

# On the Diameter of Hyperbolic Random Graphs

Tobias Friedrich<sup>1,2</sup> and Anton Krohmer<sup>1,2</sup>(✉)

<sup>1</sup> Friedrich-Schiller-Universität Jena, Jena, Germany

<sup>2</sup> Hasso Plattner Institute, Potsdam, Germany  
anton.krohmer@hpi.de

**Abstract.** Large real-world networks are typically scale-free. Recent research has shown that such graphs are described best in a geometric space. More precisely, the internet can be mapped to a hyperbolic space such that geometric greedy routing performs close to optimal (Boguná, Papadopoulos, and Krioukov. *Nature Communications*, 1:62, 2010). This observation pushed the interest in hyperbolic networks as a natural model for scale-free networks. Hyperbolic random graphs follow a power-law degree distribution with controllable exponent  $\beta$  and show high clustering (Gugelmann, Panagiotou, and Peter. *ICALP*, pp. 573–585, 2012).

For understanding the structure of the resulting graphs and for analyzing the behavior of network algorithms, the next question is bounding the size of the diameter. The only known bound is  $\mathcal{O}((\log n)^{32/((3-\beta)(5-\beta))})$  (Kiwi and Mitsche. *ANALCO*, pp. 26–39, 2015). We present two much simpler proofs for an improved upper bound of  $\mathcal{O}((\log n)^{2/(3-\beta)})$  and a lower bound of  $\Omega(\log n)$ . If the average degree is bounded from above by some constant, we show that the latter bound is tight by proving an upper bound of  $\mathcal{O}(\log n)$  for the diameter.

## 1 Introduction

Large real-world networks are almost always sparse and non-regular. Their degree distribution typically follows a *power law*, which is synonymously used for being *scale-free*. Since the 1960's, large networks have been studied in detail and hundreds of models were suggested. In the past few years, a new line of research emerged, which showed that scale-free networks can be modeled more realistically when incorporating *geometry*.

**Euclidean Random Graphs.** It is not new to study graphs in a geometric space. In fact, graphs with *Euclidean geometry* have been studied intensively for more than a decade. The standard Euclidean model are random geometric graphs which result from placing  $n$  nodes independently and uniformly at random on an Euclidean space, and creating edges between pairs of nodes if and only if their distance is at most some fixed threshold  $r$ . These graphs have been studied in relation to subjects such as cluster analysis, statistical physics, hypothesis testing, and wireless sensor networks [23]. The resulting graphs are more or less regular and hence do not show a scale-free behavior with power-law degree distribution as observed in large real-world graphs.

**Table 1.** Known diameter bounds for various random graphs. In all cases the diameter depends on the choice of the model parameters. Here we consider a constant average degree. For scale-free networks, we also assume a power law exponent  $2 < \beta < 3$ .<sup>1</sup>

Random Graph Model	Diameter
Sparse Erdős-Rényi [5]	$\Theta(\log n)$ [24]
$d$ -dim. Euclidean [23]	$\Theta(n^{1/d})$ [15]
Watts-Strogatz [26]	$\Theta(\log n)$ [6]
Kleinberg [18]	$\Theta(\log n)$ [21]
Chung-Lu [8]	$\Theta(\log n)$ [8]
Pref. Attachment [1]	$\Theta(\log \log n)$ [10]
Hyperbolic [19]	$\mathcal{O}((\log n)^{\frac{32}{(3-\beta)(5-\beta)}})$ [17]

} power-law graphs

**Hyperbolic Random Graphs.** For modeling scale-free graphs, it is natural to apply a non-Euclidean geometry with negative curvature. Krioukov et al. [19] introduced a new graph model based on *hyperbolic geometry*. Similar to euclidean random graphs, nodes are uniformly distributed in a hyperbolic space and two nodes are connected if their hyperbolic distance is small. The resulting graphs have many properties observed in large real-world networks. This was impressively demonstrated by Boguná et al. [4]: They computed a maximum likelihood fit of the internet graph in the hyperbolic space and showed that greedy routing in this hyperbolic space finds nearly optimal shortest paths in the internet graph. The quality of this embedding is an indication that hyperbolic geometry naturally appears in large scale-free graphs.

**Known Properties.** A number of properties of hyperbolic random graphs have been studied. Gugelmann et al. [16] compute exact asymptotic expressions for the expected number of vertices of degree  $k$  and prove a constant lower bound for the clustering coefficient. They confirm that the clustering is non-vanishing and that the degree sequence follows a power-law distribution with controllable exponent  $\beta$ . For  $2 < \beta < 3$ , the hyperbolic random graph has a giant component of size  $\Omega(n)$  [2, 3], similar to other scale-free networks like Chung-Lu [8]. Other studied properties include the clique number [14], bootstrap percolation [7]; as well as algorithms for efficient generation of hyperbolic random graphs [25] and efficient embedding of real networks in the hyperbolic plane [22].

