Cliques in Hyperbolic Random Graphs

Tobias Friedrich Anton Krohmer
Faculty of Mathematics and Computer Science
Friedrich Schiller University Jena, Germany
{friedrich, anton.krohmer}@uni-jena.de

***Abstract*—Most complex real-world networks display scale-free features. This motivated the study of numerous random graph models with a power-law degree distribution. There is, however, no established and simple model which also has a high clustering of vertices as typically observed in real data. Hyperbolic random graphs bridge this gap. This natural model has recently been introduced by Papadopoulos, Krioukov, Boguñá, Vahdat (INFOCOM, pp. 2973–2981, 2010) and has shown theoretically and empirically to fulfill all typical properties of real-world networks, including power-law degree distribution and high clustering.**

We study cliques in hyperbolic random graphs G and present new results on the expected number of k-cliques $\mathbb{E}[K_k]$ and the size of the largest clique $\omega(G)$. We observe that there is a phase transition at power-law exponent $\gamma = 3$. More precisely, for $\gamma \in (2, 3)$ we prove $\mathbb{E}[K_k] = n^{k(3-\gamma)/2}\Theta(k)^{-k}$ and $\omega(G) = \Theta(n^{(3-\gamma)/2})$ while for $\gamma \geqslant 3$ we prove $\mathbb{E}[K_k] = n\,\Theta(k)^{-k}$ and $\omega(G) = \Theta(\log(n)/\log\log n)$.

We empirically compare the $\omega(G)$ value of several scale-free random graph models with real-world networks. Our experiments show that the $\omega(G)$-predictions by hyperbolic random graphs are much closer to the data than other scale-free random graph models.

I. Introduction

Scale-free networks are ubiquitous in nature and society. They appear as a large array of real-world graphs that (mostly) have been formed by autonomous agents. Popular examples include social networks, protein-protein interactions, sexual networks, electricity circuits, the WWW, the internet, and many more [17]. Even though the term "scale-free network" has never been well-defined [15], it has been observed that all of these graphs share similar characteristics. They have hub nodes (nodes that interconnect the graph), community structures (subgraphs with high edge density), very low diameter (longest shortest path), a giant component (a connected component containing almost all vertices) and—probably most importantly—their degree distribution follows a power law: $P(k) \sim k^{-\gamma}$, where $P(k)$ is the fraction of nodes having degree k.

Over the course of the last decade, research has been striving to produce generative models for these types of networks that are able to accurately predict properties of real-world graphs. Popular models include the preferential attachment graphs [3] and variants of inhomogeneous random graphs [24]. The latter generalizes the Erdős-Rényi random graphs $G_{n,p}$ by using non-uniform edge probabilities. These models excel at modeling the power-law degree distribution; and they have a giant component, hub nodes and low diameter. Due to their independent edge probabilities, they are accessible to rigorous studies. Independent edge probabilities also imply, however, that the graphs have low clustering, meaning that there exist no community structures.

In contrast, most real-world graphs do have high clustering. In the case of social networks this is easy to envision: Two people are much more likely to be connected if they already have a friend in common. A number of fixes to the above models have been proposed to incorporate that intuition [16, 18, 25] (e.g. first construct a random graph, and then replace all nodes with k-cliques). Often times, however, these fixes seem artificial and can not convincingly explain why clustering occurs.

Krioukov et al. [13] took a different approach by assuming an underlying hyperbolic geometry to the network. Similarly to the well-known geometric random graphs in Euclidean space [21], they introduced *hyperbolic random graphs* in which all nodes are placed in the hyperbolic plane, and two nodes are connected whenever they have a small distance from each other. The clustering then follows from the geometric interpretation: When two nodes are close to a third node, it is likely that they are also close to each other. In fact, [10, 13] showed that the clustering coefficient in these networks is constant, and community structures emerge as natural reflections of the hyperbolic geometry. Furthermore, the hyperbolic geometry enforces a power-law degree distribution, and the presence of hub nodes.

The model achieved remarkable results for greedy forwarding: Embedding the internet graph in such a hyperbolic plane, an autonomous system can forward packets using only the hyperbolic location of the destination and its own neighbors [4, 20, 23]. Doing away with the currently used routing tables, they achieved an

978-1-4799-8381-0/15/$31.00 ©2015 IEEE

		Hyperbolic Random Graphs	
		$\frac{1}{2} < \alpha < 1$	$\alpha \geqslant 1$
$\mathbb{E}[K_k]$	$\leqslant$	$\frac{n^{(1-\alpha)k}}{k^k \exp(k(\alpha\frac{C}{2}+1))}\left(\frac{\alpha k c^{k-1}}{(1-\alpha)k+1}+1\right)$	$nk^{-k} \cdot \frac{\alpha k e (c e^{-C/2+1})^{k-1}}{(\alpha-1)k+1} \cdot (1+o(1))$
	$\geqslant$	$\left(\frac{e^{-\alpha C/2} n^{1-\alpha}}{k}\right)^k$	$\frac{n}{k}\left(\frac{1}{\sqrt{2e^C}\pi k}(1 \pm o(1))\right)^{k-1}$
	$=$	$n^{(1-\alpha)k}\Theta(k)^{-k}$	$n \cdot \Theta(k)^{-k}$
$\omega(G)$	$\leqslant$	$(1+\varepsilon)ce^{-\alpha\frac{C}{2}+1}n^{1-\alpha}$	$(1+\varepsilon \pm o(1))\frac{\log n}{\log\log n}$
	$\geqslant$	$e^{-\alpha C/2}n^{1-\alpha}(1-o(1))$	$(1-o(1))\frac{\log n}{\log\log n}$
	$=$	$\Theta(n^{1-\alpha})$	$\Theta\left(\frac{\log n}{\log\log n}\right)$

Table I: New results on the expected number of k-cliques $\mathbb{E}[K_k]$ and the size of the largest clique $\omega(G)$ in hyperbolic random graphs. Section II defines the parameters used in this table. Sections III and IV prove the upper and lower bounds on $\mathbb{E}[K_k]$, respectively. Section V proves the bounds on $\omega(G)$.

almost optimal routing that is just 10% slower on average than the optimal routing path. This result suggests that there is an underlying hyperbolic metric to at least the internet graph; and the hyperbolic geometry might be what unites most of the scale-free networks.

While the intuition behind hyperbolic random graphs is elegant, rigorous treatments can quickly become mathematically challenging. Since most of existing research has focused on the power-law exponent, average degree and clustering coefficient [8, 10], there are currently still a number of important questions open regarding this model, e.g. rigorous proofs for the existence of a giant component or for the low diameter.

