



# AVERAGE DISTANCE IN A GENERAL CLASS OF SCALE-FREE NETWORKS

KARL BRINGMANN,\* *Saarland University and Max Planck Institute for Informatics*  
RALPH KEUSCH,\*\* AND  
JOHANNES LENGLER,\*\* \*\*\* *ETH Zürich*

## Abstract

In Chung–Lu random graphs, a classic model for real-world networks, each vertex is equipped with a weight drawn from a power-law distribution, and two vertices form an edge independently with probability proportional to the product of their weights. Chung–Lu graphs have average distance  $O(\log \log n)$  and thus reproduce the small-world phenomenon, a key property of real-world networks. Modern, more realistic variants of this model also equip each vertex with a random position in a specific underlying geometry. The edge probability of two vertices then depends, say, inversely polynomially on their distance.

In this paper we study a generic augmented version of Chung–Lu random graphs. We analyze a model where the edge probability of two vertices can depend arbitrarily on their positions, as long as the marginal probability of forming an edge (for two vertices with fixed weights, one fixed position, and one random position) is as in a Chung–Lu random graph. The resulting class contains Chung–Lu random graphs, hyperbolic random graphs, and geometric inhomogeneous random graphs as special cases.

Our main result is that every random graph model in this general class has the same average distance as a Chung–Lu random graph, up to a factor of  $1 + o(1)$ . This shows in particular that specific choices, such as taking the underlying geometry to be Euclidean, do not significantly influence the average distance. The proof also shows that every random graph model in our class has a giant component and polylogarithmic diameter with high probability.

*Keywords:* Small-world graphs; complex networks; average distances

2010 Mathematics Subject Classification: Primary 05C82

Secondary 05C12

## 1. Introduction

Many large real-world networks, such as social networks or the internet infrastructure, are *scale-free*, i.e., their degree distribution follows a power law with parameter  $2 < \beta < 3$  [22].

Received 28 February 2022; accepted 6 June 2024.

\* Postal address: Saarland Informatics Campus E1 3, Raum 414, 66123 Saarbrücken, Germany.

\*\* Postal address: Andreasstrasse 5, OAT Z14.1, ETH Zürich, 8050 Zürich, Switzerland.

\*\*\* Email address: [johannes.lengler@inf.ethz.ch](mailto:johannes.lengler@inf.ethz.ch)

© The Author(s), 2024. Published by Cambridge University Press on behalf of Applied Probability Trust. This is an Open Access article, distributed under the terms of the Creative Commons Attribution licence (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted re-use, distribution, and reproduction in any medium, provided the original work is properly cited.

Such networks have been studied in detail since the 1960s. One of the key findings concerning them is the *small-world phenomenon*, which is the observation that two nodes in a network typically have very small graph-theoretic distance. In the nineties, this phenomenon was explained by theoretical models of random graphs. Since then, random graph models have been the basis for the statistical study of real-world networks, as they provide a macroscopic perspective and reproduce structural properties observed in real data. In this line of research, one studies the diameter of a graph, i.e., the largest distance between any pair of vertices in the largest component, and its average distance, i.e., the expected distance between two random nodes of the largest component. A random graph model is said to be *small-world* if its diameter is bounded by  $(\log n)^{O(1)}$  or even  $O(\log n)$ , and *ultra-small-world* if its average distance is only  $O(\log \log n)$ .

Chung–Lu random graphs are a prominent model of scale-free networks [15, 16]. In this model, every vertex  $v$  is equipped with a weight  $w_v$ , and two vertices  $u, v$  are connected independently with probability  $\min\{1, w_u w_v / W\}$ , where  $W$  is the sum over all weights  $w_v$ . The weights are typically assumed to follow a power-law distribution with power-law exponent  $\beta > 2$ . Chung–Lu random graphs have the ultra-small-world property, since in the range  $2 < \beta < 3$  the average distance is  $(2 \pm o(1)) \frac{\log \log(n)}{|\log(\beta-2)|}$  [15, 16].

However, Chung–Lu random graphs fail to capture other important features of real-world networks, such as high clustering or navigability. This is why dozens of papers have proposed more realistic models which also possess some local structure, many of which combine Chung–Lu random graphs (or other classic models such as preferential attachment [3]) with an underlying geometry; see, e.g., hyperbolic random graphs [8, 28, 40], geometric inhomogeneous random graphs [12–14, 26, 30, 33], and many others [2, 9–11, 19, 20, 29, 42]. In these models, each vertex is additionally equipped with a random position in some underlying geometric space, and the edge probability of two vertices depends on their weights as well as the geometric distance between their positions. Typical choices for the geometric space are the unit square, circle, and torus, and typical choices for the dependence on the distance are inverse polynomial, exponential, and threshold functions. Such models can naturally yield a large clustering coefficient, since there are many edges among geometrically close vertices. For some of these models the average distance has been studied and shown to be the same as in Chung–Lu graphs, up to a factor  $1 + o(1)$ ; see, e.g., [1, 9, 19, 27].

It is unclear how much these results depend on the particular choice of the underlying geometry. In particular, it is not known whether any of the important properties of Chung–Lu random graphs transfer to versions with a *non-metric* underlying space. Such spaces are well motivated in the context of social networks, where two persons are likely to know each other if they share a feature (e.g., they are in the same sports club) regardless of their differences in other features (e.g., their profession), which gives rise to a non-metric distance (see Section 3).

**Our contribution:** As the main result of this paper, we prove that in the regime  $2 < \beta < 3$ , all geometric variants of Chung–Lu random graphs have the same average distance  $(2 \pm o(1)) \frac{\log \log(n)}{|\log(\beta-2)|}$ , showing *universality* of the ultra-small-world property.

We do this by analyzing a generic augmented and very general version of Chung–Lu random graphs. Here, each vertex is equipped with a power-law weight  $w_v$  and an independently random position  $\mathbf{x}_v$  in some ground space  $\mathcal{X}$ . Two vertices  $u, v$  form an edge independently with probability  $p_{uv}$  that only depends on the positions  $\mathbf{x}_u, \mathbf{x}_v$  (and  $u, v$  as well as model parameters like the weight sequence). The dependence on  $\mathbf{x}_u, \mathbf{x}_v$  may be arbitrary, as long

as the edge probability has the same marginal probabilities as in Chung–Lu random graphs. Specifically, for fixed  $\mathbf{x}_u$  and random  $\mathbf{x}_v$  we require that the marginal edge probability  $\mathbb{E}_{\mathbf{x}_v}[p_{uv}|\mathbf{x}_u]$  is within a constant factor of the Chung–Lu edge probability  $\min\{1, w_u w_v / W\}$ . This is a natural property for any augmented version of Chung–Lu random graphs. Note that our model is stripped of any geometric specifics. In fact, the ground space is not even required to be metric. We retain only the most important features, namely power-law weights and the right marginal edge probabilities. Hence the main result also demonstrates that there exist random graph models with non-metric underlying geometry that still satisfy the ultra-small-world property.

Beyond the average distance, we establish that this general model is scale-free and has a giant component and polylogarithmic diameter. Thus all instantiations of augmented Chung–Lu random graphs share some basic properties that are considered important for models of real-world networks.

It is quite surprising that the average distance can be computed so precisely in this generality. For example, the clustering coefficient varies drastically between different instantiations of the model, as it encompasses the classic Chung–Lu random graphs, which have clustering coefficient  $n^{-\Omega(1)}$ , as well as geometric variants that have constant clustering coefficient [14]. Therefore, our results hold on graphs with very different local structure. Note that by the scale-free property, all variants of the model contain  $\Theta(n)$  edges. If an instance has high clustering, many edges are *local* edges inside well-connected subgraphs, and therefore useless for finding short paths between far vertices. Nevertheless, our main result implies that in such graphs the average distance is asymptotically the same as in Chung–Lu random graphs, where we have no clustering and every edge is potentially helpful when searching for short paths.

We also remark that our statements fail to hold for  $\beta > 3$ . Indeed, graphs in this regime can look rather diverse depending on the model. For example, some instantiations in this regime do not even have a giant component; the largest component is of polynomial size  $n^{1-\Omega(1)}$  [7]. On the other hand, it is also not hard to construct models for  $\beta > 3$  which do have a giant component, but still have polynomially large average distance; see Remark 1. This variety for  $\beta > 3$  makes it even more surprising that in the regime  $2 < \beta < 3$  the average distance can be determined precisely for all instances at once. Moreover, models in which the marginal edge probability between  $u$  and  $v$  is not a function of the product  $w_u w_v$ , as in the Chung–Lu model also look rather diverse; see [27] for a systematic exploration.

A common property of all models in our general class is that for a set  $S$  of vertices whose weights sum to  $W_S$  (often called *volume* in the literature), the expected number of half-edges going out from  $S$  is  $\Theta(W_S)$ . For the classic Chung–Lu random graphs without geometry, the targets of these half-edges are independent of each other. Thus, the quantity  $W_S$  is essentially sufficient to determine the size and the volume of the neighborhood  $\Gamma(S)$  of  $S$ , and the analyses of Chung–Lu random graphs are based on this property. However, for non-trivial geometries the size of the neighborhood crucially depends on the geometric position of the vertices in  $S$ . For example, if the clustering coefficient is constant, then even if  $S$  consists of only two adjacent vertices there is already a non-negligible probability that they share some neighbors.

Note that this problem only affects the *upper bound* on the average distance, and indeed the *lower bound* carries over trivially from the analysis of Chung–Lu graphs: it is based on a first-moment method for counting the number of self-avoiding paths of length  $(2 - \varepsilon) \frac{\log \log(n)}{|\log(\beta - 2)|}$  with prescribed weight profile, and shows that the expected number of such paths between two random vertices is  $o(1)$  [21]. The expected number of such paths depends on the marginal

probabilities  $\mathbb{E}_{x_v}[p_{uv}|x_u]$ , which we assume to be independent of  $x_u$  up to constant factors; see Section 7 for details. It is thus not surprising that the lower bound was already known.

However, the upper-bound proofs for classic Chung–Lu random graphs do not carry over to our general setting. Similarly, existing proofs for geometric scale-free networks [1, 19] cannot be transferred to our geometry-agnostic setting either, since they rely rather heavily on the specifics of the underlying geometry. Thus we need to develop new methods. Our key ingredient is the ‘bulk lemma’ (Lemma 9), which shows that it is unlikely that a vertex has a  $k$ -neighborhood of size at least  $k$  without having any vertex of high weight in this  $k$ -neighborhood. For Chung–Lu graphs, this statement is rather trivial, since each of the vertices in the  $k$ -neighborhood *independently* has a certain chance of connecting to a vertex of large weight. However, in geometric settings, these events are not independent: in an extreme case, all vertices in the  $k$ -neighborhood could occupy the same position  $x \in \mathcal{X}$ , and a high-weight vertex might connect either to all of them or to none of them (cf. the threshold model in Example 3). So a priori it is not clear that increasing the size of the  $k$ -neighborhood will increase the chances that the  $k$ -neighborhood connects to a high-weight vertex. In previously studied geometric scale-free settings such as hyperbolic geometric graphs [1] or scale-free percolation [19], this problem can be avoided because vertices in these geometries do not cluster with arbitrary density. However, we do not have any such restrictions in our model. Instead, we prove Lemma 9 by a delicate and subtle analysis of the tail bounds of the degree distribution of vertices that would hold if the lemma were false; see the beginning of the proof of Lemma 9 for a more detailed proof sketch.

Once we have established the bulk lemma, the rest of the proof is more similar to previous proofs. By the bulk lemma, a vertex in the giant component can reach a vertex of polylogarithmic weight within  $k$  steps, with failure probability decaying in  $k$ . Then we use a ‘greedy path’ argument (Lemma 7), which constructs a path by always choosing the neighbor of highest weight. In this way, any vertex of polylogarithmic weight can be connected by an ultra-short path to the ‘heavy core’, which is well-connected and contains the vertices of highest weight. The length of this greedy path is so concentrated that it can be estimated by a worst-case estimate. For the initial part of the path, we bound its average length by applying the bulk lemma repeatedly for different values of  $k$  and carefully summing up the resulting terms.

**Organization of the paper:** In Section 2 we present the details and a precise definition of the model, and we formally state the results. In Section 2.3 we introduce some notation that we use throughout the paper. Then we discuss several special cases of the model in Section 3, including geometric inhomogeneous random graphs (GIRGs), the distance model with minimum component distance, and the threshold version of GIRGs, which includes threshold hyperbolic random graphs. We conclude this part of the paper with some remarks in Section 4.

The rest of the paper is dedicated to the proof of our main results. In Section 5 we prove a concentration inequality which will be used later in the proofs. After some basic and preliminary results in Section 6, we prove the connectivity properties and the main result in Section 7, and determine the degree distribution of our model in Section 8. Finally, we prove in Section 9 that the examples from Section 3 are indeed covered by our model.

## 2. Model and results

### 2.1. Definition of the model

In this paper we study the properties of a very general random graph model, where both the set of vertices  $V$  and the set of edges  $E$  are random. Each vertex  $v$  comes with a weight

$w_v$ , which will essentially be the expected degree of  $v$ , and with a random position  $\mathbf{x}_v$  in a geometric space  $\mathcal{X}$ . We now give the full definition, first for the weight sequence and then for the resulting random graph. In the discussion below and throughout the paper, we assume that the model parameters  $\alpha, \beta, d, c_1, c_2, W_{\min}$  are constants while  $n \rightarrow \infty$ , and the hidden factors in Landau notation  $O(\cdot), \Omega(\cdot), \dots$  are universal for all vertices, edges, random variables, and non-constant parameters; see Section 2.3 for more detail.

**Power-law weights:** For every  $n \in \mathbb{N}$  let  $\mathbf{w} = (w_1, \dots, w_n)$  be a non-increasing sequence of positive weights. We call  $W := \sum_{v=1}^n w_v$  the *total weight*. Throughout this paper we will assume that the weights follow a *power law*: the fraction of vertices with weight at least  $w$  is  $\approx w^{1-\beta}$  for some  $\beta > 2$  (the *power-law exponent* of  $\mathbf{w}$ ). More precisely, we assume that for some  $\bar{w} = \bar{w}(n)$  with  $n^{\omega(1/\log \log n)} \leq \bar{w} \leq n^{(1-\Omega(1))/(\beta-1)}$ , the sequence  $\mathbf{w}$  satisfies the following conditions:

(PL1) The minimum weight is bounded by constants, i.e.,

$$W_{\min} := \min\{w_v \mid 1 \leq v \leq n\} = \Theta(1).$$

(PL2) For all  $\eta > 0$  there exist constants  $c_1, c_2 > 0$  such that

$$c_1 \frac{n}{w^{\beta-1+\eta}} \leq \#\{1 \leq v \leq n \mid w_v \geq w\} \leq c_2 \frac{n}{w^{\beta-1-\eta}},$$

where the first inequality holds for all  $W_{\min} \leq w \leq \bar{w}$  and the second for all  $w \geq W_{\min}$ .

We remark that these are standard assumptions for power-law graphs with average degree  $\Theta(1)$ . Note that since  $\bar{w} \leq (1 - \Omega(1))/(\beta - 1)$ , there are  $n^{\Omega(1)}$  vertices with weight at least  $\bar{w}$ . On the other hand, no vertex has weight larger than  $(c_2 n)^{1/(\beta-1-\eta)}$ .

**Random graph model:** Let  $\mathcal{X}$  be a non-empty set, and assume we have a measure  $\mu$  on  $\mathcal{X}$  that allows us to sample elements from  $\mathcal{X}$ . We call  $\mathcal{X}$  the *ground space* of the model and the elements of  $\mathcal{X}$  *positions*. The random graph  $\mathcal{G}(n, \mathcal{X}, \mathbf{w}, p)$  has vertex set  $V = [n] = \{1, \dots, n\}$ . For any vertex  $v$  we independently draw a position  $\mathbf{x}_v \in \mathcal{X}$  according to the measure  $\mu$ . Conditional on  $\mathbf{x}_1, \dots, \mathbf{x}_n$ , we connect any two vertices  $u \neq v$  independently with probability

$$p_{uv} := p_{uv}(\mathbf{x}_u, \mathbf{x}_v) := p_{uv}(\mathbf{x}_u, \mathbf{x}_v; n, \mathcal{X}, \mathbf{w}),$$

where  $p = (p_{uv})_{u,v \in V, u \neq v}$  is a collection of measurable functions  $\mathcal{X} \times \mathcal{X} \rightarrow [0, 1]$  that is symmetric in  $u, v$  (i.e.  $p_{uv}(\mathbf{x}_u, \mathbf{x}_v) = p_{vu}(\mathbf{x}_v, \mathbf{x}_u)$ ) and satisfies the following condition:

(EP1) For any  $u, v$ , if we fix position  $\mathbf{x}_u \in \mathcal{X}$  and draw position  $\mathbf{x}_v$  from  $\mathcal{X}$  according to  $\mu$ , then the marginal edge probability is

$$\mathbb{E}_{\mathbf{x}_v} [p_{uv}(\mathbf{x}_u, \mathbf{x}_v) \mid \mathbf{x}_u] = \Theta\left(\min\left\{1, \frac{W_u W_v}{W}\right\}\right),$$

where the hidden constants are uniform over all  $u, v$  and  $\mathbf{x}_u$ .

Note that the function  $p_{uv}$  does not depend on  $\mathbf{x}_{v'}$  for  $v' \neq u, v$ .

For most of our results we also need an additional condition, to ensure the existence of a unique giant component:

(EP2) There exist  $\eta > 0$  and a function  $\rho = \rho(n) \geq \omega(1)$  such that for any  $u, v$  with  $W_u, W_v \geq \bar{w}$ , and any fixed positions  $\mathbf{x}_u, \mathbf{x}_v \in \mathcal{X}$ , we have

$$p_{uv}(\mathbf{x}_u, \mathbf{x}_v) \geq \left(\frac{n}{\bar{w}^{\beta-1+\eta}}\right)^{-1+\rho(n)/\log \log n}.$$

**Discussion of the conditions (EP1) and (EP2):** Let us first argue why (EP1) alone is not sufficient to yield a unique giant component. Suppose we have an instantiation of our model  $G$  on a space  $\mathcal{X}$ . We will see in this paper that with high probability  $G$  has a giant component that contains all high-degree vertices. Now make a copy  $\mathcal{X}'$  of  $\mathcal{X}$ , and consider a graph where all vertices draw geometric positions from  $\mathcal{X} \cup \mathcal{X}'$ . Vertices in  $\mathcal{X}$  are never connected to vertices in  $\mathcal{X}'$ , but within  $\mathcal{X}$  and  $\mathcal{X}'$  we use the same connection probabilities as for  $G$ . Then the resulting graph will satisfy all properties of our model except for (EP2), but it will have two giant components, one in  $\mathcal{X}$  and one in  $\mathcal{X}'$ . Consequently, in general our result on average distances cannot hold without a condition like (EP2).

To see another property that can go wrong without the condition (EP2), let us slightly vary the above construction so that between any vertex in  $\mathcal{X}$  and any vertex in  $\mathcal{X}'$  there is an edge with probability  $1/n^2$ . (This probability is so small that it does not affect (EP1).) Then with probability  $\Omega(1)$  we get no connecting edge and two giant components, while with probability  $\Omega(1)$  we get a single edge connecting the two parts, and thus a single giant component. In the former case, the two giant components have internally the same average distances as a Chung–Lu random graph, up to a factor  $1 \pm o(1)$ . However, in the latter case the average distance in the unique giant component is larger by a factor of 1.5 than that in Chung–Lu random graphs (twice as large for vertices in different copies, equally large for vertices in the same copy). Thus, without the condition (EP2), the average distance is not even concentrated around a single value in general.

