

Geometric Inhomogeneous Random Graphs

Karl Bringmann*

Ralph Keusch†

Johannes Lengler‡

Abstract

Real-world networks, like social networks or the internet infrastructure, have structural properties such as their large clustering coefficient that can best be described in terms of an underlying geometry. This is why the focus of the literature on theoretical models for real-world networks shifted from classic models without geometry, such as Chung-Lu random graphs, to modern geometry-based models, such as hyperbolic random graphs.

With this paper we contribute to the theoretical analysis of these modern, more realistic random graph models. However, we do not directly study hyperbolic random graphs, but replace them by a more general model that we call *geometric inhomogeneous random graphs* (GIRGs). Since we ignore constant factors in the edge probabilities, our model is technically simpler (specifically, we avoid hyperbolic cosines), while preserving the qualitative behaviour of hyperbolic random graphs, and we suggest to replace hyperbolic random graphs by our new model in future theoretical studies.

We prove the following fundamental structural and algorithmic results on GIRGs. (1) We provide a sampling algorithm that generates a random graph from our model in expected linear time, improving the best-known sampling algorithm for hyperbolic random graphs by a factor $O(\sqrt{n})$, (2) we establish that GIRGs have a constant clustering coefficient, (3) we show that GIRGs have small separators, i.e., it suffices to delete a sublinear number of edges to break the giant component into two large pieces, and (4) we show how to compress GIRGs using an expected linear number of bits.

arXiv:1511.00576v2 [cs.SI] 18 Feb 2016

*Max-Planck-Institute for Informatics, Saarbrücken, Germany, kbringma@mpi-inf.mpg.de

†Institute of Theoretical Computer Science, ETH Zurich, Switzerland, rkeusch@inf.ethz.ch

‡Institute of Theoretical Computer Science, ETH Zurich, Switzerland, lenglerj@inf.ethz.ch

1 Introduction

Real-world networks, like social networks or the internet infrastructure, have structural properties that can best be described using *geometry*. For instance, in social networks two people are more likely to know each other if they live in the same region and share hobbies, both of which can be encoded as spatial information. This geometric structure may be responsible for some of the key properties of real-world networks, e.g., an underlying geometry naturally induces a large number of triangles, or large *clustering coefficient*: Two of one’s friends are likely to live in one’s region and have similar hobbies, so they are themselves similar and thus likely to know each other.

Classic mathematical models of real-world networks are *scale-free* (i.e., have a power-law degree distribution) and *small worlds* (i.e., most pairs of vertices have small graph-theoretic distance), but since they have no underlying geometry their clustering coefficient is as small as $n^{-\Omega(1)}$; this holds in particular for preferential attachment graphs [3] and Chung-Lu random graphs [19, 20, 21] (and their variants [10, 36]). In order to close this gap between the empirically observed clustering coefficient and theoretical models, much of the recent work on models for real-world networks focussed on scale-free random graph models that are equipped with an underlying geometry, such as hyperbolic random graphs [8, 37], spatial preferred attachment [2], and many others [10, 11, 12, 23, 29]. The basic properties – scale-freeness, small-world, and large clustering coefficient – have been rigorously established for most of these models. Beyond the basics, experiments suggest that these models have some very desirable properties.

In particular, hyperbolic random graphs are a promising model, as Boguñá et al. [8] computed a (heuristic) maximum likelihood fit of the internet graph into the hyperbolic random graph model and demonstrated its quality by showing that greedy routing in the underlying geometry of the fit finds near-optimal shortest paths. Further properties that have been studied on hyperbolic random graphs, mostly agreeing with empirical findings on real-world networks, are scale-freeness and clustering coefficient [28, 16], existence of a giant component [6], diameter [31, 27], average distance [1], bootstrap percolation [16], and clique number [26]. Algorithmic aspects include sampling algorithms [41] and compression schemes [40].

The main drawback of many modern models is their technical difficulty: The definition of hyperbolic random graphs involves hyperbolic sines and cosines, and proofs tend to be full of integrals involving such terms. For other models such as spatial preferred attachment, the difficulty mainly stems from the missing independence of edges. Thus, papers tend to be long and tedious.

Our contribution Our goal was to further study structural and algorithmic questions on the promising model of hyperbolic random graphs. However, it turned out to be beneficial to work with a more general model, that we introduce in this paper. In our model, which we call *geometric inhomogeneous random graph* (GIRG), every vertex v comes with a weight w_v (which we assume to follow a power law in this paper) and picks a uniformly random position \mathbf{x}_v in the d -dimensional torus* \mathbb{T}^d . Two vertices u, v then form an edge independently with probability p_{uv} , which is proportional to $w_u w_v$ and inversely proportional to some power of their distance $\|\mathbf{x}_u - \mathbf{x}_v\|$, see Section 2 for details. Our model is a geometric variant of the classic Chung-Lu random graphs, and similar in spirit to some other recent random graph models, see, e.g., [23].

A major difference to hyperbolic random graphs, which we prove to be a special case of GIRGs, is that we ignore constant factors in the edge probabilities p_{uv} . This allows to greatly simplify the edge probability expressions, thus reducing the technical overhead. In particular, proving the results of this paper directly for hyperbolic random graphs would have been much more tedious. This is why we suggest GIRGs as a replacement for hyperbolic random graphs in future theoretical studies.

The basic connectivity properties of GIRGs follow from more general considerations in [14], where we study a model of generic augmented Chung-Lu graphs containing GIRGs as a special case.

*We choose a toroidal ground space for the technical simplicity that comes with its symmetry. The results of this paper stay true if \mathbb{T}^d is replaced, say, by the d -dimensional solid unitcube.

In particular, with high probability[†] GIRGs have a giant component, polylogarithmic diameter, and doubly-logarithmic average distance. However, general studies such as [14] are limited to properties that do not depend on the specific underlying geometry. This is why we study the following properties specifically on GIRGs.

As our main result, we present a sampling algorithm that generates a random graph from our model in expected linear time. This improves the trivial sampling algorithm by a factor $O(n)$ and the best-known algorithm for hyperbolic random graphs by a factor $O(\sqrt{n})$ [41]. We also prove that our intuition is correct and the underlying geometry indeed causes GIRGs to have a constant clustering coefficient. Moreover, we show that GIRGs have small separators of expected size $n^{1-\Omega(1)}$; this is in agreement with empirical findings on real-world networks [5]. We then use the small separators to prove that GIRGs have low entropy, specifically, we show how to store a GIRG using $O(n)$ bits in expectation.

We present the details of our model and results in Section 2. After preliminaries (Section 3) and basic properties (Section 4), we prove our main result on sampling algorithms in Section 5. We show that hyperbolic random graphs are a special case of GIRGs in Section 6, analyze the clustering coefficient in Section 7, and determine instability and entropy in Section 8.

2 Model and Results

2.1 Definition of the Model

We start by defining the by-now classical Chung-Lu model and then describe the changes that yield our variant with underlying geometry.

Chung-Lu random graph For $n \in \mathbb{N}$ let $\mathbf{w} = (w_1, \dots, w_n)$ be a sequence of positive weights. We call $W := \sum_{v=1}^n w_v$ the *total weight*. The Chung-Lu random graph $G(n, \mathbf{w})$ has vertex set $V = [n] = \{1, \dots, n\}$, and two vertices $u \neq v$ are connected by an edge independently with probability $p_{uv} = \Theta\left(\min\left\{1, \frac{w_u w_v}{W}\right\}\right)$ [19, 20]. Note that the term $\min\{1, \cdot\}$ is necessary, as the product $w_u w_v$ may be larger than W . Classically, the Θ simply hides a factor 1, but by introducing the Θ the model also captures similar random graphs, like the Norros-Reittu model [36], while important properties stay asymptotically invariant.

Geometric inhomogeneous random graph (GIRG) Note that we obtain a circle by identifying the endpoints of the interval $[0, 1]$. Then the distance of $x, y \in [0, 1]$ along the circle is $|x - y|_C := \min\{|x - y|, 1 - |x - y|\}$. We fix a dimension $d \geq 1$ and use as our *ground space* the d -dimensional torus $\mathbb{T}^d = \mathbb{R}^d / \mathbb{Z}^d$, which can be described as the d -dimensional cube $[0, 1]^d$ where opposite boundaries are identified. As distance function we use the ∞ -norm on \mathbb{T}^d , i.e., for $x, y \in \mathbb{T}^d$ we define $\|x - y\| := \max_{1 \leq i \leq d} |x_i - y_i|_C$.

As for Chung-Lu graphs, we consider the vertex set $V = [n]$ and a weight sequence \mathbf{w} (in this paper we require the weights to follow a power law with exponent $\beta > 2$, see next paragraph). Additionally, for any vertex v we draw a point $\mathbf{x}_v \in \mathbb{T}^d$ uniformly and independently at random. Again we connect vertices $u \neq v$ independently with probability $p_{uv} = p_{uv}(r)$, which now depends not only on the weights w_u, w_v but also on the positions $\mathbf{x}_u, \mathbf{x}_v$, more precisely, on the distance $r = \|\mathbf{x}_u - \mathbf{x}_v\|$. We require for some constant $\alpha > 1$ the following edge probability condition:

$$p_{uv} = \Theta\left(\min\left\{\frac{1}{\|\mathbf{x}_u - \mathbf{x}_v\|^{\alpha d}} \cdot \left(\frac{w_u w_v}{W}\right)^\alpha, 1\right\}\right). \quad (\text{EP1})$$

We also allow $\alpha = \infty$ and in this case require that

$$p_{uv} = \begin{cases} \Theta(1) & \text{if } \|\mathbf{x}_u - \mathbf{x}_v\| \leq O\left(\left(\frac{w_u w_v}{W}\right)^{1/d}\right) \\ 0 & \text{if } \|\mathbf{x}_u - \mathbf{x}_v\| \geq \Omega\left(\left(\frac{w_u w_v}{W}\right)^{1/d}\right), \end{cases} \quad (\text{EP2})$$

[†]We say that an event holds *with high probability* (whp) if it holds with probability $1 - n^{-\omega(1)}$.

where the constants hidden by O and Ω do not have to match, i.e., there can be an interval $[c_1(\frac{w_u w_v}{W})^{1/d}, c_2(\frac{w_u w_v}{W})^{1/d}]$ for $\|x_u - x_v\|$ where the behaviour of p_{uv} is arbitrary. This finishes the definition of GIRGs. The free parameters of our model are $\alpha \in (1, \infty]$, $d \in \mathbb{N}$, the concrete weights w with power-law exponent $\beta > 2$ and average weight W/n , the concrete function $f_{uv}(x_u, x_v)$ replacing the Θ in p_{uv} , and for $\alpha = \infty$ the constants hidden by O, Ω in the requirement for p_{uv} . We will typically hide the constants $\alpha, d, \beta, W/n$ by O -notation.

Power-law weights As is often done for Chung-Lu graphs, in this paper we assume that the weights w follow a power law with exponent $\beta > 2$. We define this in a very general way by requiring (PL1) $w_{\min} := \min\{w_v \mid v \in V\} = \Omega(1)$ and (PL2) there exists $\bar{w} = \bar{w}(n) \geq n^{\omega(1/\log \log n)}$ such that for all constants $\eta > 0$ there are $c_1, c_2 > 0$ with

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

where the first inequality holds for all $w_{\min} \leq w \leq \bar{w}$ and the second holds for all $w \geq w_{\min}$. In particular, (PL2) implies that the average weight W/n is $\Theta(1)$. An example is the widely used weight function $w_v := \delta \cdot (n/v)^{1/(\beta-1)}$ with parameter $\delta = \Theta(1)$.

Discussion of the model The choice of the ground space \mathbb{T}^d is in the spirit of the classic random geometric graphs [39]. We prefer the torus to the hyper-cube for technical simplicity, as it yields symmetry. However, one could replace \mathbb{T}^d by $[0, 1]^d$ or any other manifold like the d -dimensional sphere; we claim that our results will still hold verbatim. Moreover, since in fixed dimension all L_p -norms on \mathbb{T}^d are equivalent (up to constant factors) and since the edge probabilities p_{uv} have a constant factor slack, our choice of the L_∞ -norm is without loss of generality (among all L_p -norms).

