View metadata, citation and similar papers at core.ac.uk
brought to you by CORE
provided by Dagstuhl Research Online Publication Server

Efficient Shortest Paths in Scale-Free Networks with Underlying Hyperbolic Geometry

Thomas Bläsius
Hasso Plattner Institute, University of Potsdam, Potsdam, Germany
thomas.blaesius@hpi.de

Cedric Freiberger
Hasso Plattner Institute, University of Potsdam, Potsdam, Germany
cedric.freiberger@student.hpi.de

Tobias Friedrich
Hasso Plattner Institute, University of Potsdam, Potsdam, Germany
tobias.friedrich@hpi.de

Maximilian Katzmann
Hasso Plattner Institute, University of Potsdam, Potsdam, Germany
maximilian.katzmann@hpi.de

Felix Montenegro-Retana
Hasso Plattner Institute, University of Potsdam, Potsdam, Germany
felix.montenegro-retana@student.hpi.de

Marianne Thieffry
Hasso Plattner Institute, University of Potsdam, Potsdam, Germany
marianne.thieffry@student.hpi.de

Abstract

A common way to accelerate shortest path algorithms on graphs is the use of a bidirectional search, which simultaneously explores the graph from the start and the destination. It has been observed recently that this strategy performs particularly well on scale-free real-world networks. Such networks typically have a heterogeneous degree distribution (e.g., a power-law distribution) and high clustering (i.e., vertices with a common neighbor are likely to be connected themselves). These two properties can be obtained by assuming an underlying hyperbolic geometry.

To explain the observed behavior of the bidirectional search, we analyze its running time on hyperbolic random graphs and prove that it is $\tilde{\mathcal{O}}(n^{2-1/\alpha} + n^{1/(2\alpha)} + \delta_{\max})$ with high probability, where $\alpha \in (0.5, 1)$ controls the power-law exponent of the degree distribution, and $\delta_{\max}$ is the maximum degree. This bound is sublinear, improving the obvious worst-case linear bound. Although our analysis depends on the underlying geometry, the algorithm itself is oblivious to it.

2012 ACM Subject Classification Theory of computation → Shortest paths, Theory of computation → Random network models, Theory of computation → Computational geometry

Keywords and phrases random graphs, hyperbolic geometry, scale-free networks, bidirectional shortest path

Digital Object Identifier 10.4230/LIPIcs.ICALP.2018.20

Related Version A full version of the paper is available at `https://arxiv.org/abs/1805.03253`.

Acknowledgements This research has received funding from the German Research Foundation (DFG) under grant agreement no. FR 2988 (ADLON, HYP).

© Thomas Bläsius, Cedric Freiberger, Tobias Friedrich, Maximilian Katzmann, Felix Montenegro-Retana, and Marianne Thieffry;
licensed under Creative Commons License CC-BY

45th International Colloquium on Automata, Languages, and Programming (ICALP 2018).
Editors: Ioannis Chatzigiannakis, Christos Kaklamanis, Dániel Marx, and Donald Sannella;
Article No. 20; pp. 20:1–20:14

Leibniz International Proceedings in Informatics
Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

1 Introduction

Finding shortest paths between nodes in a network is among the most basic graph problems. Besides being of independent interest, many algorithms use shortest path queries as a subroutine. On unweighted graphs, such queries can be answered in linear time using a *breadth-first search (BFS)*. Though this is optimal in the worst case, it is not efficient enough when dealing with large networks or problems involving many shortest path queries.

A way to heuristically improve the run time, is to use a *bidirectional BFS* [16]. It runs two searches, simultaneously exploring the graph from the start and the destination. The shortest path is then found once the two search spaces touch. Though this heuristic does not improve the worst-case running time, recent experiments by Borassi and Natale [6] suggest that it achieves a significant speedup on scale-free real-world networks. They also try to explain the observed run times by proving that the bidirectional BFS runs in sublinear expected time on different random graph models. Though this is a great result, we do not think that it provides a satisfying explanation for the good practical performance for two reasons.

First, the bidirectional search performs particularly well on networks with a heterogeneous degree distribution (i.e., few vertices with high degree, many vertices with low degree). A common assumption is that the degree distribution follows a *power-law*, i.e., the number of vertices of degree k is proportional to $k^{-\beta}$. The constant β is called the *power-law exponent* and is typically between 2 and 3. The above mentioned proof predicts a shorter execution time for homogeneous graphs (e.g., for Erdős-Rényi graphs) than for heterogeneous graphs (e.g., for Chung-Lu graphs), which contradicts the observed behavior.

Second, the proof relies on the independence of edges. In fact, this is the only assumption, which makes the same proof hold for multiple different models. However, this assumption is unrealistic for most real-world networks. The dependence between edges is typically measured with the *clustering coefficient*. The *local clustering coefficient* of a vertex v is the probability that two randomly chosen neighbors of v are adjacent. The clustering coefficient of the graph is the average of all local coefficients. The assumption of independent edges thus implies a clustering coefficient close to 0. In contrast, the three best performing instances in [6, Figure 2] have comparatively high clustering coefficients 0.47, 0.49, and 0.57 [14].

In this paper, we analyze the bidirectional BFS on hyperbolic random graphs, which are generated by randomly placing vertices in the hyperbolic plane and connecting each pair that is geometrically close. This model was introduced by Krioukov et al. [13] with the aim to generate graphs that closely resemble real-world networks. Hyperbolic random graphs in particular have a power-law degree distribution and high clustering [12, 13]. Moreover, as these properties emerge naturally from the hyperbolic geometry, the model is conceptually simple, which makes it accessible to mathematical analysis. It has thus gained popularity in different research areas and has been studied from different perspectives.

From the network-science perspective, the goal is to gather knowledge about real-world networks. This is for example achieved by assuming that a real-world network has a hidden underlying hyperbolic geometry, which can be revealed by embedding it into the hyperbolic plane [1, 5]. From the mathematical perspective, the focus lies on studying structural properties. Beyond degree distribution and clustering [12], the diameter [11], the component structure [4], the clique size [3], and separation properties [2] have been studied successfully.

