A Bound for the Diameter of Random Hyperbolic Graphs

Marcos Kiwi, Dieter Mitsche

A Bound for the Diameter of Random Hyperbolic Graphs*

Marcos Kiwi† Dieter Mitsche‡

Abstract

Random hyperbolic graphs were recently introduced by Krioukov et. al. [KPK+10] as a model for large networks. Gugelmann, Panagiotou, and Peter [GPP12] then initiated the rigorous study of random hyperbolic graphs using the following model: for $\alpha > \frac{1}{2}$, $C \in \mathbb{R}$, $n \in \mathbb{N}$, set $R = 2\ln n + C$ and build the graph $G = (V, E)$ with $|V| = n$ as follows: For each $v \in V$, generate i.i.d. polar coordinates (r_v, θ_v) using the joint density function $f(r, \theta)$, with θ_v chosen uniformly from $[0, 2\pi)$ and r_v with density $f(r) = \frac{\alpha \sinh(\alpha r)}{\cosh(\alpha R)-1}$ for $0 \leq r < R$. Then, join two vertices by an edge, if their hyperbolic distance is at most R. We prove that in the range $\frac{1}{2} < \alpha < 1$ a.a.s. for any two vertices of the same component, their graph distance is $O(\log^{C_0+1+o(1)} n)$, where $C_0 = 2/(\frac{1}{2} - \frac{3}{4}\alpha + \frac{\alpha^2}{4})$, thus answering a question raised in [GPP12] concerning the diameter of such random graphs. As a corollary from our proof we obtain that the second largest component has size $O(\log^{2C_0+1+o(1)} n)$, thus answering a question of Bode, Fountoulakis and Müller [BFM13]. We also show that a.a.s. there exist isolated components forming a path of length $\Omega(\log n)$, thus yielding a lower bound on the size of the second largest component.

Keywords: Random hyperbolic graphs; Complex networks.

1 Introduction

Building mathematical models to capture essential properties of large networks has become an important objective in order to better understand them. An interesting new proposal in this direction is the model of random hyperbolic graphs recently introduced by Krioukov et. al. [KPK+10] (see also [PKBnV10]). A good model should on the one hand replicate the characteristic properties that are observed in real world networks (e.g., power law degree distributions, high clustering and small diameter), but on the other hand it should also be susceptible to mathematical analysis. There are models that partly succeed in the first task but are hard to analyze rigorously. Other models, like the classical Erdös-Renyi $G(n, p)$ model, can be studied mathematically, but fail to capture certain aspects observed in real-world networks. In contrast, the authors of [PKBnV10] argued empirically and via some non-rigorous methods that random hyperbolic graphs have many of the desired properties. Actually, Boguñá, Papadopoulos and Krioukov [BnPK10] computed explicitly a maximum likelihood fit of the Internet graph, convincingly illustrating that this model is adequate for reproducing the structure of real networks with high accuracy. Gugelmann, Panagiotou, and Peter [GPP12] initiated the rigorous study of random hyperbolic graphs. They compute exact asymptotic expressions for the maximum degree, the degree distribution (confirming rigorously that the degree sequence follows a power-law distribution with controllable exponent), and also estimated the expectation of the clustering coefficient.

In words, the random hyperbolic graph model is a simple variant of the uniform distribution of n vertices within a disc of radius R of the hyperbolic plane, where two vertices are connected if their hyperbolic distance

*A preliminary version of this paper will appear in the Proceedings of the Twelfth Workshop on Analytic Algorithmics and Combinatorics (ANALCO, 2015).

†Depto. Ing. Matemática & Ctr. Modelamiento Matemático (CNRS UMI 2807), U. Chile. Beauchef 851, Santiago, Chile, Email: mkiwi@dim.uchile.cl. Gratefully acknowledges the support of Millennium Nucleus Information and Coordination in Networks ICM/FIC P10-024F and CONICYT via Basal in Applied Mathematics.

‡Université de Nice Sophia-Antipolis, Laboratoire J-A Dieudonné, Parc Valrose, 06108 Nice cedex 02, Email: dmitsche@unice.fr. The main part of this work was performed during a stay at Depto. Ing. Matemática & Ctr. Modelamiento Matemático (CNRS UMI 2807), U. Chile, and the author would like to thank them for their hospitality.

is at most R. Formally, the random hyperbolic graph model $G_{\alpha,C}(n)$ is defined in [GPP12] as described next: for $\alpha > \frac{1}{2}$, $C \in \mathbb{R}$, $n \in \mathbb{N}$, set $R = 2\ln n + C$, and build $G = (V, E)$ with vertex set $V = [n]$ as follows:

- For each $v \in V$, polar coordinates (r_v, θ_v) are generated identically and independently distributed with joint density function $f(r, \theta)$, with θ_v chosen uniformly at random in the interval $[0, 2\pi)$ and r_v with density:

$$f(r) = \begin{cases} \dfrac{\alpha \sinh(\alpha r)}{C(\alpha, R)}, & \text{if } 0 \le r < R, \\ 0, & \text{otherwise,} \end{cases}$$

  where $C(\alpha, R) = \cosh(\alpha R) - 1$ is a normalization constant.

- For $u, v \in V$, $u \neq v$, there is an edge with endpoints u and v provided $d(r_u, r_v, \theta_u - \theta_v) \le R$, where $d = d(r, r', \theta - \theta')$ denotes the hyperbolic distance between two vertices whose native representation polar coordinates are (r, θ) and (r', θ'), obtained by solving

$$\cosh(d) = \cosh(r)\cosh(r') - \sinh(r)\sinh(r')\cos(\theta - \theta'). \tag{1}$$

The restriction $\alpha > 1/2$ and the role of R, informally speaking, guarantee that the resulting graph has a bounded average degree (depending on α and C only). If $\alpha < 1/2$, then the degree sequence is so heavy tailed that this is impossible.

Research in random hyperbolic graphs is in a sense in its infancy. Besides the results mentioned above, very little else is known. Notable exceptions are the emergence and evolution of giant components [BFM13], connectedness [BFM], results on the global clustering coefficient of the so called binomial model of random hyperbolic graphs [CF13], and on the evolution of graphs on more general spaces with negative curvature [Fou12].

Notation. As typical in random graph theory, we shall consider only asymptotic properties as $n \to \infty$. We say that an event in a probability space holds asymptotically almost surely (a.a.s.) if its probability tends to one as n goes to infinity. We follow the standard conventions concerning asymptotic notation. In particular, we write $f(n) = o(g(n))$, if $\lim_{n\to\infty} |f(n)|/|g(n)| = 0$. We will nevertheless also use $1 - o(\cdot)$ when dealing with probabilities. We also say that an event holds *with extremely high probability*, w.e.p., if it occurs with probability at least $1 - e^{-\omega(\log n)}$. Throughout this paper, $\log n$ always denotes the natural logarithm of n.

The constants α, C used in the model and the constants C_0, δ defined below have only one special meaning, other constants such as $C', C'', c, c_1, c_2, c_3$ change from line to line. Since we are interested in asymptotic results only, we ignore rounding issues throughout the paper.

1.1 Results

The main problem we address in this work is the natural question, explicitly stated in [GPP12, page 6], that asks to determine the expected diameter of the giant component of a random hyperbolic graph G chosen according to $G_{\alpha,C}(n)$ for $\frac{1}{2} < \alpha < 1$. We look at this range, since for $\alpha < \frac{1}{2}$ a.a.s. a very small central configuration yielding a diameter of at most 3 exists (see [BFM]): consider a ball of sufficiently small radius around the origin and partition it into 3 sectors. It can be shown that in each of the sectors a.a.s. there will be at least one vertex, and also that every other vertex is connected to at least one of the three vertices. For $\alpha = \frac{1}{2}$ the probability of this configuration to exist depends on C (see [BFM]). For $\alpha > 1$, there exists no giant component (see [BFM13]); the case $\alpha = 1$ is a matter requiring further study.

We show (see Theorem 12) that for $\frac{1}{2} < \alpha < 1$, a.a.s., for any two vertices of the same component, their graph distance is $O(\log^{C_0+1+o(1)} n)$, where $C_0 = 2/(\frac{1}{2} - \frac{3}{4}\alpha + \frac{\alpha^2}{4})$. To establish our main result we rely on the known, and easily established fact, that for the range of α we are concerned with, a graph $G_{\alpha,C}(n)$ has a "center" clique whose size is w.e.p. $\Theta(n^{1-\alpha})$. Then, we show that, depending on how "far away" from the center clique a vertex Q is, there is either a very high or at least non-negligible probability that the vertex connects to the center clique through a path of polylogarithmic (in n) length, or otherwise all paths starting from Q have at most polylogarithmic length. It immediately follows that two vertices in the same connected component, a.a.s., either connect to the center clique through paths of polylogarithmic length, or belong to paths of size at most polylogarithmic. Either way, a bound on the diameter of $G_{\alpha,C}(n)$ follows. Rigorously developing the preceding argument requires overcoming significant obstacles, not only technical but also in terms of developing the insight to appropriately define the relevant typical events which are also amenable to a rigorous study of their probabilities of occurrence. Our main result's proof argument also yields (see Corollary 13) that the size of the second largest component is $O(\log^{2C_0+1+o(1)} n)$, thus answering a question of Bode, Fountoulakis and Müller [BFM13]. As a complementary result (see Theorem 16), we establish that a.a.s. there exists a component forming an induced path of length $\Theta(\log n)$. This last result pinpoints a region of hyperbolic space (a constant width band around the origin of sufficiently large radius) where it is likely to find a path of length $\Theta(\log n)$.

Another contribution of our work is that it proposes at least two refinements and variants of the *breadth exploration process* introduced by Bode, Fountoulakis and Müller [BFM13]. Specifically, we strengthen the method by identifying more involved strategies for exploring hyperbolic space, not solely dependent on the angular coordinates of its points, and not necessarily contiguous regions of space. We hope these refinements will be useful in tackling other problems concerning the newly minted (and captivating) hyperbolic random graph model.

1.2 Organization

This work is organized as follows. In Section 2 we introduce the geometric framework and give background results. Section 3 deals with the upper bound on the diameter as well as the upper bound on the size of the second largest component. Section 4 is dedicated to the lower bound.

2 Conventions, background results and preliminaries

In this section we first fix some notational conventions. We then recall, as well as establish, a few results we rely on throughout the following sections. All claims in some sense reflect the nature of hyperbolic space, either by providing useful approximations for the angles formed by two adjacent sides of a triangle whose vertices are at given distances, or by establishing good approximations for the mass of regions of hyperbolic space obtained by some set algebraic manipulation of balls. We conclude the section by discussing the de-Poissonization technique on which this work heavily relies.

Henceforth, for a point P in hyperbolic plane, we let (r_P, θ_P) denote its polar coordinates ($0 \leq r_P < R$ and $0 \leq \theta_P < 2\pi$). The point with polar coordinates $(0,0)$ is called the origin and is denoted by O.

By (1), the hyperbolic triangle formed by the geodesics between points A, B, and C, with opposing side segments of length a, b, and c respectively, is such that the angle formed at C is (see Figure 1):

$$\theta_c(a,b) = \arccos\Big(\frac{\cosh(a)\cosh(b) - \cosh(c)}{\sinh(a)\sinh(b)}\Big).$$

Clearly, $\theta_c(a,b) = \theta_c(b,a)$.