Our model is motivated since it generalizes hyperbolic random graphs (see Section 6). Let us nevertheless discuss why our choice of edge probabilities is *natural*: The term $\min\{\cdot, 1\}$ is necessary, as in the Chung-Lu model, because p_{uv} is a probability. To obtain a geometric model, where adjacent vertices are likely to have small distance, p_{uv} should decrease with increasing distance $\|x_u - x_v\|$, and an inverse polynomial relation seems reasonable. The constraint $\alpha > 1$ is necessary to cancel the growth of the volume of the ball of radius r proportional to r^d , so that we expect most neighbors of a vertex to lie close to it. Finally, the factor $(\frac{w_u w_v}{W})^\alpha$ ensures that the marginal probability of vertices u, v with weights w_u, w_v forming an edge is $\Pr[u \sim v] = \Theta(\min\{\frac{w_u w_v}{W}, 1\})$, as in the Chung-Lu model, and this probability does not change by more than a constant factor if we fix either x_u or x_v . This is why we see our model as a geometric variant of Chung-Lu random graphs. Note that the expected degree of a vertex $v \in V$ is $\Theta(w_v)$. The main reason why GIRGs are also *technically easy* is that for any vertex u with fixed position x_u the incident edges $\{u, v\}$ are independent. The details of these basic properties can be found in Section 4.

Sampling the weights In our definition we assume that the weight sequence w is fixed. However, if we sample the weights according to an appropriate distribution, then the sampled weights will follow a power law with probability $1 - n^{-o(1)}$, so that a model with sampled weights is almost surely included in our model. For the precise statement, see Lemma 4.6.

2.2 Properties of geometric inhomogeneous random graphs

Hyperbolic random graphs We establish that hyperbolic random graphs are a special case of GIRGs (for a formal statement see Section 6). We obtain hyperbolic random graphs from GIRGs by setting the dimension $d = 1$, the weights to a specific power law (with adjustable exponent and average weight), and the Θ in the edge probability p_{uv} to a specific, complicated function. The often studied special case of threshold hyperbolic graphs is obtained by moreover setting $\alpha = \infty$. Thus, our results on GIRGs generalize and extend the understanding of hyperbolic random graphs. Moreover, as our proofs are much less technical than typical proofs for hyperbolic random graphs, we suggest to switch from studying hyperbolic random graphs to studying GIRGs or similar models.

Connectivity properties In [14] we studied a class of generic augmented Chung-Lu random graphs with weaker assumptions on the underlying geometry than GIRGs, and we proved that GIRGs are a special case of this model (see Examples 7.1 and 7.4 in [14]). In the following we list the results of [14] transferred to GIRGs.

Theorem 2.1 (Theorem 2.1 in [14]). *Whp the degree sequence of a GIRG follows a power law with exponent β and average degree $\Theta(1)$.*

The next result determines basic connectivity properties. For $\beta > 3$, they are not well-behaved, in particular since in this case even threshold hyperbolic random graphs do not possess a giant component of linear size [7]. Hence, we restrict our attention to the regime $2 < \beta < 3$. In case $\alpha = \infty$, the following theorem requires the additional assumption $\bar{w} = \omega(n^{1/2})$.

Theorem 2.2 (Theorems 2.2 and 2.3 in [14]). *Let $2 < \beta < 3$. Whp the largest component of a GIRG has linear size and diameter $\log^{O(1)} n$, while all other components have size $\log^{O(1)} n$. Moreover, the average distance of vertices in the largest component is $(2 \pm o(1)) \frac{\log \log n}{|\log(\beta-2)|}$ in expectation and with probability $1 - o(1)$.*

We remark that most of our results in this paper crucially depend on an underlying geometry, and thus do *not* hold in the more general model from [14] where no underlying geometry is required.

Sampling Sampling algorithms, that generate a random graph from a fixed distribution, are known for Chung-Lu random graphs and others, running in expected linear time [4, 35]. As our main result, we present such an algorithm for GIRGs. This greatly improves the trivial $O(n^2)$ sampling algorithm (throwing a biased coin for each possible edge), as well as an algorithm for threshold hyperbolic random graphs with expected time $O(n^{3/2})$ [41]. It allows to run experiments on much larger graphs than the ones with $\approx 10^4$ vertices in [8]. In addition to our model assumptions, here we assume that the Θ in our requirement on p_{uv} is sufficiently explicit, i.e., we can compute p_{uv} exactly and we know a constant $c > 0$ such that replacing Θ by c yields an upper bound on p_{uv} , see Section 5 for details.

Theorem 2.3 (Section 5). *There is an algorithm for sampling a GIRG in expected time $O(n)$.*

Clustering In social networks, two friends of the same person are likely to also be friends with each other. This property of having many triangles is captured by the clustering coefficient, defined as the probability when choosing a random vertex v and two random neighbors $v_1 \neq v_2$ of v that v_1 and v_2 are adjacent (if v does not have two neighbors then its contribution to the clustering coefficient is 0). While Chung-Lu random graphs have a very small clustering coefficient of $n^{-\Omega(1)}$, it is easy to show that the clustering coefficient of GIRGs is constant. This is in accordance with empirical data of real-world networks [25] and the constant clustering coefficient of hyperbolic random graphs determined in [16, 28, 40].

Theorem 2.4 (Section 7). *Whp the clustering coefficient of a GIRG is $\Theta(1)$.*

Stability For real-world networks, a key property to analyze is their stability under attacks. It has been empirically observed that many real-world networks have small separators of size n^c , $c < 1$ [5]. In contrast, Chung-Lu random graphs are unrealistically stable, since any deletion of $o(n)$ nodes or edges reduces the size of the giant component by at most $o(n)$ [10]. We show that GIRGs agree with the empirical results much better. Specifically, if we cut the ground space \mathbb{T}^d into two halves along one of the axes then we roughly split the giant component into two halves, but the number of edges passing this cut is quite small, namely $n^{1-\Omega(1)}$. Thus, GIRGs are prone to (quite strong) adversarial attacks, just as many real-world networks. Furthermore, their small separators are useful for many algorithms, e.g., the compression scheme of the next paragraph. We remark that stability was not studied for (threshold) hyperbolic random graphs before.

Theorem 2.5 (Section 8). *Almost surely it suffices to delete $n^{\max\{2-\alpha, 3-\beta, 1-1/d\}+o(1)}$ edges of a GIRG to split its giant component into two parts of linear size each.*

Since we assume $\alpha > 1$, $\beta > 2$, and $d = \Theta(1)$, the number of deleted edges is indeed $n^{1-\Omega(1)}$.

Entropy The internet graph has empirically been shown to be well compressible, using less than one bit per edge [5, 9]. This is not the case for the Chung-Lu model, as its entropy is $\Theta(n \log n)$ [18]. We show that GIRGs have linear entropy, as is known for threshold hyperbolic random graphs [40].

Theorem 2.6 (Section 8). *We can store a GIRG using $O(n)$ bits in expectation. This compression allows to query the degree of any vertex and its i -th neighbor in time $O(1)$.*

3 Preliminaries and Notation

3.1 Notation

For $w \in \mathbb{R}_{\geq 0}$, we use the notation $V_{\geq w} := \{v \in V \mid w_v \geq w\}$ and $V_{\leq w} := \{v \in V \mid w_v \leq w\}$, as well as $W_{\geq w} := \sum_{v \in V_{\geq w}} w_v$ and $W_{\leq w} := \sum_{v \in V_{\leq w}} w_v$. For $u, v \in V$ we write $u \sim v$ if u and v are adjacent, and for $A, B \subseteq V$ we write $A \sim v$ if there exists $u \in A$ such that $u \sim v$, and we write $A \sim B$ if there exists $v \in B$ such that $A \sim v$. For a vertex $v \in V$, we denote its neighborhood by $\Gamma(v)$, i.e. $\Gamma(v) := \{u \in V \mid u \sim v\}$. We say that an event holds *with high probability* (whp) if it holds with probability $1 - n^{-\omega(1)}$.

3.2 Tools

Le Cam's theorem allows us to bound the total variation distance of a binomial distribution to a Poisson distribution with the same mean.

Theorem 3.1 (Le Cam, Proposition 1 in [34]). *Suppose X_1, \dots, X_n are independent Bernoulli random variables s.t. $\Pr[X_i = 1] = p_i$ for $i \in [n]$, $\lambda_n = \sum_{i \in [n]} p_i$ and $S_n = \sum_{i \in [n]} X_i$. Then*

$$\sum_{k=0}^{\infty} \left| \Pr[S_n = k] - \frac{\lambda_n^k e^{-\lambda_n}}{k!} \right| < 2 \sum_{i=1}^n p_i^2.$$

In particular, if $\lambda_n = \Theta(1)$ and $\max_{i \in [n]} p_i = o(1)$, then $\Pr[S_n = k] = \Theta(1)$ for $k = O(1)$.

We will need a concentration inequality which bounds large deviations taking into account some bad event \mathcal{B} . We will use the following theorem.

Theorem 3.2 (Theorem 3.3 in [14]). *Let X_1, \dots, X_m be independent random variables over $\Omega_1, \dots, \Omega_m$. Let $X = (X_1, \dots, X_m)$, $\Omega = \prod_{k=1}^m \Omega_k$ and let $f: \Omega \rightarrow \mathbb{R}$ be measurable with $0 \leq f(\omega) \leq M$ for all $\omega \in \Omega$. Let $\mathcal{B} \subseteq \Omega$ such that for some $c > 0$ and for all $\omega \in \overline{\mathcal{B}}, \omega' \in \overline{\mathcal{B}}$ that differ in at most two components we have*

$$|f(\omega) - f(\omega')| \leq c. \tag{1}$$

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

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

Cells Consider the ground space \mathbb{T}^d , split it into 2^d equal cubes, and repeat this process with each created cube; we call the resulting cubes *cells*. Cells are cubes of the form $C = [x_1 2^{-\ell}, (x_1 + 1) 2^{-\ell}] \times \dots \times [x_d 2^{-\ell}, (x_d + 1) 2^{-\ell}]$ with $\ell \geq 0$ and $0 \leq x_i < 2^\ell$. We represent cell C by the tuple (ℓ, x_1, \dots, x_d) . The volume of C is $\text{VOL}(C) = 2^{-\ell \cdot d}$. For $0 < x \leq 1$ we let $\lceil x \rceil_{2^d}$ be the smallest number larger or equal to x that is realized as the volume of a cell, or in other words x rounded up to a power of 2^d , $\lceil x \rceil_{2^d} = \min\{2^{-\ell \cdot d} \mid \ell \in \mathbb{N}_0: 2^{-\ell \cdot d} \geq x\}$. Note that the cells of a fixed level ℓ partition the ground space. We obtain a *geometric ordering* of these cells by following the recursive construction of cells in a depth-first-search manner, yielding the following lemma.

Lemma 3.3 (Geometric ordering). *There is an enumeration of the cells $C_1, \dots, C_{2^{\ell d}}$ of level ℓ such that for every cell C of level $\ell' < \ell$ the cells of level ℓ contained in C form a consecutive block C_i, \dots, C_j in the enumeration.*

Proof. We construct the geometric ordering by induction on the level ℓ . For $\ell = 0$ there is only one cell to enumerate, so let $\ell > 0$. Given an enumeration $C_1, \dots, C_{2^{(\ell-1)d}}$ of the cells of level $\ell - 1$, we first enumerate all cells of level ℓ contained in C_1 , starting with the cell which is smallest in all d coordinates, and ending with the cell which is largest in all d coordinates. Then we enumerate all cells of level ℓ contained in C_2 (starting with smallest coordinates, and ending with largest coordinates), and so on. Evidently this gives us a geometric ordering of the cells of level ℓ . \square

4 Basic Properties

In this section, we list some basic properties about GIRGs which repeatedly occur in our proofs. In particular we consider the expected degree of a vertex and the marginal probability that an edge between two vertices with given weights is present. The proofs of all statements follow from more general considerations and can be found in [14]. Let us start with the following abstract statement.

Lemma 4.1 (Lemma 4.1 in [14]). *Let $f : \mathbb{R} \rightarrow \mathbb{R}$ be a continuously differentiable function. Then for any weights $0 \leq w_0 \leq w_1$,*

$$\sum_{v \in V, w_0 \leq w_v \leq w_1} f(w_v) = f(w_0) \cdot |V_{\geq w_0}| - f(w_1) \cdot |V_{> w_1}| + \int_{w_0}^{w_1} f'(w) \cdot |V_{\geq w}| dw.$$

Recall the assumptions on power-law weights in Section 2.1. In the next lemma we calculate the partial weight sums $W_{\leq w}$ and $W_{\geq w}$.

Lemma 4.2 (Lemma 4.2 in [14]). *The total weight satisfies $W = \Theta(n)$. Moreover, for all sufficiently small $\eta > 0$,*

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

Next we consider the marginal edge probability of two vertices u, v with weights w_u, w_v . In GIRGs, this probability is essentially the same as in Chung-Lu random graphs. Furthermore, the marginal probability does not change by more than a constant factor if we fix the position \mathbf{x}_u or \mathbf{x}_v (but not both!). Moreover, conditioned on a fixed position $\mathbf{x}_v \in \mathbb{T}^d$, all edges $\{u, v\}$ are independently present. This is a central feature of our model.

