Flooding in Weighted Random Graphs

Hamed Amini*   Moez Draief†   Marc Lelarge‡

Abstract

In this paper, we study the impact of 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 weighted flooding time defined as the minimum time needed to reach all nodes from one uniformly chosen node, and the weighted diameter corresponding to the largest distance between any pair of vertices. 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.

1   Introduction

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 [18].

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 there are other parameters, besides the network topology, to be taken into account such as the communication delays between nodes due for example 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 paper, 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 paper, explained in detail in Section 2, 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: initially, one of the nodes obtains some piece of information. Then every node which has the information passes it to another 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 paper, 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 3). To the best of our knowledge, our work is the first to study this model in an asyn-

---
*INRIA-ENS, Hamed.Amini@ens.fr
†Imperial College London, M.Draief@imperial.ac.uk
‡INRIA-ENS, Marc.Lelarge@ens.fr

chronous 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 connected random graphs. The analysis of the asymptotic of distances in edge weighted graphs has received much interest. In [13], Janson considered the special case of the complete graph with fairly general i.i.d. weights on edges, including the exponential distribution with parameter 1. It is shown that, when n goes to infinity, the asymptotic distance is $\log n / n$ for two given points, that the maximum if one point is fixed and the other varies is $2 \log n / n$, and the maximum over all pairs of points (i.e. the weighted diameter) is $3 \log n / n$. Janson 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. In particular, in [3, 20], the authors derive 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. More closely related to the present work, Bhamidi, van der Hofstad and Hooghiemstra [4] study first passage percolation on random graphs with finite average degree, minimum degree greater than 2 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 Section 2.2. We also explain in Section 4.3 why the analysis made in [4] is not sufficient to obtain results for the diameter and how we extend it.

The remainder of the paper is organized as follows. In Section 2, we define our model for the random graphs and state our main result for the weighted diameter and flooding time. We also compare it to existing works on distances in random graphs. In Section 3, we restrict ourselves to the important class of random regular graphs and analyze a model for asynchronous randomized broadcast. We also compare it to its more classical synchronized version. The main ideas of the proof are given in Section 4, while technical lemmas are deferred to the Appendix.

2   Models and Results

Given a finite connected graph $G = (V, E)$, the distance $\text{dist}(a, b)$ between two nodes a and b in V is the number of edges in the shortest path connecting these two vertices. The diameter of G, denoted by $\text{diam}(G)$, is the maximum graph distance between any pair of vertices in V, i.e.

$$(2.1) \quad \text{diam}(G) = \max\{\text{dist}(a, b),\, a, b \in V\}\,.$$

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

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

where a is chosen uniformly at random in V.

In this paper, we study the impact of the introduction of edge weights on the distances in the 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 [22, Chapter 16].

For a graph $G = (V, E)$, we assign to each edge $e \in E$ a weight w_e. For any $a, b \in V$, a path between a and b is a sequence $\pi = (e_1, e_2, \ldots e_k)$ 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 edge e_i belongs to the path π. For $a, b \in V$, we define

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

where $\Pi(a, b)$ denotes the set of all paths from a to b in the graph. The *weighted diameter* and the *weighted flooding time* are given by

$$
\begin{aligned}
\text{diam}_w(G) &= \max\{\text{dist}_w(a, b),\, a, b \in V\}, \quad \text{and,} \\
\text{flood}_w(G) &= \max\{\text{dist}_w(a, b),\, b \in V\},
\end{aligned}
$$

where in the definition of flooding, a is chosen uniformly at random in V.

Our main results consist of precise asymptotic expressions for the weighted diameter and weighted flooding time of sparse random graphs on n vertices with (i) a given degree sequence satisfying asymptotic properties similar to those imposed in [17], and (ii) i.i.d. exponentially distributed weights with parameter 1.

2.1   Configuration model. For $n \in \mathbb{N}$, let $(d_i)_1^n = (d_i^{(n)})_1^n$ be a sequence of non-negative integers such that $\sum_{i=1}^n d_i$ is even. For notational simplicity, we will usually not show the dependency on n explicitly. By means of the configuration model [5], 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 a uniform random matching. 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. 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 $G(n, (d_i)_1^n)$, [14].

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

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

From now on, we assume that the sequence $(d_i)_1^n$ satisfies the following regularity conditions analogous to the ones introduced in [17] and an additional constraint on the minimal degree.

CONDITION 1. *There exists a distribution $p = \{p_k\}_{k=0}^\infty$ such that*

(i) $u_r^{(n)}/n \to p_r$ for every $r \geq 0$ as $n \to \infty$;

(ii) $\lambda := \sum_r r p_r \in (0, \infty)$;

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

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

(v) $\min_{i=1...n} d_i = d_{\min} \geq 3$, and $p_{d_{\min}} > 0$.

We define $q = \{q_r\}_{r=0}^\infty$ the size-biased probability mass function corresponding to p, by

$$(2.2) \qquad q_r = \frac{(r+1)p_{r+1}}{\lambda},$$

and let ν denote its mean, i.e. $\nu = \sum_{r=0}^\infty r q_r \in (0, \infty)$.

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. [15, 17]). By our Condition 1 point (v), we have $\nu \geq 2$ and actually, the following lemma, proved in Section A.4, shows that under Condition 1 the constructed graph is connected with high probability (w.h.p.). We say that an event $\mathcal{E}_n$ holds w.h.p. if $\mathbb{P}(\mathcal{E}_n) \to 1$ when n tends to infinity.

LEMMA 2.1. *Consider a random graph $G(n, (d_i)_1^n)$ where the degree sequence $(d_i)_1^n$ satisfies Condition 1. Then $G(n, (d_i)_1^n)$ is connected w.h.p.*

2.2 Weighted flooding time and diameter. We now state our main first result concerning the weighted flooding time and diameter of $G(n, (d_i)_1^n)$ with i.i.d. exponential 1 weights on the graph edges.

THEOREM 2.1. *Consider a random graph $G(n, (d_i)_1^n)$ with i.i.d. exponential 1 weights on its edges, where the degree sequence $(d_i)_1^n$ satisfies Condition 1. Then we have*

$$\frac{diam_w(G(n, (d_i)_1^n))}{\log n} \xrightarrow{p} \frac{1}{\nu - 1} + \frac{2}{d_{\min}}, \quad and$$

$$\frac{flood_w(G(n, (d_i)_1^n))}{\log n} \xrightarrow{p} \frac{1}{\nu - 1} + \frac{1}{d_{\min}},$$

where $\xrightarrow{p}$ is the convergence in probability.

In the particular case where G is a random r-regular graph with $r \geq 3$, we recover a result first proved in [6] concerning the weighted diameter. We refer to [1] for an extension of this result when $d_{\min}$ can be smaller than 3.

In order to further compare our result with existing ones, we reproduce here a result of Bhamidi, van der Hofstad, Hooghiemstra [4]:

THEOREM 2.2. *([4]) Consider a random graph $G(n, (d_i)_1^n)$ where the degrees d_i satisfy Condition 1. Let a, b be two uniformly chosen vertices in this graph. Then there exists a random variable V such that*

$$(2.3) \qquad dist_w(a, b) - \frac{\log n}{\nu - 1} \xrightarrow{d} V,$$

where $\xrightarrow{d}$ is the convergence in distribution.

