SIAM J. DISCRETE MATH.
Vol. 38, No. 2, pp. 1943–2000

CLIQUES IN HIGH-DIMENSIONAL GEOMETRIC INHOMOGENEOUS RANDOM GRAPHS*

TOBIAS FRIEDRICH†, ANDREAS GÖBEL†, MAXIMILIAN KATZMANN‡, AND LEON SCHILLER§

Abstract. A recent trend in the context of graph theory is to bring theoretical analyses closer to empirical observations by focusing the studies on random graph models that are used to represent practical instances. There, it was observed that geometric inhomogeneous random graphs (GIRGs) yield good representations of complex real-world networks by expressing edge probabilities as a function that depends on (heterogeneous) vertex weights and distances in some underlying geometric space that the vertices are distributed in. While most of the parameters of the model are understood well, it was unclear how the dimensionality of the ground space affects the structure of the graphs. In this paper, we complement existing research into the dimension of geometric random graph models and the ongoing study of determining the dimensionality of real-world networks by studying how the structure of GIRGs changes as the number of dimensions increases. We prove that, in the limit, GIRGs approach nongeometric inhomogeneous random graphs and present insights on how quickly the decay of the geometry impacts important graph structures. In particular, we study the expected number of cliques of a given size as well as the clique number and characterize phase transitions at which their behavior changes fundamentally. Finally, our insights help in better understanding previous results about the impact of the dimensionality on geometric random graphs.

Key words. random graphs, geometry, dimensionality, cliques, clique number, scale-free networks

MSC codes. 68Q25, 68R10, 68U05

DOI. 10.1137/23M157394X

1. Introduction.

Networks are a powerful tool to model all kinds of processes that we interact with in our day-to-day lives. From connections between people in social networks, to the exchange of information on the internet, and on to how our brains are wired, networks are everywhere. Consequently, they have been in the focus of computer science for decades. There, one of the most fundamental techniques used to model and study networks is *random graph models.* Such a model defines a probability distribution over graphs, which is typically done by specifying a random experiment on how to construct the graph. By analyzing the rules of the experiment, we can then derive structural and algorithmic properties of the resulting graphs. If the results match what we observe on real-world networks, i.e., if the model represents the graphs we encounter in practice well, then we can use it to make further predictions that help us understand real graphs and utilize them more efficiently.

The quest to find a good model started several decades ago, with the famous Erdős–Rényi (ER) random graphs [19, 24]. There, all edges in the graph exist independently with the same probability. Due to its simplicity, this model has been studied extensively. However, because the degree distribution of the resulting graphs is

*Received by the editors May 19, 2023; accepted for publication (in revised form) March 24, 2024; published electronically June 25, 2024.
https://doi.org/10.1137/23M157394X
Funding: The first author was funded by project 390859508 of the German Research Foundation. The second author was funded by the project PAGES (project 467516565) of the German Research Foundation (DFG).
†Hasso Plattner Institute, University of Potsdam, Potsdam, Brandenburg, Germany (tobias.friedrich@hpi.de, andreas.goebel@hpi.de).
‡Karlsruhe Institute of Technology, Karlsruhe, 76131, Germany (maximilian.katzmann@kit.edu).
§ETH Zürich, Switzerland (leon.schiller@inf.ethz.ch).

rather homogeneous and they lack clustering (due to the independence of the edges), the model is not considered to yield good representations of real graphs. In fact, many networks we encounter in practice feature a degree distribution that resembles a power-law [3, 39, 40], and the clustering coefficient (the probability for two neighbors of a vertex to be adjacent) is rather high [36, 41]. To overcome these drawbacks, the initial random graph model has been adjusted in several ways.

In *inhomogeneous random graphs* (IRGs)*,* often referred to as *Chung–Lu random graphs,* each vertex is assigned a weight, and the probability for two vertices to be connected by an edge is proportional to the product of the weights [1, 11, 12]. As a result, the expected degrees of the vertices in the resulting graphs match their weight. While assigning weights that follow a power-law distribution yields graphs that are closer to the complex real-world networks, the edges are still drawn independently, leading to vanishing clustering coefficients.

A very natural approach to facilitate clustering in a graph model is to introduce an underlying geometry. This was done first in *random geometric graphs* (RGGs)*,* where vertices are distributed uniformly at random in the Euclidean unit square and any two are connected by an edge if their distance lies below a certain threshold; i.e., the neighborhood of a vertex lives in a disk centered at that vertex [37]. Intuitively, two vertices that connect to a common neighbor cannot be too far away from each other, increasing the probability that they are connected by an edge themselves. In fact, random geometric graphs feature a nonvanishing clustering coefficient [13]. However, since all neighborhood disks have the same size, they all have roughly the same expected degree, again leading to a homogeneous degree distribution.

To get a random graph model that features a heterogeneous degree distribution *and* clustering, the two mentioned adjustments were recently combined to obtain *geometric inhomogeneous random graphs* (GIRGs) [29]. There, vertices are assigned a weight *and* a position in some underlying geometric space and the probability for two vertices to connected increases with the product of the weights but decreases with increasing geometric distance between them. As a result, the generated graphs have a nonvanishing clustering coefficient, and, with the appropriate choice of the weight sequence, they feature a power-law degree distribution. Additionally, recent empirical observations indicate that GIRGs represent real-world networks well with respect to certain structural and algorithmic properties [5].

We note that GIRGs are not the first model that exhibits a heterogeneous degree distribution and clustering. In fact, *hyperbolic random graphs* (HRGs) [31] feature these properties as well and have been studied extensively (see, e.g., [7, 20, 21, 23, 26]). However, in the pursuit of finding good models to represent real-world networks, GIRGs introduce a parameter that sets them apart from prior models: the choice of the underlying geometric space and, more importantly, the dimensionality of that space.