Finally, there is the algorithmic perspective, which is the focus of this paper. Usually algorithms are analyzed by proving worst-case running times. Though this is the strongest possible performance guarantee, it is rather pessimistic as practical instances rarely resemble worst-case instances. Techniques leading to a more realistic analysis include parameterized

or average case complexity. The latter is based on the assumption that instances are drawn from a certain probability distribution. Thus, its explanatory power depends on how realistic the distribution is. For hyperbolic random graphs, the maximum clique can be computed in polynomial time [3], and there are several algorithmic results based on the fact that hyperbolic random graphs have sublinear tree width [2]. Moreover, there is a compression algorithm that can store a hyperbolic random graph using on $O(n)$ bits in expectation [8, 15]. Finally, a close approximation of the shortest path between two nodes can be found using greedy routing, which visits only $\mathcal{O}(\log \log n)$ nodes for most start-destination pairs [9]. The downside of all these algorithms is that they need to know the underlying geometry, i.e., the coordinates of each vertex. Unfortunately, this is a rather unrealistic assumption for real-world networks. To the best of our knowledge, we present the first analysis of an algorithm on hyperbolic random graphs that is oblivious to the underlying geometry.

Contribution and Outline. After an introduction to hyperbolic random graphs in Section 2, we analyze the bidirectional BFS in Section 3. We first prove in Section 3.1 that a certain greedy strategy for deciding when to alternate between the forward and the backward search is not much worse than any other alternation strategy. We note that this result is interesting in its own right and does not depend on the input. In Section 3.2 we analyze the bidirectional BFS on hyperbolic random graphs. We show that, for any pair of vertices, it computes a shortest path in $\tilde{\mathcal{O}}(n^{2-1/\alpha} + n^{1/(2\alpha)} + \delta_{\max})$ time with high probability[1], where $\alpha \in (0.5, 1)$ controls the power-law exponent and $\delta_{\max}$ is the maximum degree of the graph (which is $\tilde{\mathcal{O}}(n^{1/(2\alpha)})$ almost surely [12]). We note that drawing the hyperbolic random graph is the only random choice here; once this is done our analysis always assumes the worst case. Thus, the bound in particular holds for every start-destination pair. Section 4 contains concentration bounds that were left out in Section 3 to improve readability. In Section 5, we conclude by comparing our theoretical results to empirical data.

2 Preliminaries

Let $G = (V, E)$ be an undirected, unweighted, and connected graph. We denote the number of vertices and edges with n and m, respectively. With $N(v) = \{w \in V \mid \{v, w\} \in E\}$, we denote the *neighborhood* of a vertex $v \in V$. The *degree* of v is $\deg(v) = |N(v)|$. We denote the maximum degree with $\delta_{\max}$. The soft O-notation $\tilde{\mathcal{O}}$ suppresses poly-logarithmic factors.

The Hyperbolic Plane. The major difference between hyperbolic and Euclidean geometry is the exponential expansion of the hyperbolic plane. A circle of radius r has area $2\pi(\cosh(r)-1)$ and circumference $2\pi \sinh(r)$, with $\cosh(x) = (e^x + e^{-x})/2$ and $\sinh(x) = (e^x - e^{-x})/2$, both growing as $e^x/2 \pm o(1)$. To identify points, we use radial coordinates with respect to a designated origin O and a ray starting at O. A point p is uniquely determined by its *radius* r, which is the distance to O, and the *angle* (or *angular coordinate*) φ between the reference ray and the line through p and O. In illustrations, we use the *native representation*, obtained by interpreting the hyperbolic coordinates as polar coordinates in the Euclidean plane, see Figure 1 (left). Due to the exponential expansion, line segments bend towards the origin O. Let $p_1 = (r_1, \varphi_1)$ and $p_2 = (r_2, \varphi_2)$ be two points. The *angular distance* between p_1 and p_2 is the angle between the rays from the origin through p_1 and p_2. Formally, it is

[1] *With high probability* and *almost surely* refer to probabilities $1 - O(1/n)$ and $1 - o(1)$, respectively.

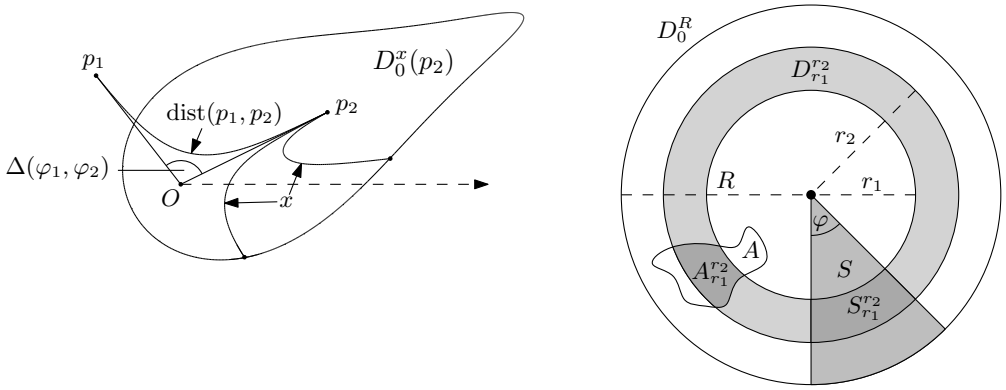

Figure 1 Left: A circle and several line segments in the native representation of the hyperbolic plane. A disk of radius x is centered at p_2. Right: Geometric shapes and their intersections.

$\Delta(\varphi_1, \varphi_2) = \pi - |\pi - |\varphi_1 - \varphi_2||$. The hyperbolic distance $\text{dist}(p_1, p_2)$ is given by

$$\cosh(\text{dist}(p_1, p_2)) = \cosh(r_1)\cosh(r_2) - \sinh(r_1)\sinh(r_2)\cos(\Delta(\varphi_1, \varphi_2)).$$

Note how the angular coordinates make simple definitions cumbersome as angles are considered modulo 2π, leading to a case distinction depending on where the reference ray lies. Whenever possible, we implicitly assume that the reference ray was chosen such that we do not have to compute modulo 2π. Thus, the above angular distance between p_1 and p_2 simplifies to $|\varphi_1 - \varphi_2|$. A third point $p = (r, \varphi)$ *lies between* p_1 and p_2 if $\varphi_1 \leq \varphi \leq \varphi_2$ or $\varphi_2 \leq \varphi \leq \varphi_1$.