**Diameter.** The diameter, the length of the longest shortest path, is a fundamental property of a network. It also sets a worst-case lower bound on the number of steps required for all communication processes on the graph. In contrast to the average distance, it is determined by a single—atypical—long path. Due to this sensitivity to small changes, it is notoriously hard to analyze. Even subtle

<sup>1</sup> Note that the table therefore refers to a non-standard Preferential Attachment version with adjustable power law exponent  $2 < \beta < 3$  (normally,  $\beta = 3$ ).

changes to the graph model can make an exponential difference in the diameter, as can be seen when comparing Chung-Lu (CL) random graphs [8] and Preferential Attachment (PA) graphs [1] in the considered range of the power law exponent  $2 < \beta < 3$ : On the one hand, we can embed a CL graph in the PA graph and they behave effectively the same [13]; on the other hand, the diameter of CL graphs is  $\Theta(\log n)$  [8] while for PA graphs it is  $\Theta(\log \log n)$  [10]. Table 1 provides an overview over existing results. It was open so far how the diameter of hyperbolic random graphs compares to the aforementioned bounds for other scale-free graph models. The only known result for their diameter is  $\mathcal{O}((\log n)^{\frac{32}{(3-\beta)^2(5-\beta)}})$  by Kiwi and Mitsche [17].

**Our Contribution.** We improve upon the result of Kiwi and Mitsche [17] in the three directions, as described by the following theorems. First, we present a much simpler proof which also shows polylogarithmic upper bound for the diameter, but with a better (i.e. smaller) exponent.

**Theorem 1.** *Let  $2 < \beta < 3$ . The diameter of the giant component in the hyperbolic random graph  $\mathcal{G}(n, \alpha, C)$  is  $\mathcal{O}((\log n)^{\frac{2}{3-\beta}})$  with probability  $1 - \mathcal{O}(n^{-3/2})$ .*

The proof of Theorem 1 is presented in Section 3. It serves as an introduction to the more involved proof of a logarithmic upper bound for the diameter presented in Section 4. There we show with more advanced techniques that for small average degrees the following theorem holds.

**Theorem 2.** *Let  $2 < \beta < 3$ , and  $C$  be a large enough constant. Then, the diameter of the giant component in the hyperbolic random graph  $\mathcal{G}(n, \alpha, C)$  is  $\mathcal{O}(\log n)$  with probability  $1 - \mathcal{O}(n^{-3/2})$ .*

The logarithmic upper bound is best possible. In particular, we show that Theorem 2 is tight by presenting the following matching lower bound.

**Theorem 3.** *Let  $2 < \beta < 3$ . Then, the diameter of the giant component in the hyperbolic random graph  $\mathcal{G}(n, \alpha, C)$  is  $\Omega(\log n)$  with probability  $1 - n^{-\Omega(1)}$ .*

Due to space constraints, the proof of Theorem 3 can be found in the long version. We point out that although we prove all diameter bounds on the giant component, our proofs will make apparent that the giant component is in fact the component with the largest diameter in the graph.

**Used Techniques.** Our formal analysis of the diameter has to deal with a number of technical challenges. First, in contrast to proving a bound on the average distance, it is not possible to average over all path lengths. In fact, it is not even sufficient to exclude a certain kind of path with probability  $1 - \mathcal{O}(n^{-c})$ ; as this has to hold for all possible  $\Omega(n!)$  paths. This makes a union bound inapplicable. We solve this by introducing *upwards paths* (cf. Definition 12), which are in a sense “almost” shortest paths, and of which there are only two per node. We prove deterministically that their length asymptotically bounds the diameter. Then, we bound the length of a single upwards path by a multiplicative drift argument known from evolutionary computation [20]; and show that the length of conjunctions of upwards paths follows an Erlang distribution.

A second major challenge is the fact that a probabilistic analysis of shortest paths (and likewise, upwards paths) typically uncovers the probability space in a consecutive fashion. Revealing the positions of nodes on the path successively introduces strong stochastic dependencies that are difficult to handle with probabilistic tail bounds [11]. Instead of studying the stochastic dependence structure in detail, we use the geometry and model the hyperbolic random graph as a Poisson point process. This allows us to analyze different areas in the graph independently, which in turn supports our stochastic analysis of shortest paths.

## 2 Notation and Preliminaries

In this section, we briefly introduce hyperbolic random graphs. Although this paper is self-contained, we recommend to a reader who is unfamiliar with the notion of hyperbolic random graphs the more thorough investigations [16, 19].

Let  $\mathbb{H}_2$  be the hyperbolic plane. Following [19], we use the *native* representation; in which a point  $v \in \mathbb{H}_2$  is represented by polar coordinates  $(r_v, \varphi_v)$ ; and  $r_v$  is the hyperbolic distance of  $v$  to the origin.<sup>2</sup>

To construct a hyperbolic random graph  $G(n, \alpha, C)$ , consider now a circle  $D_n$  with radius  $R = 2 \ln n + C$  that is centered at the origin of  $\mathbb{H}_2$ . Inside  $D_n$ ,  $n$  points are distributed independently as follows. For each point  $v$ , draw  $\varphi_v$  uniformly at random from  $[0, 2\pi)$ , and draw  $r_v$  according to the probability density function

$$\rho(r) := \frac{\alpha \sinh(\alpha r)}{\cosh(\alpha R) - 1} \approx \alpha e^{\alpha(r-R)}.$$

Next, connect two points  $u, v$  if their hyperbolic distance is at most  $R$ , i.e. if

$$d(u, v) := \cosh^{-1}(\cosh(r_u) \cosh(r_v) - \sinh(r_u) \sinh(r_v) \cos(\Delta\varphi_{u,v})) \leq R. \quad (1)$$

By  $\Delta\varphi_{u,v}$  we describe the small relative angle between two nodes  $u, v$ , i.e.  $\Delta\varphi_{u,v} := \cos^{-1}(\cos(\varphi_u - \varphi_v)) \leq \pi$ .