Unfortunately, this additional parameter that sets GIRGs apart from previous models has not gained much attention at all. In fact, it comes as a surprise that, while the underlying dimensionality of real-world networks is actively researched [2, 8, 15, 25, 32] and there is a large body of research examining the impact of the dimensionality on different homogeneous graph models [13, 17, 18] with some advancements being made on hyperbolic random graphs [42], the effects of the dimension on the structure of GIRGs have only been studied sparsely. For example, while it is known that GIRGs exhibit a clustering coefficient of $\Theta(1)$ for any fixed dimension [29], it is not known how the hidden constants scale with the dimension.

In this paper, we initiate the study of the impact of the dimensionality on GIRGs. In particular, we investigate the influence of the underlying geometry as the dimensionality increases, proving that GIRGs converge to their nongeometric counterpart (IRGs) in the limit. With our results we are able to explain seemingly disagreeing insights from prior research on the impact of dimensionality on geometric graph models. Moreover, by studying the clique structure of GIRGs and its dependence on the dimension d, we are able to quantify how quickly the underlying geometry vanishes. In the following, we discuss our results in greater detail. We note that, while we give general proof sketches for our results, the complete proofs are deferred to the full version [22].

1.1. (Geometric) inhomogeneous random graphs. Before stating our results in greater detail, let us recall the definitions of the two graph models we mainly work with throughout the paper.

Inhomogeneous random graphs (IRGs). The model of inhomogeneous random graphs was introduced by Chung and Lu [1, 11, 12] and is a natural generalization of the Erdős–Rényi model. Starting with a vertex set V of n vertices, each $v \in V$ is assigned a weight w_v. Each edge $\{u,v\} \in \binom{V}{2}$ is then independently present with probability

$$\Pr[u \sim v] = \min\left\{1, \frac{\lambda w_u w_v}{n}\right\}$$

for some constant $\lambda > 0$ controlling the average degree of the resulting graph. Note that assigning the same weight to all vertices yields the same connection probability as in Erdős–Rényi random graphs. For the sake of simplicity, we define $\kappa_{uv} = \min\{\lambda w_u w_v, n\}$ such that $\Pr[u \sim v] = \kappa_{uv}/n$. Additionally, for a set of vertices $U_k = \{v_1, \dots, v_k\}$ with weights $w_1, \dots, w_k$, we introduce the shorthand notation $\kappa_{ij} = \kappa_{v_i v_j}$ and write $\{\kappa\}^{(k)} = \{\kappa_{ij} \mid 1 \le i < j \le k\}$.

Throughout the paper, we mainly focus on inhomogeneous random graphs that feature a power-law degree distribution in expectation, which is obtained by sampling the weights accordingly. More precisely, for each $v \in V$, we sample a weight w_v from the Pareto distribution $\mathcal{P}$ with parameters $1-\beta, w_0$ and distribution function

$$\Pr[w_v \le x] = 1 - \left(\frac{x}{w_0}\right)^{1-\beta}.$$

Then the density of w_v is $\rho_{w_v}(x) = (\beta-1)x^{-\beta}/w_0^{1-\beta}$. Here, $w_0 > 0$ is a constant that represents a lower bound on the weights in the graph and β denotes the power-law exponent of the resulting degree distribution. Throughout the paper, we assume $\beta > 2$ such that a single weight has finite expectation (and thus the average degree in the graph is constant) but possibly infinite variance. We denote a graph obtained by utilizing the above weight distribution and connection probabilities with $\mathrm{IRG}(n, \beta, w_0)$. For a fixed weight sequence $\{w\}_1^n$, we denote the corresponding graph by $\mathrm{IRG}(\{w\}_1^n)$.

Geometric inhomogeneous random graphs (GIRGs). Geometric inhomogeneous random graphs are an extension of IRGs, where in addition to the weight, each vertex v is also equipped with a position $\mathbf{x}_v$ in some geometric space and the probability for edges to form depends on their weights and the distance in the underlying space [29]. While, in its raw form, the GIRG framework is rather general, we align our paper with existing analysis on GIRGs [6, 30, 35] and consider the d-dimensional torus $\mathbb{T}^d$ equipped with L_∞-norm as the geometric ground space, whereby $\mathbb{T}^d$ can be described as the d-dimensional hypercube $[0,1]^d$ with opposite boundaries identified with each other. More precisely, in what we call the *standard* GIRG model, the positions $\mathbf{x}$ of the vertices are drawn independently and uniformly at random from $\mathbb{T}^d$, according to the standard Lebesgue measure. We denote the ith component of $\mathbf{x}_v$ by $\mathbf{x}_{vi}$. Additionally, the geometric distance between two points $\mathbf{x}_u$ and $\mathbf{x}_v$, is given by

$$d(\mathbf{x}_u, \mathbf{x}_v) = \|\mathbf{x}_u - \mathbf{x}_v\|_\infty = \max_{1 \le i \le d}\{|\mathbf{x}_{ui} - \mathbf{x}_{vi}|_C\},$$

where $|\cdot|_C$ denotes the distance on the circle, i.e.,

$$|\mathbf{x}_{ui} - \mathbf{x}_{vi}|_C = \min\{|\mathbf{x}_{ui} - \mathbf{x}_{vi}|, 1 - |\mathbf{x}_{ui} - \mathbf{x}_{vi}|\}.$$

In a standard GIRG, two vertices $u \neq v$ are adjacent if and only if their distance $d(\mathbf{x}_u, \mathbf{x}_v)$ in the torus is less than or equal to a *connection threshold* t_{uv}, which is given by

$$t_{uv} = \frac{1}{2}\left(\frac{\lambda w_u w_v}{n}\right)^{1/d} = \left(\frac{w_u w_v}{\tau n}\right)^{1/d},$$

where $\tau = 2^d/\lambda$. Using L_∞ is motivated by the fact that it is the most widely used metric in the literature because it is arguably the most natural metric on the torus. In particular, it has the "nice" property that the ball of radius r is a cube and "fits" entirely into $\mathbb{T}^d$ for all $0 \le r \le 1$.