As we will see, (EP2) ensures that the high-weight vertices form a single dense network, so that the graph indeed has a unique giant component. Now let us turn to the condition (EP1). Since its right-hand side is the edge probability of Chung–Lu graphs, (EP1) is a natural condition for any augmented version of Chung–Lu graphs. In particular, (EP1) ensures that the expected degree of a vertex  $v$  with weight  $w_v$  is indeed  $\Theta(w_v)$ . For reasons similar to those discussed for (EP2), we cannot further relax (EP1) to a condition on the marginal probability over random positions  $\mathbf{x}_u$  and  $\mathbf{x}_v$ , i.e., a condition like  $\mathbb{E}_{\mathbf{x}_u, \mathbf{x}_v}[p_{uv}(\mathbf{x}_u, \mathbf{x}_v)] = \Theta(\min\{1, \frac{w_u w_v}{W}\})$ . Indeed, consider the same setup as above, with  $G$ ,  $\mathcal{X}$ , and copy  $\mathcal{X}'$ . For two vertices of weight at most  $\bar{w}$ , connect them only if they are in the same copy of  $\mathcal{X}$ . For two vertices of weight larger than  $\bar{w}$ , always treat them as if they came from the same copy (then the condition (EP2) is satisfied). For a vertex  $u$  of weight at most  $\bar{w}$  and  $v$  of weight larger than  $\bar{w}$ , connect them only if  $u$  is in  $\mathcal{X}'$ . Then the high-weight vertices form a unique component, but it is only connected to vertices in  $\mathcal{X}'$ , while the low-weight vertices in  $\mathcal{X}$  may form a second giant component. Thus, in (EP1) it is necessary to allow any fixed  $\mathbf{x}_u$ .

**Sampling the weights:** In the definition we assume that the weight sequence  $\mathbf{w}$  is fixed. However, if we sample the weights according to an appropriate distribution, then the sampled weights will follow a power law with probability  $1 - n^{-\Omega(1)}$ , so that a model with sampled weights is almost surely included in our model. For the precise statement, see Lemma 5.

**Examples:** We regain the Chung–Lu model as a special case by setting  $\mathcal{X} = \{x\}$  (the trivial ground space) and  $p_{uv} = \min\{1, \frac{w_u w_v}{W}\}$ , since then (EP1) is trivially satisfied and (EP2) is satisfied for  $2 < \beta < 3$ .

We discuss more examples in Sections 3. In particular, the model includes *geometric inhomogeneous random graphs* (GIRGs), which were introduced in [13, 14]. Consider the  $d$ -dimensional ground space  $\mathcal{X} = [0, 1]^d$  with the standard (Lebesgue) measure, where  $d \geq 1$  is a (constant) parameter of the model. Let  $\alpha \neq 1$  be a second parameter that determines how strongly the geometry influences edge probabilities. Finally, let  $\|\cdot\|$  be the Euclidean distance

on  $[0, 1]^d$ , where we identify 0 and 1 in each coordinate (i.e., we take the distance on the torus). We show in Theorem 4 that every edge probability function  $p$  satisfying

$$p_{uv} = \Theta\left(\min\left\{1, (\|\mathbf{x}_u - \mathbf{x}_v\|)^{-d\alpha} \cdot \left(\frac{W_u W_v}{W}\right)^{\max\{\alpha, 1\}}\right\}\right) \quad (1)$$

follows (EP1) and (EP2), so it is a special case of our model. As was shown in [14], an instance of hyperbolic random graphs satisfies (1) asymptotically almost surely (over the choice of random weights  $\mathbf{w}$ ). Thus, hyperbolic random graphs, which have attracted considerable theoretical and experimental interest during the last few years (see, e.g., [6, 8, 28, 36]), are also a special case of our general model.

In Section 3 we will see that GIRGs can be varied as follows. As before, let  $\mathcal{X} = [0, 1]^d$ . For  $x = (x_1, \dots, x_d)$  and  $y = (y_1, \dots, y_d) \in \mathcal{X}$ , we define the *minimum component distance*  $\|x - y\|_{\min} := \min\{|x_i - y_i| \mid 1 \leq i \leq d\}$ , where the differences  $x_i - y_i \in [-1/2, 1/2)$  are computed modulo 1, or, equivalently, on the circle. This distance reflects the property of social networks that two individuals may know each other because they are similar in only one feature (e.g., they share a hobby), regardless of their differences in other features. Note that the minimum component distance is not a metric, since there are  $x, y, z \in \mathcal{X}$  such that  $x$  and  $y$  are close in one component,  $y$  and  $z$  are close in one (different) component, but  $x$  and  $z$  are not close in any component. Let  $V(r)$  be the volume of the ball  $B_r(0) := \{x \in \mathcal{X} \mid \|x\|_{\min} \leq r\}$ . Then any  $p$  satisfying

$$p_{uv} = \Theta\left(\min\left\{1, V(\|\mathbf{x}_u - \mathbf{x}_v\|)^{-\alpha} \cdot \left(\frac{W_u W_v}{W}\right)^{\max\{\alpha, 1\}}\right\}\right)$$

satisfies the conditions (EP1) and (EP2), so it is a special case of our model.

These examples also show that our model is incomparable to the (also very general) model of inhomogeneous random graphs studied by Bollobás, Janson, and Riordan [9]. Their model requires sufficiently many long-range edges, so that setting  $\alpha > 1$  in (1) yields an edge probability that is not supported by their model. Similarly, the example with the minimum component distance is also not supported by their model.

## 2.2. Results of this paper

Our results generalize and improve the understanding of Chung–Lu random graphs, hyperbolic random graphs, and other models, as they are special cases of our fairly general model. We study the following fundamental structural questions.

**Scale-freeness:** Since we plug in power-law weights  $\mathbf{w}$ , it is not surprising that our model is scale-free. More precisely, the degrees follow a power law with exponent  $\beta$ . We say that an event holds *with high probability* (w.h.p.) if it holds with probability  $1 - n^{-\omega(1)}$ .

**Theorem 1.** (Section 8.) *Consider a random instance of our model that satisfies (EP1), but not necessarily (EP2). Then for all  $\eta > 0$ , w.h.p., we have*

$$\Omega(nd^{1-\beta-\eta}) \leq \#\{v \in V \mid \deg(v) \geq d\} \leq O(nd^{1-\beta+\eta}),$$

where the first inequality holds for all  $1 \leq d \leq \bar{w}$  and the second inequality holds for all  $d \geq 1$ . Moreover, w.h.p. the average degree is  $\Theta(1)$ .

**Giant component and diameter:** The connectivity properties of the model for  $\beta > 3$  are not very well behaved, in particular since in this case even threshold hyperbolic random graphs do

not possess a giant component of linear size [7]. Hence, for connectivity properties we restrict our attention to the regime  $2 < \beta < 3$ , which holds for most real-world networks [22].

**Theorem 2.** (Section 7.) *Let  $2 < \beta < 3$ . W.h.p. the largest component of our random graph model has linear size, while all other components have size at most  $\log^{O(1)} n$ . Moreover, w.h.p. the diameter is at most  $\log^{O(1)} n$ .*

A better bound of  $\Theta(\log n)$  holds for the diameter of Chung–Lu graphs [17] and for hyperbolic random graphs [25, 39]. It remains an open question whether the upper bound  $O(\log n)$  holds in general for our model.

**Average distance:** As our main result, we determine the average distance between two randomly chosen nodes in the giant component to be the same as in Chung–Lu random graphs up to a factor  $1 + o(1)$ , showing that the underlying geometry is negligible for this graph parameter.

**Theorem 3.** (Section 7.) *Let  $2 < \beta < 3$ . Then the average distance of our random graph model is  $(2 \pm o(1)) \frac{\log \log n}{|\log(\beta-2)|}$  in expectation and with probability  $1 - o(1)$ .*

### 2.3. Notation

In this section we collect the notation that we will use throughout the paper. For  $w, w' \in \mathbb{R}_{\geq 0}$ ,  $w \leq w'$ , we use the notation  $V_{\geq w} := \{v \in V \mid w_v \geq w\}$  and  $V_{\leq w} := \{v \in V \mid w_v \leq w\}$  and  $V_{[w, w']} := V_{\geq w} \cap V_{\leq w'}$ . Similarly, we write  $W_{\geq w} := \sum_{v \in V_{\geq w}} w_v$ ,  $W_{\leq w} := \sum_{v \in V_{\leq w}} w_v$ , and  $W_{[w, w']} := \sum_{v \in V_{[w, w']}} w_v$  for sums of weights. Recall that  $w_{\min} = \min\{w_v \mid 1 \leq v \leq n\}$ ; similarly we put  $w_{\max} := \max\{w_v \mid 1 \leq v \leq n\}$  for the maximum weight. For  $u, v \in V$  we write  $u \sim v$  if  $u$  and  $v$  are adjacent, and for  $A, B \subseteq V$  we write  $A \sim v$  if there exists  $u \in A$  such that  $u \sim v$ ; we write  $A \sim B$  if there exists  $v \in B$  such that  $A \sim v$ . For a vertex  $v \in V$ , we denote its neighborhood by  $\Gamma(v)$ , i.e.  $\Gamma(v) := \{u \in V \mid u \sim v\}$ . For  $A, B \subseteq V$ , we denote by  $E(A, B) := \{uv \in E \mid u \in A, v \in B\}$  the set of edges from  $A$  to  $B$ .

Throughout the paper we assume that  $n \rightarrow \infty$ , while the model parameters  $\alpha, \beta, d, c_1, c_2, w_{\min}$  are constants. (The parameters  $\alpha$  and  $d$  only appear in examples.) In particular, Landau notation  $O(\cdot), o(\cdot), \Omega(\cdot), \omega(\cdot), \Theta(\cdot)$  may hide any factor that depends on these constants, but these factors must be universal for all vertices and edges of the graph, for all values of  $(x_v)_{v \in V}$ , and for all values of non-constant parameters in the statement (such as  $w$  or  $k$ , which will occur in some lemmas and theorems). For example, (EP1) can be rephrased as follows: there exist  $C, C' > 0$  and  $n_0 > 0$  such that for all  $n \geq n_0$ , all  $u, v \in V = [n]$ , and all  $x_u \in \mathcal{X}$ ,

$$C \cdot \min \left\{ 1, \frac{W_u W_v}{W} \right\} \leq \mathbb{E}_{x_v} [p_{uv}(x_u, x_v) \mid x_u] \leq C' \cdot \min \left\{ 1, \frac{W_u W_v}{W} \right\}.$$

As previously stated, we say that an event holds *with high probability* (w.h.p.) if it holds with probability  $1 - n^{-\omega(1)}$ .

In the model, note that  $w$  and  $p_{uv}$  are parameters and thus fixed, while the probability space is over the positions  $x_u$  and the coin flips for the edges (with probability  $p_{uv}(x_u, x_v)$  for edge  $uv$ ). Sometimes we make statements where we condition on part of the search space to be fixed. In such cases, we emphasize this restricted probability space by including the random components as subscripts on the symbols  $\mathbb{E}$  and  $\mathbb{P}$ . For example, we write  $\mathbb{E}_{x_v} [p_{uv}(x_u, x_v) \mid x_u]$  if  $x_u$  is fixed and  $x_v$  is random.



### 3. Example: GIRGs and generalizations

In this section, we further discuss the special cases of our model mentioned in Section 2.1. Mainly, we study a class which is still fairly general, the so-called *distance model*. We show that the GIRG model introduced in [14] is a special case, and we also discuss a non-metric example. In addition, with the *threshold model* we consider a variation which includes in particular threshold hyperbolic random graphs.

#### The distance model:

We consider the following situation, which will cover both GIRGs and the non-metric example. As our underlying geometry we specify the ground space  $\mathcal{X} = [0, 1]^d$ , where  $d \geq 1$  is a (constant) parameter of the model. We sample uniformly from this set according to the standard (Lebesgue) measure. This is in the spirit of the classical random geometric graphs [41].

To describe the distance between two points  $x, y \in \mathcal{X}$ , assume we have some measurable function  $\|\cdot\|: [-1/2, 1/2]^d \rightarrow \mathbb{R}_{\geq 0}$  such that  $\|0\| = 0$  and  $\|-x\| = \|x\|$  for all  $x \in [-1/2, 1/2]^d$ . Note that  $\|\cdot\|$  does not need to be a norm or seminorm. We extend  $\|\cdot\|$  to  $\mathbb{R}^d$  via  $\|z\| := \|z - u\|$ , where  $u \in \mathbb{Z}^d$  is the unique lattice point such that  $z - u \in [-1/2, 1/2]^d$ . For  $r \geq 0$  and  $x \in \mathcal{X}$ , we define the  $r$ -ball around  $x$  to be  $B_r(x) := \{x \in \mathcal{X} \mid \|x - y\| \leq r\}$ , and we denote by  $V(r)$  the volume of the  $r$ -ball around 0. Intuitively,  $B_r(x)$  is the ball around  $x$  in  $[0, 1]^d$  with the torus geometry, i.e., with 0 and 1 identified in each coordinate. Assume that  $V: \mathbb{R}_{\geq 0} \rightarrow [0, 1]$  is surjective, i.e., for each  $V_0 \in [0, 1]$  there exists  $r$  such that  $V(r) = V_0$ .

Finally, let  $\alpha \in \mathbb{R}_{>0}$  be a long-range parameter. Since the case  $\alpha = 1$  deviates slightly from the general case, we assume  $\alpha \neq 1$ . Let  $p$  be any edge probability function that satisfies, for all  $u, v$  and  $\mathbf{x}_u, \mathbf{x}_v \in \mathcal{X} = [0, 1]^d$ ,

$$p_{uv}(\mathbf{x}_u, \mathbf{x}_v) = \Theta\left(\min\left\{1, V(\|\mathbf{x}_u - \mathbf{x}_v\|)^{-\alpha} \cdot \left(\frac{\mathbf{W}_u \mathbf{W}_v}{\mathbf{W}}\right)^{\max\{\alpha, 1\}}\right\}\right). \quad (2)$$

Then, as we will prove later in Theorem 4,  $p$  satisfies the conditions (EP1) and (EP2), so it is a special case of our model.

**Example 1.** If we choose  $\|\cdot\|$  to be the Euclidean distance  $\|\cdot\|_2$  (or any equivalent norm, such as  $\|\cdot\|_\infty$ ) then we obtain the GIRG model introduced in [14, 42], where the distance between two points  $x, y$  in  $[0, 1]^d$  is given by their geometric distance on the torus. In [14] it was shown that a graph from such a GIRG model w.h.p. has clustering coefficient  $\Omega(1)$ , that it can be stored with  $O(n)$  bits in expectation, and that it can be sampled in expected time  $O(n)$ . Moreover, it was shown that hyperbolic random graphs are contained in the one-dimensional GIRG model. These models have been intensively studied with respect to numerous properties, including component structure [32], clustering [14], diameter [25, 39], spectral gap [31], separators [5, 14, 38], clique number [6], and treewidth [5], as well as processes such as first-passage percolation [34, 35], bootstrap percolation [33], greedy routing [12], bidirectional search [4], and infection processes [26, 30]. With appropriate scaling, their infinite limits can also be defined [35, 45], and they are related to scale-free percolation [19]. In this latter model, the vertices are placed on an infinite grid, so it is not covered by our results. However, it behaves similarly to GIRGs in many respects [34, 35, 45], including average distances [19, 44].

The next distance measure, the minimum component distance, is particularly useful for modeling social networks, because it captures the following property: if two individuals share one feature (e.g., they are in the same sports club), but are very different in many other features (e.g., work, music), then they are still likely to know each other.

**Example 2.** Let the minimum component distance be defined by

$$\|x\|_{\min} := \min\{x_i \mid 1 \leq i \leq d\} \text{ for } x = (x_1, \dots, x_d) \in [-1/2, 1/2]^d.$$

Note that the minimum component distance is not a metric for  $d \geq 2$ , since there are  $x, y, z \in \mathcal{X}$  such that  $x$  and  $y$  are close in one component,  $y$  and  $z$  are close in one (different) component, but  $x$  and  $z$  are not close in any component. Thus the triangle inequality is not satisfied. However, it still satisfies the requirements specified above, so the results of this paper apply. In subsequent work, one of the authors has shown that the resulting graphs for  $d \geq 2$  have a giant component which does not have any separators of sublinear size [38]. This is different from the case of the GIRG model, where it is possible to split the giant component into two halves by removing  $n^{1-\Omega(1)}$  edges [14].

**Theorem 4.** In the distance model described above, let  $p$  be any function that satisfies Equation (2). Then the conditions (EP1) and (EP2) are satisfied, and we obtain an instance of the general model.

**The threshold model:**

Finally, we discuss a variation of Example 1 where we let  $\alpha \rightarrow \infty$  and thus obtain a threshold function for  $p$ .

**Example 3.** Let  $\|\cdot\|$  be the Euclidean distance  $\|\cdot\|_2$  and let  $p$  again satisfy (2), but this time we assume that  $\alpha = \infty$ . More precisely, we require

$$p_{uv}(\mathbf{x}_u, \mathbf{x}_v) = \begin{cases} \Theta(1) & \text{if } \|\mathbf{x}_u - \mathbf{x}_v\| \leq O\left(\left(\frac{w_u w_v}{W}\right)^{1/d}\right), \\ 0 & \text{if } \|\mathbf{x}_u - \mathbf{x}_v\| \geq \Omega\left(\left(\frac{w_u w_v}{W}\right)^{1/d}\right), \end{cases} \tag{3}$$

where the constants hidden by  $O$  and  $\Omega$  do not have to match, i.e., there can be an interval  $[c_1(\frac{w_u w_v}{W})^{1/d}, c_2(\frac{w_u w_v}{W})^{1/d}]$  for  $\|\mathbf{x}_u - \mathbf{x}_v\|$  where the behavior of  $p_{uv}(\mathbf{x}_u, \mathbf{x}_v)$  is arbitrary. This function  $p$  yields the case  $\alpha = \infty$  of the GIRG model introduced in [14]. In [14] it was shown that threshold hyperbolic random graphs are contained in this class of models, and furthermore that the model w.h.p. has clustering coefficient  $\Omega(1)$ , it can be stored with  $O(n)$  bits in expectation, and it can be sampled in expected time  $O(n)$ .

Notice that the volume of a ball with radius  $r_0 = \Theta\left(\left(\frac{w_u w_v}{W}\right)^{\frac{1}{d}}\right)$  around any fixed  $x \in \mathcal{X}$  is  $\Theta(\min\{1, \frac{w_u w_v}{W}\})$ . Thus, by (3), for fixed  $\mathbf{x}_u$  it follows directly that

$$\mathbb{E}_{\mathbf{x}_v}[p_{uv}(\mathbf{x}_u, \mathbf{x}_v) \mid \mathbf{x}_u] = \Theta(\mathbb{P}_{\mathbf{x}_v}[\|\mathbf{x}_u - \mathbf{x}_v\| \leq r \mid \mathbf{x}_u]) = \Theta\left(\min\left\{1, \frac{w_u w_v}{W}\right\}\right).$$

Since (EP1) is satisfied, Theorem 1 for the degree sequence already applies. In order to also fulfill (EP2), we additionally require that  $2 < \beta < 3$  and  $\bar{w} = \omega(n^{1/2})$ . Then for all  $w_u, w_v \geq \bar{w}$  we have  $\frac{w_u w_v}{W} = \omega(1)$ . For all positions  $\mathbf{x}_u, \mathbf{x}_v \in \mathcal{X}$  we thus obtain  $p_{uv}(\mathbf{x}_u, \mathbf{x}_v) = \Theta(1)$  by (3).

**Remark 1.** It follows from the definition that the low-weight vertices in a GIRG contain ordinary random geometric graphs as subgraphs, i.e., every pair of vertices connects with probability  $\Omega(1)$  if the distance between the vertices is at most  $cn^{-1/d}$ , where  $c$  is a constant that depends on the minimal weight  $w_{\min}$ . If  $w_{\min}$  is sufficiently large, then these subgraphs are supercritical, i.e., they have a giant component. On the other hand, in the threshold model for  $\beta > 3$  sufficiently large, all edges cover a polynomially small distance  $n^{-\Omega(1)}$ . Thus, by combining these conditions we get a random graph model in the regime  $\beta > 3$  with giant components where the average distance is polynomially large.

#### 4. Conclusion of first part

We have studied a class of random graphs that generically augment Chung–Lu random graphs by an underlying ground space, i.e., every vertex has a random position in the ground space and edge probabilities may depend arbitrarily on the vertex positions, as long as marginal edge probabilities are preserved. Since our model is very general, it contains numerous well-known models as special cases, including hyperbolic random graphs [8, 40] and geometric inhomogeneous random graphs [14]. Beyond these well-studied models, our model also includes graphs with non-metric ground spaces, which are motivated by social networks, where, for example, two persons are likely to know each other if they share one hobby, regardless of their other hobbies.