Lemma 4.3 (Lemma 4.3 in [14]). *Fix $u \in [n]$ and $\mathbf{x}_u \in \mathbb{T}^d$. All edges $\{u, v\}$, $u \neq v$, are independently present with probability*

$$\Pr[u \sim v \mid \mathbf{x}_u] = \Theta(\Pr[u \sim v]) = \Theta\left(\min\left\{1, \frac{w_u w_v}{W}\right\}\right).$$

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

Lemma 4.4 (Lemma 4.4 in [14]). *For any $v \in [n]$ we have $\mathbb{E}[\deg(v)] = \Theta(\mathbf{w}_v)$.*

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

Lemma 4.5 (Lemma 4.5 in [14]). *The following properties hold whp for all $v \in [n]$.*

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

We conclude this section by proving that if we sample the weights randomly from an appropriate distribution, then almost surely the resulting weights satisfy our conditions on power-law weights.

Lemma 4.6 (Lemma 4.6 in [14]). *Let $\mathbf{w}_{\min} = \Theta(1)$ and $F = F_n : \mathbb{R} \rightarrow [0, 1]$ be non-decreasing such that $F(z) = 0$ for all $z \leq \mathbf{w}_{\min}$, and $F(z) = 1 - \Theta(z^{1-\beta})$ for all $z \in [\mathbf{w}_{\min}, n^{1/(\beta-1-\varepsilon)}]$, where $\varepsilon > 0$. Suppose that for every vertex $v \in [n]$, we choose the weight \mathbf{w}_v independently according to the cumulative probability distribution F . Then asymptotically almost surely the resulting weight vector \mathbf{w} satisfies the power-law conditions (PL1) and (PL2) with $\bar{w} = (n/\log^2 n)^{1/(\beta-1)}$. Moreover, for any fixed function $1 \geq \lambda(n) \geq n^{-o(1)}$ the error probability is bounded by $\lambda(n)$ for sufficiently large n .*

5 Sampling Algorithm

In this section we show that GIRGs can be sampled in expected time $O(n)$. The running time depends exponentially on the (fixed) dimension d . In addition to our model assumptions, in this section we require that (1) edge probabilities p_{uv} can be computed in constant time (given any vertices u, v and positions $\mathbf{x}_u, \mathbf{x}_v$) and (2) we know an explicit constant $c > 0$ such that if $\alpha < \infty$ we have

$$p_{uv} \leq \min \left\{ c \frac{1}{\|\mathbf{x}_u - \mathbf{x}_v\|^{\alpha d}} \cdot \left(\frac{\mathbf{w}_u \mathbf{w}_v}{W} \right)^\alpha, 1 \right\},$$

and if $\alpha = \infty$ we have

$$p_{uv} \leq \begin{cases} 1 & \text{if } \|\mathbf{x}_u - \mathbf{x}_v\| < c \left(\frac{\mathbf{w}_u \mathbf{w}_v}{W} \right)^{1/d} \\ 0 & \text{otherwise.} \end{cases}$$

Note that existence of c follows from our model assumptions. In the remainder of this section we introduce building blocks of our algorithm (Section 5.1) and present our algorithm (Section 5.2) and its analysis (Section 5.3).

5.1 Building Blocks

Data structures Recall the definition of cells from Section 3. We first build a basic data structure on a set of points P that allows to access the points in a given cell C (of volume at least ν) in constant time.

Lemma 5.1. *Given a set of points P and $0 < \nu \leq 1$, in time $O(|P| + 1/\nu)$ we can construct a data structure $\mathcal{D}_\nu(P)$ supporting the following queries in time $O(1)$:*

- given a cell C of volume at least ν , return $|C \cap P|$,

- given a cell C of volume at least ν and a number k , return the k -th point in $C \cap P$ (in a fixed ordering of $C \cap P$ depending only on P and ν).

Proof. Let $\mu = \lceil \nu \rceil_{2^d} = 2^{-\ell-d}$, so that $\nu \leq \mu \leq O(\nu)$. Following the recursive construction of cells, we can determine a geometric ordering of the cells of volume μ as in Lemma 3.3 in time $O(1/\mu) = O(1/\nu)$; say $C_1, \dots, C_{1/\mu}$ are the cells of volume μ in the geometric ordering. We store this ordering by storing a pointer from each cell $C_i = (\ell, x_1, \dots, x_d)$ to its successor $C_{i+1} = (\ell, x'_1, \dots, x'_d)$, which allows to scan the cells $C_1, \dots, C_{1/\mu}$ in linear time. For any point $x \in P$, using the floor function we can determine in time $O(1)$ the cell (ℓ, x_1, \dots, x_d) of volume μ that x belongs to. This allows to determine the numbers $|C_i \cap P|$ for all i in time $O(|P| + 1/\nu)$. We also compute each prefix sum $s_i := \sum_{j < i} |C_j \cap P|$ and store it at cell $C_i = (\ell, x_1, \dots, x_d)$. Using an array $A[\cdot]$ of size $|P|$, we store (a pointer to) the k -th point in $C_i \cap P$ at position $A[s_i + k]$. Note that this preprocessing can be performed in time $O(|P| + 1/\nu)$.

A given cell C of volume at least ν may consist of several cells of volume μ . By Lemma 3.3, these cells form a contiguous subsequence $C_i, C_{i+1}, \dots, C_{j-1}, C_j$ of $C_1, \dots, C_{1/\mu}$, so that the points $C \cap P$ form a contiguous subsequence of A . For constant access time, we store for each cell C of volume at least ν the indices s_C, e_C of the first and last point of $C \cap P$ in A . Then $|C \cap P| = e_C - s_C + 1$ and the k -th point in $C \cap P$ is stored at $A[s_C + k]$. Thus, both queries can be answered in constant time. Note that the ordering $A[\cdot]$ of the points in $C \cap P$ is a mix of the geometric ordering of cells of volume μ and the given ordering of P within a cell of volume μ , in particular this ordering indeed only depends on P and ν . \square

Next we construct a partitioning of $\mathbb{T}^d \times \mathbb{T}^d$ into products of cells $A_i \times B_i$. This partitioning allows to split the problem of sampling the edges of a GIRG into one problem for each $A_i \times B_i$, which is beneficial, since each product $A_i \times B_i$ has one of two easy types. For any $A, B \subseteq \mathbb{T}^d$ we denote the distance of A and B by $d(A, B) = \inf_{a \in A, b \in B} \|a - b\|$.

Lemma 5.2. *Let $0 < \nu \leq 1$. In time $O(1/\nu)$ we can construct a set $\mathcal{P}_\nu = \{(A_1, B_1), \dots, (A_s, B_s)\}$ such that*

- (1) A_i, B_i are cells with $\text{VOL}(A_i) = \text{VOL}(B_i) \geq \nu$,
- (2) for all i , either $d(A_i, B_i) = 0$ and $\text{VOL}(A_i) = \lceil \nu \rceil_{2^d}$ (type I) or $d(A_i, B_i) \geq \text{VOL}(A_i)^{1/d}$ (type II),
- (3) the sets $A_i \times B_i$ partition $\mathbb{T}^d \times \mathbb{T}^d$,
- (4) $s = O(1/\nu)$.

Proof. Note that for cells A, B of equal volume we have $d(A, B) = 0$ if and only if either $A = B$ or (the boundaries of) A and B touch. For a cell C of level ℓ we let $\text{par}(C)$ be its *parent*, i.e., the unique cell of level $\ell - 1$ that C is contained in. Let $\mu = \lceil \nu \rceil_{2^d}$. We define \mathcal{P}_ν as follows. For any pair of cells (A, B) with $\text{VOL}(A) = \text{VOL}(B) \geq \nu$, we add (A, B) to \mathcal{P}_ν if either (i) $\text{VOL}(A) = \text{VOL}(B) = \mu$ and $d(A, B) = 0$, or (ii) $d(A, B) > 0$ and $d(\text{par}(A), \text{par}(B)) = 0$.

Property (1) follows by definition. Regarding property (2), the pairs (A, B) added in case (i) are clearly of type I. Observe that two cells A, B of equal volume that are not equal or touching have distance at least the sidelength of A , which is $\text{VOL}(A)^{1/d}$. Thus, in case (ii) the lower bound $d(A, B) > 0$ implies $d(A, B) \geq \text{VOL}(A)^{1/d}$, so that (A, B) is of type II.

For property (3), consider $(x, y) \in \mathbb{T}^d \times \mathbb{T}^d$ and let A, B be the cells of volume μ containing x, y . Let $A^{(0)} := A$ and $A^{(i)} := \text{par}(A^{(i-1)})$ for any $i \geq 1$, until $A^{(k)} = \mathbb{T}^d$. Similarly, define $B = B^{(0)} \subset \dots \subset B^{(k)} = \mathbb{T}^d$ and note that $\text{VOL}(A^{(i)}) = \text{VOL}(B^{(i)})$. Observe that each set $A^{(i)} \times B^{(i)}$ contains (x, y) . Moreover, any set $A' \times B'$, where A', B' are cells with $\text{VOL}(A') = \text{VOL}(B')$ and $(x, y) \in A' \times B'$, is of the form $A^{(i)} \times B^{(i)}$. Thus, to show that \mathcal{P}_ν partitions $\mathbb{T}^d \times \mathbb{T}^d$ we need to show that it contains exactly one of the pairs $(A^{(i)}, B^{(i)})$ (for any x, y). To show this, we use the monotonicity $d(A^{(i)}, B^{(i)}) \geq d(A^{(i+1)}, B^{(i+1)})$ and consider two cases. If $d(A, B) = 0$ then we add (A, B) to \mathcal{P}_ν in case (i), and we add no further $(A^{(i)}, B^{(i)})$, since $d(A^{(i)}, B^{(i)}) = 0$ for all i .

If $d(A, B) > 0$ then since $d(A^{(k)}, B^{(k)}) = d(\mathbb{T}^d, \mathbb{T}^d) = 0$ there is a unique index $0 \leq i < k$ with $d(A^{(i)}, B^{(i)}) > 0$ and $d(A^{(i+1)}, B^{(i+1)}) = 0$. Then we add $(A^{(i)}, B^{(i)})$ in case (ii) and no further $(A^{(j)}, B^{(j)})$. This proves property (3).

Property (4) follows from the running time bound of $O(1/\nu)$, which we show in the following. Note that we can enumerate all $1/\mu = O(1/\nu)$ cells of volume μ , and all of the at most $3^d = O(1)$ touching cells of the same volume, in time $O(1/\nu)$, proving the running time bound for case (i). Moreover, we can enumerate all $2^{\ell-d}$ cells C in level ℓ , together with all of the at most $3^d = O(1)$ touching cells C' in the same level. Then we can enumerate all $2^d = O(1)$ cells A that have C as parent as well as all $O(1)$ cells B that have C' as parent. This enumerates (a superset of) all possibilities of case (ii). Summing the running time $O(2^{\ell-d})$ over all levels ℓ with volume $2^{-\ell-d} \geq \nu$ yields a total running time of $O(1/\nu)$. \square

Weight layers We set $w_0 := w_{\min}$ and $w_i := 2w_{i-1}$ for $i \geq 1$. This splits the vertex set $V = [n]$ into *weight layers* $V_i := \{v \in V \mid w_{i-1} \leq v < w_i\}$ for $1 \leq i \leq L$ with $L = O(\log n)$. We write V_i^C for the restriction of weight layer V_i to cell C , $V_i^C := \{v \in V_i \mid x_v \in C\}$.

Geometric random variates For $0 < p \leq 1$ we write $\text{Geo}(p)$ for a geometric random variable, taking value $i \geq 1$ with probability $p(1-p)^{i-1}$. $\text{Geo}(p)$ can be sampled in constant time using the simple formula $\lceil \frac{\log(R)}{\log(1-p)} \rceil$, where R is chosen uniformly at random in $(0, 1)$, see [24][‡].