Note that, as a consequence of the above choice, the marginal connection probability $\Pr[u \sim v]$ is the same as in the IRG model, i.e., $\Pr[u \sim v] = \kappa_{uv}/n$. However, while the probability that any given edge is present is the same as in the IRG model, the edges in the GIRG model are *not* drawn independently. We denote a graph obtained by the procedure described above with $\mathrm{GIRG}(n, \beta, w_0, d)$. As for IRGs, we write $\mathrm{GIRG}(\{w\}_1^n, d)$ when considering standard GIRGs with a fixed weight sequence $\{w\}_1^n$.

As mentioned above, the standard GIRG model is a commonly used instance of the more general GIRG framework [29]. There, different geometries and distance functions may be used. For example, instead of L_∞-norm, any L_p-norm for $1 \le p < \infty$ may be used. Then, the distance between two vertices u, v is measured as

$$\|\mathbf{x}_u - \mathbf{x}_v\|_p := \begin{cases} \left(\sum_{i=1}^d |\mathbf{x}_{ui} - \mathbf{x}_{vi}|^p\right)^{1/p} & \text{if } p < \infty, \\ \max_{1 \le i \le d}\{|\mathbf{x}_{ui} - \mathbf{x}_{vi}|\} & \text{otherwise.} \end{cases}$$

With this choice, the volume (Lebesgue measure) of the ball $B_p(r)$ of radius r under L_p-norm is equal to the probability that a vertex u falls within distance at most r of v (if $r = o(1)$). We denote this volume by $\nu(r)$. We call the corresponding graphs *standard GIRGs with any L_p-norm* and note that some of our results extend to this more general model. Finally, whenever our insights consider an even broader variant of the model (e.g., variable ground spaces, distances functions, weight distributions), we say that they hold for *any GIRG* and mention the constraints explicitly.

1.2. Asymptotic equivalence. Our first main observations is that large values of d diminish the influence of the underlying geometry until, at some point, our model becomes strongly equivalent to its nongeometric counterpart, where edges are sampled independently of each other. We prove that the *total variation distance* between the distribution over all graphs of the two models tends to zero as n is kept fixed and

$d \to \infty$. We define the total variation distance of two probability measures P and Q on the measurable space $(\Omega, \mathcal{F})$ as

$$\|P, Q\|_{\mathrm{TV}} = \sup_{A \in \mathcal{F}} |P(A) - P(B)| = \frac{1}{2} \sum_{\omega \in \Omega} |P(\omega) - Q(\omega)|,$$

where the second equality holds if Ω is countable. In our case, Ω is the set $\mathcal{G}(n)$ of all possible graphs on n vertices, and P, Q are distributions over these graphs. If G_1, G_2 are two random variables mapping to Ω, we refer to $\|G_1, G_2\|_{\mathrm{TV}}$ as the total variation distance of the induced probability measures by G_1 and G_2, respectively. Informally, this measures the maximum difference in the probability that any graph G is sampled by G_1 and G_2.

THEOREM 1.1. *Let $\mathcal{G}(n)$ be the set of all graphs with n vertices, let $\{w\}_1^n$ be a weight sequence, and consider $G_{IRG} = \mathrm{IRG}(\{w\}_1^n) \in \mathcal{G}(n)$ and a standard GIRG $G_{GIRG} = \mathrm{GIRG}(\{w\}_1^n, d) \in \mathcal{G}(n)$ with any L_p-norm. Then,*

$$\lim_{d \to \infty} \|G_{GIRG}, G_{IRG}\|_{TV} = 0.$$

We note that this theorem holds for arbitrary weight sequences that do not necessarily follow a power-law and for arbitrary L_p-norms used to define distances in the ground space. For $p \in [1, \infty)$, the proof is based on the application of a multivariate central limit theorem [38], in a similar way as used to prove a related statement for *spherical random geometric graphs* (SRGGs), i.e., random geometric graphs with a hypersphere as ground space [17]. Our proof generalizes this argument to arbitrary L_p-norms and arbitrary weight sequences. For the case of L_∞-norm, we present a proof based on the inclusion-exclusion principle and the bounds we develop in the full version [22, section 4].

Remarkably, while similar behavior was previously established for SRGGs, there exist works indicating that RGGs on the hypercube do not converge to their nongeometric counterpart [13, 18] as $d \to \infty$. We show that this apparent disagreement is due to the fact that the torus is a homogeneous space while the hypercube is not. In fact, our proof shows that GIRGs on the hypercube *do* converge to a nongeometric model in which edges are, however, not sampled independently. This lack of independence is because, on the hypercube, there is a positive correlation between the distances from two vertices to a given vertex, leading to a higher tendency to form clusters, as was observed experimentally [18]. Due to the homogeneous nature of the torus, the same is not true for GIRGs, and the model converges to the plain IRG model with independent edges.

1.3. Clique structure.

To quantify for which dimensions d the graphs in the GIRG model start to behave similarly to IRGs, we investigate the number and size of cliques. Previous results on SRGGs indicate that the dimension of the underlying space heavily influences the clique structure of the model [4, 17]. However, it was not known how the size and the number of cliques depends on d if we use the torus as our ground space, and how the clique structure in high dimensions behaves for inhomogeneous weights.

We give explicit bounds on the expected number of cliques of a given size k, which we afterwards turn into bounds on the *clique number* $\omega(G)$, i.e., the size of the largest clique in the graph G. While the expected number of cliques in the GIRG model was previously studied by Michielan and Stegehuis [35] when the power-law exponent of the degree distribution satisfies $\beta \in (2, 3)$, to the best of our knowledge, the clique number of GIRGs remains unstudied even in the case of constant (but arbitrary) dimensionality. We close this gap, reproduce the existing results, and extend them to the case $\beta \geq 3$ and the case where d can grow as a function of the number of vertices n in the graph. Furthermore, our bounds for the case $\beta \in (2,3)$ are more explicit and complement the work of Michielan and Stegehuis, who expressed the (rescaled) asymptotic number of cliques as converging to a nonanalytically solvable integral. Furthermore, we show that the clique structure in our model eventually behaves asymptotically like that of an IRG if the dimension is sufficiently large. In summary, our main contributions are outlined in Tables 1, 2, and 3.

