 0 otherwise. Note that for sufficiently large n

$$\Pr[Z_x = 1] = \prod_{x' \in X\setminus\{x\}} \left(1 - \frac{w_x w_{x'}}{W_{[n]}}\right) \overset{(w_x w_{x'} \text{ bounded})}{\geqslant} e^{-2w_x}.$$

Let $Z := \sum_{x\in X} Z_x$. To apply the second moment method, first note that

$$\mathrm{Var}[Z] = \sum_{x\in X} \mathrm{Var}[Z_x] + \sum_{x\neq x'} \Pr[Z_x = 1 \wedge Z_{x'} = 1] - \Pr[Z_x = 1]\Pr[Z_{x'} = 1].$$

Note that each term in the sum above equals

$$\Pr[Z_x = 1]\left(\Pr[Z_{x'} = 1 \mid Z_x = 1] - \Pr[Z_{x'} = 1]\right).$$

The last term can be estimated by

$$\prod_{y\in X\setminus\{x',x\}} \left(1 - \frac{w_y w_{x'}}{W_{[n]}}\right) - \prod_{y\in X\setminus\{x'\}} \left(1 - \frac{w_y w_{x'}}{W_{[n]}}\right) \left(\frac{w_x w_{x'}}{W_{[n]}}\right) \cdot \prod_{y\in X\setminus\{x',x\}} \left(1 - \frac{w_y w_{x'}}{W_{[n]}}\right).$$

Since w_x and $w_{x'}$ are bounded, by putting everything together we infer that

$$\mathrm{Var}[Z] \leqslant \mathbb{E}[Z] + O(1) \cdot \frac{|X|^2}{W_{[n]}} = O(\mathbb{E}[Z]).$$

Using Chebyshev's inequality we infer for sufficiently large n with room to spare that

$$\Pr[|Z - \mathbb{E}[Z]| \geqslant \mathbb{E}[Z]/2] \leqslant n^{-1/2},$$

Set $\tilde{X} = \{x \in X : Z_x = 1\}$ and let us now address the second and third condition of the lemma. Define the set $X_{\geqslant i} := \{x \in \tilde{X} : x \text{ is a } (\alpha^{\lambda-i}, d)\text{-efficient connector}\}$, where $1 \leqslant i \leqslant \lambda$. By Lemma 3.1, where we set $d = 2$,

$$\mathbb{E}[X_{\geqslant i}] \geqslant |\tilde{X}| \cdot \left(\frac{\alpha^{\lambda-i}}{64}\right)^8.$$

Since $i \leqslant \lambda = O(\log\log\log n)$, the expectation above is bounded from below by some polynomial in n. By applying the Chernoff bounds, it follows that

$$\Pr\left[X_{\geqslant i} \geqslant \frac{1}{2}|\tilde{X}| \cdot \left(\frac{\alpha^{\lambda-i}}{64}\right)^8\right] \geqslant 1 - n^{-\omega(1)}.$$

Taking the union bound, this holds for all $1 \leqslant i \leqslant \lambda$ with probability at least $1 - n^{-\omega(1)}$. To simplify notation, let us write for brevity $c' = (1/2)\cdot(64)^{-8}$ and $\sigma = \alpha^8 < 1$.

We now define the sets X_i recursively. We begin with X_0 and put $c'n \cdot (1-\sigma)\cdot\sigma^{\lambda}$ elements of $X_{\geqslant 0}$ into the set X_0. Similarly, we put $c'n\cdot(1-\sigma)\cdot\sigma^{\lambda-1}$ elements of $X_{\geqslant 1}\setminus X_0$ into the set X_1 and so on. We now prove by induction that this recursive definition works. So assume that we have completed the construction of $X_0, X_1, \ldots, X_k$ where $|X_i| = c'n\cdot(1-\sigma)\cdot\sigma^{\lambda-i}$ for every $0 \leqslant i \leqslant k$. Then,

$$\begin{aligned} |X_{\geqslant k}| - \sum_{i=0}^{k-1} |X_i| &\geqslant c'n\cdot\sigma^{\lambda-k} - \sum_{i=0}^{k-1} c'n\cdot(1-\sigma)\cdot\sigma^{\lambda-i} \\ &\geqslant c'n\cdot\sigma^{\lambda-k} - c'n\cdot(1-\sigma)\cdot\sigma^{\lambda-(k-1)}\sum_{i=0}^{\infty}\sigma^i \\ &= c'n\cdot\sigma^{\lambda-k}\cdot(1-\sigma). \end{aligned}$$

Hence we can put $c'n\cdot\sigma^{\lambda-k}\cdot(1-\sigma)$ elements from the set $X_{\geqslant k}\setminus\cup_{i=0}^{k-1}X_i$ to the set X_k. This completes the induction and the proof of the lemma.

6.3 Proof of Lemma 3.3 Our first crucial step towards the proof of Lemma 3.3 is the following statement, which asserts that H contains a Norros-Reittu graph, where all edges appear independently.

LEMMA 6.1. *Assume that* $(\mathbf{w}'(u))_{u\in V_H} = (T_0^{-\beta+2}(\log T_0)^{-9}\, w_u)_{u\in V_H}$. *Then there is a coupling such that with high probability* $NR(\mathbf{w}') \subseteq H$.

Proof. Let $v \in X_0$. By our construction, see Lemma 3.2, v is a $((\log T_0)^{-1}, 2)$-efficient connector, and has no neighbors in X. Moreover, $|X_0| = \Omega(n(\log T_0)^{-8}) = \Omega(n(\log\log n)^{-8})$. Finally, the choice of X_0 is guaranteed to be independent of all edges that have at most one endpoint in X.

Before we construct the coupling, our first objective is to estimate the probable number of edges in H. Let us begin with bounding from below the probability p_v that v has exactly two neighbors in V_H, and no other neighbor in $[n]\setminus X$. Let $x, x', y, y' \in V_H$, not necessarily all distinct. Note that the events that $N(v) = \{x, y\}$ and $N(v) = \{x', y'\}$ are disjoint if $\{x, y\} \neq \{x', y'\}$.

Hence

(6.6)
$$p_v = \sum_{\{x,y\}\in\binom{V_H}{2}} p_v(x,y), \text{ where}$$
$$p_v(x,y) = \Pr[N(v)\cap V_H = \{x,y\} \wedge N(v)\cap([n]\setminus(V_H\cup X)) = \emptyset].$$

Note that $p_v(x,y)$ is the probability that the neighbors of v are x, y. To estimate the above expression, first note that for any $u \in [n]$ we have that $w_v w_u = o(n)$ and therefore for sufficiently large n we may infer that

$$\Pr[N(v)\cap([n]\setminus(V_H\cup X)) = \emptyset] \geqslant \prod_{u\in[n]}\left(1-\frac{w_v w_u}{W_{[n]}}\right) \geqslant e^{-2w_v}. \tag{6.7}$$

As $w_v \leqslant 2\gamma_2 = \Theta(1)$, we infer that there is a $c_1 > 0$ such that for all v

$$\Pr[N(v)\cap([n]\setminus X) = \emptyset] \geqslant c_1. \tag{6.8}$$

We next estimate $\Pr[N(v)\cap V_H = \{x,y\}]$. As different pairs appear independently as edges, we have that

$$\begin{aligned}\Pr[N(v)\cap V_H = \{x,y\}] &= \frac{w_x w_y w_v^2}{W_{[n]}^2}\prod_{z\in V_H, z\neq x,y}\left(1-\frac{w_v w_z}{W_{[n]}}\right)\\ &\geqslant \frac{w_x w_y w_v^2}{W_{[n]}^2}\cdot\left(1-\frac{w_v}{W_{[n]}}\sum_{z\in V_H} w_z\right).\end{aligned}$$