5.2 The Algorithm

Algorithm 1 Sampling algorithm for GIRGs in expected time $O(n)$

```

1:  $E := \emptyset$ 
2: sample the positions  $x_v, v \in V$ , and determine the weight layers  $V_i$ 
3: for all  $1 \leq i \leq L$  do build data structure  $\mathcal{D}_{\nu(i)}(\{x_v \mid v \in V_i\})$  with  $\nu(i) := \frac{w_i w_0}{W}$ 
4: for all  $1 \leq i \leq j \leq L$  do
5:   construct partitioning  $\mathcal{P}_{\nu(i,j)}$  with  $\nu(i,j) := \frac{w_i w_j}{W}$ 
6:   for all  $(A, B) \in \mathcal{P}_{\nu(i,j)}$  of type I do
7:     for all  $u \in V_i^A$  and  $v \in V_j^B$  do with probability  $p_{uv}$  add edge  $\{u, v\}$  to  $E$ 
8:   for all  $(A, B) \in \mathcal{P}_{\nu(i,j)}$  of type II do
9:      $\bar{p} := \min \left\{ c \cdot \frac{1}{d(A,B)^{\alpha d}} \cdot \left(\frac{w_i w_j}{W}\right)^\alpha, 1 \right\}$ 
10:     $r := \text{Geo}(\bar{p})$ 
11:    while  $r \leq |V_i^A| \cdot |V_j^B|$  do
12:      determine the  $r$ -th pair  $(u, v)$  in  $V_i^A \times V_j^B$ 
13:      with probability  $p_{uv}/\bar{p}$  add edge  $\{u, v\}$  to  $E$ 
14:       $r := r + \text{Geo}(\bar{p})$ 
15:   if  $i = j$  then remove all edges with  $u > v$  sampled in this iteration

```

We first assume $\alpha < \infty$ and discuss the simpler case $\alpha = \infty$ in Section 5.4. Given the model parameters, our Algorithm 1 samples the edge set E of a GIRG. To this end, we first sample all vertex positions x_v uniformly at random in \mathbb{T}^d . Given weights w_1, \dots, w_n we can determine the weight layers V_i in linear time (we may use counting sort or bucket sort since there are only $L = O(\log n)$ layers). Then we build the data structure from Lemma 5.1 for the points in V_i setting $\nu = \nu(i) = \frac{w_i w_0}{W}$, i.e., we build $\mathcal{D}_{\nu(i)}(\{x_v \mid v \in V_i\})$ for each i . In the following, for each pair of weight layers V_i, V_j we sample the edges between V_i and V_j . To this end, we construct the partitioning $\mathcal{P}_{\nu(i,j)}$ from Lemma 5.2 with $\nu(i,j) = \frac{w_i w_j}{W}$. Since $\mathcal{P}_{\nu(i,j)}$ partitions $\mathbb{T}^d \times \mathbb{T}^d$, every

[‡]To evaluate this formula exactly in time $O(1)$ we need to assume the RealRAM model of computation. However, also on a bounded precision machine like the WordRAM $\text{Geo}(p)$ can be sampled in expected time $O(1)$ [13].

pair of vertices $u \in V_i, v \in V_j$ satisfies $x_u \in A, x_v \in B$ for exactly one $(A, B) \in \mathcal{P}_{\nu(i,j)}$. Thus, we can iterate over all $(A, B) \in \mathcal{P}_{\nu(i,j)}$ and sample the edges between V_i^A and V_j^B .

If (A, B) is of type I, then we simply iterate over all vertices $u \in V_i^A$ and $v \in V_j^B$ and add the edge $\{u, v\}$ with probability p_{uv} ; this is the trivial sampling algorithm. Note that we can efficiently enumerate V_i^A and V_j^B using the data structure $\mathcal{D}_{\nu(i)}(\{x_v \mid v \in V_i\})$ that we constructed above.

If (A, B) is of type II, then the distance $\|x - y\|$ of any two points $x \in A, y \in B$ satisfies $d(A, B) \leq \|x - y\| \leq d(A, B) + \text{VOL}(A)^{1/d} + \text{VOL}(B)^{1/d} \leq 3d(A, B)$, by the definition of type II. Thus, $\bar{p} = \min \left\{ c \cdot \frac{1}{d(A, B)^{\alpha d}} \cdot \left(\frac{w_i w_j}{W} \right)^\alpha, 1 \right\}$ is an upper bound on the edge probability p_{uv} for any $u \in V_i^A, v \in V_j^B$, and it is a good upper bound since $d(A, B)$ is within a constant factor of $\|x_u - x_v\|$ and w_i, w_j are within constant factors of w_u, w_v . Now we first sample the set of edges \bar{E} between V_i^A and V_j^B that we would obtain if all edge probabilities were equal to \bar{p} , i.e., any $(u, v) \in V_i^A \times V_j^B$ is in \bar{E} independently with probability \bar{p} . From this set \bar{E} , we can then generate the set of edges with respect to the true edge probabilities p_{uv} by throwing a coin for each $\{u, v\} \in \bar{E}$ and letting it survive with probability p_{uv}/\bar{p} . Then in total we choose a pair (u, v) as an edge in E with probability $\bar{p} \cdot (p_{uv}/\bar{p}) = p_{uv}$, proving that we sample from the correct distribution. Note that here we used $p_{uv} \leq \bar{p}$. It is left to show how to sample the ‘‘approximate’’ edge set \bar{E} . First note that the data structure $\mathcal{D}_{\nu}(\{x_v \mid v \in V_i\})$ defines an ordering on V_i^A , and we can determine the ℓ -th element in this ordering in constant time, similarly for V_j^B . Using the lexicographic ordering, we obtain an ordering on $V_i^A \times V_j^B$ for which we can again determine the ℓ -th element in constant time. In this ordering, the first pair $(u, v) \in V_i^A \times V_j^B$ that is in \bar{E} is geometrically distributed, according to $\text{Geo}(\bar{p})$. Since geometric random variates can be generated in constant time, we can efficiently generate \bar{E} , specifically in time $O(1 + |\bar{E}|)$.

Finally, the case $i = j$ is special. With the algorithm described above, for any $u, v \in V_i$ we sample whether they form an edge twice, once for $x_u \in A, x_v \in B$ (for some $(A, B) \in \mathcal{P}_{\nu(i,j)}$) and once for $x_v \in A', x_u \in B'$ (for some $(A', B') \in \mathcal{P}_{\nu(i,j)}$). To fix this issue, in the case $i = j$ we only accept a sampled edge $(u, v) \in V_i^A \times V_j^B$ if $u < v$; then only one way of sampling edge $\{u, v\}$ remains. This changes the expected running time only by a constant factor.

5.3 Analysis

Correctness of our algorithm follows immediately from the above explanations. In the following we show that Algorithm 1 runs in expected linear time. This is clear for lines 1-2. For line 3, since building the data structure from Lemma 5.1 takes time $O(|P| + 1/\nu)$, it takes total time $\sum_{i=1}^L O(|V_i| + W/(w_i w_0))$. Clearly, the first summand $|V_i|$ sums up to n . Using $w_0 = w_{\min} = \Omega(1)$, $W = O(n)$, and that w_i grows exponentially with i , implying $\sum_i 1/w_i = O(1)$, also the second summand sums up to $O(n)$. For line 5, all invocations in total take time $O(\sum_{i,j} W/(w_i w_j))$, which is $O(n)$, since again $W = O(n)$ and $\sum_i 1/w_i = O(1)$. We claim that for any weight layers V_i, V_j the expected running time we spend on any $(A, B) \in \mathcal{P}_{\nu(i,j)}$ is $O(1 + \mathbb{E}[|E_{i,j}^{A,B}|])$, where $E_{i,j}^{A,B}$ is the set of edges in $V_i^A \times V_j^B$. Summing up the first summand $O(1)$ over all $(A, B) \in \mathcal{P}_{\nu(i,j)}$ sums up to $1/\nu(i, j) = W/(w_i w_j)$. As we have seen above, this sums up to $O(n)$ over all i, j . Summing up the second summand $O(\mathbb{E}[|E_{i,j}^{A,B}|])$ over all $(A, B) \in \mathcal{P}_{\nu(i,j)}$ and weight layers V_i, V_j yields the total expected number of edges $O(\mathbb{E}[|E|])$, which is $O(n)$, since the average weight $W/n = O(1)$ and thus the expected average degree is constant.

It is left to prove the claim that for any weight layers V_i, V_j the expected time spent on $(A, B) \in \mathcal{P}_{\nu(i,j)}$ is $O(1 + \mathbb{E}[|E_{i,j}^{A,B}|])$. If (A, B) is of type I, then any pair of vertices $(u, v) \in V_i^A \times V_j^B$ has probability $\Theta(1)$ to form an edge: Since the volume of A and B is $w_i w_j / W$, their diameter is $(w_i w_j / W)^{1/d}$ and we obtain $\|x_u - x_v\| \leq (w_i w_j / W)^{1/d} = O((w_u w_v / W)^{1/d})$, which yields $p_{uv} = \Theta\left(\min \left\{ \left(\frac{w_u w_v}{\|x_u - x_v\|^{\alpha d} W} \right)^\alpha, 1 \right\}\right) = \Theta(1)$. As we spend time $O(1)$ for any $(u, v) \in V_i^A \times V_j^B$, we stay in the desired running time bound $O(\mathbb{E}[|E_{i,j}^{A,B}|])$.

If (A, B) is of type II, we first sample edges \bar{E} with respect to the larger edge probability \bar{p} , and then for each edge $e \in \bar{E}$ sample whether it belongs to E . This takes total time $O(1 + |\bar{E}|)$.

Note that any edge $e \in \bar{E}$ has constant probability $p_{uv}/\bar{p} = \Theta(1)$ to survive: It follows from $w_u = \Theta(w_i)$, $w_v = \Theta(w_j)$, and $\|x_u - x_v\| = \Theta(d(A, B))$ that $p_{uv} = \Theta(\bar{p})$. Hence, we obtain $\mathbb{E}[|\bar{E}|] = O(\mathbb{E}[|E_{i,j}^{A,B}|])$, and the running time $O(1 + |\bar{E}|)$ is in expectation bounded by $O(1 + \mathbb{E}[|E_{i,j}^{A,B}|])$. This finishes the proof of the claim.

5.4 Case $\alpha = \infty$

For $\alpha = \infty$, edges only exist between vertices in distance $\|x_u - x_v\| < c(w_u w_v / W)^{1/d}$. We change Algorithm 1 by setting $\nu(i, j) = \max\{1, c\}^d \cdot w_i w_j / W$. Then for any $u \in V_i, v \in V_j$ and $(A, B) \in \mathcal{P}_{\nu(i,j)}$ of type II we have $d(A, B) \geq \text{VOL}(A)^{1/d} \geq \nu^{1/d} \geq c(w_u w_v / W)^{1/d}$, so there are no edges between V_i^A and V_j^B for type II. This allows to simplify the algorithm by completely ignoring type II pairs; the rest of the algorithm stays unchanged.

Additionally, we have to slightly change the running time analysis, since it no longer holds that all pairs of vertices $(u, v) \in V_i^A \times V_j^B$ satisfy $p_{uv} = \Theta(1)$. However, a variant of this property still holds: If we only uncovered that $x_u \in A$ and $x_v \in B$, but not yet where exactly in A, B they lie, then the marginal probability of (u, v) forming an edge is $\Theta(1)$, since for any $\varepsilon > 0$ a constant fraction of all pairs of points in $A \times B$ are within distance $\varepsilon(w_{i-1} w_{j-1} / W)^{1/d}$, guaranteeing edge probability $\Theta(1)$ for sufficiently small ε . This again allows to check all pairs of vertices in $V_i^A \times V_j^B$ whether they form an edge, which yields expected linear running time.

6 Comparison with Hyperbolic Random Graphs

In this section we show that hyperbolic random graphs are a special case of GIRGs. We start by defining hyperbolic random graphs. This model was first introduced by Krioukov et al. [32] and has attracted a lot of attention during the last years. As underlying geometry it uses the hyperbolic plane. There exist several different representations of hyperbolic geometry, all with advantages and disadvantages. For introducing this random graph model, it is most convenient to use the native representation. It can be described by a disk H of radius R around the origin 0 , where the position of every point x is given by its polar coordinates (r_x, θ_x) . The model is isotropic around the origin. The hyperbolic distance between two points x and y is given by the non-negative solution $d = d(x, y)$ of the equation

$$\cosh(d) = \cosh(r_x) \cosh(r_y) - \sinh(r_x) \sinh(r_y) \cos(\phi_x - \phi_y). \quad (2)$$

In the following definition, we follow the notation introduced by Gugelmann et al. [28].