We observe that the structure of the cliques undergoes three phase transitions in the size of the cliques k, the dimension d, and the power-law exponent β.

Transition in k. When $\beta \in (2,3)$ and $d \in o(\log(n))$, the first transition is at $k = \frac{2}{3-\beta}$, as was previously observed for hyperbolic random graphs [7] and for GIRGs of constant dimensionality [35]. The latter work explains this behavior by showing that for $k < \frac{2}{3-\beta}$, the number of cliques is strongly dominated by "geometric" cliques forming among vertices whose distance is of order $n^{-1/d}$ regardless of their weight. For $k > \frac{2}{3-\beta}$, on the other hand, the number of cliques is dominated by "nongeometric" cliques forming among vertices with weights on the order of $\sqrt{n}$. This behavior is in contrast to the behavior of cliques in the IRG model, where this phase transition does not exist and where the expected number of k cliques is $\Theta(n^{\frac{k}{2}(3-\beta)})$ for all $k \geq 3$ (if $\beta \in (2,3)$) [14].

Table 1

*Asymptotic behavior of the expected number of k-cliques. Results marked with * were previously known for constant k. For all depicted regimes, K_k concentrates well around its expectation, i.e., $K_k/\mathbb{E}[K_k]$ converges in probability to 1 if k is sufficiently small: for cells marked in light-gray, this holds for $k = o(n^{(3-\beta)/4})$; for cells marked in dark-gray, it holds for $k = o(\log(n)/(\log\log(n) + d))$; for white cells, it holds for all k (cf.* Theorem 1.4*).*

$\mathbb{E}[K_k]$ for $k \geq 4$			
	$d = \Theta(1)$	$d = o(\log(n))$	$d = \omega(\log(n))$
$2 < \beta < 3, k > \frac{2}{3-\beta}$	$n^{\frac{k}{2}(3-\beta)}\Theta(k)^{-k}$*	$n^{\frac{k}{2}(3-\beta)}\Theta(k)^{-k}$	$n^{\frac{k}{2}(3-\beta)}\Theta(k)^{-k}$
$2 < \beta < 3, k < \frac{2}{3-\beta}$	$n\Theta(k)^{-k}$*	$ne^{-\Theta(1)dk}\Theta(k)^{-k}$	$n^{\frac{k}{2}(3-\beta)}\Theta(k)^{-k}$
$\beta > 3$	$n\Theta(k)^{-k}$	$ne^{-\Theta(1)dk}\Theta(k)^{-k}$	$o(1)$
	equivalent to HRGs [7]		equivalent to IRGs [28]

Table 2

Asymptotic behavior of the expected number of triangles. The case $\beta = \infty$ refers to the case of constant weights. While in the case $\beta < 3$, the number of triangles already behaves like that of the IRG model if $d = \omega(\log(n))$, in the case $\beta > 3$, the number of triangles remains superconstant as long as $d = o(\log^{3/2}(n))$.

Expected number of triangles $\mathbb{E}[K_3]$			
	$d = o(\log(n))$	$d = \omega(\log(n))$	$d = \omega(\log^2(n))$
$2 < \beta < \frac{7}{3}$	$n^{\frac{3}{2}(3-\beta)}\Theta(1)$	$n^{\frac{3}{2}(3-\beta)}\Theta(1)$	$n^{\frac{3}{2}(3-\beta)}\Theta(1)$
$\frac{7}{3} < \beta < 3$	$ne^{-\Theta(1)d}\Theta(1)$	$n^{\frac{3}{2}(3-\beta)}\Theta(1)$	$n^{\frac{3}{2}(3-\beta)}\Theta(1)$
$\beta > 3$	$ne^{-\Theta(1)d}\Theta(1)$	$\Omega\left(\exp\left(\frac{\ln^3(n)}{d^2}\right)\right)$	$\Theta(1)$
$\beta = \infty$	$ne^{-\Theta(1)d}\Theta(1)$	$\Theta\left(\exp\left(\frac{\ln^3(n)}{d^2}\right)\right)$	$\Theta(1)$

TABLE 3

Asymptotic behavior of the clique number of G for different values of d in the GIRG model. The behavior of the first column is the same as in hyperbolic random graphs established in [7]*, and the behavior in the third column is the same as that of IRG graphs established in* [28]*. All results hold a.a.s. and under L_∞-norm.*

	$\omega(G)$		
	$d=\mathcal{O}(\log\log(n))$	$d=o(\log(n))$	$d=\omega(\log(n))$
$\beta<3$	$\Theta\left(n^{(3-\beta)/2}\right)$	$\Theta\left(n^{(3-\beta)/2}\right)$	$\Theta\left(n^{(3-\beta)/2}\right)$
$\beta=3$	$\Theta\left(\frac{\log(n)}{\log\log(n)}\right)$	$\Omega\left(\frac{\log(n)}{d}\right)$	$\mathcal{O}(1)$
$\beta>3$	$\Theta\left(\frac{\log(n)}{\log\log(n)}\right)$	$\Theta\left(\frac{\log(n)}{d}\right)$	≤ 3
	equivalent to HRGs [7]		equivalent to IRGs [28]

Transition in d. Still assuming $\beta\in(2,3)$, the second phase transition occurs as d becomes superlogarithmic. More precisely, we show that in the high-dimensional regime, where $d=\omega(\log(n))$, the phase transition in k vanishes, as the expected number of cliques of size $k\geq 4$ behaves asymptotically like its counterpart in the IRG model. Nevertheless, we can still differentiate the two models as long as $d=o(\log^{3/2}(n))$ by counting triangles among low degree vertices as can be seen in Table 2.