This results in a graph whose degree distribution follows a power law with exponent  $\beta = 2\alpha + 1$ , if  $\alpha \geq \frac{1}{2}$ , and  $\beta = 2$  otherwise [16]. Since most real-world networks have been shown to have a power law exponent  $2 < \beta < 3$ , we assume throughout the paper that  $\frac{1}{2} < \alpha < 1$ . Gugelmann et al. [16] proved that the average degree in this model is then  $\delta = (1 + o(1)) \frac{2\alpha^2 e^{-C/2}}{\pi(\alpha-1/2)^2}$ .

We now present a handful of Lemmas useful for analyzing the hyperbolic random graph. Most of them are taken from [16]. We begin by an upper bound for the angular distance between two connected nodes. Consider two nodes with radial coordinates  $r, y$ . Denote by  $\theta_r(y)$  the maximal radial distance such that these two nodes are connected. By equation (1),

$$\theta_r(y) = \arccos \left( \frac{\cosh(y) \cosh(r) - \cosh(R)}{\sinh(y) \sinh(r)} \right). \quad (2)$$

This terse expression is closely approximated by the following Lemma.

<sup>2</sup> Note that this seemingly trivial fact does not hold for conventional models (e.g. Poincaré halfplane) for the hyperbolic plane.

**Lemma 4 ([16]).** *Let  $0 \leq r \leq R$  and  $y \geq R - r$ . Then,*

$$\theta_r(y) = \theta_y(r) = 2e^{\frac{R-r-y}{2}} (1 \pm \Theta(e^{R-r-y})).$$

For most computations on hyperbolic random graphs, we need expressions for the probability that a sampled point falls into a certain area. To this end, Gugelmann et al. [16] define the *probability measure* of a set  $S \subseteq D_n$  as

$$\mu(S) := \int_S f(y) \, dy,$$

where  $f(r)$  is the probability mass of a point  $p = (r, \varphi)$  given by  $f(r) := \frac{\rho(r)}{2\pi} = \frac{\alpha \sinh(\alpha r)}{2\pi(\cosh(\alpha R) - 1)}$ . We further define the *ball* with radius  $x$  around a point  $(r, \varphi)$  as

$$B_{r,\varphi}(x) := \{(r', \varphi') \mid d((r', \varphi'), (r, \varphi)) \leq x\}.$$

We write  $B_r(x)$  for  $B_{r,0}(x)$ . Note that  $D_n = B_0(R)$ . Using these definitions, we can formulate the following Lemma.

**Lemma 5 ([16]).** *For any  $0 \leq r \leq R$  we have*

$$\mu(B_0(r)) = e^{-\alpha(R-r)}(1 + o(1)) \tag{3}$$

$$\mu(B_r(R) \cap B_0(R)) = \frac{2\alpha e^{-r/2}}{\pi(\alpha - 1/2)} \cdot (1 \pm \mathcal{O}(e^{-(\alpha-1/2)r} + e^{-r})) \tag{4}$$

Since we often argue over sequences of nodes on a path, we say that a node  $v$  is *between* two nodes  $u, w$ , if  $\Delta\varphi_{u,v} + \Delta\varphi_{v,w} = \Delta\varphi_{u,w}$ . Recall that  $\Delta\varphi_{u,v} \leq \pi$  describes the small angle between  $u$  and  $v$ . E.g., if  $u = (r_1, 0), v = (r_2, \frac{\pi}{2}), w = (r_3, \pi)$ , then  $v$  lies between  $u$  and  $w$ . However,  $w$  does not lie between  $u$  and  $v$  as  $\Delta\varphi_{u,v} = \pi/2$  but  $\Delta\varphi_{u,w} + \Delta\varphi_{w,v} = \frac{3}{4}\pi$ .

Finally, we define the area  $B_I := B_0(R - \frac{\log R}{1-\alpha} - c)$  as the *inner band*, and  $B_O := D_n \setminus B_I$  as the *outer band*, where  $c \in \mathbb{R}$  is a large enough constant.

**The Poisson Point Process.** We often want to argue about the probability that an area  $S \subseteq D_n$  contains one or more nodes. To this end, we usually apply the simple formula

$$\Pr[\exists v \in S] = 1 - (1 - \mu(S))^n \geq 1 - \exp(-n \cdot \mu(S)). \tag{5}$$

Unfortunately, this formula significantly complicates once the positions of some nodes are already known. This introduces conditions on  $\Pr[\exists v \in S]$  which can be hard to grasp analytically. To circumvent this problem, we use a Poisson point process  $\mathcal{P}_n$  [23] which describes a different way of distributing nodes inside  $D_n$ . It is fully characterized by the following two properties:

- If two areas  $S, S'$  are disjoint, then the number of nodes that fall within  $S$  and  $S'$  are independent random variables.
- The expected number of points that fall within  $S$  is  $\int_S n\mu(S)$ .