Lemma 4.1, applied to the last sum, implies that its value is in $\Theta(nT_0^{-\beta+2})$. So, the expression in the brackets is $1-o(1) \geqslant \frac{1}{2}$, whenever n is large enough. By plugging this together with (6.8) into (6.6) we infer that

$$\begin{aligned}&p_v(x,y) \geqslant \frac{1}{2}c_1\, w_v^2\frac{w_x w_y}{W_{[n]}^2}\\ &\Longrightarrow\ p_v \geqslant \frac{1}{4}c_1\,\frac{w_v^2}{W_{[n]}^2}\left(W_{V_H}^2 - \sum_{u\in V_H} w_u^2\right).\end{aligned} \tag{6.9}$$

By applying Lemma 4.1 twice and using the fact that the maximum weight of a vertex in the kernel is $\leqslant n^{\alpha(n)}$, where $\alpha(n) \leqslant \frac{1}{\beta-1}$, we obtain the bounds

$$\begin{aligned}&W_{V_H}^2 = \Theta(n^2T_0^{-2\beta+4})\\ &\quad\text{and}\\ &\sum_{u\in V_H} w_u^2 = O(n\cdot n^{(3-\beta)/(\beta-1)}) = o(n^2T_0^{-2\beta+4}).\end{aligned} \tag{6.10}$$

Hence, $W_{V_H}^2 - \sum_{u\in V_H} w_u^2 = (1-o(1))W_{V_H}^2$, and therefore, if n is large enough there is a $c_2 > 0$ such that

$$p_v \geqslant c_2\, w_v^2\,\frac{W_{V_H}^2}{W_{[n]}^2}. \tag{6.11}$$

As it will be needed later, we also can obtain similarly an upper bound on p_v. To achieve this, we estimate the probability for the event "$N(v)\cap([n]\setminus X) = \emptyset$" from above by 1:

$$\begin{aligned}p_v &\leqslant \sum_{\{x,y\}\in\binom{V_H}{2}}\Pr[N(v)\cap V_H = \{x,y\}]\\ &\leqslant w_v^2\,\frac{W_{V_H}^2 - \sum_{u\in V_H} w_u^2}{W_{[n]}^2}.\end{aligned} \tag{6.12}$$

With the estimates (6.11) and (6.10) at hand, we can obtain a lower bound for the number of edges e_H in H that holds with high probability. Indeed, let $m = \sum_{v\in X_0} p_v$. By using (6.10) we infer that there is a constant $c_3 > 0$ such that

$$\begin{aligned}\mathbb{E}\,[e_H] = m &\geqslant c_2\frac{W_{V_H}^2}{W_{[n]}^2}\cdot\sum_{v\in X_0} w_v^2\\ &\overset{(w_v=\Theta(1))}{\geqslant} c_3\,T_0^{-2\beta+4}\,|X_0|.\end{aligned}$$

We shall now show that e_H is concentrated around its mean by applying the second moment method. If we denote by p_{uv} the probability that $u, v \in X_0$ have exactly two neighbors in V_H and no neighbors otherwise, then, as in (6.6)

$$\begin{aligned}p_{uv} = &\sum_{\{x,y\},\{x',y'\}\in\binom{V_H}{2}}\\ &\Pr[N(v)\cap V_H = \{x,y\}]\Pr[N(u)\cap V_H = \{x',y'\}]\\ &\times\Pr[N(v)\cap([n]\setminus X) = N(u)\cap([n]\setminus X) = \emptyset].\end{aligned}$$

The last probability equals $\Pr[N(v)\cap([n]\setminus X) = \emptyset]\Pr[N(u)\cap([n]\setminus X) = \emptyset](1-\frac{w_u w_v}{W_{[n]}})$, and so, by using (6.6), $p_{uv} = (1+o(1))p_u p_v$. A straightforward application of Chebyshev's inequality, together with the fact $|X_0| = \Omega(n(\log\log n)^{-8})$ then implies that with high probability $e_H \geqslant m/2 \geqslant c_4 T_0^{-2\beta+4}|X_0|$, where $c_4 = c_3/2$.

We now are ready to construct the sought coupling. To do so, we will define a sequence of (simple) couplings, whose concatenation will yield the desired result. We begin with defining an auxiliary multigraph H' such that $H' \subseteq H$ with probability one (where, as usual, containment is defined by replacing multi-edges by edges). Let $\mathcal{C} \subseteq X_0$ be the set of efficient connectors, that have exaxtly two neighbors in V_H, and no other neighbors. Let $v \in \mathcal{C}$. Note that $e_H = |\mathcal{C}|$. We estimate the probability that v is joined to x, y in V_H. Note that by (6.9) and (6.12), if we abbreviate $W_{V_H}^{(2)} = \sum_{u \in V_H} w_u^2$, then we find a $c_5 > 0$ such that

$$\begin{aligned}
&\Pr[N(v) \cap V_H = \{x, y\} \mid v \in \mathcal{C}] = p_v(x, y) \cdot p_v^{-1} \\
&\geqslant 2c_5\, w_v^2 \frac{w_x w_y}{W_{[n]}^2} \left(w_v^2 \frac{W_{V_H}^2 - W_{V_H}^{(2)}}{W_{[n]}^2} \right)^{-1} \\
&= 2c_5 \frac{w_x w_y}{W_{V_H}^2 - W_{V_H}^{(2)}} =: p_{xy}.
\end{aligned}$$

As this probability is independent of v, we define H' as follows. We start with an initially empty graph on vertex set V_H. Then we repeat $|\mathcal{C}|$ times independently: with probability p_{xy} add the edge xy, and otherwise, with the remaining probability $1 - c_5$, skip this step. With this definition we obviously always have $H' \subseteq H$.

In the second step, we construct another auxiliary graph H'' by getting rid of the “idle” steps in the construction of H'. The expected number of edges in H' is $c_5|\mathcal{C}| = c_5\, e_H$. Thus, by applying the Chernoff bounds, and using the fact that $e_H \geqslant c_4\, T_0^{-2\beta+4}|X_0|$ with high probability, we infer that there is an a $c_6 > 0$ such that with high probability the number of edges in H' is at least $c_6\, T_0^{-2\beta+4}|X_0| =: \ell_{H'}$. We construct H'' by adding independently $\ell_{H'}$ edges to an initially empty graph with vertex set V_H, such that in each step the edge xy is added with probability

$$p'_{xy} = \frac{p_{xy}}{\sum_{x',y' \in V_H, x' \neq y'} p_{x'y'}} = \frac{2 w_x w_y}{W_{V_H}^2 - W_{V_H}^{(2)}}.$$

It follows that there is a coupling such that $H'' \subseteq H'$ with high probability, as p'_{xy} equals the probability that the edge xy is included in H' in a specific step, conditional on the event that an edge is added in this specific step.

Finally, we couple H'' to $NR(\mathbf{w}'(n))$ such that $NR(\mathbf{w}'(n)) \subseteq H''$ with high probability. Note that

$$W' := \sum_{x \in V_H} w'_x = T_0^{-\beta+2} (\log T_0)^{-9} \sum_{x \in V_H} w_x.$$

By applying Lemma 4.1 we infer that the last sum is $O(n T_0^{-\beta+2})$. Hence, as $|X_0| = \Omega(n (\log T_0)^{-8})$, for sufficiently large n it follows that $W' \leqslant \frac{\ell_{H'}}{2}$. The probability space of the coupling is

$$\Omega = \prod_{\{x,y\} \in \binom{V_H}{2}} \Omega_{xy},$$

where Ω_{xy} is a Poisson process (or a Poisson *clock*) with intensity $(2w'_x(n))(2w'_y(n))/2W'$. We start with two initially empty graphs G_1, G_2, and we will show that in the end $G_1 \sim NR(\mathbf{w}')$ and with high probability $G_2 \sim H''$ and $G_1 \subseteq G_2$.