The reason for this behavior is that the number of cliques in the case $d=\omega(\log(n))$ is already dominated by cliques forming among vertices of weight close to $\sqrt{n}$. For those, the probability that a clique is formed already behaves like in an IRG, although, for vertices of small weight, said probability it is still significantly larger as long as $d=o(\log(n)^{3/2})$.

Regarding the clique number, in the case $\beta>3$, we observe a similar phase transition in d. For constant d, the clique number of a GIRG is $\Theta(\log(n)/\log\log(n))=\omega(1)$. We find that this asymptotic behavior remains unchanged if $d=\mathcal{O}(\log\log(n))$. However, if $d=\omega(\log\log(n))$ but $d=o(\log(n))$, the clique number scales as $\Theta(\log(n)/d)$, which is still superconstant. Additionally, if $d=\omega(\log(n))$, we see that, again, GIRGs show the same behavior as IRGs. That is, there are asymptotically no cliques of size larger than 3.

Transition in β. The third phase transition in the high-dimensional case occurs at $\beta=3$, which is in line with the fact that networks with a power-law exponent $\beta\in(2,3)$ contain with high probability (w.h.p., meaning with probability $1-O(1/n)$) a densely connected "heavy core" of $\Theta(n^{\frac{1}{2}(3-\beta)})$ vertices with weight $\sqrt{n}$ or above, which vanishes if β is larger than 3. This heavy core strongly dominates the number of cliques of sufficient size and explains why the clique number is $\Theta(n^{\frac{1}{2}(3-\beta)})$ regardless of d if $\beta\in(2,3)$. As β grows beyond 3, the core disappears and leaves only very small cliques. Accordingly for $\beta>3$ IRGs contain asymptotically almost surely (a.a.s., meaning with probability $1-o(1)$) no cliques of size greater than 3. In contrast to that, for GIRGs of dimension $d=o(\log(n))$ (and HRGs), the clique number remains superconstant and so does the number of k-cliques for any constant $k\geq 3$. If $d=\omega(\log(n))$, there are no cliques of size greater than 3 like in an IRG. However, as noted before, GIRGs feature many more triangles than IRGs as long as $d=o(\log^{3/2}(n))$.

Characterizing the typical clique. Our analysis also yields insights into where cliques typically form within the graph. Previously known in this regard was that, for constant $d,\beta\in(2,3)$ and constant k, cliques of size $k>\frac{2}{3-\beta}$ form dominantly among vertices of weight on the order of $\sqrt{n}$, whereas for $k<\frac{2}{3-\beta}$, they form among vertices of pairwise distance on the order of $n^{-1/d}$, as shown in [35]. We extend these results to the case where k and d are allowed to be superconstant, where $\beta>3$, and we provide a characterization in terms of the weights of the vertices associated to a clique that extend the known results even for the previously studied parameter regimes.

To be more precise, we denote by $w_{\min}, w_{\max}$ the minimal and maximal vertex weights associated to a randomly chosen clique and study for which weights w, $w_{\min}$ and $w_{\max}$ are arbitrarily likely to be on the order of w. To this end, we define the following.

Definition 1.2. *For any $w \in \mathbb{R}$ and $\varepsilon > 0$ define*

$$M_\varepsilon^{(+)}(w) := \{x \in \mathbb{R} \mid x \le w/\varepsilon\} \text{ and } M_\varepsilon^{(-)}(w) := \{x \in \mathbb{R} \mid x \ge \varepsilon w\}.$$

Furthermore, define

$$M_\varepsilon(w) := M_\varepsilon^{(+)}(w) \cap M_\varepsilon^{(-)}(w).$$

We proceed by studying for which w we can make the conditional probabilities that $w_{\min}$ or $w_{\max}$ is in $M_\varepsilon^{(+)}(w), M_\varepsilon^{(-)}(w)$, or $M_\varepsilon(w)$ arbitrarily large by adjusting ε.

Our results are summarized in Table 4. The central result of these two tables is that assuming $d = o(\log(n))$, if $k < \frac{2}{3-\beta}$ or if k is arbitrary and $\beta > 3$, then cliques dominantly form among vertices of very small weight, more precisely among vertices of weight at most $k^{\frac{1}{2-\beta}} n^{o(1)}$, which (if you take our results on the clique number into account) is $n^{o(1)}$ in total for cliques of all sizes that appear in the model with nonvanishing probability. On the other hand, if $d = \omega(\log(n))$ and $\beta \in (2,3)$, then cliques of all sizes dominantly form among very high-weight vertices, more precisely among vertices of weight at least on the order of $\sqrt{n}$. Again, this is the same behavior as in IRGs. We formalize this result in the following theorem.

Theorem 1.3. *Let U_k be a set of k randomly chosen vertices. If $\beta \in (2,3)$, $k < \frac{2}{3-\beta}$, and $d = o(\log(n))$, then there is a function $f(n) = e^{\Theta(1)d} = n^{o(1)}$ such that for all $p \in (0,1)$, there is an $\varepsilon > 0$ such that*

$$\Pr\left[w_{\max} \le f(n)/\varepsilon \mid U_k \text{ is clique}\right] \ge p.$$

If $\beta \in (2,3)$ and $d = \omega(\log(n))$, then for all (potentially superconstant) $k \ge 3$ and all $p \in (0,1)$, there is an $\varepsilon > 0$ such that

$$\Pr\left[w_{\min} \ge \varepsilon\sqrt{n} \mid U_k \text{ is clique}\right] \ge p.$$

Table 4

Dominant regimes for the minimum/maximum vertex weight associated to a clique. An entry of $M_\varepsilon(w)$ (as defined in Definition 1.2) means that for every $p \in (0,1)$, there is an $\varepsilon > 0$ such that $\Pr[w_{\min} \in M_\varepsilon(w) \mid U_k$ is clique$] \ge p$ (resp., $\Pr[w_{\max} \in M_\varepsilon(w) \mid U_k$ is clique$] \ge p$), where U_k is a set of k vertices chosen u.a.r.