Throughout the paper, we regularly use different geometric shapes, which are mostly based on disks centered at the origin O, as can be seen in Figure 1 (right). With $D_{r_1}^{r_2}$ we denote the set of points that have radius between r_1 and r_2. Note that D_0^r is the disk of radius r centered at O. The restriction of a disk D_0^r to all points with angular coordinates in a certain interval is called *sector*, which we usually denote with the letter S. Its *angular width* is the length of this interval. For an arbitrary set of points A, $A_{r_1}^{r_2}$ denotes the restriction of A to points with radii in $[r_1, r_2]$, i.e., $A_{r_1}^{r_2} = A \cap D_{r_1}^{r_2}$.

Hyperbolic Random Graphs. A *hyperbolic random graph* is generated by drawing n points uniformly at random in a disk of the hyperbolic plane and connecting pairs of points whose distance is below a threshold. More precisely, the model depends on two parameters C and α. The generated graphs have a power-law degree distribution with power-law exponent $\beta = 2\alpha + 1$ and with an average degree depending on C. The n points are sampled within the disk D_0^R of radius $R = 2\log n + C$. For each vertex, the angular coordinate is drawn uniformly from $[0, 2\pi]$. The radius r is sampled according to the probability density function

$$f(r) = \frac{1}{2\pi}\frac{\alpha\sinh(\alpha r)}{\cosh(\alpha R) - 1} = \Theta(e^{\alpha(r-R)}), \tag{1}$$

for $r \in [0, R]$. For $r > R$, $f(r) = 0$. Two vertices are connected by an edge if and only if their hyperbolic distance is less than R. The above probability distribution is a natural choice as the probability for a vertex ending up in a certain region is proportional to its area (at least for $\alpha = 1$). Note that the exponential growth in r reflects the fact that the area of a disk grows exponentially with the radius. It follows that a hyperbolic random graph typically has few vertices with high degree close to the center of the disk and many vertices with low degree near its boundary. The following lemma is common knowledge; see the full version of the paper for a proof.

▶ **Lemma 1.** *Let G be a hyperbolic random graph. Furthermore, let v_1, v_2 be two nodes with radii $r_1 \leq r_2 \leq R$, respectively, and with the same angular coordinate. Then $N(v_2) \subseteq N(v_1)$.*

Given two vertices with fixed radii r_1 and r_2, their hyperbolic distance grows with increasing angular distance. The maximum angular distance such that they are still adjacent [12, Lemma 3.1] is

$$\theta(r_1, r_2) = \arccos\left(\frac{\cosh(r_1)\cosh(r_2) - \cosh(R)}{\sinh(r_1)\sinh(r_2)}\right) = 2e^{\frac{R-r_1-r_2}{2}}(1 + \Theta(e^{R-r_1-r_2})). \tag{2}$$

The probability that a sampled node falls into a given subset $A \subseteq D_0^R$ of the disk is given by its probability measure $\mu(A) = \int_A f(r)\,\mathrm{d}r$, which can be thought of as the area of A. There are two types of regions we encounter regularly: disks D_0^r with radius r centered at the origin and disks $D_0^R(r, \varphi)$ of radius R centered at a point (r, φ). Note that the measure of $D_0^R(r, \varphi)$ gives the probability that a random vertex lies in the neighborhood of a vertex with position (r, φ). Gugelmann et al. [12, Lemma 3.2] showed that

$$\mu(D_0^r) = e^{\alpha(r-R)}(1 + o(1)), \text{ and} \tag{3}$$

$$\mu(D_0^R(r, \varphi)) = \Theta(e^{-r/2}). \tag{4}$$

For a given region $A \subseteq D_0^R$ of the disk, let $X_1, \dots, X_n$ be random variables with $X_i = 1$ if $i \in A$ and $X_i = 0$ otherwise. Then $X = \sum_{i=1}^n X_i$ is the number of vertices lying in A. By the linearity of expectation, we obtain that the expected number of vertices in A is $\mathbb{E}[X] = \sum_{i=1}^n \mathbb{E}[X_i] = n\mu(A)$. To bound the number of vertices in A *with high probability*, we regularly use the following Chernoff bound.

▶ **Theorem 2** (Chernoff Bound [10, A.1])**.** *Let $X_1, \dots, X_n$ be n independent random variables with $X_i \in \{0, 1\}$ and let X be their sum. For any $\delta > 0$,*

$$\Pr[X > (1 + \delta)\mathbb{E}[X]] < \exp\left(-\frac{\delta^2}{3}\mathbb{E}[X]\right).$$

▶ **Corollary 3.** *Let $X_1, \dots, X_n$ be n independent random variables with $X_i \in \{0, 1\}$ and let X be their sum. Let $f(n) = \Omega(\log n)$. If $f(n)$ is an upper bound for $\mathbb{E}[X]$, then for any constant c there is a constant c' such that $X \leq c' f(n)$ holds with probability $1 - \mathcal{O}(n^{-c})$.*

3 Bidirectional BFS

In this section, we analyze the running time of the bidirectional BFS on hyperbolic random graphs. Our results are summarized in the following main theorem.

▶ **Theorem 4.** *Let G be a hyperbolic random graph. With high probability the shortest path between any two vertices in G can be computed in $\tilde{\mathcal{O}}(n^{2-1/\alpha} + n^{1/(2\alpha)} + \delta_{\max})$ time.*

To prove this, we make use of the hyperbolic geometry in the following way; see Figure 2. As long as the two searches visit only low-degree vertices, all explored vertices lie within a small region, i.e., the searches operate locally. Once the searches visit high-degree vertices closer to the center of the hyperbolic disk (gray area in Figure 2), it takes only few steps to complete the search, as hyperbolic random graphs have a densely connected core. Thus, we split our analysis in two phases: a first phase in which both searches advance towards the center and a second phase in which both searches meet in the center. Note that this strategy