Despite its generality, we show that all instantiations of our model have similar connectivity properties, assuming that vertex weights follow a power law with exponent  $2 < \beta < 3$ . In particular, there exists a unique giant component of linear size, and the diameter is polylogarithmic. Surprisingly, for all instantiations of our model the average distance is the same as in Chung–Lu random graphs, namely  $(2 \pm o(1)) \frac{\log \log n}{|\log(\beta-2)|}$ . In some sense, this shows universality of ultra-small worlds.

We leave it as an open problem to determine whether the diameter of our model is  $O(\log n)$  for  $2 < \beta < 3$ .

#### 5. Concentration inequalities

The rest of the paper is dedicated to proofs. We will use the following concentration inequalities.

**Theorem 5.** (Chernoff–Hoeffding bound [24, Theorem 1.1].) *Let  $X := \sum_{i \in [n]} X_i$ , where, for all  $i \in [n]$ , the random variables  $X_i$  are independently distributed in  $[0, 1]$ . Then the following hold:*

- (i)  $\mathbb{P}[X > (1 + \varepsilon)\mathbb{E}[X]] \leq \exp\left(-\frac{\varepsilon^2}{3}\mathbb{E}[X]\right)$  for all  $0 < \varepsilon < 1$ ,
- (ii)  $\mathbb{P}[X < (1 - \varepsilon)\mathbb{E}[X]] \leq \exp\left(-\frac{\varepsilon^2}{2}\mathbb{E}[X]\right)$  for all  $0 < \varepsilon < 1$ , and
- (iii)  $\mathbb{P}[X > t] \leq 2^{-t}$  for all  $t > 2e\mathbb{E}[X]$ .

We will need a concentration inequality which bounds large deviations taking into account some bad event  $\mathcal{B}$ . We start with the following variant of McDiarmid’s inequality as given in [37].

**Theorem 6.** (Theorem 3.6 in [37], slightly simplified.) *Let  $X_1, \dots, X_m$  be independent random variables over  $\Omega_1, \dots, \Omega_m$ . Let  $X = (X_1, \dots, X_m)$ ,  $\Omega = \prod_{k=1}^m \Omega_k$ , and let  $f: \Omega \rightarrow \mathbb{R}$  be measurable with  $0 \leq f(\omega) \leq M$  for all  $\omega \in \Omega$ . Let  $\mathcal{B} \subseteq \Omega$  be such that, for some  $c > 0$  and for all  $\omega \in \overline{\mathcal{B}}, \omega' \in \Omega$  that differ in only one component, we have  $|f(\omega) - f(\omega')| \leq c$ . Then for all  $t > 0$ ,*

$$\mathbb{P}[|f(X) - \mathbb{E}[f(X)]| \geq t] \leq 2e^{-\frac{t^2}{8mc^2}} + 2\frac{mM}{c}\mathbb{P}[\mathcal{B}]. \quad (4)$$

Our improved version of this theorem is the following, where in the Lipschitz condition both  $\omega$  and  $\omega'$  come from the good set  $\overline{\mathcal{B}}$ , but we have to consider changes of *two* components at once. A similar inequality has recently been proven by Combes [18]; see also [46].

**Theorem 7.** Let  $X_1, \dots, X_m$  be independent random variables over  $\Omega_1, \dots, \Omega_m$ . Let  $X = (X_1, \dots, X_m)$ ,  $\Omega = \prod_{k=1}^m \Omega_k$ , and let  $f: \Omega \rightarrow \mathbb{R}$  be measurable with  $0 \leq f(\omega) \leq M$  for all  $\omega \in \Omega$ . Let  $\mathcal{B} \subseteq \Omega$  be such that, for some  $c > 0$  and for all  $\omega \in \overline{\mathcal{B}}$ ,  $\omega' \in \overline{\mathcal{B}}$  that differ in at most two components, we have

$$|f(\omega) - f(\omega')| \leq c. \tag{5}$$

Then for all  $t \geq 2M\mathbb{P}[\mathcal{B}]$

$$\mathbb{P}[|f(X) - \mathbb{E}[f(X)]| \geq t] \leq 2e^{-\frac{t^2}{32mc^2}} + (2\frac{mM}{c} + 1)\mathbb{P}[\mathcal{B}].$$

*Proof.* We say that  $\omega, \omega' \in \Omega$  are neighbors if they differ in exactly one component  $\Omega_k$ . Given a function  $f$  as in the statement, we define a function  $f'$  as follows. On  $\overline{\mathcal{B}}$  the functions  $f$  and  $f'$  coincide. Let  $\omega \in \mathcal{B}$ . If  $\omega$  has a neighbor  $\omega' \in \overline{\mathcal{B}}$ , then choose any such  $\omega'$  and set  $f'(\omega) := f(\omega')$ . Otherwise set  $f'(\omega) := f(\omega)$ .  $\square$

The constructed function  $f'$  satisfies the precondition of Theorem 6. Indeed, let  $\omega \in \overline{\mathcal{B}}$  and  $\omega' \in \Omega$  differ in only one position. If  $\omega' \in \overline{\mathcal{B}}$ , then since  $f'(\omega) = f(\omega)$  and  $f'(\omega') = f(\omega')$ , and by the assumption on  $f$ , we obtain  $|f'(\omega) - f'(\omega')| \leq c$ . Otherwise we have  $\omega' \in \mathcal{B}$ , and since  $\omega'$  has at least one neighbor in  $\overline{\mathcal{B}}$ , namely  $\omega$ , we have  $f'(\omega') = f(\omega'')$  for some neighbor  $\omega'' \in \overline{\mathcal{B}}$  of  $\omega'$ . Note that both  $\omega$  and  $\omega''$  are in  $\overline{\mathcal{B}}$ , and as they are both neighbors of  $\omega'$  they differ in at most two components. Thus, by the assumption on  $f$  we have  $|f'(\omega) - f'(\omega')| = |f(\omega) - f(\omega'')| \leq c$ . Hence, we can use Theorem 6 on  $f'$  and obtain concentration of  $f'(X)$ . Specifically, since  $\mathbb{P}[f(X) \neq f'(X)] \leq \mathbb{P}[\mathcal{B}]$ , and thus  $|\mathbb{E}[f(X)] - \mathbb{E}[f'(X)]| \leq M\mathbb{P}[\mathcal{B}]$ , we obtain

$$\begin{aligned} \mathbb{P}[|f(X) - \mathbb{E}[f(X)]| \geq t] &\leq \mathbb{P}[\mathcal{B}] + \mathbb{P}[|f'(X) - \mathbb{E}[f'(X)]| \geq t - M\mathbb{P}[\mathcal{B}]] \\ &\leq \mathbb{P}[\mathcal{B}] + \mathbb{P}[|f'(X) - \mathbb{E}[f'(X)]| \geq t/2], \end{aligned}$$

since  $t \geq 2M\mathbb{P}[\mathcal{B}]$ , which together with Theorem 6 proves the claim.

### 6. Basic properties

In this section, we prove some basic properties of our random graph model which occur repeatedly in our proofs. In particular we calculate the expected degree of a vertex and the marginal probability that an edge between two vertices with given weights is present. Let us start by calculating the partial weight sums  $W_{\leq w}$  and  $W_{\geq w}$ . The values of these sums will follow from the assumptions on power-law weights in Section 2.1.

**Lemma 1.** The total weight satisfies  $W = \Theta(n)$ . Moreover, for all sufficiently small  $\eta > 0$ ,

- (i)  $W_{\geq w} = O(nw^{2-\beta+\eta})$  for all  $w \geq w_{\min}$ ,
- (ii)  $W_{\geq w} = \Omega(nw^{2-\beta-\eta})$  for all  $w_{\min} \leq w \leq \bar{w}$ ,
- (iii)  $W_{\leq w} = O(n)$  for all  $w$ , and
- (iv)  $W_{\leq w} = \Omega(n)$  for all  $w = \omega(1)$ .

*Proof.* Let  $w_1 \geq w_0 \geq 0$  be two fixed weights. We start by summing up all vertex-weights between  $w_0$  and  $w_1$ . By Fubini's theorem, we can rewrite this sum as

$$\begin{aligned} \sum_{v \in V, w_0 \leq w_v \leq w_1} w_v &= \int_0^\infty |V_{\geq \max\{w_0, x\}} \setminus V_{> w_1}| dx \\ &= w_0 \cdot |V_{\geq w_0}| + \int_{w_0}^{w_1} |V_{\geq x}| dx - w_1 \cdot |V_{> w_1}|. \end{aligned} \tag{6}$$

We start with (i) and apply (6) with  $w_0 = w$  and  $w_1 = w_{\max}$ . Then the set  $V_{> w_1}$  is empty, and we have  $W_{\geq w} = w \cdot |V_{\geq w}| + \int_w^{w_{\max}} |V_{\geq x}| dx$ ; thus the assumption (PL2) implies that  $W_{\geq w}$  equals

$$|V_{\geq w}| \cdot w + \int_w^\infty |V_{\geq x}| dx = O\left(nw^{2-\beta+\eta} + \int_w^\infty nx^{1-\beta+\eta} dx\right) = O\left(nw^{2-\beta+\eta}\right).$$

For (ii) we similarly obtain

$$W_{\geq w} = \Omega\left(nw^{2-\beta-\eta} + \int_w^{w_{\max}} nx^{1-\beta-\eta} dx\right) = \Omega\left(nw^{2-\beta-\eta}\right).$$

For (iii), we see that if  $w < w_{\min}$ , then clearly  $W_{\leq w} = 0$ . Otherwise, Equation (6) with  $w_0 = w_{\min}$  and  $w_1 = w$  implies

$$\begin{aligned} W_{\leq w} &= |V_{\geq w_{\min}}| \cdot w_{\min} + \int_{w_{\min}}^w |V_{\geq x}| dx - |V_{> w}| \cdot w \\ &\leq nw_{\min} + O\left(\int_{w_{\min}}^w nx^{1-\beta+\eta} dx\right) = O(n). \end{aligned}$$

For (iv) we obtain

$$\begin{aligned} W_{\leq w} &\geq \int_{w_{\min}}^w |V_{\geq x}| dx - |V_{> w}| \cdot w = \Omega\left(\int_{w_{\min}}^w nx^{1-\beta-\eta} dx\right) - O\left(nw^{2-\beta+\eta}\right) \\ &= \Omega(n) - o(n) = \Omega(n). \end{aligned}$$

In particular, with the choice  $w = w_{\max}$ , the property  $W = \Theta(n)$  follows from (iii) and (iv).  $\square$

Next we consider the marginal edge probability  $\mathbb{P}[u \sim v]$  of two vertices  $u, v$  with weights  $w_u, w_v$ . For a fixed position  $\mathbf{x}_u \in \mathcal{X}$ , we already know this probability by (EP1).

**Lemma 2.** *Let  $u \in [n]$  and let  $\mathbf{x}_u \in \mathcal{X}$  be any fixed position. Then all edges  $\{u, v\}$ ,  $u \neq v$ , are independently present, and we have*

$$\mathbb{P}_{\mathbf{x}_u, \mathbf{x}_v}[u \sim v] = \Theta(\mathbb{P}_{\mathbf{x}_v}[u \sim v | \mathbf{x}_u]) = \Theta\left(\min\left\{1, \frac{w_u w_v}{W}\right\}\right).$$

*Proof.* Let  $u, v \in [n]$ . Then, from (EP1), it follows directly that

$$\begin{aligned} \mathbb{P}[u \sim v] &= \mathbb{E}_{\mathbf{x}_u}[\mathbb{P}_{\mathbf{x}_v}[u \sim v | \mathbf{x}_u]] = \mathbb{E}_{\mathbf{x}_u}\left[\Theta\left(\min\left\{1, \frac{w_u w_v}{W}\right\}\right)\right] \\ &= \Theta\left(\min\left\{1, \frac{w_u w_v}{W}\right\}\right). \end{aligned} \quad \square$$

Furthermore, for every fixed  $\mathbf{x}_u \in \mathcal{X}$  the edges incident to  $u$  are independently present with probability  $\mathbb{P}_{\mathbf{x}_v}[u \sim v | \mathbf{x}_u]$ , as the event ‘ $u \sim v$ ’ only depends on  $\mathbf{x}_v$  and an independent random choice for the edge  $uv$  (after fixing  $\mathbf{x}_v$ ).

**Remark 2.** We will often use Lemma 2 in the following form. Let  $S, T \subseteq [n]$ ,  $S \cap T = \emptyset$ , and let  $\mathcal{E}$  be any event of non-zero probability that is determined (i.e., measurable) by  $(\mathbf{x}_v)_{v \in S}$ .

Moreover, let  $v_0 \in S$ . Then conditional on  $\mathbf{x}_{v_0}$ , the events ‘ $u \sim v_0$ ’ for  $u \in T$  are independent of each other, and have probability  $\Theta(\min\{1, w_u w_{v_0}/W\})$  respectively. Since these bounds are independent of  $x_{v_0}$ , they also hold conditional on  $\mathcal{E}$ . In particular, using  $1 - x \leq e^{-x}$  for all  $x \in \mathbb{R}$ ,

$$\mathbb{P}[\exists u \in T: u \sim v_0 \mid \mathcal{E}] = 1 - \prod_{u \in T} \left(1 - \Theta\left(\min\left\{1, \frac{w_u w_{v_0}}{W}\right\}\right)\right) \tag{7}$$

$$\geq 1 - \exp\left\{-\Theta\left(\sum_{u \in T} \min\left\{1, \frac{w_u w_{v_0}}{W}\right\}\right)\right\}, \tag{8}$$

where we can exchange the sum and  $\Theta$ -notation in the last step because the hidden constants are uniform over all  $u$ . We express this by saying that the lower bounds  $\Theta(\min\{1, w_u w_{v_0}/W\})$  hold independently, even after conditioning on  $\mathcal{E}$ .

The following lemma shows that the expected degree of a vertex is of the same order as the weight of the vertex; thus we can interpret a given weight sequence  $\mathbf{w}$  as a sequence of expected degrees.

**Lemma 3.** *For any  $v \in [n]$  and any  $\mathbf{x}_v \in \mathcal{X}$  we have  $\mathbb{E}[\deg(v) \mid \mathbf{x}_v] = \Theta(\mathbb{E}[\deg(v)]) = \Theta(w_v)$ .*

*Proof.* Let  $v$  be any vertex and  $\mathbf{x}_v \in \mathcal{X}$ . We show first that  $\mathbb{E}[\deg(v) \mid \mathbf{x}_v] = \Theta(w_v)$  by estimating the expected degree both from below and from above. By Lemma 2, the expected degree of  $v$  is at most

$$\begin{aligned} \sum_{u \neq v} \mathbb{P}[u \sim v \mid \mathbf{x}_v] &= \Theta\left(\sum_{u \neq v} \min\left\{1, \frac{w_u w_v}{W}\right\}\right) = O\left(\sum_{u \in V} \frac{w_u w_v}{W}\right) \\ &= O\left(\frac{w_v}{W} \sum_{u \in V} w_u\right) = O(w_v). \end{aligned}$$

For the lower bound,  $\mathbb{P}[u \sim v \mid \mathbf{x}_v] = \Theta(\frac{w_u w_v}{W})$  holds for all  $w_u \leq \frac{W}{w_v}$ . We set  $w' := \frac{W}{w_v}$  and observe that  $w' = \omega(1)$  by (PL2). Using Lemma 1(iv), we obtain

$$\mathbb{E}[\deg(v)] \geq \sum_{u \neq v, u \in V_{\leq w'}} \mathbb{P}[u \sim v] = \Omega\left(\frac{w_v}{W} W_{\leq w'}\right) = \Omega(w_v).$$

The formula  $\mathbb{E}[\deg(v)] = \Theta(w_v)$  follows from an analogous computation, or by observing that  $\min_{x_v} \mathbb{E}[\deg(v) \mid \mathbf{x}_v] \leq \mathbb{E}[\deg(v)] \leq \max_{x_v} \mathbb{E}[\deg(v) \mid \mathbf{x}_v]$ .

As the expected degree of a vertex is roughly the same as its weight, it is no surprise that w.h.p. the degrees of all vertices with sufficiently large weight are concentrated around the expected value. The following lemma gives a precise statement.

**Lemma 4.** *The following properties hold w.h.p.:*

- (i)  $\deg(v) = O(w_v + \log^2 n)$  for all  $v \in [n]$ ;
- (ii)  $\deg(v) = (1 + o(1))\mathbb{E}[\deg(v)] = \Theta(w_v)$  for all  $v \in V_{\geq \omega(\log^2 n)}$ ;
- (iii)  $\sum_{v \in V_{\geq w}} \deg(v) = \Theta(W_{\geq w})$  for all  $w = \omega(\log^2 n)$ . □

*Proof.* Let  $v \in V$  with fixed position  $\mathbf{x}_v \in \mathcal{X}$  and let  $\mu := \mathbb{E}[\deg(v) \mid \mathbf{x}_v] = \Theta(\mathbf{w}_v)$ , where the last equation holds by Lemma 3. In particular, there exists a constant  $C$  such that  $2e\mu < C \log^2 n$  holds for all vertices  $v \in V_{\leq \log^2 n}$  and all positions  $\mathbf{x}_v \in \mathcal{X}$ . Moreover, by the definition of the model, conditioned on the position  $\mathbf{x}_v$ , the degree of  $v$  is a sum of independent Bernoulli random variables. Thus, if  $v \in V_{\leq \log^2 n}$ , we may apply a Chernoff bound (Theorem 5(iii)) and obtain  $\mathbb{P}[\deg(v) > C \log^2 n] \leq 2^{-C \log^2 n} = n^{-\omega(1)}$ . If  $v \in V_{\geq \log^2 n}$ , we similarly obtain  $\mathbb{P}[\deg(v) > 3\mu/2] \leq e^{-\Theta(\mu)} = n^{-\omega(1)}$  and  $\mu = \Theta(\mathbf{w}_v)$  by Lemma 3. Then (i) follows from applying a union bound over all vertices.

For (ii), let  $v \in V$  such that  $\mathbf{w}_v = \omega(\log^2 n)$ , let  $\mu$  be as defined above and put  $\varepsilon = \frac{\log n}{\sqrt{\mu}} = o(1)$ . Thus by the Chernoff bound,

$$\mathbb{P}[|\deg(v) - \mu| > \varepsilon \cdot \mu] \leq e^{-\Theta(\varepsilon^2 \cdot \mu)} = n^{-\omega(1)},$$

and we obtain (ii) by applying Lemma 3 and a union bound over all such vertices. Finally, from (ii) we infer  $\sum_{v \in V_{\geq w}} \deg(v) = \sum_{v \in V_{\geq w}} \Theta(\mathbf{w}_v) = \Theta(\mathbf{W}_{\geq w})$  for all  $w = \omega(\log^2 n)$ , which shows (iii).  $\square$

We conclude this section by proving that if we sample the weights randomly from an appropriate distribution, then with probability  $1 - o(1)$  the resulting weights satisfy our conditions on power-law weights.

**Lemma 5.** *Let  $\mathbf{w}_{\min}, C_1, C_2 > 0$  be constants, let  $F: \mathbb{R} \rightarrow [0, 1]$  be non-decreasing such that  $F(z) = 0$  for all  $z \leq \mathbf{w}_{\min}$ , and assume  $C_1 z^{1-\beta} \leq 1 - F(z) \leq C_2 z^{1-\beta}$  for all  $z \geq \mathbf{w}_{\min}$ . Suppose that for every vertex  $v \in [n]$  we choose the weight  $\mathbf{w}_v$  independently according to the cumulative probability distribution  $F$ . Then with  $\bar{w} = (n/\log^2 n)^{1/(\beta-1)}$ , the resulting weight vector  $\mathbf{w}$  satisfies (PL1) deterministically, the lower bound of (PL2) w.h.p., and the upper bound of (PL2) with probability  $1 - n^{-\Omega(\eta)}$  for all  $\eta = \eta(n) = \omega(\log \log n / \log n)$ .*

In particular, this lemma proves that for all small constants  $\eta > 0$ , with probability  $1 - n^{-\Omega(1)}$ , (PL1) and (PL2) are fulfilled for weights sampled according to  $F(\cdot)$ . Moreover, it follows that any property which holds with probability  $1 - q$  for weights satisfying (PL1) and (PL2) also holds in a model of sampled weights with probability at least  $1 - q - n^{-\Omega(1)}$ .

