

# Maximal Cliques in Scale-Free Random Graphs

Thomas Bläsius and Maximillian Katzmann

Karlsruhe Institute of Technology, Karlsruhe, Germany

and

Clara Stegehuis

Faculty of Electrical Engineering, Mathematics and Computer Science, University of Twente

## Abstract

We investigate the number of maximal cliques, i.e., cliques that are not contained in any larger clique, in three network models: Erdős–Rényi random graphs, inhomogeneous random graphs (also called Chung–Lu graphs), and geometric inhomogeneous random graphs. For sparse and not-too-dense Erdős–Rényi graphs, we give linear and polynomial upper bounds on the number of maximal cliques. For the dense regime, we give super-polynomial and even exponential lower bounds. Although (geometric) inhomogeneous random graphs are sparse, we give super-polynomial lower bounds for these models. This comes from the fact that these graphs have a power-law degree distribution, which leads to a dense subgraph in which we find many maximal cliques. These lower bounds seem to contradict previous empirical evidence that (geometric) inhomogeneous random graphs have only few maximal cliques. We resolve this contradiction by providing experiments indicating that, even for large networks, the linear lower-order terms dominate, before the super-polynomial asymptotic behavior kicks in only for networks of extreme size.

## 1 Introduction

While networks appear in many different applications, many real-world networks were found to share some important characteristics. First of all, often their degree distribution is heavy-tailed, which is sometimes denoted as the network being scale-free. Secondly, they often have a high clustering coefficient, implying that it is likely that two neighbors of a vertex are connected themselves as well. For this reason, random graph models that can achieve both scale-freeness and a high clustering coefficient have been at the center of attention over the last years.

One example of such a model is the popular hyperbolic random graph (HRG) [19], which has for example been used to model the network of world wide trade [13] or the Internet on the Autonomous Systems level [6, 18]. This random graph model embeds the vertices in an underlying hyperbolic space and connects them with probabilities depending on their distances, where nearby vertices are more likely to connect. The triangle inequality then ensures the presence of many triangles, while the hyperbolic space ensures the presence of a scale-free degree distribution. Recently, the geometric inhomogeneous random graph (GIRG) was proposed as a generalization of HRG. It combines power-law distributed weights with Euclidean space, making the model simpler to analyze [7].

While the hyperbolic random graph and the GIRG have been designed to exhibit high clustering and a scale-free degree distribution, the question remains whether other properties of this model match real-world data. For this reason, many properties of the GIRG or hyperbolic random graph have been analyzed mathematically, such as the maximum clique size [2], number of  $k$ -cliques [21], spectral gap [17] and separator size [5, 20].

In this paper, we focus on another network property: the number of maximal cliques, i.e., cliques that are not part of any larger clique. Cliques in general are an important indicator for structural properties of a network. Indeed, the number of large cliques is a measure of the tendency of a network to cluster into groups. Small cliques of size 3 (triangles) on the other hand, can form an indication of the transitivity of a network or its clustering coefficient.

To study these structural clique-based properties, however, all cliques of a given size need to be listed, which can be a computationally expensive process. To list all network cliques, it suffices to list only all maximal cliques, as all smaller cliques can be generated from at least one maximal clique. For this reason, enumerating all maximal cliques of a graph is at the heart of our understanding of cliques in general.

For enumerating all maximal cliques, an output-polynomial algorithm [22] exists, which can enumerate all maximal cliques efficiently if the graph contains only few of them. This creates a link between enumeration and counting: if the maximal clique count is low, then it is possible to efficiently enumerate them. There also exist highly efficient implementations to enumerate all maximal cliques [9, 10, 11]. However, for a given graph, it is usually not known a priori how many maximal cliques it has. If this number is large, enumerating all maximal cliques can still take exponential time. However, in practice, enumerating the number of maximal cliques often takes a short amount of time for many real-world instances as well as in realistic network models [3]. In this paper, we therefore focus on the number of maximal cliques in the GIRG random graph, that is, the maximal clique count. As the GIRG possesses the two main characteristics that are essential to many real-world networks, scale-freeness and an underlying geometry, we believe that investigating the number of maximal cliques in the GIRG can provide insights into why enumerating the number of maximal cliques can often be done efficiently for many real-world networks.

To investigate the influence of the different properties of scale-freeness and clustering, we investigate the number of maximal cliques in three steps. First, we investigate a model without heavy-tailed degrees and with a small clustering coefficient, the Erdős–Rényi model  $G(n, p)$ ; see Section 2. We then investigate the GIRG model (Section 3), which has both clustering and scale-free degrees. Finally, in Section 4, we investigate the Inhomogeneous Random Graph (IRG), a model that is scale-free but has a small clustering coefficient. We complement our theoretical bounds with experiments in Section 5. In all models, we will be interested in the large  $n$  limit. That is, we investigate how the number of maximal cliques scales in the number of nodes  $n$  when  $n$  grows large. Our main findings can be summarized as follows; also see Table 1 for an overview of our results.

- There is a strong dependence on the density of the network. For the Erdős–Rényi model ( $G(n, p)$ ) we obtain a linear upper bound for sparse graphs ( $O(n)$  edges) and a polynomial upper bound for non-dense graphs ( $O(n^{2-\varepsilon})$  edges for any  $\varepsilon > 0$ ). For dense graphs on the other hand ( $\Omega(n^2)$  edges), we obtain a super-polynomial lower bound. If the density is high enough, our lower bound is even exponential.
- This insight carries over to the IRG and GIRG models. Though they are overall sparse, they contain sufficiently large dense subgraphs that allow us to obtain super-polynomial lower bounds.
- In the IRG model with power-law exponent  $\tau \in (2, 3)$  the small maximal cliques localize: asymptotically maximal cliques of constant size  $k > 2$  are formed by  $k - 2$  hubs of high degree proportional to  $n^{1/(\tau-1)}$  and two vertices of lower degree proportional to  $n^{(\tau-2)/(\tau-1)}$ .
- We complement our theoretical lower bounds with experiments showing that the super-polynomial growth becomes only relevant for very large networks.

**Discussion and Related Work.** Although cliques themselves have been studied extensively in the literature, there is, to the best of our knowledge, only little previous work on the number of *maximal* cliques in network models. In fact, the only theoretical analysis we are aware of is the recent preprint by Yamaji [23], giving bounds for hyperbolic random graphs (HRG) and random geometric graph (RGG), which are also shown in Table 1. Interestingly, this includes the upper bound of  $\exp(O(n^{\frac{3-\tau}{6}+\varepsilon}))$  for the HRG model. In contrast to that, we give the asymptotically larger lower bound  $\exp(\Omega(n^{\frac{3-\tau}{4}-\varepsilon}))$  for the corresponding GIRG variant. Thus, there is an asymptotic difference between the HRG and the GIRG model.

This is surprising as the GIRG model is typically perceived as a generalization of the HRG model. More precisely, there is a mapping between the two models such that for every HRG with average degree  $d_{\text{HRG}}$  there exist GIRGs with average degree  $d_{\text{GIRG}}$  and  $D_{\text{GIRG}}$  with  $d_{\text{GIRG}} \leq d_{\text{HRG}} \leq D_{\text{GIRG}}$  that are sub- and supergraphs of the HRG, respectively. Moreover,  $d_{\text{GIRG}}$  and  $D_{\text{GIRG}}$  are only a constant factor apart and experiments indicate that  $d_{\text{HRG}} = d_{\text{GIRG}} \cdot (1 + o(1))$ , i.e., every HRG has a corresponding GIRG that is missing only a sublinear number of edges [4]. In the case of maximal cliques, however, this minor difference between the models leads to an asymptotic difference.

Besides this theoretical analysis, it has been observed empirically that the number of maximal cliques in most real-world networks as well as in the GIRG and the IRG model is smaller than the number of edges of the graph [3]. This indicates linear scaling in the graph size with low constant factors and small lower-order terms, which seems to be a stark contradiction to the super-polynomial lower bounds we prove here. We resolve this contradiction with our experiments in Section 5, where we observe that the graph size has to be quite large before the asymptotic behavior kicks in, i.e., we observe the super-polynomial scaling as predicted by our theorems but on such a low level that it is overshadowed by the linear lower-order terms.

**Notation and setting.** In the rest of this paper, we will be interested in results in the large  $n$  limit, where  $n$  denotes the number of nodes in the random graph. We therefore use classical asymptotic notation, in terms of the graph size  $n$ . For any two non-negative functions  $f(n), g(n)$  we will write  $f(n) \in o(g(n))$  if  $\lim_{n \rightarrow \infty} f(n)/g(n) = 0$ ;  $f(n) \in O(g(n))$  if  $\limsup_{n \rightarrow \infty} f(n)/g(n) < \infty$ ;  $f(n) \in \Omega(g(n))$  if  $\liminf_{n \rightarrow \infty} f(n)/g(n) > 0$ ;  $f(n) \in \Theta(g(n))$  if  $f(n) \in O(g(n))$  and  $f(n) \in \Omega(g(n))$ . Moreover, we will say that a sequence of events  $\{\mathcal{E}_n\}_{n \geq 1}$  happens with high probability (w.h.p.) if  $\lim_{n \rightarrow \infty} \mathbb{P}(\mathcal{E}_n) = 1$ .

## 2 Erdős–Rényi Random Graph

An Erdős–Rényi random graph [14, 12]  $G(n, p)$  has  $n$  vertices and each pair of vertices is connected independently with probability  $p$ . We give bounds on the number of maximal cliques in a  $G(n, p)$  depending on  $p$ . Roughly speaking, we give super-polynomial lower bounds for the dense regime and polynomial upper bounds for a sparser regime. Specifically, we first give a general lower bound that is super-polynomial if  $p$  is non-vanishing for growing  $n$ , i.e., if  $p \in \Omega(1)$ . Note that  $p \in \Omega(1)$  yields a dense graph with a quadratic number of edges in expectation. For super-dense graph with  $p = 1 - c/n$  for a constant  $c$ , we strengthen this lower bound to exponential. In contrast to this, we give a polynomial upper if  $p \in O(n^{-a})$  for any constant  $a > 0$ . For sparse graphs with  $p \in O(n^{-1})$ , yielding graphs with  $\Theta(n)$  edges in expectation, our upper bound on the number of maximal cliques is linear. We start with the general lower bound.

**Theorem 2.1.** *Let  $N$  be the number of maximal cliques in a  $G(n, p)$ . Then, for  $n$  sufficiently large,*

$$\mathbb{E}[N] \geq n^{\frac{\log(n)/2 - \log \log n + \log \log(1/p)}{\log(1/p)}} \cdot \frac{1 - o(1)}{e}. \quad (1)$$

	Model	Maximal cliques	Reference
$G(n, p)$	$p = 1 - \Theta(\frac{1}{n})$	$2^{\Omega(n)}$	Theorem 2.3
	$p \in \Theta(1)$	$n^{\Omega(\log n)}$	Theorem 2.1
	$p \in O(\frac{1}{n^a})$	$n^{O(1)}$	Theorem 2.4
	$p \in O(\frac{1}{n})$	$O(n)$	Theorem 2.4
IRG		$\exp(\Omega(n^{\frac{3-\tau}{4}-\varepsilon} \log n))$	Theorem 4.1
GIRG	$d$ -dim torus, $T = 0$	$\exp(\Omega(n^{\frac{3-\tau}{4}-\varepsilon}))$	Corollary 3.6
	$d$ -dim torus, $T > 0$	$\exp(\Omega(n^{\frac{3-\tau}{5}} (\varepsilon \log n)^{-(1/2)}))$	Corollary 3.9
	2-dim square, $T = 0$	$\exp(\Omega(n^{\frac{3-\tau}{10}-\varepsilon}))$	Theorem 3.7
	2-dim square, $T > 0$	$\exp(\Omega(n^{\frac{3-\tau}{10}-\varepsilon}))$	Theorem 3.10
RGG	2-dim, dense	$\exp(\Omega(n^{\frac{1}{3}}))$	[23]
	2-dim, dense	$\exp(O(n^{\frac{1}{3}+\varepsilon}))$	[23]
HRG		$\exp(\Omega(n^{\frac{3-\tau}{6}}))$	[23]
		$\exp(O(n^{\frac{3-\tau}{6}+\varepsilon}))$	[23]

Table 1: Summary of our and other results on the number of maximal cliques in different random graph models and their scaling in the number of vertices.