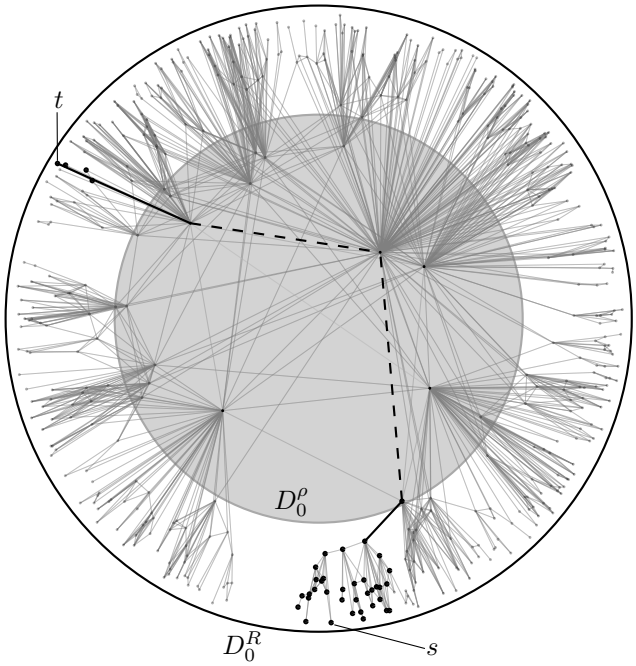

Figure 2 Visualization of the two phases of each BFS in a hyperbolic random graph. Nodes that are visited during the first phase are bold. The bold black edges denote the first encounter of a node in the inner disk D_0^ρ (gray region). This corresponds to the first step in the second phase. The last step then leads to a common neighbor via the dashed edges.

assumes that we know the coordinates of the vertices as we would like to stop a search once it reached the center. To resolve this issue, we first show in Section 3.1 that there exists an alternation strategy that is oblivious to the geometry but performs not much worse than any other alternation strategy. We note that this result is independent of hyperbolic random graphs and thus interesting in its own right. Afterwards, in Section 3.2, we actually analyze the bidirectional search in hyperbolic random graphs.

3.1 Bidirectional Search and Alternation Strategies

In an unweighted and undirected graph $G = (V, E)$, a BFS finds the shortest path between two vertices $s, t \in V$ by starting at s and exploring the graph level after level, where the ith level L_i^s contains the vertices with distance i to s. More formally, the BFS starts with the set $L_0^s = \{s\}$ on level 0. Assuming the first i levels $L_1^s, \dots, L_i^s$ have been computed already, one obtains the next level L_{i+1}^s as the set of neighbors of vertices in level L_i^s that are not contained in earlier layers. Computing L_{i+1}^s from L_i^s is called an *exploration step*, which is obtained by *exploring the edges* between vertices in L_i^s and L_{i+1}^s.

The bidirectional BFS runs two BFSs simultaneously. The *forward search* starts at s and the *backward search* starts at t. The shortest path between the two vertices can then be obtained, once the search spaces of the forward and backward search touch. Since the two searches cannot actually be run simultaneously, they alternate depending on their progress. When exactly the two searches alternate is determined by the *alternation strategy*. Note that we only swap after full exploration steps, i.e., we never explore only half of level i of one search before continuing with the other. This has the advantage that we can be certain to know the shortest path once a vertex is found by both searches.

In the following we define the *greedy alternation strategy* as introduced by Borassi and Natale [6] and show that it is not much worse than any other alternation strategy. Assume the latest layers of the forward and backward searches are L_i^s and L_j^t, respectively. Then the next exploration step of the forward search would cost time proportional to $c_i^s := \sum_{v \in L_i^s} \deg(v)$, while the cost for the backward search is $c_j^t := \sum_{v \in L_j^t} \deg(v)$. The *greedy alternation strategy* then greedily continues with the search that causes the fewer cost in the next exploration step, i.e., it continues with the forward search if $c_i^s \leq c_j^t$ and with the backward search otherwise.

▶ **Theorem 5.** *Let G be a graph with diameter d. If there exists an alternation strategy such that the bidirectional BFS explores $f(n)$ edges, then the bidirectional BFS with greedy alternation strategy explores at most $d \cdot f(n)$ edges.*

Proof. Let A be the alternation strategy that explores only $f(n)$ edges. First note that the number of explored edges only depends on the number of layers explored by the two different searches and not on the actual order in which they are explored. Thus, if the greedy alternation strategy is different from A, we can assume without loss of generality that the greedy strategy performed more exploration steps in the forward search and fewer in the backward search compared to A. Let c^s and c^t be the number of edges explored by the forward and backward search, respectively, when using the greedy strategy. Moreover, let j be the last layer of the backward search (which is actually not explored) and, accordingly, let c_j^t be the number of edges the next step in the backward search would have explored. Then $c^t + c_j^t \leq f(n)$ as, when using A, the backward search still explores layer j. Moreover, the forward search with the greedy strategy explores at most $c^t + c_j^t$ (and therefore at most $f(n)$) edges in each step, as exploring the backward search would be cheaper otherwise. Consequently, each step in the forward and backward search costs at most $f(n)$. As there are at most d steps in total, we obtain the claimed bound. ◀

3.2 Bidirectional Search in Hyperbolic Random Graphs

To analyze the size of the search space of the bidirectional BFS in hyperbolic random graphs, we separate the whole disk D_0^R into two partitions. One is the *inner disk* D_0^ρ centered at the origin. Its radius ρ is chosen in such a way that any two vertices in D_0^ρ have a common neighbor with high probability. The second part is the *outer band* D_ρ^R, the remainder of the whole disk. A single BFS now explores the graph in two phases. In the first phase, the BFS explores vertices in the outer band. The phase ends, when the next vertex to be encountered lies in the inner disk. Once both BFSs completed the first phase, they only need at most two more steps for their search spaces to share a vertex. One step to encounter the vertex in the inner disk and another step to meet at their common neighbor that any two vertices in the inner disk have with high probability; see Figure 2.

For our analysis we assume an alternation strategy in which each search stops once it explored one additional layer after finding the first vertex in the inner disk D_0^ρ. Of course, this cannot be implemented without knowing the underlying geometry of the network. However, by Theorem 5 the search space explored using the greedy alternation strategy is only a poly-logarithmic factor larger, as the diameter of hyperbolic random graphs is poly-logarithmic with high probability [11]. The following lemma shows for which choice of ρ the above sketched strategy works.