For completeness, we also include results of van der Hofstad, Hooghiemstra, Mieghem [21] and Fernholz, Ramachandran [8] for the typical distance and the diameter in $G(n, (d_i)_1^n)$:

THEOREM 2.3. *([21, 8]) Consider a random graph $G(n, (d_i)_1^n)$ where the degree sequence $(d_i)_1^n$ satisfies Condition 1. Let a, b be two uniformly chosen vertices in this graph. Then we have*

$$\frac{dist(a, b)}{\log n} \xrightarrow{p} \frac{1}{\log \nu}, \quad and \quad \frac{diam(G(n, (d_i)_1^n))}{\log n} \xrightarrow{p} \frac{1}{\log \nu}.$$

Our Theorem 2.1 is obviously consistent with the asymptotic for the typical weighted distance on random graphs given by (2.3). However contrary to the case without weight, we see that the asymptotics for the weighted flooding time or diameter are not the same as

for the typical weighted distance between two uniformly chosen vertices. The appearance of the common factor $\frac{\log n}{\nu-1}$ is quite easy to understand at an heuristic level: 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 (see [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 the size of the order of $\sqrt{n}$. In particular, if one considers two such exploration processes started from a and b, then by that time they should intersect with great probability. This explains why the typical weighted distance is of the order $\frac{\log n}{\nu-1}$. We give a more precise statement of this heuristic in Section 4.3. 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 $d_{\min}$ and large weights on all its incident edges. This event gives the contribution $\frac{\log n}{d_{\min}}$. We refer to Section 4.5 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 and then obtains the contribution $2\frac{\log n}{d_{\min}}$. We see that if one is interested in passing the information between two typical vertices, it can be achieved in time of the order $\frac{\log n}{\nu-1}$ and there is a price (in time) of $\frac{\log n}{d_{\min}}$ to pay if one wishes to pass the information to everyone from a typical vertex and another price (in time) of $\frac{\log n}{d_{\min}}$ to pay if one wishes to pass the information to everyone from a vertex in a worst case scenario. As shown by Theorem 2.3, there is no such discrepancy when there are no weights, i.e. if the vertices are synchronized so that the process can be run in slotted rounds. In such a case, the exploration process behaves like a standard Galton-Watson branching process and bad scenarios (corresponding to slow growth) have very low probability [9] so that they do not contribute in the large n limit.

3 Broadcasting in random regular graphs

In this section we elaborate on another aspect of our result. By comparing our main Theorem 2.1 with Theorem 2.3 in the case of random regular graphs, we see that the weighted flooding time is actually smaller than the graph distance flooding time. As for the diameter, this is also valid if $r \geq 6$. With previous discussion, the heuristic explanation of this phenomenon is clear: the

random weights introduce variance that allows for the branching process (approximating the exploration process) to grow faster than without weights. Even if the weights have an average of one, weights with small values allow the branching process to grow faster than with constant weight equal to one. Of course the variability of the weights has also a drawback when one looks at worst case scenario which correspond to the factors $\frac{\log n}{d_{\min}}$. 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 bigger than $d_{\min}$. We now concentrate on one important practical implication of this phenomenon.

We consider the asynchronous analogue of the standard *phone call model* [19]. 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. We consider the well-studied push model. In this model, if a node i holds the message it passes it to its randomly chosen neighbor regardless of its state. Note that this may yield an unnecessary transmission (if the receiver already had 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 $\text{ABT}(G)$ for asynchronous broadcast time with one initial informed node chosen uniformly at random among the vertices of G.

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. As shown in Section 4.6, the dynamic evolution of informed nodes then corresponds to the flooding time with i.i.d. weights on edges distributed according to an exponential distribution with mean r. The fact that the graph is regular, is crucial to get this property and this is the reason why we require G to be a r-regular graph. Hence, our Theorem 2.1 allows us to analyze the asynchronous broadcast algorithm for these graphs and we get the following corollary:

COROLLARY 3.1. *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 [12]. 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 [12] and Pittel [19] 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. [10] 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 [7] 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 [11] 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.
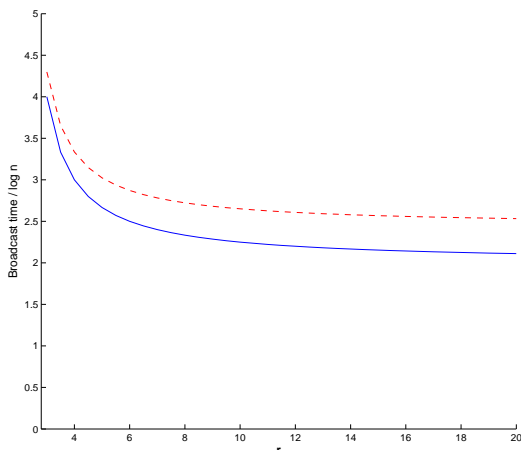

Figure 1: Comparison of the time to broadcast in the synchronized version (dashed) and with exponential random weights (plain)

Note that if instead of independent Poisson clocks of rate one, we take a deterministic process with slots of size one, our model is exactly the one studied in [11]. Hence locally, both processes behave similarly: when a node receives 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 1 shows the comparison between results in [11] and our Theorem 3.1: 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. This result while surprising is in agreement with the discussion comparing our Theorem 2.1 with Theorem 2.3. 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!

4   Proof of Theorem 2.1

In this section we present the main steps of the proof of Theorem 2.1. The main idea of the proof is to grow simultaneously balls from each vertex so that the diameter is twice the time when the last two balls intersect. Indeed, instead of taking a graph at random and then analyzing the balls, we use a standard coupling argument in random graphs theory consisting in building the balls and in the same time the graph. We present this coupling in the following Section 4.1 and Section 4.2. We then present the branching process approximation in Section 4.3. 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 the upper bound in Section 4.4 and the lower bound in Section 4.5.

4.1   The exploration process. For $t > 0$, we define the t-radius neighborhood of a vertex a as

$$B_w(a,t) \quad = \quad \{b,\ \text{dist}_w(a,b) \leq t\},$$

and the first time at which the set $B_w(a,t)$ reaches size k is denoted by:

$$T_a(k) = \min\{t : |B_w(a,t)| \geq k\}.$$

Fix a vertex a, and consider the 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 the process:

- Start with $B = \{a\}$, where a has d_a half-edges. Reveal any matchings and weights of these d_a half-edges connecting them amongst themselves (creating self-loops at a). The remaining unmatched half-edges are stored in a list L.

- 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 parameter ℓ. 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 match is also in B. More precisely, let d be the degree of this new vertex and $2x$ the number of matched half-edges in B (including the matched half-edges h_i and h). There is a total of $m - 2x$ unmatched half-edges. 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 same 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 that $B_w(a,t) = B(a,t)$ for all t.

Let τ_i denote the time of the i-th exploration step in the above continuous-time exploration process, i.e. $T_a(i) = \tau_{i-1}$. Assuming $L(a,\tau_i)$ is not empty, at time τ_{i+1}, we match a uniformly chosen half-edge from the set $L(a,\tau_i)$ to a uniformly chosen half-edge among all other half-edges, excluding those in $L(a,\tau_i)$. Let $\mathcal{F}_t$ be the σ-field generated by the above process until time t. Given $\mathcal{F}_{\tau_i}$, $\tau_{i+1} - \tau_i$ is an exponential random variable with rate $S_i(a) = |L(a,\tau_i)|$ the size of the list consisting of unmatched half-edges in $B(a,\tau_i)$. Let $i^* = \min\{i, \ S_i(a) = 0\} \leq n$, then we set $S_i(a) = 0$ for all $i^* \leq i \leq n$.

For $i \in [1, i^*]$, we define $\hat{d}_i$ the forward degree i.e. the degree minus one, of the vertex added during the i-th exploration step and let

$$(4.4) \qquad \hat{S}_i(a) := d_a + \hat{d}_1 + \ldots + \hat{d}_i - i.$$

For a connected set H, we denote by $\mathrm{tx}(H)$ the tree excess of H which is the maximum number of edges that can be deleted from the induced subgraph on H while still keeping it connected. Let $X_i(a) := \mathrm{tx}(B(a,\tau_i))$ for $i \leq i^*$, so that we have for $i \leq i^*$:

$$(4.5) \qquad S_i(a) = \hat{S}_i(a) - 2X_i(a).$$

Note that i^* and the sequence $\hat{d}_i$ depend on a.

We define

$$(4.6) \quad \alpha_n = \log^3 n \text{ and, } \beta_n = 3\sqrt{\frac{\lambda}{\nu - 1} n \log n}.$$

An important ingredient in the proof will be the comparison of the variables $\{\hat{d}_1, \ldots, \hat{d}_k\}$, for an appropriately chosen k, to an i.i.d. sequence.

4.2 Coupling the forward degrees sequence $\hat{d}_i$.

The sequence $(\hat{d}_i)_{i \leq i^*}$ can be constructed as follows. Initially, associate to vertex j a bin containing a set of d_j white balls. At step 0, color the balls corresponding to vertex a in red. Subsequently, at step $k \leq i^*$, choose a ball uniformly at random among all white balls; if the ball is drawn from node j's bin then set $\hat{d}_k = d_j - 1$, and color all the balls in the bin in red. If $i^* < n$, there are still white balls at step $i^* + 1$ and we complete the sequence $\hat{d}_i$ for $i \in [i^*+1, n]$ by continuing the sampling described above, so that we obtain a sequence $(\hat{d}_i)_{i=1}^n$ coinciding with the sequence defined in Section 4.1 for $i \leq i^*$. We also extend the sequence $\hat{S}_i(a)$ for $i > i^*$ thanks to (4.4) and we set $X_i(a) = X_{i^*}(a)$ for all $i > i^*$. Note that with these conventions, the relation (4.5) is not valid for $i > i^*$ but we have $S_i(a) \leq \hat{S}_i(a) - 2X_i(a)$.