We construct the graphs as follows. Suppose that we observe a tick at time $t < 1/2$ originating from the clock xy. Then we add xy to both graphs. Moreover, unless the total number of ticks observed up to $t = 1/2$ is more than $\ell_{H'}$, we continue adding edges at the subsequent ticks to G_2 until exactly $\ell_{H'}$ were added.

We argue that indeed $G_1 \sim NR(\mathbf{w}')$. Note that an edge $xy \in G_1$ iff the corresponding Poisson clock ticked before $1/2$, independently of all other edges. So,

$$\begin{aligned}
&\Pr[xy \in G_1] = \Pr\left[\mathrm{Exp}\left(\frac{2W'}{(2w'_x)(2w'_y)} \right) < 1/2 \right] \\
&= 1 - e^{-w'_x w'_y / W'}.
\end{aligned}$$

So, $G_1 \sim NR(\mathbf{w}')$. To see that $G_1 \subseteq G_2$ holds with high probability, note that the expected number of edges in G_1 is

$$\sum_{x,y \in \binom{V_H}{2}} 1 - e^{-w'_x w'_y / W'} \leqslant \sum_{x,y \in \binom{V_H}{2}} \frac{w'_x w'_y}{W'} \leqslant W' \leqslant \frac{\ell_{H'}}{2}.$$

A straightforward application of the Chernoff bounds then guarantees that with high probability the number of edges in G_1 is less than$\ell_{H'}$. It remains to check that under this condition $G_2 \sim H''$. Note that the property that the exponential distribution is memoryless guarantees that the probability that the edge x, y is added at the point in time where some edge is added is independent of all other ticks and equals

$$\begin{aligned}
&\Pr\left[\mathrm{Exp}\Big(\frac{2W'}{(2w'_x)(2w'_y)} \Big) = \min_{\{x',y'\} \in \binom{V_H}{2}} \Big\{ \mathrm{Exp}\Big(\frac{2W'}{(2w'_{x'})(2w'_{y'})} \Big) \Big\} \right] \\
&= \frac{w'_x w'_y}{\sum_{\{x',y'\} \in \binom{V_H}{2}} w'_{x'} w'_{y'}} = p'_{x,y}.
\end{aligned}$$

With the above result at hand we are ready to prove Lemma 3.3.

Proof. [Proof of Lemma 3.3] By applying Lemma 6.1 to H, we infer that there is a coupling such that with high probability $NR(\mathbf{w}') \subseteq H$, where $\mathbf{w}' = (T_0^{-\beta+2} (\log T_0)^{-9} w_u)_{u \in V_H}$. Thus, it is enough to show the claim for $NR(\mathbf{w}')$, as the considered property is not violated if edges are added to the graph.

Let us write $N = |V_H| \leqslant n$. Recall that $T_0 = (\log n)^A$, where $A = 3(3-\beta)^{-2}$. This implies for sufficiently large n that the smallest weight in $\mathbf{w}'$ is at least

$$\begin{aligned}
&T_0^{-\beta+2} (\log T_0)^{-9} \cdot T_0 \geqslant T_0^{3-\beta} (\log T_0)^{-9} \\
&\geqslant (\log n)^{2/(3-\beta)} \geqslant (\log N)^{2/(3-\beta)}.
\end{aligned}$$

So, the sequence of weights $\mathbf{w}'(N)$ follows a (truncated) power law with minimum weight $(\log N)^{2/(3-\beta)}$. By applying Theorem 9.24 from [29] we infer that the diameter of $NR(\mathbf{w}')$ is bounded from above by $\frac{2+o(1)}{|\log(\beta-2)|}\log\log n$, which completes the proof.

6.4 Connecting vertices of high weight through efficient connectors – Proof of Lemma 3.4 Applying Lemma 3.2 we obtain a sequence of sets $X_1, X_2, \ldots, X_\lambda \subseteq X$, such that each X_i contains at least $c\,\alpha^{8(\lambda-i)}\,n$ nodes, each $u \in X_i$ is a $(\alpha^{\lambda-i}, 2)$-efficient connector, and X_i has no edge to any other node in X.

Let $i \geqslant 1$ and suppose that we have found sets $(\tilde{S}_j)_{j\leqslant i-1}$ with the desired properties, where additionally the efficient connectors joining vertices in $\tilde{S}_{j-1}$ to $\tilde{S}_j$ are only from X_j. We will show that there is a $\tilde{S}_i \subseteq S_i$ such that $|\tilde{S}_i| \geqslant |S_i|/2$, and every vertex in $\tilde{S}_i$ is connected by an efficient connector $x \in X_i$ to some vertex in $\tilde{S}_{i-1}$, and $\deg(x) = 2$. To this end, define for $u \in S_i$ the set,

$$H_u := \left\{x \in X_i\colon\ \deg(x) = 2 \wedge x \in N(u) \cap N(\tilde{S}_{i-1})\right\},$$

where $N(R)$ denotes the neighborhood of R. Let us first compute the probability that a fixed node $x \in X_i$ is in H_u. Note that

(6.13)
$$\begin{aligned}&\Pr[\,x \in H_u\,] \geqslant \\ &\frac{w_u w_x}{W_{[n]}} \cdot \sum_{y\in\tilde{S}_{i-1}} \frac{w_x w_y}{W_{[n]}} \cdot \prod_{r\in[n]\setminus(\{u,y\}\cup X)} \left(1 - \frac{w_x w_r}{W_{[n]}}\right) \\ &\overset{(w_x=\Theta(1))}{\geqslant} \frac{w_x^2}{W_{[n]}^2} \cdot e^{-2w_x} \cdot w_u \cdot W_{\tilde{S}_{i-1}}.\end{aligned}$$

In order to get an estimate for the above expression, note that all $z \in S_{i-1}$ satisfy $T_{i-1} \leqslant w_z < T_{i-2}$. Hence, $W_{\tilde{S}_{i-1}}$ is at least the sum of the weights of the $|S_{i-1}|/2$ nodes with smallest weights in S_{i-1}. Our assumption that $(F_n)_{n\geqslant 1}$ is of power law type β guarantees that there is a $\gamma > 0$ such that

$$\begin{aligned}|S_{i-1}| &= n((1 - F_n(T_{i-1})) - (1 - F_n(T_{i-2}))) \\ &\geqslant \gamma\, T_{i-1}^{-\beta+1}\, n.\end{aligned}$$

As the minimal weight in S_{i-1} is T_{i-1}, we readily infer that $W_{\tilde{S}_{i-1}} \geqslant \gamma\, T_{i-1}^{-\beta+2}\, n$. Let us now compute the expected size of H_u. As for any $x \in X_i$ it holds that $w_x = \Theta(1)$, Equation (6.13) implies that there is a $\gamma' > 0$ such that

$$\begin{aligned}\mathbb{E}[|H_u|] &\geqslant \sum_{x\in X_i} \frac{w_x^2}{W_{[n]}^2} \cdot e^{-2w_x} \cdot w_u \cdot \gamma\, T_{i-1}^{-\beta+2}\, n \\ &\geqslant \gamma' \cdot \frac{|X_i|}{n}\, w_u\, T_{i-1}^{-\beta+2}.\end{aligned}$$

Using that $|X_i| \geqslant c\alpha^{8(\lambda-i)} \cdot n$, and $w_u \geqslant T_i = T_0^{\alpha^i}$ we infer that there is a $\gamma'' > 0$ such that

$$\begin{aligned}\mathbb{E}[|H_u|] &\geqslant \gamma''\alpha^{8(\lambda-i)} T_0^{\alpha^{i-1}(-\beta+2)} \cdot T_0^{\alpha^i} \\ &\overset{(\alpha=2(\beta-2)/(\beta-1))}{=} \gamma''\, \alpha^{8(\lambda-i)}\, T_i^{(3-\beta)/2}.\end{aligned}$$

Easy calculus proves there is a constant $\kappa_1 > 0$ such that for all $i \leqslant \lambda - \kappa_1$ the above expression is decreasing in i and lower bounded by $T_i^{(3-\beta)/4}$.