▶ **Lemma 6.** *Let G be a hyperbolic random graph. With high probability, G contains a vertex that is adjacent to every other vertex in D_0^ρ, for $\rho = \frac{1}{\alpha}(\log n - \log\log n)$.*

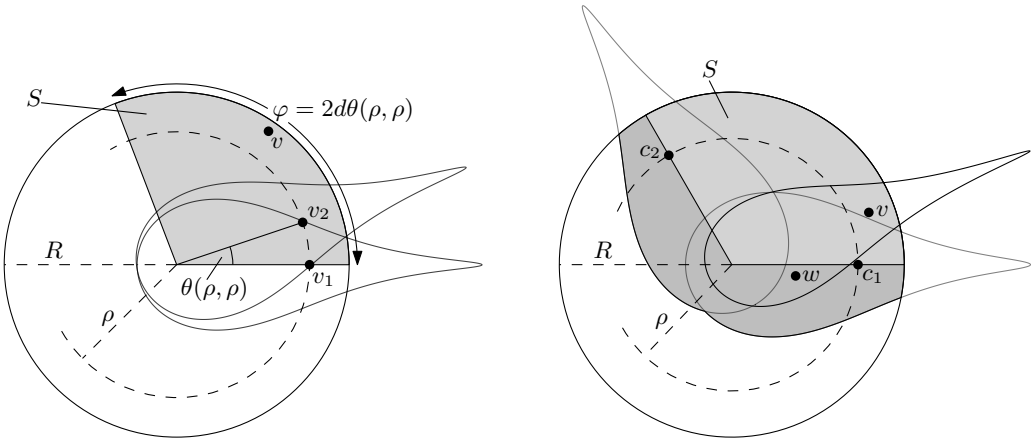

Figure 3 Left: The sector S of angular width φ contains the search space of a BFS in the outer band D_ρ^R starting at v. The vertices v_1 and v_2 are at maximum angular distance to still be adjacent. Right: Neighbor w of vertex v is in S (gray) or a neighbor of c_1 or c_2 (dark gray).

Proof. Assume v is a vertex with radius at most $R-\rho$. Note that the distance between two points is upper bounded by the sum of their radii. Thus, every vertex in D_0^ρ has distance at most R to v, and is therefore adjacent to v. Hence, to prove the claim, it suffices to show the existence of this vertex v with radius at most $R-\rho$. As described in Section 2, the probability for a single vertex to have radius at most $R-\rho$ is given by the measure $\mu(D_0^{R-\rho})$. Using Equation (3) we obtain

$$\begin{aligned}\mu(D_0^{R-\rho}) &= e^{-\alpha\rho}(1+o(1)) \\ &= \frac{\log n}{n}(1+o(1)).\end{aligned}$$

Thus, the probability that none of the n vertices lies in $D_0^{R-\rho}$ is $O((1-\frac{\log n}{n})^n) = O(\frac{1}{n})$. Hence, there is at least one vertex with radius at most $R-\rho$ with high probability. ◀

In the following, we first bound the search space explored in the first phase, i.e., before we enter the inner disk D_0^ρ. Afterwards we bound the search space explored in the second phase, which consists of two exploration steps. The first one to enter D_0^ρ and the second one to find a common neighbor, which exists due to Lemma 6.

3.2.1 Search Space in the First Phase

To bound the size of the search space in the outer band, we make use of the network geometry in the following way. For two vertices in the outer band to be adjacent, their angular distance has to be small. Moreover, the number of exploration steps is bounded by the diameter of the graph. Thus, the maximum angular distance between vertices visited in the first phase cannot be too large. Note that following lemma restricts the search to a sublinear portion of the disk, which we later use to show that also the number of explored edges is sublinear.

▶ **Lemma 7.** *With high probability, all vertices a BFS on a hyperbolic random graph explores before finding a vertex with radius at most* $\rho = \frac{1}{\alpha}(\log n - \log\log n)$ *lie within a sector of angular width* $\tilde{\mathcal{O}}(n^{1-1/\alpha})$.

Proof. For an illustration of the proof see Figure 3 (left). Recall from Section 2 that $\theta(r_1, r_2)$ denotes the maximum angular distance between two vertices of radii r_1 and r_2 such that

they are still adjacent. As this angle increases with decreasing radii, $\theta(r_1, r_2) \leq \theta(\rho, \rho)$ holds for all vertices in the outer band D_ρ^R.

Now assume we start a BFS at a vertex $v \in D_\rho^R$ and perform d exploration steps without leaving the outer band D_ρ^R. Then no explored vertex has angular distance more than $d\theta(\rho, \rho)$ from v. Thus, the whole search space lies within a disk sector of angular width $2d\theta(\rho, \rho)$. The number of steps d is at most poly-logarithmic as the diameter of a hyperbolic random graph is poly-logarithmic with high probability [11]. Using Equation (2) for $\theta(\rho, \rho)$, we obtain

$$\begin{aligned}\theta(\rho, \rho) &= 2e^{\frac{R-2\rho}{2}}(1 + \Theta(e^{R-2\rho})) \\ &= 2e^{C/2}n^{1-1/\alpha}\log^{1/\alpha} n(1 + \Theta((\log n/n^{1-\alpha})^{2/\alpha}) \\ &= \mathcal{O}(n^{1-1/\alpha}\log^{1/\alpha} n),\end{aligned}$$

which proves the claimed bound. ◀

Note that the expected number of vertices in a sector S of angular width φ is linear in $n\varphi$ due to the fact that the angular coordinate of each vertex is chosen uniformly at random. Thus, Lemma 7 already shows that the expected number of vertices visited in the first phase of the BFS is $\tilde{\mathcal{O}}(n^{2-1/\alpha})$, which is sublinear in n. It is also not hard to see that this bound holds with high probability (see Corollary 3). To also bound the number of explored edges, we sum the degrees of vertices in S. It is not surprising that this yields the same asymptotic bound in expectation, as the expected average degree in a hyperbolic random graph is constant. However, showing that this value is concentrated around its expectation is more involved. Though we can use techniques similar to those that have been used to show that the average degree of the whole graph is constant with high probability [7, 12], the situation is complicated by the restriction to a sublinear portion of the disk. Nonetheless, we obtain the following theorem.