We now present a coupling of the variables $\{\hat{d}_1, \ldots, \hat{d}_k\}$ valid for $k \leq \beta_n$ defined in (4.6), with an i.i.d. sequence, that we now define. First, we denote the order statistics of the degrees of the nodes of the graph by

$$d_{(1)} \leq d_{(2)} \leq \cdots \leq d_{(n)} .$$

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

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

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

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

Note that by Condition 1, we have $\beta_n \Delta_n = o(n)$ implying that both distributions $\underline{\pi}^{(n)}$ and $\bar{\pi}^{(n)}$ converge to size biased distribution q defined in (2.2).

For two real-valued random variables A and B, we write $A \leq_{st} B$ if for all x, we have $\mathbb{P}(A > x) \leq \mathbb{P}(B > x)$. If C is another random variable, we write $A \leq_{st} B|C$ if for all x, $\mathbb{P}(A > x) \leq \mathbb{P}(B > x|C)$ a.s.

LEMMA 4.1. *For an uniformly chosen vertex a and for $i \leq \beta_n$, we have*

$$(4.7) \quad \underline{D}_i^{(n)} \leq_{st} \hat{d}_i | (d_a, \hat{d}_1, \dots, \hat{d}_{i-1}) \leq_{st} \bar{D}_i^{(n)}$$

where $\underline{D}_i^{(n)}$ (resp. $\bar{D}_i^{(n)}$) are i.i.d. with distribution $\underline{\pi}^{(n)}$ (resp. $\bar{\pi}^{(n)}$). In particular, we have

$$\sum_{k=1}^i \underline{D}_k^{(n)} \leq_{st} \sum_{k=1}^i \hat{d}_k \leq_{st} \sum_{k=1}^i \bar{D}_k^{(n)}.$$

Proof. We fix the sequence of degrees $d_i^{(n)}$ and the initial vertex a. We now prove that conditionally on the values of $(d_a, \hat{d}_1, \dots, \hat{d}_{j-1})$, the random variable $\hat{d}_j$ is stochastically smaller than $\bar{D}_j^{(n)}$. This can be seen by a simple coupling argument as follows. First order the balls from 1 to m consistently with the order statistics, i.e. start by numbering the balls in the bin with the fewest balls and then move to the larger ones as ordered by the number of balls they contain.

Given the sequence $(d_a, \hat{d}_1, \dots, \hat{d}_{j-1})$, color in red the balls of bins of the corresponding sizes. In order to get a sample for $\bar{D}_j^{(n)}$, pick a ball at random among all balls in the last $n - (\beta_n + 1)\Delta_n$ bins and set $\bar{D}_j^{(n)}$ to be equal to the size of the selected bin minus one. If the ball picked is white set $\tilde{d}_j = \bar{D}_j^{(n)}$. If there are red balls in the last $n - (\beta_n + 1)\Delta_n$ bins and if such a ball is picked, say this is the ℓ-th ball among these red balls for the induced order, then set $\tilde{d}_j$ to be the size of the bin containing the ℓ-th white ball, minus one. Since $d_a + \hat{d}_1 + \cdots + \hat{d}_{j-1} \leq \beta_n \Delta_n$, this ball is in one of the first $(\beta_n + 1)\Delta_n$ bins. In all cases, we have: $\hat{d}_j \leq_{st} \tilde{d}_j \leq \bar{D}_j^{(n)}$ given the sequence $(d_a, \hat{d}_1, \dots, \hat{d}_{j-1})$. A similar argument allows to prove that $\underline{D}_j^{(n)} \leq_{st} \hat{d}_j$ given the sequence $(d_a, \hat{d}_1, \dots, \hat{d}_{j-1})$.

The second statement follows from the following lemma [8, Lemma A.3] .

LEMMA 4.2. *Let $X_1, \dots, X_t$ be a random process adapted to a filtration $\mathcal{F}_0 = \sigma[\emptyset], \mathcal{F}_1, \dots, \mathcal{F}_t$, and let $\Sigma_t = X_1 + \dots + X_t$. 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 Σ_t is stochastically greater (resp. smaller) than the sum of t i.i.d. μ-distributed random variables.*

4.3 Branching process approximation. In this section, we consider the continuous-time Markovian branching process Z_t approximating the exploration process defined in Section 4.1. As we will see, the tree excess $X_i(a)$ remains small (compared to $\hat{S}_i(a)$) with very high probability at least when i is not too large.

The branching process approximation consists in neglecting this term and considering that the sequence of $\hat{d}$ is a sequence of i.i.d. random variables with distribution given by (2.2) (which is true asymptotically). Z_t is started with one ancestor. Each of the members of the population has an exponential lifetime (with mean one) and upon her death she gives birth to a random number $\hat{D}$ of particles, where $\hat{D}$ has distribution (2.2). We assume that $d_{\min} \geq 3$ so that $\hat{D} \geq 2$ a.s. We now define the split times: the times at which the particles split (see [2] III.9). Let $\hat{\Sigma}_i = D + \sum_{j=1}^i \hat{D}_j - i$ (where D is the degree of the vertex the process starts from, i.e. $\mathbb{P}(D = r) = p_r$) and E_i a sequence of independent exponential random variables with mean one. Under the assumption $d_{\min} \geq 3$, the split times are defined by $T_0 = 0$ and for $i \geq 1$,

$$(4.8) \qquad T_i = \sum_{j=0}^{i-1} \frac{E_i}{\hat{\Sigma}_i}.$$

Note that $\mathbb{E}[\hat{\Sigma}_i] \approx (\nu - 1)i$ and it is shown in [2] that: $\lim_{n\to\infty} \frac{T_n}{\log n} = \frac{1}{\nu-1}$. In particular at time $\frac{\log n}{2(\nu-1)}$, the process reaches size $\sqrt{n}$, so that for two given vertices of the graph, there is a high probability that the two balls intersect (see Proposition 4.1 below). This heuristic argument allows to understand the typical distance given by (2.3) in [4].

In order to be able to compute the diameter, we need to find x such that $\mathbb{P}(T_{\sqrt{n+}} \geq x) \approx \frac{1}{n}$ (here $n+$ is an informal notation to denote a sequence growing slightly faster than n, like $n \log n$, see the parameter β_n). Hence we need to study large deviations results for the sequence $\{T_i\}_{i\in\mathbb{N}}$. Moreover we have to take care of the error introduced by the branching process approximation. This is done by a coupling argument given in Section 4.2. We now give the main technical steps of the proof.

4.4 Proof of the upper bound. The following proposition, proved in Section A.2, shows that to bound the distance between any pair of nodes it suffices to bound the time when $\beta_n := 3\sqrt{\frac{\lambda}{\nu-1} n \log n}$ nodes have been reached in the exploration process defined in Section 4.1 starting from these two nodes.

PROPOSITION 4.1. *We have w.h.p.*

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

Now we give an upper bound (which holds w.h.p.) for the time needed to explore β_n nodes from any vertex a.