Since $|H_u|$ is a binomial random variable (recall that by definition, there are no internal edges between vertices in X_i), it follows by a Chernoff bound that

$$\begin{aligned}\Pr[\,|H_u| = 0\,] &\leqslant \Pr\left[\,|H_u| \leqslant \frac{1}{2}\mathbb{E}[|H_u|]\,\right] \\ &\leqslant \exp(-\mathbb{E}[|H_u|]/8) \leqslant \exp(-T_i^{(3-\beta)/4}/8).\end{aligned}$$

Now define

$$\tilde{S}_i := \{u \in S_i\colon |H_u| > 0\}.$$

Hence $\mathbb{E}\left[|S_i \setminus \tilde{S}_i|\right] \leqslant |S_i| \cdot \exp(-T_i^{(3-\beta)/4}/8)$, and Markov's inequality yields

$$\begin{aligned}&\Pr\left[\,|S_i \setminus \tilde{S}_i| \geqslant \exp(-T_i^{(3-\beta)/4}/16) \cdot |S_i|\,\right] \\ &\leqslant \exp(-T_i^{(3-\beta)/4}/16).\end{aligned}$$

Let κ_2 be such that for all $i \leqslant \lambda - \kappa_2$ we have that $\exp(-T_i^{(3-\beta)/4}/16) \leqslant 1/2$. Now choose the constant κ in the conclusion of the lemma large enough, in particular, at least as large as $\max\{\kappa_1, \kappa_2\}$, so that we also have

$$\sum_{i=0}^{\lambda-\kappa} \exp(-T_i^{(3-\beta)/4}/16) \leqslant \varepsilon.$$

holds. With this choice, the statement of the lemma follows since the probability that any of the constructed sets $\tilde{S}_i$, where $i \leqslant \lambda - \kappa$, fails to satisfy the desired properties is bounded by ε.

6.5 Passing the information to the periphery – Proof of Lemma 3.5 To prove Lemma 3.5, we need to focus on the local structure of $CL[R]$, where $R := [n] \setminus \{S \cup X\}$. In particular, we will focus on the subgraph of it that is induced by the vertices that have degree bounded by some constant, which we will specify during our proof. In particular, bounding the total number of vertices that lie within a certain distance from vertices of high degree will show that for most of the vertices in R their neighbourhoods consist of small degree vertices. Being more precise, we first show the following lemma.

LEMMA 6.2. *Fix $D > 0$. Let $G = (V, E)$ be a graph so that for all $0 \leqslant x \leqslant D$,*

$$\gamma_1 x^{-\beta+1} \leqslant \frac{1}{n}\sum_{i=1}^{n} \mathbf{1}[\deg(i) > x] \leqslant \gamma_2 x^{-\beta+1}, \tag{6.14}$$

where $0 < \gamma_1' \leqslant \gamma_2'$ are constants. Define $N_0(D) := \{u \in V\colon \deg(u) > D\}$ and $N_i(D) := N(N_{i-1}(D))$. Then

$$|N_i(D)| \leqslant c \cdot n \cdot (D)^{(1-\beta)\cdot\left(\frac{\beta-2}{\beta-1}\right)^i},$$

where $c > 0$ is a constant independent of D and i.

Proof. Using the same arguments as in Lemma 4.1 (the only difference is that we work with degrees instead of weights), we obtain that

$$\mathrm{vol}(N_0(D)) = \sum_{i\colon\ \deg(i) > D} \deg(i) \leqslant cnD^{-\beta+2},$$

where $c > 0$ is a constant. Hence, $|N_1(D)| \leqslant \mathrm{vol}(N_0(D)) \leqslant cnD^{-\beta+2}$. More generally, let $s_i := |N_i(D)|$. Since $|N_{i+1}(D)| \leqslant \mathrm{vol}(N_i(D))$, we want to find a value ℓ_i so that

$$\sum_{j=1}^{n} \mathbf{1}[\deg(j) > \ell_i] \geqslant s_i.$$

Using Lemma 6.2, we find that $\ell_i := \tilde{c}(n/s_i)^{1/(\beta-1)}$, where $\tilde{c} > 0$ is a constant, satisfies the above inequality. Hence,

$$\begin{aligned} s_{i+1} &:= N_{i+1}(D) \leqslant \mathrm{vol}(N_i(D)) \\ &= \sum_{j=1}^{n} \mathbf{1}[\deg(j) > \ell_i]\deg(j) \leqslant c \cdot n \cdot \ell_i^{-\beta+2} \\ &\leqslant c \cdot n \cdot \left(\frac{n}{s_i}\right)^{(-\beta+2)/(\beta-1)} = \tilde{c} n^{1/(\beta-1)} s_i^{(\beta-2)/(\beta-1)}, \end{aligned}$$

where c is a constant independent of D and i. Now define $a_i := s_i/n$ and $N := (\beta-2)/(\beta-1) \in (0,1)$. Then,

$$a_{i+1} \leqslant c \cdot a_i^N.$$

Solving this recursion yields

$$a_i \leqslant \tilde{c}^{1/(1-N)} \cdot a_1^{N^i},$$

which implies the lemma.

Recall that by Theorem 4.1 $CL(\mathbf{w})$ satisfies 6.14 with probability $1 - o(1)$. Also, $N = nF_n(C) - \lceil \varepsilon n \rceil$. Therefore we obtain the following corollary.

LEMMA 6.3. *For any real $\delta > 0$ and any positive integer r, there exists a positive integer $D = D(\delta, r)$ such that with probability $1 - o(1)$ there are at most δN vertices in $CL(\mathbf{w})$ which lie within distance r from the set of vertices that have degree larger than D.*

The next step in our proof has to do with structure of $CL[R]$. Firstly, we will show that with high probability $CL[R]$ has a unique giant component, which essentially is almost all of the giant component of $CL(\mathbf{w})$. Thereafter, we apply the above lemma in order to deduce that most of the vertices of the giant component do not have any vertices of degree larger than D (in R) within distance r, where D and r will be specified during the proof. In other words, for most of the vertices of the giant component of $CL[R]$ their balls of radius r consist of bounded degree vertices. These balls contain at least r vertices. Thereafter, we show that most of these balls consist of vertices that have a small number of edges to S and also have at least one edge to $\tilde{S}$. This edge connects the centre of the ball to $\tilde{S}$ through a path where every internal vertex has bounded degree. As we shall see, such a path is able to "communicate" the information between its endpoints in a small number of rounds. Of course, we need to bound the degree of the vertices of R to the set X. But we do separately at the end of this section, as here we will be conditioning on a certain realisation of the set $\tilde{S}$ as this is specified in Lemma 3.4, and this requires the exposure of the edges between X and R. We now proceed with the detailed exposition of these steps.