▶ **Theorem 8.** *Let G be a hyperbolic random graph. The degrees of vertices in every sector of angular width φ sum to $\tilde{\mathcal{O}}(\varphi n + \delta_{\max})$ with high probability if $\varphi = \Omega(n^{1-1/\alpha}\log n)$.*

We note that $\delta_{\max}$ has to be included here, as the theorem states a bound for every sector, and thus in particular for sectors containing the vertex of maximum degree. Recall, that $\delta_{\max} = \tilde{\mathcal{O}}(n^{1/(2\alpha)})$ holds almost surely [12]. Moreover, we note that the condition $\varphi = \Omega(n^{1-1/\alpha}\log n)$ is crucial for our proof, i.e., the angular width of the sector has to be sufficiently large for the concentration bound to hold. Fortunately, this matches the bound found in Lemma 7. As the proof for Theorem 8 is rather technical, we defer it to Section 4. Together with Lemma 7, we obtain the following corollary.

▶ **Corollary 9.** *On a hyperbolic random graph, the first phase of the bidirectional search explores with high probability only $\tilde{\mathcal{O}}(n^{2-1/\alpha} + \delta_{\max})$ many edges.*

3.2.2 Search Space in the Second Phase

The first phase of the BFS is completed when the next vertex to be encountered lies in the inner disk. Thus, the second phase consists of only two exploration steps. One step to encounter the vertex in the inner disk and another step to meet the other search. Thus, to bound the running time of the second phase, we have to bound the number of edges explored in these two exploration steps. To do this, let V_1 be the set of vertices encountered in the first phase. Recall that all these vertices lie within a sector S of angular width $\varphi = \tilde{\mathcal{O}}(n^{1-1/\alpha})$ (Lemma 7). The number of explored edges in the second phase is then bounded by the sum of

degrees of all neighbors $N(V_1)$ of vertices in V_1. To bound this sum, we divide the neighbors of V_1 into two categories: $N(V_1) \cap S$ and $N(V_1) \setminus S$. Note that we already bounded the sum of degrees of vertices in S for the first phase (see Theorem 8), which clearly also bounds this sum for $N(V_1) \cap S$. Thus, it remains to bound the sum of degrees of vertices in $N(V_1) \setminus S$.

To bound this sum, we introduce two *hypothetical vertices* (i.e., vertices with specific positions that are not actually part of the graph) c_1 and c_2 such that every vertex in $N(V_1) \setminus S$ is a neighbor of c_1 or c_2. Then it remains to bound the sum of degrees of neighbors of these two vertices. To define c_1 and c_2, recall that the first phase was not only restricted to the sector S but also to points with radius greater than ρ, i.e., all vertices in V_1 lie within S_ρ^R. The hypothetical vertices c_1 and c_2 are basically positioned at the corners of this region, i.e., they both have radius ρ, and they assume the maximum and minimum angular coordinate within S, respectively. Figure 3 (right) shows these positions. We obtain the following.

▶ **Lemma 10.** *Let G be a hyperbolic random graph and let $v \in S_\rho^R$ for a sector S. Then, every neighbor of v lies in S or is a neighbor of one of the hypothetical vertices c_1 or c_2.*

Proof. Let $v = (r, \varphi) \in S_\rho^R$ and $w \in N(v) \setminus S$. Without loss of generality, assume that c_1 lies between v and w, as is depicted in Figure 3 (right). Now consider the point $v' = (\rho, \varphi)$ obtained by moving v to the same radius as c_1. According to Lemma 1 we have $N(v) \subseteq N(v')$. In particular, it holds that $w \in N(v')$ and therefore $\operatorname{dist}(v', w) \leq R$. Since v' and c_1 have the same radial coordinate and c_1 is between v' and w, it follows that $\operatorname{dist}(c_1, w) \leq R$. ◀

By the above argumentation, it remains to sum the degrees of neighbors of c_1 and c_2. It is not hard to see that the degrees of the neighbors of a node with radius r sum to $\mathcal{O}(ne^{-(\alpha-1/2)r})$ in expectation. For c_1 and c_2, which both have radius ρ, the degrees of their neighbors thus sum to $\tilde{\mathcal{O}}(n^{1/(2\alpha)})$ in expectation. Note that this matches the claimed bound in Theorem 4. However, to actually prove Theorem 4, we need to show that this bound holds with high probability for every possible angular coordinates of c_1 and c_2. Again, showing this concentration bound is rather technical and thus deferred to Section 4. Together with the bounds on the sum of degrees in a sector of width $\varphi = \tilde{\mathcal{O}}(n^{1-1/\alpha})$ (Theorem 8), we obtain the following corollary, which concludes the proof of Theorem 4.

▶ **Corollary 11.** *On a hyperbolic random graph, the second phase of the bidirectional search explores with high probability only $\tilde{\mathcal{O}}(n^{2-1/\alpha} + n^{1/(2\alpha)} + \delta_{\max})$ many edges.*

4 Concentration Bounds for the Sum of Vertex Degrees

Here we prove the concentration bounds that were announced in the previous section. For the first phase, we already know that the search space is contained within a sector S of sublinear width (Lemma 7). Thus, the running time in the first phase is bounded by the sum of vertex degrees in this sector. Moreover, all edges explored in the second phase also lie within the same sector S or are incident to neighbors of the two hypothetical vertices c_1 and c_2 (Lemma 10). Thus, the running time of the second phase is bounded by the sum of vertex degrees in S and in the neighborhood of c_1 and c_2. We start by proving Theorem 8 to bound the sum of degrees in a given sector. Afterwards, we consider the neighborhood of c_1 and c_2. To improve readability, we restate Theorem 8 here.

▶ **Theorem 8.** *Let G be a hyperbolic random graph. The degrees of vertices in every sector of angular width φ sum to $\tilde{\mathcal{O}}(\varphi n + \delta_{\max})$ with high probability if $\varphi = \Omega(n^{1-1/\alpha} \log n)$.*

Due to space constraints, we only sketch the proof by explaining the overall strategy and stating the core lemmas. A full proof can be found in the full version of the paper. The proof of Theorem 8 basically works as follows. For each degree, we want to compute the number of vertices of this degree and multiply it with the degree. As all vertices with a certain degree have roughly the same radius, we can separate the disk into small bands, one for each degree. Then summing over all degrees comes down to summing over all bands and multiplying the number of vertices in this band with the corresponding degree. If we can prove that each of these values is highly concentrated (i.e., probability $1 - O(n^{-2})$), we obtain that the sum is concentrated as well (using the union bound). Unfortunately, this fails in two situations. For large radii the degree is too small to be concentrated around its expected value. Moreover, for small radii, the number of vertices within the corresponding band (i.e., the number of high degree vertices) is too small to be concentrated.