Definition 6.1. *Let $\alpha_H > 0, C_H \in \mathbb{R}, T_H > 0, n \in \mathbb{N}$, and set $R = 2 \log n + C_H$. Then the random hyperbolic graph $G_{\alpha_H, C_H, T_H}(n)$ is a graph with vertex set $V = [n]$ and the following properties:*

- *Every vertex $v \in [n]$ independently draws random coordinates (r_v, ϕ_v) , where the angle π_v is chosen uniformly at random in $[0, 2\pi)$ and the radius $r_v \in [0, R]$ is random with density $f(r) := \frac{\alpha_H \sinh(\alpha_H r)}{\cosh(\alpha_H R) - 1}$.*
- *Every potential edge $e = \{u, v\}$, $u, v \in [n]$, is independently present with probability*

$$p_H(d(u, v)) = \left(1 + e^{\frac{1}{2T_H}(d(u,v)-R)}\right)^{-1}.$$

In the limit $T_H \rightarrow 0$, we obtain the threshold hyperbolic random graph $G_{\alpha_H, C_H}(n)$, where every edge $e = \{u, v\}$ is present if and only if $d(u, v) \leq R$.

We will show that hyperbolic random graphs are almost surely contained in our general framework. To this end, we embed the disk of the native hyperbolic model into our model with dimension 1, hence we reduce the geometry of the hyperbolic disk to the geometry of a circle, but gain additional freedom as we can choose the weights of vertices. Notice that a single point on the hyperbolic

disk has measure zero, so we can assume that no vertex has radius $r_v = 0$. For the parameters, we put

$$d := 1, \quad \beta := 2\alpha_H + 1, \quad \alpha := 1/T_H.$$

Furthermore, we define the mapping

$$\mathbf{w}_v := e^{\frac{R-r_v}{2}} \quad \text{and} \quad \mathbf{x}_v := \frac{\phi_v}{2\pi}.$$

Since this is a bijection between $H \setminus \{0\}$ and $[1, e^{R/2}) \times \mathbb{T}^1$, there exists as well an inverse function $g(\mathbf{w}_u, \mathbf{x}_u) = (r_u, \phi_u)$. Finally for any two vertices $u \neq v$ on the torus, we set

$$p_{uv} := p_H(d(g(\mathbf{w}_u, \mathbf{x}_u), g(\mathbf{w}_v, \mathbf{x}_v))).$$

This finishes our embedding. The following lemma, which we prove near the end of this section, demonstrates that under this mapping almost surely the weights will follow a power law.

Lemma 6.2. *Let $\alpha_H > \frac{1}{2}$. Then almost surely the induced weight sequence \mathbf{w} follows a power law with parameter $\beta = 2\alpha_H + 1$. Moreover, for any fixed function $1 \geq \lambda(n) \geq n^{-o(1)}$ the error probability is bounded by $\lambda(n)$ for sufficiently large n .*

Now we come to the main statement of this section. In the following we assume that if we sample an instance of the hyperbolic random graph model, we first sample the radii, then the angles and at last the edges.

Theorem 6.3. *Let $\alpha_H > \frac{1}{2}$, $n \in \mathbb{N}$ and fix a set of radii $(r_1, \dots, r_n) \in [0, R]^n$ inducing a power-law weight sequence \mathbf{w} with parameter $\beta = 2\alpha_H + 1$. Then the random positions \mathbf{x}_u and the edge probabilities $p_{uv}(\mathbf{x}_u, \mathbf{x}_v)$ produced by our mapping satisfy the properties of the GIRG model, i.e., for fixed radii inducing power-law weights, hyperbolic random graphs are a special case of GIRGs.*

Note that the precondition of Theorem 6.3 holds almost surely by Lemma 6.2. Therefore an instance of random hyperbolic graphs is almost surely included in our GIRG model with parameters as set above. In particular, any property that holds with probability $1 - q$ for GIRGs also holds for hyperbolic random graphs with probability at least, say, $1 - q - n^{1/\log \log \log n}$. Before proving Lemma 6.2 and Theorem 6.3, we consider the following basic property of hyperbolic random graphs.

Lemma 6.4. *Let $\alpha_H > \frac{1}{2}$. Then with probability $1 - n^{-\Omega(1)}$ every vertex has radius at least $r_0 := (1 - \frac{1}{2\alpha_H}) \log n$. Furthermore, for all $r = \omega(1)$, $r \leq R$ and $v \in V$, we have*

$$\Pr[r_v \leq r] = e^{-\alpha_H(R-r)}(1 + o(1)).$$

Proof. Let $v \in V$. By the given density f it follows immediately that

$$\Pr[r_v \leq r] = \int_0^r f(x) dx = \alpha_H \int_0^r \frac{\sinh(\alpha_H x)}{\cosh(\alpha_H R) - 1} dx = \frac{\cosh(\alpha_H r) - 1}{\cosh(\alpha_H R) - 1} = e^{-\alpha_H(R-r)}(1 + o(1)), \quad (3)$$

where we used $\cosh(x) = \frac{e^x + e^{-x}}{2} = \frac{e^x}{2}(1 + o(1))$ whenever $x = \omega(1)$. Now let X_{r_0} be the random variable counting the vertices of radius at most r_0 . We observe that (3) implies

$$\mathbb{E}[X_{r_0}] = n e^{-\alpha_H(R-r_0)}(1 + o(1)) = e^{-\alpha_H C_H} n^{1/2 - \alpha_H} (1 + o(1)) = n^{-\Omega(1)}.$$

By Markov's inequality, we have $X_{r_0} = 0$ with probability $1 - n^{-\Omega(1)}$. □

Proof of Lemma 6.2. For every vertex of the random hyperbolic graph, the radius is chosen independently and uniformly according to $f(r)$. Hence under the mapping g , we sample weights independently. We will prove that we fulfil the prerequisites of Theorem 4.6. Let $0 < \varepsilon < 1$. By

Lemma 6.4, the probability that a vertex v has radius at most $r \geq \varepsilon \log n$ is $\Theta(e^{-\alpha_H(R-r)})$. Let $1 \leq z \leq o(n^{1-\varepsilon/2})$. Then $R - 2 \log z \geq \varepsilon \log n$, and

$$\begin{aligned} F(z) &:= \Pr[\mathbf{w}_v \leq z] = 1 - \Pr[r_v \leq R - 2 \log z] = 1 - \Theta(e^{-2\alpha_H \log z}) \\ &= 1 - \Theta(z^{-2\alpha_H}) = 1 - \Theta(z^{1-\beta}). \end{aligned}$$

Furthermore, for $z < 1$ we get

$$F(z) := \Pr[\mathbf{w}_v \leq z] = \Pr[r_v \geq R - 2 \log z] = 0.$$

Clearly, F is non-decreasing and therefore satisfies the preconditions of Theorem 4.6 with $\mathbf{w}_{\min} = 1$. It follows from this theorem that almost surely the weight sequence \mathbf{w} follows a power law with parameter β , and the error probability is at most $\lambda(n)$. \square

Proof of Theorem 6.3. Let us start by considering the sampling process of a random hyperbolic graph. First we sample the radii of the vertices, for which the precondition of the theorem assumes that they induce a power-law weight sequence. Secondly we sample the angles, which in our transformation correspond to coordinates chosen independently and uniformly at random on \mathbb{T}^1 . It remains to prove that p_{uv} as defined above satisfies conditions (EP1) and (EP2).

Let $u \neq v$ be two vertices of the random hyperbolic graph with coordinates (r_u, ϕ_u) and (r_v, ϕ_v) and consider their mappings $(\mathbf{x}_u, \mathbf{x}_v)$ and $(\mathbf{w}_u, \mathbf{w}_v)$. Since the hyperbolic model is isotropic around the origin, we can assume without loss of generality that $r_u \geq r_v$, $\phi_v = 0$ and $\phi_u \leq \pi$.

Let us first consider the threshold model, corresponding to $\alpha = \infty$. We claim that there exist constants $M > m > 0$ such that whenever $\|\mathbf{x}_u - \mathbf{x}_v\| \geq M \frac{\mathbf{w}_u \mathbf{w}_v}{W}$, then $p_{uv} = 0$, and whenever $\|\mathbf{x}_u - \mathbf{x}_v\| \leq m \frac{\mathbf{w}_u \mathbf{w}_v}{W}$, then $p_{uv} = 1$. This will imply (EP2), as we set $d = 1$. Recall that in the threshold model, two vertices u and v are connected if and only if $d(u, v) \leq R$. When $r_u + r_v \leq R$, this is the case for all angles ϕ_u and ϕ_v . Otherwise, for $\phi = 0$ and $\phi_u \leq \pi$, the distance between u and v is increasing in ϕ_u and there exists a critical value ϕ_0 such that $d((r_u, \phi_u), (r_v, 0)) \leq R$ if and only if $\phi_u \leq \phi_0$. The following lemma estimates ϕ_0 .

Lemma 6.5 (Lemma 3.1 in [28]). *Let $0 \leq r_u \leq R$, $r_u + r_v \geq R$ and assume $\phi_v = 0$. Then*

$$\phi_0 = 2e^{\frac{R-r_u-r_v}{2}} (1 + \Theta(e^{R-r_u-r_v})).$$

Suppose $\|\mathbf{x}_u - \mathbf{x}_v\| \geq M \frac{\mathbf{w}_u \mathbf{w}_v}{W}$. Notice that by our transformation we have $\|\mathbf{x}_u - \mathbf{x}_v\| = \frac{\phi_u}{2\pi}$ and

$$\frac{\mathbf{w}_u \mathbf{w}_v}{W} = \Theta\left(\frac{\mathbf{w}_u \mathbf{w}_v}{n}\right) = \Theta\left(\frac{e^{R-(r_u+r_v)/2}}{n}\right) = \Theta(e^{(R-r_u-r_v)/2}), \quad (4)$$

where we used $W = \Theta(n)$ by Lemma 4.2. Hence, we have $\frac{\phi_u}{2\pi M} = \Omega(e^{(R-r_u-r_v)/2})$. Since $\phi_u \leq 1$, this implies $r_u + r_v > R$, if we choose the constant M sufficiently large. Moreover, for sufficiently large n we obtain $\phi_u > 2e^{\frac{R-r_u-r_v}{2}} (1 + \Theta(e^{R-r_u-r_v}))$. Thus, by Lemma 6.5 the two vertices u and v are not connected and indeed $p_{uv} = 0$.

On the other hand, assume $\|\mathbf{x}_u - \mathbf{x}_v\| \leq m \frac{\mathbf{w}_u \mathbf{w}_v}{W}$. Then either $r_u + r_v < R$ and thus $\{u, v\} \in E$ follows directly, or $r_u + r_v \geq R$ and $\phi_u < 2e^{\frac{R-r_u-r_v}{2}} (1 + \Theta(e^{R-r_u-r_v}))$. In the second case, Lemma 6.5 implies $p_{uv} = 1$.

We now turn to the case $\alpha < \infty$. By the identity $\cosh(x \pm y) = \cosh(x) \cosh(y) \pm \sinh(x) \sinh(y)$ and our assumptions on ϕ_u and ϕ_v we can rewrite (2) as

$$\cosh(d) = \cosh(r_u - r_v) + (1 - \cos(\phi_u)) \sinh(r_u) \sinh(r_v). \quad (5)$$

Next we observe that $\cosh(x) = \Theta(e^{|x|})$ for all x and $\sinh(x) = \Theta(e^x)$ for all $x = \omega(1)$. Observe that (PL2) and $\mathbf{w}_v = e^{(R-r_v)/2}$ imply $r_v = \Theta(\log n)$ for all vertices v . Furthermore, we perform a Taylor approximation of $1 - \cos(\phi_u)$ around 0 and get $1 - \cos(\phi_u) = \frac{\phi_u^2}{2} - \frac{\phi_u^4}{24} + \dots = \Theta(\phi_u^2)$, as

ϕ_u is at most a constant. Combining these observations with (5) and the assumption $r_u \geq r_v$, we deduce

$$e^{d-R} = \Theta(\cosh(d)e^{-R}) = \Theta(e^{r_u-r_v-R} + \phi_u^2 e^{r_u+r_v-R}). \quad (6)$$

In the condition (EP1) on p_{uv} the minimum is obtained by the second term whenever $\|x_u - x_v\| \leq \frac{w_u w_v}{W}$. Mapping u and v to the hyperbolic disk, this implies $\phi_u = O(e^{(R-r_u-r_v)/2})$. We claim that whenever $\phi_u = O(e^{(R-r_u-r_v)/2})$, the two vertices u and v are connected with constant probability and therefore $p_{uv} = \Theta(1)$. Indeed, in this case by (6) we have $e^{d-R} = O(1)$, and using Definition 6.1 we deduce

$$p_{uv} = p_H(d(u, v)) = \left(1 + (e^{d-R})^{(1/(2T_H))}\right)^{-1} = \Theta(1).$$

On the other hand, suppose $\|x_u - x_v\| \geq \frac{w_u w_v}{W}$, which implies $\phi_u = \Omega(e^{(R-r_u-r_v)/2})$. In this case by (6) we have $e^{d-R} = \Theta(\phi_u^2 e^{r_u+r_v-R}) = \Omega(1)$. However, if $e^{d-R} = \Omega(1)$, we can use Definition 6.1 and (4) to obtain

$$\begin{aligned} p_{uv} &= \left(1 + e^{\frac{1}{2T}(d-R)}\right)^{-1} = \Theta\left(e^{-\frac{1}{2T_H}(d-R)}\right) = \Theta\left(\left(\phi_u^{-1} e^{(R-r_u-r_v)/2}\right)^{1/T_H}\right) \\ &= \Theta\left(\left(\frac{1}{\|x_u - x_v\|} \cdot \frac{w_u w_v}{W}\right)^{1/T_H}\right) = \Theta\left(\left(\frac{1}{\|x_u - x_v\|} \cdot \frac{w_u w_v}{W}\right)^\alpha\right). \end{aligned}$$

This finishes the case $\alpha < \infty$ and thus the proof. \square

7 Clustering

Before proving Theorem 2.4, we give the formal definition of the clustering coefficient.