In fact, although some of the proofs hold for a wider range of α, we will always assume $\frac{1}{2} < \alpha < 1$, since as already discussed, we are interested in this regime. In order to avoid unnecessary repetitions, we henceforth omit this restriction from the statement of this article's results.

First, we recall some useful estimates. A very handy approximation for $\theta_c(\cdot,\cdot)$ is given by the following result.

Lemma 1 ([GPP12, Lemma 3.1])**.** *If* $0 \leq \min\{a,b\} \leq c \leq a+b$*, then*

$$\theta_c(a,b) = 2e^{\frac{1}{2}(c-a-b)}\big(1 + \Theta(e^{c-a-b})\big).$$

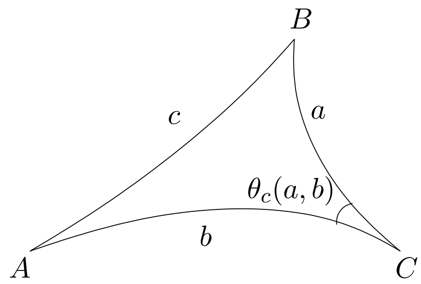

Figure 1: Hyperbolic triangle.

Remark 1. *We will use the previous lemma also in this form: let A and B be two points such that $r_A, r_B > R/2$, $0 \le \min\{r_A, r_B\} \le d := d(A,B) \le R$. Then, taking $c = d$, $a = r_A$, $b = r_B$ in Lemma 1 and noting that $r_A + r_B \ge R \ge d$, we get*

$$|\theta_A - \theta_B| = \theta_d(r_A, r_B) = 2e^{\frac{1}{2}(d - r_A - r_B)}\big(1 + \Theta(e^{d - r_A - r_B})\big).$$

Note also that for fixed $r_A, r_B > R/2$, $|\theta_A - \theta_B|$ is increasing as a function of d (for d satisfying the constraints). Below, when aiming for an upper bound, we always use $d = R$.

Throughout, we will need estimates for measures of regions of the hyperbolic plane, and more specifically, for regions obtained by performing some set algebra involving a couple of balls. Hereafter, for a point P of the hyperbolic plane, ρ_P will be used to denote the radius of a ball centered at P, and $B_P(\rho_P)$ to denote the closed ball of radius ρ_P centered at P. Also, we denote by $\mu(S)$ the probability measure of a set S, i.e.,

$$\mu(S) = \int_S f(r,\theta) dr d\theta.$$

We collect now a few results for such measures.

Lemma 2 ([GPP12, Lemma 3.2])**.** *For any $0 \le \rho_O \le R$, $\mu(B_O(\rho_O)) = e^{-\alpha(R-\rho_O)}(1 + o(1))$.*

The following result gives an approximation for the mass of the intersection of two balls of "reasonable size", one of which being centered at the origin O and the other one containing the origin.

Lemma 3. *Let $C_\alpha = 2\alpha/(\pi(\alpha - \frac{1}{2}))$. For $r_A \le \rho_A$ and $\rho_O + r_A \ge \rho_A$,*

$$\mu(B_A(\rho_A) \cap B_O(\rho_O)) = C_\alpha e^{-\alpha(R-\rho_O) - \frac{1}{2}(\rho_O - \rho_A + r_A)} + O(e^{-\alpha(R - \rho_A + r_A)}).$$

Proof. We want to bound (see Figure 2):

$$\mu(B_A(\rho_A) \cap B_O(\rho_O)) = \mu(B_O(\rho_A - r_A)) + 2\int_{\rho_A - r_A}^{\rho_O} \int_0^{\theta_{\rho_A}(r, r_A)} \frac{f(r)}{2\pi} d\theta dr.$$

Relying on the approximation of $\theta_{\rho_A}(r, r_A)$ given by Lemma 1,

$$\begin{aligned}&\mu(B_A(\rho_A) \cap B_O(\rho_O)) \\ &\quad = \mu(B_O(\rho_A - r_A)) + \frac{\alpha e^{\frac{1}{2}(\rho_A - r_A)}}{\pi C(\alpha, R)} \times \int_{\rho_A - r_A}^{\rho_O} (e^{(\alpha - \frac{1}{2})r} - e^{-(\alpha + \frac{1}{2})r})(1 + \Theta(e^{\rho_A - r - r_A}))dr.\end{aligned}$$

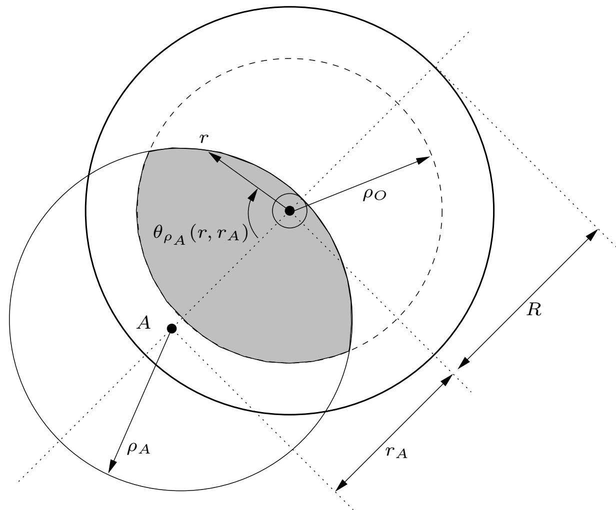

Figure 2: the shaded area corresponds to $B_A(\rho_A) \cap B_O(\rho_0)$.

We will first solve the integral without the error term $\Theta(e^{\rho_A - r - r_A})$. Since $C_\alpha = 2\alpha/(\pi(\alpha - \frac{1}{2}))$, and recalling that $C(\alpha, R) = \cosh(\alpha R) - 1 = \frac{1}{2}e^{\alpha R} + O(e^{-\alpha R})$, for this part we obtain

$$\begin{aligned}
&\frac{\alpha e^{\frac{1}{2}(\rho_A - r_A)}}{\pi C(\alpha, R)} \Big(\frac{e^{(\alpha - \frac{1}{2})\rho_O} - e^{(\alpha - \frac{1}{2})(\rho_A - r_A)}}{\alpha - \frac{1}{2}} + O(1) \Big) \\
&\qquad = C_\alpha \Big(e^{-\alpha(R - \rho_O) - \frac{1}{2}(\rho_O - \rho_A + r_A)} - e^{-\alpha(R - \rho_A + r_A)} + \Theta(e^{\frac{1}{2}(\rho_A - r_A) - \alpha R}) \Big) \\
&\qquad = C_\alpha \Big(e^{-\alpha(R - \rho_O) - \frac{1}{2}(\rho_O - \rho_A + r_A)} + O(e^{-\alpha(R - \rho_A + r_A)}) \Big),
\end{aligned}$$

where the last identity is because $e^{\frac{1}{2}(\rho_A - r_A) - \alpha R} = O(e^{-\alpha(R - \rho_A + r_A)})$ for $\alpha > \frac{1}{2}$. Now, for the error term, since $\alpha < \frac{3}{2}$, we obtain

$$\frac{\alpha e^{\frac{3}{2}(\rho_A - r_A)}}{\pi C(\alpha, R)} \Big(\frac{\Theta(e^{(\alpha - \frac{3}{2})\rho_O} - e^{(\alpha - \frac{3}{2})(\rho_A - r_A)})}{\alpha - \frac{3}{2}} + O(1) \Big) = \frac{\alpha e^{\frac{3}{2}(\rho_A - r_A)}}{\pi C(\alpha, R)} O(1) = O(e^{-\alpha(R - \rho_A + r_A)}).$$

Applying Lemma 2 to $\mu(B_O(\rho_A - r_A))$, the desired conclusion follows. ☐

From now on, denote $R_0 := R/2$ and $R_i := Re^{-\alpha^i/2}$ for $i \geq 1$. Observe that $R_{i-1} \leq R_i$ for all i. Next, we show that for $R/2 < r_A \leq R - O(1)$, a constant fraction of the mass of the intersection of the balls $B_A(R)$ and $B_O(R_i)$ is explained by the mass of the intersection of $B_A(R)$ with just the band $B_O(R_i) \setminus B_O(R_{i-1})$.

Lemma 4. *Let $\xi > 0$ be some constant and $i \geq 1$ be such that $R_i < R - \xi$. If $R_i < r_A \leq R_{i+1}$, then*

$$\mu(B_A(R) \cap \big(B_O(R_i) \setminus B_O(R_{i-1})\big)) = \Omega(\mu(B_A(R) \cap B_O(R_i))),$$

where the constant $C' = C'(\xi, \alpha)$ hidden inside the asymptotic notation can be assumed nondecreasing as a function of ξ.

Proof. Since $R_{i-1} \leq R_i$, we have that $B_O(R_{i-1}) \subseteq B_O(R_i)$, so the left hand side of the stated identity can be re-written as $\mu(B_A(R) \cap B_O(R_i)) - \mu(B_A(R) \cap B_O(R_{i-1}))$. Now, observe that by Lemma 3 applied with $\rho_A = R$, $\rho_O = R_i$ (so $R_i + r_A > 2R_i \geq R$),

$$\mu(B_A(R) \cap B_O(R_i)) = (1 + o(1)) C_\alpha e^{-\alpha(R - R_i) - \frac{1}{2}(R_i - R + r_A)}, \tag{2}$$

where $C_\alpha = 2\alpha/(\pi(\alpha-\frac{1}{2})$. By the same argument, this time applied with $\rho_A = R$, $\rho_O = R_{i-1}$, we also have

$$\mu(B_A(R)\cap B_O(R_{i-1})) = (1+o(1))C_\alpha e^{-\alpha(R-R_{i-1})-\frac{1}{2}(R_{i-1}-R+r_A)}. \tag{3}$$

It suffices to show that the ratio $\rho(R,i)$ between the right hand sides of the expressions in (2) and (3) is at least a constant. We claim that the assertion regarding $\rho(R,i)$ holds. Indeed, note that

$$\rho(R,i) = (1+o(1))e^{(\alpha-\frac{1}{2})(R_i-R_{i-1})} = (1+o(1))e^{(\alpha-\frac{1}{2})R(e^{-\alpha^i/2}-e^{-\alpha^{i-1}/2})},$$

and let $i_1 = O(1)$ be large enough so that $1-\alpha \geq \alpha^{i_1-1}/2$. Since $\frac{1}{2} < \alpha < 1$, $\beta := e^{-\alpha^{i_1}/2} - e^{-\alpha^{i_1-1}/2} > 0$ is a constant such that if $i \leq i_1$, then $\rho(R,i) \geq \rho(R,i_1) = (1+o(1))e^{(\alpha-\frac{1}{2})\beta R}$. Thus, the claim holds for $i \leq i_1$.