To overcome this issue, we partition the sector S into three parts. An inner part $S_0^{\rho_1}$, containing all points of radius at most ρ_1, an outer part $S_{\rho_2}^R$, containing all points of radius at least ρ_2, and a central part $S_{\rho_1}^{\rho_2}$, containing all points in between. We choose ρ_2 in such a way that the smallest degree in the central part $S_{\rho_1}^{\rho_2}$ is $\Omega(\log n)$, which ensures that all vertex degrees in $S_{\rho_1}^{\rho_2}$ are concentrated. Moreover, we choose ρ_1 such that the number of vertices with maximum degree in $S_{\rho_1}^{\rho_2}$ is $\Omega(\log n)$, which ensures that for each vertex degree, the number of vertices with this degree is concentrated. To achieve this, we set

$$\rho_1 = 2\log n - \frac{\log(\varphi n) - \log\log n}{\alpha} \quad \text{and} \quad \rho_2 = \frac{\log n}{\alpha},$$

and show concentration separately for the three parts.

The Inner Part of a Sector. The inner part $S_0^{\rho_1}$ contains vertices of high degree. It is not hard to see that there are only poly-logarithmically many vertices with radius at most ρ_1. Thus, we obtain the following lemma.

▶ **Lemma 12.** *Let G be a hyperbolic random graph. For every sector S of angular width φ, the degrees of the nodes in $S_0^{\rho_1}$ sum to $\tilde{\mathcal{O}}(\delta_{\max})$ with high probability.*

The Central Part of a Sector. For each possible vertex degree k, we want to compute the number of vertices with this degree in the central part $S_{\rho_1}^{\rho_2}$. First note, that by Equation (4) a vertex with fixed radius has expected degree $\Theta(k)$ if this radius is $2\log(n/k)$. Motivated by this, we define $r_k = 2\log(n/k)$. To bound the sum of degrees in the central part $S_{\rho_1}^{\rho_2}$, we use that vertices with radius significantly larger than r_k also have a smaller degree. More formally, one can show that there exists a constant ε such that all vertices of degree k have radius at most $r_k + \varepsilon$ with high probability. From this, we can derive a bound $g(k)$ for the number of vertices with degree at least k by bounding the number of vertices with radius at most $r_k + \varepsilon$. Then summing the vertex degrees boils down to integrating over $g(k)$, which yields the following lemma.

▶ **Lemma 13.** *Let G be a hyperbolic random graph. For every sector S of angular width φ, the degrees of the nodes in $S_{\rho_1}^{\rho_2}$ sum to $\mathcal{O}(\varphi n^{3-2\alpha-1/(2\alpha)})$ with high probability.*

Note that $3 - 2\alpha - 1/(2\alpha) \leq 1$ for $\alpha \in [0.5, 1]$. Thus, the lemma in particular shows that $S_{\rho_1}^{\rho_2}$ contains at most $\mathcal{O}(\varphi n)$ edges, as claimed in Theorem 8.

The Outer Part of a Sector. The outer part $S_{\rho_2}^R$ contains many vertices, all of which have low expected degree. To bound their sum with high probability, we consider the coordinates of the vertices as random variables and the sum of their degrees as a function in these

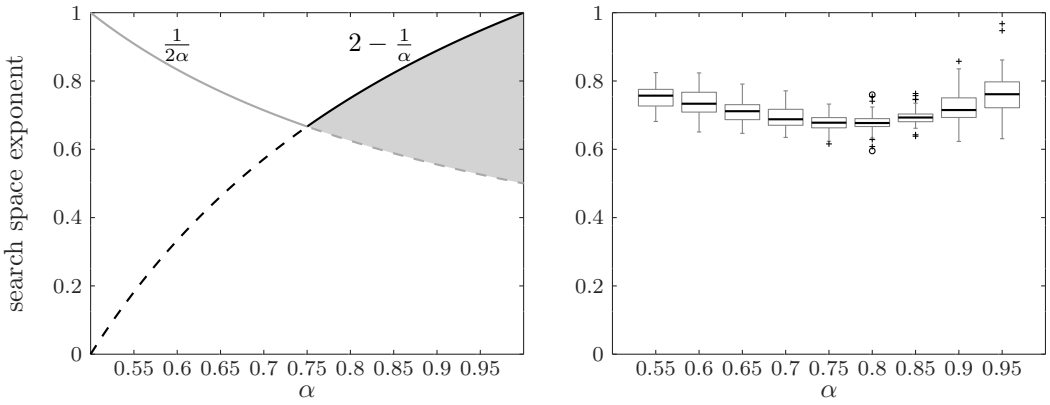

Figure 4 Left: The exponent of our theoretical bound depending on α. Right: The corresponding empirically measured search spaces. The data was obtained by generating 20 hyperbolic random graphs with average degree roughly 8 for each shown α and each $n \in \{100\text{k}, 200\text{k}, 300\text{k}\}$. For each graph we sampled 300k start-destination pairs and report the maximum number of edges explored in one search. The numbers are normalized with the total number of edges m of the graph such that x is plotted for a search space of size m^x.

variables. Then, our plan to show concentration is to apply a method of average bounded differences [10, Theorem 7.2]. It is based on the fact that changing the value of a single random variable (i.e., moving the position of a single vertex) has only little effect on the function (i.e., on the sum of degrees). To make sure that this is actually true, we exclude certain bad events that happen only with low probability: First, the maximum degree in $S^R_{\rho_2}$ should not be too high such that moving a single vertex can increase its degree only slightly. Second, there should not be too many vertices in $S^R_{\rho_2}$ such that the sum of degrees actually changes only for few vertices (as we do not count vertices not in $S^R_{\rho_2}$). Overall we obtain the following lemma.