Definition 7.1. *In a graph $G = (V, E)$ the clustering coefficient of a vertex $v \in V$ is defined as*

$$\text{CC}(v) := \text{CC}_G(v) := \begin{cases} \#\{\{u, u'\} \subseteq \Gamma(v) \mid u \sim u'\} / \binom{\deg v}{2}, & \text{if } \deg(v) \geq 2 \\ 0, & \text{otherwise,} \end{cases}$$

and the (mean) clustering coefficient of G is

$$\text{CC}(G) := \frac{\sum_{v \in V} \text{CC}(v)}{|V|}.$$

Proof of Theorem 2.4. Let $V' := V_{\leq n^{1/8}}$ and $G' = G[V']$. We first show that for the subgraph G' we have $\mathbb{E}[\text{CC}(G')] = \Omega(1)$. Let $w_0 = \Theta(1)$ be a weight such that there are linearly many vertices with weight at most w_0 . Since $\text{CC}(G') = \frac{1}{|V'|} \sum_{v \in V'} \text{CC}_{G'}(v) = \Theta\left(\frac{1}{n} \sum_{v \in V'} \text{CC}_{G'}(v)\right)$, it suffices to show that a vertex v of weight at most w_0 fulfills $\mathbb{E}[\text{CC}_{G'}(v)] = \Omega(1)$. For this we consider the set $\bar{V} := V_{\leq w_0}$ of vertices of weight at most w_0 . Fix such a vertex v at position $x_v \in \mathbb{T}^d$, and let $U(v)$ be the ball around x_v with radius $cn^{-1/d}$ for a sufficiently small constant $c > 0$. Clearly the volume of this ball is $\Theta(n^{-1})$. Thus, the expected number of vertices in \bar{V} with position in $U(v)$ is $\Theta(1)$. Consider the event $\mathcal{E} = \mathcal{E}(v)$ that the following three properties hold.

- (i) v has at least two neighbors in \bar{V} with positions in $U(v)$.
- (ii) v does not have neighbors in \bar{V} with positions in $\mathbb{T}^d \setminus U(v)$.
- (iii) v does not have neighbors in $[n] \setminus \bar{V}$.

We claim that $\Pr[\mathcal{E}] = \Theta(1)$. For (i), note that the expected number of vertices in \bar{V} with position in $U(v)$ is $\Theta(1)$. Since the position of every vertex is independent, by Le Cam's theorem (Theorem 3.1) the probability that there are at least two vertices in \bar{V} with position in $U(v)$ is $\Theta(1)$. Moreover, by (EP1) and (EP2), for each such vertex the probability to connect to v is $\Theta(1)$ (in the case $\alpha = \infty$

this follows because we have chosen c small enough), so (i) holds with probability $\Theta(1)$. For (ii), for any vertex $u \in \bar{V} \setminus \{v\}$, we can bound

$$\Pr[v \sim u, \mathbf{x}_v \in \mathbb{T}^d \setminus U(v)] \leq \Pr[v \sim u] = \Theta(1/n).$$

Hence, by Le Cam's theorem, (ii) holds with probability $\Theta(1)$, and this probability can only increase if we condition on (i). Finally, for every fixed position x , (iii) holds independently of (i) or (ii) with probability $\Theta(1)$, again by Le Cam's theorem. This proves the claim that $\Pr[\mathcal{E}] = \Theta(1)$.

Conditioned on \mathcal{E} , let v_1 and v_2 be two random neighbors of v . Then $\mathbf{x}_{v_1}, \mathbf{x}_{v_2} \in U(v)$, and $\mathbf{w}_{v_1}, \mathbf{w}_{v_2} \leq w_0$. By the triangle inequality we obtain $\|\mathbf{x}_{v_1} - \mathbf{x}_{v_2}\| \leq 2cn^{-1/d}$. For c sufficiently small, we deduce from (EP1) and (EP2) that $v_1 \sim v_2$ with probability $\Theta(1)$. Thus we have shown that $\mathbb{E}[\text{CC}_{G'}(v) \mid \mathcal{E}(v)] = \Omega(1)$ for all $v \in \bar{V}$. Since $\Pr[\mathcal{E}(v)] = \Theta(1)$, this proves $\mathbb{E}[\text{CC}_{G'}(v)] = \Omega(1)$ for all $v \in \bar{V}$, which implies $\mathbb{E}[\text{CC}(G')] = \Omega(1)$.

Next we show that $\text{CC}(G')$ is concentrated. We aim to do this via an Azuma-type inequality with error event, as given in Theorem 3.2. Note that in our graph model, we apply two different randomized processes to create the geometric graph. First, for every vertex v we choose \mathbf{x}_v independently at random. Afterwards, every edge is present with some probability p_{uv} . Recall that we can apply the concentration bound only if all random variables are independent, which is not the case so far.

The n random variables $\mathbf{x}_1, \dots, \mathbf{x}_n$ define the vertex set and the edge probabilities p_{uv} . We introduce a second set of $n - 1$ independent random variables. For every $u \in \{2, \dots, n\}$ we let $Y_u := (Y_u^1, \dots, Y_u^{u-1})$, where every Y_u^v is independently and uniformly at random from $[0, 1]$. Then for $v < u$, we include the edge $\{u, v\}$ in the graph if and only if

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

We observe that indeed this implies $\Pr[u \sim v \mid \mathbf{x}_u, \mathbf{x}_v] = p_{uv}(\mathbf{x}_u, \mathbf{x}_v)$ as desired. Furthermore, the $2n - 1$ random variables $\mathbf{x}_1, \dots, \mathbf{x}_n, Y_2, \dots, Y_n$ are independent and define a probability space Ω . Then G, G' and $\text{CC}(G')$ are all random variables on Ω . Consider the following bad event:

$$\mathcal{B} := \{\omega \in \Omega : \text{the maximum degree in } G'(\omega) \text{ is at least } n^{1/4}\}. \quad (7)$$

We observe that $\Pr[\mathcal{B}] = n^{-\omega(1)}$, since whp every vertex $v \in V'$ has degree at most $O(w_v + \log^2 n) = o(n^{1/4})$ by Lemma 4.5. Let $\omega, \omega' \in \bar{\mathcal{B}}$ such that they differ in at most two coordinates. We observe that changing one coordinate \mathbf{x}_i or Y_i can influence only the local clustering coefficients of i itself and of the vertices which are neighbors of i either before or after the coordinate change. Unless \mathcal{B} holds, every vertex in G' has degree at most $n^{1/4}$ and therefore every coordinate of the probability space has effect at most $2n^{1/4}/n$ onto $\text{CC}(G')$. Thus, we obtain $|\text{CC}(G'(\omega)) - \text{CC}(G'(\omega'))| \leq 4n^{-3/4}$. We apply Theorem 3.2 with $t = n^{-1/8}$ and $c := 4n^{-3/4}$ and deduce

$$\Pr[|\text{CC}(G') - \mathbb{E}[\text{CC}(G')]| \geq t] \leq 2e^{-\Omega(t^2/n^{1-3/2})} + n^{O(1)} \Pr[\mathcal{B}] = n^{-\omega(1)},$$

where we used $\frac{t^2}{n^{1-3/2}} = n^{1/4}$. Hence, we have $\text{CC}(G') = (1 + o(1))\mathbb{E}[\text{CC}(G')] = \Omega(1)$ whp.

In order to compare $\text{CC}(G)$ with $\text{CC}(G')$, we observe that every additional edge $e = \{u, v\}$ which we add to G' can decrease only $\text{CC}(u)$ and $\text{CC}(v)$, both by at most one. Thus,

$$\text{CC}(G) \geq \frac{|V'|}{n} \text{CC}(G') - \frac{2}{n} \sum_{v \in V \setminus V'} \deg(v).$$

By Lemma 4.5, whp $\frac{2}{n} \sum_{v \in V \setminus V'} \deg(v) = \Theta(n^{-1}W_{>n^{1/8}}) = o(1)$. Together with $|V'| = \Theta(n)$, this concludes the argument and proves that $\text{CC}(G) = \Omega(1)$ whp. \square

8 Stability of the Giant, Entropy, and Compression Algorithm

In this section we prove Theorems 2.6 and 2.5. More precisely, we show that whp the graph (and its giant) has separators of sublinear size, and we make use of these small separators to devise a compression algorithm that can store the graph using a linear number of bits in expectation. Note that the compression maintains only the graph up to isomorphism, not the underlying geometry. The main idea is to enumerate the vertices in an ordering that reflects the geometry, and then storing for each vertex i the differences $i - j$ for all neighbors j of i . We start with a technical lemma that gives the number of edges intersecting an axis-parallel, regular grid. (For $\gamma > 0$ with $1/\gamma \in \mathbb{N}$, the axis-parallel, regular grid with side length γ is the union of all $d - 1$ -dimensional hyperplanes that are orthogonal to an axis and that are in distance $k\gamma$ from the origin for a $k \in \mathbb{Z}$.) Both the existence of small separators and the efficiency of the compression algorithm follow easily from that formula.

Lemma 8.1. *Let $\eta > 0$. Let $1 \leq \mu \leq n^{1/d}$ be an integer, and consider an axis-parallel, regular grid with side length $1/\mu$ on \mathbb{T}^d . Then in expectation the grid intersects at most $O(n \cdot (n/\mu^d)^{2-\beta+\eta} + (n^{2-\alpha}\mu^{d(\alpha-1)} + n^{1-1/d}\mu)(1 + \log(n/\mu^d)))$ edges.*

Proof. For $u, v \in V$, let ρ_{uv} be the probability that the edge uv exists and cuts the grid. Let $r_{\max} := 1/2$ be the diameter of \mathbb{T}^d . We write

$$\rho_{uv} = \int_0^{r_{\max}} \Pr[\|x_u - x_v\| = r] \cdot p_{uv}(r) \cdot \Pr[x_u, x_v \text{ in different cells of } \mu\text{-grid}] dr. \quad (8)$$

Observe that u and v have distance r with probability density $\Pr[\|x_u - x_v\| = r] = O(r^{d-1})$. Furthermore, setting $\gamma_{uv} := \min\{(w_u w_v / W)^{1/d}, r_{\max}\}$ we have

$$p_{uv}(r) = \begin{cases} \Theta(1) & \text{if } r \geq \gamma_{uv} \\ \Theta((\gamma_{uv}/r)^{\alpha d}) & \text{otherwise} \end{cases}$$

Additionally, in the case $\alpha = \infty$, by increasing γ_{uv} by at most a constant factor we may assume $p_{uv}(r) = 0$ for all $r \geq \gamma_{uv}$. For the last term in (8), for a fixed axis of \mathbb{T}^d consider the hyperplanes $\{h_i\}_{1 \leq i \leq \mu}$ of the grid perpendicular to that axis. If the edge $e = uv$ has length $\|x_u - x_v\| = r$, then after a random shift along the axis, the edge e intersects one of the h_i with probability at most $\min\{\mu r, 1\}$. By symmetry of the underlying space, a random shift does not change the probability to intersect one of the h_i , so any edge of length r has probability at most $\min\{\mu r, 1\}$ to intersect one of the h_i . By the union bound over all (constantly many) axes, the probability for u, v to lie in different cells of the grid is $O(\min\{\mu r, 1\})$.