PROPOSITION 4.2. *For a uniformly chosen vertex u and any $\epsilon > 0$, we have*

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

A proof of this proposition is given in Section A.1. We give here a heuristic based on the branching process approximation defined in the last section. Note that for T_i defined by (4.8), we have

$$\mathbb{E}\left[e^{\theta T_n}|\hat{\Sigma}\right] = \prod_{i=0}^{n-1}\left(1 + \frac{\theta}{\hat{\Sigma}_i - \theta}\right) \leq \exp\left(\sum_{i=0}^{n-1} \frac{\theta}{\hat{\Sigma}_i - \theta}\right).$$

Then for small values of $i \leq \alpha_n$, where $\alpha_n = \log^3 n$ grows slower than any power of n, we use the lower bound:

$$\hat{\Sigma}_i \geq i(d_{\min} - 2) + d_{\min},$$

so that we get with $\theta = d_{\min}$:

$$\mathbb{P}(T_{\alpha_n} \geq x\log n) \leq \mathbb{E}\left[e^{d_{\min}T_{\alpha_n}}\right]\exp(-xd_{\min}\log n)$$
$$\approx \alpha_n^{\frac{d_{\min}}{d_{\min}-2}} n^{-xd_{\min}}.$$

Now for larger values of i, we can use the approximation $\hat{\Sigma}_i \geq i(\nu - 1 - \epsilon)$ so that we get

$$\mathbb{P}(T_{\beta_n} - T_{\alpha_n} \geq y\log n) \leq \mathbb{E}\left[e^{\theta(T_{\beta_n} - T_{\alpha_n})}\right]\exp(-\theta y \log n)$$
$$\approx \left(\frac{\beta_n}{\alpha_n}\right)^{\frac{\theta}{\nu-1-\epsilon}} n^{-y\theta} \approx n^{\theta\left(\frac{1}{2(\nu-1-\epsilon)}-y\right)},$$

so if we choose $y \approx \frac{1}{2(\nu-1)}$, we get

$$\mathbb{P}\left(T_{\beta_n} - T_{\alpha_n} \geq \left(\frac{1}{2(\nu-1)}\right)\log n\right) \approx n^{-1}.$$

We refer to Section A.1 for the technical details.
By Proposition 4.2, we have for a uniformly chosen vertex u and any $\epsilon > 0$:

$$\mathbb{P}\left(T_u(\beta_n) \geq \left(\frac{1+\epsilon}{2(\nu-1)}\right)\log n\right) = o(1), \text{ and}$$

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

By taking a union bound over u, it follows that

$$\mathbb{P}\left(T_u(\beta_n) \leq \left(\frac{1+\epsilon}{2(\nu-1)} + \frac{1+\epsilon}{d_{\min}}\right)\log n, \text{ for all } u\right)$$
$$= 1 - o(1).$$

Combining this with Proposition 4.1 finishes the proof of the upper bound.

4.5 Proof of the lower bound. To prove the lower bound, it suffices to prove that for any $\epsilon > 0$, there exists w.h.p. two vertices u and v such that

$$\text{dist}_w(u,v) \geq \left(\frac{1}{\nu-1} + \frac{2}{d_{\min}}\right)(1-\epsilon)\log n.$$

Let $G_n \sim G(n, (d_i)_1^n)$, and

$$B'_w(u,t) := \{v : \text{dist}_w(N(u), v) \leq t\},$$

where $N(u)$ denotes the neighbors of u in G_n. Let $V_{d_{\min}}$ be the set of vertices with degree $d_{\min}$. The following proposition, proved in Section A.3, gives a lower bound for the distance between the neighbors of a uniformly chosen vertex $u \in V_{d_{\min}}$ from the neighbors of another uniformly chosen vertex $v \in V_{d_{\min}}$.

PROPOSITION 4.3. *Let u, v be two uniformly chosen vertices of the graph G_n, with degree $d_{\min}$, i.e. $u, v \in V_{d_{\min}}$, and let $t_n = \frac{(1-\epsilon)\log n}{2(\nu-1)}$. We have w.h.p.*

$$B'_w(u,t_n) \cap B'_w(v,t_n) = \emptyset.$$

Now let $s_n := \frac{1-\epsilon}{d_{\min}}\log n$, and call a vertex in $V_{d_{\min}}$ *bad* if the weights on all the $d_{\min}$ edges connected to it are larger than s_n. Let A_u denote the event that u is bad, and let $Y := \sum_u \mathbf{1}_{A_u}$ be the number of bad vertices. Then we have for $u \in V_{d_{\min}}$:

$$\mathbb{P}(A_u) = \mathbb{P}(\text{Exp}(d_{min}) \geq s_n) = n^{-1+\epsilon}.$$

Then it is easy to see that

$$\mathbb{E}(Y) = \sum_u \mathbb{P}(A_u) = p_{d_{\min}}(1 + o(1))n^{\epsilon}, \text{ and}$$

$$\text{Var}(Y) = \sum_{u,v\in V_{d_{\min}}} \text{Cov}\left(\mathbf{1}_{A_u}, \mathbf{1}_{A_v}\right)$$
$$= \sum_{u\in V_{d_{\min}}} \text{Var}\left(\mathbf{1}_{A_u}\right) + \sum_{uv\in E} \text{Cov}\left(\mathbf{1}_{A_u}, \mathbf{1}_{A_v}\right)$$
$$\leq (d_{\min} + 1)\mathbb{E}(Y).$$

Then by Chebyshev's inequality, we have

$$Y \geq \frac{2}{3}p_{d_{\min}}n^{\epsilon}$$

with high probability.

Let Y' denote the number of bad vertices that are of distance at most $s_n + \frac{(1-\epsilon)}{\nu-1}\log n$ from vertex a (chosen uniformly). By Proposition 4.3 for a uniformly chosen vertex i we have w.h.p. $B'_w(a,t_n) \cap B'_w(i,t_n) = \emptyset$. In particular, conditioning on the event A_i, the probability that $B'_w(i,t_n)$ does not intersect $B'_w(a,t_n)$ remains the same. Hence, for a uniformly chosen vertex i we have

$$\mathbb{P}\left(A_i, B'_w(a,t_n) \cap B'_w(i,t_n) \neq \emptyset\right) = o(\mathbb{P}(A_i)),$$

and then we deduce

$$\mathbb{E}(Y') = o(\mathbb{E}(Y)) = o(\log n).$$

By Markov's inequality, $Y' \leq \frac{1}{3}p_{d_{\min}}n^\epsilon$ w.h.p., and hence $Y - Y'$ is w.h.p. positive. This implies the existence of a vertex i whose distance from a is at least $\left(\frac{1}{\nu-1} + \frac{1}{d_{\min}}\right)(1-\epsilon)\log n$. Then for any $\epsilon > 0$ we have w.h.p.

$$\begin{aligned}
\text{flood}_w(G_n) &:= \max_{1 \leq i \leq n} \text{dist}_w(a, i) \\
&\geq \left(\frac{1}{\nu-1} + \frac{1}{d_{\min}}\right)(1-\epsilon)\log n.
\end{aligned}$$

Let R denote the number of pairs of distinct bad vertices. Then $Y \geq \frac{2}{3}p_{d_{\min}}n^\epsilon$ gives

$$R \geq \frac{1}{4}p_{d_{\min}}^2 n^{2\epsilon}$$

w.h.p. By Proposition 4.3 for two uniformly chosen vertices u, v we have w.h.p.

$$B'_w(u, t_n) \cap B'_w(v, t_n) = \emptyset.$$

In particular, conditioning on the events A_u, and A_v, the probability that $B'_w(u, t_n)$ does not intersect $B'_w(v, t_n)$ remains the same. Hence, for two uniformly chosen vertices u, v we have,

$$\mathbb{P}\left(A_u, A_v, B'_w(u, t_n) \cap B'_w(v, t_n) \neq \emptyset\right) = o(\mathbb{P}(A_u, A_v)).$$

Let R' denote the number of pairs of bad vertices that are of distance at most $2s_n + \frac{(1-\epsilon)}{\nu-1}\log n$. Then we have

$$\mathbb{E}R' = o(\mathbb{E}Y^2) = o(n^{2\epsilon}).$$

By Markov's inequality, $R' \leq \frac{1}{6}p_{d_{\min}}^2 n^{2\epsilon}$ w.h.p, and hence $R - R'$ is w.h.p positive. This implies that for any $\epsilon > 0$ we have w.h.p.

$$\begin{aligned}
\text{diam}_w(G_n) &:= \max_{1 \leq i,j \leq n} \text{dist}_w(i, j) \\
&\geq \left(\frac{1}{\nu-1} + \frac{2}{d_{\min}}\right)(1-\epsilon)\log n,
\end{aligned}$$

which completes the proof.

4.6   Proof of Corollary 3.1. First we prove that for a 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 1 corresponds exactly to the flooding time with exponential random weights on edges with mean r. Let $\mathcal{I}(a, t)$ denote the set of informed nodes at time t when we start the broadcast process from node $a \in V$.

Indeed we show that random map $\mathcal{I}(a, .)$ from $[0, \infty)$ to subsets of V has the same law as $B_w(a, .)$ when the weights are exponential with mean r using a coupling argument: 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.