▶ **Lemma 14.** *Let G be a hyperbolic random graph. For every sector S of angular width φ, the degrees of the nodes in $S^R_{\rho_2}$ sum to $\mathcal{O}(\varphi n)$ with high probability if $\varphi = \Omega(n^{1-1/\alpha} \log n)$.*

The Neigborhood of a Vertex with Radius ρ. For the second phase, we showed in Section 3.2.2 that it remains to bound the sum of degrees in the neighborhood of the corner vertices. Recall that they both have radius $\rho = 1/\alpha(\log n - \log\log n)$. Let v be a vertex with radius ρ and let A be the disk with radius R around v. Note that we already know from Section 3.2 that the maximum angular distance of neighbors of v with radius at least ρ is $\tilde{\mathcal{O}}(n^{1-1/\alpha})$. Thus, A^R_ρ is contained within a sector of this width and we can use Theorem 8 to obtain the desired bound for this part. Moreover, as in Lemma 12, we can bound the number of vertices with small radius. For all radii in between, A contains a sector of angular width $\Omega(n^{1-1/\alpha})$. It is thus not surprising that for each degree occurring in this part, the number of vertices of this degree is concentrated around its expectation. Hence, similar arguments as for Lemma 13, lead to the following lemma.

▶ **Lemma 15.** *Let G be a hyperbolic random graph and let v be a hypothetical vertex with radius $\rho = 1/\alpha(\log n - \log\log n)$ and arbitrary angular coordinate. The degrees of neighbors of v sum to $\tilde{\mathcal{O}}(n^{2-1/\alpha} + n^{1/(2\alpha)} + \delta_{\max})$ with high probability.*

5 Conclusion

To conclude, we discuss why we think that the bound $\tilde{\mathcal{O}}(n^{2-1/\alpha} + n^{1/(2\alpha)} + \delta_{\max})$ is rather tight; see Figure 4 (left) for a plot of the exponents. Clearly, the maximum degree of the graph is a lower bound, i.e., we cannot improve the $\delta_{\max}$. As $\delta_{\max} = \tilde{\Theta}(n^{1/(2\alpha)})$ holds almost surely [12], we also cannot improve below $\tilde{\mathcal{O}}(n^{1/(2\alpha)})$. For the term $n^{2-1/\alpha}$ we do not have a lower bound. Thus, the gray region in Figure 4 (left) is the only part where our bound can potentially be improved. However, by only making a single step from a vertex with radius $\rho = 1/\alpha(\log n - \log\log n)$, we can already reach vertices with angular distance $\Theta(n^{1-1/\alpha})$. Thus, it seems likely, that there exists a start-destination pair such that all vertices within a sector of this angular width are actually explored. As such a sector contains $\Theta(n^{2-1/\alpha})$ vertices, our bound seems rather tight (at least asymptotically and up to poly-logarithmic factors). For a comparison of our theoretical bound with actual search-space sizes in hyperbolic random graphs, see Figure 4.

References

1 Gregorio Alanis-Lobato, Pablo Mier, and Miguel A. Andrade-Navarro. Manifold learning and maximum likelihood estimation for hyperbolic network embedding. *Applied Network Science*, 1(10):1–14, 2016. doi:10.1007/s41109-016-0013-0.

2 Thomas Bläsius, Tobias Friedrich, and Anton Krohmer. Hyperbolic random graphs: Separators and treewidth. In *24th ESA*, volume 57 of *LIPIcs*, pages 15:1–15:16, 2016. doi:10.4230/LIPIcs.ESA.2016.15.

3 Thomas Bläsius, Tobias Friedrich, and Anton Krohmer. Cliques in hyperbolic random graphs. *Algorithmica*, 80:1–21, 2017. doi:10.1007/s00453-017-0323-3.

4 Michel Bode, Nikolaos Fountoulakis, and Tobias Müller. On the giant component of random hyperbolic graphs. In *7th EuroComb*, pages 425–429, 2013. doi:10.1007/978-88-7642-475-5_68.

5 Marián Boguñá, Fragkiskos Papadopoulos, and Dmitri Krioukov. Sustaining the internet with hyperbolic mapping. *Nature Communications*, 1:1–8, 2010. doi:10.1038/ncomms1063.

6 Michele Borassi and Emanuele Natale. KADABRA is an adaptive algorithm for betweenness via random approximation. In *24th ESA*, pages 20:1–20:18, 2016. doi:10.4230/LIPIcs.ESA.2016.20.

7 Karl Bringmann, Ralph Keusch, and Johannes Lengler. Average distance in a general class of scale-free networks with underlying geometry. *CoRR*, abs/1602.05712, 2016. URL: http://arxiv.org/abs/1602.05712.

8 Karl Bringmann, Ralph Keusch, and Johannes Lengler. Sampling geometric inhomogeneous random graphs in linear time. In *25th ESA*, volume 87 of *LIPIcs*, pages 20:1–20:15, 2017. doi:10.4230/LIPIcs.ESA.2017.20.

9 Karl Bringmann, Ralph Keusch, Johannes Lengler, Yannic Maus, and Anisur Rahaman Molla. Greedy routing and the algorithmic small-world phenomenon. In *PODC '17*, pages 371–380, 2017. doi:10.1145/3087801.3087829.

10 Devdatt P. Dubhashi and Alessandro Panconesi. *Concentration of Measure for the Analysis of Randomized Algorithms*. Cambridge University Press, 2012.

11 Tobias Friedrich and Anton Krohmer. On the diameter of hyperbolic random graphs. In *42nd ICALP*, pages 614–625, 2015. doi:10.1007/978-3-662-47666-6_49.

12 Luca Gugelmann, Konstantinos Panagiotou, and Ueli Peter. Random hyperbolic graphs: Degree sequence and clustering. In *39th ICALP*, pages 573–585, 2012. doi:10.1007/978-3-642-31585-5_51.

13 Dmitri Krioukov, Fragkiskos Papadopoulos, Maksim Kitsak, Amin Vahdat, and Marián Boguñá. Hyperbolic geometry of complex networks. *Phys. Rev. E*, 82:036106, 2010. doi:10.1103/PhysRevE.82.036106.

14 Jure Leskovec and Andrej Krevl. SNAP Datasets: Stanford large network dataset collection. http://snap.stanford.edu/data, 2014.

15 Ueli Peter. *Random Graph Models for Complex Systems*. PhD thesis, ETH Zürich, 2014.

16 Ira Sheldon Pohl. *Bi-directional and Heuristic Search in Path Problems*. PhD thesis, Stanford University, 1969.