*Proof of Lemma 5.* The condition (PL1) is fulfilled by definition of  $F$ , and we only need to prove (PL2). For all  $z > \mathbf{w}_{\min}$ , denote by  $Y_z$  the number of vertices with weight at least  $z$  and observe that

$$\mathbb{E}[Y_z] = n(1 - F(z)) = \Theta(nz^{1-\beta}). \tag{9}$$

Let us first consider the case  $z \in [\mathbf{w}_{\min}, \bar{w}]$ . For all  $z$  in this range we have  $\mathbb{E}[Y_z] = \Omega(\log^2 n)$ , so for any  $z \in [\mathbf{w}_{\min}, \bar{w}]$  the Chernoff bound (Theorem 5(i)–(ii)) yields

$$\mathbb{P}[|Y_z - \mathbb{E}[Y_z]| \geq 0.5\mathbb{E}[Y_z]] \leq \exp(-\Omega(\mathbb{E}[Y_z])) = n^{-\Omega(\log n)}.$$

We will argue in the following that we may use a union bound over a finite number of  $z$  to deduce that with probability  $1 - n^{-\Omega(\log n)}$ , the bound  $|Y_z - \mathbb{E}[Y_z]| < 0.5\mathbb{E}[Y_z]$  holds for all  $z \in [\mathbf{w}_{\min}, \bar{w}]$ , even though there are infinitely many such  $z$ . Consider the sets

$$\begin{aligned} S_1 &:= \{ \inf\{z \in [\mathbf{w}_{\min}, \bar{w}]: 0.5\mathbb{E}[Y_z] \geq i\} \mid i \in \{0, \dots, n\} \}, \\ S_2 &:= \{ \sup\{z \in [\mathbf{w}_{\min}, \bar{w}]: 1.5\mathbb{E}[Y_z] \leq i\} \mid i \in \{0, \dots, n\} \}, \\ S &:= \{\mathbf{w}_{\min}, \bar{w}\} \cup S_1 \cup S_2, \end{aligned}$$

where we discard any infimum or supremum taken over the empty set. By construction, the set  $S$  has at most  $2n + 4 = O(n)$  elements. Thus, by a union bound, with probability  $1 - n^{-\Omega(\log n)}$ , the bound  $|Y_z - \mathbb{E}[Y_z]| < 0.5\mathbb{E}[Y_z]$  holds for all  $z \in S$ . Assume that this event occurs. We want to show that then deterministically the same bound also holds for all other  $z \in [W_{\min}, \bar{w}]$ . To this end, let  $z \in [W_{\min}, \bar{w}] \setminus S$ , and let  $z^- := \max\{z' \in S \mid z' < z\}$  and  $z^+ := \min\{z' \in S \mid z' > z\}$  be its closest lower and upper neighbors in  $S$ . Then  $Y_{z^+} > 0.5\mathbb{E}[Y_{z^+}]$  since  $z^+ \in S$ . Since  $Y_{z^+}$  is an integer, we can strengthen this to  $Y_{z^+} \geq i := \min\{j \in \mathbb{N} \mid j > 0.5\mathbb{E}[Y_{z^+}]\}$ , and note that  $0.5\mathbb{E}[Y_{z^+}] < i$ . By the definition of  $S$ , we have  $0.5\mathbb{E}[Y_{z'}] < i$  for all  $z' \in (z^-, z^+)$ , since otherwise  $S$  would need to contain another point in  $(z^-, z^+)$  where  $0.5\mathbb{E}[Y_{z'}]$  crosses the integer  $i$ . Since  $Y_z$  is non-increasing in  $z$ , we may conclude that  $Y_z \geq Y_{z^+} \geq i > 0.5\mathbb{E}[Y_z]$ , which is half of the claim. The other inequality  $Y_z \leq 1.5\mathbb{E}[Y_z]$  follows analogously from considering  $z^-$ .

Summarizing, we have shown that with probability  $1 - n^{-\omega(1)}$ ,  $Y_z = \Theta(nz^{1-\beta})$  holds for all  $z \in [W_{\min}, \bar{w}]$ . In this case, all  $z$  in our range satisfy both the lower and the upper bound of (PL2) even for  $\eta = 0$ . In particular, this proves that with probability  $1 - n^{-\omega(1)}$ , the lower bound of (PL2) holds for all  $\eta \geq 0$ .

It only remains to prove the upper bound of (PL2) for  $z \geq \bar{w}$ . Let  $z \geq \bar{w}$  and  $\eta = \eta(n) = \omega(\log \log n / \log n)$ . By Markov's inequality and (9),

$$\mathbb{P}[Y_z \geq nz^{1-\beta+\eta}] \leq \mathbb{P}[Y_z \geq \Omega(z^\eta)\mathbb{E}[Y_z]] \leq O(z^{-\eta}) \leq n^{-\Omega(\eta)}. \tag{10}$$

By the same argument as above, we can restrict  $z$  to  $\bar{w}$  and values where the intended bound  $\Omega(z^\eta)\mathbb{E}[Y_z]$  is integral, which happens only for  $O(\log^2 n)$  values of  $z$  above  $\bar{w}$ . Hence we can use the union bound to obtain error probability

$$O(n^{-\Omega(\eta)} \log^2 n) = n^{-\Omega(\eta)},$$

since  $\eta(n) = \omega(\log \log n / \log n)$ . In particular it also follows that with probability  $1 - n^{-\Omega(\eta)}$ , the maximum weight satisfies  $W_{\max} \leq n^{1/(\beta-1-\eta)}$ . □

### 7. Giant component, diameter, and average distance

Under the assumption  $2 < \beta < 3$ , we prove that w.h.p. the general model has a giant component with diameter at most  $(\log n)^{O(1)}$ , and that all other components are only of polylogarithmic size. We further show that the average distance between any two vertices in the giant component is  $(2 + o(1)) \log \log n / |\log(\beta - 2)|$  in expectation and with probability  $1 - o(1)$ . The same formula has been shown to hold for various graph models, including Chung–Lu random graphs [17] and hyperbolic random graphs [1]. The lower bound follows from using the first-moment method on the number of paths of different types. Note that the probability that a fixed path  $P = (v_1, \dots, v_k)$  exists in our model is the same as in Chung–Lu random graphs, since the marginal probability of the event  $v_i \sim v_{i+1}$  conditioned on the positions of  $v_1, \dots, v_i$  is  $\Theta(\min\{1, w_{v_i}w_{v_{i+1}}/W\})$ , as in the Chung–Lu model. In particular, the expected number of paths coincides for the two models (save the factors coming from the  $\Theta(\cdot)$ -notation). Not surprisingly, the lower bound for the expected average distance follows from general statements on power-law graphs, bounding the expected number of too-short paths by  $o(1)$  [21, Theorem 2]; see also Section 6.2 in [43]. The main contribution of this section is to prove a matching upper bound for the average distance.

The proof strategy is as follows. We first prove that, w.h.p., for every vertex of weight at least  $(\log n)^C$  there exists an ultra-short path to the ‘heavy core’, which has diameter  $o(\log \log n)$  and contains the vertices of highest weight. Afterwards, we show that a random low-weight



vertex has a large probability of connecting to a vertex of weight at least  $(\log n)^C$  within a small number of steps. The statement is formalized below as the ‘bulk lemma’ (Lemma 9). This lemma is the crucial step of the main proof, and it is new relative to previous studies of Chung–Lu random graphs and similar models. It contains a delicate analysis of the  $k$ -hop neighborhood of a random vertex, restricted to small weights. Thus, the underlying geometry is used implicitly in order to make the argument applicable for the fairly general model that we study.

Throughout this section, let  $G$  be a graph sampled from our model. For the result on the heavy core and for the bulk lemma, we do not need the condition  $\beta < 3$ . This is only needed for the greedy paths. We start by considering the subgraph induced by the *heavy vertices*  $\bar{V} := V_{\geq \bar{w}}$ , where  $\bar{w}$  is given by the definition of power-law weights; see the condition (PL2). In particular, recall that  $n^{\omega(1/\log \log n)} \leq \bar{w} \leq n^{(1-\Omega(1))/(\beta-1)}$ . We call the induced subgraph  $\bar{G} := G[\bar{V}]$  the *heavy core*.

**Lemma 6.** (Heavy core) *W.h.p.  $\bar{G}$  is connected and has diameter  $o(\log \log n)$ .*

*Proof.* Let  $\bar{n}$  be the number of vertices in the heavy core, and let  $\eta > 0$  be small enough. Since  $\bar{w} \leq n^{(1-\Omega(1))/(\beta-1)}$ , we may bound  $\bar{n} = \Omega(n\bar{w}^{1-\beta-\eta}) = n^{\Omega(1)}$ . By (EP2), if  $\eta$  is sufficiently small, the connection probability for any heavy vertices  $u, v$ , regardless of their position, is at least

$$p_{uv}(\mathbf{x}_u, \mathbf{x}_v) \geq \left( \frac{n}{\bar{w}^{\beta-1+\eta}} \right)^{-1+\omega(1/\log \log n)} \geq \bar{n}^{-1+\omega(1/\log \log n)}.$$

Therefore, the diameter of the heavy core is at most the diameter of an Erdős–Rényi random graph  $G(\bar{n}, p)$ , with  $p = \bar{n}^{-1+\omega(1/\log \log n)}$ . However, with probability  $1 - \bar{n}^{-\omega(1)}$  this diameter is  $\Theta(\log \bar{n} / \log(p\bar{n})) = o(\log \log n)$  [23], and in particular the graph is connected in this case. Since  $\bar{n} = n^{\Omega(1)}$ , this proves the lemma.  $\square$

Next we show that if we start at a vertex of weight  $w$ , going greedily to the neighbors of largest weight yields a short path to the heavy core with a probability that approaches 1 as  $w$  increases.

**Lemma 7.** (Greedy path) *Let  $2 < \beta < 3$ .*

- (i) *Let  $0 < \varepsilon < 1$  be constant and let  $v$  be a vertex of weight  $2 \leq w < \bar{w}$ . Then with probability at least  $1 - O(\exp(-w^{\Omega(\varepsilon)}))$  there exists a weight-increasing path of length at most  $(1 + \varepsilon) \frac{\log \log n}{|\log(\beta-2)|}$  from  $v$  to the heavy core.*
- (ii) *For every constant  $\varepsilon > 0$  there exists a constant  $C = C(\varepsilon) > 0$  such that w.h.p. for all  $v \in V_{\geq (\log n)^C}$  there exists a weight-increasing path of length at most  $(1 + \varepsilon) \frac{\log \log n}{|\log(\beta-2)|}$  from  $v$  to the heavy core.*

*Proof.* Let  $0 < \varepsilon < 1$ , let  $v$  be a vertex of weight  $2 \leq w \leq \bar{w}$ , and let

$$\tau = \tau(\varepsilon) := (\beta - 2)^{-1/(1+\varepsilon/2)}.$$

Note that  $1 < \tau < 1/(\beta - 2)$ , and that  $1/\log \tau = (1 + \varepsilon/2)/|\log(\beta - 2)|$ . Moreover, we define an increasing weight sequence  $w_0, w_1, \dots, w_{i_{\max}} := \bar{w}$  such that for all  $1 \leq i \leq i_{\max}$  it holds that  $w_i := w_{i-1}^\tau$ , and such that  $w_0 \leq \mathbf{w}_v < w_1$ . For all  $i < i_{\max}$  we put  $V_i := V_{\geq w_i} \setminus V_{\geq w_{i+1}}$ . Furthermore, we put  $V_{i_{\max}} := \bar{V} = V_{\geq \bar{w}}$  and  $v_0 := v$ . We will show that with sufficiently high probability, for all  $0 \leq i < i_{\max}$  the vertex  $v_i$  has at least one neighbor  $v_{i+1} \in V_{i+1}$ . By picking such a  $v_{i+1}$ , we iteratively define a weight-increasing path. Note that

$i_{\max} = \lceil \log_{\tau} (\log \bar{w} / \log w_v) \rceil$ , so this implies that there is a path from  $v$  to the heavy core of length at most

$$i_{\max} \leq (1 + \varepsilon/2) \frac{\log \log n}{|\log(\beta - 2)|} + 1 \leq (1 + \varepsilon) \frac{\log \log n}{|\log(\beta - 2)|},$$

for sufficiently large  $n$ , and thus proves the statement (i). □

Let  $0 \leq i < i_{\max}$  and assume that there exists a weight-increasing path from  $v_0$  to some vertex  $v_i \in V_i$ . Note that this event only depends on the random graph induced by the vertex set  $V_{<w_{i+1}}$ . We want to verify that  $v_i$  connects to at least one vertex  $v_{i+1} \in V_{i+1}$ . First, observe that by the condition (PL2), each layer  $V_i$  contains at least  $\Omega(nw_i^{1-\beta-\eta})$  and at most  $O(nw_i^{1-\beta+\eta})$  vertices. Next, by the condition (EP1), for any position  $\mathbf{x}_{v_i}$  the edges from  $v_i$  to vertices  $v$  with  $v \in V_{i+1}$ , are independently present with probability  $\Omega(\min\{w_v w_i / W, 1\})$ , respectively; see also Lemma 2. Note that this bound is independent of  $\mathbf{x}_{v_i}$  and thus holds for any realization of  $V_{<w_{i+1}}$ ; see also Remark 2. If  $w_i w_{i+1} \geq W$ , this probability is  $\Omega(1)$ . However, then  $w_i \geq n^{1/(1+\tau)}$  and we deduce  $|V_{i+1}| = n^{\Omega(1)}$ . In this case, the probability that  $v_i$  connects to at least one vertex of the next weight layer is  $1 - \exp(-n^{\Omega(1)}) = 1 - \exp(-w_i^{\Omega(\varepsilon)})$ , where the second step is a trivial upper bound that holds since  $\varepsilon$  is constant and  $w_i \leq \bar{w} \leq n^{O(1)}$ . So assume  $w_i w_{i+1} < W$ , where we can lower-bound the edge probability by  $\Omega(w_i w_{i+1} / W)$ . Thus, for any  $\eta > 0$  the probability that  $v_i$  does *not* connect to a vertex in  $V_{i+1}$  is at most

$$\begin{aligned} p_i &:= \prod_{v \in V, w_v \geq w_{i+1}} \left( 1 - \Omega\left(\frac{w_i w_{i+1}}{W}\right) \right) \leq \exp\left(-\Omega\left(\frac{w_i w_{i+1}}{W} \cdot |V_{i+1}|\right)\right) \\ &\leq \exp\left(-\Omega(w_i w_{i+1}^{2-\beta-\eta})\right), \end{aligned}$$

where we used Lemma 1 in the last step. Since  $w_{i+1} \leq w_i^{\tau}$ , we obtain

$$p_i \leq \exp\left(-\Omega(w_i^{1-\tau(\beta-2+\eta)})\right).$$

Note that as  $\tau < 1/(\beta - 2)$ , the exponent of  $w_i$  in this expression is positive for sufficiently small  $\eta > 0$ . More precisely, we have

$$1 - \tau(\beta - 2) = 1 - (\beta - 2)^{\varepsilon/(2+\varepsilon)} = \Omega(\varepsilon),$$

and thus for  $\eta > 0$  sufficiently small compared to  $\varepsilon$ ,

$$p_i \leq \exp\left(-w_i^{\Omega(\varepsilon)}\right). \tag{11}$$

By induction and a union bound, for  $0 \leq j \leq i_{\max}$  it follows that there is a weight-increasing path from  $v_0$  to  $V_j$  of length  $j$  with probability at least  $1 - \sum_{i=0}^{j-1} \exp(-w_i^{\Omega(\varepsilon)}) = 1 - \exp(-w^{\Omega(\varepsilon)})$ . With  $j = i_{\max}$  this proves the first claim. For the following step, let us denote the hidden constant in the last expression by  $C$ , so that the probability is at least  $1 - \exp(-w^{C\varepsilon})$ .

For the second statement, let  $C = C(\varepsilon) := 2/(C'\varepsilon)$ . If a vertex  $v$  has weight at least  $(\log n)^C$ , then the error probability estimated above is at least  $1 - e^{-(\log n)^2} = 1 - n^{-\omega(1)}$ . The claim now follows from a union bound over all vertices of weight at least  $(\log n)^C$ .

Lemma 7(i) implies that any vertex of sufficiently large constant weight has probability  $\Omega(1)$  of being connected to the heavy core by a weight-increasing path. In particular,

this means that the connected component of the heavy core has size  $\Omega(n)$  in expectation. In the following lemma, we show that the same holds w.h.p., which is less obvious, and requires our improved variant of McDiarmid's inequality in Theorem 7.

**Lemma 8.** (Giant component.) *Let  $2 < \beta < 3$ . W.h.p. there are  $\Omega(n)$  vertices in the same component as the heavy core.*

*Proof.* Let  $\varepsilon > 0$  be a sufficiently small constant, let  $\eta > 0$  be a constant that is sufficiently small compared to  $\varepsilon$ , and let  $\tau := (\beta - 2)^{-1/(1+\varepsilon/2)}$ . We use a system of weight layers as in the proof of Lemma 7(i), starting with  $w_0 := 2$ , so let  $w_i := 2^{\tau^i}$  and  $V_i := V_{\geq w_i} \setminus V_{\geq w_{i+1}}$  for all  $i \geq 0$ . We remark already here that we will only need to study indices  $i$  for which  $w_i \leq (\log n)^C$ , where  $C = C(\varepsilon)$  is the constant from Lemma 7(ii), since we already know by Lemma 7(ii) that otherwise all vertices in  $V_i$  are in the same component as the heavy core, w.h.p. Moreover, by (PL2), if we choose  $i_0 = O(1)$  sufficiently large, then for all  $i \geq i_0$  with  $w_i \leq (\log n)^C$  we have  $|V_{\geq w_{i+1}}| \leq |V_{\geq w_i}|/2$ , and hence  $|V_i| \geq |V_{\geq w_i}|/2$ . In particular, again by (PL2) we have  $|V_i| \geq n/(\log n)^{O(1)}$  for all  $i \geq i_0$  with  $w_i \leq (\log n)^C$ , and we have  $|V_{i_0}| = \Omega(n)$  since we assumed  $i_0 = O(1)$ . We will increase  $i_0$  further in the proof, but maintain  $i_0 = O(1)$ .  $\square$

As already mentioned, the tricky part of the lemma is that it should hold w.h.p. The general strategy will be to define recursively for each  $i$  two sets  $B_i, B'_i$  of 'bad' vertices, such that the following hold:

- Deterministically, every vertex in  $V_i \setminus (B_i \cup B'_i)$  has a weight-increasing path to a vertex of weight at least  $(\log n)^C$ .
- W.h.p.,  $B_i$  and  $B'_i$  are small. (This is made precise in Claim 1 and Equation (12) below.) In particular, if  $i_0 = O(1)$  is sufficiently large then w.h.p.  $|B_i \cup B'_i| \leq |V_i|/2$ .

It then follows that w.h.p., at least half of the vertices in  $V_{i_0}$  are connected to the heavy core. Since  $|V_{i_0}| = \Omega(n)$ , this proves the lemma.

So let  $i \geq i_0$  be such that  $w_i \leq (\log n)^C$ . For every  $v \in V_i$ , let  $\Gamma_i(v) := \{u \in V_{i+1} \mid v \sim u\}$ , and let  $E_i(v) := \mathbb{E}[|\Gamma_i(v)|]$ . Moreover, let  $\gamma := \tau(2 - \beta - \eta) + 1 > 0$ . Then for every  $v \in V_i$ , by (EP1) and Lemma 2,

$$E_i(v) \geq \Omega\left(n w_{i+1}^{1-\beta-\eta} \cdot \frac{w_i w_{i+1}}{\mathbf{W}}\right) \geq \Omega(w_{i+1}^{2-\beta-\eta} w_i) \geq \Omega(w_i^{\tau(2-\beta-\eta)+1}) \geq \Omega(w_i^\gamma).$$

As this lower bound is independent of  $v$ , we also have

$$E_i := \min_{v \in V_i} E_i(v) = \Omega(w_i^\gamma).$$

Let  $\lambda := \min\{\gamma, \frac{1}{2C\tau}\}$ . Furthermore put  $B_i := \{v \in V_i \mid |\Gamma_i(v)| \leq E_i/2\}$ . This is the first set of 'bad' vertices.