Now we distinguish several cases. For $\gamma_{uv} > 1/\mu$ and $\alpha < \infty$, we may estimate

$$\rho_{uv} \leq O\left(\int_0^{1/\mu} r^{d-1} \cdot \mu r dr + \int_{1/\mu}^{\gamma_{uv}} r^{d-1} dr + \underbrace{\int_{\gamma_{uv}}^{r_{\max}} r^{d-1-d\alpha} \gamma_{uv}^{d\alpha} dr}_{=O(\gamma_{uv}^d), \text{ since } d-d\alpha < 0}\right) \leq O(\mu^{-d} + \gamma_{uv}^d) \leq O(\gamma_{uv}^d). \quad (9)$$

For $\gamma_{uv} > 1/\mu$ and $\alpha = \infty$, equation (9) remains true, except that the third integral is replaced by 0 by our choice of γ_{uv} . So in this case we still get $\rho_{uv} \leq O(\gamma_{uv}^d)$.

The case $\gamma_{uv} \leq 1/\mu$ is a bit more complicated. Again we consider first $\alpha < \infty$. Then we may bound

$$\rho_{uv} \leq O\left(\underbrace{\int_0^{\gamma_{uv}} r^{d-1} \cdot \mu r dr}_{=: I_1} + \underbrace{\int_{\gamma_{uv}}^{1/\mu} r^{d-1} \cdot \mu r \cdot r^{-d\alpha} \gamma_{uv}^{d\alpha} dr}_{=: I_2} + \underbrace{\int_{1/\mu}^{r_{\max}} r^{d-1-d\alpha} \gamma_{uv}^{d\alpha} dr}_{=: I_3}\right). \quad (10)$$

Similarly as before, $I_1 \leq O(\gamma_{uv}^{d+1}\mu)$ and $I_3 \leq O(\mu^{d\alpha-d}\gamma_{uv}^{d\alpha})$. Note that both terms are bounded from above by $O((\gamma_{uv}\mu)^{d\tilde{\alpha}}\mu^{-d})$, where $\tilde{\alpha} := \min\{\alpha, 1 + 1/d\}$, since $\gamma_{uv}\mu \leq 1$. For I_2 , the inverse derivative of $r^{d-d\alpha}$ is either $\Theta(r^{1+d-d\alpha})$, or $\log r$, or $-\Theta(r^{1+d-d\alpha})$, depending on whether $1+d-d\alpha$ is positive, zero, or negative, respectively. Therefore, we obtain

$$I_2 \leq \begin{cases} O(\gamma_{uv}^{d\alpha}\mu^{d\alpha-d}) & = O((\gamma_{uv}\mu)^{d\alpha}\mu^{-d}), & \text{if } d-d\alpha > -1 \\ O(\gamma_{uv}^{d\alpha}\mu(\log(1/\mu) - \log(\gamma_{uv}))) & = O((\gamma_{uv}\mu)^{d+1}\mu^{-d}|\log(\gamma_{uv}\mu)|), & \text{if } d-d\alpha = -1 \\ O(\gamma_{uv}^{d+1}\mu) & = O((\gamma_{uv}\mu)^{d+1}\mu^{-d}), & \text{if } d-d\alpha < -1. \end{cases}$$

In particular, we can bound all terms (including I_1 and I_3) in a unified way by $O((\gamma_{uv}\mu)^{d\tilde{\alpha}}\mu^{-d})(1 + |\log(\gamma_{uv}\mu)|)$. Moreover, since $\gamma_{uv} \geq (w_{\min}^2/W)^{1/d} = \Omega(n^{-1/d})$, the second factor is bounded by $O(1 + \log(n^{1/d}/\mu)) = O(1 + \log(n/\mu^d))$. Also, in the case $\alpha = \infty$ the same calculation applies, except that I_2 and I_3 are replaced by 0. Note that naturally $\tilde{\alpha} = 1 + 1/d$ for $\alpha = \infty$. So altogether we have shown that

$$\rho_{uv} \leq \begin{cases} O(\gamma_{uv}^d) & \text{if } \gamma_{uv} \geq 1/\mu, \\ O((\gamma_{uv}\mu)^{d\tilde{\alpha}}\mu^{-d}(1 + \log(n^d/\mu))) & \text{if } \gamma_{uv} \leq 1/\mu. \end{cases}$$

Therefore, the expected number of edges intersecting the grid is in $O(S_1 + S_2)$, where

$$S_1 := \sum_{u,v \in V, \gamma_{uv} > 1/\mu} \gamma_{uv}^d \quad \text{and} \quad S_2 := \sum_{u,v \in V, \gamma_{uv} \leq 1/\mu} (\gamma_{uv}\mu)^{d\tilde{\alpha}}\mu^{-d}(1 + \log(n^d/\mu)).$$

Let $0 < \eta' < \eta < \beta - 2$ be (sufficiently small) constants. Then we may use the power-law assumption (PL2), Lemma 4.2, and Lemma 4.1 to bound S_1 :

$$\begin{aligned} S_1 &\leq \sum_{u,v \in V, w_u w_v > W/\mu^d} \frac{w_u w_v}{W} = \sum_{u \in V} \frac{w_u}{W} \cdot W_{\geq W/(\mu^d w_u)} \stackrel{4.2}{\leq} O\left(\sum_{u \in V} \frac{w_u}{W} \cdot n \left(\frac{W}{\mu^d w_u}\right)^{2-\beta+\eta}\right) \\ &\leq O\left(\left(\frac{\mu^d}{n}\right)^{\beta-2-\eta} \sum_{u \in V} w_u^{\beta-1-\eta}\right) \stackrel{4.1}{=} O\left(\left(\frac{\mu^d}{n}\right)^{\beta-2-\eta} \int_{w_{\min}}^{\infty} n w^{1-\beta+\eta'} w^{\beta-2-\eta} dw\right) \\ &= O\left(n \cdot \left(\mu^d/n\right)^{\beta-2-\eta}\right). \end{aligned}$$

To tackle S_2 , we again use Lemma 4.1, let $\lambda_u := W/(w_u \mu^d)$ and obtain

$$S'_2 := \sum_{\substack{u,v \in V \\ \gamma_{uv} \leq 1/\mu}} (\gamma_{uv})^{d\tilde{\alpha}} = \sum_{u \in V} \sum_{v \in V_{\leq \lambda_u}} \left(\frac{w_u w_v}{W}\right)^{\tilde{\alpha}} \stackrel{4.1}{\leq} O\left(\sum_{u \in V} \left(\frac{w_u}{n}\right)^{\tilde{\alpha}} \int_{w_{\min}}^{\lambda_u} n w^{1-\beta+\eta} w^{\tilde{\alpha}-1} dw\right) \quad (11)$$

Now we distinguish two cases, because the integral behaves differently for exponents larger or smaller than -1 . If $\tilde{\alpha} \geq \beta - 1$, then for $0 < \eta' < \eta$ equation (11) evaluates to

$$\begin{aligned} S'_2 &\leq O\left(\sum_{u \in V} \left(\frac{w_u}{n}\right)^{\tilde{\alpha}} n \lambda_u^{1+\tilde{\alpha}-\beta+\eta}\right) = O\left(\frac{n^{2-\beta+\eta}}{\mu^{d(1+\tilde{\alpha}-\beta+\eta)}} \sum_{u \in V} w_u^{\beta-1-\eta}\right) \\ &\stackrel{4.1}{\leq} O\left(\frac{n^{2-\beta+\eta}}{\mu^{d(1+\tilde{\alpha}-\beta+\eta)}} \int_{w_{\min}}^{\infty} n w^{1-\beta+\eta'} w^{\beta-2-\eta} dw\right) = O\left(n \frac{n^{2-\beta+\eta}}{\mu^{d(1+\tilde{\alpha}-\beta+\eta)}}\right). \end{aligned}$$

Therefore, $S_2 = \mu^{d\tilde{\alpha}-d}(1 + \log(n^d/\mu))S'_2 \leq O(n \cdot (n/\mu^d)^{2-\beta+\eta})$, which is one of the terms in the lemma. On the other hand, if $\tilde{\alpha} < \beta - 1$ then for $0 < \eta < \beta - \tilde{\alpha} - 1$ we obtain from (11)

$$S_2 \leq O\left(n^{1-\tilde{\alpha}} \sum_{u \in V} w_u^{\tilde{\alpha}}\right) \stackrel{4.1}{\leq} O\left(n^{1-\tilde{\alpha}} \int_{w_{\min}}^{\infty} n w^{1-\beta+\eta} w^{\tilde{\alpha}-1} dw\right) \leq O(n^{2-\tilde{\alpha}}),$$

and again $S_2 = \mu^{d\tilde{\alpha}-d}(1 + \log(n^d/\mu))S'_2$ corresponds to terms in the lemma after plugging in $\tilde{\alpha}$. This concludes the proof. \square

From Lemma 8.1, we immediately obtain that there is a sublinear set of vertices that disconnects the giant component.

Proof of Theorem 2.5. By Lemma 8.1 for $\mu = 2$, there are $m = O(n^{\max\{3-\beta, 2-\alpha, 1-1/d\}+\eta})$ edges intersecting a grid of side length $1/2$ in expectation, and two hyperplanes of this grid suffice to split \mathbb{T}^d into two halves. Whp there are $\Omega(n)$ vertices in each grid cell, and whp the weights of the vertices in each half satisfy a power law. If $2 < \beta < 3$ then whp each halfspace gives rise to a giant component of linear size, this follows from more general considerations in [14]. Hence, whp the two hyperplanes split the giant of G into two parts of linear size, although almost surely they only intersect $n^{1-\Omega(1)}$ edges. Finally, since the bound $m = O(n^{\max\{3-\beta, 2-\alpha, 1-1/d\}+\eta})$ holds for all $\eta > 0$ we may conclude that it also holds with η replaced by $o(1)$. \square

Compression algorithm With Lemma 8.1 at hand, we are ready to give a compression algorithm that stores the graph with $O(n)$ bits, i.e., with $O(1)$ bits per edge, proving Theorem 2.6. We remark that our result does not directly follow from the general compression scheme on graphs with small separators in [5], since our graphs only have small separators in expectation, in particular, small subgraphs of size $O(\sqrt{\log n})$ can form expanders and thus not have small separators. However, our algorithm loosely follows their algorithm as well as the practical compression scheme of [9], see also [17].

We first enumerate the vertices as follows. Recall the definition of cells from Section 3, and consider all cells of volume $2^{-\ell_0 d}$, where $\ell_0 := \lfloor \log n/d \rfloor$. Note that the boundaries of these cells induce a grid as in Lemma 8.1. Since each such cell has volume $\Theta(1/n)$, the expected number of vertices in each cell is constant. We fix a geometric ordering of these cells as in Lemma 3.3, and we enumerate the vertices in the order of the cells, breaking ties (between vertices in the same cell) arbitrarily. For the rest of the section we will assume that the vertices are enumerated in this way, i.e., we identify $V = [n]$, where $i \in [n]$ refers to the vertex with index i .

Having enumerated the vertices, for each vertex $i \in [n]$ we store a block of $1 + \deg(i)$ sub-blocks. The first sub-block consists of a single dummy bit (to avoid empty sequences arising from isolated vertices). In the other $\deg(i)$ sub-blocks we store the differences $i - j$ using $\log_2 |i - j| + O(1)$ bits, where j runs through all neighbors of i . We assume that the information for all vertices is stored in a big successive block B in the memory. Moreover, we create two more blocks B_V and B_E of the same length. Both B_V and B_E have a one-bit whenever the corresponding bit in B is the first bit of the block of a vertex, and B_E has also a one-bit whenever the corresponding bit in B is the first bit of an edge (i.e., the first bit encoding a difference $i - j$). All other bits in B_V and B_E are zero.

It is clear that with the data above the graph is determined. To handle queries efficiently, we replace B_V and B_E each with a rank/select data structure. This data structure allows to handle in constant time queries of the form “Rank(b)”, which returns the number of one-bits up to position b , and “Select(i)”, which returns the position of the i -th one-bit [30, 22, 38]. Given $i, s \in \mathbb{N}$, we can find the index of the s -th neighbor of i in constant time by Algorithm 2, and the degree of i by Algorithm 3. In particular, it is also possible for Algorithm 2 to first check whether $s \leq \deg(i)$.

Algorithm 2 Finding the s -th neighbor of vertex i

- | | |
|---|---|
| 1: $b := \text{SELECT}(i, B_V)$ | ▷ starting position of vertex i |
| 2: $k := \text{RANK}(b, B_E)$ | ▷ number of edges and vertices before b |
| 3: $b_1 := \text{SELECT}(k + s, B_E)$ | ▷ starting position of s -th edge of vertex i |
| 4: $b_2 := \text{SELECT}(k + s + 1, B_E)$ | ▷ bit after ending position of s -th edge of vertex i |
| 5: return $B[b_1 : b_2 - 1]$ | ▷ block that stores s -th edge of vertex i |
-

We need to show that the data structure needs $O(n)$ bits in expectation. There are n dummy bits, so we must show that we require $O(n)$ bits to store all differences $i - j$, where ij runs through all edges of the graph. We need $\log_2 |i - j| + O(1)$ bits for each edge, and the $O(1)$ terms sum up to $O(|E|)$, which is $O(n)$ in expectation. Thus, it remains to prove the following.