Dominant Regimes for $w_{\min}$			
	$d = \Theta(1)$	$d = o(\log(n))$	$d = \omega(\log(n))$
$2 < \beta < 3, k > \frac{2}{3-\beta}$	$M_\varepsilon(\sqrt{n})$	$M_\varepsilon(\sqrt{n})$	$M_\varepsilon(\sqrt{n})$
$2 < \beta < 3, k < \frac{2}{3-\beta}$	$M_\varepsilon(1)$	$M_\varepsilon^{(+)}(n^{o(1)})$	$M_\varepsilon(\sqrt{n})$
$\beta > 3$	$M_\varepsilon(1)$	$M_\varepsilon^{(+)}(n^{o(1)})$	
Dominant Regimes for $w_{\max}$			
	$d = \Theta(1)$	$d = o(\log(n))$	$d = \omega(\log(n))$
$2 < \beta < 3, k > \frac{2}{3-\beta}$	$M_\varepsilon(k^{\frac{1}{\beta-1}}\sqrt{n})$	$M_\varepsilon(k^{\frac{1}{\beta-1}}\sqrt{n})$	$M_\varepsilon(k^{\frac{1}{\beta-1}}\sqrt{n})$
$2 < \beta < 3, k < \frac{2}{3-\beta}$	$M_\varepsilon(k^{\frac{1}{\beta-2}})$	$M_\varepsilon^{(+)}(k^{\frac{1}{\beta-2}} n^{o(1)})$	$M_\varepsilon(k^{\frac{1}{\beta-1}}\sqrt{n})$
$\beta > 3$	$M_\varepsilon(k^{\frac{1}{\beta-2}})$	$M_\varepsilon^{(+)}(k^{\frac{1}{\beta-2}} n^{o(1)})$	–

Moreover, it is worth noting that for d,k constant, cliques of size $k < \frac{2}{3-\beta}$, if $\beta \in (2,3)$, and cliques of size $k \geq 3$, if $\beta > 3$, dominantly form among vertices of only constant weight. Additionally, we remark that the dependence of $w_{\max}$ on k as given in Table 4 is the same as one would expect in a star centered at the vertex of minimal weight $v_{\min}$. That is, if the weight of $v_{\min}$ is much smaller than $\sqrt{n}$, then each neighbor u of $v_{\min}$ has a weight that is essentially a sample of a Pareto distribution with exponent $\beta - 1$ instead of β, since conditioning on $u \sim v_{\min}$ imposes a bias towards a higher weight of u. Since there are $\Theta(k)$ neighbors, the maximum weight among these is essentially the maximum of $\Theta(k)$ independent samples from this distribution, which is typically of order $k^{\frac{1}{\beta-2}}$. If $w_{\min}$ is already of order $\sqrt{n}$, the situation is similar; however, conditioning on $u \sim v_{\min}$ no longer imposes a bias towards a higher weight of u as vertices with weight in this range are all connected with probability $\Omega(1)$. Thus, the weight of the neighbors of $v_{\min}$ continues to follow a Pareto distribution with exponent β, and the maximal weight among them is of order $k^{\frac{1}{\beta-1}}$. Our results show that this known behavior for stars remains essentially true for cliques; that is, conditioning on having a clique does not induce a bias towards much larger weights than conditioning on having a star.

Concentration bounds. The above analysis not only gives insights into where cliques dominantly form but also allows us to establish concentration bounds on the number of cliques in a similar way as done in [35]. More precisely, it allows us to establish that K_k rescaled by its expectation converges in probability to 1 for almost all the regimes we consider and almost all relevant sizes of k. We write $K_k/\mathbb{E}[K_k] \to_p 1$ to denote convergence in probability and formalize in our statement in the following theorem.

THEOREM 1.4. *We have $K_k/\mathbb{E}[K_k] \to_p 1$; that is for all $\delta > 0$,*

$$\Pr\left[\left|\frac{K_k}{\mathbb{E}[K_k]} - 1\right| \geq \delta\right] = o(1)$$

if one of the following conditions holds.
(i) $d = o(\log(n))$, $\beta \in (2,3)$, $k \neq \frac{2}{3-\beta}$, *and* $k = o(n^{(3-\beta)/4})$.
(ii) $d = \omega(\log(n))$, $\beta \in (2,3)$, $k = o(n^{(3-\beta)/4})$.
(iii) $d = o(\log(n))$, $\beta > 3$, *and* $k = o(\log(n)/(\log\log(n) + d))$.

We remark that even for values of k larger than the ones stated above, our results imply (slightly weaker) concentration bounds. General bounds on the variance of cliques are given in subsection 3.3.

Proof techniques. The proofs of our results (i.e., the ones in the above tables) are mainly based on bounds on the probability that a set of k randomly chosen vertices forms a clique. To obtain concentration bounds on the number of cliques as needed for deriving bounds on the clique number, we use the second moment method and Chernoff bounds.

For the case of $d = \omega(\log(n))$, many of our results are derived from the following general insight. We show that for all $\beta > 2$, the probability that a set of vertices forms a clique already behaves similarly to that in the IRG model if the weights of the involved nodes are sufficiently large. For $d = \omega(\log(n)^2)$, this holds in the entire graph, that is, regardless of the weights of the involved vertices. In fact our statement holds even more generally. That is, the described behavior applies not only to the probability that a clique is formed but also to the probability that any set of edges (or a superset thereof) is created.