Using now $1-x \leq e^{-x} \leq 1-x+x^2/2$, by our choice of i_1, and recalling that $\alpha < 1$,

$$R\Big(e^{-\alpha^i/2} - e^{-\alpha^{i-1}/2}\Big) \geq R\frac{\alpha^{i-1}}{2}\Big(1-\alpha-\frac{\alpha^{i-1}}{4}\Big) \geq R\frac{\alpha^{i-1}}{2}\cdot\frac{1-\alpha}{2}.$$

Since by assumption $R_i < R-\xi$, we also have $\alpha^i \geq \frac{2\xi}{R}$, and thus $\rho(R,i) \geq (1+o(1))e^{(\alpha-\frac{1}{2})(1-\alpha)\xi/(2\alpha)}$, finishing the proof. □

In order to simplify our proofs, we will make use of a technique known as de-Poissonization, which has many applications in geometric probability (see [Pen03] for a detailed account of the subject). Throughout the paper we work with a Poisson point process on the hyperbolic disc of radius R and denote its point set by $\mathcal{P}$. Recall that $R = 2\log n + C$ for $C \in \mathbb{R}$. The intensity function at polar coordinates (r,θ) for $0 \leq r < R$ and $0 \leq \theta < 2\pi$ is equal to

$$g(r,\theta) := \delta e^{R/2} f(r,\theta) = \delta e^{R/2}\frac{\alpha\sinh(\alpha r)}{2\pi C(\alpha,R)}$$

with $\delta = e^{-C/2}$. Note that $\int_{r=0}^{R}\int_{\theta=0}^{2\pi} g(r,\theta)d\theta dr = \delta e^{R/2} = n$, and thus $\mathbb{E}|\mathcal{P}| = n$.

The main advantage of defining $\mathcal{P}$ as a Poisson point process is motivated by the following two properties: the number of points of $\mathcal{P}$ that lie in any region $S \cap B_O(R)$ follows a Poisson distribution with mean given by $\int_S g(r,\theta)drd\theta = n\mu(S\cap B_O(R))$, and the number of points of $\mathcal{P}$ in disjoint regions of the hyperbolic plane are independently distributed. Moreover, by conditioning $\mathcal{P}$ upon the event $|\mathcal{P}| = n$, we recover the original distribution. Note that for any k, by standard estimates for the Poisson distribution, we have

$$\mathbb{P}(|\mathcal{P}| = k) \leq \mathbb{P}(|\mathcal{P}| = n) = O(1/\sqrt{n}). \tag{4}$$

In some of the arguments below, we will add a set of vertices $\mathcal{Q}$ of size $|\mathcal{Q}| = m$, that are all chosen independently according to the same probability distribution as every vertex in $G_{\alpha,C}(n)$, we add them to $\mathcal{P}$ and consider $\mathcal{P}\cup\mathcal{Q}$. Throughout this work all vertices named by Q, or in case there are more, $Q_1,\ldots,Q_m$ are vertices added in this way. For $m = o(n^{1/2})$, by Stirling's approximation, we have

$$\mathbb{P}(|\mathcal{P}| = n-m) = e^{-n}\frac{n^{n-m}}{(n-m)!} = e^{-m}\Big(\frac{n}{n-m}\Big)^{n-m}\Theta(n^{-1/2}) = e^{-m}\Big(1+\frac{m}{n-m}\Big)^{n-m}\Theta(n^{-1/2}). \tag{5}$$

Using that $e^{x/(1+x)} \leq 1+x$ if $x > -1$ and recalling that $m = o(\sqrt{n})$, we get

$$\mathbb{P}(|\mathcal{P}| = n-m) \geq e^{-m}e^{(n-m)\frac{m}{n}}\Theta(n^{-1/2}) = e^{-o(1)}\Theta(n^{-1/2}) = \Omega(1/\sqrt{n}).$$

Thus, for such an m, by (5),

$$\mathbb{P}(|\mathcal{P}| = n-m) = \Theta(1/\sqrt{n}). \tag{6}$$

By conditioning upon $|\mathcal{P}| = n-m$, we recover the original distribution of n points. Moreover, for any $m = o(\sqrt{n})$, any event holding in $\mathcal{P}$ or $\mathcal{P}\cup\mathcal{Q}$ with probability at least $1-o(f_n)$ must hold in the original setup with probability at least $1-o(f_n\sqrt{n})$, and in particular, any event holding with probability at least $1-o(1/\sqrt{n})$ holds asymptotically almost surely (a.a.s.), that is, with probability tending to 1 as $n\to\infty$, in the original model. We identify below points of $\mathcal{P}$ with vertices. We prove all results below for the Poisson model and then transform the results to the original model.

3 The upper bound on the diameter

In this section we prove the main result of this work, a polylogarithmic upper bound in n on the diameter $G_{\alpha,C}(n)$ that holds asymptotically almost surely.

We begin by showing that there cannot be long paths with all vertices being close to the boundary (i.e., at distance $R - O(1)$ from the origin O).

Lemma 5. *Let $\xi > 0$. Let $Q \in \mathcal{Q}$ be a vertex added, and suppose $r_Q > R - \xi$. There is a constant $C' = C'(\xi, C)$ such that with probability at least $1 - o(n^{-3/2})$, there is no path $Q =: v_0, v_1, \ldots, v_k$, with $k \geq C' \log n$ and $r_{v_i} > R - \xi$ for $i = 1, \ldots, k$.*

Proof. First, note that by Remark 1, for two vertices v_i, v_j with $r_{v_i}, r_{v_j} > R - \xi$ and $d(v_i, v_j) \leq R$, we must have $|\theta_{v_i} - \theta_{v_j}| \leq (1+o(1))2e^{-R/2+\xi} = C''/n$ for $C'' = (2+o(1))e^{-C/2+\xi}$. Partition the disc of radius R into $\Theta(n)$ equal sized sectors of angle $2\pi C''/n$ and order them counterclockwise. Note that any path containing only vertices v with $r_v > R - \xi$ satisfies the following property: if it contains a vertex v_i in a sector S_i and a vertex v_j in a sector S_j, either all sectors between S_i and S_j in the counterclockwise ordering or all sectors between S_i and S_j in the clockwise ordering have to contain at least one vertex from the path. By symmetry, for each sector, the expected number of vertices inside the sector is C'', and therefore this is also an upper bound for the expected number of vertices v inside this sector with the additional restriction of $r_v > R - \xi$. Thus, the probability of having at least one vertex inside a sector is at most $1 - e^{-C''} < 1$. The probability to have at least one vertex in $C''' \log n$ given consecutive sectors starting from the sector that contains Q is thus at most $2(1 - e^{-C''})^{C''' \log n} = o(n^{-3/2})$ for $C''' = C'''(C'')$ sufficiently large. Thus, with probability at least $1 - o(n^{-3/2})$, any such path starting at Q contains only vertices from at most $C''' \log n$ sectors. By Chernoff bounds together with a union bound, for C''' large enough, with probability at least $1 - o(n^{-3/2})$, any consecutive set of $C''' \log n$ sectors contains $O(\log n)$ vertices, and the statement follows. □

Because of the peculiarities of hyperbolic space, two points close to the boundary R that are within distance at most R from each other subtend a small angle at the origin. In fact, the angle is smaller the closer both points are to the boundary. This fact, coupled with the previous lemma, implies that with high probability there is no path all of whose vertices are within a constant distance of the boundary whose extreme vertices subtend a large angle at the origin. Formally, we have the following corollary:

Corollary 6. *Let $\xi > 0$. Let $Q \in \mathcal{Q}$ be a vertex added, and suppose $r_Q > R - \xi$, and let $Q =: v_0, v_1, \ldots, v_k$ be a path with $r_{v_i} > R - \xi$ for $i = 1, \ldots, k$. Then, with probability at least $1 - o(n^{-3/2})$, for any $0 \leq i \leq j \leq k$,*

$$|\theta_{v_i} - \theta_{v_j}| = O(\log n / n).$$

Proof. By Lemma 5, with probability at least $1 - o(n^{-3/2})$, $k = O(\log n)$. By the proof of Lemma 5, for any two v_i, v_j with $r_{v_i}, r_{v_j} > R - \xi$ and $d(v_i, v_j) \leq R$, we must have $|\theta_{v_i} - \theta_{v_j}| = O(1/n)$. The corollary follows by the triangle inequality for angles. □

Recall that $C_0 = 2/(\frac{1}{2} - \frac{3}{4}\alpha + \frac{\alpha^2}{4})$ and note that $C_0 > 8$. Define throughout the whole section $i_0 := \log_\alpha((2C_0 \log R)/R)$. Note that $R_{i_0} = Re^{-\alpha^{i_0}/2} = R(1 - (1+o(1))\frac{\alpha^{i_0}}{2}) = R - (1+o(1))C_0 \log R$.

As already pointed out, the farther from the origin two vertices within distance at most R from each other are, the smaller the angle they subtend at the origin. In particular, for any two adjacent vertices v_i, v_{i+1} with $r_{v_i}, r_{v_{i+1}} > R_{i_0}$, by Remark 1,

$$|\theta_{v_i} - \theta_{v_{i+1}}| \leq (2+o(1))e^{\frac{1}{2}(R - r_{v_i} - r_{v_{i+1}})} \leq (2+o(1))e^{\frac{1}{2}(R - 2R_{i_0})} = e^{-\frac{1}{2}R} R^{C_0(1+o(1))} = O(\tfrac{1}{n} \log^{C_0+o(1)} n). \quad (7)$$

For a vertex $v \in \mathcal{P}$ with $r_v > R_{i_0}$, let ℓ be such that $R_\ell < r_v \leq R_{\ell+1}$. Define a *center path* from v to be a sequence of vertices $v =: w_0, w_1, \ldots, w_s$ such that $d(w_i, w_{i+1}) \leq R$, $R_{\ell-i} < r_{w_i} \leq R_{\ell-i+1}$ for $0 \leq i \leq s-1$, and $R_{i_0-1} < r_{w_s} \leq R_{i_0}$. In words, $w_0, w_1, \ldots, w_s$ is a center path from v provided $w_0 = v \in B_O(R_{\ell+1}) \setminus B_O(R_\ell)$, $w_s \in B_O(R_{i_0})$, and for every $0 < i < s$ the vertex w_i is in the band $B_O(R_{\ell-i+1}) \setminus B_O(R_{\ell-i})$ around the origin; note that each band is closer to the origin than the previous one.

We now establish the fact that if a vertex is "far" from the origin but some constant distance away from the boundary, then its probability of connecting via a center path to a vertex inside $B_O(R_{i_0})$ is at least a constant.