Let $T(v)$ denote the time at which node v becomes informed in the asynchronous broadcast model and let $\tau_i(u, v)$, for $(u, v) \in E$, denote the i-th time that node u contacts node v. Now we define the weight of the edge $e = (u, v)$ as follows:

- if $T(u) \leq T(v)$ then

$$w_e := \min_i\{ \tau_i(u, v) - T(u) \mid \tau_i(u, v) > T(u) \}.$$

- if $T(u) > T(v)$ then

$$w_e := \min_i\{ \tau_i(v, u) - T(v) \mid \tau_i(v, u) > T(v) \}.$$

Thanks to the memoryless property of the Poisson process, $\{w_e, e \in E\}$ are independent exponential random variables with mean r and are such 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 and it is easy to conclude the proof by Theorem 2.1, that is w.h.p.,

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

References

[1] H. Amini and M. Lelarge. The diameter of weighted random graphs. *in preparation.*

[2] K. B. Athreya and P. E. Ney. *Branching processes.* Dover publications, 2004.

[3] S. Bhamidi. First passage percolation on locally treelike networks. I. dense random graphs. *Journal of Mathematical Physics,* 49(12):125218, 2008.

[4] 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.

[5] B. Bollobás. *Random Graphs.* Cambridge University Press, 2001.

[6] J. Ding, J. H. Kim, E. Lubetzky, and Y. Peres. Diameters in supercritical random graphs via first passage percolation. *http://arxiv.org/abs/0906.1840,* 2009.

[7] R. Elsässer and T. Sauerwald. On the runtime and robustness of randomized broadcasting. *Theoretical Computer Science*, 410(36):3414–3427, 2009.

[8] D. Fernholz and V. Ramachandran. The diameter of sparse random graphs. *Random Structures and Algorithms*, 31(4):482–516, 2007.

[9] K. Fleischmann and V. Wachtel. Lower deviation probabilities for supercritical Galton-Watson processes. *Ann. Inst. H. Poincaré Probab. Statist.*, 43(2):233–255, 2007.

[10] N. Fountoulakis, A. Huber, and K. Panagiotou. Reliable broadcasting in random networks and the effect of density. *in Proc. Infocom*, 2010.

[11] N. Fountoulakis and K. Panagiotou. Rumor spreading on random regular graphs and expanders. manuscript available at www.mpi-inf.mpg.de/ fountoul/RRRS.pdf, 2010.

[12] A. Frieze and G. Grimmett. The shortest path problem for graphs with random arc-lengths. *Discrete Applied Mathematics*, 10:57–77, 1985.

[13] 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.

[14] S. Janson. The probability that a random multigraph is simple. *Combinatorics, Probability and Computing*, 18(1-2):205–225, 2009.

[15] S. Janson and M. J. Luczak. A new approach to the giant component problem. *Random Structures and Algorithms*, 34(2):197–216, 2009.

[16] A. Klenke and L. Mattner. Stochastic ordering of classical discrete distributions. *To appear in Advances in Applied probability*, 2010.

[17] 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.

[18] D. Peleg. Distributed computing: A locality-sensitive approach. *SIAM Monographs on Discrete Mathematics and Applications*, 5.

[19] B. Pittel. On spreading a rumor. *SIAM Journal on Applied Mathematics*, 1:213–223, 1987.

[20] R. van der Hofstad, G. Hooghiemstra, and P. V. Mieghem. The flooding time in the random graph. *Extremes*, 5(2):111–129, 2002.

[21] R. van der Hofstad, G. Hooghiemstra, and P. V. Mieghem. Distances in random graphs with finite variance degrees. *Random Structures and Algorithms*, 27(1):76–123, 2005.

[22] P. van Mieghem. *Performance Analysis of Communications Networks and Systems*. Cambridge University Press, 2006.

A Proofs

Recall that $\beta_n = 3\sqrt{\frac{\lambda}{\nu-1}n\log n}$, and $\alpha_n = \log^3 n$. Now we consider the exploration process defined in Section 4.1.

A.1 Proof of Proposition 4.2. Before getting to the main part of the proof, we need to prove some technical lemmas.

We start by some simple remarks. The process $i \mapsto X_i(a)$ is non-decreasing in $i \in [1, n]$. Moreover, given $\mathcal{F}_i$, the increment $X_{i+1}(a) - X_i(a)$ is stochastically dominated by a binomial variable:

$$(A.1) \quad \mathrm{Bin}\left(\hat{d}_{i+1}, \frac{(S_i(a)-1)^+}{m^{(n)} - 2(X_i(a)+i)}\right).$$

Note that if $i > i^*$, then $S_i(a) = 0$ and $X_{i+1}(a) - X_i(a) = 0$ so that (A.1) is still valid. For $i < n/2$, we have

$$\frac{(S_i(a)-1)^+}{m^{(n)} - 2(X_i(a)+i)} \leq \frac{\hat{S}_i(a) - 2X_i(a)}{m^{(n)} - 2(X_i(a)+i)}$$
$$\leq \frac{\hat{S}_i(a)}{m^{(n)} - 2i}$$
$$\leq \frac{\max_{\ell \leq i}\hat{S}_\ell(a)}{n - 2i}.$$

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

$$(A.2)\ \ X_i(a) \leq_{st} \mathrm{Bin}\left(\max_{\ell \leq i}\hat{S}_\ell(a) + i, \frac{\max_{\ell \leq i}\hat{S}_\ell(a)}{n - 2i}\right).$$

LEMMA A.1. *For any $k < n/2$, we have*

$$\mathbb{P}\left(2X_k(a) \geq x \mid \hat{S}_k(a), i^* \geq k\right) \leq$$
$$\mathbb{P}\left(\mathrm{Bin}\left(\hat{S}_k(a), \sqrt{\hat{S}_k(a)/n}\right) \geq x \mid \hat{S}_k(a)\right).$$

Proof. Note that given a set of k edges connecting the $k + 1$ vertices in $B_w(a, T_a(k+1))$ and given $\hat{S}_k(a)$, the remaining edges of the graph are obtained by an uniform matching of the remaining half-edges; the total remaining number of half-edges is $m^{(n)} - 2k$ (which is greater than n by $d_{\min} \geq 3$) and the number of unmatched half-edges in $B_w(a, T_a(k+1))$ is exatly $\hat{S}_k(a)$.

The proof follows from the following lemma proved in [8, Lemma 3.2]:

LEMMA A.2. *Let A be a set of m points, i.e. $|A| = m$, and let F be a uniform random matching of elements of A. For $a \in A$, we denote by $F(a)$ the point matched to a, and similarly for $B \subset A$, we write $F(B)$ for the set of points matched to B. Now assume $B \subset A$, and let $k = |B|$. We have*

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

We define the events :

$$R(a) := \{S_k(a) \geq d_{\min} + (d_{\min} - 2)k, \text{ for all } 1 \leq k \leq \alpha_n\},$$
$$R'(a) := \{S_k(a) \geq 1 + (d_{\min} - 2)k, \text{ for all } 1 \leq k \leq \alpha_n\}.$$

LEMMA A.3. *For a uniformly chosen vertex a, we have*

$$(A.3) \qquad \mathbb{P}(R(a)) \geq 1 - o((\log^{10} n / n)),$$
$$(A.4) \qquad \mathbb{P}(R'(a)) \geq 1 - o(n^{-3/2}).$$

Proof. Note that Since $d_{\min} \geq 3$, the sequence $\hat{S}_k(a)$ is non-decreasing in k. We also have for all $k \leq \alpha_n$,

$$d_{\min} + (d_{\min} - 2)k \leq \hat{S}_k(a) \leq \alpha_n \Delta_n = o(n).$$

Hence we have

$$\{X_k(a) = 0\} \subset R(a), \quad \{X_k(a) \leq 1\} \subset R'(a).$$

We distinguish two cases:

- **Case 1:** $\hat{S}_{\alpha_n}(a) < 2d_{\min}\alpha_n$. Then given this event denoted by Q_1, by (A.2), we have

$$X_{\alpha_n}(a) \leq_{st} \text{Bin}\left((2d_{\min} + 1)\alpha_n, \frac{2d_{\min}\alpha_n}{n - 2\alpha_n}\right).$$

Hence we have

$$\mathbb{P}\left(X_{\alpha_n}(a) \geq 1 \mid Q_1\right) \leq O(\alpha_n^2/n),$$

and

$$\mathbb{P}\left(X_{\alpha_n}(a) \geq 2 \mid Q_1\right) \leq O(\alpha_n^4/n^2).$$