6.5.1 $CL[R]$ as an inhomogeneous random graph We will express $CL[R]$ in the framework of inhomogeneous random graphs, as this will be essential at various steps during our proof. Now, for distinct $i, j \in R$, the probability of the edge $\{i, j\}$ being present is equal to

$$p_{ij}(\mathbf{w}) = \frac{w_i w_j}{W_n(\mathbf{w})}.$$

Note that all vertices in R have weight at most $C := T_0^{i_{max}}$ and note that this can become as large as we need. Set $N := |R|$ and recall that $N = nF_n(C) - \lceil \varepsilon n \rceil$. For future use, let us set right now $\varepsilon' := \lceil \varepsilon n \rceil / n$. Also, recall that W_n is the random variable which is the weight of a randomly chosen vertex from $[n]$. Then we write $W_n(\mathbf{w})/n = \mathbb{E}[\,W_n\,]$ and, in turn, the above probability can by written as

$$p_{ij}(\mathbf{w}) = \frac{F_n(C) - \varepsilon'}{N} \frac{w_i w_j}{\mathbb{E}[\,W_n\,]}.$$

We consider the points $i/N \in (0,1]$, for all $i \in [N]$. For every $i \in [n]$ we set $w_i = [F_n]^{-1}(i/n)$, where $[F_n]^{-1}$ is the generalised inverse of the function $F_n(x)$ and it is defined as $[F_n]^{-1}(u) := \sup\{s : F_n(s) \leqslant u\}$, for all $u \in (0,1]$.

We define the kernel function $\kappa_{n,R}(x,y)$ for $(x,y) \in (0,1]^2$ to be

$$\kappa_{n,R}(x,y) = (F_n(C) - \varepsilon') \, \frac{\psi_n(x) \cdot \psi_n(y)}{\mathbb{E}[\, W_n \,]},$$

where $\psi_n(x) = [F_n]^{-1}(\varepsilon' + x(F_n(C) - \varepsilon'))$. Therefore, since $N = n(F_n(C)-\varepsilon')$, we have $\psi_n(i/N) = [F_n]^{-1}(\varepsilon' + i/n)$. We will be assuming throughout this section that C is a point of continuity of F - this assumption excludes only countably many choices for C.

Note that this sequence of functions has a limit κ_R which is a real-valued function on $(0,1]^2$. This is

$$\kappa_R(x,y) = (F(C) - \varepsilon) \, \frac{\psi(x) \cdot \psi(y)}{\mathbb{E}[\, W_F \,]},$$

where $\psi(x) = [F]^{-1}(\varepsilon + x(F(C) - \varepsilon))$ and W_F is a real-valued random variable whose distribution is F. This is a pointwise limit of the sequence $\{\kappa_{n,R}\}_{n\in\mathbb{N}}$ on the points (x,y) of continuity of $[F]^{-1}(\varepsilon + x(F(C) - \varepsilon)) \cdot [F]^{-1}(\varepsilon + y(F(C) - \varepsilon))$, that is, for almost every $(x,y) \in (0,1]^2$.

We need to verify a number of conditions regarding the functions $\kappa_{n,R}$ as well as the limiting function κ. Namely, we need to verify that $\{\kappa_{n,R}\}_{n\in\mathbb{N}}$ is a *graphical sequence with limit* κ_R (see [29] Def. 9.2 p. 202). We have already shown that κ is the pointwise limit of the sequence $\{\kappa_{n,R}\}_{n\in\mathbb{N}}$. We also need to verify the following properties:

i. κ_R is continuous almost everywhere - this follows from the definition of κ_R;

ii. $\int_{[0,1]^2} \kappa_R(x,y)dxdy < \infty$ - this follows as $\psi(x) = [F]^{-1}(\varepsilon + x(F(C) - \varepsilon)) \leqslant C$ for all $x \in (0,1]$;

iii. if $e(R)$ denotes the number of edges of $CL[R]$, then we have

$$\lim_{N\to\infty} \frac{1}{N}\mathbb{E}[\, e(R) \,] = \frac{1}{2}\int_{(0,1]^2} \kappa_R(x,y)dxdy.$$

To see this, we write:

$$\begin{aligned}\mathbb{E}[\, e(R) \,] &= \frac{1}{2}\sum_{i\neq j\in[N]} \frac{\kappa_{n,R}(i/N, j/N)}{N} \\ &= N\Bigg(\frac{1}{2}\int_{(0,1]^2} \kappa_{n,R}(x,y)dxdy \\ &\quad - \frac{1}{N}\frac{(F(C)-\varepsilon)}{\mathbb{E}[\, W_n \,]}\int_{(0,1]} \psi(x)^2 dx\Bigg).\end{aligned}$$

But the functions $\kappa_{n,R}(x,y)$ are uniformly bounded and furthermore their sequence converges pointwise almost everywhere to κ_R, which is integrable. Thus, by the Bounded Convergence Theorem, the first integral converges to $\int_{[0,1]^2} \kappa_R(x,y)dxdy$, while the second term is vanishing as $N \to \infty$.

It is also clear that κ_R is *irreducible*, that is, there exists no subset $A \subset [0,1]$ of positive Lebesque measure for which $\kappa_R = 0$ almost everywhere in $A \times ([0,1] \setminus A)$. The above properties are minimal assumptions on the asymptotic behaviour of kernel sequences usually required in the theory of inhomogeneous random graphs (see e.g. [3] or [29]).

6.5.2 $CL[R]$ and its giant component We now show that $CL[R]$ itself has a giant component with high probability. We will use the standard theory of inhomogeneous random graphs to show that, with high probability, $CL[R]$ in fact has a unique component whose number of vertices is a certain fraction of the vertices of R, whereas every other component has at most logarithmic size.

Since the sequence $\{\kappa_{n,R}\}_{n\in\mathbb{N}}$ is graphical with limit κ_P, which is irreducible, by Theorem 3.1 in [3], it suffices to show the following.

LEMMA 6.4. *The norm of the linear operator* $T_{\kappa_R} : L^2((0,1]) \to L^2((0,1])$ *defined as*

$$(T_{\kappa_R} f)(x) = \int_{(0,1]} \kappa_R(x,y)f(y)dy,$$

is strictly larger than 1. (Here and below integration is meant to be Lebesque integration.)

Proof. To show this, it is enough to show that there exists a real-valued function f on $(0,1]$ whose L^2-norm is equal to 1 and $T_{\kappa_R} f$ has L^2-norm strictly larger than 1. We set $f(x) := \psi(x)/\left\|\psi\right\|_2$. Thus, for any $x \in (0,1]$ we have

$$\begin{aligned}(T_{\kappa_R} f)(x) = \frac{1}{\|\psi\|_2}\int_{(0,1]} \kappa_R(x,y)\psi(y)dy = \\ \frac{(F(C)-\varepsilon)\psi(x)}{\|\psi\|_2\,\mathbb{E}\,[W_F]}\int_{(0,1]} \psi^2(y)dy.\end{aligned} \tag{6.15}$$

But

$$\begin{aligned}\int_{(0,1]} \psi^2(y)dy &= \int_{(0,1]} \left([F]^{-1}(\varepsilon + y(F(C)-\varepsilon))\right)^2 dy \\ &= \frac{1}{(F(C)-\varepsilon)}\int_{(\varepsilon,F(C)]} \left([F]^{-1}(y)\right)^2 dy \\ &= \frac{1}{(F(C)-\varepsilon)}\int_{(0,1)} \left([F]^{-1}(y)\right)^2 \mathbf{1}_{\{\varepsilon<y\leqslant F(C)\}}dy.\end{aligned}$$