*Proof.* Let  $N_k$  be the number of maximal cliques of size  $k$ . To estimate  $\mathbb{E}[N_k]$ , note that the probability that a fixed subset  $C \subseteq V$  of  $|C| = k$  vertices forms a clique is  $p^{k(k-1)/2}$ . Moreover, it is maximal if none of the other  $n - k$  vertices is connected to all  $k$  vertices of  $C$ , which happens with probability  $(1 - p^k)^{n-k}$ . As the two events are independent and there are  $\binom{n}{k}$  vertex sets of size  $k$ , we obtain

$$\mathbb{E}[N_k] = \binom{n}{k} p^{k(k-1)/2} (1 - p^k)^{n-k}. \quad (2)$$

Using that  $\binom{n}{k} \geq (n/k)^k$  and increasing the exponents of the probabilities, we obtain

$$\mathbb{E}[N_k] \geq \left(\frac{n}{k}\right)^k p^{k^2/2} (1 - p^k)^n.$$

We now set  $k = \log(n)/\log(1/p) = -\log(n)/\log(p)$ , which yields  $p^k = n^{-1}$ . Thus, in the above bound, the term  $n^k p^{k^2/2}$  simplifies to  $n^k n^{-k/2} = n^{k/2}$ . Moreover, the term  $(1 - p^k)^n$  simplifies to  $(1 - 1/n)^n$ , which converges to  $1/e$  for  $n \rightarrow \infty$ . Thus, we obtain

$$\begin{aligned} \mathbb{E}[N_k] &\geq n^{k/2} \frac{1}{e k^k} (1 - o(1)) \\ &= n^{\frac{\log(n)/2}{\log(1/p)}} \cdot \left(\frac{\log(n)}{\log(1/p)}\right)^{-\frac{\log(n)}{\log(1/p)}} \cdot \frac{1 - o(1)}{e}. \end{aligned}$$

Changing the base of the second factor yields

$$= n^{\frac{\log(n)/2}{\log(1/p)}} \cdot e^{-\frac{\log(n)}{\log(1/p)} \cdot \log\left(\frac{\log(n)}{\log(1/p)}\right)} \cdot \frac{1 - o(1)}{e}$$

$$\begin{aligned}
&= n^{\frac{\log(n)/2}{\log(1/p)}} \cdot n^{-\frac{\log \log n - \log \log(1/p)}{\log(1/p)}} \cdot \frac{1 - o(1)}{e} \\
&= n^{\frac{\log(n)/2 - \log \log n + \log \log(1/p)}{\log(1/p)}} \cdot \frac{1 - o(1)}{e}.
\end{aligned}$$

As there are clearly at least as many maximal cliques as maximal cliques of size  $k$ , claimed bound for  $\mathbb{E}[N]$  follows.  $\square$

This means that in a dense Erdős–Rényi random graph (constant  $p$ ), the expected number of maximal cliques is super-polynomial in  $n$ . In the following, we show that, when the graph gets even denser, the number of maximal cliques even grows exponentially. For this, we prove the existence of an induced subgraph that has many maximal cliques. Specifically, we aim to find a large *co-matching*, i.e., the complement graph of a matching (or equivalently, a co-matching).

**Lemma 2.2.** *Let  $G$  be a co-matching on  $2k$  vertices. Then  $G$  has  $2^k$  maximal cliques.*

*Proof.* The complement  $\overline{G}$  of  $G$  is a matching with  $k$  edges. The maximal independent sets of  $\overline{G}$  are the vertex sets that contain for each edge exactly one of its vertices. Thus,  $\overline{G}$  has  $2^k$  maximal independent sets, which implies that  $G$  has  $2^k$  maximal cliques.  $\square$

With this, we can show an exponential lower bound for super-dense Erdős–Rényi graphs.

**Theorem 2.3.** *For every  $c > 0$ , there exists a  $\zeta > 0$  and  $n' > 0$  such that  $G(n, 1 - c/n)$  contains at least  $2^{\zeta n}$  cliques with high probability for all  $n \geq n'$ .*

*Proof.* A co-matching in  $G(n, 1 - c/n)$  corresponds to an induced matching in  $G(n, c/n)$ . Now fix  $M > 1$ . Then, by [15, Theorem 5.12], with high probability the Erdős–Rényi random graph contains a linear number of vertices of degree at most  $M$  and at least 1. Denote the reduced graph with only vertices of degree at most  $M$  by  $G_{\leq M}$ , which has a linear number of edges. Now we construct an induced matching of linear size in  $G_{\leq M}$  as follows. Start with any edge  $\{u, v\}$  in  $G_{\leq M}$ , and add it to the matching. Then, remove  $u, v$  and all neighbors of  $u$  and  $v$  from  $G_{\leq M}$ . This removes at most  $2M^2$  edges from  $G_{\leq M}$ , as all degrees are bounded by  $M$ . Then, pick another edge and continue this process until  $G_{\leq M}$  contains no more edges. As this process removes only a constant number of edges after picking a new edge, at least a linear number of edges will be added before the process finishes. Thus, there is an induced matching of at least  $\zeta n$  with high probability, which yields the claim due to Lemma 2.2.  $\square$

Next we consider less dense Erdős–Rényi graphs with  $p \in O(n^{-a})$  for a constant  $a \in (0, 1]$  and prove a polynomial upper bound on the number of maximal cliques. The degree of the polynomial depends on  $a$ . For sparse graphs with  $p \in O(n^{-1})$ , our bound is linear.

**Theorem 2.4.** *Let  $p = (c/n)^a$  for constants  $c > 0$  and  $a \in (0, 1]$  and let  $N$  be the number of maximal cliques in a  $G(n, p)$ . Then  $\mathbb{E}[N] \in O(n^x)$  with*

$$x = \left\lceil \frac{1}{a} \right\rceil - a \cdot \left( \left\lceil \frac{1}{a} \right\rceil \right).$$

*Proof.* As in Theorem 2.1, let  $N_k$  be the number of maximal cliques of size  $k$ . Note that the number of maximal cliques is upper bounded by the number of (potentially non-maximal) cliques. Thus, we obtain

$$\mathbb{E}[N_k] \leq \binom{n}{k} p^{\frac{k(k-1)}{2}}.$$

Using that  $\binom{n}{k} \leq (en/k)^k$ , inserting  $p = (c/n)^a$ , and rearranging yields

$$\begin{aligned} \mathbb{E}[N_k] &\leq \left(\frac{en}{k}\right)^k \left(\frac{c}{n}\right)^{a\frac{k(k-1)}{2}} \\ &= \left(\frac{ce}{k}\right)^k \left(\frac{c}{n}\right)^{a\frac{k(k-1)}{2} - k}. \end{aligned} \quad (3)$$

We first argue that we can focus on the case where  $k$  is constant as the above term vanishes sufficiently quickly for growing  $k$ . For this, note that  $ak(k-1)/2 - k \geq k$  if  $k \geq 4/a + 1$ . Thus, as  $c/n < 1$  for sufficiently large  $n$ , the second factor of Equation (3) is upper bounded by  $(c/n)^k$ . For  $k \geq 4/a + 1$ , it then follows that  $\mathbb{E}[N_k] \leq (c^2 e/(kn))^k$ . For sufficiently large  $n$ , the fraction is smaller than 1 and thus the sum over all  $N_k$  for larger values of  $k$  is upper bounded by a constant due to the convergence of the geometric series.

Focusing on  $k \in \Theta(1)$  and ignoring constant factors, we obtain

$$\mathbb{E}[N] \in O\left(\max_{k \in \mathbb{N}^+} \left\{n^{x(k)}\right\}\right) \text{ with } x(k) = k - a\frac{k(k-1)}{2}.$$

To evaluate the maximum, note that  $x(k)$  describes a parabola with its maximum at  $k_0 = 1/a + 1/2$ . However,  $k_0$  may not be integral. To determine the integer  $k$  that maximizes  $x(k)$ , note that for  $a \in [\frac{1}{i}, \frac{1}{i-1}]$  with  $i \in \mathbb{N}^+$ , we get  $k_0 \in [i - \frac{1}{2}, i + \frac{1}{2}]$ . Thus,  $i$  is the closest integer to  $k_0$ . As the parabola is symmetric at its maximum  $k_0$ , the exponent  $x(k)$  is maximized for the integer  $k = i = \lceil \frac{1}{a} \rceil$ . Substituting  $k(k-1)/2 = \binom{k}{2}$  yields the claim.  $\square$

### 3 Geometric Inhomogeneous Random Graphs (GIRG)

While the Erdős–Rényi random graph is homogeneous, and does not contain geometry, we now investigate the number of maximal cliques in a model that contains both these properties, the Geometric Inhomogeneous Random Graph (GIRG) [7]. We will use similar notation as in [7], except for the parameters  $\alpha$  and  $\beta$ , which we will replace by  $1/T$  and  $\tau$  respectively, to be more consistent with the literature on other similar models [19]. In this model, each vertex  $v$  has a weight,  $w_v$  and a position  $x_v$ . The weights are independent copies of a power-law random variable  $W$  with exponent  $\tau$ , i.e.,

$$1 - F(w) := \mathbb{P}(W > w) = w^{1-\tau}, \quad (4)$$

for all  $w \geq 1$ . We impose the condition  $\tau \in (2, 3)$ , to ensure that the weights have finite mean but unbounded variance. The parameter  $\mu$  denotes the mean of this distribution, and can be computed as  $\mu = (\tau - 2)^{-1}$ . The vertex positions  $x_1, \dots, x_n$  are independent copies of a uniform random variable on the  $d$ -dimensional torus  $\mathbb{T}^d = \mathbb{R}^d/\mathbb{Z}^d$ .

An edge between any two vertices  $u, v \in V$  of the GIRG appears independently with a probability  $p_{uv}$  determined by the weights and the positions of the vertices

$$p_{uv} = \min\left\{\left(\frac{w_u w_v}{n\mu \|x_u - x_v\|^d}\right)^{1/T}, 1\right\}, \quad (5)$$

where  $\|\cdot\|$  denotes the maximum norm on the torus,  $\mu$  is a parameter controlling the average degree, and  $0 < T < 1$  is the *temperature* and controls the influence of the geometry. We say that  $T = 0$  is the *threshold* case of the GIRG. That is, when  $T = 0$ ,

$$p_{uv} = \begin{cases} 1 & \frac{w_u w_v}{n\mu \|x_u - x_v\|^d} \geq 1 \\ 0 & \text{else.} \end{cases} \quad (6)$$

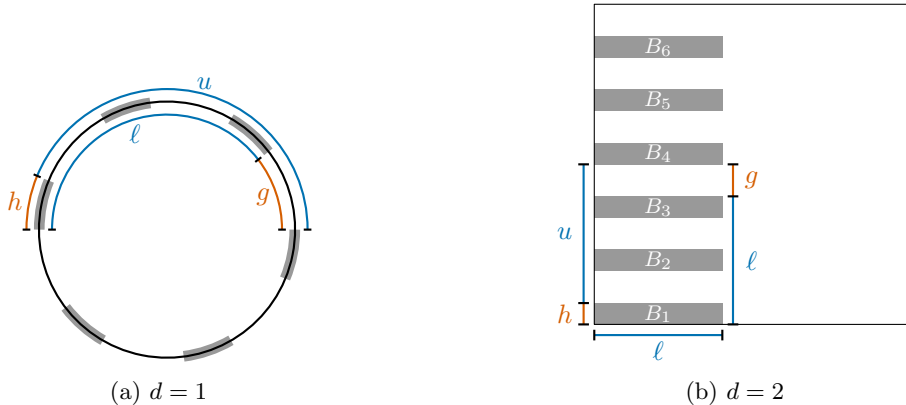


Figure 1: Illustration of the gray shaded boxes  $B_i$  on the 1 and 2-dimensional torus.

In general, we will be interested in results for the GIRG model when the number of nodes,  $n$ , tends to infinity. We will then often refer to the family of GIRGs generated for varying  $n$  by  $G^{(n)}$ , where we assume that all other parameters ( $\mu, \tau, \gamma, d$ ) remain fixed.

In the following, we first give a lower bound for the threshold case (Section 3.1). The proof makes use of the toroidal structure of the ground space. To prove that this is not essential to obtain a super-polynomial number of maximal cliques, we additionally give a lower bound for a variant of the model where the ground space is a 2-dimensional unit square with Euclidean norm (Section 3.2). Finally, in Section 3.3, we show how to extend these results to the general case with non-zero temperatures.

### 3.1 Threshold Case

Here we show that a  $d$ -dimensional threshold GIRG  $G = (V, E)$  has, with high probability, a super-polynomial number of maximal cliques. To achieve this, we proceed as follows to show that  $G$  has a large co-matching as induced subgraph (also see Lemma 2.2). We consider the vertex set  $S \subseteq V$  containing all vertices whose weight lies between a lower bound  $w_\ell$  and an upper bound  $w_u$ . As a co-matching is quite dense, it makes sense to think of these as rather large weights. We then define disjoint regions  $B_1, \dots, B_{2k}$ . For  $i \in [k]$ , we call  $B_i$  and  $B_{i+k}$  a pair of *opposite* regions. These regions will satisfy the following three properties. First, every  $B_i$  contains a vertex from  $S$  with high probability. Secondly, pairs of vertices from  $S$  in opposite regions are not connected. And thirdly, vertices from  $S$  that do not lie in opposite regions are connected. Note that these properties imply the existence of a co-matching on  $2k$  vertices, as choosing an arbitrary vertex of  $S$  for each region  $B_i$  makes it so that each chosen vertex has exactly one partner from the opposite region to which it is not connected, while it is connected to the vertices from all other regions.

In the following we first give a parameterized definition of the regions  $B_i$  and then show how to choose the parameters for the above strategy to work; also see Figure 1. Each  $B_i$  is an axis-aligned box, i.e., the cross product of intervals. Let  $g(n), h(n) > 0$  such that  $1/(g(n) + h(n))$  is an even number. Think of  $h(n)$  of as the *height* of each box and of  $g(n)$  as the *gap* between the boxes, yielding  $2k = 1/(g(n) + h(n))$  boxes. Now we define  $B_i = [(i-1) \cdot (g(n) + h(n)), (i-1) \cdot (g(n) + h(n)) + h(n)] \times [0, \frac{1}{2} - g(n)]^{d-1}$  for  $i \in [2k]$ . We call the resulting regions  $B_1, \dots, B_{2k}$  the *evenly spaced boxes of height  $h(n)$  and gap  $g(n)$* , see Figure 1 for an illustration for  $d = 1$  and  $d = 2$ . As before,  $B_i$  and  $B_{i+k}$  for  $i \in [k]$  are *opposite* boxes.

With this, note that the distance between any pair of points in opposite boxes is at least  $u = \frac{1}{2} - h(n)$  (recall that we assume the infinity norm). Moreover, the distance between any pair of points in non-opposite regions is at most  $\ell = \frac{1}{2} - g(n)$ . This yields the following lemma.