One can show that these properties imply that the number of nodes inside  $S$  follows a Poisson distribution with mean  $n\mu(S)$ . In particular, we obtain that the number of nodes  $|\mathcal{P}_n|$  inside  $D_n$  is distributed as  $\text{Po}(n)$ , i.e.  $\mathbb{E}[|\mathcal{P}_n|] = n$ , and

$$\Pr(|\mathcal{P}_n| = n) = \frac{e^{-n}n^n}{n!} = \Theta(n^{-1/2}).$$

Let the random variable  $\mathcal{G}(\mathcal{P}_n, n, \alpha, C)$  denote the resulting graph when using the Poisson point process to distribute nodes inside  $D_n$ . Since it holds

$$\Pr[\mathcal{G}(\mathcal{P}_n, n, \alpha, C) = G \mid |\mathcal{P}_n| = n] = \Pr[\mathcal{G}(n, \alpha, C) = G],$$

we have that every property  $p$  with  $\Pr[p(\mathcal{G}(\mathcal{P}_n, n, \alpha, C))] \leq \mathcal{O}(n^{-c})$  holds for the hyperbolic random graphs with probability  $\Pr[p(\mathcal{G}(n, \alpha, C))] \leq \mathcal{O}(n^{\frac{1}{2}-c})$ .

We explicitly state whenever we use the Poisson point process  $\mathcal{G}(\mathcal{P}_n, n, \alpha, C)$  instead of the normal hyperbolic random graph  $\mathcal{G}(n, \alpha, C)$ . In particular, we can use a matching expression for equation (5):  $\Pr[\exists v \in S] = 1 - \exp(-n \cdot \mu(S))$ .

### 3 Polylogarithmic Upper Bound

As an introduction to the main proof, we first show a simple polylogarithmic upper bound on the diameter of the hyperbolic random graph. We start by investigating nodes in the inner band  $B_I$  and show that they are connected by a path of at most  $\mathcal{O}(\log \log n)$  nodes. We prove this by partitioning  $D_n$  into  $R$  layers of constant thickness 1. Then, a node in layer  $i$  has radial coordinate  $\in (R - i, R - i + 1]$ . We denote the layer  $i$  by  $L_i := B_0(R - i + 1) \setminus B_0(R - i)$ .

**Lemma 6.** *Let  $1 \leq i, j \leq R/2$ , and consider two nodes  $v \in L_i, w \in L_j$ . Then,*

$$\frac{2}{e} e^{\frac{i+j-R}{2}} (1 - \Theta(e^{i+j-R})) \leq \theta_{r_u}(r_v) \leq 2e^{\frac{i+j-R}{2}} (1 + \Theta(e^{i+j-R})),$$

Furthermore, we have  $\mu(L_j \cap B_R(v)) = \Theta(e^{-\alpha j + \frac{i+j-R}{2}})$ , and, if  $(i+j)/R < 1 - \varepsilon$  for some constant  $\varepsilon > 0$ , we have for large  $n$

$$\frac{1}{e} e^{-\alpha j + \frac{i+j-R}{2}} \leq \mu(L_j \cap B_R(v)) \leq 4e^{-\alpha j + \frac{i+j-R}{2}}.$$

*Proof.* The statements follow directly from Lemmas 4 and 5 and the fact that we have  $R - i < r_v \leq R - i + 1$  for a node  $v \in L_i$ . □

Using Lemma 6, we can now prove that a node  $v \in B_I$  has a path of length  $\mathcal{O}(\log \log n)$  that leads to  $B_0(R/2)$ . Recall that the inner band was defined as  $B_I := B_0(R - \frac{\log R}{1-\alpha} - c)$ , where  $c$  is a large enough constant.

**Lemma 7.** *Consider a node  $v$  in layer  $i$ . With probability  $1 - \mathcal{O}(n^{-3})$  it holds*

1. *if  $i \in [\frac{\log R}{1-\alpha} + c, \frac{2 \log R}{1-\alpha} + c]$ , then  $v$  has a neighbor in layer  $L_{i+1}$ , and*
2. *if  $i \in [\frac{2 \log R}{1-\alpha} + c, R/2]$ , then  $v$  has a neighbor in layer  $L_j$  for  $j = \frac{\alpha}{2\alpha-1}i$ .*

*Proof.* The probability that node  $v \in L_i$  does not contain a neighbor in  $L_{i+1}$  is

$$(1 - \Theta(e^{-\alpha(i+1) + i + \frac{1-R}{2}}))^n \leq \exp(-\Theta(1) \cdot e^{\log R + c(1-\alpha)}).$$

Since  $R = 2 \log n + C$  and  $c$  is a large enough constant, this proves part (1) of the claim. An analogous argument shows part (2). □

Lemma 7 shows that there exists a path of length  $\mathcal{O}(\log \log n)$  from each node  $v \in B_I$  to some node  $u \in B_0(R - \frac{2 \log R}{1-\alpha} - c)$ . Similarly, from  $u$  there exists a path of length  $\mathcal{O}(\log \log n)$  to  $B_0(R/2)$  with high probability. Since we know that the nodes in  $B_0(R/2)$  form a clique by the triangle inequality, we therefore obtain that all nodes in  $B_I$  form a connected component with diameter  $\mathcal{O}(\log \log n)$ .