Algorithm 3 Finding the degree of vertex i

- | | |
|---|---|
| 1: $b := \text{SELECT}(i, B_V)$ | \triangleright starting position of vertex i |
| 2: $b' := \text{SELECT}(i + 1, B_V)$ | \triangleright starting position of vertex $i + 1$ |
| 3: $\Delta := \text{RANK}(b', B_E) - \text{RANK}(b, B_E)$ | \triangleright block in B_E contains $\deg(i) + 1$ one-bits |
| 4: return $\Delta - 1$ | |
-

Lemma 8.2. *Let the vertices in V be enumerated by the geometric ordering. Then,*

$$\mathbb{E} \left[\sum_{ij \in E} \log(|i - j|) \right] = O(n). \quad (12)$$

Proof. We abbreviate the expectation in (12) by R . Note that the geometric ordering puts all the vertices that are in the same cell of a $2^{-\ell}$ -grid in a consecutive block, for all $1 \leq \ell \leq \ell_0$. Therefore, if $e = ij$ does not intersect the $2^{-\ell}$ -grid then $|i - j| \leq \#\{\text{vertices in the cell of } e\}$. For $1 \leq \ell \leq \ell_0$, let \mathcal{E}_ℓ be the set of edges intersecting the $2^{-\ell}$ -grid. For convenience, let $\mathcal{E}_0 := \emptyset$, and let $\mathcal{E}_{\ell_0+1} := E$ be the set of all edges. Then, using concavity of log in the third step,

$$\begin{aligned} R &\leq \mathbb{E} \left[\sum_{\ell=0}^{\ell_0} \sum_{e=ij \in \mathcal{E}_{\ell+1} \setminus \mathcal{E}_\ell} \log(\#\{\text{vertices in the cell of } e\}) \right] \\ &= \sum_{\ell=0}^{\ell_0} \sum_{u < v} \Pr[uv \in \mathcal{E}_{\ell+1} \setminus \mathcal{E}_\ell] \mathbb{E}[\log(\#\{\text{vertices in the cell of } u\}) \mid uv \in \mathcal{E}_{\ell+1} \setminus \mathcal{E}_\ell] \\ &\leq \sum_{\ell=0}^{\ell_0} \sum_{u < v} \Pr[uv \in \mathcal{E}_{\ell+1}] \log \left(\underbrace{\mathbb{E}[\#\{\text{vertices in the cell of } u\} \mid uv \in \mathcal{E}_{\ell+1} \setminus \mathcal{E}_\ell]}_{=: T_\ell} \right) \end{aligned}$$

The term T_ℓ is at most $T_\ell \leq 2 + (n - 2)2^{-\ell d} \leq 3n2^{-\ell d}$ for $\ell \leq \ell_0$ (where we count 2 for u and v and use independence of the other vertex positions). Thus it remains to show that $\mathbb{E}[\sum_{\ell=0}^{\ell_0} |\mathcal{E}_{\ell+1}| \log(3n2^{-\ell d})] = O(n)$. From Lemma 8.1 we know that $\mathbb{E}[|\mathcal{E}_\ell|] \leq E_\ell$, where $E_\ell = n \cdot (2^{d\ell}/n)^{\beta-2-o(1)} + (n^{2-\alpha}2^{d\ell(\alpha-1)} + n^{1-1/d}2^\ell)(1 + \log(n2^{-d\ell}))$. Since E_ℓ increases exponentially in ℓ , we obtain

$$\mathbb{E} \left[\sum_{\ell=0}^{\ell_0} |\mathcal{E}_{\ell+1}| \log(3n2^{-d\ell}) \right] \leq O \left(\sum_{\ell=1}^{\ell_0+1} E_\ell \log(3n2^{-d\ell+d}) \right) = O(E_{\ell_0+1} \log(3n2^{-d\ell_0})) = O(n),$$

where the last equality follows since $1/n \leq 2^{-d\ell_0} \leq O(1/n)$ by our choice of ℓ_0 . This proves the lemma, and hence shows that we need $O(n)$ bits in expectation to store the graph. \square

This concludes the proof of Theorem 2.6.

9 Conclusion

To cope with the technical shortcomings of hyperbolic random graphs, we introduced a new model of scale-free random graphs with underlying geometry – geometric inhomogeneous random graphs – and theoretically analyzed their fundamental structural and algorithmic properties. Scale-freeness and basic connectivity properties of our model follow from more general considerations [14]. We established that (1) hyperbolic random graphs are a special case of GIRGs, (2) GIRGs have a constant clustering coefficient, and (3) GIRGs have small separators and are very well compressible. As our main result, (4) we presented an expected-linear-time sampling algorithm. This improves the best-known sampling algorithm for hyperbolic random graphs by a factor $O(\sqrt{n})$.

The most important experimental finding for hyperbolic random graphs is that greedily constructed paths are very close to shortest paths [8]. Hence, we will study greedy routing on GIRGs in future work, for which we laid the foundations in the present paper.

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

References

- [1] M. A. Abdullah, M. Bode, and N. Fountoulakis. Typical distances in a geometric model for complex networks. *arXiv preprint arXiv:1506.07811*, 2015.
- [2] W. Aiello, A. Bonato, C. Cooper, J. Janssen, and P. Prałat. A spatial web graph model with local influence regions. *Internet Mathematics*, 5(1-2):175–196, 2008.
- [3] A. Barabási and R. Albert. Emergence of scaling in random networks. *Science*, 286(5439):509–512, 1999.
- [4] V. Batagelj and U. Brandes. Efficient generation of large random networks. *Physical Review E*, 71(3):036113, 2005.
- [5] D. K. Blandford, G. E. Blelloch, and I. A. Kash. Compact representations of separable graphs. In *Proceedings of the fourteenth annual ACM-SIAM Symposium on Discrete algorithms*, pages 679–688. Society for Industrial and Applied Mathematics, 2003.
- [6] M. Bode, N. Fountoulakis, and T. Müller. On the giant component of random hyperbolic graphs. In *7th European Conference on Combinatorics, Graph Theory and Applications (EUROCOMB)*, pages 425–429. Springer, 2013.
- [7] M. Bode, N. Fountoulakis, and T. Müller. On the geometrisation of the Chung-Lu model and its component structure. *Preprint*, 2014.
- [8] M. Boguñá, F. Papadopoulos, and D. Krioukov. Sustaining the Internet with hyperbolic mapping. *Nature Communications*, 1(6), September 2010.
- [9] P. Boldi and S. Vigna. The WebGraph framework I : compression techniques. In *13th International Conference on World Wide Web (WWW)*, pages 595–602, 2004.
- [10] B. Bollobás, S. Janson, and O. Riordan. The phase transition in inhomogeneous random graphs. *Random Structures & Algorithms*, 31(1):3–122, 2007.
- [11] A. Bonato, J. Janssen, and P. Prałat. A geometric model for on-line social networks. In *1st International Workshop on Modeling Social Media (WOSM)*, 2010.
- [12] M. Bradonjić, A. Hagberg, and A. G. Percus. The structure of geographical threshold graphs. *Internet Mathematics*, 5(1-2):113–139, 2008.
- [13] K. Bringmann and T. Friedrich. Exact and efficient generation of geometric random variates and random graphs. In *40th International Colloquium on Automata, Languages, and Programming (ICALP)*, pages 267–278, 2013.
- [14] K. Bringmann, R. Keusch, and J. Lengler. Average distance in a general class of scale-free networks with underlying geometry. *arxiv*, 2016.
- [15] K. Bringmann, R. Keusch, and J. Lengler. An Azuma-Hoeffding inequality with error events and two-sided Lipschitz condition. *manuscript*, 2016.
- [16] E. Candellero and N. Fountoulakis. Clustering and the hyperbolic geometry of complex networks. In *11th International Workshop on Algorithms and Models for the Web Graph (WAW)*, pages 1–12, 2014.
- [17] F. Chierichetti, R. Kumar, S. Lattanzi, M. Mitzenmacher, A. Panconesi, and P. Raghavan. On compressing social networks. In *15th International Conference on Knowledge Discovery and Data Mining (KDD)*, pages 219–228, 2009.
- [18] F. Chierichetti, R. Kumar, S. Lattanzi, A. Panconesi, and P. Raghavan. Models for the Compressible Web. In *50th Annual IEEE Symposium on Foundations of Computer Science (FOCS)*, pages 331–340, 2009.
- [19] F. Chung and L. Lu. The average distances in random graphs with given expected degrees. *Proceedings of the National Academy of Sciences (PNAS)*, 99(25):15879–15882, 2002.
- [20] F. Chung and L. Lu. Connected components in random graphs with given expected degree sequences. *Annals of Combinatorics*, 6(2):125–145, 2002.
- [21] F. Chung and L. Lu. The average distance in a random graph with given expected degrees. *Internet Mathematics*, 1(1):91–113, 2004.
- [22] D. R. Clark and I. Munro. Efficient suffix trees on secondary storage. In *7th Symposium on Discrete Algorithms (SODA)*, pages 383–391, 1996.
- [23] M. Deijfen, R. van der Hofstad, and G. Hooghiemstra. Scale-free percolation. *Annales de l’Institut Henri Poincaré, Probabilités et Statistiques*, 49(3):817–838, 2013.
- [24] L. Devroye. *Nonuniform random variate generation*. Springer, New York, 1986.
- [25] S. N. Dorogovtsev and J. F. F. Mendes. Evolution of networks. *Advances in Physics*, 51(4):1079–1187, 2002.
- [26] T. Friedrich and A. Krohmer. Cliques in hyperbolic random graphs. In *2015 IEEE Conference on Computer Communication (INFOCOM)*, pages 1544–1552, 2015.

- [27] T. Friedrich and A. Krohmer. On the diameter of hyperbolic random graphs. In *42nd International Colloquium on Automata, Languages, and Programming (ICALP)*, Lecture Notes in Computer Science, 2015. To appear.
- [28] L. Gugelmann, K. Panagiotou, and U. Peter. Random hyperbolic graphs: degree sequence and clustering. In *9th International Colloquium on Automata, Languages, and Programming (ICALP)*, pages 573–585, 2012.
- [29] E. Jacob and P. Mörters. A spatial preferential attachment model with local clustering. In *Algorithms and Models for the Web Graph*, pages 14–25. Springer, 2013.
- [30] G. Jacobson. Space-efficient static trees and graphs. In *30th Annual Symposium on Foundations of Computer Science (FOCS)*, pages 549–554, 1989.
- [31] Marcos Kiwi and Dieter Mitsche. A bound for the diameter of random hyperbolic graphs. In *2015 Proceedings of the Twelfth Workshop on Analytic Algorithmics and Combinatorics (ANALCO)*, pages 26–39. SIAM, 2015.
- [32] D. Krioukov, F. Papadopoulos, M. Kitsak, A. Vahdat, and M. Boguñá. Hyperbolic geometry of complex networks. *Phys. Rev. E*, 82:036106, 2010.
- [33] S. Kutin. Extensions to McDiarmid’s inequality when differences are bounded with high probability. Technical report, Department of Computer Science, University of Chicago, 2002.
- [34] L. Le Cam. An approximation theorem for the poisson binomial distribution. *Pacific J. Math*, 10(4):1181–1197, 1960.
- [35] J. C. Miller and A. Hagberg. Efficient generation of networks with given expected degrees. In *8th International Conference on Algorithms and Models for the Web Graph (WAW)*, pages 115–126, 2011.
- [36] I. Norros and H. Reittu. On a conditionally Poissonian graph process. *Advances in Applied Probability*, 38(1):59–75, 2006.
- [37] F. Papadopoulos, D. Krioukov, M. Boguñá, and A. Vahdat. Greedy forwarding in dynamic scale-free networks embedded in hyperbolic metric spaces. In *INFOCOM 2010. 29th IEEE International Conference on Computer Communications*, pages 1–9, March 2010.
- [38] M. Pătraşcu. Succincter. In *49th Annual Symposium on Foundations of Computer Science (FOCS)*, pages 305–313, 2008.
- [39] M. Penrose. *Random geometric graphs*, volume 5. Oxford University Press Oxford, 2003.
- [40] U. Peter. *Random Graph Models for Complex Systems*. PhD thesis, ETH Zurich, 2014.
- [41] M. Von Looz, C. L. Staudt, H. Meyerhenke, and R. Prutkin. Fast generation of dynamic complex networks with underlying hyperbolic geometry. *CoRR*, abs/1501.03545, 2015.