Then we have

$$\mathbb{P}(R(a)^c \mid Q_1) \leq O(\alpha_n^2/n),$$
$$\mathbb{P}(R'(a)^c \mid Q_1) \leq O(\alpha_n^4/n^2).$$

- **Case 2.** if $\hat{S}_{\alpha_n}(a) \geq 2d_{\min}\alpha_n$, we have

$$\hat{S}_{\alpha_n}(a) \leq \alpha_n \Delta_n = o(n).$$

Let $k = \max\{i, \hat{S}_i(a) < 2d_{\min}\alpha_n\}$. Then on the event Q_1^c, we have $k \leq \alpha_n$ and by a similar argument as in case 1, given the event Q_1^c, we obtain

$$X_k(a) \leq_{st} \text{Bin}\left((2d_{\min} + 1)\alpha_n, \frac{2d_{\min}\alpha_n}{n - 2\alpha_n}\right).$$

Then given the event Q_1^c, we have by (A.2):

$$X_{\alpha_n}(a) \leq_{st} \text{Bin}\left(\alpha_n(\Delta_n + 1), \frac{\alpha_n \Delta_n}{n - 2\alpha_n}\right).$$

Then letting $M = \lceil 2\tau^{-1} \rceil$, and by Chernoff's inequality we have

$$\mathbb{P}\left(X_{\alpha_n}(a) \geq M \mid Q_1^c\right) \leq O\left((\Delta_n^2 \alpha_n^2/n)^M\right) = o(n^{-3}).$$

Note that for n large enough

$$2d_{\min}\alpha_n - 2M > d_{\min} + (d_{\min} - 2)\alpha_n.$$

Hence we have

$$\{X_k(a) = 0, X_{\alpha_n}(a) \leq M, Q_1^c\} \subset R(a) \cap Q_1^c,$$
$$\{X_k(a) \leq 1, X_{\alpha_n}(a) \leq M, Q_1^c\} \subset R'(a) \cap Q_1^c$$

Then we have

$$\mathbb{P}\left(R(a)^c \mid Q_1^c\right) \leq O(\alpha_n^2/n),$$
$$\mathbb{P}\left(R'(a)^c \mid Q_1^c\right) \leq O(\alpha_n^4/n^2).$$

The lemma follows.

We will use the following properties: if Y is an exponential random variable with mean μ^{-1}, then for any $\theta < \mu$, we have $\mathbb{E}\left[e^{\theta Y}\right] = \frac{\mu}{\mu - \theta}$. Given the sequence $S_k(a)$, for $k < i^*$, the random variables $\tau_{k+1} - \tau_k$ are iid exponential random variables with mean $S_k(a)^{-1}$.

LEMMA A.4. *For a uniformly chosen vertex a, and any $\epsilon > 0$, we have*

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

Proof. Assume that $R'(a)$ holds and consider the following two cases:
Case 1: $R(a)$ **holds.**
In this case, for any $k < \alpha_n$ we have conditioning on $R(a) \cap R'(a)$,

$$\tau_{k+1} - \tau_k \leq_{st} Y_k = \text{Exp}\left(d_{\min} + (d_{\min} - 2)k\right),$$

and all Y_k's are independent. Hence, we have

$$\mathbb{E}\left[e^{d_{\min}(T_a(\alpha_n) - T_a(1))} \mid R(a) \cap R'(a)\right]$$
$$\leq \prod_{i=1}^{\alpha_n - 1}\left(1 + \frac{d_{\min}}{(d_{\min} - 2)i}\right)$$
$$\leq \exp\left[\frac{d_{\min}}{d_{\min} - 2}\sum_{i=1}^{\alpha_n - 1}\frac{1}{i}\right]$$
$$\leq \alpha_n^{d_{\min}} = (\log n)^{3d_{\min}},$$

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

$$\mathbb{P}\left(T_a(\alpha_n) - T_a(1) \geq \epsilon \log n + l \mid R(a) \cap R'(a)\right)$$
$$\leq (\log n)^{3d_{\min}}\exp(-d_{\min}(\epsilon \log n + l))$$
$$= \frac{(\log n)^{3d_{\min}}}{n^{\epsilon d_{\min}}}e^{-d_{\min}l} = o(e^{-d_{\min}l}).$$

We also have $T_a(1) \leq_{st} \text{Exp}(d_{\min})$; hence

$$\mathbb{P}(T_a(1) \geq \epsilon \log n + l) \leq \frac{\exp(-ld_{\min})}{n^{\epsilon d_{\min}}},$$

and we have

$$\mathbb{P}\left(T_a(\alpha_n) \geq \epsilon \log n + l \mid R(a) \cap R'(a)\right) = o(e^{-d_{\min}l}).$$

Case 2: $R(a)$ **doesn't hold.**
In this case, for any $k < \alpha_n$ we have conditioning on $R(a)^c \cap R'(a)$,

$$\tau_{k+1} - \tau_k \leq_{st} Y_k = \text{Exp}(1 + (d_{\min} - 2)k),$$

and all the Y_k's are independent. We have

$$\mathbb{E}\left[e^{(T_a(\alpha_n) - T_a(1))} \mid R(a)^c \cap R'(a) \right]$$
$$\leq \prod_{i=1}^{\alpha_n - 1} \left(1 + \frac{1}{(d_{\min} - 2)i} \right)$$
$$\leq \exp\left[\frac{1}{d_{\min} - 2} \sum_{i=2}^{\alpha_n - 1} \frac{1}{i} \right]$$
$$\leq \alpha_n = \log^3 n,$$

for n large enough. Again by Markov's inequality, we have

$$\mathbb{P}\left(T_a(\alpha_n) - T_a(1) \geq \epsilon \log n + l \mid R(a)^c \cap R'(a)\right)$$
$$\leq \log^3 n \exp(-\epsilon \log n - l) = o(n^{-\epsilon/4}).$$

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

$$\mathbb{P}\left(T_a(\alpha_n) \geq \epsilon \log n + l \mid R(a)^c \cap R'(a)\right) = o(n^{-\epsilon/4}).$$

Putting all these together we have

$$\mathbb{P}\left(T_a(\alpha_n) \geq \epsilon \log n + l\right) \leq 1 - \mathbb{P}(R'(a)) +$$
$$(1 - \mathbb{P}(R(a)))n^{-\epsilon/4} + o(e^{-d_{\min}l})$$
$$\leq o(n^{-1} + e^{-d_{\min}l}),$$

as desired.

We continue with a simple large deviation estimate.

LEMMA A.5. *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*

$$(A.5) \quad \mathbb{P}\left(\underline{D}_1^{(n)} + \cdots + \underline{D}_k^{(n)} \leq k\eta\right) \leq e^{-\gamma k}.$$

Proof. Let D^* be a random variable with distribution $\mathbb{P}(D^* = k) = q_k$ given in (2.2) 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 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 $\underline{\phi}^{(n)}(\theta) = \mathbb{E}[e^{-\theta \underline{D}_1^{(n)}}]$. We have for $\theta > 0$,

$$\mathbb{P}\left(\underline{D}_1^{(n)} + \cdots + \underline{D}_k^{(n)} \leq \eta k\right) \leq \exp\left(k\left(\theta\eta + \log \underline{\phi}^{(n)}(\theta)\right)\right).$$

Fix $\theta < \theta_0$ and let n be sufficiently large so that $\log \underline{\phi}^{(n)}(\theta) \leq \log \phi(\theta) + \epsilon$. This yields

$$\mathbb{P}\left(\underline{D}_1^{(n)} + \cdots + \underline{D}_k^{(n)} \leq \eta k\right) \leq \exp\left(k\left(\theta\eta + \log \phi(\theta) + \epsilon\theta\right)\right)$$
$$\leq \exp\left(k\theta\left(\eta - \nu + 2\epsilon\right)\right),$$

which concludes the proof.