**Corollary 8.** *Let  $\frac{1}{2} < \alpha < 1$ . With probability  $1 - \mathcal{O}(n^{-3})$ , all nodes  $u, v \in B_I$  in the hyperbolic random graph are connected by a path of length  $\mathcal{O}(\log \log n)$ .*

### 3.1 Outer Band

By Corollary 8, we obtain that the diameter of the graph induced by nodes in  $B_I$  is at most  $\mathcal{O}(\log \log n)$ . In this section, we show that each component in  $B_O$  has a polylogarithmic diameter. Then, one can easily conclude that the overall diameter of the giant component is polylogarithmic, since all nodes in  $B_0(R/2)$  belong to the giant component [3]. We begin by presenting one of the crucial Lemmas in this paper that will often be reused.

**Lemma 9.** *Let  $u, v, w \in V$  be nodes such that  $v$  lies between  $u$  and  $w$ , and let  $\{u, w\} \in E$ . If  $r_v \leq r_u$  and  $r_v \leq r_w$ , then  $v$  is connected to both  $u$  and  $w$ . If  $r_v \leq r_u$  but  $r_v \geq r_w$ , then  $v$  is at least connected to  $w$ .*

*Proof.* By [3, Lemma 5.28], we know that if two nodes  $(r_1, \varphi_1), (r_2, \varphi_2)$  are connected, then so are  $(r'_1, \varphi_1), (r'_2, \varphi_2)$  where  $r_1 \leq r'_1$  and  $r'_2 \leq r_2$ . Since the distance between nodes is monotone in the relative angle  $\Delta\varphi$ , this proves the first part of the claim. The second part can be proven by an analogous argument.  $\square$

For convenience, we say that an edge  $\{u, w\}$  passes under  $v$  if one of the requirements of Lemma 9 is fulfilled. Using this, we are ready to show Theorem 1. In this argument, we investigate the angular distance a path can at most traverse until it passes under a node in  $B_I$ . By Lemma 9, we then have with high probability a short path to the center  $B_0(R/2)$  of the graph.

*(Proof of Theorem 1).* Partition the hyperbolic disc into  $n$  disjoint sectors of equal angle  $\Theta(1/n)$ . The probability that  $k$  consecutive sectors contain no node in  $B_I$  is

$$\begin{aligned} (1 - \Theta(k/n) \cdot \mu(B_0(R - \frac{\log R}{1-\alpha} - c)))^n &\leq \exp(-\Theta(1) \cdot k \cdot e^{-\alpha \log R / (1-\alpha)}) \\ &= \exp(-\Theta(1) \cdot k \cdot (\log n)^{-\frac{\alpha}{1-\alpha}}). \end{aligned}$$

Hence, we know that with probability  $1 - \mathcal{O}(n^{-3})$ , there are no  $k := \Theta((\log n)^{\frac{1}{1-\alpha}})$  such consecutive sectors. By a Chernoff bound, the number of nodes in  $k$  such consecutive sectors is  $\Theta((\log n)^{\frac{1}{1-\alpha}})$  with probability  $1 - \mathcal{O}(n^{-3})$ . Applying a union bound, we get that with probability  $1 - \mathcal{O}(n^{-2})$ , every sequence of  $k$  consecutive sectors contains at least one node in  $B_I$  and at most  $\Theta(k)$  nodes in total. Consider now a node  $v \in B_O$  that belongs to the giant component. Then, there must exist a path from  $v$  to some node  $u \in B_I$ . By Lemma 9, this path can

visit at most  $k$  sectors—and therefore use at most  $\Theta(k)$  nodes—before reaching  $u$ . From  $u$ , there is a path of length  $\mathcal{O}(\log \log n)$  to the center  $B_0(R/2)$  of the hyperbolic disc by Corollary 8. Since this holds for all nodes, and the center forms a clique, the diameter is therefore  $\mathcal{O}((\log n)^{\frac{1}{1-\alpha}}) = \mathcal{O}((\log n)^{\frac{2}{3-2\beta}})$ .  $\square$

This bound slightly improves upon the results in [17] who show an upper bound of  $\mathcal{O}((\log n)^{\frac{8}{(1-\alpha)(2-\alpha)}})$ . As we will see in Theorem 3, however, the lower bound on the diameter is only  $\Omega(\log n)$ . We bridge this gap in the remaining part of the paper by analyzing the behavior in the outer band more carefully.

### 4 Logarithmic Upper Bound

In this section, we show that the diameter of the hyperbolic random graph is actually  $\mathcal{O}(\log n)$ , as long as the average degree is a small enough constant. We proceed by the following proof strategy. Consider a node  $v \in B_O$ . We investigate the *upwards path* from this node, which is intuitively constructed as follows: Each node on an upwards path has the smallest radial coordinate among all neighbors of the preceding node.

We first show that the diameter is asymptotically bounded by the longest upwards path in the graph. Afterwards, we prove that an upwards path is at most of length  $\mathcal{O}(\log n)$  with high probability by investigating a random walk whose hitting time dominates the length of the upwards path. A simple union bound over all nodes will conclude the proof.