Our contribution. Closely related to clustering and community structures, we analyze the emergence of cliques in hyperbolic random graphs. In particular, we present bounds on the expected number of k-cliques and the size of the largest clique. The results are summarized in Table I. We observe that there is a phase transition at power-law exponent $\gamma = 3$, with smaller exponents yielding polynomial-size cliques and larger exponents yielding logarithmic-size cliques. While clique is NP- and W[1]-complete for general graphs, we argue that the largest clique of hyperbolic random graphs can be found in polynomial time. Finally, we validate our results experimentally by observing that the hyperbolic random graph provides better predictions than several other models on the size of the largest cliques of real-world graphs.

Comparison with other scale-free models. Using the results of Janson et al. [11], we compare the *clique numbers*, i.e. the size of the largest clique, of some popular scale-free network models to the hyperbolic random graph in Table II.

We notice that the (asymptotic) clique number is almost the same for Chung-Lu [1], Norros-Reittu [19] and hyperbolic random graphs in the case where the power law exponent is $2 < \gamma < 3$. An intuitive explanation for this phenomenon is that all these models have a tightly connected *core*: A subgraph of polynomial size in which the edge probability is $1 - o(1)$ or even 1. Large cliques emerge as a consequence of this core.

But even when such a core does not exist in the graph (which is the case for $\gamma \geqslant 3$), one would expect to have small communities and therefore cliques in the graph. In particular, due to the large clustering coefficient it is likely that a node's neighbors (or a subset of the neighbors) form a clique. Consequently, the hyperbolic random graph has approximately $\frac{\log n}{\log\log n}$ as the largest clique. Existing models with independent edge probabilities in this case predict a largest clique of size $\leqslant 3$ with high probability (i.e., with probability $\geqslant 1 - o(1)$), which seems unlikely. Our experimental findings in the next section confirm this intuition.

Experiments. We conducted preliminary experiments to validate our theoretical results. We used all scale-free networks from [27] that have $n \geqslant 50\,000$ and $\gamma \geqslant 2$. The power-law exponents γ were estimated by a maximum likelihood estimation [2, 6] which minimizes the Kolmogorov-Smirnov (K–S) error. We discarded all networks with K–S error $D > 0.02$ and where the clique finding algorithm by Eppstein and Strash [7] did not terminate within one day (LiveJournal, Friendster).

To determine the quality of predictions by the various models in Table II, we compare for each real graph $\widetilde{G}$ the true clique number $\omega(\widetilde{G})$ for the real graphs to the clique numbers $\omega(\widetilde{G}')$ that the various models achieve with high probability. We chose the adjustable model parameters (e.g. power-law exponent, average degree) to match the respective values from the real graph $\widetilde{G}$. The results

	Power-Law Exponent		
Random Graph Model	$2<\gamma<3$	$\gamma=3$	$\gamma>3$
Hyperbolic (new results)	$\Theta(n^{(3-\gamma)/2})$	$\Theta\big(\frac{\log n}{\log\log n}\big)$	$\Theta\big(\frac{\log n}{\log\log n}\big)$
Chung-Lu	$\Theta(n^{(3-\gamma)/2})$	$\Theta(1)$	3
Norros-Reittu	$\Theta(n^{(3-\gamma)/2}\log^{-(\gamma-1)/2} n)$	$\Theta(1)$	3
Generalized RG	$\Omega(n^{\frac{3-\gamma}{1+\gamma}})$, $\mathcal{O}(n^{\frac{3-\gamma}{1+\gamma}}\log^{\frac{\gamma-1}{\gamma+1}} n)$	$\Theta(1)$	3
Pref. Attachment	$\Theta(1)$	$\Theta(1)$	$\Theta(1)$

Table II: Comparison of our new results on the clique number $\omega(G)$ of hyperbolic random graphs to known results by Janson et al. [11] for other scale-free random graph models. All bounds hold with high probability.

are plotted as the multiplicative error $\omega(\widetilde{G}')/\omega(\widetilde{G})$ in Figure 1.

In all cases, the hyperbolic random graph's predictions are close to the true value, and for five graphs it beats all existing models. Chung-Lu and Norros-Reittu have a multiplicative error of up to two orders of magnitude; and while the Preferential Attachment model seems to be close to reality in some cases, its clique number can never exceed $|E|/|V|$, giving it limited prediction value.

Especially for networks with power-law exponent $\gamma \geqslant 3$—where the clique number of the hyperbolic random graph differs strongly from previous results—its prediction is surprisingly accurate; the other models (unrealistically) report a constant value for $\omega(G)$. These experiments indicate that the hyperbolic random graph is a further step in the right direction in the analysis and understanding of complex real-world networks.

II. Preliminaries

The hyperbolic random graph is constructed as follows. Let $\mathbb{H}^2$ be the hyperbolic plane with negative curvature $K=-\zeta$. For simplicity, we assume in this paper that $\zeta=1$, but our results easily translate to arbitrary $\zeta>0$. To obtain a graph G with n nodes, let D_n be a disc in $\mathbb{H}^2$ of radius $R=2\ln n+C$, where C adjusts the average degree of G. The disc is centered in the point of origin. Afterwards, n points are sampled in D_n as follows. Let $\alpha>0$ be some constant. The probability density for the radial coordinate r of a point $p=(r,\phi)$ is given by

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

and the angular coordinate ϕ is sampled uniformly from $[0,2\pi]$. We write r_u and ϕ_u to refer to the polar coordinates of a point u.

In the most general model, the probability that two nodes u,v with relative angle $\Delta\theta$ connect is $(\exp(\frac{\beta}{2}(\mathrm{d}(u,v)-R))+1)^{-1}$, where

$$\cosh(\mathrm{d}(u,v)) := \cosh r_u \cosh r_v - \sinh r_u \sinh r_v \cos\Delta\theta$$

defines the distance between two points in $\mathbb{H}^2$. This produces a power-law graph with exponent $\gamma=2\alpha+1$ if $\alpha\geqslant\frac{1}{2}$, and $\gamma=2$ else. We assume that $\alpha>\frac{1}{2}$, i.e. $\gamma>2$. Observe that when $\beta\to\infty$, $p(\cdot)$ becomes a step function that connects two nodes if they have distance at most R from each other. We call this case the *step model*; and the case $\beta<\infty$ the *binomial model*. In this paper, we focus on the step model. If not explicitly stated otherwise, we refer with G to a random graph that was sampled according to the step model.

Note that when $\alpha=1$, the points are uniformly sampled from D_n, and a power-law graph with exponent $\gamma=3$ arises. Gugelmann et al. [10] showed that the average degree in the step model is $\delta=(1+o(1))\frac{2\alpha^2e^{-C/2}}{\pi(\alpha-1/2)^2}$.