We now prove:

LEMMA A.6. *For any* $\epsilon > 0$, *we define the event*

$$R''(a) := \left\{ S_k(a) \geq \frac{\nu - 1}{1 + \epsilon} k, \text{ for all } \alpha_n \leq k \leq \beta_n \right\}.$$

For a uniformly chosen vertex a, *we have*

$$\mathbb{P}(R''(a)) \geq 1 - o(n^{-3/2}).$$

To prove this, we need the following intermediate result proved in [16, Theorem 1]:

LEMMA A.7. *Let* $n_1, n_2 \in \mathbb{N}$ *and* $p_1, p_2 \in (0, 1)$. *We have* $\text{Bin}(n_1, p_1) \leq_{st} \text{Bin}(n_2, p_2)$ *if and only if the following conditions hold*

(i) $n_1 \leq n_2$,

(ii) $(1 - p_1)^{n_1} \geq (1 - p_2)^{n_2}$.

In particular, we have

COROLLARY A.1. *If* $x \leq y = o(n)$, *we have*

$$x - \text{Bin}(x, \sqrt{x/n}) \leq_{st} y - \text{Bin}(y, \sqrt{y/n}).$$

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 A.6.

Proof. [Proof of Lemma A.6] Note that by Lemma A.3 we have

$$\mathbb{P}(i^* \geq \alpha_n) \geq \mathbb{P}(R'(a)) \geq 1 - o(n^{-3/2}).$$

Then we get

$$\mathbb{P}(R''(a)) \geq 1 - \mathbb{P}(i^* < \alpha_n) - \mathbb{P}(R''(a)^c,\ i^* \geq \alpha_n)$$
$$\geq 1 - o(n^{-3/2}) - \mathbb{P}(R''(a)^c \mid i^* \geq \alpha_n).$$

Then to prove Lemma A.6 it suffices to prove

$$\mathbb{P}(R''(a) \mid i^* \geq \alpha_n) \geq 1 - o(n^{-3/2}).$$

By Lemmas 4.1 and A.5, for any $\epsilon > 0$, $k \geq \alpha_n$ and n large enough, we have

$$\mathbb{P}\left(\hat{d}_1 + ... + \hat{d}_k \leq \frac{\nu}{1+\epsilon/2}k\right) \leq e^{-\gamma k} = o(n^{-6}).$$

Then with probability at least $1 - o(n^{-6})$, for any $k \leq \beta_n$,

$$d_{\min} + \frac{\nu-1}{1+\epsilon/2}k < d_a + \hat{d}_1 + ... + \hat{d}_k - k < (k+1)\Delta_n = o(n).$$

By the union bound on k, with probability at least $1 - o(n^{-5})$, we have for all $\alpha_n \leq k \leq \beta_n$,

$$(A.6)\quad d_{\min} + \frac{\nu-1}{1+\epsilon/2}k < \hat{S}_k(a) < (k+1)\Delta_n = o(n).$$

Then in the remaining of the proof we can assume that the above condition is satisfied.

By Lemma A.1, Corollary A.1 and (A.6), conditioning on $\hat{S}_k(a)$ and $\{i^* \geq k\}$, we have:

$$S_k(a) \geq_{st} d_{\min} + \frac{\nu-1}{1+\epsilon/2}k -$$
$$\mathrm{Bin}\left(d_{\min} + \frac{\nu-1}{1+\epsilon}k, \sqrt{\left(d_{\min} + \frac{\nu-1}{1+\epsilon}k\right)/n}\right)$$
$$\geq_{st} d_{\min} + \frac{\nu-1}{1+\epsilon/2}k - \mathrm{Bin}\left(\nu k, \sqrt{\nu k/n}\right).$$

Corollary A.1 and (A.6) imply that

$$S_k(a) \mid \{i^* \geq k\} \geq_{st} d_{\min} + \frac{\nu-1}{1+\epsilon}k -$$
$$\mathrm{Bin}\left(d_{\min} + \frac{\nu-1}{1+\epsilon}k, \sqrt{\left(d_{\min} + \frac{\nu-1}{1+\epsilon}k\right)/n}\right)$$
$$\geq_{st} d_{\min} + \frac{\nu-1}{1+\epsilon}k - \mathrm{Bin}\left(\nu k, \sqrt{\nu k/n}\right).$$

By Chernoff's inequality, and as $k\sqrt{k/n} = o(k/\sqrt{\alpha_n})$, we have

$$\mathbb{P}(\mathrm{Bin}(\nu k, \sqrt{\nu k/n}) \geq k/\sqrt{\alpha_n}) \leq \exp\left(-\frac{1}{3}k/\sqrt{\alpha_n}\right)$$
$$= o(n^{-6}).$$

Moreover, with probability at least $1 - o(n^{-6})$, conditioned on $\{i^* \geq k\}$, we have

$$S_k(a) \geq d_{\min} + \frac{\nu-1}{1+\epsilon}k - \frac{k}{\sqrt{\alpha_n}} \geq \frac{\nu-1}{1+2\epsilon}k,$$

for n large enough. Then letting

$$R_k'' = \{S_k(a) \geq \frac{\nu-1}{1+2\epsilon}k\},$$

we have

$$(A.7)\qquad \mathbb{P}(R_k'' \mid i^* \geq k) \geq 1 - o(n^{-6}).$$

Then $R''(a) = \bigcap_{k=\alpha_n}^{\beta_n} R_k''(a)$, and by using the fact that $R_{k-1}''(a) \subset \{i^* \geq k\}$ we get

$$\mathbb{P}(R''(a)) \mid i^* \geq \alpha_n) = 1 - \mathbb{P}\left(\bigcup_{k=\alpha_n}^{\beta_n} R_k''(a)^c \mid i^* \geq \alpha_n\right)$$
$$= 1 - \mathbb{P}\left(R_{\alpha_n}''(a)^c \bigcup_{k=\alpha_n+1}^{\beta_n} \left(R_k''(a)^c \cap R_{k-1}''(a)\right) \mid i^* \geq \alpha_n\right)$$
$$\geq 1 - \mathbb{P}\left(R_{\alpha_n}''(a)^c \bigcup_{k=\alpha_n+1}^{\beta_n} \left(R_k''(a)^c \cap \{i^* \geq k\}\right) \mid i^* \geq \alpha_n\right)$$
$$\geq 1 - \sum_{k=\alpha_n}^{\beta_n} \mathbb{P}(R_k''(a)^c \mid i^* \geq k) \geq 1 - o(n^{-5}),$$

which concludes the proof.

LEMMA A.8. *For a uniformly chosen vertex a and any $\epsilon > 0$, we have*

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

Proof. Conditioning on the event $R''(a)$ defined in Lemma A.6, we have that, for any $\alpha_n \leq k \leq \beta_n$,

$$\tau_{k+1} - \tau_k \leq_{st} Y_k \sim \mathrm{Exp}(S_k(a)) \leq_{st} \mathrm{Exp}\left(\frac{\nu-1}{1+\epsilon}k\right),$$

and all the Y_k's are independent.

Then we let $s = \sqrt{\alpha_n}$, and for n large enough we obtain that

$$\mathbb{E}\left[e^{s(T_a(\beta_n)-T_a(\alpha_n))}\mid R''(a)\right] \le \prod_{i=\alpha_n}^{\beta_n-1}\left(1+\frac{s}{\frac{(\nu-1)i}{1+\epsilon}-s}\right)$$

$$\le \prod_{i=\alpha_n}^{\beta_n-1}\left(1+\frac{s(1+2\epsilon)}{(\nu-1)i}\right)$$

$$\le \exp\left[\frac{s(1+2\epsilon)}{\nu-1}\sum_{i=\alpha_n}^{\beta_n-1}\frac{1}{i}\right]$$

$$\le \exp\left[\frac{s(1+3\epsilon)\log n}{2(\nu-1)}\right].$$

Then we have by Markov's inequality

$$\mathbb{P}\left(T_a(\beta_n)-T_a(\alpha_n)\ge\frac{(1+4\epsilon)\log n}{2(\nu-1)}\right) \le 1-\mathbb{P}(R''(a))$$

$$+\mathbb{P}(R''(a))\mathbb{E}e^{s(T_a(\beta_n)-T_a(\alpha_n))}\exp\left(-\frac{s(1+4\epsilon)\log n}{2(\nu-1)}\right)$$

$$\le \exp\left(-\frac{s\epsilon\log n}{2(\nu-1)}\right)+o(n^{-5})=o(n^{-1}),$$

which concludes the proof.

Then combining Lemma A.8 and A.4 finishes the proof.