**Claim 1.** There is a constant  $c > 0$  such that w.h.p., for all  $i \geq i_0$  with  $w_i \leq (\log n)^C$ , it holds that  $|B_i| \leq 2 \exp(-c w_i^\lambda) \cdot |V_i|$ .

We postpone the proof of Claim 1 (and Claim 2 below) until we have finished the main argument. We uncover the sets  $V_i$  one by one, starting with the largest weights. Let  $\delta > 0$  be so small that  $\tau(\lambda - \delta) > \lambda$ . We will show by downwards induction in  $i$  that w.h.p., for all  $i \geq i_0$  the fraction of vertices in  $V_i$  with a weight-increasing path to the heavy core is at least  $1 - \exp(-c w_i^{\lambda-\delta})$ . Recall that if  $w_i \geq (\log n)^C$  then we already know that w.h.p. all vertices in

$V_i$  are connected to the heavy core with weight-increasing paths, which gives the base case of the induction. For the remaining values  $i \geq i_0$ ,

- (a) denote by  $V'_i$  the set of vertices in  $V_i$  for which there is no weight-increasing path to the heavy core that uses exactly one vertex per layer;
- (b) let  $\Lambda_i \subseteq V_i$  be the subset  $V_i$  of size  $\exp(-cw_{i-1}^{\lambda-\delta})|V_i|$  consisting of the vertices in  $V_i$  with the smallest number of neighbors in  $V_{i+1} \setminus V'_{i+1}$  (where we break ties according to some previously fixed order).

Recall that  $E(V_i, \Lambda_i)$  denotes the set of edges from  $V_i$  to  $\Lambda_i$ . We then use the following claim (whose proof we also postpone).

**Claim 2.** There exists  $D > 0$  such that w.h.p., for all  $i \geq i_0$  with  $w_i \leq (\log n)^C$ , it holds that

$$|E(V_i, \Lambda_{i+1})| \leq \exp(-cw_{i+1}^{\lambda-\delta}) \cdot |V_i| \cdot E_i \cdot w_i^D.$$

The induction now runs as follows. Consider some  $i \geq i_0$  such that  $w_i \leq (\log n)^C$ , and assume by induction that for sufficiently many vertices of  $V_{i+1}$  there is a weight-increasing path to the heavy core; that is, assume  $|V'_{i+1}| \leq \exp(-cw_{i+1}^{\lambda-\delta}) \cdot |V_{i+1}|$ . By the construction of  $\Lambda_{i+1}$ , this implies  $V'_{i+1} \subseteq \Lambda_{i+1}$ . Now we consider  $B'_i := \{v \in V_i \mid |E(\{v\}, V'_{i+1})| \geq E_i/2\}$ . If the low-probability event of Claim 2 does not occur, using  $w_i^{\lambda-\delta} = w_i^{\lambda+\Omega(1)}$  and increasing  $i_0$  if necessary, we find that

$$|B'_i| \leq \frac{2|E(V_i, V'_{i+1})|}{E_i} \leq 2 \exp(-cw_{i+1}^{\lambda-\delta}) \cdot |V_i| \cdot w_i^D \leq 2 \exp(-cw_i^\lambda) \cdot |V_i|, \tag{12}$$

provided that  $i_0 = i_0(c, D)$  (and thus  $w_{i_0}$ ) is a sufficiently large constant. Note that we may increase  $i_0$  because Claims 1 and 2 become weaker as  $i_0$  increases. It remains to observe that every vertex in  $V_i \setminus (B_i \cup B'_i)$  has at least one edge into  $V_{i+1} \setminus V'_{i+1}$ . Since the latter vertices are all connected to the heavy core, we have at least  $|V_i| - |B_i| - |B'_i|$  vertices in  $V_i$  that are connected to the heavy core. By Claim 1 and Equation (12), w.h.p.  $B_i$  and  $B'_i$  both have size at most  $2 \exp(-cw_i^\lambda)|V_i|$ . So, together they have size at most  $4 \exp(-cw_i^\lambda)|V_i| \leq \exp(-cw_i^{\lambda-\delta})|V_i|$ , where the last step holds for all  $i \geq i_0$  if  $i_0$  (and thus  $w_i \geq w_{i_0}$ ) is sufficiently large. This proves the induction hypothesis and thus concludes the induction modulo Claims 1 and 2. To conclude the proof, if  $i_0 = O(1)$  is sufficiently large then  $\exp(-cw_{i_0}^{\lambda-\delta}) \leq 1/2$ . The existence of the giant component now follows since  $|V_{i_0} \setminus (B_{i_0} \cup B'_{i_0})| \geq (1 - \exp(-cw_{i_0}^{\lambda-\delta}))|V_{i_0}| \geq |V_{i_0}|/2$ . Hence, w.h.p. at least half of  $V_{i_0}$  is connected to the heavy core, where  $|V_{i_0}| = \Omega(n)$ .

*Proof of Claim 1.* Let  $i \geq i_0$  be such that  $w_i \leq (\log n)^C$ , and recall that  $|V_i| = n/(\log n)^{O(1)}$ . For a single  $v \in V_i$  with given location, the events ‘ $v \sim u$ ’ are independent for all  $u \in V_{i+1}$  by Lemma 2. So by the Chernoff bound (Theorem 5), there is a constant  $c > 0$  such that  $\mathbb{P}[v \in B_i] \leq \exp(-cw_i^\lambda)$  and  $\mathbb{E}[|B_i|] \leq \exp(-cw_i^\lambda)|V_i|$ . Let  $G_i$  be the subgraph induced by  $V_i$  and  $V_{i+1}$ , and observe that the size of  $B_i$  only depends on  $G_i$ . In order to prove concentration of  $|B_i|$  we will use Theorem 7. For this, we need to argue that we can write the distribution of  $G_i$  as a product probability space, i.e., we need to describe it via *independent* random variables. Recall that two different random processes are applied to create the geometric graph. First, we choose the positions  $x_v \in \mathcal{X}$  independently at random. Afterwards, every edge  $\{u, v\}$  is inserted with some probability  $p_{uv}$ . So far, these random variables are *not* independent, so this description does not yield a product probability space.

Without loss of generality, let us assume that the vertices are sorted by weights in decreasing order. For every vertex  $u \in V_i \cup V_{i+1}$  we first have the random variable  $\mathbf{x}_u$  for its position. Now, for each  $u \in V_i \cup V_{i+1}$  we introduce a second, independent random variable  $Y_u := (Y_u^1, \dots, Y_u^{u-1})$  (which is the empty tuple for  $u = 1$ ), where each  $Y_u^v$  is a real number chosen independently and uniformly at random from the interval  $[0, 1]$ . Then for  $v < u$ , we include the edge  $\{u, v\}$  in the graph if and only if

$$p_{uv} > Y_u^v.$$

We observe that indeed this implies  $\mathbb{P}[u \sim v \mid \mathbf{x}_u, \mathbf{x}_v] = p_{uv}(\mathbf{x}_u, \mathbf{x}_v)$ , as desired. Moreover, conditional on the vertex positions, the events  $u \sim v$  and  $u' \sim v'$  are independent as long as  $\{u, v\} \neq \{u', v'\}$ . This matches exactly the definition of  $G_i$  as (induced subgraph of) a GIRG. Thus, if we denote by  $\Omega$  the product space of all random variables over all components in  $(\mathbf{x}_u, Y_u)_{u \in V_i \cup V_{i+1}}$ , then the distribution of  $G_i$  is described by  $\Omega$ : every  $\omega \in \Omega$  defines a graph  $G_i(\omega)$ , and the probability distribution of  $G_i(\omega)$  equals the probability distribution of  $G_i$  as subgraph of a GIRG. Note that in the definition of the product space  $\Omega$ , the random variable  $Y_u$  is counted as a single coordinate, and the outcomes of this coordinate are vectors in  $[0, 1]^{u-1}$ . Thus  $\Omega$  has  $2(|V_i| + |V_{i+1}|)$  coordinates. This convention is useful since then  $\Omega$  is a product space of at most  $O(n)$  coordinates, rather than  $O(n^2)$ .

Our aim is to apply Theorem 7 to the function  $f(\omega) := |B_i(\omega)|$ , which depends on  $\omega$  since the graph  $G_i$  depends on  $\omega$ . In particular,  $0 \leq f(\omega) \leq n =: M$ . We study the bad event  $\mathcal{B}$  that there exists a vertex  $v \in V_i \cup V_{i+1}$  with degree larger than  $(\log n)^{2C\tau^2}$  in  $G_i$ . By Lemma 4(i) we have  $\mathbb{P}[\mathcal{B}] = n^{-\omega(1)}$ , since  $w_i \leq (\log n)^C$  and therefore  $w_v \leq (\log n)^{C\tau^2}$  for all  $v \in V_i \cup V_{i+1}$ . Let  $\omega, \omega' \in \bar{\mathcal{B}}$  be elements that differ in at most two coordinates in our product probability space  $\Omega$ . We observe that changing one coordinate  $\mathbf{x}_u$  or  $Y_u$  (the latter means changing a vector in  $[0, 1]^{u-1}$ ) can only influence the neighborhoods of  $u$  itself and of the vertices that are neighbors of  $u$  before or after the coordinate change. In formula, changing one coordinate  $\mathbf{x}_u$  or  $Y_u$  can add or remove at most

$$1 + \deg_{G_i(\omega)}(u) + \deg_{G_i(\omega')}(u) \leq 1 + 2(\log n)^{2C\tau^2}$$

vertices from  $B_i$ , where the latter inequality follows from  $\omega, \omega' \in \bar{\mathcal{B}}$ . Therefore,

$$||B_i(\omega)| - |B_i(\omega')|| \leq (\log n)^{O(1)}.$$

We pick  $t = \exp(-cw_i^\lambda) \cdot (|V_i| + |V_{i+1}|)$  and observe that  $w_i^\lambda \leq (\log n)^{1/2}$  by our choice of  $\lambda = \min\{\gamma, \frac{1}{2C\tau}\}$ . In particular,  $t \geq 2M\mathbb{P}[\mathcal{B}]$  is satisfied for sufficiently large  $n$ , which allows us to apply Theorem 7 with  $f(\omega) = |B_i(\omega)|$ ,  $c = (\log n)^{O(1)}$ ,  $m = 2(|V_i| + |V_{i+1}|)$ , and  $M = n$  and get

$$\begin{aligned} \mathbb{P}[|B_i| - \mathbb{E}[|B_i|] \geq t] &\leq 2 \exp\left(-\frac{t^2}{64(|V_i| + |V_{i+1}|)(\log n)^{O(1)}}\right) + n^{O(1)}\mathbb{P}[\mathcal{B}] \\ &\leq 2 \exp\left(-\frac{e^{-2cw_i^\lambda}(|V_i| + |V_{i+1}|)}{64(\log n)^{O(1)}}\right) + n^{O(1)}\mathbb{P}[\mathcal{B}] = n^{-\omega(1)}, \end{aligned}$$

where the first summand is  $2 \exp(-n^{1-o(1)}) \leq n^{-\omega(1)}$  because  $|V_i| + |V_{i+1}| \geq n/(\log n)^{O(1)}$  and  $e^{-2cw_i^\lambda} \geq n^{-o(1)}$ .

Hence, w.h.p. we have  $|B_i| \leq \mathbb{E}[|B_i|] + t \leq (\exp(-cw_i^\gamma) + \exp(-cw_i^\lambda)) \cdot |V_i|$ . For fixed  $i \geq i_0$ , the statement now follows since  $\lambda < \gamma$  and  $w_i > 1$ , and then the proof of the claim is finished by a union bound over all  $O(\log \log n)$  choices of  $i$ .  $\square$

*Proof of Claim 2.* Let  $i \geq i_0$  be such that  $w_i \leq (\log n)^C$ . We assume that the subgraph induced by  $V_{\geq w_{i+2}}$  is given, and now we uncover  $V_i$  and  $V_{i+1}$  to obtain the subgraph induced by  $V_{\geq w_i}$ . Similarly as in the proof of Claim 1, we can assume that this probability space  $\Omega$  is a product probability space with  $2(|V_i| + |V_{i+1}|)$  coordinates. Recall that  $G_i$  denotes the subgraph induced by  $V_i \cup V_{i+1}$ . We consider the same bad event  $\mathcal{B}$  as in the proof of Claim 1, i.e.,  $\mathcal{B}$  denotes the event that the maximum degree in  $G_i$  is larger than  $(\log n)^{2c\tau^2}$ . Note that  $\mathcal{B}$  is independent of  $V_{\geq w_{i+2}}$ , so indeed Lemma 4 can again be applied to deduce  $\mathbb{P}[\mathcal{B}] = n^{-\omega(1)}$ .

Let  $Z_i := |E(V_i, \Lambda_{i+1})|$ . We will apply Theorem 7 to the random variable  $f(\omega) = Z_i(\omega)$ , so let  $\omega, \omega' \in \bar{\mathcal{B}}$  be such that they differ in at most two coordinates of  $\Omega$ . If we change a coordinate of  $\Omega$  that stems from a vertex  $v \in V_i$ , under  $\bar{\mathcal{B}}$  the influence on  $Z_i$  is at most  $(\log n)^{O(1)}$ . If a coordinate belonging to a vertex  $v \in V_{i+1}$  is changed, this may result in a different set  $\Lambda_{i+1}$ , because either  $v$  might enter the set  $\Lambda_{i+1}$  and replace another vertex  $v'$  in  $\Lambda_{i+1}$  (namely the vertex with the largest number of neighbors in  $V_{i+2} \setminus V'_{i+2}$ ; see (b) on page 19 for the definition of  $\Lambda_i$ ); or  $v$  might be removed from  $\Lambda_{i+1}$  and be replaced by some other vertex  $v'$ . In either case, the symmetric difference between the old  $\Lambda_{i+1}$  and the new  $\Lambda_{i+1}$  is at most two, and  $E(V_i, \Lambda_{i+1})$  can only change in edges incident to  $v$  or  $v'$ . Hence, under  $\bar{\mathcal{B}}$  changing the coordinate of a vertex  $v \in V_{i+1}$  can change  $Z_i$  by at most  $c = (\log n)^{O(1)}$ . Note that the same is also true if the set  $\Lambda_{i+1}$  does not change. Since  $\omega, \omega'$  differ in at most two coordinates, we conclude that  $|Z_i(\omega) - Z_i(\omega')| \leq (\log n)^{O(1)}$ .

Next, we want to upper-bound  $\mathbb{E}[Z_i]$ . First, we uncover  $V_{i+1}$  to obtain the subgraph induced by  $V_{\geq w_{i+1}}$ . Then the set  $\Lambda_{i+1}$  is determined. In a second step, we uncover  $V_i$ . By (EP1) and linearity of expectation, we deduce that

$$\begin{aligned} \mathbb{E}[Z_i] &\leq |\Lambda_{i+1}| \cdot |V_i| \cdot O\left(\frac{w_{i+1}^{1+\eta} w_i^{1+\eta}}{W}\right) \\ &= O\left(\exp(-cw_{i+1}^{\lambda-\delta}) n w_{i+1}^{1-\beta+\eta} \cdot |V_i| \cdot \frac{w_{i+1}^{1+\eta} w_i^{1+\eta}}{W}\right) \\ &\leq \exp(-cw_{i+1}^{\lambda-\delta}) \cdot |V_i| \cdot O(w_{i+1}^{2-\beta+2\eta} w_i^{1+\eta}) \\ &\leq \exp(-cw_{i+1}^{\lambda-\delta}) \cdot |V_i| \cdot O(w_i^{\tau(2-\beta)+1+\eta(2\tau+1)}) \\ &\leq \exp(-cw_{i+1}^{\lambda-\delta}) \cdot |V_i| \cdot E_i \cdot O(w_i^{\eta(3\tau+1)}). \end{aligned}$$

Since we assumed  $w_i \geq 2$ , we may upper-bound the  $O(\cdot)$ -term by  $0.5w_i^D$  for a sufficiently large  $D > 0$ .

Now we can apply Theorem 7 with  $t = 0.5 \exp(-cw_{i+1}^{\lambda-\delta}) \cdot |V_i| \cdot E_i \cdot w_i^D$ . It follows similarly as in the proof of Claim 1 that  $\mathbb{P}[|Z_i - \mathbb{E}[Z_i]| \geq t] = n^{-\omega(1)}$ , and we conclude that with probability  $1 - n^{-\omega(1)}$  it holds that

$$|Z_i| \leq \mathbb{E}[Z_i] + t \leq \exp(-cw_{i+1}^{\lambda-\delta}) \cdot |V_i| \cdot E_i \cdot w_i^D.$$

Now the claim follows by a union bound over all  $O(\log \log n)$  choices of  $i$ .  $\square$

By Lemma 7(ii), w.h.p. every vertex of weight at least  $(\log n)^C$  has small distance from the heavy core. It remains to show that every vertex in the giant component has a large probability

of connecting to such a high-weight vertex in a small number of steps. The next lemma shows that the more vertices of small weight a vertex  $v$  is connected to, the more likely it is that  $v$  has small distance to a vertex of large weight.

**Lemma 9.** (Bulk lemma). *Let  $\varepsilon > 0$ . Let  $w_{\min} \leq w \leq \bar{w}$  be a weight such that there are  $\Omega(n)$  nodes with weight less than  $w$ , and let  $k \geq \max\{2, w^{\beta-1+\varepsilon}\}$  be an integer. For a vertex  $v \in V_{<w}$ , let  $C_v \subseteq V_{<w}$  be the connected component of  $v$  in the graph  $G_{<w}$ . Then for a random vertex  $v \in V_{<w}$ ,*

$$\mathbb{P}_{G,v}[\text{dist}(v, V_{\geq w}) \leq k \text{ or } |C_v| < k] \geq 1 - O(e^{-w^{\Omega(1)}}).$$

*Proof.* For any node  $v \in V_{<w}$ , we define  $N_v \subseteq V_{<w}$  to be the set of all vertices within distance less than  $k$  from  $v$  in the graph  $G_{<w}$ . Observe that  $|C_v| < k$  holds if and only if  $|N_v| < k$ : indeed, if  $|C_v| < k$ , then  $C_v$  contains no path of length  $k$  and thus all nodes in  $C_v$  have distance less than  $k$  to  $v$ , i.e.,  $C_v = N_v$  and thus  $|N_v| = |C_v| < k$ . If  $|C_v| \geq k$ , then either  $C_v = N_v$  and thus  $|N_v| = |C_v| \geq k$ , or  $C_v$  contains some node  $u$  at distance at least  $k$  from  $v$ , but then the first  $k$  nodes on a path from  $v$  to  $u$  are all contained in  $N_v$  and thus  $|N_v| \geq k$ .  $\square$

It follows that proving the bulk lemma is equivalent to proving the following inequality:

$$\mathbb{P}_{G,v}[\text{dist}(v, V_{\geq w}) > k \text{ and } |N_v| \geq k] \leq O(e^{-w^{\Omega(1)}}). \quad (13)$$

Before proving (13), let us sketch some of the ideas of the proof. We first uncover the graph  $G_{<w}$  induced by vertices of weight less than  $w$ . For a fixed realization of  $G_{<w}$ , we consider the probability  $P_v$  that a random node in  $V_{[w,\bar{w}]}$  has a neighbor in  $N_v$ , for an appropriately chosen  $\bar{w} > w$ . We consider two cases: that  $P_v$  is large or that it is small.

If  $P_v$  is large, then it is very likely that there is at least one edge between  $N_v$  and  $V_{[w,\bar{w}]}$ , and thus it is very likely that  $\text{dist}(v, V_{\geq w}) \leq k$ , which proves (13). In a geometric setting, e.g. in hyperbolic random graphs or GIRGs (see Section 3), intuitively this case occurs if the nodes in  $N_v$  are spread out in the geometric space. Note that this argument holds for every fixed realization of  $G_{<w}$  for which  $P_v$  is large.