**Lemma 3.1.** *Let  $B_1, \dots, B_{2k}$  be evenly spaced boxes of height  $h(n)$  and gap  $g(n)$  in  $\mathbb{T}^d$ . Let  $w_\ell = (\frac{1}{2} - g(n))^{d/2} \sqrt{\mu n}$  and  $w_u = (\frac{1}{2} - h(n))^{d/2} \sqrt{\mu n}$ . If a GIRG on  $n$  vertices with  $T = 0$ ,  $\tau \in (2, 3)$  and  $\mu$  places one vertex of weight in  $[w_\ell, w_u]$  in each box  $B_i$ , then these vertices form a co-matching.*

*Proof.* As observed above, the vertices in opposite boxes have distance at least  $u = \frac{1}{2} - h(n)$ . Moreover, the vertices considered here have weight less than  $w_u = u^{d/2} \sqrt{\mu n}$ . As  $w_u^2 / (\mu n u^d) = 1$ , these vertices are not connected because the weight interval  $[w_\ell, w_u]$  is open at  $w_u$  (see Equation (6)). Similarly, vertices in non-opposite boxes have distance at most  $\ell = \frac{1}{2} - g(n)$  and weight at least  $w_\ell = \ell^{d/2} \sqrt{\mu n}$ . As  $w_\ell^2 / (\mu n \ell^d) = 1$ , such vertices are connected. Hence, we get a co-matching.  $\square$

It now remains to choose  $g(n)$  and  $h(n)$  appropriately. First observe that, for the weight range in Lemma 3.1 to be non-empty, we need  $w_\ell < w_u$  and thus  $g > h$ . Beyond that, we want to achieve the following three goals. First, the weight range needs to be sufficiently large such that we actually have a sufficient number of vertices in this range. For this, we want to choose  $g(n)$  substantially larger than  $h(n)$ . Secondly, we want to make each box  $B_i$  sufficiently large for it to contain a vertex with high probability. For this, we mainly want  $h(n)$  to be large. Thirdly, we want the number of boxes  $2k = 1/(g(n) + h(n))$  to be large to obtain a large co-matching. For this, we want  $g(n)$  and  $h(n)$  to be small.

Note that the restrictions of choosing  $h(n)$  large,  $g(n)$  larger than  $h(n)$ , and  $g(n) + h(n)$  small are obviously conflicting. In the following, we show how to balance these goals out to obtain a co-matching of polynomial size. We start by estimating the number of vertices in the given weight range in the following lemma, which is slightly more general than we need.

**Lemma 3.2.** *Let the vertex weights independently be sampled as in (4), with  $\tau \in (2, 3)$ . Let  $a, b > 0$  be constants and let  $g(n), h(n)$  be functions of  $n$  such that  $g(n), h(n) \in o(1)$ . Let  $S$  be the set of vertices with weight in  $[(a - g(n))^b \sqrt{\mu n}, (a - h(n))^b \sqrt{\mu n}]$ . Then*

$$\mathbb{E}[|S|] = n^{\frac{3-\tau}{2}} \cdot \mu^{\frac{1-\tau}{2}} b a^{b(1-\tau)-1} \cdot (g(n) - h(n) \pm O(g(n)^2 + h(n)^2)). \quad (7)$$

*Proof.* Recall from (4) that the cumulative distribution function for the weights is  $F(x) = 1 - x^{1-\tau}$ . Thus, we get

$$\begin{aligned} \mathbb{E}[|S|] &= n \cdot (F((a - h(n))^b \sqrt{\mu n}) - F((a - g(n))^b \sqrt{\mu n})) \\ &= n \cdot (((a - g(n))^b \sqrt{\mu n})^{1-\tau} - ((a - h(n))^b \sqrt{\mu n})^{1-\tau}) \\ &= \mu^{\frac{1-\tau}{2}} n^{\frac{3-\tau}{2}} \left( (a - g(n))^{b(1-\tau)} - (a - h(n))^{b(1-\tau)} \right). \end{aligned} \quad (8)$$

We can now use the Taylor expansion of  $f(x) = (a - x)^c$  at 0 to obtain the bound  $f(x) = a^c - ca^{c-1}x \pm O(x^2)$ , which is valid for  $x \in o(1)$ . Since  $g(n), h(n) \in o(1)$  we can thus bound the above term in parentheses for  $c = b(1 - \tau)$  as

$$\begin{aligned} (a - g(n))^c - (a - h(n))^c &= -ca^{c-1}g(n) + ca^{c-1}h \pm O(g(n)^2 + h(n)^2) \\ &= -ca^{c-1}(g(n) - h(n) \pm O(g(n)^2 + h(n)^2)) \\ &= b(\tau - 1)a^{b(1-\tau)-1}(g(n) - h(n) \pm O(g(n)^2 + h(n)^2)). \end{aligned} \quad (9)$$

Equations (8) and (9) together yield the claim.  $\square$



Note that, if additionally  $h(n) \in o(g(n))$ , we can write the last factor as  $g(n)(1 - h(n)/g(n) \pm O(g(n))) = g(n)(1 \pm o(1))$  and obtain the following corollary.

**Corollary 3.3.** *Let  $a, b > 0$  be constants and let  $g(n), h(n)$  be functions of  $n$  such that  $g(n) \in o(1)$  and  $h(n) \in o(g(n))$ . Let  $S$  be the set of vertices with weight in  $[(a - g(n))^b \sqrt{\mu n}, (a - h(n))^b \sqrt{\mu n}]$ . Then*

$$\mathbb{E}[|S|] = g(n)n^{\frac{3-\tau}{2}} \cdot \mu^{\frac{1-\tau}{2}} b a^{b(1-\tau)-1} \cdot (1 \pm o(1)). \quad (10)$$

Consider again the weights  $w_\ell$  and  $w_u$  as given in Lemma 3.1 and let  $S$  be the set of vertices in  $[w_\ell, w_u]$ . Then Corollary 3.3 in particular implies that  $S$  contains  $\Theta(g(n) \cdot n^{\frac{3-\tau}{2}})$  vertices in expectation.

With this, we turn to our second goal mentioned above, namely that each box  $B_i$  should be sufficiently large.

**Lemma 3.4.** *Let  $B_1, \dots, B_{2k}$  be evenly spaced boxes of height  $h(n)$  and gap  $g(n)$  in  $\mathbb{T}^d$ . If  $g(n) \in o(1)$  then each box  $B_i$  has volume  $h/2^{d-1} \cdot (1 - o(1))$ .*

*Proof.* Recall that the height of  $B_i$  is  $h(n)$  while its extent in all other dimensions is  $\ell = \frac{1}{2} - g(n)$ . Thus its volume is  $h(n) \cdot (\frac{1}{2} - g(n))^{d-1} = h(n)/2^{d-1} \cdot (1 - 2g(n))^{d-1}$ . The claim follows from the fact that  $(1 - 2g(n))^{d-1}$  approaches 1 from below for  $n \rightarrow \infty$  as  $g(n) \in o(1)$  and  $d$  constant.  $\square$

Corollary 3.3 and Lemma 3.4 together tell us that the expected number of vertices in each box that have a weight in the desired range is in  $\Theta(h(n) \cdot g(n) \cdot n^{\frac{3-\tau}{2}})$ . Recall we want to choose  $h(n)$  and  $g(n)$  as small as possible such that each box still contains a vertex with high probability. We set  $h(n) = c \cdot n^{-\frac{3-\tau}{4}}$  and  $g(n) = c \cdot n^{-\frac{3-\tau}{4} + \varepsilon}$  for arbitrary constants  $c > 0$  and  $\varepsilon > 0$ . Note that this satisfies the condition  $h(n) \in o(g(n))$  of Corollary 3.3 and yields an expected number of  $\Theta(n^\varepsilon)$  vertices with the desired weight in each box. Since the number of vertices in a given box follows a binomial distribution and since  $n^\varepsilon \in \omega(\log(n))$ , we can apply a Chernoff bound to conclude that actual number of vertices matches the expected value (up to constant factors) with probability  $1 - O(n^{-c'})$  for any  $c' > 0$  [1, Corollaries 2.3 and 2.4]. Together with a union bound, it follows that every box contains  $\Theta(n^\varepsilon)$  vertices (and thus at least one vertex) with probability  $1 - O(2k \cdot n^{-c'})$ . By choosing  $g(n), h(n)$ , and  $k$  appropriately, we obtain the following theorem.

**Theorem 3.5.** *Let  $G^{(n)}$  be a  $d$ -dimensional GIRG with  $T = 0$ ,  $\mu > 0$  and  $\tau \in (2, 3)$  and let  $s > 0$  and  $\varepsilon > 0$  be arbitrary constants. Then, with high probability,  $G^{(n)}$  contains a co-matching of size  $s \cdot n^{\frac{3-\tau}{4} - \varepsilon}$  as induced subgraph.*

*Proof.* Let  $B_1, \dots, B_{2k}$  be evenly spaced boxes of height  $h(n) = c \cdot n^{-\frac{3-\tau}{4}}$  and gap  $g(n) = c \cdot n^{-\frac{3-\tau}{4} + \varepsilon}$  (for appropriately chosen  $c > 0$ , which will be determined later). Let  $w_\ell$  and  $w_u$  be defined as in Lemma 3.1. As argued above, Corollary 3.3 and Lemma 3.4 imply that, with high probability, each box  $B_i$  includes at least one vertex with weight in  $[w_\ell, w_u]$ . By Lemma 3.1 any set that contains exactly one vertex of each box forms a co-matching of size  $2k$ .

Recall that  $2k = 1/(g(n) + h(n))$ . Thus, we can choose  $c$  such that  $2k = s \cdot n^{\frac{3-\tau}{4} - \varepsilon}$ . Again, by the above argumentation, it follows that every box contains at least one vertex with probability  $1 - O(2kn^{-c'}) = 1 - O(n^{\frac{3-\tau}{4} - \varepsilon - c'})$  for any constant  $c' > 0$ . Choosing  $c'$  sufficiently large then yields the claim.  $\square$

This theorem together with Lemma 2.2 directly imply the following corollary.

**Corollary 3.6.** *Let  $G^{(n)}$  be a  $d$ -dimensional GIRG with  $T = 0$ ,  $\mu > 0$  and  $\tau \in (2, 3)$ , and let  $b > 0$  and  $\varepsilon > 0$  be arbitrary constants. Then, with high probability, the number of maximal cliques in  $G^{(n)}$  is at least  $b^{n^{(3-\tau)/4 - \varepsilon}}$ .*

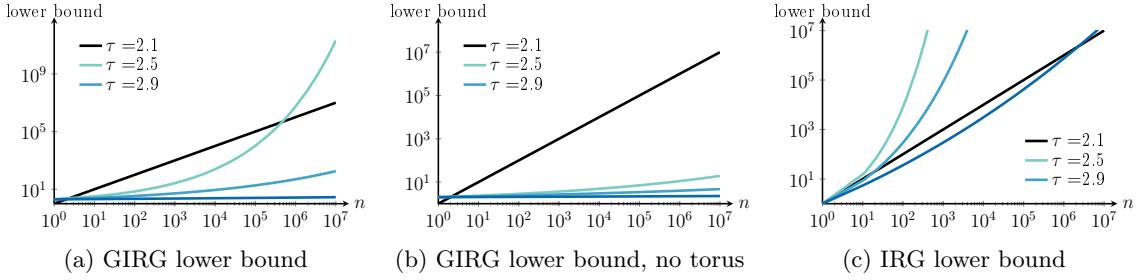


Figure 2: Lower bound on the number of maximal cliques of Corollary 3.6 (with  $b = 2$ ,  $\varepsilon \downarrow 0$ ), 3.7 (for  $\varepsilon \downarrow 0$ ,  $b = 2$  and  $C = 1$ ), 4.1 (for  $\varepsilon \downarrow 0$  and  $b = 2$ ) against  $n$  for different values of  $\tau$ . The black line is the line  $n$ .

Figure 2a shows this lower bound for  $b = 2$  against  $n$ . Interestingly, while Corollary 3.6 shows that the number of maximal cliques grows super-polynomially in  $n$ , for  $\tau > 2$ , this growth may still be slower than the linear slope  $n$  for large geometric networks. This is of particular importance as the smaller order terms of the number of maximal cliques contain terms of at least  $\Theta(n)$ . Indeed, the number of maximal 2-cliques is lower bounded by the number of vertices of degree 1, which scales linearly by Equation (4). Thus, for practical purposes, the dominant term could be the linear term instead of the super-polynomial term, especially if the degree exponent is close to 3.

### 3.2 GIRG with 2-Dimensional Square

Our previous lower bound for the number of maximal cliques relies on the toroidal structure of the underlying space. We now show that even if the vertex positions are constrained to be positioned in the square  $[0, 1]^2$  instead, the GIRG still contains a super-polynomial number of maximal cliques. In this setting, we will also switch from the infinity norm to the 2-norm. We will discuss possible extensions to other norms in Section 6.

**Theorem 3.7.** *For any  $\varepsilon > 0$  and  $b > 0$ , a 2-dimensional GIRG  $G^{(n)}$  on  $n$  vertices with vertex positions uniformly distributed over  $[0, 1]^2$  equipped with the 2-norm and  $T = 0$ ,  $\mu > 0$  and  $\tau \in (2, 3)$  contains with high probability at least*

$$Cb^n^{\frac{3-\tau}{10}-\varepsilon} \quad (11)$$

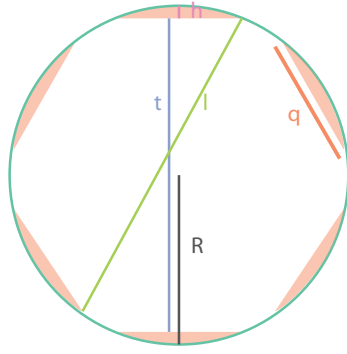
maximal cliques for some  $C > 0$ .