We start by stating a bound that shows that if  $v$  is between two nodes  $u, w$  that are connected by an edge, then  $v$  is either connected to  $u$  or  $v$ , or one of these nodes has a radial coordinate at least 1 smaller than  $v$ . Due to space constraints, this and all following proofs can be found in the long version.

**Lemma 10.** *Let  $u, v, w$  be nodes in the outer band such that  $v$  lies between  $u$  and  $w$ . Furthermore, let  $\{u, w\} \in E$ , but  $\{u, v\}, \{v, w\} \notin E$ . Then, for large  $n$ , at least one of the following holds:  $r_u \leq r_v - 1$  or  $r_w \leq r_v - 1$ .*

Similarly to Lemma 9, we say that an edge  $\{u, w\}$  *passes over*  $v$ , if the requirements of Lemma 10 are fulfilled. Before we introduce the formal definition of an upwards path, we define the notion of a straight path.

**Definition 11.** *Let  $\pi = [v_1, \dots, v_k]$  be a path in the hyperbolic random graph where  $\forall i, v_i \in B_O$ . We say that  $\pi$  is straight, if  $\forall i \in \{2, \dots, k - 1\}$  the node  $v_i$  lies between  $v_{i-1}$  and  $v_{i+1}$ .*

The definition of a straight path captures the intuitive notion that the path does not “jump back and forth”. Next, we define an upwards path, which is a special case of a straight path.

**Definition 12.** *Let  $v \in B_O$  be a node in the hyperbolic random graph and define  $\tilde{\varphi}_u := (\pi + \varphi_u - \varphi_v) \bmod 2\pi$ . Furthermore, we define the neighbors to the right of  $u$  as*

$$\tilde{\Gamma}(u) := \Gamma(u) \cap \{w \in B_O \mid \tilde{\varphi}_w \geq \tilde{\varphi}_u\}. \tag{6}$$



Then we say that  $\pi_v = [v = v_0, v_1, \dots, v_k]$  is an upwards path from left to right from  $v$  if  $\forall i \in \{0, \dots, k - 1\}: v_{i+1} = \operatorname{argmax}_{u \in \tilde{r}(v_i)} \{r_u\}$ , and there is no longer upwards path  $\pi'_v \supseteq \pi_v$ .

Analogously, we define an upwards path from right to left by replacing  $\tilde{\varphi}_w \geq \tilde{\varphi}_u$  by  $\tilde{\varphi}_w \leq \tilde{\varphi}_u$  in equation (6).

Observe that there are two upwards paths from each node: One from right to left, and the other from left to right. An upwards path also only uses nodes in  $B_O$ . The exclusion of  $B_I$  can only increase the diameter of a component in the outer band.

The next Lemma shows that the length of the longest upwards path asymptotically bounds the length of all straight shortest paths in the outer band.

**Lemma 13.** *Assume that for all nodes  $v \in B_O$ , the upwards paths in both directions are of length  $|\pi_v| \leq f(n)$ . Let  $\pi = [u_1, u_2, \dots, u_k]$  be a straight shortest path. Then,  $|\pi| \leq 2 \cdot f(n) + 1 = \mathcal{O}(f(n))$ .*

We proceed by arguing that all upwards paths in the outer band are of length  $\mathcal{O}(\log n)$  at most. In fact, we show a stronger statement by deriving an exponential tail bound on the length of an upwards path. To this end, we model an upwards path as a random walk. Consider for some node  $v$  all neighbors to the right of  $v$ . Among those, the neighbor in the largest layer (or equivalently, the smallest radial coordinate) is the neighbor on which any upwards path from left to right will continue. We formulate a probability that the upwards path jumps into a certain layer and analyze the probability that after  $T$  steps, the random walk modeled by this process is absorbed, i.e. we reach a node that has no further neighbors in this direction.

Let the random variables  $[u = V_0, V_1, \dots]$  describe the upwards path from  $u$ , and let  $X_i := \ell$  if  $V_i$  is in layer  $L_\ell$ , and  $X_i := 0$  if the upwards path consists of  $< i$  nodes. Without loss of generality, the upwards path is from left to right. Then, we have

$$\Pr[X_{i+1} = m \mid X_1, \dots, X_i] \leq \frac{1}{2} \Pr[\exists w \in L_m \text{ such that } d(V_i, w) \leq R] \cdot \Pr[\nexists m' > m \text{ such that } \exists w' \in L_{m'} \text{ with } d(V_i, w') \leq R] \quad (7)$$

Note that in  $\Pr[X_{i+1} = m \mid X_1, \dots, X_i]$  we implicitly condition on the fact that some preceding node  $V_{i'}$  with  $i' < i$  on the upwards path was *not* connected to  $V_{i+1}$ ; which technically excludes some subset of  $B_R(V_i)$ . We fix this issue by considering the Poisson point process and exposing the randomness as follows. First, we assume that there are no preconditions on  $\Pr[X_{i+1} = m]$  (i.e. the upwards path begins in  $V_i$ ). Then, the above formula is exact. We now expose all neighbors of  $V_i$ , and obtain  $w \in L_m$  as the neighbor in the uppermost layer. Now we expose the conditions. This is possible, since in the Poisson point model, each area disjoint from other areas can be treated independently. The exposing of the conditions can only delete nodes. In this process,  $w$  might be deleted, lowering the probability of  $\Pr[X_{i+1} = m]$ .