The case of small  $P_v$  intuitively occurs if the nodes in  $N_v$  form a bulk, i.e., they are concentrated in a small geometric region. For this case, it suffices to bound the number of *bad nodes*, i.e., the size of the set  $B = \{v \in V_{<w} \mid |N_k| \geq k \text{ and } P_v \text{ is small}\}$ . Note that  $B$  is a random variable that is determined by  $G_{<w}$ . In the most technical part of our proof, we show that if  $|B|$  is large, then a node in  $V_{[w,\bar{w}]}$  has a significantly increased probability of having large degree. Since node-degrees in our model are Chernoff-concentrated, this shows that  $|B|$  is very unlikely to be large. So in contrast to the case of large  $P_v$ , we argue (indirectly) about the distribution of  $G_{<w}$ , i.e., we show that  $G_{<w}$  is unlikely to have the property that  $|B|$  is large. See Claims 4, 5, and 6 for the technical details of this argument. We deduce that a random node  $v$  is very unlikely to be bad, which proves the desired inequality (13).

For the formal proof, we let  $0 < \eta \leq \varepsilon/4$  be sufficiently small. Then by the power-law assumption (PL2) we may choose  $\tilde{w} = O(w^{1+\eta})$  such that there are at least  $\Omega(n/w^{\beta-1+\eta})$  vertices with weights between  $w$  and  $\tilde{w}$ . We denote by  $v^*$  a node chosen uniformly at random from  $V_{[w,\tilde{w}]}$ .

We first uncover the graph  $G_{<w}$  induced by vertices of weight less than  $w$ , i.e., we uncover the positions of these vertices and the edges in the induced subgraph. Note that uncovering  $G_{<w}$  in particular yields the sets  $N_v$ ,  $v \in V_{<w}$ . Once  $G_{<w}$  is fixed, consider the probability  $P_v$  that a random node  $v^* \in V_{[w,\tilde{w}]}$  is a neighbor of at least one node in  $N_v$ :

$$P_v := \mathbb{P}[\Gamma(v^*) \cap N_v \neq \emptyset \mid G_{<w}].$$

Note that  $P_v$  depends on the choice of the node  $v$ , and it is a random variable depending on the random choices in  $G_{<w}$ , i.e.,  $P_v$  is a number after uncovering the graph  $G_{<w}$ . The probability in the definition of  $P_v$  runs over the random choice of  $v^* \in V_{[w, \tilde{w}]}$ , the positions of the nodes in  $V_{\geq w}$ , and the edges in  $G$  excluding the induced subgraph  $G_{<w}$ .

Now we choose a parameter  $\tau := w^{\beta-1+2\eta}/n$  and split into two cases:

$$\begin{aligned} &\mathbb{P}_{G,v}[\text{dist}(v, V_{\geq w}) > k \text{ and } |N_v| \geq k] \\ &= \mathbb{P}_{G,v}[\text{dist}(v, V_{\geq w}) > k \text{ and } |N_v| \geq k \text{ and } P_v \geq \tau] \\ &\quad + \mathbb{P}_{G,v}[\text{dist}(v, V_{\geq w}) > k \text{ and } |N_v| \geq k \text{ and } P_v < \tau] \end{aligned} \tag{14}$$

*Case 1: Large  $P_v$ .* To bound the first summand, recall that the size of  $V_{[w, \tilde{w}]}$  is at least  $\Omega(n/w^{\beta-1+\eta})$ . Since the probability that a random node in  $V_{[w, \tilde{w}]}$  has a neighbor in  $N_v$  is equal to  $P_v \geq \tau$ , we obtain

$$\mathbb{E}_{G,v}[\mathbb{1}\{u \in V_{[w, \tilde{w}]} \mid \Gamma(u) \cap N_v \neq \emptyset\} \mid P_v \geq \tau] \geq \Omega\left(\frac{n}{w^{\beta-1+\eta}} \cdot \tau\right) \geq \Omega(w^\eta),$$

where we used  $\tau = w^{\beta-1+2\eta}/n$ . Since every  $u \in V_{[w, \tilde{w}]}$  draws its position and its edges to  $V_{<w}$  independently from the other nodes in  $V_{[w, \tilde{w}]}$ , the Chernoff bound yields

$$\mathbb{P}_{G,v}[\exists u \in V_{[w, \tilde{w}]} : \Gamma(u) \cap N_v \neq \emptyset \mid P_v \geq \tau] \geq 1 - O(e^{-\Omega(w^\eta)}) \geq 1 - O(e^{-w^{\Omega(1)}}). \tag{15}$$

Since every node in  $N_v$  has distance less than  $k$  to  $v$ , if there exists a node  $u \in V_{[w, \tilde{w}]}$  with  $\Gamma(u) \cap N_v \neq \emptyset$  then the distance from  $v$  to  $V_{\geq w} \supseteq V_{[w, \tilde{w}]}$  is at most  $k$ . Thus, (15) implies

$$\mathbb{P}_{G,v}[\text{dist}(v, V_{\geq w}) > k \mid P_v \geq \tau] \leq O(e^{-w^{\Omega(1)}}).$$

This allows us to bound the first summand of Equation (14) by

$$\begin{aligned} &\mathbb{P}_{G,v}[\text{dist}(v, V_{\geq w}) > k \text{ and } |N_v| \geq k \text{ and } P_v \geq \tau] \\ &\leq \mathbb{P}_{G,v}[\text{dist}(v, V_{\geq w}) > k \text{ and } P_v \geq \tau] \leq \mathbb{P}_{G,v}[\text{dist}(v, V_{\geq w}) > k \mid P_v \geq \tau] \leq O(e^{-w^{\Omega(1)}}). \end{aligned}$$

*Case 2: Small  $P_v$ .* For the second summand of Equation (14), we have the bound

$$\begin{aligned} &\mathbb{P}_{G,v}[\text{dist}(v, V_{\geq w}) > k \text{ and } |N_v| \geq k \text{ and } P_v < \tau] \\ &\leq \mathbb{P}_{G,v}[|N_v| \geq k \text{ and } P_v < \tau] = \mathbb{P}_{G_{<w},v}[|N_v| \geq k \text{ and } P_v < \tau], \end{aligned}$$

since  $N_v$  and  $P_v$  depend only on the random choices in  $G_{<w}$ . Now we define the set of *bad nodes*  $B$  as

$$B := \{v \in V_{<w} \mid |N_v| \geq k \text{ and } P_v < \tau\}.$$

Note that  $B$  is a random variable that depends only on the random choices in  $G_{<w}$ . Since  $v$  is a random node in  $V_{<w}$ , and  $|V_{<w}| = \Omega(n)$  by assumption on  $w$ , we obtain

$$\mathbb{P}_{G_{<w},v}[|N_v| \geq k \text{ and } P_v < \tau] = \mathbb{E}_{G_{<w}}[|B|/|V_{<w}|] = O\left(\frac{1}{n}\right) \cdot \mathbb{E}_{G_{<w}}[|B|].$$

It now remains to prove that  $\mathbb{E}_{G_{<w}}[|B|] \leq O(n \cdot e^{-\Omega(w)})$ , which we will do in Claim 6 below. Taken together, these considerations yield a bound of  $O(e^{-\Omega(w)})$  on the second summand of Equation (14).



We have thus shown, modulo Claim 6, that both summands of (14) are  $O(e^{-w^{\Omega(1)}})$ . This yields (13), which as discussed above is equivalent to the lemma statement.

To finish the proof of the bulk lemma it remains to prove  $\mathbb{E}_{G_{<w}}[|B|] \leq O(n \cdot e^{-\Omega(w)})$ ; see Claim 6 below. We prove this through a series of claims.

**Claim 3.** Fix  $G_{<w}$  (and thus the sets  $B$  and  $N_v$  for  $v \in V_{<w}$ ). There exist a set  $\tilde{B} \subseteq B$  and sets  $\tilde{N}_v \subseteq N_v$  for all  $v \in \tilde{B}$  such that the following hold:

- (i) the sets  $\tilde{N}_v$  for  $v \in \tilde{B}$  are disjoint,
- (ii)  $\sum_{v \in \tilde{B}} |\tilde{N}_v| \geq |B|$ , and
- (iii)  $|\tilde{N}_v| \geq k/2$  for all  $v \in \tilde{B}$ .

*Proof.* We greedily construct  $\tilde{B}$  as follows. Initially let  $L := B$ . While  $L$  is non-empty, pick any  $v \in L$ , add  $v$  to  $\tilde{B}$ , and remove all nodes  $u \in N_v$  from  $L$ . □

After constructing  $\tilde{B}$ , assign each node  $u \in \bigcup_{v \in \tilde{B}} N_v$  to its closest node  $v \in \tilde{B}$ , i.e., to the node  $v \in \tilde{B}$  minimizing the distance from  $u$  to  $v$  in  $G_{<w}$  (breaking ties arbitrarily). For any  $v \in \tilde{B}$  let  $\tilde{N}_v$  be the set of all nodes  $u$  that were assigned to  $v$ .

This finishes the construction. We clearly have  $\tilde{B} \subseteq B$ . Since each node  $u \in \bigcup_{v \in \tilde{B}} N_v$  is contained in some set  $N_v$  and thus is at distance less than  $k$  from some  $v \in \tilde{B}$ , and since we assign  $u$  to its closest node in  $\tilde{B}$ , each node  $u$  is assigned to a node within distance  $k$ . It follows that  $\tilde{N}_v \subseteq N_v$ . It remains to show the properties (i)–(iii).

The property (i) holds since we assign each node  $u \in \bigcup_{v \in \tilde{B}} N_v$  to a unique node  $v \in \tilde{B}$ .

For (ii), observe that  $B \subseteq \bigcup_{v \in \tilde{B}} N_v$ , since each  $b \in B$  is initially in  $L$ , and  $b$  is removed from  $L$  only if we add  $b$  to  $\tilde{B}$  or we add another node  $v$  to  $\tilde{B}$  such that  $b \in N_v$ ; in both cases  $b$  ends up in  $\bigcup_{v \in \tilde{B}} N_v$ . Our assignment satisfies  $\bigcup_{v \in \tilde{B}} N_v = \bigcup_{v \in \tilde{B}} \tilde{N}_v$ , and thus  $B \subseteq \bigcup_{v \in \tilde{B}} \tilde{N}_v$ . Since the latter is a disjoint union by (i), we obtain  $|B| \leq \sum_{v \in \tilde{B}} |\tilde{N}_v|$ .

For (iii), fix a node  $v \in \tilde{B}$  and recall that  $N_v = N_v(k, w)$  is the set of all nodes with distance less than  $k$  to  $v$  in  $G_{<w}$ . We similarly define  $N'_v$  as the set of all nodes with distance less than  $k/2$  to  $v$  in  $G_{<w}$ . Note that by construction any two nodes in  $\tilde{B}$  have distance at least  $k$ . Since any node  $u \in N'_v$  is at distance less than  $k/2$  from  $v$ , by the triangle inequality  $u$  has distance at least  $k/2$  from any other node in  $\tilde{B}$ , and thus  $u$  is assigned to  $v$ . This shows that  $N'_v \subseteq \tilde{N}_v$ .

Now if  $N'_v = N_v$ , then we have  $|\tilde{N}_v| \geq |N'_v| = |N_v| \geq k \geq k/2$ , where we used the fact that  $v \in B$  and the definition of  $B = \{v \in V_{<w} \mid |N_v| \geq k \text{ and } P_v < \tau\}$ . Otherwise there exists a node  $u \in N_v \setminus N'_v$ . Consider any path from  $v$  to  $u$  in  $G_{<w}$ . Since  $u \notin N'_v$  this path has length at least  $k/2$ , and the first  $k/2$  of the nodes on the path are contained in  $N'_v$ , so we obtain  $|\tilde{N}_v| \geq |N'_v| \geq k/2$ . In both cases we arrive at  $|\tilde{N}_v| \geq k/2$ .

In what follows, for any outcome of  $G_{<w}$  we fix sets  $\tilde{B}$  and  $\tilde{N}_v$  for all  $v \in \tilde{B}$ , as guaranteed by the above claim. Let  $c > 0$  be such that for every vertex  $v$  of weight at least  $w$ , every vertex  $u \in V$ , and every fixed position  $\mathbf{x}_u \in \mathcal{X}$  we have  $\mathbb{P}[v \sim u \mid \mathbf{x}_u] \geq cw/n$ , i.e.,  $c$  is the hidden constant of the condition (EP1). Recall that  $v^*$  is a random node in  $V_{[w, \tilde{w}]}$ . We set

$$\lambda := \frac{c}{2} w^{2-\beta-2\eta}$$

and define the random variable

$$S := \sum_{v \in \tilde{B}} \lambda |\tilde{N}_v| \cdot \mathbb{1}\{|\Gamma(v^*) \cap \tilde{N}_v| \geq \lambda |\tilde{N}_v|\},$$

where the indicator random variable  $\mathbb{1}\{E\}$  is 1 if the event  $E$  is true and 0 otherwise. In the following we prove upper and lower bounds on  $\mathbb{E}[S]$  and then we combine them to bound  $\mathbb{E}[|B|]$ .

**Claim 4.** *We have  $\mathbb{E}_G[S] \leq O(e^{-\Omega(w)})$ .*

*Proof.* We will first argue that

$$\mathbb{E}_G[S] = \sum_{i \geq 1} i\lambda \cdot \mathbb{P}_G[S = i\lambda] = \sum_{i \geq k/2} i\lambda \cdot \mathbb{P}_G[S = i\lambda] \leq \sum_{i \geq k/2} i\lambda \cdot \mathbb{P}_G[\deg(v^*) \geq i\lambda].$$

Indeed, the first equation is true because  $S$  is an integer multiple of  $\lambda$ . For the second equation, recall that by Claim 3(iii) we have  $|\tilde{N}_v| \geq k/2$  for all  $v \in \tilde{B}$ . It follows that  $S > 0$  implies  $S \geq \lambda k/2$ , so smaller summands are 0. For the last inequality, recall that by Claim 3(i) the sets  $\tilde{N}_v$  are disjoint. Thus,  $S$  is a lower bound on the degree of  $v^*$ .  $\square$

Now consider any node  $u \in V_{[w, \tilde{w}]}$  and the random variable  $\deg(u)$ . Since  $u$  has weight at most  $\tilde{w}$ , its expected degree is  $O(\tilde{w})$ . Moreover, even after conditioning on the position  $\mathbf{x}_u$ , the expected degree of  $u$  is  $O(\tilde{w})$ , and it is the sum of independent random variables; see Lemmas 2 and 3. Therefore, by the Chernoff–Hoeffding bound (Theorem 5), there is a constant  $C > 0$  (independent of  $\mathbf{w}_u$  and  $\mathbf{x}_u$ ) such that  $\mathbb{P}_G[\deg(u) \geq i\lambda] \leq e^{-\Omega(i\lambda)}$  for all  $i \geq C\tilde{w}/\lambda$ . Since this holds for all  $u \in V_{[w, \tilde{w}]}$ , for the random node  $v^* \in V_{[w, \tilde{w}]}$  we also have  $\mathbb{P}_G[\deg(v^*) \geq i\lambda] \leq e^{-\Omega(i\lambda)}$  for all  $i \geq C\tilde{w}/\lambda$ .

Recall that we assume  $k \geq w^{\beta-1+\varepsilon} \geq w^{\beta-1+4\eta}$ . Therefore, we have  $\lambda k/2 = \Omega(w^{2-\beta-2\eta} \cdot w^{\beta-1+4\eta}) = \Omega(w^{1+2\eta}) = \Omega(\tilde{w} \cdot w^\eta)$ , since  $\tilde{w} = O(w^{1+\eta})$ . We first consider the case that  $w$  is at least a sufficiently large constant  $c' > 0$ . Then the  $w^\eta$  factor is large enough to dominate any constant factor, so we can conclude  $\lambda k/2 \geq C\tilde{w}$ . This yields

$$\mathbb{E}_G[S] \leq \sum_{i \geq C\tilde{w}/\lambda} i\lambda \cdot \mathbb{P}_G[\deg(v^*) \geq i\lambda] \leq \sum_{i \geq C\tilde{w}/\lambda} i\lambda \cdot e^{-\Omega(i\lambda)}.$$

To estimate this further, for  $j \in \mathbb{N}$  consider the summands with  $j/\lambda \leq i < (j+1)/\lambda$ . There are  $O(1 + 1/\lambda)$  such summands, each of them having value  $i\lambda e^{-\Omega(i\lambda)} = O(je^{-\Omega(j)})$ . Hence, we may continue as follows:

$$\mathbb{E}_G[S] \leq \sum_{j \geq C\tilde{w}} O((1 + \frac{1}{\lambda}) \cdot je^{-\Omega(j)}) \leq O((1 + \frac{1}{\lambda}) \cdot e^{-\Omega(C\tilde{w})}) \leq O(e^{-\Omega(w)}),$$

where the last inequality uses  $w \leq \tilde{w}$  and  $1/\lambda \leq O(w^{O(1)})$ .

It remains to consider the case that  $w < c'$ , meaning  $w$  is bounded by a constant. In this case it suffices to show that  $\mathbb{E}_G[S] = O(1)$ . This follows from the fact that  $\mathbb{E}_G[S] \leq \mathbb{E}_G[\deg(v^*)] \leq O(\tilde{w}) \leq O(w^{1+\eta}) \leq O(1)$ , since  $w$  is bounded by a constant.

**Claim 5.** *We have  $\mathbb{E}_G[S] \geq \Omega(\frac{1}{mw^{O(1)}}) \cdot \mathbb{E}_{G_{<w}}[|B|]$ .*

*Proof.* Let  $X$  be a random variable taking values in  $[0, M]$ , and let  $x \in [0, M]$ . Then

$$\mathbb{E}[X] \leq \mathbb{P}[X \geq x] \cdot M + \mathbb{P}[0 < X < x] \cdot x \leq \mathbb{P}[X \geq x] \cdot M + \mathbb{P}[X > 0] \cdot x.$$

Rearranging this yields

$$\mathbb{P}[X \geq x] \geq \frac{1}{M} (\mathbb{E}[X] - x \cdot \mathbb{P}[X > 0]).$$

We first uncover  $G_{<w}$ , and then for any  $v \in \tilde{B}$  we apply the above inequality on the random variable  $X = |\Gamma(v^*) \cap \tilde{N}_v|$  given  $G_{<w}$ , which takes values in  $[0, M] = [0, |\tilde{N}_v|]$ . Setting  $x := \lambda|\tilde{N}_v|$ , we obtain

$$\begin{aligned} & \mathbb{P}[|\Gamma(v^*) \cap \tilde{N}_v| \geq \lambda|\tilde{N}_v| \mid G_{<w}] \\ & \geq \frac{1}{|\tilde{N}_v|} \left( \mathbb{E}[|\Gamma(v^*) \cap \tilde{N}_v| \mid G_{<w}] - \lambda|\tilde{N}_v| \cdot \mathbb{P}[|\Gamma(v^*) \cap \tilde{N}_v| \neq 0 \mid G_{<w}] \right). \end{aligned}$$

Recall that we defined  $\lambda = \frac{c}{2}w^{2-\beta-2\eta}$ , where the constant  $c > 0$  is such that for any node  $v$  of weight at least  $w$ , any node  $u$ , and every fixed position  $\mathbf{x}_u$  we have  $\mathbb{P}[v \sim u \mid \mathbf{x}_u] \geq cw/n$ . It follows that even after fixing  $G_{<w}$ , any node  $v^* \in V_{[w, \bar{w}]}$  has probability at least  $cw/n$  of connecting to any node in  $V_{<w}$ . This gives us the bound

$$\mathbb{E}[|\Gamma(v^*) \cap \tilde{N}_v| \mid G_{<w}] \geq |\tilde{N}_v| \cdot \frac{cw}{n}.$$

Using the definition of  $P_v$  and the property  $P_v < \tau$  of any  $v \in \tilde{B} \subseteq B$ , we have

$$\mathbb{P}[|\Gamma(v^*) \cap \tilde{N}_v| \neq 0 \mid G_{<w}] \leq \mathbb{P}[|\Gamma(v^*) \cap N_v| \neq 0 \mid G_{<w}] = P_v < \tau.$$

Combining the three inequalities above yields

$$\mathbb{P}[|\Gamma(v^*) \cap \tilde{N}_v| \geq \lambda|\tilde{N}_v| \mid G_{<w}] \geq \frac{cw}{n} - \lambda\tau.$$

Since we defined  $\lambda = \frac{c}{2}w^{2-\beta-2\eta}$  and  $\tau = w^{\beta-1+2\eta}/n$ , we have  $\lambda\tau = \frac{cw}{2n}$ , and thus

$$\mathbb{P}[|\Gamma(v^*) \cap \tilde{N}_v| \geq \lambda|\tilde{N}_v| \mid G_{<w}] \geq \frac{cw}{2n} \geq \Omega\left(\frac{1}{n}\right).$$

Now we have the bound

$$\mathbb{E}[S \mid G_{<w}] = \sum_{v \in \tilde{B}} \lambda|\tilde{N}_v| \cdot \mathbb{P}[|\Gamma(v^*) \cap \tilde{N}_v| \geq \lambda|\tilde{N}_v| \mid G_{<w}] \geq \Omega\left(\frac{\lambda}{n}\right) \cdot \sum_{v \in \tilde{B}} |\tilde{N}_v|.$$

By Claim 3(ii), the last sum is at least  $|B|$ . Since we have  $\lambda \geq \Omega(1/w^{O(1)})$ , we obtain

$$\mathbb{E}[S \mid G_{<w}] \geq \Omega\left(\frac{1}{nw^{O(1)}}\right) \cdot |B|.$$

So far we assumed that  $G_{<w}$  is fixed. Taking the expectation over  $G_{<w}$  finally yields

$$\mathbb{E}_G[S] = \mathbb{E}_{G_{<w}}[\mathbb{E}[S \mid G_{<w}]] \geq \Omega\left(\frac{1}{nw^{O(1)}}\right) \cdot \mathbb{E}_{G_{<w}}[|B|].$$

□

We are now ready to prove the remaining claim needed in the proof of the bulk lemma.