*Proof.* Let  $S$  be the set of vertices with weights within  $[a\sqrt{\mu n(1-c \cdot n^{-\beta})}, a\sqrt{\mu n}]$  for some  $0 < a < 1/4$ ,  $\beta > 0$  (and appropriately chosen  $c > 0$ , which will be determined later). By Corollary 3.3,

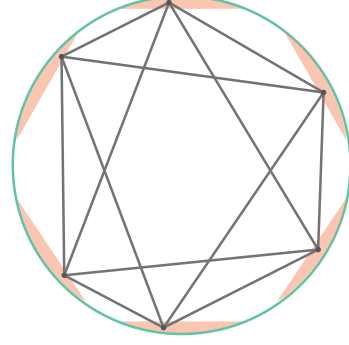
$$\mathbb{E}[|S|] = n^{(3-\tau)/2-\beta}(1+o(1)) \quad (12)$$

Let  $\mathcal{C}$  be a circle on  $[0, 1]^2$  of constant radius  $R < 1/4$ . We now create an even number of areas  $B_1, \dots, B_{2k}$  of height  $h(n)$ , evenly distributed over  $\mathcal{C}$  as illustrated in Figure 3a. That is, we consider  $2k$  identical and evenly spaced circular segments  $B_1, \dots, B_{2k}$  of height  $h$  and chord length  $q$ . We ensure that any pair of vertices in two opposite areas  $B_i$  and  $B_{i+k}$  are disconnected. That is, the distance  $t$  between the two ends of these areas should equal

$$t = a. \quad (13)$$



(a) A circle of radius  $R$  with several orange areas of height  $h(n)$ .



(b) Any set of vertices with one in each orange area form a co-matching.

Figure 3: Clique minus a matching in the 2-dimensional GIRG.

We also ensure that any pair of vertices in non-opposite areas connect. This means that the distance  $\ell$  between the rightmost part of  $B_i$  and the leftmost part of any non-opposite  $B_j$  is at most

$$\ell = a\sqrt{1 - c \cdot n^{-\beta}}. \quad (14)$$

The width  $q$  of one of the  $B_i$ 's is given by

$$\frac{q}{2} = \sqrt{R^2 - (R - h)^2} = \sqrt{h(2R - h)} = \sqrt{h(n)(h(n) + t)} = \sqrt{h(n)(h(n) + a)}. \quad (15)$$

Thus, the area of  $B_i$  is given by

$$c_1 h(n)^{3/2} \sqrt{h(n) + a}, \quad (16)$$

for some constant  $c_1 > 0$ . The probability that a given area contains no vertices from  $S$  is given by

$$\left(1 - c_1 h(n)^{3/2} \sqrt{h(n) + a}\right)^{|S|} = \exp(-c_2 n^{(3-\tau)/2-\beta} h(n)^{3/2})(1 + o(1)), \quad (17)$$

for some  $c_2 > 0$ .

We now calculate the maximal number of areas that we can pack on  $\mathcal{C}$ . The circumference of  $\mathcal{C}$  is  $2\pi R$ . The arc length of a single area is at most  $q\pi/2$ . Furthermore, the arc length of the section with a chord of length  $\ell$ ,  $a(\ell)$ , is given by

$$\begin{aligned} a(\ell) &= 2R \sin^{-1}(\ell/(2R)) = 2R \sin^{-1}\left(\frac{a\sqrt{1 - c \cdot n^{-\beta}}}{a + 2h(n)}\right) \\ &= 2R \sin^{-1}\left(1 - \frac{a + 2h(n) - a\sqrt{1 - c \cdot n^{-\beta}}}{a + 2h(n)}\right) \\ &= \pi R - 2^{3/2} R \sqrt{\frac{a + 2h(n) - a\sqrt{1 - c \cdot n^{-\beta}}}{a + 2h(n)}}(1 + o(1)), \end{aligned} \quad (18)$$

using the Taylor series of  $\sin^{-1}(1 - x)$  around  $x = 0$ . Now take  $h(n) = c \cdot n^{-\gamma(n)}$ . Then, by Equation (18),  $a(\ell)$  scales as

$$a(\ell) = \pi R - R\Theta\left(\sqrt{(1 - \sqrt{1 - c \cdot n^{-\beta}}) + c \cdot n^{-\gamma(n)}}\right) = \pi R - R\Theta(\sqrt{c \cdot n^{-\beta} + c \cdot n^{-\gamma(n)}}). \quad (19)$$

Now the arc length between two adjacent sections  $B_i$  and  $B_{i+1}$  is equal to  $\pi R - a(\ell)$ . This means that the arc length between  $B_i$  and  $B_{i+1}$  scales as  $\pi R - (\pi R - R\Theta(\sqrt{c \cdot n^{-\beta} + c \cdot n^{-\gamma(n)}})) = R\Theta(\sqrt{c \cdot n^{-\beta} + c \cdot n^{-\gamma(n)}})$ .

The maximal value of the number of possible areas  $2k$ , is the total circumference of  $\mathcal{C}$  divided by the arc length of an interval  $B_i$  and the arc length between  $B_i$  and  $B_{i+1}$ , which yields

$$2k = \frac{2\pi R}{R\Theta(\sqrt{c \cdot n^{-\beta} + c \cdot n^{-\gamma(n)}}) + \sqrt{h(n)(h(n) + a)}} \in \Theta\left(\min(n^{\beta/2}, n^{\gamma(n)/2})\right). \quad (20)$$

Thus, by choosing  $c$  correctly, we can let  $2k = s \cdot \min(n^{\beta/2}, n^{\gamma(n)/2})$  for any  $s > 0$ . When all  $i \in [2k]$  contain at least one vertex in  $S$ , any set of  $2k$  vertices with exactly one vertex in each  $B_i$  forms a co-matching, as illustrated in Figure 3b. Furthermore (17) shows that with high probability, all  $B_i$  are non-empty, as long as  $n^{(3-\tau)/2-\beta}h(n)^{3/2} \rightarrow \infty$  as  $n \rightarrow \infty$ .

We therefore choose  $\beta = (3 - \tau)/5 - \varepsilon$  and  $\gamma(n) = (3 - \tau)/5$ . Then, with high probability there is a co-matching of size  $2k = s \cdot n^{(3-\tau)/10-\varepsilon}$ . Thus, by Lemma 2.2 and choosing  $s$  sufficiently large yields that for fixed  $b > 0$  the number of maximal cliques can be bounded from below by

$$b^{n^{(3-\tau)/10-\varepsilon}} \quad (21)$$

□

Figure 2b shows the lower bound of Theorem 3.7 against  $n$ . As for the toroidal case, the super-polynomial growth may be dominated by lower-order linear terms.

### 3.3 Non-Threshold Case

We now show how our constructions extend to the non-threshold GIRG, where the connection probability is given by Equation (5) instead of Equation (6).

**Theorem 3.8.** *Let  $G^{(n)}$  be a  $d$ -dimensional GIRG on  $n$  vertices with  $T > 0$ ,  $\mu > 0$  and  $\tau \in (2, 3)$  and let  $s > 0$  be an arbitrary constant. Then, there exists an  $\varepsilon > 0$  such that, with high probability,  $G$  contains a co-matching of size  $s \cdot n^{(3-\tau)/5} \cdot (\varepsilon \log n)^{-(1/2)}$  as induced subgraph.*

*Proof.* As before, we consider  $2k$  boxes  $B_1, \dots, B_{2k}$  with height  $h(n)$  and gap  $g(n)$ , though now we choose

$$g(n) = h(n) \cdot (\varepsilon \log n)^{1/2} \quad \text{and} \quad h(n) = \frac{1}{2s} \cdot n^{-\frac{3-\tau}{5}},$$

for a constant  $\varepsilon > 0$ , which we determine below. We again focus on the vertex set  $S$  containing all vertices with weights in  $[w_\ell, w_u]$ , though our choice for  $w_u$  is slightly different. In particular, we choose

$$w_\ell = (1/2 - g(n))^{d/2} \sqrt{\mu n} \quad \text{and} \quad w_u = (1/2 - (T/d + 1)h)^{d/2} \sqrt{\mu n}.$$

Our goal now is to show that, with high probability, there exists at least one co-matching that contains one vertex from each box. That is, if  $M$  denotes the number of such co-matchings, we want to show that  $M > 0$  with high probability.

We start by bounding the number of vertices from  $S$  that lie in a given box  $B_i$ , denoted by  $S(B_i)$ . Since  $T$  and  $d$  are constants and  $(\varepsilon \log n)^{1/2} \in \omega(1)$ , we have  $(T/d + 1)h(n) \in o(g(n))$ , allowing us to bound  $\mathbb{E}[|S|]$  using Corollary 3.3, which yields

$$\mathbb{E}[|S|] \in \Theta\left(gn^{\frac{3-\tau}{2}}\right).$$

Moreover, since the vertices are distributed uniformly at random in the ground space, the expected fraction of vertices from  $S$  that lie in the box  $B_i$  is proportional to its volume, which is  $\Theta(h)$  according to Lemma 3.4. It follows that

$$\mathbb{E}[|S(B_i)|] \in \Theta\left(ghn^{\frac{3-\tau}{2}}\right) \in \Theta\left(h(n)^2 n^{\frac{3-\tau}{2}} (\varepsilon \log n)^{1/2}\right) \in \Theta\left(n^{(3-\tau)(1/2-2/5)} (\varepsilon \log n)^{1/2}\right).$$

Analogous to the proof of Lemma 3.4 we can apply a Chernoff bound to conclude that the number of vertices in  $S(B_i)$  matches the expected value (up to constant factors) with probability  $1 - O(n^{-c})$  for any  $c > 0$ . Note that the number of boxes is given by

$$2k = \frac{1}{g(n) + h(n)} = \frac{1}{h((\varepsilon \log n)^{1/2} + 1)} = \frac{2s \cdot n^{\frac{3-\tau}{5}}}{(\varepsilon \log n)^{1/2} + 1}, \quad (22)$$

which is at most  $n$ . Thus, applying the union bound yields that with high probability *every* box contains  $n' \in \Theta(n^{(3-\tau)(1/2-2/5)} (\varepsilon \log n)^{1/2})$  vertices. In the following, we implicitly condition on this event to happen. Now recall that a co-matching consisting of one vertex from each box forms if each vertex is adjacent to the vertices in all other boxes, *except* the vertex from the opposite box.

Despite the temperature, vertices in non-opposite boxes are still adjacent with probability 1, since the weight of two such vertices  $i$  and  $j$  is at least  $w_\ell$  and their distance is at most  $\frac{1}{2} - g(n)$  and, thus, according to Equation 5

$$p_{ij} = \min \left\{ \left( \frac{w_i w_j}{n\mu \|x_i - x_j\|_2^d} \right)^{1/T}, 1 \right\} \geq \min \left\{ \left( \frac{w_\ell^2}{n\mu (1/2 - g(n))^d} \right)^{1/T}, 1 \right\} = 1.$$

In contrast to the threshold case, however, the probability for vertices in opposite boxes to be adjacent is no longer 0. Since two such vertices  $i$  and  $j$  have distance at least  $\frac{1}{2} - h(n)$  and weight at most  $w_u = (\frac{1}{2} - (T/d + 1)h)^{d/2} \sqrt{\mu n}$ , we can bound the probability for them to be adjacent using Equation 5, which yields

$$\begin{aligned} p_{ij} &= \min \left\{ \left( \frac{w_i w_j}{n\mu \|x_i - x_j\|_2^d} \right)^{1/T}, 1 \right\} \\ &\leq \min \left\{ \left( \frac{w_u^2}{n\mu (\frac{1}{2} - h(n))^d} \right)^{1/T}, 1 \right\} \\ &= \min \left\{ \left( \frac{1 - (T/d + 1)2h(n)}{1 - 2h(n)} \right)^{d/T}, 1 \right\} \\ &= \min \left\{ \left( 1 - \frac{T}{d} \cdot \frac{2h(n)}{1 - 2h(n)} \right)^{d/T}, 1 \right\} \\ &\leq \min \left\{ \left( 1 - \frac{T}{d} \cdot 2h(n) \right)^{d/T}, 1 \right\}. \end{aligned}$$

Since  $1 - x \leq e^{-x}$ , we obtain  $p_{ij} \leq e^{-2h(n)}$ .