But recall that almost everywhere

$$\left([F_n]^{-1}(y)\right)^2 \mathbf{1}_{\{\varepsilon'<y\leqslant F_n(C)\}} \to \left([F]^{-1}(y)\right)^2 \mathbf{1}_{\{\varepsilon<y\leqslant F(C)\}}$$

as $n \to \infty$. As the functions $\left([F_n]^{-1}(y)\right)^2 \mathbf{1}_{\{\varepsilon'<y\leqslant F_n(C)\}}$ are uniformly bounded by C^2, the Bounded Convergence Theorem implies that

$$\lim_{n\to\infty} \int_{(0,1]} \left([F_n]^{-1}(y)\right)^2 \mathbf{1}_{\{\varepsilon'<y\leqslant F_n(C)\}} dy = \int_{(0,1]} \left([F]^{-1}(y)\right)^2 \mathbf{1}_{\{\varepsilon<y\leqslant F(C)\}} dy.$$

To bound the maximum weight in X we define as $c_X := \sup\{s > 0 \ : \ F_n(s) \leqslant \varepsilon\}$. Thus for any s in this set we have $1-F_n(s) \geqslant 1-\varepsilon$. Also, $\gamma_2 s^{-\beta+1} \geqslant 1-F_n(s)$. These two inequalities yield

$$s \leqslant \left(\frac{\gamma_2}{1-\varepsilon}\right)^{\frac{1}{\beta-1}} \ \Rightarrow \ c_X \leqslant \left(\frac{\gamma_2}{1-\varepsilon}\right)^{\frac{1}{\beta-1}}. \tag{6.16}$$

So for all n

$$\int_{(0,1]} \left([F_n]^{-1}(y)\right)^2 \mathbf{1}_{\{\varepsilon'<y\leqslant F_n(C)\}} dy \geqslant \frac{1}{n} \sum_{i \ : \ \left(\frac{\gamma_2}{1-\varepsilon}\right)^{\frac{1}{\beta-1}} \leqslant w_i \leqslant C} w_i^2 = \Theta(C^{3-\beta}),$$

by Lemma 4.1. Thus, if n and C are large enough, we have

$$\frac{F(C)-\varepsilon}{\mathbb{E}[\,W_F\,]} \int_{(0,1]} \psi^2(y) dy > 1$$

and (6.15) implies that the L^2-norm of $T_{\kappa_R} f$ is strictly greater than 1.

Thus, Theorem 3.1 in [3] implies the existence of a giant component in $CL[R]$. If $C_1(R)$ denotes the (lexicographically first) largest component in $CL[R]$ and $|C_1(R)|$ its number of vertices, we have the following.

Corollary 6.1. *There exists a constant $\zeta_R \in (0,1)$ such that*

$$\frac{|C_1(R)|}{N} \to \zeta_R,$$

in probability *as $n \to \infty$, whereas every other component contains $O(\log n)$ vertices.*

6.5.3 Balls of bounded degree vertices Taking $\delta = \varepsilon\zeta_R/4$ in Lemma 6.3 and letting r be sufficiently large (we will specify this later) we obtain a $D = D(\varepsilon, r)$ such that with probability $1-o(1)$ there are at most δN vertices in $CL(\mathbf{w})$ and, therefore in $CL[R]$ as well, which lie within distance r from the set of vertices that have degree larger than D. Thus, for all but at most δN of the vertices of $C_1(R)$, the ball of radius r around each such vertex contains only vertices of degree at most D in R. For a vertex v, we denote by $B_r(v)$ the ball of radius r around v.

Therefore, we obtain the following lemma.

Lemma 6.5. *For every $\varepsilon > 0$ and every positive integer r there exists a positive integer D such that with probability $1 - o(1)$, all but at most $\varepsilon|C_1(R)|/2$ vertices of $C_1(R)$ are such that for every such vertex v we have $|B_r(v)| \geqslant r$.*

We will denote this subset of vertices of $V(C_1(R))$ by G_R.

6.5.4 Thin paths between $C_1(R)$ and $\tilde{S}$ Here, we will argue that most of the vertices of G_R are such that the ball around each of them does not have many neighbors in S and at least one vertex there is adjacent to $\tilde{S}$.

We call a vertex $v \in G_R$ *efficient* if each vertex $u \in B_r(v)$ has degree at most $\hat{\Delta}$ in S and there is a vertex in $B_r(v)$ that has at least 1 neighbor in $\tilde{S}$. Moreover, we associate each vertex in G_R with an indicator random variable I_v which is equal to 1 if and only if v is efficient. In the sequel we show that the probability that I_v is equal to 1 is close to 1. Thereafter, we will show that indeed with high probability the sum of these indicators is very close to $|C_1(R)|$. We will do this by means of Markov's inequality.

Let v be a vertex of G_R. We will bound from below $\Pr[\, I_v = 1\,]$. We first bound the probability that a vertex $u \in B_r(v)$ has at least $\hat{\Delta}$ neighbours in S. To this end, we need to bound the weight of S. As S consists of all vertices that have weight at least C, by Lemma 4.1 we have

$$W_{[S]} = O\left(nC^{-\beta+2}\right).$$

Therefore,

$$\mathbb{E}\left[\deg_S(u)\right] = O\left(w_u C^{-\beta+2}\right) = O\left(C^{3-\beta}\right).$$

Thus, by Markov's inequality

$$\Pr\left[\deg_S(u) \geqslant \hat{\Delta}\right] = O\left(\frac{C^{3-\beta}}{\hat{\Delta}}\right).$$

As $|B_r(v)| \leqslant D^r$, we make $\hat{\Delta}$ large enough, we obtain

$$\Pr\left[\exists\, u \in B_r(v) \ : \ \deg_S(u) \geqslant \hat{\Delta}\right] \leqslant \varepsilon^2/8. \tag{6.17}$$

Finally, we will bound from above the probability that u has no neighbours in $\tilde{S}$. The probability that this is not adjacent to $\tilde{S}$ is:

$$\prod_{i\in\tilde{S}} \left(1 - \frac{w_u w_i}{W_{[n]}}\right) \leqslant \exp\left(-w_u \frac{W_{[\tilde{S}]}}{W_{[n]}}\right). \tag{6.18}$$

We now need to bound $W_{[\tilde{S}]}$ from below. Recall that $\tilde{S}$ contains at least half of the vertices in S. Thus, to bound its total weight from below it is enough to bound from above the total weight of the $\lceil \frac{1}{2}|S| \rceil$ vertices of highest weight in $[n]$. We do so in the proof of the following claim.

Claim 6.1. *We have*

$$W_{[\tilde{S}]} \geqslant W_{[S]} \left(1 - O\left(\left(\frac{\gamma_1}{\gamma_2} \right)^{\frac{\beta-2}{\beta-1}} \right) \right).$$

Proof. Recall that $|S| \geqslant n\gamma_1 C^{-\beta+1}$. This means that it is enough to bound from above the total weight of the $\lceil \frac{n}{2}\gamma_2 C^{-\beta+1} \rceil$ nodes of highest weight. Assuming that the weight of these vertices is bounded from below by C', we will determine the largest such C'. For this, we again use the assumption that $\gamma_1 C'^{-\beta+1} \leqslant 1 - F_n(C')$, this time having

$$n(1 - F_n(C')) = \lceil \frac{n}{2}\gamma_2 C^{-\beta+1} \rceil \leqslant n\gamma_2 C^{-\beta+1}.$$

Thus, we obtain:

$$C' \geqslant \left(\frac{\gamma_1}{\gamma_2} \right)^{\frac{1}{\beta-1}} C.$$

But by Lemma 4.1, all vertices that have weight at least C' have total weight $\Theta\left(n \left(\frac{\gamma_1}{\gamma_2} \right)^{\frac{\beta-2}{\beta-1}} C^{-\beta+2} \right)$.