**Claim 6.** We have  $\mathbb{E}_{G_{<w}}[|B|] \leq O(n \cdot e^{-\Omega(w)})$ .

*Proof.* Combining Claims 4 and 5 and rearranging yields

$$\mathbb{E}_{G_{<w}}[|B|] \leq O(nw^{O(1)}e^{-\Omega(w)}) \leq O(ne^{-\Omega(w)}),$$

as required. □

The upper bounds on the diameter and the average distance now follow easily from the lemmas we have proved so far. We collect the results in the following theorem, which reformulates and specifies Theorem 2 and Theorem 3.

**Theorem 8.** (Components and distances). Let  $2 < \beta < 3$ .

- (i) *W.h.p., there is a giant component, i.e., a connected component which contains  $\Omega(n)$  vertices.*
- (ii) *W.h.p., all other components have at most polylogarithmic size.*
- (iii) *W.h.p., the giant component has polylogarithmic diameter.*
- (iv) *In expectation and with probability  $1 - o(1)$ , the average distance (i.e., the expected distance between two uniformly random vertices in the largest component) is  $\frac{2+o(1)}{\lceil \log \beta - 2 \rceil} \log \log n$ .*
- (v) *With probability  $1 - o(1)$ , a  $(1 - o(1))$ -fraction of all pairs of vertices in the giant component have distance at most  $\frac{2+o(1)}{\lceil \log(\beta-2) \rceil} \log \log n$ .*

*Proof.* (i). This was proved in Lemma 8 and is simply stated here for completeness.

(ii), (iii). We fix a sufficiently small constant  $\varepsilon > 0$  and conclude from the same lemma that w.h.p. the giant component contains all vertices of weight at least  $w := (\log n)^C$ , for a suitable constant  $C > 0$ , and that w.h.p. all such vertices have distance at most  $\frac{1+\varepsilon}{\lceil \log(\beta-2) \rceil} \log \log n$  from the heavy core  $\bar{V}$ . We apply Lemma 9 with  $k = w^{\beta-1+\varepsilon}$ . Then a random vertex in  $V_{<w}$  has probability at least  $1 - e^{-w^{\Omega(1)}}$  of either being at distance at most  $k$  of  $V_{\geq w}$ , or being in a component of size less than  $k$ . Note that for sufficiently large  $C$  this probability is at least  $1 - n^{-\omega(1)}$ . By the union bound, w.h.p. one of the two possibilities happens for all vertices in  $V_{<w}$ . This already shows that w.h.p. all non-giant components are of size less than  $k = (\log n)^{O(1)}$ . For the diameter of the giant component, recall that w.h.p. the heavy core has diameter  $o(\log \log n)$  by Lemma 6. Therefore, w.h.p. the diameter of the giant component is  $O(k + \log \log n) = (\log n)^{O(1)}$ .

(iv). For the average distance, let  $\varepsilon > 0$ , and let  $v \in V$  be a vertex chosen uniformly at random. Fix  $k \geq 3$ ,  $k = n^{o(1)}$ , and let  $w := w(k) = k^{1/\beta}$  so that Lemma 9 is applicable as  $k = w^\beta$ . We sort the vertices by weight and uncover the graph vertex by vertex in increasing order, until either (1) we see for the first time a vertex  $v' \in V_{\geq w}$  such that in the subgraph induced by  $V_{\leq w_{v'}}$  there exists a path of length at most  $k$  from  $v$  to  $v'$ , or (2) we have uncovered the full graph and (1) never happened. If  $w_{v'} \geq w$ , then (1) trivially occurs. Otherwise, by Lemma 9, with probability  $1 - O(\exp(-w^{\Omega(1)}))$  either the event (1) happens or the connected component of  $v$  in  $G$  has size less than  $k$ . In the latter case,  $v$  is not connected to the core and there is nothing to show. Otherwise, we have uncovered only the vertices of weight at most  $w_{v'}$ , which allows us to apply Lemma 7(i) since its statement only depends on vertices of higher weight. By Lemma 7(i), with probability  $1 - O(\exp(-w^{\Omega(\varepsilon)}))$  there is a weight-increasing path from  $v'$  to the heavy core of length at most  $\lambda_\varepsilon := (1 + \varepsilon) \frac{\log \log n}{\lceil \log(\beta-2) \rceil}$ . Summarizing, we have shown that for a random vertex  $v$  and every  $k \geq 3$  with  $k = n^{o(1)}$ ,

$$\mathbb{P}[\infty > \text{dist}(v, V_{\text{core}}) \geq k + \lambda_\varepsilon] \leq e^{-\Omega(w(k)^{\Omega(\varepsilon)})} = O(e^{-k^{\Omega(\varepsilon)}}). \tag{16}$$

Let us first consider the expectation of the average distance; i.e., if  $u, u'$  denote random vertices in the largest component of a random graph  $G$ , then we consider  $\mathbb{E}_G[\mathbb{E}_{u, u'}[\text{dist}(u, u')]]$ . Since  $\text{dist}(u, u') \leq n$  we can condition on any event happening with probability  $1 - n^{-\omega(1)}$ ; in particular we can condition on the event  $\mathcal{E}$  that  $G$  has a giant component containing  $V_{\text{core}}$ , all other components have size  $(\log n)^{O(1)}$ ,  $G$  has diameter  $(\log n)^{O(1)}$ , and finally the core has

diameter  $d_{\text{core}} = o(\log \log n)$ . Moreover, by the bound

$$\text{dist}(u, u') \leq \text{dist}(u, V_{\text{core}}) + \text{dist}(u', V_{\text{core}}) + d_{\text{core}},$$

it suffices to bound  $2 \cdot \mathbb{E}_G[\mathbb{E}_u[\text{dist}(u, V_{\text{core}})] \mid \mathcal{E}] + d_{\text{core}}$ . Now we use the fact  $\mathbb{E}[X] = \sum_{k>0} \mathbb{P}[X \geq k]$  for a random variable  $X$  taking values in  $\mathbb{N}_{\geq 0}$  to obtain the bound

$$\mathbb{E}_u[\text{dist}(u, V_{\text{core}})] \leq \lambda_\varepsilon + \sum_{k=1}^{(\log n)^{O(1)}} \mathbb{P}_u[\text{dist}(u, V_{\text{core}}) \geq k + \lambda_\varepsilon].$$

Note that conditioned on  $\mathcal{E}$ , since  $u$  is chosen uniformly at random from the giant component,  $\text{dist}(u, V_{\text{core}}) < \infty$ . Taking the expectation over  $G$ , conditioned on  $\mathcal{E}$ , we may use (16) to bound the probability that  $\text{dist}(v, V_{\text{core}})$  is too large for a vertex chosen uniformly at random from  $V$ . Since the giant component has size  $\Omega(n)$ , this probability increases at most by a constant factor if we instead choose  $v$  uniformly at random from the giant component. Hence, for every constant  $\varepsilon > 0$  we obtain

$$\mathbb{E}_G[\mathbb{E}_u[\text{dist}(u, V_{\text{core}})] \mid \mathcal{E}] \leq \lambda_\varepsilon + \sum_{k=1}^{(\log n)^{O(1)}} O(e^{-k^{\Omega(\varepsilon)}}) + n^{-\omega(1)}. \tag{17}$$

We now use the inequality

$$\sum_{k=1}^{\infty} e^{-k^\kappa} \leq \int_{x=0}^{\infty} e^{-x^\kappa} dx = \Gamma(1 + 1/\kappa),$$

where  $\Gamma$  is Euler’s gamma function. Since  $\Gamma(x)$  is monotonically increasing on the real axis for  $x \geq 2$  and  $\Gamma(1 + n) = n!$ , we have  $\Gamma(1 + 1/\kappa) \leq \lceil 1/\kappa \rceil! \leq (1/\kappa)^{O(1/\kappa)}$  for  $\kappa \leq 1$ . Plugging this into Equation (17) yields

$$\mathbb{E}_G[\mathbb{E}_u[\text{dist}(u, V_{\text{core}})] \mid \mathcal{E}] \leq \lambda_\varepsilon + O(1/\varepsilon)^{O(1/\varepsilon)} + n^{-\omega(1)}.$$

Note that for sufficiently slowly falling  $\varepsilon = \varepsilon(n) = o(1)$  we have

$$O(1/\varepsilon)^{O(1/\varepsilon)} = o(\log \log n).$$

This yields the desired bound on the expected average distance of

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

For the concentration, we want to show  $\mathbb{P}_G[\mathbb{E}_{u,u'}[\text{dist}(u, u')] \geq 2\lambda_\varepsilon] = o(1)$ , where we choose the same  $\varepsilon(n) = o(1)$  as before. Similarly as before, we may upper-bound  $\mathbb{P}_G[\mathbb{E}_{u,u'}[\text{dist}(u, u')] \geq 2\lambda_\varepsilon]$  by

$$n^{-\omega(1)} + \mathbb{P}_G[2 \cdot \mathbb{E}_u[\text{dist}(u, V_{\text{core}})] + d_{\text{core}} \geq 2\lambda_\varepsilon \mid \mathcal{E}].$$

Let  $\gamma > 0$  be a sufficiently small constant, and let

$$\rho = \rho(n) = \frac{\varepsilon}{3|\log(\beta - 2)|} \cdot \log \log n = o(\log \log n).$$

We claim that for sufficiently large  $n$ ,  $2 \cdot \mathbb{E}_u[\text{dist}(u, V_{\text{core}})] + d_{\text{core}} \geq 2\lambda_\varepsilon$  can only happen if for some  $k > \rho$  we have

$$\mathbb{P}_u[\text{dist}(u, V_{\text{core}}) \geq k + \lambda_{\varepsilon/3}] \geq e^{-2k^{\gamma \cdot \varepsilon}}. \tag{18}$$

Indeed, otherwise we have (conditioned on  $\mathcal{E}$ ), similarly as before,

$$\begin{aligned} \mathbb{E}_u[\text{dist}(u, V_{\text{core}})] &\leq \lambda_{\varepsilon/3} + \rho + \sum_{k=\rho}^{(\log n)^{O(1)}} \mathbb{P}_u[\text{dist}(u, V_{\text{core}}) \geq k + \lambda_{\varepsilon/3}] \\ &\leq \lambda_{2\varepsilon/3} + O(1/\varepsilon)^{O(1/\varepsilon)}, \end{aligned}$$

and thus, indeed  $\mathbb{E}_{u,u'}[\text{dist}(u, u')]$  is at most

$$2 \cdot \mathbb{E}_u[\text{dist}(u, V_{\text{core}})] + d_{\text{core}} \leq 2\lambda_{2\varepsilon/3} + o(\log \log n) < 2\lambda_\varepsilon = 2\lambda_{o(1)},$$

if  $\varepsilon = \varepsilon(n) = o(1)$  decreases sufficiently slowly. However, using the union bound over all  $\rho \leq k \leq (\log n)^{O(1)}$ , we see that the probability that  $G$  is such that (18) holds for some  $k > \rho$  is bounded from above by

$$\sum_{k=\rho}^{(\log n)^{O(1)}} \mathbb{P}_G[\mathbb{P}_u[\text{dist}(u, V_{\text{core}}) \geq k + \lambda_{\varepsilon/3}] \geq e^{-k^{\gamma \cdot \varepsilon}} \mid \mathcal{E}].$$

By (16) it follows that

$$\mathbb{E}_G[\mathbb{P}_u[\infty > \text{dist}(u, V_{\text{core}}) \geq k + \lambda_{\varepsilon/3}]] \leq O(\exp(-2k^{\gamma \cdot \varepsilon})),$$

for  $\gamma > 0$  sufficiently small. We apply Markov's inequality and deduce that

$$\mathbb{P}_G[\mathbb{P}_u[\infty > \text{dist}(u, V_{\text{core}}) \geq k + \lambda_{\varepsilon/3}] > e^{-k^{\gamma \cdot \varepsilon}}] \leq O(e^{-k^{\Omega(\varepsilon)}}).$$

Because the giant component has linear size, this probability increases at most by a constant factor if we instead draw  $v$  from the giant component (conditioned on  $\mathcal{E}$ ). Thus, the desired probability is bounded by

$$\sum_{k=\rho}^{(\log n)^{O(1)}} O(e^{-k^{\Omega(\varepsilon)}}), \tag{19}$$

which is  $o(1)$ , since  $\rho = \varepsilon \cdot \log \log n$  grows sufficiently quickly compared to a sufficiently slowly falling  $\varepsilon = o(1)$ . This shows the concentration of the average distance and proves the statement (iv).

(v). Regarding the last statement (v), (19) shows that with probability  $1 - o(1)$   $G$  is such that (18) does not hold for any  $k > \rho$ . However, in this case the fraction of pairs  $\{u, u'\}$  of vertices in the giant component that have distance at least  $2k + 2\lambda_{\varepsilon/3}$  is at most  $e^{-4k^{\gamma \cdot \varepsilon}}$ . Taking  $k = 2\rho$  and assuming that  $\rho = \varepsilon \log \log n$  grows sufficiently quickly compared to  $\varepsilon = o(1)$ , we see that a  $(1 - o(1))$ -fraction of pairs  $\{u, u'\}$  have distance at most  $2\lambda_{o(1)}$ , given that  $\rho = \varepsilon \cdot \log \log n$  grows sufficiently fast compared to  $\varepsilon = o(1)$ . This finishes the proof of Theorem 8.

## 8. Degree sequence

By definition of the model, we are assuming that the weight sequence  $\mathbf{w}$  follows a power law. Since the expected degree of a vertex with weight  $w_v$  is  $\Theta(w_v)$  by Lemma 3, it is not surprising that the degree sequence of the random graph will also follow a power law. In this section, we give details and prove Theorem 1, using Theorem 7 to show concentration. Some of the ideas in our proof are based on [28]. We start with the maximum degree  $\Delta(G)$ , which is a simple corollary of Lemma 4.

**Corollary 1.** *W.h.p.,  $\Delta(G) = \Theta(\mathbf{w}_{\max})$ , where  $\mathbf{w}_{\max} = \max\{w_v \mid v \in V\}$ . In particular, for all  $\eta > 0$ , w.h.p.,  $\Delta(G) = \Omega(\bar{w})$  and  $\Delta(G) = O(n^{1/(\beta-1-\eta)})$ .*

*Proof.* We deduce from the model definition that  $\omega(\log^2 n) \leq \bar{w} \leq \mathbf{w}_{\max} = O(n^{1/(\beta-1-\eta)})$ . Then Lemma 4 directly implies the statement.  $\square$

Next, we calculate the expected number of vertices that have degree at least  $d$ .

**Lemma 10.** *Let  $\eta > 0$  be sufficiently small. Then for all  $d \geq 1$ ,  $d = d(n) = o(\bar{w})$ , we have*

$$\Omega(nd^{1-\beta-\eta}) \leq \mathbb{E}[\#\{v \in V \mid \deg(v) \geq d\}] \leq O(nd^{1-\beta+\eta}).$$

*Proof.* Let  $\eta > 0$  be sufficiently small. Recall that by Lemma 3, it holds that  $\mathbb{E}[\deg(v)] = \Theta(w_v)$  for every vertex  $v \in V$ . Let  $1 \leq d \ll \bar{w}$  and let  $v$  be any vertex with weight  $w_v \geq \Omega(d)$  large enough so that  $\mathbb{E}[\deg(v)] \geq 2d$ . Then by a Chernoff bound

$$\mathbb{P}[\deg(v) < d] \leq \mathbb{P}[\deg(v) < 0.5\mathbb{E}[\deg(v)]] \leq e^{-\mathbb{E}[\deg(v)]/8} \leq e^{-d/4} \leq e^{-1/4}.$$

By the power-law assumption (PL2), there are  $\Omega(nd^{1-\beta-\eta})$  vertices with weight  $\Omega(d)$ , and a single vertex in this set has degree at least  $d$  with probability at least  $1 - e^{-1/4}$ . By linearity of expectation,  $\mathbb{E}[\#\{v \in V \mid \deg(v) \geq d\}] = \sum_{v \in [n]} \mathbb{P}[\deg(v) \geq d] = \Omega(nd^{1-\beta-\eta})$ .  $\square$

Next let  $v$  be a vertex with weight  $w_v \leq O(d)$  small enough so that  $2e\mathbb{E}[\deg(v)] \leq 3d/4$ . By a Chernoff bound (Theorem 5(iii)), we obtain

$$\mathbb{P}[\deg(v) \geq d] \leq \mathbb{P}[\deg(v) > 3d/4] \leq 2^{-3d/4}.$$

Thus, for the upper bound it follows that

$$\begin{aligned} \mathbb{E}[\#\{v \in V \mid \deg(v) \geq d\}] &= \sum_{v \in [n]} \mathbb{P}[\deg(v) \geq d] \\ &\leq |V_{\geq O(d)}| + \sum_{v \in V_{\leq O(d)}} \mathbb{P}[\deg(v) \geq d] \\ &\leq O(nd^{1-\beta+\eta}) + n \cdot 2^{-3d/4}. \end{aligned}$$

Note that  $d^2 \leq 3 \cdot 2^{3d/4}$  holds for all  $d \geq 1$ . Hence  $n \cdot 2^{-3d/4} \leq 3nd^{-2} < 3nd^{1-\beta+\eta}$  and indeed it holds that  $\mathbb{E}[\#\{v \in V \mid \deg(v) \geq d\}] = O(nd^{1-\beta+\eta})$ .

After these preparations we come to the main theorem of this section, which is a more precise formulation of Theorem 1 and states that the degree sequence follows a power law with the same exponent  $\beta$  as the weight sequence.

**Theorem 9.** *For all  $\eta > 0$ , w.h.p. we have*

$$\Omega(nd^{1-\beta-\eta}) \leq \#\{v \in V \mid \deg(v) \geq d\} \leq O(nd^{1-\beta+\eta}),$$

where the first inequality holds for all  $1 \leq d \leq \bar{w}$  and the second inequality holds for all  $d \geq 1$ .

Before we prove Theorem 9, we note that by combining it with the standard calculations from Lemma 1 we immediately obtain the average degree in the graph.