With this we are now ready to bound the probability  $\mathbb{P}(M > 0)$ , that at least one co-matching forms that contains one vertex from each box. To this end, we need to find one non-edge in each pair of opposite boxes, i.e., each such pair needs to contain two vertices (one from each box) that are not adjacent. Conversely, the only way to *not* find a co-matching is if there exists one pair of opposite

boxes such that all vertices in one box are adjacent to all vertices in the other. This happens with probability at most  $(p_{ij})^{(n')^2}$ . Since there are  $k$  pairs of opposite boxes, applying the union bound yields

$$\mathbb{P}(M = 0) \leq k(p_{ij})^{(n')^2} \leq k \exp(-2h(n)(n')^2). \quad (23)$$

Now recall that  $n' \in \Theta(n^{(3-\tau)(1/2-2/5)}(\varepsilon \log n)^{1/2})$  and that  $h(n) = 1/(2s) \cdot n^{-(3-\tau)/5}$ . Consequently, we obtain

$$\begin{aligned} \mathbb{P}(M = 0) &\leq k \exp\left(-\Theta\left(n^{-(3-\tau)/5} \cdot n^{(3-\tau)(1-4/5)} \cdot \varepsilon \log n\right)\right) \\ &= k \exp(-\Theta(\varepsilon \log n)) \\ &= kn^{-\Theta(\varepsilon)}. \end{aligned}$$

Moreover, since  $k \in O(n^{(3-\tau)/5})$  (see Equation 22), we have

$$\mathbb{P}(M = 0) \in O\left(n^{(3-\tau)/5-\Theta(\varepsilon)}\right),$$

meaning, for sufficiently large  $n$ , we can choose  $\varepsilon$  such that  $\mathbb{P}(M = 0) \in O(n^{-1})$  and, conversely,  $\mathbb{P}(M > 0) = 1 - O(n^{-1})$ . So with high probability there exists at least one co-matching of size

$$2k = \frac{2s \cdot n^{\frac{3-\tau}{5}}}{(\varepsilon \log n)^{1/2} + 1} \geq \frac{2s \cdot n^{\frac{3-\tau}{5}}}{2(\varepsilon \log n)^{1/2}} = \frac{s \cdot n^{\frac{3-\tau}{5}}}{(\varepsilon \log n)^{1/2}},$$

where the inequality holds for sufficiently large  $n$ .  $\square$

Together with Lemma 2.2 we obtain the following corollary.

**Corollary 3.9.** *Let  $G^{(n)}$  be a  $d$ -dimensional GIRG on  $n$  vertices with  $T > 0$ ,  $\mu > 0$  and  $\tau \in (2, 3)$  and let  $b > 0$  be an arbitrary constant. Then there exists an  $\varepsilon > 0$  such that, with high probability, the number of maximal cliques in  $G^{(n)}$  is at least*

$$bn^{(3-\tau)/5 \cdot (\varepsilon \log n)^{-(1/2)}}. \quad (24)$$

We can extend Theorem 3.7 to non-zero temperature in a very similar fashion (proof is in Appendix A)

**Theorem 3.10.** *For any  $\varepsilon > 0$  and  $b > 1$ , a 2-dimensional GIRG  $G^{(n)}$  on  $n$  vertices with  $T > 0$  and vertex positions uniformly distributed over  $[0, 1]^2$  contains with high probability at least*

$$bn^{\frac{3-\tau}{12} - \varepsilon \log(n)^{1-\varepsilon}} \quad (25)$$

*maximal cliques.*

## 4 Inhomogeneous Random Graphs (IRG)

We now turn to a random graph model that is scale-free, but does not contain a source of geometry, the inhomogeneous random graph (IRG), or Chung-Lu random graph [8]. We show that also in this

model, the number of maximal cliques scales super-polynomially in the network size  $n$ . Again, every vertex  $i$  draws its weight  $w_i$  independently from the power-law distribution (4), where we will again assume that  $\tau \in (2, 3)$ . Then, all pairs of vertices  $u$  and  $v$  connect independently with probability

$$p(w_u, w_v) = \min\left(\frac{w_u w_v}{\mu n}, 1\right), \quad (26)$$

where  $\mu$  controls the expected average degree.

To show a lower bound on the number of maximal cliques, we make use of the fact that an IRG contains a not too small rather dense subgraph with high probability. The following theorem is obtained by looking just at the subgraph induced by vertices with weights in a certain range. We chose the specific range to satisfy three criteria. First, the range is sufficiently large, such that the subgraph contains many vertices. Second, the range is sufficiently small such that all vertex pairs in the subgraph are connected with a similar probability. And third, the weights are large enough such that a densely connected subgraph forms, but not so large that the vertices merge into a single clique.

**Theorem 4.1.** *Let  $G^{(n)}$  be an IRG on  $n$  vertices with  $\tau \in (2, 3)$  and  $\mu > 0$  and let  $b > 1$  and  $\varepsilon \in (0, \frac{3-\tau}{4})$  be arbitrary constants. Then, the expected number of maximal cliques in  $G^{(n)}$  is in*

$$\Omega(b^{n^{(3-\tau)/4-\varepsilon} \log n}). \quad (27)$$

*Proof.* We show that already the subgraph  $G'$  induced by the vertices in a certain weight range has the claimed expected number of maximal cliques. To define  $G'$ , we consider weights in  $[w_\ell, w_u]$  with  $w_\ell = \sqrt{(1-g(n))\mu n}$  and  $w_u = \sqrt{(1-h(n))\mu n}$ . To abbreviate notation, let

$$\gamma(n) = n^{\frac{3-\tau}{4}}.$$

For constants  $a$  and  $c$  we determine later, we choose  $g(n)$  and  $h(n)$  as

$$g(n) = ah(n) \quad \text{and} \quad h(n) = cn^\varepsilon \log(n)\gamma(n)^{-1}.$$

Note that  $w_\ell < w_u$  if and only if  $a > 1$ . Let  $n'$  be the number of vertices in  $G'$ . From Lemma 3.2 it follows

$$\mathbb{E}[n'] \in \Theta(\gamma(n)^2 \cdot n^\varepsilon \log(n)\gamma(n)^{-1}) \in \Theta(n^\varepsilon \gamma(n) \log n).$$

As every vertex has independently the same probability to be in  $G'$ , a Chernoff bound implies that  $n' \in \Theta(n^\varepsilon \gamma(n) \log n)$  holds with high probability. Thus, in the following, we implicitly condition on this event to happen.

To give a lower bound on the number of maximal cliques in  $G'$ , we only count the number  $N_k$  of maximal cliques of size  $k$  with

$$k = \frac{3\varepsilon}{c} n^{-\varepsilon} \gamma(n).$$

We note that this is the same constant  $c$  as in the definition of  $h(n)$  above. As the number of maximal cliques in  $G'$  is a lower bound for the number of maximal cliques in  $G^{(n)}$ , we lower bound the expectation of  $N_k$ , by the expected number of maximal cliques in  $G'$ ,

$$\mathbb{E}[N_k] \geq \binom{n'}{k} \mathbb{P}(C \text{ is a clique}) \mathbb{P}(C \text{ maximal} \mid C \text{ is a clique}). \quad (28)$$

In the following, we give estimates for the three terms individually.

We start with the event that  $C$  is a clique. Due to the lower and upper bound on the weights in  $G'$ , it follows that any pair of vertices in  $G'$  is connected with probability at least  $p_\ell = 1 - g(n)$  and at most  $p_u = 1 - h(n)$ . Thus, for a fixed subset  $C$  of vertices of size  $|C| = k$ , the probability that all  $k$  vertices are pairwise connected is at least  $p_\ell^{k(k-1)/2} = (1 - g(n))^{k(k-1)/2}$ . As  $g(n)$  goes to 0 for growing  $n$  and  $1 - x \in \Omega(\exp(-x))$  in this case, we get

$$\mathbb{P}(C \text{ is a clique}) \geq (1 - g(n))^{k(k-1)/2} \in \Omega\left(\exp\left(-\frac{gk(k-1)}{2}\right)\right). \quad (29)$$

For  $C$  to be a maximal clique (conditioning on it being a clique), additionally no other vertex can be connected to all vertices from  $C$ . This probability is at least  $(1 - p_u^k)^{n'-k} = (1 - (1 - h(n))^k)^{n'-k}$ . As  $1 - x \leq \exp(-x)$ , it follows that  $(1 - h(n))^k \leq \exp(-hk) = n^{-3\varepsilon}$ , where the last equality follows from plugging in the values we chose for  $h(n)$  and  $k$ . Again using  $1 - x \in \Omega(\exp(-x))$  for sufficiently small  $x$ , we can conclude that

$$\begin{aligned} \mathbb{P}(C \text{ maximal} \mid C \text{ is a clique}) &\geq (1 - (1 - h(n))^k)^{n'-k} \\ &\geq (1 - n^{-3\varepsilon})^{n'-k} \\ &\in \Omega\left(\exp(-n^{-3\varepsilon}(n' - k))\right). \end{aligned} \quad (30)$$

Finally, for the binomial coefficient, we get

$$\binom{n'}{k} \geq \left(\frac{n'}{k}\right)^k = \exp\left(\log\left(\frac{n'}{k}\right)k\right). \quad (31)$$

Our goal is to show that  $\log(\mathbb{E}[N_k]) \geq n^{(3-\tau)/4-\varepsilon} \log(n) \log(b)$ . Thus, plugging Equations (29), (30), and (31) into the logarithm of (28) yields that we need to show that for every constant  $b > 1$ , we can choose the constants  $a > 1$  and  $c$  in the definitions of  $g(n)$  and  $h(n)$  such that

$$\log\left(\frac{n'}{k}\right)k - \frac{gk(k-1)}{2} - n^{-3\varepsilon}(n' - k) \stackrel{\text{(to be shown)}}{>} n^{-\varepsilon} \gamma(n) \log n \log b = n^{\frac{3-\tau}{4}-\varepsilon} \log n \log b. \quad (32)$$

This can be achieved by simply plugging in the values for  $n'$ ,  $k$ , and  $g(n)$ . For the first (and only positive) term, we obtain

$$\begin{aligned} \log\left(\frac{n'}{k}\right)k &= \log\left(\frac{\Theta(n^\varepsilon \log(n) \gamma(n))}{3\varepsilon/cn^{-\varepsilon} \gamma(n)}\right) \frac{3\varepsilon}{c} n^{-\varepsilon} \gamma(n) \\ &= \log(n^{2\varepsilon} \Theta(\log n)) \frac{3\varepsilon}{c} n^{-\varepsilon} \gamma(n) \end{aligned}$$

and thus for sufficiently large  $n$

$$\geq \frac{6\varepsilon^2}{c} n^{-\varepsilon} \gamma(n) \log n.$$

For the negative terms, we start with the latter and obtain

$$n^{-3\varepsilon}(n' - k) \in \Theta(n^{-3\varepsilon} n^\varepsilon \gamma(n) \log n) = \Theta(n^{-2\varepsilon} \gamma(n) \log n).$$

This is asymptotically smaller than the positive term and can thus be ignored. For the other negative term, first note that  $gk = 3a\varepsilon \log n$ . Thus, we obtain

$$\frac{gk(k-1)}{2} \leq \frac{3a\varepsilon}{2} \log(n)k = \frac{3a\varepsilon}{2} \log(n) \frac{3\varepsilon}{c} n^{-\varepsilon} \gamma(n) = \frac{9a\varepsilon^2}{2c} n^{-\varepsilon} \gamma(n) \log n.$$



Together with the positive term, we obtain that for sufficiently large  $n$ , it holds

$$\begin{aligned} \log\left(\frac{n'}{k}\right)k - \frac{gk(k-1)}{2} &\geq \frac{6\varepsilon^2}{c}n^{-\varepsilon}\gamma(n)\log n - \frac{9a\varepsilon^2}{2c}n^{-\varepsilon}\gamma(n)\log n \\ &= \left(6 - \frac{9a}{2}\right)\frac{\varepsilon^2}{c}n^{-\varepsilon}\gamma(n)\log n. \end{aligned}$$

With this, we can choose  $a > 1$  such that the first factor is positive and we can choose  $c$  such that  $\varepsilon^2/c = \log b$ , which proves (32), as then

$$\left(6 - \frac{9a}{2}\right)\frac{\varepsilon^2}{c}n^{-\varepsilon}\gamma(n)\log n = Cn^{(3-\tau)/4-\varepsilon}\log(n)\log(b).$$

for some  $C > 0$ . □

Figure 2c shows that the lower bound provided by Theorem 4.1 may still be smaller than linear for networks that are quite large, especially when  $\tau \approx 3$ .

#### 4.1 Small Maximal Cliques are Rare

We now focus on the maximal cliques of a fixed size in the IRG. How many maximal cliques of size  $k$  are present in an IRG?

Let  $N(K_k)$  denote the number of maximal cliques of size  $k$ . Furthermore, let  $M_n(\varepsilon)$  denote

$$M_n(\varepsilon) = \{(v_1, \dots, v_k) : w_i \in [\varepsilon, 1/\varepsilon](\mu n)^{\frac{\tau-2}{\tau-1}} \text{ for } i = 1, 2 \text{ and } w_i \in [\varepsilon, 1/\varepsilon](\mu n)^{\frac{1}{\tau-1}} \forall i \in \{3, \dots, k\}\}. \quad (33)$$

Thus,  $M_n(\varepsilon)$  is the set of sets of  $k$  vertices such that two vertices have weight proportional to  $n^{(\tau-2)/(\tau-1)}$ , and all other vertices have weights proportional to  $n^{1/(\tau-1)}$ . Denote the number of maximal  $k$ -cliques with sets of vertices in  $M_n(\varepsilon)$  by  $N(K_k, M_n(\varepsilon))$ . Then, the following theorem shows that these ‘typical’ maximal cliques are asymptotically all maximal cliques. Furthermore, it shows that all maximal cliques of size  $k > 2$  occur equally frequently in scaling, and they also appear on the same types of vertices. Here we use  $\xrightarrow{\mathbb{P}}$  to denote convergence in probability.