For most computations on hyperbolic random graphs, one needs close approximations of the probability that a sampled point falls in a certain area. To this end, Gugelmann et al. [10] define the probability measure of a set $S\subseteq D_n$ as

$$\mu(S) := \int_S f(y)\,\mathrm{d}y,$$

where $f(r)$ is the probability mass of a point $p=(r,\theta)$ given by $f(r):=\frac{\alpha\sinh(\alpha r)}{2\pi(\cosh(\alpha R)-1)}$. In other words, given a point $p=(r,\theta)$ that was sampled as described above,

$$\mu(S) = \Pr[p\in S\colon p=(r,\theta)].$$

We define the *ball* with radius x around a certain point (r,θ) as

$$B_{r,\theta}(x) := \{(r',\theta') \mid \mathrm{d}((r',\theta'),(r,\theta)) \leqslant x\}.$$

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

n	m	γ	$\omega(G)$	Network
62 586	147 892	4.80	4	**Gnutella31**
3 774 768	16 518 948	4.02	11	**cit-Patents**
334 863	925 872	3.58	7	**com-Amazon**
403 394	3 387 388	3.15	11	**Amazon0601**
196 591	950 327	2.82	29	**loc-Gowalla**
875 713	5 105 039	2.72	44	**web-Google**
3 072 441	117 185 083	2.66	51	**com-Orkut**
1 134 890	2 987 624	2.53	17	**com-Youtube**
1 696 415	11 095 298	2.34	67	**as-Skitter**

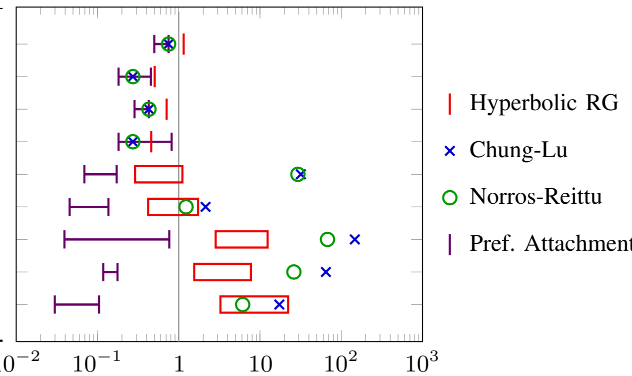

Figure 1: Predictions of the clique number by various models. The plot shows the multiplicative error of the estimations, i.e. a value of 1 is a perfect prediction. Exact bounds (up to a factor of $(1 \pm o(1))$ are known for Chung-Lu and Norros-Reittu, so they are represented by a point [11]. For Preferential Attachment Graphs we use $\omega(G) \in [2, \lceil \frac{\delta}{2} \rceil]$, where δ is the average degree. For the hyperbolic random graph, the box represents our upper and lower bound. The predictions of the Generalized Random Graph [11] are not shown, as they contain unspecified constants. The predictions of the hyperbolic random graphs are typically closest to the behavior observed in the nine real networks.

Lemma II.1. *For any $0 \leqslant r \leqslant R$ we have*

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

$$\mu(B_r(R) \cap B_0(R)) = \frac{2\alpha e^{-r/2}}{\pi(\alpha - 1/2)} \cdot \mathcal{E}_1 \tag{2}$$

$$\mu(B_r(R) \cap B_0(r)) = \tag{3}$$
$$\begin{cases} \frac{2\alpha}{\pi(\alpha-\frac{1}{2})} e^{-R(\alpha-1/2)+r(\alpha-1)} \cdot \mathcal{E}_2 & \text{if } r > R/2, \\ \mu(B_0(r)) & \text{if } r \leqslant R/2, \end{cases}$$

with error terms $\mathcal{E}_1 = 1 \pm \mathcal{O}(e^{-(\alpha-1/2)r} + e^{-r})$ and $\mathcal{E}_2 = 1 \pm \Theta(e^{(R-2r)(\alpha-\frac{1}{2})})$ if $\alpha \neq \frac{3}{2}$ and $\mathcal{E}_2 = 1 \pm \Theta(e^{R-2r}(2r - R))$ otherwise.

For the proof of (1) and (2) we refer to [10]. The proof of (3) is analogous and rather technical, so we postpone it to the appendix.

Using Lemma II.1, we compute an upper bound on the expected number of k-cliques in the hyperbolic random graph.

III. Proof of the Upper Bound

In a clique, each pair of nodes is connected. To compute an upper bound on the probability that k nodes form a clique, we examine a relaxed condition; namely that all nodes connect to one specific node v.

For a set U of k independently sampled points, let $v \in U$ be the node with $r_v = \max_{u \in U}\{r_u\}$. We begin by computing the probability density function of r_v which we call $\rho_v(r)$. By the definition of the cumulative distribution function, we have

$$\begin{aligned} \Pr[r_v \leqslant x] &= \Pr[\forall u \in U \colon r_u \leqslant x] \\ &= \left(\int_0^x \frac{\alpha \sinh(\alpha r)}{\cosh(\alpha R) - 1} \, \mathrm{d}r \right)^k \\ &= \left(\frac{\cosh(\alpha x) - 1}{\cosh(\alpha R) - 1} \right)^k . \end{aligned}$$

The resulting probability density function is given by

$$\begin{aligned} \rho_v(r) &= \frac{\partial}{\partial r} \left(\frac{\cosh(\alpha r) - 1}{\cosh(\alpha R) - 1} \right)^k . \\ &= \alpha k \sinh(\alpha r) \frac{(\cosh(\alpha r) - 1)^{k-1}}{(\cosh(\alpha R) - 1)^k} . \end{aligned}$$

Following the explanation above, we know that the probability that a set U of k independently sampled nodes forms a clique is at most the probability that all nodes are connected to v. Formally,

$$\begin{aligned} &\Pr[U \text{ is clique}] \leqslant \Pr[\forall u \in U \colon \mathrm{d}(u,v) \leqslant R] \\ &= \int_0^R \rho_v(r) \cdot \Pr[\forall u \in U \colon u \in B_r(R) \mid r_v = r] \\ &= \int_0^R \rho_v(r) \cdot \left(\frac{\mu(B_r(R) \cap B_0(r))}{\mu(B_0(r))} \right)^{k-1} \mathrm{d}r \end{aligned}$$