Using the bound of this claim in (6.18) we deduce that the probability that u is not adjacent to $\tilde{S}$ is at least $1 - O(C^{-\beta+2})$. Therefore, since $B_r(v)$ contains at least r vertices, the probability that none of its vertices is adjacent to $\tilde{S}$ is at most $(1 - O(C^{-\beta+2}))^r \leqslant \varepsilon^2/8$, provided that r is chosen large enough compared to C. This together with (6.17) imply that

$$\Pr [I_v = 0] \leqslant \varepsilon^2/2.$$

Therefore the expected number of vertices of G_R that are not efficient is at most $\varepsilon^2|G_R|/4 \leqslant \varepsilon^2|C_1(R)|/4$ and Markov's inequality implies that with probability at least $1 - \varepsilon/2$ there are at most $\varepsilon|C_1(R)|/2$ of them. To sum up, together with Lemma 6.5 we have arrived at the following fact.

Lemma 6.6. *For every $\varepsilon > 0$ if C is sufficiently large, then there exist integers $r, \hat{\Delta}$ such that, with probability at least $1 - \varepsilon/2$, at least $(1-\varepsilon/2)|C_1(R)|$ of the vertices in $C_1(R)$ are such that there is a path emanating from each of them to $\tilde{S}$ of length at most r whose internal vertices have degree at most $\hat{\Delta}$ to S.*

Lemma 6.6 almost completes the proof of Lemma 3.5. We finally need to argue about the vertices that belong to small components of $CL[R]$.

Outside the giant component of $CL[R]$ We will show that we can effectively ignore the vertices that lie outside the giant component. Let us assume that we have a C as in Lemma 6.6, which we can make as large as we please. We prove the following lemma.

Lemma 6.7. *For every $\varepsilon > 0$ and for any C large enough (depending on ε), with probability at least $1 - \varepsilon$ at most εn vertices that do not belong to $C_1(R)$ belong to components directly connected to S.*

Proof. We will distinguish the remaining components of $CL[R]$ (that is, all components apart from the largest one) into three classes. Let $\mathcal{C}^1$ be the subset of these components that contain at most $C^{(\beta-2)/6}$ vertices all of them having weight less than $C^{(\beta-2)/6}$. We also let $\mathcal{C}^2$ be the subset of components that also contain at most $C^{(\beta-2)/6}$ vertices but contain at least one vertex of weight at least $C^{(\beta-2)/6}$. Finally, let $\mathcal{C}^3$ be the remaining components (of course, apart from the largest one). We will argue about each one of these subsets separately.

Regarding $\mathcal{C}^1$, we will show that most of the components contained there are not connected to S. Indeed, let us consider a component $B \in \mathcal{C}^1$. The probability that none of its vertices is adjacent to S is at least

$$\prod_{v \in V(B)} \left(1 - \frac{w_v W_{[S]}}{W_{[n]}} \right) \geqslant 1 - \frac{W_{[S]}}{W_{[n]}} \sum_{v \in V(B)} w_v$$
$$\geqslant 1 - \frac{W_{[S]}}{W_{[n]}} C^{\frac{2(\beta-2)}{6}} = 1 - \Theta\left(C^{-\frac{2(\beta-2)}{3}} \right),$$

using Lemma 4.1. Let us set ε_1^2 to be the above error term multiplies by the bound on the number of vertices in each component of $\mathcal{C}^1$, that is, $\varepsilon_1^2 = \Theta\left(C^{-\frac{(\beta-2)}{2}} \right)$. Thus, the expected number of vertices of $CL[R]$ which belong to components in $\mathcal{C}^1$ that are connected to S is at most $\varepsilon_1^2|\mathcal{C}^1| \leqslant \varepsilon_1^2 n$. Thus, using Markov's inequality we deduce that with probability at least $1 - \varepsilon_1$ there are at most $\varepsilon_1 n$ vertices contained in such components.

For $\mathcal{C}^2$, we will simply show that $|\mathcal{C}^2|$ itself is small. As $\mathcal{C}^2$ consists of components that contain at least one vertex of weight at least $C^{(\beta-2)/6}$, the size of $\mathcal{C}^2$ is at most the number of vertices of weight at least $C^{(\beta-2)/6}$. But the number of these vertices is $n(1 - F_n(C^{(\beta-2)/6})) = O(nC^{\frac{(\beta-2)(1-\beta)}{6}})$. Thus, for every $\varepsilon_2 > 0$, there exists C large enough such that the total number of vertices in components that belong to $\mathcal{C}^2$ is at most $\varepsilon_2 n$, for any n large enough.

Now, regarding $\mathcal{C}^3$, we will use more precise bounds on the counts of components of size at least $C^{(\beta-2)/6}$ from the theory of inhomogeneous random graphs as developed in [3]. In Lemma 6.4, we proved that

$CL[R]$ is supercritical and therefore it has a unique giant component of linear order, whereas every other component has $O(\log n)$ vertices. By Theorem 12.3 in [3], since κ is bounded by C, the distribution of the random graph induced by these components is that of an inhomogeneous random graph on $[N]$ with a certain limiting kernel $\hat{\kappa}_R$ on $(0,1]^2$ equipped with a certain measure such that $\|T_{\hat{\kappa}_R}\| < 1$. For this random graph, (12.1) and (12.2) in [3] imply that the probability that a certain vertex belongs to a component with at least $C^{(\beta-2)/6}$ vertices but not in the largest one is at most $e^{-aC^{(\beta-2)/6}}$, for some constant a, which does not depend on C. Therefore, the expected number of such vertices is at most $Ne^{-aC^{(\beta-2)/6}}$. So, for every $\varepsilon_3 > 0$ there exists C large enough such that the expected number of such vertices is at most $\varepsilon_3^2 n$. Markov's inequality implies that with probability at least $1-\varepsilon_3$, the number of vertices belonging to components of order at least $C^{\frac{\beta-2}{6}}$ but not in the largest one is at most $\varepsilon_3 n$.

These three cases complete the proof of the lemma, setting $\varepsilon := (\varepsilon_1+\varepsilon_2+\varepsilon_3)^{1/2}$ and taking C sufficiently large.

Finishing the proof of Lemma 3.5 In the above argument, we have not dealt at all with the degrees of the vertices in R to the set X. We did so, because in the above analysis we assumed that we have exposed the set $\tilde{S}$ which requires the exposure of the potential edges between the sets X and R. To complete the proof of Lemma 3.5, we consider these edges separately.

LEMMA 6.8. *For any $\varepsilon > 0$ and $C > 0$, there exists $\tilde{\Delta}$ such that with probability at least $1-\varepsilon/2$, there are at most $\varepsilon\zeta_R N/4$ vertices in R which have at least $\tilde{\Delta}$ neighbours in X.*

Proof. Firstly, we bound from below the probability that $u \in R$ has at least $\tilde{\Delta}$ neighbours in X. To do so, we first need to bound the total weight of X

CLAIM 6.2. *With X being defined as the set of the εn vertices of smallest weight in $\mathbf{w}$, we have*

$$W_{[X]} \leqslant W_{[n]} - O\left(n\left(\frac{1-\varepsilon}{\gamma_2}\right)^{\frac{\beta-2}{\beta-1}}\right).$$

Proof. Recall that the maximum weight in X is bounded by c_X as in (6.16). Thus, using Lemma 4.1

$$W_{[X]} \leqslant \sum_{i:w_i \leqslant \left(\frac{\gamma_2}{1-\varepsilon}\right)^{\frac{1}{\beta-1}}} w_i = W_{[n]} - O\left(n\left(\frac{1-\varepsilon}{\gamma_2}\right)^{\frac{\beta-2}{\beta-1}}\right),$$

which concludes the proof of the claim.