Lemma 7. *Let $\xi' > 0$ and $Q \in \mathcal{Q}$ be a vertex added. Suppose that $R_\ell < r_Q \leq R_{\ell+1}$ for $\ell \geq i_0$ and $R_\ell \leq R - \xi'$. Then, with probability bounded away from 1, there exists no center path $Q =: \omega_0, \dots, \omega_s$ from Q.*

Proof. Denote by $\mathcal{E}_0$ the event that there exists one vertex of $\mathcal{P}$ that belongs to $B_{w_0}(R) \cap (B_O(R_\ell) \setminus B_O(R_{\ell-1}))$. By Lemma 4,

$$\mu(B_{w_0}(R) \cap (B_O(R_\ell) \setminus B_O(R_{\ell-1}))) = \Omega(\mu(B_{w_0}(R) \cap B_O(R_\ell))).$$

By Lemma 3 (applied with $\rho_A = R$, $\rho_O = R_\ell$, and $R_\ell < r_A = r_{w_0} \leq R_{\ell+1}$), we have

$$\mu(B_{w_0}(R) \cap B_O(R_\ell)) = \frac{2\alpha}{\pi(\alpha - \frac{1}{2})}\Big(e^{-(\alpha-\frac{1}{2})(R-R_\ell)-\frac{1}{2}r_{w_0}} + O(e^{-\alpha r_{w_0}})\Big) = \Theta(e^{-(\alpha-\frac{1}{2})(R-R_\ell)-\frac{1}{2}r_{w_0}}),$$

where the last identity follows because $r_{w_0} > R - R_\ell$. Now, since $1 - e^{-x} \leq x$, we have $R - R_\ell = R(1 - e^{-\alpha^\ell/2}) \leq R\alpha^\ell/2$, and since for $x = o(1)$, $e^{-x} = (1 + o(1))(1 - x)$, we also have $r_{w_0} \leq R_{\ell+1} = Re^{-\alpha^{\ell+1}/2} = R(1 - (1 + o(1))\alpha^{\ell+1}/2)$. Thus

$$\mathbb{P}({\mathcal{E}_0}^c) = \exp(-\Omega(ne^{-(\alpha-\frac{1}{2})(R-R_\ell)-\frac{1}{2}r_{w_0}})) = \exp(-\Omega(ne^{-(\alpha-\frac{1}{2})\frac{R}{2}\alpha^\ell-(1-(1+o(1))\frac{R}{4}\alpha^{\ell+1})})).$$

Recalling that $e^{-R/2} = \Theta(1/n)$, we have

$$\mathbb{P}({\mathcal{E}_0}^c) = \exp(-\Omega(e^{-(\alpha-\frac{1}{2})\frac{R}{2}\alpha^\ell+(1+o(1))\frac{R}{4}\alpha^{\ell+1}})). \tag{8}$$

Note that if $\mathcal{E}_0$ holds, then a vertex inside the desired region is found. Assuming the existence of a vertex w_i with $R_{\ell-i} < r_{w_i} \leq R_{\ell-i+1}$, we continue inductively for $i = 1, \dots, s-1 = \ell - i_0$ in the same way: we define the event $\mathcal{E}_i$ that there exists one element of $\mathcal{P}$ that belongs to $B_{w_i}(R) \cap (B_O(R_{\ell-i}) \setminus B_O(R_{\ell-i-1}))$. By the same calculations (noting that Lemma 3 can still be applied and $R_{\ell-i} + r_{w_i} > R$ still holds), we obtain

$$\mathbb{P}({\mathcal{E}_i}^c) = \exp(-\Omega(e^{-(\alpha-\frac{1}{2})\frac{R}{2}\alpha^{\ell-i}+(1+o(1))\frac{R}{4}\alpha^{\ell-i+1}})).$$

Denote by $\mathcal{C}$ the event of not having a center path, we have thus by independence of the events

$$\mathbb{P}(\mathcal{C}) \leq \mathbb{P}({\mathcal{E}_0}^c) + \sum_{i=1}^{s-1}\mathbb{P}({\mathcal{E}_i}^c | \mathcal{E}_0, \dots, \mathcal{E}_{i-1}) = \sum_{i=0}^{s-1}\mathbb{P}({\mathcal{E}_i}^c).$$

Hence,

$$\begin{aligned}\mathbb{P}(\mathcal{C}) &\leq \sum_{i=0}^{\ell-i_0}\exp(-\Omega(e^{-(\alpha-\frac{1}{2})\frac{R}{2}\alpha^{\ell-i}+(1+o(1))\frac{R}{4}\alpha^{\ell-i+1}})) \leq \sum_{i=0}^{\ell-i_0} e^{-\Omega(1-(\alpha-\frac{1}{2})\frac{R}{2}\alpha^{\ell-i}+(1+o(1))\frac{R}{4}\alpha^{\ell-i+1})} \\ &\leq \sum_{i=0}^{\ell-i_0} e^{-\Omega(1+(1+o(1))\frac{R}{4}\alpha^{\ell-i}(1-\alpha))} \leq \sum_{i=0}^{\ell-i_0} e^{-C'-(1+o(1))C'\frac{R}{4}\alpha^{\ell-i}(1-\alpha)},\end{aligned}$$

where $C' = C'(\xi') > 0$ is the constant hidden in the asymptotic notation of Lemma 4. Clearly, the closer Q is to the boundary, the more difficult it is to find a center path, and we may thus assume that $R_\ell < r_Q \leq R_{\ell+1}$ is such that $R_{\ell+1} > R - \xi'$. Then, noting that $R_{\ell+1} = Re^{-\alpha^{\ell+1}/2} > R - \xi'$ implies that $\ell = \log_\alpha\big((1+o(1))2\xi'/R\big)$ must hold. Plugging this into the previous sum we get

$$\mathbb{P}(\mathcal{C}) \leq e^{-C'}\sum_{i\geq 0} e^{-(1+o(1))C'\frac{\xi'(1-\alpha)}{2\alpha^i}}.$$

Clearly, since $\alpha^{-1} > 1$, the sum converges. Note that the constant C' coming from Lemma 4 is nondecreasing as a function of ξ'. Hence, by choosing $\xi' = \xi'(C', \alpha)$ big enough, the sum is less than 1, and the statement follows. □

Define $j_0 = j_0(\alpha) \geq 1$ to be a constant sufficiently large so that $e^{-\alpha^j/2} \leq 1 - (1 - \frac{1-\alpha}{2})\frac{\alpha^j}{2}$ for $j \geq j_0$ (note that such j_0 exists because $e^{-x} \leq 1 - x + x^2$ if $|x| \leq 1$).

The following lemmas will show that for vertices v with $r_v \leq R_{i_0}$ the probability of not connecting to a vertex in $B_O(R/2)$ can be bounded by a much smaller expression than the analogous probability bound for vertices close to the boundary R given by Lemma 7. First, we consider the case of vertices in the band $B_O(R_j) \setminus B_O(R_{j-1})$ for $j_0 \leq j \leq i_0$, and show that w.e.p. they have neighbors in the next band closer to the origin. Formally, we establish the following result, whose proof argument is reminiscent of that of Lemma 7, although the calculations involved are different.

Lemma 8. *Let $Q \in \mathcal{Q}$ be a vertex added such that $R_j < r_Q \leq R_{j+1}$ with $j_0 \leq j \leq i_0$. W.e.p.,*

$$\big(B_Q(R) \cap (B_O(R_j) \setminus B_O(R_{j-1})\big) \cap \mathcal{P} \neq \emptyset.$$

Proof. As in the previous proof, by Lemma 4,

$$\mu(B_Q(R) \cap (B_O(R_j) \setminus B_O(R_{j-1}))) = \Omega(\mu(B_Q(R) \cap B_O(R_j))), \tag{9}$$

and also as before, by Lemma 3 (applied with $\rho_A = R$, $\rho_O = R_j$, and $R_j < r_A = r_Q \leq R_{j+1}$), we have

$$\mu(B_Q(R) \cap B_O(R_j)) = \Theta(e^{-(\alpha-\frac{1}{2})(R-R_j)-\frac{1}{2}r_Q}).$$

Again, $R - R_j = R(1 - e^{-\alpha^j/2}) \leq R\alpha^j/2$, and since by assumption $e^{-\alpha^j/2} \leq 1 - (1 - \frac{1-\alpha}{2})\frac{\alpha^j}{2}$, we have $r_Q \leq R_{j+1} = Re^{-\alpha^{j+1}/2} \leq R(1 - (1 - \frac{1-\alpha}{2})\frac{\alpha^{j+1}}{2})$. Thus,

$$\mu(B_z(R) \cap B_O(R_j)) = \Omega(e^{-(\alpha-\frac{1}{2})\frac{R}{2}\alpha^j - (1-(1-\frac{1-\alpha}{2})\frac{R}{4}\alpha^{j+1})}) = \Omega(e^{-R/2}e^{\frac{R}{2}\alpha^j(\frac{1}{2}-\frac{3}{4}\alpha+\frac{\alpha^2}{4})}).$$

Note that $\frac{1}{2} - \frac{3}{4}\alpha + \frac{\alpha^2}{4} > 0$ for $\frac{1}{2} < \alpha < 1$, so the last displayed expression is clearly decreasing in j. By plugging in our upper bound $j = i_0 = \log_\alpha(2C_0 \log R/R)$, we get with our choice of $C_0 = 2/(\frac{1}{2} - \frac{3}{4}\alpha + \frac{\alpha^2}{4})$,

$$\mu(B_Q(R) \cap B_O(R_j)) = \Omega\Big(\frac{R^2}{e^{R/2}}\Big) = \Omega((\log n)^2/n).$$

Hence, the expected number of vertices in $B_Q(R) \cap B_O(R_j)$ is $\Omega(\log^2 n)$, and by Chernoff bounds for Poisson random variables (see [AS08, Theorem A.1.15]), w.e.p. there are at least $\Omega(\log^2 n)$ vertices in this region. By (9) the statement follows. □

Next, we show the analogue of the preceding lemma, but for vertices within $B_O(R_{j_0})$ and a somewhat different choice of concentric bands. Indeed, for vertices $v \in \mathcal{P}$ with $R/2 < r_v \leq R_{j_0}$ we modify the definition of R_i: since $j_0 = O(1)$, there exists some $\frac{1}{2} < c < 1$ such that $R_{j_0} = Re^{-\alpha^{j_0}/2} =: cR$. Let $T \geq 1$ be the largest integer such that $c - \frac{T}{2}(1-c)(1-\alpha) > \frac{1}{2}$. Define now the new bands to be $R'_0 := cR$, and for any $i = 1, \ldots, T$, define $R'_i := R(c - \frac{i}{2}(1-c)(1-\alpha))$. Note in particular, that $R'_i \geq R/2$ for all i in the range of interest. We have the following result.

Lemma 9. *Let $Q \in \mathcal{Q}$ be a vertex added.*

- *If $R'_j < r_Q \leq R'_{j-1}$ for some $1 \leq j \leq T-1$, then w.e.p., we have*

$$\big(B_Q(R) \cap (B_O(R'_j) \setminus B_O(R'_{j+1})\big) \cap \mathcal{P} \neq \emptyset.$$

- *If $R'_T < r_Q \leq R'_{T-1}$, then w.e.p. we have*

$$\big(B_Q(R) \cap B_O(R/2)\big) \cap \mathcal{P} \neq \emptyset.$$

Proof. First, assume $1 \leq j \leq T-1$. By Lemma 3 (applied with $\rho_A = R$, $\rho_O = R'_j$, and $R'_j < r_A = r_Q \leq R'_{j-1}$), since $R'_j + r_Q > R$ still holds, we have

$$\mu(B_Q(R) \cap B_O(R'_j)) = \Theta(e^{-(\alpha-\frac{1}{2})(R-R'_j)-\frac{1}{2}r_Q}) = \Theta(e^{-\frac{R}{2}(2\alpha-1)(1-c+\frac{j}{2}(1-c)(1-\alpha))-\frac{1}{2}r_Q}),$$

and also

$$\mu(B_Q(R) \cap B_O(R'_{j+1})) = \Theta(e^{-\frac{R}{2}(2\alpha-1)(1-c+\frac{j+1}{2}(1-c)(1-\alpha))-\frac{1}{2}r_Q}).$$

Thus,

$$\mu(B_Q(R) \cap B_O(R'_{j+1})) = \Theta(e^{-\frac{R}{4}(2\alpha-1)(1-c)(1-\alpha)})\mu(B_Q(R) \cap B_O(R'_j)),$$

so we get

$$\mu(B_Q(R) \cap (B_O(R'_j) \setminus B_O(R'_{j+1}))) = (1+o(1))\mu(B_Q(R) \cap B_O(R'_j)). \tag{10}$$