For the last equality, observe that we condition on the fact that all radial coordinates are $\leqslant r$. Hence, the probability that a uniformly sampled node u is connected to v is the probability that $u \in B_r(R)$, conditioned on the fact that $r_u \leqslant r$, i.e. $u \in B_0(r)$. We split the integral in two parts. If $r < R/2$, then by triangle inequality it follows that all k nodes are connected. This agrees with Lemma II.1, and we obtain

$$\int_0^{R/2} \rho_v(r) \cdot \left(\frac{\mu(B_r(R) \cap B_0(r))}{\mu(B_0(r))}\right)^{k-1} \mathrm{d}r \quad (4)$$
$$= \int_0^{R/2} \rho_v(r)\,\mathrm{d}r \leqslant \left(\frac{\cosh(\alpha R/2)-1}{\cosh(\alpha R)-1}\right)^k \leqslant e^{-\alpha k \frac{R}{2}}.$$

For the second part of the integration, we compute

$$\int_{R/2}^{R} \rho_v(r) \cdot \left(\frac{\mu(B_r(R) \cap B_0(r))}{\mu(B_0(r))}\right)^{k-1} \mathrm{d}r =$$
$$\int_{R/2}^{R} \alpha k \sinh(\alpha r) \tfrac{(\cosh(\alpha r)-1)^{k-1}}{(\cosh(\alpha R)-1)^k} \left(\Theta(e^{R/2-r})\right)^{k-1} \mathrm{d}r$$

To obtain a tight bound, observe that

$$(1-x)^k = ((1-x)^{1/x})^{kx} \geqslant (\tfrac{1}{e}(1-x))^{kx} \geqslant e^{-2kx} \quad (5)$$

whenever $1-x \geqslant 1/e$. Note that

$$e^x/2 - 1 \leqslant \cosh(x) - 1 \leqslant \sinh(x) \leqslant e^x/2$$

whenever $x > 0$. Furthermore, we have that $(\frac{1}{2}e^{(\alpha R)} - 1)^k = \frac{1}{2}e^{\alpha R k}(1-2e^{-\alpha R})^k$. By (5) it follows $(1-2e^{-\alpha R})^k \geqslant e^{-4k\exp(-\alpha R)}$. Since $R = 2\log n + C$ and $k \leqslant n$ we obtain that $(1-2e^{-\alpha R})^k \geqslant 1/e$ for large enough n. Then, for some constant c,

$$\int_{R/2}^{R} \rho_v(r) \cdot \left(\frac{\mu(B_r(R) \cap B_0(r))}{\mu(B_0(r))}\right)^{k-1} \mathrm{d}r$$
$$\leqslant \int_{R/2}^{R} \alpha k \frac{1}{e} \frac{e^{\alpha r k}}{e^{\alpha R k}} \left(c e^{R/2-r}\right)^{k-1} \mathrm{d}r \quad (6)$$
$$= \frac{\alpha k c^{k-1}}{(\alpha-1)k+1} \left[e^{k(\alpha(r-R)-r+R/2)+r-R/2}\right]_{R/2}^{R}, \quad (7)$$

where the last equation holds if $\alpha \neq 1$ and $k \neq 1/(1-\alpha)$. To cover all possibilities for α and k, we distinguish the following cases:

(a) $\alpha = 1$. In this case, (6) evaluates to

$$kc^{k-1}[e^{-\frac{R}{2}(k+1)+r}]_{R/2}^{R} \leqslant kc^{k-1}e^{-\frac{R}{2}(k-1)}.$$

(b) $\alpha > 1$. Then,

$$(7) \leqslant \frac{\alpha k c^{k-1}}{(\alpha-1)k+1} e^{-\frac{R}{2}(k-1)}.$$

(c) $\frac{1}{2} \leqslant \alpha < 1$. In this case, we have to pay attention to the sign in front of the antiderivative.

(c.i) $k < \frac{1}{1-\alpha}$. In that case,

$$(7) \leqslant \frac{\alpha k c^{k-1}}{(\alpha-1)k+1} e^{-\frac{R}{2}(k-1)}.$$

(c.ii) $k = \frac{1}{1-\alpha}$. Consider once more (6).

$$(6) = \int_{R/2}^{R} \alpha k c^{k-1} \frac{1}{e} e^{-R(\alpha k - \frac{k}{2} + \frac{1}{2})} \mathrm{d}r$$
$$= \alpha k c^{k-1} \frac{R}{2e} e^{-\frac{R}{2}(k-1)}$$

(c.iii) $k > \frac{1}{1-\alpha}$. Here, the sign of the antiderivative is negative, and we obtain

$$(7) \leqslant \frac{\alpha k c^{k-1}}{(1-\alpha)k+1} e^{-\alpha k \frac{R}{2}}.$$

Taken together with (4), cases (a)–(c.ii) only change by a factor of $(1+o(1))$, and in the case of (c.iii) we obtain $(1 + \frac{\alpha k c^{k-1}}{(1-\alpha)k+1})e^{-\alpha k R/2}$. When $\alpha > 1$ (i.e. when the graph has a power law exponent $\gamma > 3$), the number of cliques is therefore bounded by

$$\begin{aligned}\mathbb{E}[K_k] &= \binom{n}{k} \Pr[U \text{ is clique}] \\ &\leqslant \left(\frac{ne}{k}\right)^k \frac{\alpha k c^{k-1}}{(\alpha-1)k+1} e^{-\frac{R}{2}(k-1)} \\ &= nk^{-k} \cdot \frac{\alpha k e (c e^{-C/2+1})^{k-1}}{(\alpha-1)k+1} \\ &= n \cdot \Theta(k)^{-k}.\end{aligned}$$

For $\alpha = 1$ we obtain a similar bound $\mathbb{E}[K_k] \leqslant n \cdot \Theta(k)^{-k+1} = n \cdot \Theta(k)^{-k}$.

For networks with a dense core ($\frac{1}{2} < \alpha < 1$), we have an exponential number $n^{k(1-\alpha)} \cdot \Theta(k)^{-k}$ of k-cliques, if k is large enough. Table I contains the detailed results for these cases. In the case where $k \leqslant 1/(1-\alpha)$, which is not shown in the table, there is in fact a linear number of k-cliques. This agrees e.g. with the fact that for $k = 2$ there are $\Theta(n)$ many edges in G.

IV. Proof of the Lower Bound

To obtain a matching lower bound, we consider two cases. Let us first investigate a sector of the disc D_n with angle $\theta = a/n$, for some adjustable constant a. Clearly, there are $\frac{2\pi n}{a}$ such non-overlapping circular sectors. For small enough a, such a circular sector has a diameter of $\leqslant R$, as we show in the following. The points inside such a circular sector thereby form a complete subgraph.

Obviously, the probability that a point falls in one specific circular sector of angle a/n is exactly $\frac{a}{2\pi n}$. Therefore, the probability that k independently sampled points all lie in one sector is

$$\Pr[U \text{ is clique}] \geqslant \frac{2\pi n}{a} \cdot \left(\frac{a}{2\pi n}\right)^k = \left(\frac{a}{2\pi n}\right)^{k-1}.$$

This probability is clearly maximized by choosing a as large as possible, i.e. such that for any larger a' the diameter is $> R$. In the following, we derive the largest possible value for a.