**Theorem 4.2** (Maximal clique localization). *Let  $G^{(n)}$  be an IRG on  $n$  vertices with  $\tau \in (2, 3)$  and  $\mu > 0$ . For any fixed  $k \geq 3$ ,*

(i) *For any  $\varepsilon_n$  such that  $\lim_{n \rightarrow \infty} \varepsilon_n = 0$ ,*

$$\frac{N(K_k, M_n(\varepsilon_n))}{N(K_k)} \xrightarrow{\mathbb{P}} 1. \quad (34)$$

(ii) *Furthermore, for any fixed  $0 < \varepsilon < 1$ ,*

$$\mathbb{E}[N(K_k, M_n(\varepsilon))] = \Theta(n^{(3-\tau)(2\tau-3)/(\tau-1)}). \quad (35)$$

Theorem 4.2(i) states that asymptotically all maximal  $k$ -cliques are formed between two vertices of weights proportional to  $n^{(\tau-2)/(\tau-1)}$  and all other vertices of weights proportional to  $n^{1/(\tau-1)}$ . Theorem 4.2(ii) then shows that there are proportional to  $n^{(3-\tau)(2\tau-3)/(\tau-1)}$  such maximal  $k$ -cliques. As visualized in Figure 4, this scaling is significantly smaller than the scaling of the total number of  $k$ -cliques, which scales as  $n^{k/2(3-\tau)}$  [16]. Interestingly, the scaling of the number of maximal

cliques is  $k$ -independent, contrary to the total number of cliques. In particular, the number of  $k$  maximal cliques is always  $o(n)$ , contrary to the number of  $k$ -cliques which scales larger than  $n$  when  $\tau < 3 - 2/k$ . This shows once more that the large number of maximal cliques in the IRG is caused by extremely large maximal cliques, as fixed-size maximal cliques are only linearly many.

To prove this theorem, we need the following technical lemma, which is proven in Appendix B:

**Lemma 4.3.** *When  $\tau \in (2, 3)$ , then*

$$\int_0^1 \cdots \int_0^1 x_3^{1-\tau} \cdots x_k^{1-\tau} \int_0^\infty \int_{x_1}^\infty x_1^{k-1-\tau} x_2^{1-\tau} \prod_{i=3}^k \min(x_2 x_i, 1) e^{-\mu^{1-\tau} x_1 x_2^{\tau-2}} dx_2 dx_1 \cdots dx_k < \infty. \quad (36)$$

Furthermore, we need a lemma that bounds the probability that a given clique on vertices of weights  $x_1 \leq x_2 \leq \cdots \leq x_k$  is maximal:

**Lemma 4.4.** *Let  $G$  be an IRG with  $\tau \in (2, 3)$  and  $\mu > 0$ . Then, the probability that a given clique between  $k$  vertices of weights  $x_1 \leq x_2 \leq \cdots \leq x_k$  is maximal is bounded by*

$$\begin{aligned} & \exp\left(-C_1 n^{2-\tau} \mu^{1-\tau} x_1 x_2^{\tau-2}\right) (1 + o(1)) \leq \mathbb{P}(\text{clique on weights } x_1, \dots, x_k \text{ maximal}) \\ & \leq \exp\left(-C_2 n^{2-\tau} \mu^{1-\tau} x_1 x_2^{\tau-2}\right), \end{aligned} \quad (37)$$

for some  $0 < C_1 \leq C_2 < \infty$ .

*Proof.* When  $x_1 \leq x_2 \leq \cdots \leq x_k$ , we can compute the probability that this  $k$  clique is part of a larger clique with a randomly chosen vertex as

$$\begin{aligned} & \int_1^\infty w^{-\tau} \prod_{i \in [k]} \min\left(\frac{w x_i}{\mu n}, 1\right) dw \\ & = \frac{x_1 \cdots x_k}{(\mu n)^k} \int_1^{\mu n/x_k} w^{k-\tau} dw + \frac{x_1 \cdots x_{k-1}}{(\mu n)^{k-1}} \int_{\mu n/x_k}^{\mu n/x_{k-1}} w^{k-1-\tau} dw \\ & \quad + \cdots + \frac{x_1 x_2}{(\mu n)^2} \int_{\mu n/x_3}^{\mu n/x_2} w^{2-\tau} dw + \frac{x_1}{\mu n} \int_{\mu n/x_2}^{\mu n/x_1} w^{1-\tau} dw + \int_{\mu n/x_1}^\infty w^{-\tau} dw \\ & = c_k \frac{x_1 \cdots x_k}{(\mu n)^k} \left(\frac{\mu n}{x_k}\right)^{k+1-\tau} + \cdots + c_2 \frac{x_1}{\mu n} \left(\frac{\mu n}{x_2}\right)^{2-\tau} + c_1 \left(\frac{\mu n}{x_1}\right)^{1-\tau}, \end{aligned} \quad (38)$$

for some  $c_1, \dots, c_k > 0$ . When  $x_1 \leq x_2 \leq \cdots \leq x_k$ , this term becomes

$$(\mu n)^{1-\tau} \sum_{l=1}^k c_l x_l^{-l+\tau} \prod_{i < l} x_i. \quad (39)$$

The ratio between two consecutive terms of this summation equals

$$\frac{x_l^{\tau-l} x_1 \cdots x_{l-1}}{x_{l+1}^{\tau-l-1} x_1 \cdots x_l} = \left(\frac{x_l}{x_{l+1}}\right)^{\tau-l-1}. \quad (40)$$

Now as  $x_l \leq x_{l+1}$  and  $\tau \in (2, 3)$ , this ratio is larger than 1 for  $l \geq 2$ , and smaller than one for  $l = 1$ . This means that the summation can be dominated by

$$(\mu n)^{1-\tau} \sum_{l=1}^k c_l x_l^{-l+\tau} \prod_{i < l} x_i \leq C (\mu n)^{1-\tau} x_1 x_2^{\tau-2}, \quad (41)$$

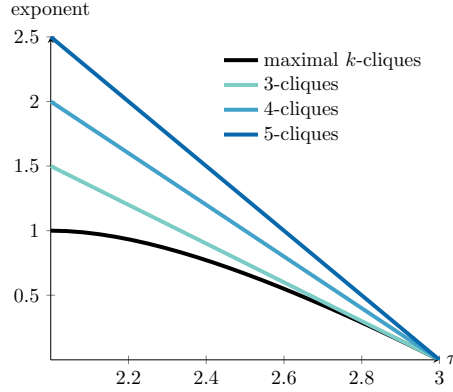


Figure 4: Scaling of the number of maximal  $k$ -cliques, and the total number of (not necessarily maximal) 3,4,5-cliques.

for some  $C > 0$ .

Thus, the probability that a clique on vertices with weights  $x_1, \dots, x_k$  is maximal can be upper bounded by

$$\begin{aligned} \mathbb{P}((x_1, \dots, x_k) \text{ clique maximal}) &\leq \left(1 - C(\mu n)^{1-\tau} x_1 x_2^{\tau-2}\right)^n \\ &\leq \exp\left(-C n^{2-\tau} \mu^{1-\tau} x_1 x_2^{\tau-2}\right). \end{aligned} \quad (42)$$

We lower bound the probability that the clique is maximal by using that

$$(\mu n)^{1-\tau} \sum_{l=1}^k c_l x_l^{-l+\tau} \prod_{i<l} x_i \geq c_2 (\mu n)^{1-\tau} x_1 x_2^{\tau-2}. \quad (43)$$

Thus,

$$\begin{aligned} \mathbb{P}((x_1, \dots, x_k) \text{ clique maximal}) &\geq \left(1 - c_2 (\mu n)^{1-\tau} x_1 x_2^{\tau-2}\right)^n \\ &\geq \exp\left(-c_2 n^{2-\tau} \mu^{1-\tau} x_1 x_2^{\tau-2} / (1 + c_2 n^{1-\tau} \mu^{1-\tau} x_1 x_2^{\tau-2})\right) \\ &= \exp\left(-c_2 n^{2-\tau} \mu^{1-\tau} x_1 x_2^{\tau-2}\right) (1 + o(1)). \end{aligned} \quad (44)$$

□

Now we are ready to prove Theorem 4.2:

*Proof of Theorem 4.2.* Fix  $\ell_i \leq u_i$  for  $i \in [k]$ . We now compute the expected number of maximal  $k$ -cliques in which the vertices have weights  $n^{(\tau-2)/(\tau-1)}[\ell_i, u_i]$  for  $i = 1, 2$ , and  $n^{1/(\tau-1)}[\ell_i, u_i]$  for  $i \geq 3$ .

We bound the expected number of such maximal copies of  $K_k$  by

$$\begin{aligned} & \sum_{\mathbf{v}} \mathbb{E} \left[ I(K_k, \mathbf{v}) \mathbb{1}_{\{w_{v_i} \in [\ell_i, u_i] n^{(\tau-2)/(\tau-1)}, i=1,2, w_{v_i} \in [\ell_i, u_i] n^{1/(\tau-1)}, i \geq 3\}} \right] \\ &= n^k \int_{\ell_1 n^{(\tau-2)/(\tau-1)}}^{u_1 n^{(\tau-2)/(\tau-1)}} \int_{\ell_2 n^{(\tau-2)/(\tau-1)}}^{u_2 n^{(\tau-2)/(\tau-1)}} \cdots \int_{\ell_k n^{1/(\tau-1)}}^{u_k n^{1/(\tau-1)}} (x_1 \cdots x_k)^{-\tau} \prod_{1 \leq i < j \leq k} \min\left(\frac{x_i x_j}{n}, 1\right) \\ & \quad \cdot \mathbb{P}((x_1, \dots, x_k) \text{ clique maximal}) dx_k \cdots dx_1, \end{aligned}$$

where  $I(K_k, \mathbf{v})$  is the indicator that a maximal  $k$ -clique is present on vertices  $\mathbf{v}$ , and the sum over  $\mathbf{v}$  is over all possible sets of  $k$  vertices. Now the probability that a clique is maximal can be upper bounded as in Lemma 4.4.

We bound the minimum in (45) by

- (a)  $x_i x_j / n$  for  $\{i, j\} = \{1, 2\}$  or  $i = 1, j \geq 3$ ;
- (b) 1 for  $i, j \geq 3$ .

Making the change of variables  $x_i = y_i n^{1/(\tau-1)}$  for  $i = 3, \dots, k$  and  $x_i = y_i / n^{(\tau-2)/(\tau-1)}$  otherwise, we obtain the bound

$$\begin{aligned} & \sum_{\mathbf{v}} \mathbb{E} \left[ I(K_k, \mathbf{v}) \mathbb{1}_{\{w_{v_i} \in [\ell_i, u_i] n^{(\tau-2)/(\tau-1)}, i=1,2, w_{v_i} \in [\ell_i, u_i] n^{1/(\tau-1)}, i \geq 3\}} \right] \\ & \leq \tilde{K} n^k n^{2(\tau-2)/(\tau-1) - k + 1} \\ & \quad \times \int_{\ell_1}^{u_1} \int_{y_1}^{u_2} \int_{\ell_3}^{u_3} \cdots \int_{\ell_k}^{u_k} y_1^{2-\tau} y_2^{1-\tau} y_3^{1-\tau} \cdots y_k^{1-\tau} \prod_{j \geq 3} \min(y_2 y_j, 1) \exp(-\mu^{1-\tau} y_1 y_2^{\tau-2}) dy_k \cdots dy_1, \end{aligned} \tag{45}$$

for some  $\tilde{K} > 0$ . Because the weights are sampled i.i.d. from a power-law distribution, the maximal weight  $w_{\max}$  satisfies that for any  $\eta_n \rightarrow 0$ ,  $w_{\max} \leq n^{1/(\tau-1)}/\eta_n$  with high probability. Thus, we may assume that  $u_i \leq 1/\eta_n$  when  $i \geq 3$ . Now suppose that at least one vertex has weight smaller than  $\varepsilon_n n^{(\tau-2)/(\tau-1)}$  for  $i = 1, 2$  or smaller than  $\varepsilon_n n^{1/(\tau-1)}$  for  $i \geq 3$ . This corresponds to taking  $u_i = \varepsilon_n$  and  $\ell_i = 0$  for at least one  $i$ , or at least one integral in (45) with interval  $[0, \varepsilon_n]$ . Similarly, when vertex 1 or 2 has weight higher than  $1/\varepsilon_n n^{(\tau-2)/(\tau-1)}$ , this corresponds to taking  $\ell_i = 1/\varepsilon_n$  and  $u_i = \infty$  for  $i = 1$  or 2, or at least one integral in (45) with interval  $[1/\varepsilon_n, \infty]$ . Lemma 4.3 then shows that these integrals tends to zero when choosing  $u_i = \eta_n$  fixed for  $i \geq 3$  and  $\varepsilon_n \rightarrow 0$ . Thus, choosing  $\eta_n \rightarrow 0$  sufficiently slowly compared to  $\varepsilon_n$  yields that

$$\sum_{\mathbf{v}} \mathbb{E} \left[ I(K_k, \mathbf{v}) \mathbb{1}_{\{\mathbf{v} \notin \Gamma_n(\varepsilon_n, \eta_n)\}} \right] \in o((n^{(3-\tau)(2\tau-3)/(\tau-1)})), \tag{46}$$

where

$$\Gamma_n(\varepsilon_n, \eta_n) = \{(v_1, \dots, v_k) : w_{v_i} \in n^{(\tau-2)/(\tau-1)}[\varepsilon_n, 1/\varepsilon_n], i = 1, 2, n^{1/(\tau-1)}[\varepsilon_n, 1/\eta_n]\}. \tag{47}$$