Now, since $r_Q \leq R'_{j-1}$,

$$\begin{aligned}
&\mu(B_Q(R) \cap B_O(R'_j)) = \Omega\big(e^{-\frac{R}{2}(2\alpha-1)(1-c+\frac{j}{2}(1-c)(1-\alpha))-\frac{1}{2}R'_{j-1}}\big) \\
&= \Omega\Big(e^{-\frac{R}{2}\big((2\alpha-1)(1-c+\frac{j}{2}(1-c)(1-\alpha))+c-\frac{j-1}{2}(1-c)(1-\alpha)\big)}\Big) = \Omega\Big(e^{-\frac{R}{2}\big(-j(1-c)(1-\alpha)^2+(1-c)(2\alpha-1+\frac{1-\alpha}{2})+c\big)}\Big).
\end{aligned}$$

Clearly, the bigger j, the bigger the last displayed term is. Thus, we plug in the smallest possible value $j = 1$, and together with (10) we obtain

$$\begin{aligned}
\mu(B_Q(R) \cap (B_O(R'_j) \setminus B_O(R'_{j+1}))) &= \Omega\Big(e^{-\frac{R}{2}\big(-(1-c)(1-\alpha)^2+(1-c)(2\alpha-1+\frac{1-\alpha}{2})+c\big)}\Big) \\
&= \Omega\Big(e^{-\frac{R}{2}\big((1-c)(\frac{7\alpha-3}{2}-\alpha^2)+c\big)}\Big).
\end{aligned}$$

Note that for $\frac{1}{2} < \alpha < 1$, we have $\frac{7\alpha-3}{2} - \alpha^2 < 1$, and thus the constant factor multiplying $-\frac{R}{2}$ in the exponent of the last term of the previously displayed equation is clearly bounded by $c < 1$. Hence, $\mu(B_Q(R) \cap (B_O(R'_j) \setminus B_O(R'_{j+1}))) = \Omega(e^{-\gamma R/2})$ for some $\gamma < 1$. Hence, the expected number of vertices inside $B_Q(R) \cap (B_O(R'_j) \setminus B_O(R'_{j+1}))$ is $n^{1-\gamma}$. By Chernoff bounds the first part of the lemma follows. For the second part, by Lemma 3 (applied with $\rho_A = R$, $\rho_O = R/2$, and $R'_T < r_A = r_Q \leq R'_{T-1}$), since $r_Q > R/2$ still holds, we have

$$\mu(B_Q(R) \cap B_O(R/2)) = \Theta(e^{-(\alpha-\frac{1}{2})\frac{R}{2}-\frac{1}{2}r_Q}) = \Omega(e^{-(\alpha-\frac{1}{2})\frac{R}{2}-\frac{1}{2}R'_{T-1}}).$$

Note that $R'_{T-1} \leq \frac{R}{2} + R(1-c)(1-\alpha)$ must hold, as otherwise $R'_{T+1} = R'_{T-1} - R(1-c)(1-\alpha) > R/2$ would hold, and then $c - \frac{T+1}{2}(1-c)(1-\alpha) > \frac{1}{2}$, contradicting the definition of T. Thus,

$$\mu(B_Q(R) \cap B_O(R/2)) = \Omega(e^{-\frac{R}{2}\big(\alpha+(1-c)(1-\alpha)\big)}).$$

Since $\alpha + (1-c)(1-\alpha) < 1$, $\mu(B_Q(R) \cap B_O(R/2)) = n^{1-\gamma'}$ for some $\gamma' > 0$, and by the same argument as before, the second part of the lemma follows. □

Intuitively, due to the nature of hyperbolic space, one expects that vertices in a center path from a vertex v that is "sufficiently far" from the origin, tend to stay close to the ray between v and the origin. One way of capturing this foreknowledge is to show that the angle subtended at the origin by the endpoints of any such center path is "rather small". The following lemma formalizes this intuition.

Lemma 10. *Let u be a vertex with $r_u > R_{i_0}$ and let $u =: u_0, u_1, \ldots, u_m$ be a center path starting at u, or the longest subpath found in an attempt of finding a center path starting at u, in case no center path is found. Then,*

$$|\theta_u - \theta_{u_m}| = o(\log^{C_0} n/n).$$

Algorithm 1 Sequence in which regions are exposed

procedure EXPOSE(Q, ξ)
  Let ℓ_0 be the smallest integer such that $R_{\ell_0} > R - \xi$
  for $j = 0, \dots, C'' \log n$ **do**
    Expose $A^0(R_{i_0})$
    $\ell \leftarrow \ell_0 + 1$
    repeat ▷ Start of $(j+1)$-th phase
      $\ell \leftarrow \ell - 1$
      if $\ell = i_0$ **then return** 'no path' and stop
      Expose $A^{j+1}(R_\ell) \setminus A^j(R_{i_0})$ ▷ $(\ell_0 - \ell + 1)$-th sub-phase of the $(j+1)$-th phase
    until $\big(\exists Q =: v_0, \dots, v_{k_{j+1}} \in \mathcal{P}$ a path, $v_i \in A^j(R_{i_0}) \cup A^{j+1}(R_\ell)$ for $i < k_{j+1}$,
      $v_{k_{j+1}} \in A^{j+1}(R_\ell) \setminus A^j(R_{i_0})$ and $r_{v_{k_{j+1}}} \le R - \xi\big)$
    if $\exists$ center path starting from $v_{k_{j+1}}$ **then**
      return 'success' and stop
  return 'failure' and stop

Proof. Let ℓ be such that $R_\ell < r_u \le R_{\ell+1}$, and note that $\ell \ge i_0$. Since by making paths longer, the regions exposed get only larger and differences in angles get larger as well, it suffices to consider the case where indeed a center path is found. Suppose there exists such a center path starting from u, that is, a sequence of vertices denoted by $u =: u_0, u_1, \dots, u_m$, satisfying $d(u_j, u_{j+1}) \le R$ for all $0 \le j \le m-1$, $R_{\ell-j} < r_{u_j} \le R_{\ell-j+1}$ for any $0 \le j \le m-1$, and $R_{i_0-1} < r_{u_m} \le R_{i_0}$. By Remark 1, $|\theta_{u_j} - \theta_{u_{j+1}}| \le 2e^{\frac{1}{2}(R - R_{\ell-j} - R_{\ell-j+1})}(1+o(1))$, and thus, by the triangle inequality for angles,

$$\begin{aligned}|\theta_u - \theta_{u_m}| &\le \sum_{j=0}^{m-1} |\theta_{u_j} - \theta_{u_{j+1}}| \le 2(1+o(1)) \sum_{j=0}^{m-1} e^{\frac{1}{2}(R - R_{\ell-j} - R_{\ell-j+1})} \\ &= 2(1+o(1)) \sum_{i=i_0+1}^{m+i_0} e^{\frac{1}{2}(R - R_{i-1} - R_i)} \le 2(1+o(1)) \sum_{i \ge i_0} 2e^{\frac{1}{2}(R - R_i - R_{i+1})}.\end{aligned}$$

Since $R_i = Re^{-\alpha^i/2} = (1+o(1))R(1 - \alpha^i/2)$ for $i \ge i_0$,

$$\begin{aligned}|\theta_u - \theta_{u_m}| &\le (2+o(1)) e^{-\frac{1}{2}R} \sum_{i \ge i_0} e^{(1+o(1))\frac{R}{4}\left(\alpha^i + \alpha^{i+1}\right)} = O(e^{-\frac{1}{2}R} R^{\frac{1}{2}(C_0 + C_0\alpha)(1+o(1))}) \\ &= O(n^{-1} (\log n)^{\frac{1}{2}C_0(1+\alpha)(1+o(1))}) = o(\log^{C_0} n/n),\end{aligned}$$

where the last equality is a consequence of $\alpha < 1$. The statement follows. ☐

We now have all the necessary ingredients to state and prove the main lemma.

Lemma 11. *Let $Q \in \mathcal{Q}$ be a vertex added. With probability at least $1 - o(n^{-3/2})$, Q either connects to a vertex in $\mathcal{P} \cap B_O(R_{i_0})$ through a path of length $O(\log^{C_0+1+o(1)} n)$, or any path starting at Q has length at most $O(\log^{C_0+1+o(1)} n)$.*

Proof. Suppose first that $r_Q \le R_{i_0}$. By applying iteratively either Lemma 8 and Lemma 9 (if $R_{j_0} < r_Q \le R_{i_0}$), or only Lemma 9 (if $r_Q \le R_{j_0}$), w.e.p., Q is connected to a vertex in $B_O(R/2)$ through a path of length $O(\log R) + O(T) = O(\log \log n)$.

Otherwise, suppose now $r_Q > R_{i_0}$. If $r_Q > R - \xi$ for some $\xi > 0$, then by Lemma 5, with probability at least $1 - o(n^{-3/2})$, either all paths starting at Q have length $O(\log n)$, or after at most $O(\log n)$ steps we reach a vertex v with $r_v \le R - \xi$. Since an additional $O(\log n)$ steps will be negligible in the following, we may thus assume $r_Q \le R - \xi$. Let ℓ_0 be the smallest integer i such that $R_i > R - \xi$.

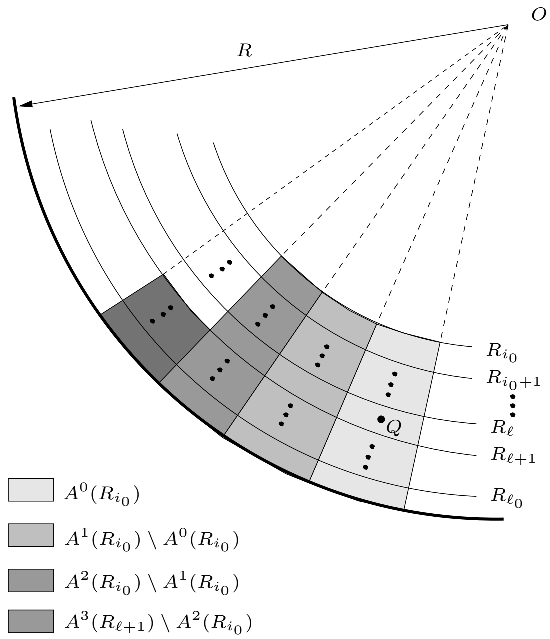

Figure 3: Illustration of sequence of exposures of procedure Expose.

Let ℓ be such that $R_\ell < r_Q \leq R_{\ell+1}$ and note that $\ell \geq i_0$. By Lemma 7, the probability of having no center path starting at $v_0 = v_{k_0} := Q$ is at most $c < 1$. If we find a center path starting at v_{k_0}, we have reached our goal and stop with success. Otherwise, we will expose regions around the location of Q in different exposing phases. Specifically, fix $\varepsilon > 0$ to be any arbitrary small constant. Then, for $0 \leq j \leq C'' \log n$, where C'' is a sufficiently large constant, define

$$A^j(\rho) := \{P \mid r_P > \rho, |\theta_P - \theta_Q| \leq \tfrac{j+1}{n} \log^{C_0+\varepsilon} n\}$$

and perform the Expose procedure whose pseudocode is given in Algorithm 1 (see illustration in Figure 3).