**Corollary 2.** *W.h.p.,  $\frac{1}{n} \sum_{v \in V} \deg(v) = \Theta(1)$  and thus  $|E| = \Theta(n)$ .*

*Proof of Theorem 9.* We first consider the case where  $d$  is larger than  $\log^3 n = o(\bar{w})$ . From the condition (PL2) on the vertex weights and Lemma 3 it follows that

$$\#\{v \in V \mid \mathbb{E}[\deg(v)] \geq 1.5d\} = \Omega(nd^{1-\beta-\eta})$$

holds for all  $\log^3 n \leq d \leq \bar{w}$ . Then, by Lemma 4, w.h.p. every vertex  $v$  with  $\mathbb{E}[\deg(v)] \geq 1.5d$  has degree at least  $(1 - o(1))1.5d \geq d$  for  $n$  large enough. Hence w.h.p. there exist at least  $\Omega(nd^{1-\beta-\eta})$  vertices with degree at least  $d$ . Vice versa, by Lemma 3 we have

$$\#\{v \in V \mid \mathbb{E}[\deg(v)] \geq 0.5d\} = O(nd^{1-\beta+\eta}).$$

By the same arguments as above, w.h.p. every vertex  $v$  with  $\mathbb{E}[\deg(v)] < 0.5d$  has degree at most  $(1 + o(1))0.5d < d$ . Thus in total there can be at most  $O(nd^{1-\beta+\eta})$  vertices with degree at least  $d$ . This proves the theorem for  $d \geq \log^3 n$ .

Let  $1 \leq d \leq \log^3 n$ , let  $\varepsilon > 0$  be sufficiently small, let  $V' := V_{\leq n^\varepsilon}$  be the set of small-weight vertices, and let  $G' := G[V']$ . First, we introduce some notation: define the two random variables

$$g_d := \#\{v \in V \mid \deg(v) \geq d\} \quad \text{and} \quad f_d := \#\{v \in V' \mid \deg_{G'}(v) \geq d\}.$$

Note that by Lemma 10, we already have

$$\Omega(nd^{1-\beta-\eta}) \leq \mathbb{E}[g_d] \leq O(nd^{1-\beta+\eta}),$$

and it remains to prove concentration. Clearly,

$$f_d \leq g_d \leq f_d + 2 \sum_{v \in V \setminus V'} \deg(v). \tag{20}$$

Next we apply Lemma 4 together with Lemma 1 and see that w.h.p.,

$$\sum_{v \in V \setminus V'} \deg(v) = \Theta(W_{\geq n^\varepsilon}) = O(n^{1+(2-\beta+\eta)\varepsilon}) = n^{1-\Omega(1)}.$$

Recall that we assume  $d \leq \log^3 n$ , so in particular

$$\mathbb{E}[g_d] = \Omega(n/(\log n)^{3(\beta-1+\eta)}).$$

It follows that  $\mathbb{E}[\sum_{v \in V \setminus V'} \deg(v)] = o(\mathbb{E}[g_d])$ . The inequality (20) thus implies  $\mathbb{E}[f_d] = (1 + o(1))\mathbb{E}[g_d]$ . Hence, it is sufficient to prove that the random variable  $f_d$  is concentrated around its expectation, because this will transfer immediately to  $g_d$ .

We aim to show this concentration result via Theorem 7. Similarly as in the proof of Claim 1, we can assume that the probability space  $\Omega$  under consideration is a product space of independent random variables. More precisely, the  $n$  independent random variables  $x_1, \dots, x_n$



define the vertex set and the  $n - 1$  independent random variables  $Y_2, \dots, Y_n$  define the edge set, where each  $Y_u$  has the form  $(Y_u^1, \dots, Y_u^{u-1})$ , each  $Y_u^v$  is a real number chosen uniformly at random from  $[0, 1]$ , and for  $v < u$ , the edge  $\{u, v\}$  is present in the graph if and only if  $p_{uv} > Y_u^v$ . The  $2n - 1$  random variables then define the product probability space  $\Omega$ , i.e., for every  $\omega \in \Omega$ , we denote by  $G(\omega)$  the resulting graph, and similarly we use  $G' = G'(\omega)$  and  $f_d = f_d(\omega)$ . We now consider the bad event:

$$\mathcal{B} := \{\omega \in \Omega: \text{the maximum degree in } G'(\omega) \text{ is at least } n^{2\varepsilon}\}. \tag{21}$$

We observe that  $\mathbb{P}[\mathcal{B}] = n^{-\omega(1)}$ , since by Lemma 4 w.h.p. every vertex  $v \in V'$  has degree at most  $O(\mathbf{w}_v + \log^2 n) = o(n^{2\varepsilon})$ . Let  $\omega, \omega' \in \bar{\mathcal{B}}$  be such that they differ in at most two coordinates. We observe that changing one coordinate  $x_i$  or  $Y_i$  can influence only the degrees of  $i$  itself and of the vertices which are neighbors of  $i$  either before or after the coordinate change. It follows that  $|f_d(\omega) - f_d(\omega')| \leq 4n^{2\varepsilon} =: c$ . Therefore,  $f_d$  satisfies the Lipschitz condition of Theorem 7 with bad event  $\mathcal{B}$ . Let  $t = n^{1-\varepsilon} = o(\mathbb{E}[f_d])$ . Then since  $n\mathbb{P}[\mathcal{B}] = n^{-\omega(1)}$ , Theorem 7 implies

$$\begin{aligned} \mathbb{P}[|f_d - \mathbb{E}[f_d]| \geq t] &\leq 2e^{-\frac{t^2}{64c^2n}} + \left(\frac{4n^2}{c} + 1\right)\mathbb{P}[\mathcal{B}] \\ &= e^{-\Omega(n^{1-4\varepsilon})} + n^{-\omega(1)} = n^{-\omega(1)}, \end{aligned}$$

which proves concentration and concludes the proof. □

### 9. Proof of Theorem 4

In this section we prove Theorem 4, which states that for any connection function  $p$  satisfying Equation 2, the corresponding distance model satisfies the conditions (EP1) and (EP2), and thus the results from Section 2.2 on degree sequences, component structure, and average distances apply.

*Proof of Theorem 4.* Fix  $u, v$ , and also the position  $\mathbf{x}_u$ . Note that  $V(r) = \mathbb{P}_{\mathbf{x}_v}(\|\mathbf{x}_u - \mathbf{x}_v\| \leq r)$ , so  $V(r)$  describes the cumulative probability distribution of the random variable  $\|\mathbf{x}_u - \mathbf{x}_v\|$ . The marginal edge probability is given by the Riemann–Stieltjes integral over  $r$ ,

$$E := \mathbb{E}_{\mathbf{x}_v}[p_{uv}(\mathbf{x}_u, \mathbf{x}_v) \mid \mathbf{x}_u] = \Theta\left(\int_0^\infty \Lambda_{u,v}(r) dV(r)\right),$$

where

$$\Lambda_{u,v}(r) := \min\left\{1, V(r)^{-\alpha} \cdot \left(\frac{\mathbf{w}_u \mathbf{w}_v}{\mathbf{w}}\right)^{\max\{\alpha, 1\}}\right\}.$$

In particular, for every sequence of partitions  $r^{(t)} = \{0 = r_0^{(t)} < \dots < r_{\ell(t)}^{(t)}\}$  with meshes tending to zero, the upper Darboux sum with respect to  $r^{(t)}$  converges to the expectation,

$$E = \Theta\left(\lim_{t \rightarrow \infty} \sum_{s=1}^{\ell(t)} \left(\sup_{r_{s-1}^{(t)} \leq r \leq r_s^{(t)}} \Lambda_{u,v}(r)\right) (V(r_s^{(t)}) - V(r_{s-1}^{(t)}))\right).$$

Since  $V$  is surjective, we may refine the meshes  $r^{(t)}$  if necessary so that the meshes of the partitions  $V^{(t)} = \{V(r_0^{(t)}), \dots, V(r_{\ell(t)}^{(t)})\} =: \{V_0^{(t)}, \dots, V_{\ell(t)}^{(t)}\}$  also tend to zero. Hence,

$$\begin{aligned}
 E &= \Theta \left( \lim_{t \rightarrow \infty} \sum_{s=1}^{\ell(t)} \min \left\{ 1, (V_s^{(t)})^{-\alpha} \cdot \left( \frac{W_u W_v}{W} \right)^{\max\{\alpha, 1\}} \right\} \left( V_s^{(t)} - V_{s-1}^{(t)} \right) \right) \\
 &= \Theta \left( \int_0^1 \min \left\{ 1, V^{-\alpha} \cdot \left( \frac{W_u W_v}{W} \right)^{\max\{\alpha, 1\}} \right\} dV \right),
 \end{aligned}$$

where the latter integral is an ordinary Riemann integral. If  $w_u w_v / W \geq 1$ , the integrand is 1 and we obtain  $E = \Theta(1) = \Theta \left( \min \left\{ 1, \frac{W_u W_v}{W} \right\} \right)$ . On the other hand, if  $w_u w_v / W < 1$ , then let  $r_0 := \left( \frac{W_u W_v}{W} \right)^{\max\{\alpha, 1\} / \alpha} < 1$ . Note that if  $r_0 = \Theta(1)$ , then also  $r_0 = \Theta(w_u w_v / W)$ . Therefore,

$$\begin{aligned}
 E &= \Theta \left( \int_0^{r_0} 1 dV + \left( \frac{W_u W_v}{W} \right)^{\max\{\alpha, 1\}} \int_{r_0}^1 V^{-\alpha} dV \right) \\
 &= \begin{cases} \Theta \left( r_0 + \frac{w_u w_v}{W} (1 - r_0^{1-\alpha}) \right) = \Theta \left( \frac{w_u w_v}{W} \right) & , \text{ if } \alpha < 1, \text{ and} \\ \Theta \left( r_0 + \left( \frac{w_u w_v}{W} \right)^\alpha (r_0^{1-\alpha} - 1) \right) = \Theta \left( \frac{w_u w_v}{W} \right) & , \text{ if } \alpha > 1, \end{cases}
 \end{aligned}$$

as required.

It remains to show that  $p$  satisfies (EP2). Since  $V(\|x_u - x_v\|) \leq 1$ , from Equation (2) we obtain the lower bound

$$p_{uv} \geq \Omega \left( \min \left\{ 1, \left( \frac{W_u W_v}{W} \right)^{\max\{\alpha, 1\}} \right\} \right).$$

If  $w_u w_v / W \geq 1$  then there is nothing to show (since the right-hand side of (EP2) is  $o(1)$  by the upper bound on  $\bar{w}$ ). Otherwise, if  $w_u w_v / W < 1$ , then

$$p_{uv} \geq \Omega \left( \left( \frac{W_u W_v}{W} \right)^{\max\{\alpha, 1\}} \right) \geq \Omega \left( \frac{\bar{w}^2}{n} \right) \geq \left( \frac{n}{\bar{w}^{\beta-1+\eta}} \right)^{-1+\omega(1/\log \log n)},$$

where the last step follows from the lower bound on  $\bar{w}$ . This concludes the proof. □

### Acknowledgements

We thank Hafsteinn Einarsson, Tobias Friedrich, and Anton Krohmer for helpful discussions.

### Funding information

There are no funding bodies to thank in relation to the creation of this article.

### Competing interests

There were no competing interests to declare which arose during the preparation or publication process of this article.

### References

- [1] ABDULLAH, M. A., BODE, M. AND MÜLLER, T. (2017). Typical distances in a geometric model for complex networks. *Internet Math.* **2017**, 33 pp.
- [2] AIELLO, W. et al. (2008). A spatial web graph model with local influence regions. *Internet Math.* **5**, 175–196.
- [3] BARABÁSI, A.-L. AND ALBERT, R. (1999). Emergence of scaling in random networks. *Science* **286**, 509–512.

- [4] BLÄSIUS, T. *et al.* (2022). Efficient shortest paths in scale-free networks with underlying hyperbolic geometry. *ACM Trans. Algorithms* **18**, 1–32.
- [5] BLÄSIUS, T., FRIEDRICH, T. AND KROHMER, A. (2016). Hyperbolic random graphs: separators and treewidth. In *24th Annual European Symposium on Algorithms (ESA 2016)*, Dagstuhl Publishing, Saarbrücken, pp. 15:1–15:16.
- [6] BLÄSIUS, T., FRIEDRICH, T. AND KROHMER, A. (2018). Cliques in hyperbolic random graphs. *Algorithmica* **80**, 2324–2344.
- [7] BODE, M., FOUNTOULAKIS, N. AND MÜLLER, T. (2015). On a geometrisation of the Chung–Lu model for complex networks. *J. Complex Networks* **3**, 361–387.
- [8] BOGUŃÁ, M., PAPADOPOULOS, F. AND KRIOUKOV, D. (2010). Sustaining the Internet with hyperbolic mapping. *Nature Commun.* **1**, article no. 62.
- [9] BOLLOBÁS, B., JANSON, S. AND RIORDAN, O. (2007). The phase transition in inhomogeneous random graphs. *Random Structures Algorithms* **31**, 3–122.
- [10] BONATO, A., JANSSEN, J. AND PRAFAT, P. (2010). A geometric model for on-line social networks. In *MSM '10: Proceedings of the International Workshop on Modeling Social Media*, Association for Computing Machinery, New York, article no. 4.
- [11] Bradonjić, M., Hagberg, A. and Percus, A. G. (2008). The structure of geographical threshold graphs. *Internet Math.* **5**, 113–139.
- [12] BRINGMANN, K. *et al.* (2022). Greedy routing and the algorithmic small-world phenomenon. *J. Comput. System Sci.* **125**, 59–105.
- [13] BRINGMANN, K., KEUSCH, R. AND LENGLER, J. (2017). Sampling geometric inhomogeneous random graphs in linear time. In *25th Annual European Symposium on Algorithms (ESA 2017)*, Dagstuhl Publishing, Saarbrücken, pp. 20:1–20:15.
- [14] BRINGMANN, K., KEUSCH, R. AND LENGLER, J. (2019). Geometric inhomogeneous random graphs. *Theoret. Comput. Sci.* **760**, 35–54.
- [15] CHUNG, F. AND LU, L. (2002). The average distances in random graphs with given expected degrees. *Proc. Nat. Acad. Sci. USA* **99**, 15879–15882.
- [16] CHUNG, F. AND LU, L. (2002). Connected components in random graphs with given expected degree sequences. *Ann. Combinatorics* **6**, 125–145.
- [17] CHUNG, F. AND LU, L. (2004). The average distance in a random graph with given expected degrees. *Internet Math.* **1**, 91–113.
- [18] COMBES, R. (2015). *An extension of McDiarmid's inequality*. Preprint. Available at <https://arxiv.org/abs/1511.05240>.
- [19] DEIFEN, M., VAN DER HOFSTAD, R. AND HOOGHIEMSTRA, G. (2013). Scale-free percolation. *Ann. Inst. H. Poincaré Prob. Statist.* **49**, 817–838.
- [20] DEPREZ, P. AND WÜTHRICH, M. V. (2019). Scale-free percolation in continuum space. *Commun. Math. Statist.* **7**, 269–308.
- [21] DEREICH, S., MÖNCH, C. AND MÖRTERS, P. (2012). Typical distances in ultrasmall random networks. *Adv. Appl. Prob.* **44**, 583–601.
- [22] DOROGOVTSSEV, S. N. AND MENDES, J. F. F. (2002). Evolution of networks. *Adv. Phys.* **51**, 1079–1187.
- [23] DRAIEF, M. AND MASSOULI, L. (2010). *Epidemics and Rumours in Complex Networks*, 1st edn. Cambridge University Press.
- [24] DUBHASHI, P. D. AND PANCONESI, A. (2009). *Concentration of Measure for the Analysis of Randomized Algorithms*. Cambridge University Press.
- [25] FRIEDRICH, T. AND KROHMER, A. (2018). On the diameter of hyperbolic random graphs. *SIAM J. Discrete Math.* **32**, 1314–1334.
- [26] GOLDBERG, L. A., JORRITSMA, J., KOMJÁTHY, J. AND LAPINSKAS, J. (2021). Increasing efficacy of contact-tracing applications by user referrals and stricter quarantining. *PLoS ONE* **16**, article no. e0250435.
- [27] GRACAR, P., GRAUER, A. AND MÖRTERS, P. (2022). Chemical distance in geometric random graphs with long edges and scale-free degree distribution. *Commun. Math. Phys.* **395**, 859–906.
- [28] GUGELMANN, L., PANAGIOTOU, K. AND PETER, U. (2012). Random hyperbolic graphs: degree sequence and clustering. In *ICALP '12: Proceedings of the 39th International Colloquium on Automata, Languages, and Programming*, Springer, Berlin, pp. 573–585.
- [29] JACOB, E. AND MÖRTERS, P. (2013). A spatial preferential attachment model with local clustering. In *Algorithms and Models for the Web Graph*, Springer, Cham, pp. 14–25.
- [30] JORRITSMA, J., HULSHOF, T. AND KOMJÁTHY, J. (2020). Not all interventions are equal for the height of the second peak. *Chaos Solitons Fractals* **139**, article no. 109965.
- [31] KIWI, M. AND MITSCHKE, D. (2018). Spectral gap of random hyperbolic graphs and related parameters. *Ann. Appl. Prob.* **28**, 941–989.

- [32] KIWI, M. AND MITSCHKE, D. (2019). On the second largest component of random hyperbolic graphs. *SIAM J. Discrete Math.* **33**, 2200–2217.
- [33] KOCH, C. AND LENGLER, J. (2021). Bootstrap percolation on geometric inhomogeneous random graphs. *Internet Math.* **2021**, 36 pp.
- [34] KOMJÁTHY, J., LAPINSKAS, J. A. AND LENGLER, J. (2021). Stopping explosion by penalising transmission to hubs in scale-free spatial random graphs. In *Ann. Inst. H. Poincaré Prob. Statist.* **57**, 1968–2016.
- [35] KOMJÁTHY, J. AND LODEWIJKS, B. (2020). Explosion in weighted hyperbolic random graphs and geometric inhomogeneous random graphs. *Stoch. Process. Appl.* **130**, 1309–1367.
- [36] KRIOUKOV, D. *et al.* (2010). Hyperbolic geometry of complex networks. *Phys. Rev. E* **82**, article no. 036106.
- [37] KUTIN, S. (2002). Extensions to McDiarmid’s inequality when differences are bounded with high probability. Tech. Rep., University of Chicago.
- [38] LENGLER, J. AND TODOROVIC, L. (2017). *Existence of small separators depends on geometry for geometric inhomogeneous random graphs*. Preprint. Available at <https://arxiv.org/abs/1711.03814>.
- [39] MÜLLER, T. AND STAPS, M. (2019). The diameter of KPKVB random graphs. *Adv. Appl. Prob.* **51**, 358–377.
- [40] PAPADOPOULOS, F., KRIOUKOV, D., BOGUŃÁ, M. AND VAHDAT, A. (2010). Greedy forwarding in dynamic scale-free networks embedded in hyperbolic metric spaces. In *2010 Proceedings IEEE INFOCOM*, Institute of Electrical and Electronics Engineers, pp. 1–9.
- [41] PENROSE, M. (2003). *Random Geometric Graphs*. Oxford University Press.
- [42] SERRANO, M. A., KRIOUKOV, D. AND BOGUŃÁ, M. (2008). Self-similarity of complex networks and hidden metric spaces. *Phys. Rev. Lett.* **100**, article no. 078701.
- [43] VAN DER HOFSTAD, R. (2024). *Random Graphs and Complex Networks*, Vol. **2**. Cambridge University Press.
- [44] VAN DER HOFSTAD, R. AND KOMJÁTHY, J. (2017). *Explosion and distances in scale-free percolation*. Preprint. Available at <https://arxiv.org/abs/1706.02597>.
- [45] VAN DER HOFSTAD, R., VAN DER HOORN, P. AND MAITRA, N. (2023). Local limits of spatial inhomogeneous random graphs. *Adv. Appl. Prob.* **55**, 793–840.
- [46] WARNKE, L. (2016). On the method of typical bounded differences. *Combinatorics Prob. Comput.* **25**, 269–299.