THEOREM 1.5. *Let G be a standard GIRG and let $k \geq 3$ be a constant. Furthermore, let $U_k = \{v_1, \ldots, v_k\}$ be a set of vertices chosen uniformly at random and let $\{\kappa\}^{(k)} = \{\kappa_{ij} \mid 1 \leq i, j \leq k\}$ describe the pairwise product of weights of the vertices in U_k. Let $E(U_k)$ denote the (random) set of edges formed among the vertices in U_k. Then, for any set of edges $\mathcal{A} \subseteq \binom{U_k}{2}$,*

$$\Pr\left[E(U_k) \supseteq \mathcal{A} \mid \{\kappa\}^{(k)}\right] = \begin{cases} (1 \pm o(1)) \prod_{\{i,j\} \in \mathcal{A}} \frac{\kappa_{ij}}{n} & \text{if } d = \omega(\log^2(n)), \\ (1 \pm o(1)) \prod_{\{i,j\} \in \mathcal{A}} \left(\frac{\kappa_{ij}}{n}\right)^{1 \mp \mathcal{O}\left(\frac{\log(n)}{d}\right)} & \text{if } d = \omega(\log(n)). \end{cases}$$

For the proof we derive elementary bounds on the probability of the described events and use series expansions to investigate their asymptotic behavior. Remarkably, in contrast to our bounds for the case $d = o(\log(n))$, the high-dimensional case requires us to pay closer attention to the topology of the torus.

We leverage the above theorem to prove that GIRGs eventually become equivalent to IRGs with respect to the total variation distance. Theorem 1.5 already implies that the expected number of cliques in a GIRG is asymptotically the same as in an IRG for all $k \geq 3$ and all $\beta > 2$ if $d = \omega(\log^2(n))$. However, we are able to show that the expected number of cliques for $\beta \in (2, 3)$ actually already behaves like that of an IRG if $d = \omega(\log(n))$. The reason for this is that the clique probability among high-weight vertices starts to behave like that of an IRG earlier than is the case for low-weight vertices, and cliques forming among these high-weight vertices already dominate the number of cliques. Moreover, the clique number behaves like that of an IRG if $d = \omega(\log(n))$ for all $\beta > 2$. However, the number of triangles among vertices of constant weight asymptotically exceeds that of an IRG as long as $d = o(\log^{3/2}(n))$, which we prove by deriving even sharper bounds on the expected number of triangles. Accordingly, convergence with respect to the total variation distance cannot occur before this point (this holds for all $\beta > 2$).

In contrast to this, for the low-dimensional case (where $d = o(\log(n))$), the underlying geometry still induces strongly notable effects regarding the number of sufficiently small cliques for all $\beta > 2$. However, even here, the expected number of such cliques decays exponentially in dk. The main difficulty in showing this is that we have to handle the case of inhomogeneous weights, which significantly influence the probability that a set of k vertices chosen uniformly at random forms a clique. To this end, we prove the following theorem, which bounds the probability that a clique among k vertices is formed if the ratio of the maximal and minimal weights is at most c^d. Note that the vertices forming a star is necessary for a clique to form. For this reason we consider the event $\mathbf{E}^c_{\text{star}}$ of the vertices forming a star centered at the lowest weight vertex. The theorem generalizes a result of Decreusefond et al. [16].

THEOREM 1.6. *Let G be a standard GIRG and consider $k \geq 3$. Furthermore, let $U_k = \{v_1, v_2, \ldots, v_k\}$ be a set of vertices chosen uniformly at random and assume without loss of generality that $w_1 \leq \cdots \leq w_k$. Let $\mathbf{E}^c_{star}$ be the event that v_1 connects to all vertices in $U_k \setminus \{v_1\}$ and that $w_k \leq c^d w_1$ for some constant $c \geq 1$ with $c^2(w_1^2/(\tau n))^{1/d} \leq 1/4$. Then, the probability that U_k is a clique conditioned on $\mathbf{E}^c_{star}$ fulfills*

$$\left(\frac{1}{2}\right)^{d(k-1)} k^d \leq \Pr\left[U_k \text{ is clique} \mid \mathbf{E}^c_{star}\right] \leq c^{d(k-2)} \left(\frac{1}{2}\right)^{d(k-1)} k^d.$$

Building on the variant by Decreusefond et al. [16], we provide an alternative proof of the original statement, showing that the clique probability conditioned on the event $\mathbf{E}^c_{\text{star}}$ is monotonous in the weight of all other vertices. Remarkably, this only holds if we condition on the event that the center of our star is of minimal weight among the vertices in U_k.

We apply Theorem 1.6 to bound the clique probability in the whole graph (where the ratio of the maximum and minimum weights of vertices in U_k is not necessarily bounded). Afterwards, we additionally use Chernoff bounds and the second moment method to bound the clique number.

1.4. Relation to previous analyses. In the following, we discuss how our results compare to insights obtained on similar graph models that (apart from not considering weighted vertices) mainly differ in the considered ground space. We note that, in the following, we consider GIRGs with uniform weights in order to obtain a valid comparison.

Random geometric graphs on the sphere. Our results indicate that the GIRG model on the torus behaves similarly to the model of spherical random geometric graphs (SRGGs) in the high-dimensional case. In this model, vertices are distributed on the surface of a $d-1$ dimensional sphere and an edge is present whenever the Euclidean distance between two points (measured by their inner product) falls below a given threshold. Analogously to the behavior of GIRGs, when keeping n fixed and considering increasing $d \to \infty$, this model converges to its nongeometric counterpart, which in their case is the Erdős–Rényi model [17]. It is further shown that the clique number converges to that of an Erdős–Rényi graph (up to a factor of $1 + o(1)$) if $d = \omega(\log^3(n))$.