If 'no path' is reported in Line 8 of Expose during phase $j+1$, then among the paths starting from v_0 with all vertices v_i satisfying $r_{v_i} > R_{i_0}$, there is no path with its last vertex $v_{k_{j+1}} \in A^{j+1}(R_{i_0}) \setminus A^j(R_{i_0})$ satisfying $r_{v_{k_{j+1}}} \leq R - \xi$. We will show now that the length of any such path is with probability at least $1 - o(n^{-3/2})$ bounded by $O(j \log^{C_0+\varepsilon} n)$: indeed, partition the disc of radius R centered at the origin into $\Theta(n/\log n)$ equal sized sectors of angle $C' \log n/n$ for $C' > 0$ sufficiently large. By Chernoff bounds for Poisson variables (see for example [AS08, Theorem A.1.15]) together with a union bound over all $\Theta(n/\log n)$ sectors, with probability at least $1-o(n^{-3/2})$, the number of vertices in each sector is $\Theta(\log n)$. Thus, with this probability, the number of vertices in $A^j(R_{i_0})$ is $O(j \log^{C_0+\varepsilon} n)$. By (7) any two adjacent vertices w, w' with $r_w, r'_w > R_{i_0}$ satisfy $|\theta_w - \theta_{w'}| = O(\log^{C_0+o(1)} n/n)$. If a path starting from v_0 with all vertices v_i satisfying $r_{v_i} > R_{i_0}$ ends with some vertex $v \in A^j(R_{i_0})$, then the path does not extend beyond $A^{j+1}(R_{i_0})$, and thus in this case, with probability at least $1-o(n^{-3/2})$, its length is at most the number of vertices in $A^j(R_{i_0})$, i.e., $O(j \log^{C_0+\varepsilon} n)$. If the path ends with a vertex $v \in A^{j+1}(R_{i_0}) \setminus A^j(R_{i_0})$, then note that $r_v > R - \xi$ must hold. Consider then the last vertex w in this path with $w \in A^j(R_{i_0})$. Since the next vertex w' satisfies $w' \in A^{j+1}(R_{i_0}) \setminus A^j(R_{i_0})$ and $r'_w > R - \xi$, by calculations as in (7), it must hold that $|\theta_w - \theta_{w'}| = O(\log^{C_0/2+o(1)} n/n)$. Moreover, by Corollary 6, with probability at least $1-o(n^{-3/2})$, any two vertices w'', w''' on a path containing only vertices inside $A^{j+1}(R_{i_0}) \setminus A^j(R_{i_0})$ satisfy $|\theta_{w''} - \theta_{w'''}| = O(\log n/n)$. Thus, also $|\theta_w - \theta_v| = O(\log^{C_0/2+o(1)} n/n)$. Since this would also have to hold for another possible neighbor z of v with $r_z > R_{i_0}$, and $z \notin A^{j+1}(R_{i_0})$, starting from v the path cannot extend to a vertex outside $A^{j+1}(R_{i_0})$ as well. Hence, with probability at least $1 - o(n^{-3/2})$, the number of vertices of such a path is bounded by $O((j+1) \log^{C_0+\varepsilon} n)$.

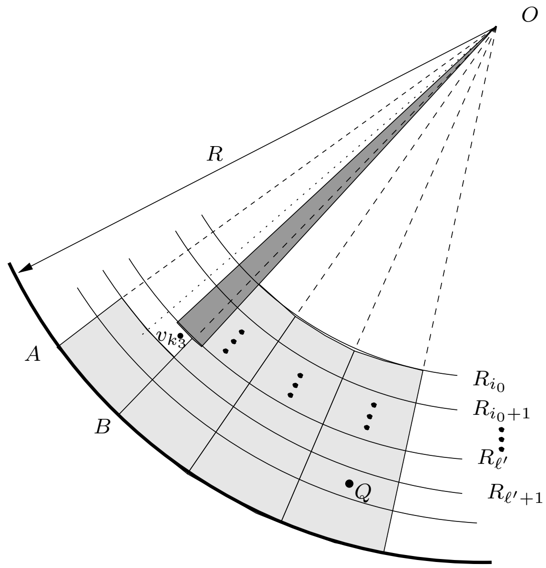

Figure 4: The region that has already been exposed is lightly shaded. The AOB slice of $B_O(R)$ corresponds to S_{j+1}. The darkly shaded area contains the region to be exposed in a center path without restrictions starting at v_{k_3}.

If on the other hand the condition of Line 10 of Expose is satisfied during phase j, then for some $i_0 \leq \ell' \leq \ell_0$ there is a path $v_0, \dots, v_{k_{j+1}} \in \mathcal{P}$ with $v_i \in A^j(R_{i_0}) \cup A^{j+1}(R_{\ell'})$ for all $i < k_{j+1}$, $v_{k_{j+1}} \in A^{j+1}(R_{\ell'}) \setminus A^j(R_{i_0})$. Recall that $r_{v_{k_{j+1}}} \leq R - \xi$, and observe that $R_{\ell'} < r_{v_{k_{j+1}}} \leq R_{\ell'+1}$. Consider then the restriction of a center path starting from $v_{k_{j+1}}$ to the region

$$S_{j+1} := \{P \mid \tfrac{j+1}{n} \log^{C_0+\varepsilon} n < |\theta_P - \theta_Q| \leq \tfrac{j+2}{n} \log^{C_0+\varepsilon} n\}.$$

Note first that since $R_{\ell'} < r_{v_{k_{j+1}}} \leq R_{\ell'+1}$, the first region to be exposed on the center path starting from $v_{k_{j+1}}$ contains vertices v with $R_{\ell'-1} < r_v \leq R_{\ell'}$, and this region when restricted to S_{j+1} has not been exposed before. In particular, the regions exposed by the center paths of v_{k_i} and $v_{k_{j+1}}$, whenever restricted to S_i and S_{j+1}, respectively, for $i < j+1$ are disjoint. Suppose without loss of generality that $|\theta_Q - \theta_{v_{k_{j+1}}}| \leq (j + \frac{3}{2}) \log^{C_0+\varepsilon} n/n$. Observe that by Lemma 10 the angle between any vertex v with $r_v > R_{i_0}$ and the terminal vertex of the center path starting at v or the longest subpath in an attempt of building a center path starting at v, in case a center path is not found, is bounded by $o(\log^{C_0} n/n)$. Therefore, at least half of the region to be exposed by a center path in each step is retained in S_{j+1}: more precisely (see Figure 4), in the first step the region containing all points P that were to be exposed in a center path without restrictions and that satisfy $|\theta_Q - \theta_P| \geq |\theta_Q - \theta_{v_{k_{j+1}}}|$ is included in S_{j+1}, and at every subsequent step on the center path starting from a vertex u, the region of all points P that were to be exposed in a center path without restrictions and that satisfy $|\theta_Q - \theta_P| \geq |\theta_Q - \theta_u|$ is always included in S_{j+1}. Lemma 7 can still be applied, and the probability of having no center path starting at $v_{k_{j+1}}$ is at most $c' < 1$, and by disjointness, this is clearly independent of all previous events.

The probability that the exposing process returns 'failure' is the probability that during $C'' \log n$ phases no center path is found. By independence, for $C''(c)$ sufficiently large, this probability is at most ${c'}^{C'' \log n} = o(n^{-3/2})$. Thus, with probability at least $1 - o(n^{-3/2})$, the exposing process ends with 'no path' in some phase j, or with 'success' in some phase j'. If the exposing process ends with 'no path' in some phase j, with probability $1 - o(n^{-3/2})$ any path starting at v_0 containing only vertices v with $r_v > R_{i_0}$ has length at most $O((j+1) \log^{C_0+\varepsilon} n) = O(\log^{C_0+1+\varepsilon} n)$. If the exposing process ends with 'success' in some phase j', a center path of length $O(\log R)$ starting from $v_{k_{j'}}$ is found, and as in the case of 'no path', the total length of the path starting at v_0 together with the center path is at most $O(\log^{C_0+1+\varepsilon} n)$. Thus, with probability at least $1-o(n^{-3/2})$, v_0 either connects to a vertex in $\mathcal{P} \cap B_O(R_{i_0})$ through a path of length $O(\log^{C_0+1+\varepsilon} n)$, or any path starting at v_0 has length at most $O(\log^{C_0+1+\varepsilon} n)$, and the statement follows, since ε can be chosen arbitrarily small. ☐

The upper bound for the diameter now follows easily.

Theorem 12. *A.a.s., any two vertices u and v belonging to the same connected component satisfy $d_G(u,v) = O(\log^{C_0+1+o(1)} n)$.*

Proof. Let $Q \in \mathcal{Q}$ be a vertex added in the Poissonized model. By Lemma 11, with probability at least $1 - o(n^{-3/2})$, either Q connects to a vertex in $B_O(R/2)$ using $O(\log^{C_0+1+o(1)} n)$ steps, or all paths starting from Q have length $O(\log^{C_0+1+o(1)} n)$. Note that all vertices inside $B_O(R/2)$ form a clique, and by Lemma 2 together with Chernoff bounds there are w.e.p. $\Theta(n^{1-\alpha})$ vertices inside $B_O(R/2)$. By de-Poissonizing the model, using (6), the same statement holds with probability at least $1 - o(n^{-1})$ for our particular choice of Q in the uniform model. Using a union bound over all n vertices, with probability $1 - o(1)$ the same holds for all vertices simultaneously. Thus, the maximum graph distance between any two vertices of the same connected component is $O(\log^{C_0+1+o(1)} n)$, and the statement follows. ☐

By the result of Bode, Fountoulakis and Müller [BFM13], a.a.s., all vertices belonging to $B_O(R/2)$ are part of a giant component. We call this the *center giant component*, and we will show that there is no other giant component.

Corollary 13. *A.a.s., the size of the second largest component of $G_{\alpha,C}(n)$ is $O(\log^{2C_0+1+o(1)} n)$.*

Proof. Having de-Poissonized some results in the previous proof, we may now work directly in the model $G_{\alpha,C}(n)$ during this proof. As in the proof of Lemma 11, partition the disc of radius R centered at the origin into $\Theta(n/\log n)$ equal sized sectors of angle $C' \log n/n$ for some large constant $C' > 0$ and order them counterclockwise. The distance between sectors is measured according to this ordering: two sectors are at distance D from each other, if the shorter of the two routes around the clock between the two sectors has exactly $D-1$ sectors in between.

Now, by the proof of Theorem 12, a.a.s., every vertex v with $r_v \le R_{i_0}$ connects to $B_O(R/2)$, so a.a.s., every such vertex is part of the center giant component. We may assume that all other vertices v in other components satisfy $r_v > R_{i_0}$, and consider only these components from now on. By Theorem 12, a.a.s., for any two vertices of the same component, their graph distance is bounded by $O(\log^{C_0+1+o(1)} n)$. Recall that for any two vertices v_i, v_j with $d(v_i, v_j) \le R$ and $r_{v_i}, r_{v_j} > R_{i_0}$, by (7), $|\theta_{v_i} - \theta_{v_j}| = O((\log n)^{C_0+o(1)}/n)$, and therefore, by the triangle inequality for angles, for any two vertices u and v belonging to the same component, a.a.s. we have $|\theta_u - \theta_v| = O((\log n)^{2C_0+1+o(1)}/n)$. Since for any two vertices u and v belonging to sectors at distance $2 \le D = o(n/\log n)$ from each other we have $|\theta_u - \theta_v| = \Theta(D \log n/n)$, it must hold that between any two vertices u and v of the same component, there are at most $O(\log^{2C_0+o(1)} n)$ sectors. By Chernoff bounds for binomial random variables together with a union bound over all sectors, for $C' > 0$ large enough, a.a.s., in any sector there are at most $O(\log n)$ vertices. Thus, a.a.s., the number of vertices of the second largest component is bounded by $O(\log^{2C_0+1+o(1)} n)$. ☐

4 The lower bound on the diameter

In this section we show that the diameter of a random hyperbolic graph is $\Omega(\log n)$. In fact, we do more, since we actually establish the existence of a component forming a path of length $\Theta(\log n)$.