A. The Diameter of a Circular Sector

Let u, v be two points inside the circular sector S of angle a/n that have maximal distance. Observe that these points have to lie on the boundary of S: Otherwise, consider the geodesic that goes through u, v and intersects S at u', v'. Clearly, $\mathrm{d}(u', v') > \mathrm{d}(u, v)$, a contradiction.

Consider now that u, v both have maximal radial coordinates $r_u = r_v = R$. As we can observe from the distance formula, the distance between two points in S can be maximized by setting $\Delta\theta$ as large as possible (since S has an angle of $a/n \ll \pi$). It is therefore sufficient to investigate pairs of points that have a maximal relative angle of a/n. Then, their distance is

$$\mathrm{d}(u, v) = \cosh^{-1}(\cosh^2(R) - \sinh^2(R)\cos(\tfrac{a}{n})).$$

A series expansion gives $\cos(x) = 1 - \frac{x^2}{2} + \mathcal{O}(x^4)$, leaving us with

$$\mathrm{d}(u, v) \leqslant \cosh^{-1}(1 + (\tfrac{a}{n})^2 \sinh^2(R)).$$

We now note that by choosing $a = 2\sqrt{a'/e^C}(1 + o(1))$ suitably we get that $(\frac{a}{n})^2 \sinh^2(R) = a'e^R$ from the definition of R. Therefore,

$$\begin{aligned}\mathrm{d}(u, v) &\leqslant \cosh^{-1}(1 + a'e^R) \\ &= \log(1 + a'e^R + \sqrt{2a'e^R + a'^2e^{2R}}) \\ &= \log(e^R \cdot 2a'(1 + o(1))) \\ &\leqslant R + \ln(2a'(1 + o(1))). \qquad (8)\end{aligned}$$

Hence, by choosing $a' < \frac{1}{2}(1 - o(1))$ and thereby $a \leqslant \sqrt{2/e^C}(1 \pm o(1))$ we obtain that two nodes u, v on the circular arc of S have distance at most R.

Finally, we show that in this case, the diameter of S is indeed at most R.

Lemma IV.1. *The set S as defined above has a diameter of at most R.*

The proof can be found in the appendix. To conclude, we notice that

$$\begin{aligned}\Pr[U \text{ is clique}] &\geqslant \left(\frac{a}{2\pi n}\right)^{k-1} \\ &= \left(\frac{1}{\sqrt{2e^C}\pi n}(1 \pm o(1))\right)^{k-1}\end{aligned}$$

implies once again that

$$\mathbb{E}[K_k] = \binom{n}{k} \Pr[U \text{ is } k\text{-clique}] \geqslant n \cdot \Theta(k)^{-k}.$$

B. Polynomial Cliques

We proved that the expected number of k-cliques in the hyperbolic random graph is at least $n \cdot \Theta(k)^{-k}$. This bound, however, does not match the upper bound for the dense case, i.e. when $\frac{1}{2} < \alpha < 1$. In this case, we need a different approach. Consider the ball $B_0(R/2)$. All nodes in this area have distance $\leqslant R$ from each other by the triangle inequality. It is therefore left to bound the number of nodes in $B_0(R/2)$. By Lemma II.1 we know that

$$\mu(B_0(x)) = e^{-\alpha(R-x)}(1 + o(1)),$$

i.e. the probability that a sampled point has at most distance x from the center of D_n is $e^{-\alpha(R-x)}(1+o(1))$. Consequently, we expect $ne^{-\alpha R/2}(1 + o(1))$ nodes in $B_0(R/2)$. Observe that for $\frac{1}{2} < \alpha < 1$ and $R = 2\ln n + C$ this amounts to $e^{-\alpha C/2}n^{1-\alpha}$, which is polynomial.

Let $K_k(G)$ be the number of k-cliques in G. Clearly, if $G' \subseteq G$, then we have that $K_k(G') \leqslant K_k(G)$. Consider for G the hyperbolic random graph and for G' the graph induced on G by only taking vertices v with $r_v \leqslant R/2$. Then, we get

$$\mathbb{E}[K_k] = \mathbb{E}[K_k(G)] \geqslant \mathbb{E}[K_k(G')] = \mathbb{E}\left[\binom{X}{k}\right],$$

where X is the random variable describing the number of nodes that drop in $B_0(R/2)$. To show the lower bound, we use the following well-known lemma, which can e.g. be found in [26, Ex. 1].

Lemma IV.2. *The function $f(x) = \binom{x}{k}$ is convex on $x \geqslant k$.*

Therefore, using Jensen's inequality [12] which says $f(\mathbb{E}[X]) \leqslant \mathbb{E}[f(X)]$ for convex functions f, we obtain

$$\begin{aligned}\mathbb{E}\left[\binom{X}{k}\right] &\geqslant \binom{\mathbb{E}[X]}{k} = \binom{e^{-\alpha C/2}n^{1-\alpha}(1 + o(1))}{k} \\ &\geqslant \left(\frac{e^{-\alpha C/2}n^{1-\alpha}}{k}\right)^k.\end{aligned}$$

Taken together with the result above, we conclude

$$\mathbb{E}[K_k] \geqslant \max\{n, n^{(1-\alpha)k}\} \cdot \Theta(k)^{-k}.$$

V. Largest Clique

In this section, we present bounds on the clique number $\omega(G)$, i.e. the size of the largest clique in G. So far, we computed an upper bound on the expected number of k-cliques in the graph. We can use this result to obtain an upper bound on $\omega(G)$ by applying the Markov inequality

$$\Pr[K_k > 1] \leqslant \mathbb{E}[K_k].$$

6

Let therefore ε be an arbitrarily small constant and set $\mathbb{E}[K_k] = n^{-\varepsilon}$. We obtain an upper bound on the clique number that holds with high probability.