Therefore, our stated formula is indeed an upper bound.

**Lemma 14.** *Let the random variables  $[u = V_0, V_1, \dots]$  describe the upwards path from  $u$ , and let  $\forall i, X_i := \ell$  if  $V_i$  is in layer  $L_\ell$ , and  $X_i := 0$  if the upwards path consists of  $< i$  nodes. Then, if  $C$  is large enough, we have  $\mathbb{E}[X_{i+1}] \leq 0.99 \cdot X_i$ .*

Lemma 14 shows that  $X_i$  has a multiplicative drift towards 0. Let  $T := \min\{i \mid X_i = 0\}$  be the random variable describing the length of an upwards path. We now bound  $T$  by a multiplicative drift theorem as presented by Lehre and Witt [20, Theorem 7] and originally developed by Doerr and Goldberg [9, Theorem 1] for the analysis of evolutionary algorithms. For the sake of completeness, we restate their result.

**Theorem 15 (from [9, 20]).** *Let  $(X_t)_{t \geq 0}$  be a stochastic process over some state space  $\{0\} \cup [x_{\min}, x_{\max}]$ , where  $x_{\min} > 0$ . Suppose that there exists some  $0 < \delta < 1$  such that  $\mathbb{E}[X_t - X_{t+1} \mid X_0, \dots, X_t] \geq \delta X_t$ . Then for the first hitting time  $T := \min\{t \mid X_t = 0\}$  it holds*

$$\Pr[T \geq \frac{1}{\delta}(\ln(X_0/x_{\min}) + r) \mid X_0] \leq e^{-r} \text{ for all } r > 0.$$

In our case,  $X_0 \leq \frac{\log R}{1-\alpha} + c$  and  $x_{\min} = 1$ . Using Lemma 14 this shows that

$$\Pr[T \geq 101 \cdot (\log \log \log n + r)] \leq e^{-r}. \tag{8}$$

Hence, with probability  $1 - \mathcal{O}(n^{-3})$  the random walk process described by  $X_i$  terminates after  $\mathcal{O}(\log n)$  steps. By a union bound we have that all upwards paths in  $G$  are of length  $\mathcal{O}(\log n)$  with probability  $1 - \mathcal{O}(n^{-2})$ . To show that all shortest paths are of length  $\mathcal{O}(\log n)$ , however, we need a slightly stronger statement, namely that the sum of  $\mathcal{O}(\log \log n)$  upwards paths is at most of length  $\mathcal{O}(\log n)$ .

**Lemma 16.** *Let  $(T_i)_{i=1 \dots X}$  be distributed according to equation (8), where  $X = c \log \log n$ . Then, with probability  $1 - \mathcal{O}(n^{-3})$ ,  $\sum_{i=1}^X T_i \leq \mathcal{O}(\log n)$ .*

To conclude our result on the diameter, it is left to investigate shortest paths in  $B_O$  that are not straight. The general proof strategy for those paths is as follows. First, we show that such a path has edges that switch directions. Such an edge must pass over all preceding (or all following) nodes, as will become apparent in the next Lemma.

**Lemma 17.** *Consider a shortest path  $\pi = [u_1, \dots, u_k]$  that is not straight. In particular,  $\pi$  then has one or more sequence of nodes  $u_{i-1}, u_i, u_{i+1}$  such that  $u_i$  is not between  $u_{i-1}$  and  $u_{i+1}$ . Then, for all such positions  $i$  it holds that either*

1.  $\forall j < i: u_j$  is between  $u_i$  and  $u_{i+1}$ , or
2.  $\forall j > i: u_j$  is between  $u_i$  and  $u_{i-1}$ .

By Lemma 10, we know that can only be  $\mathcal{O}(\log \log n)$  changes of directions; and Lemma 16 lets us conclude that the total path length is still  $\mathcal{O}(\log n)$ . This can be used to show that *all* shortest paths are of length  $\mathcal{O}(\log n)$ .

**Lemma 18.** *Let  $\pi = [u_1, u_2, \dots, u_k]$  be a shortest path where  $\forall i, u_i \in B_O$ . Then, with probability  $1 - \mathcal{O}(n^{-3/2})$ ,  $|\pi| = \mathcal{O}(\log n)$ .*

Lemma 18 in conjunction with Lemma 7 then proves Theorem 2, i.e. that the diameter of the hyperbolic random graph is  $\mathcal{O}(\log n)$  if the average degree is a small enough constant.

## 5 Conclusion

We derive a new polylogarithmic upper bound on the diameter of hyperbolic random graphs; and show that it is  $\mathcal{O}(\log n)$  if the average degree is small. We further prove a matching lower bound. This immediately yields lower bounds for any broadcasting protocol that has to reach all nodes. Processes such as bootstrap percolation or rumor spreading therefore must run at least  $\Omega(\log n)$  steps until they inform all nodes in the giant component.

Our work focuses on power law exponents  $2 < \beta < 3$ , but we believe that our proof can be extended to bound the diameter for  $\beta > 3$  by  $\Theta(\log n)$ . For other scale-free models it was also interesting to study the phase transition at  $\beta = 2$  and  $\beta = 3$ . Another natural open question is the average distance (also known as average diameter) between two random nodes. We conjecture that the average distance is  $\Theta(\log \log n)$ , but leave this open for future work.