To achieve this section's main objective we rely on additional bounds concerning the measure of regions defined in terms of set algebraic manipulation of distinct balls. In the following section we derive these bounds. The reader, however, might prefer to skip their proofs upon first reading, and come back to them once their intended use is understood.

4.1 Useful bounds

The vertices of the path component whose existence we will establish will be shown to be at distance $R-\Omega(1)$ of the origin O. This explains why below we focus on approximating the measure of regions which are close to the boundary.

Our next result establishes two facts. First, it gives an approximation for the measure of a band centered at the origin. Then, it shows that (in expectation) there are $\Omega(1)$ vertices in the intersection of two constant width bands of radius $R-\Theta(1)$, one of which is centered at the origin and the other one centered at a point that is inside the previous band.

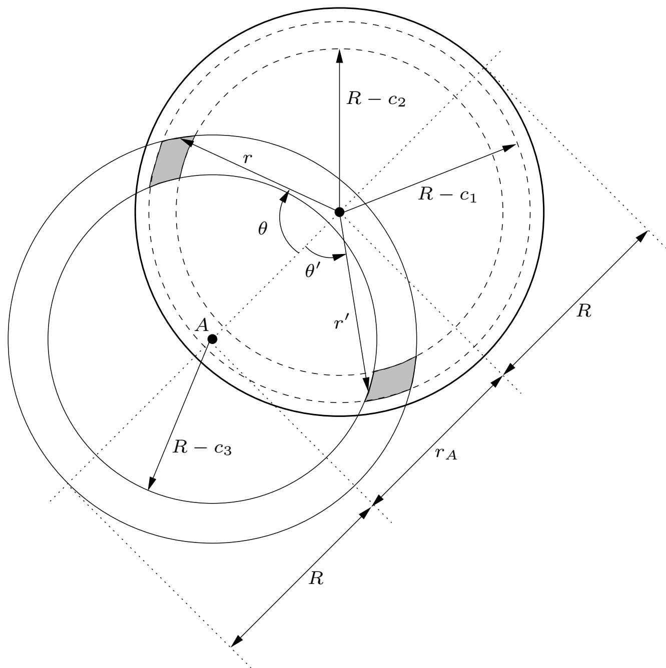

Figure 5: the shaded area corresponds to $\big(B_A(R)\setminus B_A(R-c_3)\big)\cap\big(B_O(R-c_1)\setminus B_O(R-c_2)\big)$ (also, $\theta=\theta_R(r,r_A)$ and $\theta'=\theta_{R-c_3}(r',r_A)$)

Lemma 14. *The following statements hold:*

1. If $0\le\rho'_O\le\rho_O<R$*, then*

$$\mu(B_O(\rho_O)\setminus B_O(\rho'_O))=e^{-\alpha(R-\rho_O)}(1-e^{-\alpha(\rho_O-\rho'_O)}+o(1)).$$

2. If $0<c_1<c_2$ *are two positive constants,* $R-c_2\le r_A\le R-c_1$*, and* $c_3=\Omega(1)$*, then*

$$\mu((B_A(R)\setminus B_A(R-c_3))\cap(B_O(R-c_1)\setminus B_O(R-c_2)))=\Omega(e^{-R/2}).$$

Proof. The first part of the lemma follows directly from Lemma 2. For the second part (see Figure 5), again relying on Lemma 1 we get that the second expression equals

$$2\int_{R-c_2}^{R-c_1}\int_{\theta_{R-c_3}(r,r_A)}^{\theta_R(r,r_A)}\frac{f(r)}{2\pi}d\theta dr=2\frac{e^{(R-r_A)/2}(1-e^{-c_3/2})}{\pi C(\alpha,R)}\times\int_{R-c_2}^{R-c_1}e^{-r/2}\alpha\sinh(\alpha r)\Big(1+O(e^{-r})\Big)dr$$
$$=\frac{\Omega(1)}{C(\alpha,R)}\Big(e^{(\alpha-\frac{1}{2})R}+O(e^{(\alpha-3/2)R})\Big).$$

The desired conclusion follows recalling that $C(\alpha,R) = \cosh(\alpha R) - 1 = \frac{1}{2}e^{\alpha R} + O(e^{-\alpha R})$. □

The following result implicitly establishes that provided c_1, c_2, and c_3 are appropriately chosen constants, if A is a vertex such that $R - c_2 < r_A < R - c_1$, then in expectation A has $\Omega(1)$ neighbors at distance at least $R - c_3$ in the band $B_O(R - c_1) \setminus B_O(R - c_2)$ both in the clockwise and anticlockwise direction.

Lemma 15. *Let $0 < c_3 < c_1 < c_2$ be small positive constants so that $2e^{c_1-c_2} > e^{c_3/2}$ holds. Suppose $R - c_2 \le r_A, r_B \le R - c_1$ and $R - c_3 \le d(A,B) \le R$. Then*

$$\mu\big([(B_B(R) \setminus B_B(R-c_3)) \cap (B_O(R-c_1) \setminus B_O(R-c_2))] \setminus B_A(R)\big) = \Omega(e^{-R/2}) = \Omega(1/n).$$

Proof. By Lemma 14, Part 2, we know that $B_B(R) \setminus B_B(R - c_3)$ intersects the band at distance between $R-c_2$ and $R-c_1$, i.e. $B_O(R-c_1) \setminus B_O(R-c_2)$, in a region of measure $\Omega(e^{-R/2})$. Note that the intersection comprises two disconnected regions of equal measure, say $\mathcal{D}$ and $\mathcal{D}'$. We may assume that $A \in \mathcal{D}$ and also that for all points $P \in \mathcal{D}$ and $P' \in \mathcal{D}'$ we have $\theta_P < \theta_B < \theta_{P'}$. We will show that $B_A(R)$ does not intersect $\mathcal{D}'$: suppose for contradiction that $P \in \mathcal{D}$ and $P' \in \mathcal{D}'$ are within distance R: then, by Remark 1, we would have

$$|\theta_P - \theta_{P'}| \le (2 + o(1))e^{\frac{1}{2}(R-2(R-c_2))} = (2 + o(1))e^{-R/2}e^{c_2}. \tag{11}$$

On the other hand, for any $P \in \mathcal{D}$, since $d(P,B) \ge R - c_3 > R - c_1 > R - c_2$, we have again by Remark 1, $|\theta_B - \theta_P| \ge (2 + o(1))e^{\frac{1}{2}(R-c_3-2(R-c_1))} = (2 + o(1))e^{-R/2}e^{c_1-\frac{c_3}{2}}$, and the same bound holds for $|\theta_B - \theta_{P'}|$. Since P and P' satisfy $\theta_P < \theta_B < \theta_{P'}$, we have $|\theta_P - \theta_{P'}| = |\theta_B - \theta_P| + |\theta_B - \theta_{P'}|$, and thus $|\theta_P - \theta_{P'}| \ge (4 + o(1))e^{-R/2}e^{c_1-\frac{c_3}{2}}$. Since by assumption $2e^{c_1-\frac{c_3}{2}} > e^{c_2}$, this contradicts (11). The lemma follows. □

4.2 The existence of a path component of length $\Theta(\log n)$

The main objective of this section is to show that inside a band at constant radial distance from the boundary we find $\Theta(\log n)$ vertices forming a path. Although the calculations involved require careful bookkeeping, at a high level the proof strategy is not complicated. We now informally describe it. We fix a band centered at the origin, say $B_O(R - c_1) \setminus B_O(R - c_2)$, where c_1 and c_2 are constants. Roughly, we show there are $m = o(\sqrt{n})$ vertices $Q_1, \ldots, Q_m$ in the aforementioned band that satisfy the following two properties: (1) the Q_i's are spread out throughout the band, i.e., each pair subtends an angle at the origin that is "sufficiently" large, and (2) with "not too small" probability each Q_i is the endvertex of a path of length $\Theta(\log n)$ which is completely contained in the mentioned band. The former property implies that the events considered in the latter are independent, from where it easily follows that w.e.p. there must be a path of length $\Theta(\log n)$ inside the band. However, it will become clear next that there are subtle issues that must me dealt with carefully in order to formalize this paragraph's discussion.

Theorem 16. *A.a.s., there exists a component forming a path of length $\Theta(\log n)$.*

Proof. As before, we work first in the Poissonized model and derive in the end from it the result in the uniform model. Fix throughout the proof c_1, c_2, c_3 three positive constants such that $c_3 < c_1 < c_2$ and $2e^{c_1-c_2} < e^{-c_3/2}$. First, by Lemma 14, Part 1, applied with $\rho_0 = R - c_1$ and $\rho'_0 = R - c_2$ we have

$$\mu(B_O(\rho_O) \setminus B_O(\rho'_O)) = e^{-\alpha c_1}(1 - e^{-\alpha(c_2-c_1)} + o(1)) = \Theta(1).$$

Let $\Theta \subseteq [0, 2\pi)$ be a set of forbidden angles such that $\mu(R_\Theta) < 1$, where $R_\Theta := \{(r_P, \theta_P) : 0 \le r_P < R, \theta_P \in \Theta\}$ (for a geometric interpretation of R_Θ, note that when Θ is an interval, R_Θ is a cone with vertex O). As a constant fraction of the angles is still allowed, clearly,

$$\mu\big(((B_O(\rho_O) \setminus B_O(\rho'_O)) \setminus R_\Theta\big) = \Theta(1) \tag{12}$$

still holds. For any vertex A with $R - c_2 \le r_A \le R - c_1$, by Lemma 3 (applied with $\rho_A = R$ and $\rho_O = R$)

$$\mu(B_A(R) \cap B_O(R)) = \frac{2\alpha}{\pi(\alpha - \frac{1}{2})}e^{-\frac{1}{2}r_A} + O(e^{-\alpha r_A}) = \Theta(1/n), \tag{13}$$

which together with Lemma 14, Part 2, yields

$$\mu\big(\big(B_A(R)\setminus B_A(R-c_3)\big)\cap\big(B_O(\rho_O)\setminus B_O(\rho_O')\big)\big)=\Theta(e^{-R/2})=\Theta(1/n). \tag{14}$$

For two vertices A, B with $R-c_3\le d(A,B)\le R$ that satisfy $\rho_O'\le r_A, r_B\le\rho_O$, by Lemma 15, together with (13),

$$\mu\big([(B_B(R)\setminus B_B(R-c_3))\cap(B_O(\rho_O)\setminus B_O(\rho_O'))]\setminus B_A(R)\big)=\Theta(e^{-R/2})=\Theta(1/n). \tag{15}$$

Let $\varepsilon=\varepsilon(\alpha)$ be a constant chosen small enough so that $1-\frac{1}{2\alpha}-\varepsilon>0$. Let $\nu>0$ be a sufficiently small constant. Let $m:=\nu n^{1-\frac{1}{2\alpha}-\varepsilon}$, and note that since $\alpha<1$, we have $m=o(n^{1/2})$. We will add in the following, if necessary, up to m vertices $Q_1,\dots,Q_m\in\mathcal{Q}$ to $\mathcal{P}$, all of them chosen independently and following the same distribution as in $G_{\alpha,C}(n)$. In order to be more precise, define for $1\le j\le m$, the event $\mathcal{E}_j$ that occurs when following conditions hold (if one condition fails, then stop exposing and checking further conditions and proceed with the next j; also stop if all conditions hold for one j):