A. Sparse Case $(\alpha \geqslant 1)$

For $\alpha \geqslant 1$ we have

$$\Pr[K_k > 1] \leqslant \mathbb{E}[K_k] \leqslant n \cdot (c_1 k)^{-k} \overset{!}{=} n^{-\varepsilon}$$
$$\Leftrightarrow \quad (c_1 k)^{-k} = n^{-1-\varepsilon},$$

for some appropriate constant c_1. The solution to this equation is $k = \frac{1}{c_1} e^{W(c_1 \log(n^{1+\varepsilon}))}$, where $W(\cdot)$ is the Lambert W function, defined by $W(z)e^{W(z)} = z$. To see that this holds, observe

$$\begin{aligned}(c_1 k)^{-k} &= \exp\left(-W(c_1 \log(n^{1+\varepsilon})) \tfrac{1}{c_1} e^{W(c_1 \log(n^{1+\varepsilon}))}\right) \\ &= \exp(-\log(n^{1+\varepsilon})) = n^{-1-\varepsilon}\end{aligned}$$

The Lambert W function has an asymptotic expansion yielding $W(z) = \ln z - \ln\ln z + o(1)$ for growing z, which simplifies our formula to

$$\begin{aligned}k &= \tfrac{1}{c_1} e^{W(c_1 \log(n^{1+\varepsilon}))} \\ &= \tfrac{1}{c_1} e^{\log(c_1(1+\varepsilon)\log n) - \log\log(c_1(1+\varepsilon)\log n) + o(1)} \\ &= \frac{(1+\varepsilon)\log n}{\log(c_1(1+\varepsilon)\log n)}(1+o(1)) \\ &= (1+\varepsilon \pm o(1)) \frac{\log n}{\log\log n}.\end{aligned}$$

Therefore, there is no larger clique than $(1 + \varepsilon \pm o(1)) \frac{\log n}{\log\log n}$ with high probability, proving the upper bound. To obtain a matching lower bound, observe that the analysis in Section IV corresponds to a balls-into-bins experiment: There are $\frac{2\pi n}{c}$ circular sectors (bins), and each node (ball) is uniformly sampled in one of those. Since there are n balls and $\Theta(n)$ bins, an application of [22, Theorem 1] gives that with high probability the bin with the maximum load will hold at least $\frac{\log n}{\log\log n}(1 - o(1))$ balls, proving a lower bound for the maximum clique.

B. Dense Case $\left(\frac{1}{2} < \alpha < 1\right)$

On the other hand, when $\alpha < 1$, we get

$$\Pr[K_k > 1] \leqslant \mathbb{E}[K_k] \leqslant n^{(1-\alpha)k} \cdot (c_2 k)^{-k},$$

for some appropriate constant c_2. Let c_3 be some constant such that $c_3 c_2 > 1$. If one sets $k = c_3 n^{1-\alpha}$, we get

$$\begin{aligned}\mathbb{E}[K_k] &\leqslant n^{(1-\alpha)k} \cdot (c_2 k)^{-k} \\ &= n^{(1-\alpha)c_3 n^{1-\alpha}} \cdot (c_2 c_3 n^{1-\alpha})^{-c_3 n^{1-\alpha}} \\ &= (c_2 c_3)^{-e^{-C/2} n^{1-\alpha}}.\end{aligned}$$

This term is asymptotically smaller than $n^{-\varepsilon}$ for any ε, since

$$(c_2 c_3)^{-c_3 n^{1-\alpha}} \leqslant n^{-\varepsilon} \Leftrightarrow c_3 n^{1-\alpha} \log(c_2 c_3) \geqslant \varepsilon \log n,$$

since a polynomial in n is larger than $\varepsilon \log n$ for large enough n. Therefore, we know that $\omega(G) \leqslant \Theta(n^{1-\alpha})$ in this case. Combining this with the fact that by a Chernoff bound, with high probability at least $e^{-\alpha C/2} n^{1-\alpha}(1 - o(1))$ nodes have radial coordinate $r \leqslant R/2$, we have that the largest clique is of size $\Theta(n^{1-\alpha})$ with high probability.

VI. Algorithms for Finding Cliques

So far, we showed bounds on the size of cliques in hyperbolic random graphs, but did not yet investigate on how to find them algorithmically. For the case $\alpha \geqslant 1$ we showed that there are only few cliques in the graph, and therefore a degeneracy approach as in [7, 9] or a simple enumeration algorithm finds the largest clique in polynomial or even linear time. We now argue that a polynomial runtime is also achievable for $\alpha < 1$.

Clark et al. [5] provided a polynomial time algorithm for finding cliques in unit disc graphs (the euclidean analog to hyperbolic graphs) which works as follows. For any two nodes u, v with distance $\mathrm{d}(u,v) = d \leqslant R$ consider all cliques $\mathcal{C}(u,v)$ such that u, v span the diameter of $C \in \mathcal{C}(u,v)$, i.e.

$$\forall C \in \mathcal{C}(u,v), \forall u', v' \in C\colon \mathrm{d}(u',v') \leqslant d.$$

Clearly, for any clique C there exist two nodes $u, v \in C$ such that $C \in \mathcal{C}(u,v)$. Since all nodes that belong to a clique in $\mathcal{C}(u,v)$ are closer than d to both u and v, they lie inside a lens of the form $B_d(0) \cap B_d(d)$. Now consider the upper and the lower half of the lens. In the euclidean geometry, we can easily see that all points inside the upper (and lower, resp.) half of the lens have distance at most $d \leqslant R$ from each other. Hence, the induced subgraph on all nodes inside $B_d(0) \cap B_d(d)$ is the complement of a bipartite graph. Since we can find the largest independent set on bipartite graphs in polynomial time, we can find the largest clique in the complement as well. Finding the largest clique in the whole graph then boils down to searching for the largest clique in all such lenses induced by the n edges.

This approach translates to hyperbolic random graphs: we only have to show that the diameter of such a half lens is no larger than d. This can be done using the Karush-Kuhn-Tucker conditions, similar to Lemma IV.1. Due to space constraints we omit this technical but not difficult computation and conclude that there is a polynomial-time algorithm for finding a k-clique in hyperbolic random graphs.

VII. Conclusion

We present an analysis of the emergence of cliques in the hyperbolic random graphs and suggest how to find them algorithmically. We found that the large clustering coefficient of these graphs strongly affects the clique number when $\gamma > 3$. Previous models with independent edge probabilities predicted a clique number of 3 in this case, whereas the hyperbolic random graph contains a $\frac{\log n}{\log \log n}$ size clique. Our experiments validated that this is indeed closer to reality; and even in the case where $2 < \gamma < 3$ the clique number of hyperbolic random graphs was closer to real-world graphs than other existing models.

We therefore present another piece of evidence that hyperbolic random graphs model the behavior of real-world graphs best compared to existing scale-free models. This should motivate the study of other graph properties like giant component and small diameter in even more detail.

References

[1] W. Aiello, F. Chung, and L. Lu. A random graph model for massive graphs. In *STOC*, pp. 171–180, 2000.

[2] J. Alstott, E. Bullmore, and D. Plenz. powerlaw: a python package for analysis of heavy-tailed distributions. *PloS one*, 9: e85777, 2014.

[3] A.-L. Barabási and R. Albert. Emergence of scaling in random networks. *Science*, 286:509–512, 1999.

[4] M. Boguná, F. Papadopoulos, and D. Krioukov. Sustaining the internet with hyperbolic mapping. *Nature Comm.*, 1:62, 2010.