**Acknowledgements.** We thank Konstantinos Panagiotou for many useful discussions, and suggesting a layer-based proof as in Lemma 6.

## References

1. Barabási, A.-L., Albert, R.: Emergence of scaling in random networks. *Science* **286**, 509–512 (1999)
2. Bode, M., Fountoulakis, N., Müller, T.: On the giant component of random hyperbolic graphs. In: 7th European Conference on Combinatorics, Graph Theory and Applications (EuroComb), pp. 425–429 (2013)
3. Bode, M., Fountoulakis, N., Müller, T.: The probability that the hyperbolic random graph is connected (2014). [www.math.uu.nl/Muell001/Papers/BFM.pdf](http://www.math.uu.nl/Muell001/Papers/BFM.pdf)
4. Boguná, M., Papadopoulos, F., Krioukov, D.: Sustaining the internet with hyperbolic mapping. *Nature Communications* **1**, 62 (2010)
5. Bollobás, B.: *Random graphs*. Springer (1998)
6. Bollobás, B., Chung, F.R.K.: The diameter of a cycle plus a random matching. *SIAM Journal of Discrete Mathematics* **1**, 328–333 (1988)
7. Candellero, E., Fountoulakis, N.: Bootstrap percolation and the geometry of complex networks (2014). arxiv1412.1301
8. Chung, F., Lu, L.: The average distances in random graphs with given expected degrees. *Proceedings of the National Academy of Sciences* **99**, 15879–15882 (2002)
9. Doerr, B., Goldberg, L.A.: Drift analysis with tail bounds. In: Schaefer, R., Cotta, C., Kołodziej, J., Rudolph, G. (eds.) PPSN XI. LNCS, vol. 6238, pp. 174–183. Springer, Heidelberg (2010)

10. Dommers, S., van der Hofstad, R., Hooghiemstra, G.: Diameters in preferential attachment models. *Journal of Statistical Physics* **139**, 72–107 (2010)
11. Dubhashi, D.P., Panconesi, A.: Concentration of measure for the analysis of randomized algorithms. Cambridge University Press (2009)
12. Evans, M., Hastings, N., Peacock, B.: *Statistical Distributions*, chapter 12, 3rd edn., pp. 71–73. Wiley-Interscience (2000)
13. Fountoulakis, N., Panagiotou, K., Sauerwald, T.: Ultra-fast rumor spreading in models of real-world networks (2015). Unpublished draft
14. Friedrich, T., Krohmer, A.: Cliques in hyperbolic random graphs. In: 34th IEEE Conference on Computer Communications (INFOCOM) (2015). [https://hpi.de/fileadmin/user\\_upload/fachgebiete/friedrich/publications/2015/cliques2015.pdf](https://hpi.de/fileadmin/user_upload/fachgebiete/friedrich/publications/2015/cliques2015.pdf)
15. Friedrich, T., Sauerwald, T., Stauffer, A.: Diameter and broadcast time of random geometric graphs in arbitrary dimensions. *Algorithmica* **67**, 65–88 (2013)
16. Gugelmann, L., Panagiotou, K., Peter, U.: Random hyperbolic graphs: degree sequence and clustering. In: Czumaj, A., Mehlhorn, K., Pitts, A., Wattenhofer, R. (eds.) ICALP 2012, Part II. LNCS, vol. 7392, pp. 573–585. Springer, Heidelberg (2012)
17. Kiwi, M., Mitsche, D.: A bound for the diameter of random hyperbolic graphs. In: 12th Workshop on Analytic Algorithmics and Combinatorics (ANALCO), pp. 26–39 (2015)
18. Kleinberg, J.: Navigation in a small world. *Nature* **406**, 845 (2000)
19. Krioukov, D., Papadopoulos, F., Kitsak, M., Vahdat, A., Boguñá, M.: Hyperbolic geometry of complex networks. *Physical Review E* **82**, 036106 (2010)
20. Lehre, P.K., Witt, C.: General drift analysis with tail bounds (2013). arxiv1307.2559
21. Martel, C.U., Nguyen, V.: Analyzing Kleinberg’s (and other) small-world models. In: 23rd Annual ACM Symposium on Principles of Distributed Computing (PODC), pp. 179–188 (2004)
22. Papadopoulos, F., Psomas, C., Krioukov, D.: Network mapping by replaying hyperbolic growth. *IEEE/ACM Transactions on Networking*, 198–211 (2014)
23. Penrose, M.: *Random Geometric Graphs*. Oxford scholarship online. Oxford University Press (2003)
24. Riordan, O., Wormald, N.: The diameter of sparse random graphs. *Combinatorics, Probability and Computing* **19**, 835–926 (2010)
25. von Looz, M., Staudt, C.L., Meyerhenke, H., Prutkin, R.: Fast generation of dynamic complex networks with underlying hyperbolic geometry (2015). arxiv1501.03545
26. Watts, D.J., Strogatz, S.H.: Collective dynamics of ‘small-world’ networks. *Nature* **393**, 440–442 (1998)