A.2 Proof of Proposition 4.1. Fix two vertices u and v. First consider the exploration process for $B_w(u,t)$ until reaching $t = T_u(\beta_n)$. We know by Lemma A.6 that,

$$S_{\beta_n}(u)\ge(\nu-1-o(1))\beta_n$$

with probability at least $1-o(n^{-3/2})$. Thus there are at least $(\nu-1-o(1))\beta_n$ half-edges in $B_w(u,T_u(\beta_n))$ except with probability $n^{-3/2}$.

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 with $B_w(u,T_u(\beta_n))$ is at most

$$\left(1-\frac{(\nu-1-o(1))\beta_n}{m}\right)^{\beta_n}\le \exp[-(9-o(1))\log n]<n^{-4}$$

for large n (recall that $\beta_n^2 = 9\frac{\lambda}{\nu-1}n\log n$). The union bound over u and v completes the proof.

A.3 Proof of Proposition 4.3. We fix a vertex $u\in V_{d_{\min}}$. Let $\hat{d}'_1,...,\hat{d}'_{d_{\min}}$ be the froward degree (i.e. the degree minus one) of neighbors of u. Now we consider the exploration process defined in Section 4.1 from the set $N(u)$. Let $\hat{d}'_{d_{\min}+i}$ be the forward degree

of the vertex added at i's exploration step, with $i\ge 1$, and let

(A.8) $\qquad \hat{S}'_i(u):=\hat{d}'_1+...+\hat{d}'_{d_{\min}+i}-i.$

Again let τ_i be the time of the i'th matching. We have

$$\tau_{i+1}-\tau_i\ge_{st} Y_i\sim\mathrm{Exp}\left(\hat{S}'_i(u)\right),$$

and all the Y_i's are independent. This follows from the fact that the worst case is when the explored set forms a tree. Also by Lemma 4.1, we have

$$\sum_{j=1}^{d_{\min}+i}\hat{d}'_j\le_{st}\sum_{j=1}^{d_{\min}+i}\bar{D}_j^{(n)},$$

where $\bar{D}_j^{(n)}$ are i.i.d with distribution $\bar{\pi}^{(n)}$. Let $\bar{\nu}^{(n)}$ be the expected value of $\bar{D}_1^{(n)}$ which is:

$$\bar{\nu}^{(n)}:=\sum_k k\bar{\pi}_k^{(n)},$$

and let $z_n=\sqrt{n/\log n}$. Now we show that $\tau_{z_n}\ge t_n$ with high probability.

Let us define

$$T'(k)\sim\sum_{i=1}^k\mathrm{Exp}\left(\sum_{j=1}^{d_{\min}+i}\bar{D}_j^{(n)}-i\right),$$

where all the exponential variables are independents. Then we have $\tau_{z_n}\ge_{st}T'(z_n)$.

LEMMA A.9. *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 $\mu_i=\mathbb{E}X_i$, $\Sigma_i=X_1+...+X_i$, $\Lambda_i=\mu_1+...+\mu_i$. Let $Y_i\sim Exp(\Sigma_i)$, and $Z_i\sim Exp(\Lambda_i)$, where all exponential variables are independents. Then we have*

$$Y_1+...+Y_t\ge_{st}Z_1+...+Z_t.$$

Proof. By Jensen's inequality it is easy to see that for positive random variable X, we have

$$\mathrm{Exp}(X)\ge_{st}\mathrm{Exp}(\mathbb{E}X).$$

Then by induction, it suffices to prove that for a pair of random variables X_1, X_2 we have $Y_1+Y_2\ge_{st}Z_1+Z_2$. We have

$$\mathbb{P}(Y_1+Y_2>s)=\mathbb{E}_{X_1}[\mathbb{P}(Y_1+Y_2>s|X_1)]$$
$$\ge\mathbb{E}_{X_1}[\mathbb{P}(\mathrm{Exp}(X_1)+\mathrm{Exp}(X_1+\mu_2)>s)]$$
$$\ge\mathbb{P}(Z_1+Z_2>s).$$

Then by Lemma A.9, we have

$$T'(z_n) \geq_{st} T^*(z_n) := \sum_{i=0}^{z_n} \text{Exp}\left(d_{\min}\bar{\nu}^{(n)} + (\bar{\nu}^{(n)} - 1)i\right),$$

where all exponential variables are independents. Let $b := d_{\min}\bar{\nu}^{(n)} - (\bar{\nu}^{(n)} - 1)$. Then similarly to [6], we have

$$\mathbb{P}(T^*(z_n) \leq t) \leq \int_{\sum x_i \leq t} e^{-\sum_{i=1}^{z_n}((\bar{\nu}^{(n)}-1)i+b)x_i} dx_1...dx_{z_n}$$
$$\prod_{i=1}^{z_n}((\bar{\nu}^{(n)} - 1)i + b)$$
$$= \int_{0 \leq y_1 \leq ... \leq t} e^{-(\bar{\nu}^{(n)}-1)\sum_{i=1}^{z_n} y_i} e^{-by_{z_n}} dy_1...dy_{z_n}$$
$$\prod_{i=1}^{z_n}((\bar{\nu}^{(n)} - 1)i + b),$$

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 $[0, y]$),

$$\mathbb{P}(T^*(z_n) \leq t) \leq \int_0^t e^{-(\bar{\nu}^{(n)}-1+b)y} dy \frac{\prod_{i=1}^{z_n}(i + \frac{b}{\bar{\nu}^{(n)}-1})}{(z_n - 1)!}$$
$$\cdot \int_{[0,y]^{z_n-1}} (\bar{\nu}^{(n)} - 1)^{z_n} e^{-(\bar{\nu}^{(n)}-1)\sum_{i=1}^{z_n-1} y_i} dy_1...dy_{z_n-1}$$
$$\leq \int_0^t e^{-(\bar{\nu}^{(n)}-1+b)y} dy \frac{\prod_{i=1}^{z_n}(i + \frac{b}{\bar{\nu}^{(n)}-1})}{(z_n - 1)!}$$
$$\prod_{i=1}^{z_n-1} \int_0^y (\bar{\nu}^{(n)} - 1)e^{-(\bar{\nu}^{(n)}-1)y_i} dy_i$$
$$\leq \int_0^t e^{-d_{\min}\bar{\nu}^{(n)}}(1 - e^{-(\bar{\nu}^{(n)}-1)y})^{z_n-1} dy$$
$$c(\bar{\nu}^{(n)} - 1)z_n^{\frac{b}{\bar{\nu}^{(n)}-1}+1},$$

where $c > 0$ is an absolute constant. Then we obtain

$$\mathbb{P}(T^*(z_n) \leq t_n) \leq c(\bar{\nu}^{(n)}-1)z_n^{\frac{b}{\bar{\nu}^{(n)}-1}} \int_0^t e^{-n^\epsilon} dy = o(n^{-4}).$$

Then w.h.p. we have $|B'_w(u, t_n)| \leq z_n$. Choosing another vertex v, at random, and exposing $B'_w(v, t_n)$, again w.h.p we obtain a set of size at most z_n. Now because each matching is uniform among the remaining half-edges, then its probability of hitting $B'_w(u, t_n)$ is at most $\hat{S}'_{z_n}(u)/n$.

Let $\epsilon_n := \log \log n$. By Markov's inequality we have

$$\mathbb{P}\left(\hat{S}'_{z_n}(u) \geq z_n\epsilon_n\right) \leq \mathbb{E}\hat{S}'_{z_n}(u)/z_n\epsilon_n$$
$$\leq \frac{d_{\min}\bar{\nu}^{(n)} + (\bar{\nu}^{(n)} - 1)z_n}{z_n\epsilon_n} = o(1).$$

We conclude

$$\mathbb{P}(B'_w(u, t_n)(u) \cap B'_w(v, t_n) \neq \emptyset) \leq \epsilon_n z_n^2/n = o(1),$$

which completes the proof.

A.4   Proof of Lemma 2.1. By Lemma A.3, $R'(a)$ holds with probability at leat $1 - o(n^{-1})$. Then with probability $1 - o(n^{-1})$, for an uniformly chosen vertex a, we have $S_k(a) \geq 1$ for all $1 \leq k \leq \alpha_n$. Then by union bound with probability $1 - o(1)$, for all nodes $a \in V$, the size of the cluster C_a, starting from a reaches α_n. Then we use Lemma A.6 to show that for all nodes, this cluster also reaches β_n. Now it is easy to conclude by Proposition 4.1.