[5] B. N. Clark, C. J. Colbourn, and D. S. Johnson. Unit disk graphs. *Annals of Discrete Mathematics*, 48:165–177, 1991.

[6] A. Clauset, C. R. Shalizi, and M. E. Newman. Power-law distributions in empirical data. *SIAM Review*, 51:661–703, 2009.

[7] D. Eppstein and D. Strash. Listing all maximal cliques in large sparse real-world graphs. In *SEA*, pp. 364–375. 2011.

[8] N. Fountoulakis. On the evolution of random geometric graphs on spaces of negative curvature. In *arXiv*, 1205.2923, 2012.

[9] T. Friedrich and A. Krohmer. Parameterized clique on scale-free networks. In *ISAAC*, pp. 659–668, 2012.

[10] L. Gugelmann, K. Panagiotou, and U. Peter. Random hyperbolic graphs: degree sequence and clustering. In *ICALP*, pp. 573–585. 2012.

[11] S. Janson, T. Łuczak, and I. Norros. Large cliques in a power-law random graph. *J. Appl. Probab.*, 47:1124–1135, 2010.

[12] J. Jensen. Sur les fonctions convexes et les inégalités entre les valeurs moyennes. *Acta Mathematica*, 30:175–193, 1906.

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

[14] H. W. Kuhn. Nonlinear programming: a historical view. *SIGMAP Bull.*, pp. 6–18, 1982.

[15] L. Li, D. Alderson, J. C. Doyle, and W. Willinger. Towards a theory of scale-free graphs: Definition, properties, and implications. *Internet Mathematics*, 2:431–523, 2005.

[16] D. Liben-Nowell and J. M. Kleinberg. The link-prediction problem for social networks. *J. Am. Soc. Inf. Sci.*, 58:1019–1031, 2007.

[17] M. E. Newman. The structure and function of complex networks. *SIAM Review*, 45:167–256, 2003.

[18] M. E. J. Newman. Clustering and preferential attachment in growing networks. *Phys. Rev. E*, 64:025102, 2001.

[19] I. Norros and H. Reittu. On a conditionally poissonian graph process. *Advances in Applied Probability*, 38:59–75, 2006.

[20] F. Papadopoulos, D. Krioukov, M. Bogua, and A. Vahdat. Greedy forwarding in dynamic scale-free networks embedded in hyperbolic metric spaces. In *INFOCOM*, pp. 2973–2981, 2010.

[21] M. D. Penrose. *Random Geometric Graphs*. Oxford University Press, 2003.

[22] M. Raab and A. Steger. “Balls into bins” – a simple and tight analysis. In *RANDOM*, pp. 159–170. 1998.

[23] Y. Shavitt and T. Tankel. On the curvature of the internet and its usage for overlay construction and distance estimation. In *INFOCOM*, pp. 374–384, 2004.

[24] R. van der Hofstad. Random graphs and complex networks. Lecture notes available from http://www.win.tue.nl/~rhofstad/NotesRGCN.pdf, 2014.

[25] A. Vázquez. Growing network with local rules: Preferential attachment, clustering hierarchy, and degree correlations. *Phys. Rev. E*, 67:056104, 2003.

[26] Y. Wang and Y.-N. Yeh. Polynomials with real zeros and Pólya frequency sequences. *J. Comb. Theory, Ser. A*, 109:63–74, 2005.

[27] J. Yang and J. Leskovec. Defining and evaluating network communities based on ground-truth. In *SIGKDD Workshop on Mining Data Semantics*, p. 3, 2012. See http://snap.stanford.edu/data/.

Appendix

Proof of Lemma II.1, (3). Consider first the case $r \leqslant R/2$. In that case, the ball $B_r(R)$ fully encloses $B_0(r)$. Therefore, the intersection of those two areas has probability measure exactly $\mu(B_0(r)) = e^{-\alpha(R-r)}(1+o(1))$, proving the second case.

We now assume $r \geqslant R/2$. Then, we can write

$$\mu(B_r(R) \cap B_0(r)) = \mu(B_0(R-r)) + 2\int_{R-r}^{r}\int_{0}^{\theta_r(y)} f(y)\,\mathrm{d}\theta\,\mathrm{d}y,$$

where $\theta_r(y) = \arccos\left(\frac{\cosh(r)\cosh(y)-\cosh(R)}{\sinh(r)\sinh(y)}\right)$ is given by the definition of the distance function. The first part of the sum simplifies to $(1+o(1))e^{-\alpha r}$. For the integral, we note that it is the same (apart from the integration bounds) as in the proof of (2), see [10]. We can thereby simplify

$$2\int_{R-r}^{r}\int_{0}^{\theta_r(y)} f(y)\,\mathrm{d}\theta\,\mathrm{d}y = \qquad (9)$$
$$4\int_{R-r}^{r} e^{(R-r-y)/2}(1 \pm \mathcal{O}(e^{R-r-y}))\tfrac{\alpha\sinh(\alpha y)}{2\pi(\cosh(\alpha R)-1)}\,\mathrm{d}y.$$

We first compute the integral without the error term and later add the error term. We obtain

$$4\int_{R-r}^{r} \frac{e^{(R-r-y)/2}\alpha\sinh(\alpha y)}{2\pi(\cosh(\alpha R)-1)}\,\mathrm{d}y = \qquad (10)$$
$$\left[\tfrac{4\alpha\exp(\frac{R-r-y}{2})}{\pi(4\alpha^2-1)(\cosh(\alpha R)-1)}(2\alpha\cosh(\alpha y)+\sinh(\alpha y))\right]_{R-r}^{r}$$

Setting $C_1 = \frac{4\alpha}{\pi(4\alpha^2-1)(\cosh(\alpha R)-1)}$ and inserting the integration bounds, we obtain

$$\begin{aligned}(10) &= C_1\big(e^{\frac{R}{2}-r}(2\alpha\cosh(\alpha r)+\sinh(\alpha r)) \\ &\quad - 2\alpha\cosh(\alpha(R-r))+\sinh(\alpha(R-r))\big) \\ &= C_1(\alpha+\tfrac{1}{2})\big(e^{\frac{R}{2}-r}e^{\alpha r}(1+\Theta(e^{-2\alpha r})) \\ &\quad - e^{\alpha(R-r)}(1+\Theta(e^{-2\alpha(R-r)}))\big) \\ &= C_1(\alpha+\tfrac{1}{2})e^{\frac{R}{2}+r(\alpha-1)}(1+\Theta(e^{-2\alpha r}) \\ &\quad - \Theta(e^{(R-2r)(\alpha-\frac{1}{2})})).\end{aligned}$$