Let  $\bar{\Gamma}_n(\varepsilon_n, \eta_n)$  be the complement of  $\Gamma_n(\varepsilon_n, \eta_n)$ . Denote the number of maximal cliques with vertices in  $\bar{\Gamma}_n(\varepsilon_n, \eta_n)$  by  $N(K_k, \bar{\Gamma}_n(\varepsilon_n, \eta_n))$ . Since  $w_{\max} \leq n^{1/(\tau-1)}/\eta_n$  with high probability,  $\Gamma_n(\varepsilon_n, \eta_n) = M_n(\varepsilon_n)$  with high probability. Therefore, with high probability,

$$N(K_k, \bar{M}_n(\varepsilon_n)) = N(K_k, \bar{\Gamma}_n(\varepsilon_n, \eta_n)), \tag{48}$$

where  $N(K_k, \bar{M}_n(\varepsilon_n))$  denotes the number of maximal  $k$ -cliques on vertices not in  $M_n(\varepsilon_n)$ . By (46) and the Markov inequality, we have for all  $\epsilon > 0$

$$\lim_{n \rightarrow \infty} \mathbb{P} \left( \left| \frac{N(K_k, \bar{M}_n(\varepsilon_n))}{n^{(3-\tau)(2\tau-3)/(\tau-1)}} \right| > \epsilon \right) = 0. \quad (49)$$

Furthermore, Lemma 4.3 combined with the lower bound in (45) shows that when choosing  $u_i = 1/\varepsilon$  and  $\ell_i = \varepsilon$  for some fixed  $\varepsilon > 0$  for all  $i$ ,

$$\mathbb{E}[N(K_k, M_n(\varepsilon))] \in \Theta(n^{(3-\tau)(2\tau-3)/(\tau-1)}). \quad (50)$$

Thus, for fixed  $\varepsilon > 0$ ,

$$N(K_k) = N(K_k, M_n(\varepsilon)) + N(K_k, \bar{M}_n(\varepsilon)) = \Theta_p(n^{(3-\tau)(2\tau-3)/(\tau-1)}), \quad (51)$$

which shows that

$$\frac{N(K_k, M_n(\varepsilon_n))}{N(K_k)} \xrightarrow{\mathbb{P}} 1, \quad (52)$$

as required. This completes the proof of Theorem 4.2.  $\square$

## 5 Experiments

As mentioned in the introduction, empirical evidence suggests that the number of maximal cliques in IRGs and GIRGs is small [3]. In fact, all generated networks with  $n = 50$  k nodes and expected average degree 10 have fewer maximal cliques than edges. This stands in stark contrast to our super-polynomial lower bounds. This discrepancy probably comes from the fact that  $n = 50$  k is low enough that a linear lower-order term dominates the super-polynomial terms. In this section, we complement our theoretical lower bounds with experiments<sup>1</sup> with an  $n$  that is sufficiently large to make the super-polynomial terms dominant. Additionally, we consider dense and super-dense Erdős–Rényi graphs.

### 5.1 Cliques in the Dense Subgraph of GIRGs and IRGs

Our theoretical lower bounds are based on the existence of a dense subgraph among the vertices with weights  $\Theta(\sqrt{n})$ . To experimentally observe the super-polynomial scaling, we generate IRGs and GIRGs restricted to vertices of high weight. This restriction lets us consider much larger values of  $n$ . In the following, we first describe the exact experiment setup, before describing and discussing the results.

**Experiment Setup.** We generate IRGs and GIRGs with varying number of vertices  $n$  and deterministic power-law weights where the  $v$ th vertex has weight

$$w_v = \left(\frac{n}{v}\right)^{\frac{1}{\tau-1}}.$$

Note that the minimum weight is  $w_n = 1$ .

<sup>1</sup>The corresponding code is available at: <https://github.com/thobl/maximal-cliques-scale-free-rand-graph>

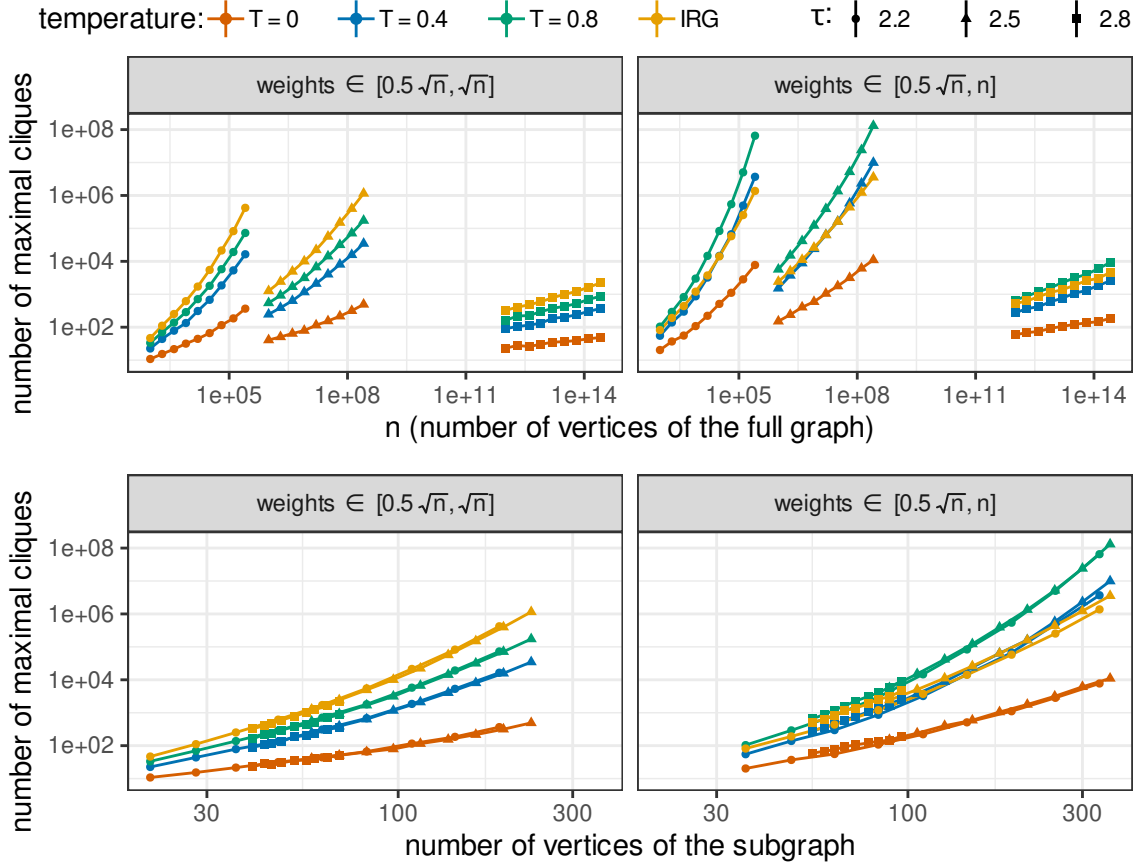


Figure 5: The number of maximal cliques of the dense subgraph of GIRGs and IRGs. The considered subgraphs contain all vertices with weights in  $[0.5\sqrt{n}, \sqrt{n}]$  (left column) and  $[0.5\sqrt{n}, n]$  (right column). The top and bottom plots show the number of cliques with respect to the size of the full graph, and with respect to the size of the considered subgraph, respectively. All axes are logarithmic. Each point is the average of 10 sampled graphs.

We use the power-law exponents  $\tau \in \{2.2, 2.5, 2.8\}$  and for GIRGs we consider the temperatures  $T \in \{0, 0.4, 0.8\}$  and dimension  $d = 1$ . For each parameter setting, we consider two subgraphs: The subgraph induced by vertices with  $0.5\sqrt{n} \leq w_i \leq \sqrt{n}$  and within the larger interval  $0.5\sqrt{n} \leq w_i \leq n$ . In preliminary experiments, we also tried constant factors other than 0.5, yielding comparable results.

As connection probability for the IRGs between the  $u$ th and  $v$ th vertex, we use  $\min\{1, w_u w_v / n\}$ , i.e., vertices of weight 1 have connection probability  $1/n$  and vertices of weight at least  $\sqrt{n}$  are deterministically connected. For GIRGs, we choose the constant factor  $\mu$  in Equation (5) such that we obtain the same expected<sup>2</sup> average degree as for the corresponding IRG in the considered subgraph. For each of these configurations, we generate 10 graphs. Figure 5 shows the average.

<sup>2</sup>We do not sample the positions before computing the expected average degree but we compute the expectation with respect to random positions.

**General Observations.** One can clearly see in Figure 5 (top row) that the scaling of the number of cliques depending on the graph size is super-polynomial (upward curves in a plot with logarithmic axes). Thus, on the one hand, this agrees with our theoretical analysis. On the other hand, the plots also explain why previous experiments [3] showed a small number of cliques: While the scaling is super-polynomial, the constant factors are quite low. In the top-left plot for  $\tau = 2.5$ , more than 200 M nodes are necessary to get just barely above 1 M maximal cliques in the dense subgraph. For  $\tau = 0.8$  this is even more extreme with  $n = 200$  T yielding only 10 k maximal cliques. Thus, unless we deal with huge graphs, the maximal cliques in the dense part of the graph are dominated by the number of cliques in the sparser parts, despite the super-polynomial growth of the former.

**Effect of the Power-Law Exponent  $\tau$ .** The top plots of Figure 5 show that a smaller power-law exponent  $\tau$  leads to more maximal cliques. The bottom plots show the number of cliques with respect to the size of the dense subgraph and not with respect to the size of the full graph. One can see that the difference for the different power-law exponents solely comes from the fact that the dense subgraph is larger for smaller  $\tau$ . For the same size of the dense subgraph, the scaling is almost independent of the power-law exponent.

**Effect of the Geometry.** In the left plots of Figure 5, we can see that geometry leads to fewer maximal cliques. For  $T = 0$ , the super-polynomial scaling is only barely noticeable. Higher temperatures lead to a larger number of cliques and we get even more cliques for IRGs. Interestingly, the scaling is slower for IRGs when additionally considering the core of vertices with weight more than  $\sqrt{n}$  (see next paragraph).

**Effect of the Core.** When not capping the weight at  $\sqrt{n}$  but also considering vertices of even higher weight (right plots), we can observe the following. The overall picture remains similar, with a slightly increased number of cliques. However, this increase is higher for GIRGs than it is for IRGs. A potential explanation for this is the following. For IRGs, the core forms a clique and adding a large clique to the graph does not change the overall number of maximal cliques by too much. For GIRGs, however, it depends on the constant  $\mu$  controlling the average degree whether this subgraph forms a clique or not. Thus, for the same average degree, the maximum clique is probably somewhat smaller for GIRGs and thus adding the vertices of weight at least  $\sqrt{n}$  leads to more additional cliques than in IRGs.

## 5.2 Cliques in the Dense and Super-Dense Erdős–Rényi Graphs

Here we count the cliques for dense Erdős–Rényi graphs with constant connection probabilities  $p \in \{0.6, 0.7, 0.8, 0.9\}$  and super-dense Erdős–Rényi graphs with connection probability  $p = 1 - c/n$  for  $c \in \{1, 2, 4, 8\}$ . Note that the complement of a super-dense Erdős–Rényi graph has constant expected average degree. The scaling of the number of cliques with respect to the number of vertices is shown in Figure 6, where each point represents 20 samples.

Note that for constant  $p$ , the left plot with logarithmic  $y$ -axis is curved downward, indicating sub-exponential scaling, while the middle plot with logarithmic  $x$ - and  $y$ -axis is bent upwards, indicating super-polynomial scaling. This is in line with our lower bound in Theorem 2.1.

For the super-dense case, the right plot indicates exponential scaling, in line with Theorem 2.3.

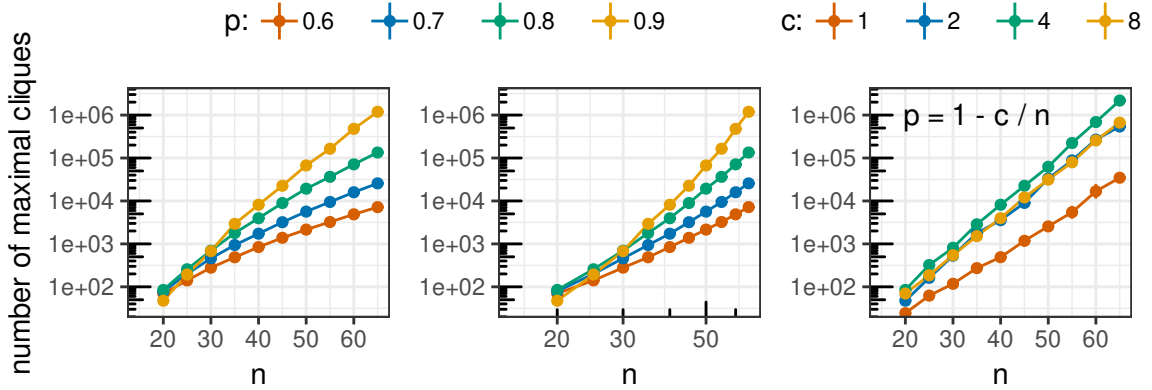


Figure 6: The number of maximal cliques in dense and super-dense  $G(n, p)$ s. For the left and middle plot,  $p$  is constant. For the right plot,  $p = 1 - c/n$  for constant  $c$ . Note that the  $y$ -axes are logarithmic and the  $x$ -axis in the middle plot is logarithmic. Each point is the average of 20 sampled graphs.

## 6 Conclusion and Discussion

In this paper, we have investigated the number of maximal cliques in three random graph models: the Erdős–Rényi random graph, the inhomogeneous random graph and the geometric inhomogeneous random graph. We have shown that sparse Erdős–Rényi random graphs only contain a polynomial amount of maximal cliques, but in the other two sparse models, the number of maximal cliques scales at least super-polynomially in the network size. This is caused by the degree-heterogeneity in these models, as many large maximal cliques are present close to the core of these random graphs. We prove that there only exist a linear amount of small maximal cliques. Interestingly, these small maximal cliques are almost always formed by two low-degree vertices, whereas all other vertices are hubs of high degree.