- **Condition 1:** Add the j-th vertex $Q_j\in\mathcal{Q}$ and let $A_0^j:=Q_j$. We require $\rho_O'\le r_{Q_j}\le\rho_O$, and for $j\ge 2$, we additionally require the coordinates of A_0^j and A_0^k for any $1\le k<j$ to be sufficiently different in terms of their angles, i.e., letting $\Theta_k=\{\theta:|\theta_{A_0^k}-\theta|\le C''n^{-(1-\frac{1}{2\alpha}-\varepsilon)}\}$ for some large constant C', we require that $\theta_{A_0^j}\notin\cup_{k=1}^{j-1}\Theta_k$.
- **Condition 2:** Expose the region $\big(B_{A_0^j}(R)\setminus B_{A_0^j}(R-c_3)\big)\cap\big(B_O(\rho_O)\setminus B_O(\rho_O')\big)$. We require exactly one vertex in this region, call it A_1^j. Let $L:=\nu'\log n$ with $\nu'>0$ being a small constant. Then, inductively, for $1\le i<L$, expose the region $\big(B_{A_i^j}(R)\setminus B_{A_i^j}(R-c_3)\big)\cap\big(B_O(\rho_O)\setminus B_O(\rho_O')\big)\setminus B_{A_{i-1}^j}(R)$. We require also exactly one vertex in this region, and name it inductively A_{i+1}^j. In other words, A_{i+1}^j belongs to the $(R-c_3,R)$-band centered at A_i^j and also to the (ρ_O',ρ_O)-band centered at the origin O. For $i=L$, expose the region $\big(B_{A_L^j}(R)\setminus B_{A_L^j}(R-c_3)\big)\cap\big(B_O(\rho_O)\setminus B_O(\rho_O')\big)\setminus B_{A_{L-1}^j}(R)$, and we require that there is no more vertex in this region.
- **Condition 3:** For any $0\le i\le L$, expose $B_{A_i^j}(R)\cap(B_O(R)\setminus(B_O(R(1-\frac{1}{2\alpha}-\varepsilon)))$, and we require that A_i^j has no other neighbor inside this region except the one(s) from the previous condition.

We will now bound from above the probability that for all $1\le j\le m$ the events $\mathcal{E}_j$ fail. Note first that for ν sufficiently small, the set of forbidden angles $\Theta^j:=\cup_{k=1}^{j-1}\Theta_k$ still is such that $\mu(R_{\Theta^j})<1$, and (12) can be applied with $R_\Theta=R_{\Theta^j}$ for any $1\le j\le m$. Hence, for any j, independently of the outcomes of previous events, there is an absolute constant $c>0$ such that for this j the probability that Condition 1 holds is at least c. Given that Condition 1 holds, by (14) applied to A_0^j and (15) applied successively to $A_1^j,\dots,A_L^j$ we get that the probability that Condition 2 holds is at least c'^{-L} for some fixed $0<c'<1$. Suppose then that Condition 2 also holds. By a union bound and by (13), we have

$$\mu\Big(\bigcup_{i=0}^{L}[B_{A_i^j}(R)\cap(B_O(R)\setminus B_O(R(1-\tfrac{1}{2\alpha}-\varepsilon))]\Big)\le\sum_{i=0}^{L}\mu(B_{A_i^j}(R)\cap B_O(R))=O(L/n),$$

and thus for ν' sufficiently small the probability that Condition 3 holds is at least $n^{-\eta}$ for some fixed value η which can be made small by choosing ν' small (in fact, part of the region might have been already exposed in Condition 2, but since we know there are no other vertices in there, this only helps). Altogether, we obtain

$$\mathbb{P}(\mathcal{E}_j)\ge n^{-\eta'}$$

for some $\eta'>0$. Again, η' can be made sufficiently small by making the constant ν' (and thus $L=\nu'\log n$) sufficiently small.

Now, since $d(A_{i-1}^j, A_i^j) \le R$ and $r_{A_{i-1}^j}, r_{A_i^j} \ge R - c_2$, by Remark 1, we have $|\theta_{A_{i-1}^j} - \theta_{A_i^j}| = O(e^{-R/2})$, and therefore $|\theta_{A_0^j} - \theta_{A_L^j}| = O(\log n/n)$. Since by construction for $j \neq j'$ we have $|\theta_{A_0^j} - \theta_{A_0^{j'}}| = \omega(\log n/n)$, for $j \neq j'$ the regions exposed in Condition 2 are disjoint. Also, for any $j \neq j'$, if the region in Condition 3 containing one of the vertices A_i^j or $A_{i'}^{j'}$ is exposed, then $B_{A_i^j}(R) \cap (B_O(R) \setminus (B_O(R(1-\frac{1}{2\alpha}-\varepsilon)))$ is disjoint from $B_{A_{i'}^{j'}}(R) \cap (B_O(R) \setminus (B_O(R(1-\frac{1}{2\alpha}-\varepsilon)))$: indeed, note that for any point $P \in B_{A_i^j}(R) \cap (B_O(R) \setminus (B_O(R(1-\frac{1}{2\alpha}-\varepsilon)))$ we have by Remark 1

$$|\theta_{A_i^j} - \theta_P| \le (2+o(1))e^{\frac{1}{2}(R-(R-c_2)-R(1-\frac{1}{2\alpha}-\varepsilon))} = (2+o(1))e^{\frac{1}{2}c_2} n^{-(1-\frac{1}{2\alpha}-\varepsilon)}.$$

By construction, we have $|\theta_{A_0^j} - \theta_{A_0^{j'}}| \ge C' n^{-(1-\frac{1}{2\alpha}-\varepsilon)}$ for some large enough $C' > 0$, and thus by the triangle inequality for angles there exists $C'' := C''(C') > 0$ such that $|\theta_{A_i^j} - \theta_{A_{i'}^{j'}}| \ge (1+o(1))C'' n^{-(1-\frac{1}{2\alpha}-\varepsilon)}$ holds for any i, i' and any $j \neq j'$. Hence, the regions exposed in Condition 3 are disjoint, and by the same reason, the region exposed in Condition 2 for some j and the one exposed in Condition 3 for some $j' \neq j$ are disjoint as well. It follows that the probabilities of the corresponding conditions to hold are thus independent. Hence, for ν' sufficiently small, η' is small enough such that $-\eta' + 1 - \frac{1}{2\alpha} - \varepsilon > 0$, and

$$\prod_{j=1}^{m} \mathbb{P}\big(\mathcal{E}_j{}^c\big) \le (1 - n^{-\eta'})^m = e^{-\Omega(n^{\xi})},$$

for some positive $\xi > 0$. Thus, w.e.p. there exists one j, for which the event $\mathcal{E}_j$ holds.

In order to de-Poissonize, let $\mathcal{E}_P^j$ be the event that in the Poissonized model there exists some j for which the event $\mathcal{E}_j$ holds, and similarly $\mathcal{E}_U^j$ the corresponding event for the uniform model. Since we have shown that $\mathbb{P}(\mathcal{E}_P^j) = 1 - e^{-\Omega(n^{\xi})}$, and since $m = o(n^{-1/2})$, using (6), we also have for some function ω_n tending to infinity arbitrarily slowly $\mathbb{P}(\mathcal{E}_U^j) \ge 1 - \omega_n \sqrt{n} e^{-\Omega(n^{\xi})} = 1 - e^{-\Omega(n^{\xi})}$. Let $\rho_O = R(1 - \frac{1}{2\alpha} - \varepsilon)$. Denote by N_U the random variable counting the number of vertices inside $B_O(\rho_O)$ in $G_{\alpha,C}(n)$. By Lemma 2, $\mu(B_O(\rho_O)) = O(e^{-R/2-\alpha\varepsilon R}) = o(1/n)$, and thus $\mathbb{E}N_U = o(1)$, and by Markov's inequality, $\mathbb{P}(N_U = 0) = 1 - o(1)$. Since

$$1 - o(1) = \mathbb{P}(N_U = 0) = \mathbb{P}(N_U = 0|\mathcal{E}_U^j)\mathbb{P}(\mathcal{E}_U^j) + \mathbb{P}(N_U = 0|\mathcal{E}_U^{j\,c})\mathbb{P}(\mathcal{E}_U^{j\,c}) = (1 - o(1))\mathbb{P}(N_U = 0|\mathcal{E}_U^j) + o(1),$$

we have $\mathbb{P}(N_U = 0|\mathcal{E}_U^j) = 1 - o(1)$. Thus, $\mathbb{P}(N_U = 0, \mathcal{E}_U^j) = 1 - o(1)$. If this event holds, this means that there is no vertex inside $B_O(\rho_O)$, and the vertices $A_0^j, \dots, A_L^j$ form an induced path, yielding the desired result. ☐

5 Conclusion

We have shown that in random hyperbolic graphs a.a.s. the diameter of the giant component is $O(polylog(n))$, the size of the second largest component $O(polylog(n))$, and at the same time there exists a path component of length $\Theta(\log n)$. It is an interesting and challenging problem to tighten these bounds by improving the exponents of the $polylog(n)$ terms established in this work.

Acknowledgement

We thank Tobias Müller for generously answering questions concerning the random hyperbolic graph model, specifically about the state-of-art as well as pointers to the relevant literature concerning the model.

References

[AS08] N. Alon and J. Spencer. *The Probabilistic Method, Third Edition.* John Wiley & Sons, 2008.

[BFM] M. Bode, N. Fountoulakis, and T. Müller. The probability that the hyperbolic random graph is connected. Preprint available at `www.math.uu.nl/~Muell001/Papers/BFM.pdf`.

[BFM13] M. Bode, N. Fountoulakis, and T. Müller. On the giant component of random hyperbolic graphs. In *Proceedings of the 7th European Conference on Combinatorics, Graph Theory and Applications*, EUROCOMB'13, pages 425–429, 2013.

[BnPK10] M. Boguñá, F. Papadopoulos, and D. Krioukov. Sustaining the internet with hyperbolic mapping. *Nature Communications*, 1:62, 2010.

[CF13] E. Candellero and N. Fountoulakis. Clustering in random geometric graphs on the hyperbolic plane. *ArXiv e-prints*, September 2013. Preprint available at `arxiv.org/abs/1309.0459`.

[Fou12] N. Fountoulakis. On the evolution of random graphs on spaces with negative curvature. *ArXiv e-prints*, May 2012.

[GPP12] L. Gugelmann, K. Panagiotou, and U. Peter. Random hyperbolic graphs: Degree sequence and clustering. In *Proceedings of the 39th International Colloquium Conference on Automata, Languages, and Programming - Volume Part II*, ICALP'12, pages 573–585. Springer-Verlag, 2012.

[KPK+10] D. Krioukov, F. Papadopoulos, M. Kitsak, A. Vahdat, and M. Boguñá. Hyperbolic geometry of complex networks. *Phys. Rev. E*, 82:036106, 2010.

[Pen03] M. Penrose. *Random Geometric Graphs.* Oxford Studies in Probability, Oxford U. P., 2003.

[PKBnV10] F. Papadopoulos, D. Krioukov, M. Boguñá, and A. Vahdat. Greedy forwarding in dynamic scale-free networks embedded in hyperbolic metric spaces. In *Proceedings of the 29th Conference on Information Communications*, INFOCOM'10, pages 2973–2981. IEEE Press, 2010.