Since $1/(\cosh(x)-1) = 2e^{-x}(1+\Theta(e^{-x}))$, we conclude

$$\begin{aligned}(10) = \frac{2\alpha}{\pi(\alpha-\frac{1}{2})}&e^{-R(\alpha-1/2)+r(\alpha-1)} \\ &\cdot(1+\Theta(e^{-\alpha r})-\Theta(e^{(R-2r)(\alpha-\frac{1}{2})})).\end{aligned}$$

It is left to bound the error term in (9). To this end, we compute

$$\begin{aligned}&\int_{R-r}^{r}\mathcal{O}(e^{\frac{3}{2}(R-r-y)})\frac{\sinh(\alpha y)}{\cosh(\alpha R)}\,\mathrm{d}y \\ &= \int_{R-r}^{r}\mathcal{O}(e^{\frac{3}{2}(R-r-y)+\alpha(y-R)})\,\mathrm{d}y\end{aligned}$$

This integral evaluates to $\mathcal{O}(e^{-R(\alpha-1/2)+r(\alpha-1)})\cdot\mathcal{E}_3$, where $\mathcal{E}_3 = e^{R-2r}$ if $\alpha > \frac{3}{2}$, $\mathcal{E}_3 = e^{-\frac{R}{2}-r(\alpha-1)}$ if $\alpha < \frac{3}{2}$ and $\mathcal{E}_3 = e^{R-2r}(2r-R)$ otherwise. Observe that in the former cases $e^{(R-2r)(\alpha-\frac{1}{2})}$ dominates the error term. Since $(R-2r)(\alpha-\frac{1}{2}) \geqslant -\alpha r$, we obtain for $\alpha \neq \frac{3}{2}$ the final solution

$$\mu(B_r(R)\cap B_0(r)) = \frac{2\alpha}{\pi(\alpha-\frac{1}{2})}e^{-R(\alpha-1/2)+r(\alpha-1)}\cdot\mathcal{E}_2,$$

where $\mathcal{E}_2 = 1\pm\Theta(e^{(R-2r)(\alpha-\frac{1}{2})})$, proving the claim. □

Proof of Lemma IV.1. We use the Karush-Kuhn-Tucker conditions [14] for finding maxima under inequality constraints:

$$\begin{aligned}&\text{max. } f(r,r') = \cosh(r)\cosh(r') - \sinh(r)\sinh(r')\cdot q \\ &\quad\text{s.t. } 0 \leqslant r, r' \leqslant R,\end{aligned}$$

where we write $q = \cos(\frac{a}{n})$. We introduce slack variables $\alpha, \beta, \lambda, \mu$ and obtain the following system of equations.

$$\begin{aligned}0 \leqslant r, r' \leqslant R &\qquad \alpha,\beta,\lambda,\mu \geqslant 0 \\ \lambda(r-R) = 0 &\qquad \alpha r = 0 \\ \mu(r'-R) = 0 &\qquad \beta r' = 0 \\ \sinh(r)\cosh(r') - q\cosh(r)\sinh(r') &= \lambda - \alpha \\ \cosh(r)\sinh(r') - q\sinh(r)\cosh(r') &= \mu - \beta\end{aligned}$$

We distinguish several cases given by the conditions that are equal to zero. We notice that either $\alpha = 0$ or $r = 0$ such that above system holds. Let $r = 0$. The case for $r' = 0$ is analog. Then, the system collapses to

$$\begin{aligned}0 \leqslant r' \leqslant R &\qquad \alpha,\beta,\mu \geqslant 0 \\ q\sinh(r') = \alpha &\qquad \mu(r'-R) = 0 \\ \sinh(r') = \mu - \beta. &\qquad \beta r' = 0\end{aligned}$$

We can now further distinguish the cases (i) $\mu = 0$ and (ii) $r' = R$. For (i) we are done immediately, since then $\sinh(r') = -\beta$ implies that $r' = \beta = 0$; and u, v have a total distance of 0, which is clearly not a maximum. For the case (ii) we get that u and v have a distance of R. Let us now assume that $\alpha = \beta = 0$. Moreover, assume w.l.o.g. that $r \geqslant r'$. Then we obtain that $\sinh(r-r') \geqslant 0$ and therefore

$$\sinh(r)\cosh(r') \geqslant \cosh(r)\sinh(r').$$

Since $\cos(\frac{a}{n}) = q < 1$ in our case, this means that $\lambda > 0$, which in turn implies that $r = R$. Therefore, we can rewrite the last condition as

$$\cosh(R)\sinh(r') - q\sinh(R)\cosh(r') = \mu.$$

Again, we distinguish two cases (i) $\mu = 0$ and (ii) $\mu > 0$. For (i), we immediately obtain $r' = r = R$, for which we can use (8) and obtain a distance $\leqslant R$ for appropriately chosen constants. For the case (ii) where $\mu > 0$, we obtain that $r' = \tanh^{-1}(q\cdot\tanh(R))$. In this case, we first compute two bounds on the above term. By the definitions of hyperbolic functions, we get

$$r' = \tanh^{-1}(q\tanh(R)) = \tfrac{1}{2}\ln\left(\tfrac{(1+q)e^R+(1-q)e^{-R}}{(1-q)e^R+(1+q)e^{-R}}\right)$$

To obtain an upper bound, remember that $q \leqslant 1$ and $q = \cos(a/n) = 1 - \frac{a}{2}e^{-R} + \mathcal{O}(e^{-2R})$, giving

$$r' \leqslant \frac{1}{2}\ln\left(\frac{2e^R + ae^{-2R}}{a - \mathcal{O}(e^{-R})}\right) \leqslant \frac{R}{2} \pm \Theta(1),$$

when R is large enough. Similarly, the lower bound is

$$r' \geqslant \frac{1}{2}\ln\left(\frac{e^R}{a + \mathcal{O}(e^{-R})}\right) \geqslant \frac{R}{2} \pm \Theta(1).$$

Putting this together, we see that for $r_u = R$ and $r_v = \tanh^{-1}(q\tanh(R)) = \frac{R}{2} \pm \mathcal{O}(1)$ we get the distance

$$\begin{aligned}&\cosh(\mathrm{d}(u,v)) \\ &= \cosh(R)\cosh(\tanh^{-1}(q\tanh(R))) - \\ &\quad q\sinh(R)\sinh(\tanh^{-1}(q\tanh(R))) \\ &\leqslant \cosh(\tfrac{R}{2}\pm\Theta(1)) + ae^{-R}\sinh(R)\sinh(\tfrac{R}{2}\pm\Theta(1)) \\ &= \mathcal{O}(e^{R/2})\end{aligned}$$

which is smaller than $\cosh(R)$ for large enough R. This concludes the proof. □