We have then shown that this dominant super-polynomial behavior of the number of maximal cliques often only kicks for extreme network sizes, and that experimentally, lower-order linear terms instead drive the scaling of the number of maximal cliques until large values of the network size. This explains the dichotomy between the theoretical super-polynomial lower bounds for these models, and the observation that in real-world networks, the amount of maximal cliques is often quite small.

Several of our results only constitute lower bounds for the number of maximal cliques. We believe that relatively close upper bounds can be constructed in a similar fashion, but leave this open for further research.

While Theorem 3.7 only holds for 2-norms, we believe that the theorem can be extended to any  $L^p$ -norm for  $p \neq 1, \infty$ , by looking at the  $L^p$  norm-cycle instead of the regular cycle. For  $p = 1, \infty$  this approach fails, shortest distance paths to non-opposing segments pass through the center of the cycle. Therefore, opposing segments are just as close as many non-opposing ones. Whether Theorem 3.7 also holds with 1 or  $\infty$  norms is therefore a question for further research. We also believe that this approach also extends to the underlying space  $[0, 1]^d$  for general  $d$ , where instead of looking at a cycle inside  $[0, 1]^2$ , one studies a  $d$ -ball inscribed in  $[0, 1]^d$  instead.



## References

- [1] T. Bläsius, C. Freiberger, T. Friedrich, M. Katzmann, F. Montenegro-Retana, and M. Thieffry. Efficient shortest paths in scale-free networks with underlying hyperbolic geometry. *ACM Trans. Algorithms*, 18(2), 2022.
- [2] T. Bläsius, T. Friedrich, and A. Krohmer. Cliques in hyperbolic random graphs. *Algorithmica*, 80(8):2324–2344, 2018.
- [3] T. Bläsius and P. Fischbeck. On the external validity of average-case analyses of graph algorithms. In *European Symposium on Algorithms (ESA)*, pages 21:1–21:14, 2022.
- [4] T. Bläsius, T. Friedrich, M. Katzmann, U. Meyer, M. Penschuck, and C. Weyand. Efficiently generating geometric inhomogeneous and hyperbolic random graphs. *Network Science*, 10(4):361–380, 2022.
- [5] T. Bläsius, T. Friedrich, and A. Krohmer. Hyperbolic random graphs: Separators and treewidth. In *European Symposium on Algorithms (ESA)*, pages 15:1–15:16, 2016.
- [6] M. Boguñá, F. Papadopoulos, and D. Krioukov. Sustaining the internet with hyperbolic mapping. *Nat. Commun.*, 1(6):1–8, 2010.
- [7] K. Bringmann, R. Keusch, and J. Lengler. Geometric inhomogeneous random graphs. *Theoretical Computer Science*, 760:35–54, feb 2019.
- [8] F. Chung and L. Lu. The average distances in random graphs with given expected degrees. *Proc. Natl. Acad. Sci. USA*, 99(25):15879–15882, 2002.
- [9] D. Eppstein, M. Löffler, and D. Strash. Listing all maximal cliques in sparse graphs in near-optimal time. In *International Symposium on Algorithms and Computation (ISAAC)*, pages 403–414, 2010.
- [10] D. Eppstein, M. Löffler, and D. Strash. Listing all maximal cliques in large sparse real-world graphs. *ACM Journal of Experimental Algorithmics (JEA)*, 18, 2013.
- [11] D. Eppstein and D. Strash. Listing all maximal cliques in large sparse real-world graphs. In *Symposium on Experimental and Efficient Algorithms (SEA)*, pages 364–375, 2011.
- [12] P. Erdős and A. Rényi. On random graphs. i. *Publicationes Mathematicae Debrecen*, 6(3–4):290–297, 1959.
- [13] G. García-Pérez, M. Boguñá, A. Allard, and M. Á. Serrano. The hidden hyperbolic geometry of international trade: World trade atlas 1870–2013. *Sci. Rep.*, 6(1), 2016.
- [14] E. N. Gilbert. Random graphs. *The Annals of Mathematical Statistics*, 30(4):1141–1144, Dec. 1959.
- [15] R. van der Hofstad. *Random Graphs and Complex Networks Vol. 1*. Cambridge University Press, 2017.
- [16] R. van der Hofstad, J. S. H. van Leeuwen, and C. Stegehuis. Optimal subgraph structures in scale-free configuration models. *The Annals of Applied Probability*, 31(2), 2021.

- [17] M. Kiwi and D. Mitsche. Spectral gap of random hyperbolic graphs and related parameters. *The Annals of Applied Probability*, 28(2):941–989, apr 2018.
- [18] R. Kleinberg. Geographic routing using hyperbolic space. In *IEEE INFOCOM 2007 - 26th IEEE International Conference on Computer Communications*. IEEE, 2007.
- [19] D. Krioukov, F. Papadopoulos, M. Kitsak, A. Vahdat, and M. Boguná. Hyperbolic geometry of complex networks. *Phys. Rev. E*, 82(3):036106, 2010.
- [20] J. Lengler and L. Todorovic. Existence of small separators depends on geometry for geometric inhomogeneous random graphs.
- [21] R. Michielan and C. Stegehuis. Cliques in geometric inhomogeneous random graphs. 10(1).
- [22] S. Tsukiyama, M. Ide, H. Ariyoshi, and I. Shirakawa. A new algorithm for generating all the maximal independent sets. *SIAM Journal on Computing*, 6(3):505–517, sep 1977.
- [23] H. Yamaji. On the number of maximal cliques in two-dimensional random geometric graphs: Euclidean and hyperbolic. Mar. 2023.

## A Proof of Theorem 3.10

**Lemma A.1.** *Let  $(A_i)_{i \in [k]}$  be a set of areas of size  $A$ , and let  $S$  be a set of vertices, such that  $A|S| > n^\varepsilon$  for some  $\varepsilon > 0$ . Then, for any  $0 < \lambda < 1$  and  $k < \exp(\lambda A|S|)$ , with high probability all areas contain at least  $(1 - \lambda)A|S|$  vertices.*

*Proof.* The Chernoff bound gives for the number of vertices from  $S$  within area  $A$ ,  $N_{S,A}$ :

$$\mathbb{P}(N_{S,A} < (1 - \lambda)A|S|) \leq \exp\left(-\lambda A|S|\right). \quad (53)$$

This implies that when  $A|S| > n^\varepsilon$  for some  $\varepsilon > 0$ , then, with high probability, all areas contain at least  $(1 - \lambda)A|S|$  vertices.  $\square$

We follow the same construction of areas and sets as in the proof of Theorem 3.7. By (20) this creates  $2k = s \cdot n^{\min(\beta/2, \gamma(n)/2)}$  areas of size  $A = n^{-3/2\gamma(n)}$ , with on average  $\mathbb{E}[|S|] = n^{(3-\tau)/2-\beta}$  vertices. Thus, Lemma A.1 shows that as long as  $\beta + 3/2\gamma(n) < (3 - \tau)/2$ , then all areas contain with high probability at least

$$n' = c_1 n^{(3-\tau)/2-\beta-3/2\gamma(n)}$$

vertices for some  $c_1 > 0$ .

From (5), it follows that any set of vertices that contains one in each given area still satisfies the requirement that all vertices in non-opposite boxes connect, as in non-opposite boxes, the connection probability equals 1 by (14). Now to form a co-matching, vertices in opposite boxes should not connect.

With high probability, a positive proportion of vertices in two opposing areas have distance at least  $t + h = a + n^{-\gamma(n)}$ , by the uniform distribution within areas, and the fact that a positive proportion of the two areas have distance  $t + h$ .

By (5), the probability that vertices  $i, j \in S$  at distance at least  $a + n^{-\gamma(n)}$  are connected is bounded by

$$p_{ij} \leq \min\left(\left(\frac{a^2(1 - n^{-\beta})}{(a + n^{-\gamma(n)})^2}\right)^{1/T}, 1\right)$$

$$\begin{aligned}
&= (1 - n^{-\beta})(1 - n^{-\gamma(n)})(1 + o(1)) \\
&= (1 - \max(n^{-\beta}, n^{-\gamma(n)}))(1 + o(1)).
\end{aligned} \tag{54}$$

Similarly as in (23),

$$\mathbb{P}(M = 0) \leq k(1 - \max(n^{-\beta}, n^{-\gamma(n)})(n')^2) \leq k \exp(-\max(n^{-\beta}, n^{-\gamma(n)})(n')^2) \tag{55}$$

Using that  $n' = c_1 n^{(3-\tau)/2-\beta-3/2\gamma(n)}$  therefore yields

$$\mathbb{P}(M = 0) \leq k \exp(-c_1^2 n^{(3-\tau)-\beta-3\gamma(n)} \max(n^{-\beta}, n^{-\gamma(n)})). \tag{56}$$

Thus, choosing  $\beta = \gamma(n) = (3 - \tau)/5 - \varepsilon$  ensures that there is a co-matching of size  $k = s \cdot n^{(3-\tau)/10-\varepsilon}$   $\square$

## B Proof of Lemma 4.3

*Proof.* This integral equals

$$\begin{aligned}
&\int_0^1 \cdots \int_0^1 x_3^{1-\tau} \cdots x_k^{1-\tau} \int_0^1 \int_0^{x_2} x_1^{k-1-\tau} x_2^{k-1-\tau} \exp\left(-\mu^{1-\tau} x_1 x_2^{\tau-2}\right) dx_1 dx_2 \cdots dx_k \\
&+ \int_0^1 \cdots \int_0^1 x_3^{2-\tau} \cdots x_k^{2-\tau} \int_1^\infty \int_0^{x_2} x_1^{k-1-\tau} x_2^{1-\tau} \prod_{i=3}^k \min(x_2 x_i, 1) \exp\left(-\mu^{1-\tau} x_1 x_2^{\tau-2}\right) dx_1 dx_2 \cdots dx_k
\end{aligned} \tag{57}$$

Now the first integral is bounded by

$$\int_0^1 \cdots \int_0^1 x_3^{2-\tau} \cdots x_k^{2-\tau} \int_0^1 \int_0^{x_2} x_1^{k-1-\tau} x_2^{k-1-\tau} dx_2 dx_1 \cdots dx_k < \infty, \tag{58}$$

as  $2 - \tau > -1$ , and  $k - 1 - \tau > -1$  for  $k \geq 3$  as well. We now turn to the second integral. The second integral is finite if

$$\int_0^1 \cdots \int_0^1 x_3^{1-\tau} \cdots x_k^{1-\tau} \int_1^\infty \int_0^{x_2} x_1^{k-1-\tau} x_2^{1-\tau} \prod_{i=3}^k \min(x_2 x_i, 1) \mathbb{1}_{\{x_2^{\tau-2} x_1 < 1\}} dx_1 dx_2 \cdots dx_k < \infty. \tag{59}$$

This results in

$$\begin{aligned}
&\int_0^1 \cdots \int_0^1 x_3^{1-\tau} \cdots x_k^{1-\tau} \int_1^\infty \int_0^{x_2^{2-\tau}} x_1^{k-1-\tau} x_2^{1-\tau} \prod_{i=3}^k \min(x_2 x_i, 1) dx_1 \cdots dx_k \\
&= \int_0^1 \cdots \int_0^1 x_3^{1-\tau} \cdots x_k^{1-\tau} \int_1^\infty x_2^{(2-\tau)(k+1-\tau)-1} \prod_{i=3}^k \min(x_2 x_i, 1) dx_2 \cdots dx_k
\end{aligned} \tag{60}$$

W.l.o.g. we assume that  $x_3 > x_4 > \cdots > x_k$ . Then, the inner integral evaluates to

$$\int_1^\infty x_2^{(2-\tau)(k+1-\tau)-1} \prod_{i=3}^k \min(x_2 x_i, 1) dx_2$$

$$\begin{aligned}
&= \int_1^{1/x_3} x_2^{(2-\tau)(k+1-\tau)+k-3} x_3 \cdots x_k dx_2 + \cdots + \int_{1/x_k}^{\infty} x_2^{(2-\tau)(k+1-\tau)-1} dx_2 \\
&= C_3 x_3^{(\tau-2)(k+1-\tau)+3-k} x_4 \cdots x_k + C_4 x_4^{(\tau-2)(k+1-\tau)+4-k} x_5 \cdots x_k + \cdots + C_k x_k^{(\tau-2)(k+1-\tau)} \quad (61)
\end{aligned}$$

We now show that all these terms evaluate to a finite integral when plugged into (60). Indeed,

$$\begin{aligned}
&\int_0^1 \int_0^{x_3} \cdots \int_0^{x_{k-1}} x_l^{(\tau-2)(k+1-\tau)+l-k} x_{l+1} \cdots x_k x_3^{1-\tau} \cdots x_k^{1-\tau} dx_k dx_{k-1} \cdots dx_3 \\
&= \int_0^1 \int_0^{x_3} \cdots \int_0^{x_{l-1}} x_l^{(\tau-2)(l-\tau)-1} x_3^{1-\tau} \cdots x_{l-1}^{1-\tau} dx_l dx_{l-1} \cdots dx_3 \\
&= \int_0^1 \int_0^{x_3} \cdots \int_0^{x_{l-2}} x_{l-1}^{(\tau-2)(l-1-\tau)-1} x_3^{1-\tau} \cdots x_{l-2}^{1-\tau} dx_l dx_{l-2} \cdots dx_3 < \infty \quad (62)
\end{aligned}$$

as the index  $l - k$  remains at least 3. Therefore, (36) is finite as well.  $\square$