The above claim immediately implies that the expected degree of u in X, which we denote by $\deg_X(u)$ is at most

$$\mathbb{E}\left[\deg_X(u)\right] \leqslant \frac{w_u W_{[X]}}{W_{[n]}} = w_u\left(1 - O\left(\left(\frac{1-\varepsilon}{\gamma_2}\right)^{\frac{\beta-2}{\beta-1}}\right)\right)$$
$$\leqslant C\left(1 - O\left(\left(\frac{1-\varepsilon}{\gamma_2}\right)^{\frac{\beta-2}{\beta-1}}\right)\right).$$

Therefore, by Markov's inequality

$$\Pr\left[\deg_X(u) \geqslant \tilde{\Delta}\right] = O\left(\frac{C}{\tilde{\Delta}}\right).$$

Thus making $\tilde{\Delta}$ large enough, we deduce that

$$\Pr\left[\deg_X(u) \geqslant \tilde{\Delta}\right] \leqslant \varepsilon^2\zeta_R/8. \tag{6.19}$$

Thus, the expected number of such vertices is at most $\varepsilon^2\zeta_R N/8$. Applying Markov's inequality, the probability that there are at least $\varepsilon\zeta_R N/4$ such vertices in R is at most $\varepsilon/2$.

By Corollary 6.1, we deduce that for n large enough, with probability at least $1-\varepsilon$, there are at most $\varepsilon|C_1(R)|/2$ vertices in $C_1(R)$ that have degree at least $\tilde{\Delta}$ in X. Now, combining this lemma together with Lemmas 6.6 and 6.7 and setting $\Delta = \hat{\Delta} + \tilde{\Delta}$, we finally deduce Lemma 3.5.

7 Concluding remarks

The present work establishes for a class of random graphs ultrafast time bounds on the running time of the synchronous push-pull protocol that is needed until the majority of the vertices are informed. This class of random graphs has power law degree distribution with exponent $\beta \in (2,3)$. On the other hand, we show that when β exceeds 3, then the synchronous version of the push-pull protocol needs logarithmic to spread the rumor even on a relatively small part of the underlying graph. Thus, the exponent $\beta = 3$ is the critical point, such that when β crosses this value we have an exponential speedup of the synchronous push-pull protocol. We believe that this is a universal phenomenon and ultrafast dissemination of the rumor for both versions of the push-pull protocol occurs, when the underlying graph is a random graph that has power law degree distribution with exponent $\beta \in (2,3)$. In a forthcoming paper, we establish this for preferential attachment random graphs, proving that the statement of Theorem 1.1 as well as that of Theorem 1.2 do also hold in this case.

References

[1] www.internetworldstats.com.

[2] J.P. Bagrow, D. Wang and A. Barabási, *Collective Response of Human Populations to Large-Scale Emergencies*, PLoS ONE, 6 (2011), e17680.

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

[4] S. Boyd, A. Ghosh, B. Prabhakar and D. Shah, *Randomized gossip algorithms*, IEEE Transactions on Information Theory, 52 (2006), pp. 2508–2530.

[5] J. Candia, M.C. González, P. Wang, T. Schoenhar, G. Madey and A. Barabási, *Uncovering individual and collective human dynamics from mobile phone records*, J. Phys. A: Math. Theor., 41 (2008), 224015.

[6] K. Censor-Hillel and H. Shachnai, *Partial information spreading with application to distributed maximum coverage*, in Proceedings of the 29th ACM Symposium on Principles of Distributed Computing, 2010, pp. 161–170.

[7] F. Chierichetti, S. Lattanzi and A. Panconesi, *Rumor Spreading in Social Networks*, in Proceedings of the 36th International Colloquium on Automata, Languages and Programming, 2009, pp. 375–386.

[8] F. Chierichetti, S. Lattanzi and A. Panconesi, *Almost tight bounds for rumour spreading with conductance*, in Proceedings of the 42nd Symposium on the Theory of Computing, 2010, pp. 399–408.

[9] F. Chierichetti, S. Lattanzi and A. Panconesi, *Rumour spreading and graph conductance*, in Proceedings of the 21st ACM-SIAM Symposium on Discrete Algorithms, 2010, pp. 1657–1663.

[10] F. Chung and L. Lu, *Connected components in random graphs with given expected degree sequences*, Annals of Combinatorics, 6 (2002), pp. 125–145.

[11] F. Chung and L. Lu, *The Average Distance in a Random Graph with Given Expected Degrees*, Internet Mathematics, 1 (2003), pp. 91–113.

[12] F. Chung L. Lu and V. Vu, *The spectra of random graphs with given expected degrees*, Proc Natl Acad Sci U S A.,100 (2003), pp. 6313-6318.

[13] A. Demers, D. Greene, C. Hauser, W. Irish, J. Larson, S. Shenker, H. Sturgis, D. Swinehart and D. Terry, *Epidemic algorithms for replicated database maintenance*, In Proceedings of the 6th ACM Symposium on Principles of Distributed Computing, 1987, pp. 1–12.

[14] S. Dereich and C. Mönch and P. Mörtens, *Typical distances in ultrasmall random networks*, 2011 (available at http://arxiv.org/abs/1102.5680).

[15] B. Doerr, M. Fouz and T. Friedrich, *Social Networks Spread Rumors in Sublogarithmic Time*, In Proceedings of the 43rd ACM Symposium on Theory of Computing, 2011, pp. 21–30.

[16] R. Elsässer, *On Randomized Broadcasting in Power Law Networks*, In Proceedings of the 20th Symposium on Distributed Computing, 2006, pp. 370–384.

[17] M. Faloutsos, P. Faloutsos and C. Faloutsos, *On power-law relationships of the Internet topology*, In Proceedings of the SIGCOM Conference on Applications, Technologies, Architectures and Protocols for Computer Communication, 1999, pp. 251–262.

[18] G. Giakkoupis, *Tight bounds for rumor spreading in graphs of a given conductance*, In Proceedings of the 28th Symposium on Theoretical Aspects of Computer Science, 2011, pp. 57–68.

[19] B. Haeupler, *Analyzing Network Coding Gossip Made Easy*, In Proceedings of the 43rd Symposium on the Theory of Computing, 2011, pp. 293–302.

[20] U. Kang, C.E. Tsourakakis, A.P. Appel, C. Faloutsos, J. Leskovec, *HADI: Mining Radii of Large Graphs*, ACM Trans. Knowl. Discov. Data, 5 (2011), pp. 1–24.

[21] R. Karp, C. Schindelhauer, S. Shenker and B. Vöcking, *Randomized Rumor Spreading*, In Proceedings of the 41st IEEE Symposium on the Foundations of Computer Science, 2000, pp. 565–574.

[22] M. Mitzenmacher, *A brief history of generative models for power law and lognormal distributions*, Internet Mathematics, 1 (2004), pp. 226–251.

[23] D. Mosk-Aoyama and D. Shah, *Fast Distributed Algorithms for Computing Separable Functions*, IEEE Transactions on Information Theory, 54 (2008), pp. 2997–3007.

[24] M. E. J. Newman, *Power Laws, Pareto Distributions and Zipfs Law*, Contemporary Physics, 46 (2005), pp. 323–351.

[25] I. Norros and H. Reittu, *On a conditionally Poissonian graph process*, Adv. in Appl. Probab., 38 (2006), pp. 59–275.

[26] T. Sauerwald, *On mixing and edge expansion properties in randomized broadcasting*, Algorithmica, 56 (2010), pp. 51–88.

[27] B. Söderberg, *General formalism for inhomogeneous random graphs*, Phys. Rev. E (3), 66 (2002), 066121.

[28] E. Upfal and M. Mitzenmacher, *Probability and Computing: Randomized Algorithms and Probabilistic Analysis*, Cambridge University Press, 2005.

[29] R. van der Hofstad, *Random Graphs and Complex Networks*, 2011, (manuscript available at www.win.tue.nl/rhofstad/NotesRGCN2011.pdf).