Although the overall behavior of SRGGs is similar to that of GIRGs, the magnitude of d in comparison to n at which nongeometric features become dominant seems to differ. In fact, it is shown in [10, proof of Theorem 3] that the expected number of triangles in sparse SRGGs still grows with n as long as $d = o(\log^3(n))$, whereas its expectation is constant in the nongeometric, sparse case (as for Erdős–Rényi graphs). On the other hand, in the GIRG model, we show that the expected number of triangles in the sparse case converges to the same (constant) value as that of the nongeometric model if only $d = \omega(\log^{3/2}(n))$. This indicates that, in the high-dimensional regime, differences in the nature of the underlying geometry result in notably different behavior, although in the case of constant dimensionality, the models are often assumed to behave very similarly.

Random geometric graphs on the hypercube. The work of Dall and Christensen [13] and the recent work of Erba et al. [18] show that RGGs on the hypercube do *not* converge to Erdős–Rényi graphs as n is fixed and $d \to \infty$. However, our results imply that this is the case for RGGs on the torus. These apparent disagreements are despite the fact that Erba et al. use a similar central limit theorem for conducting their calculations and simulations [18].

The tools established in our paper yield an explanation for this behavior. Our proof of Theorem 1.1 relies on the fact that, for independent zero-mean variables $Z_1, \dots, Z_d$, the covariance matrix of the random vector $Z = \sum_{i=1}^d Z_i$ is the identity matrix. This, in turn, is based on the fact that the torus is a *homogeneous space,* which implies that the probability measure of a ball of radius r (proportional to its Lebesgue measure or volume, respectively) is the same, regardless of where this ball is centered. It follows that the random variables $Z_{(u,v)}$ and $Z_{(u,s)}$, denoting the normalized distances from u to s and v, respectively, are independent. As a result their covariance is 0 although both "depend" on the position of u.

For the hypercube, this is not the case. Although one may analogously define the distance of two vertices as a sum of independent, zero-mean random vectors over all dimensions just like we do in this paper, the random variables $Z_{(u,v)}$ and $Z_{(u,s)}$ do *not* have a covariance of 0.

1.5. Conjectures and future work. While making the first steps towards understanding GIRGs and sparse RGGs on the torus in high dimensions, we encountered several questions whose investigation does not fit into the scope of this paper. In the following, we give a brief overview of our conjectures and possible starting points for future work.

Noisy GIRGs. It would be interesting to extend our results to the *temperate* version of GIRGs, where the threshold is softened using a *temperature* parameter. That is, while the probability for an edge to exist still decreases with increasing distance, we can now have longer edges and shorter nonedges with certain probabilities. The motivation of this variant of GIRGs is based on the fact that real data is often noisy as well, leading to an even better representation of real-world graphs. In this regard, we remark that most of our proof techniques for the case of constant dimension carry over quite directly to temperate GIRGs. However, when considering nonconstant dimension, having an additional temperature parameter seems to complicate things significantly, which is the reason why we concentrate on threshold GIRGs in this work. Nevertheless, we note that both temperature and dimensionality affect the influence of the underlying geometry, so it would be interesting to further investigate whether a sufficiently high temperature has additional impact on how quickly GIRGs converge to IRGs.

Testing thresholds for detecting underlying geometry. Another crucial question is under which circumstances the underlying geometry of our model remains detectable by means of statistical testing, and when (i.e., for which values of d) our model converges in total variation distance to its nongeometric counterpart. A large body of work has already been devoted to this question for SRGGs [17, 10, 9, 34, 33] and recently also for random intersection graphs [9]. While the question of when these graphs lose their geometry in the dense case is already largely answered, it remains open for the sparse case (where the marginal connection probability is proportional to $1/n$), and progress has only been made recently [9, 33]. It would be interesting to study this question for our model, both for the case of constant and for the case of inhomogeneous weights. Our work indicates that GIRGs and RGGs on the torus might lose their geometry earlier than SRGGs as the number of triangles is in expectation already the same as in an Erdős–Rényi graph if $d = \omega(\log^{3/2}(n))$, while for SRGGs this only happens if $d = \omega(\log^3(n))$ [10].

Furthermore, it remains to investigate dense RGGs on the torus in this regard, where the marginal connection probability of any pair of vertices is constant and does not decrease with n. For dense SRGGs, an analysis of the high-dimensional case has shown that the underlying geometry remains detectable as long as $d = o(n^3)$, while for sparse SRGGs it is conjectured that the respective threshold is only at $d = \log(n)^3$. In the dense case, this is accomplished by counting so-called *signed triangles* [10]. Although for the sparse case, signed triangles have no advantage over ordinary triangles, they are much more powerful in the dense case and might prove useful for analyzing dense RGGs on the torus as well. Additionally, as GIRGs contain both very sparse and very dense parts, it is an interesting question whether inhomogeneous weights actually result in different testing thresholds somewhere between that of the dense and the sparse cases.

2. Preliminaries. We let $G = (V, E)$ be a (random) graph on n vertices. We let $\binom{U}{2}$ be the set of all *possible* edges among vertices of a subset $U \subseteq V$ and denote the *actual* set of edges between them by $E(U) = E \cap \binom{U}{2}$. A k-clique in G is a complete induced subgraph on k vertices of G. We let K_k denote the random variable representing the number of k-cliques in G. We typically use $U_k = \{v_1, \ldots, v_k\}$ to denote a set of k vertices chosen independently and uniformly at random from the graph. For the sake of brevity, we simple write that U_k is a set of *random vertices* to denote that U_k is obtained that way. Further, we write $w_1, \ldots, w_k$ for the weights of the vertices in U_k. The probability that U_k forms a clique is denoted by q_k. Additionally, $\mathbf{E}_{\text{star}}$ is the event that U_k is a star with center v_1.

We use standard Landau notation to describe the asymptotic behavior of functions for sufficiently large n. That is, for functions f, g, we write $f(n) = \mathcal{O}(g(n))$ if there is a constant $c > 0$ such that for all sufficiently large n, $f(n) \leq cg(n)$. Similarly, we write $f(n) = \Omega(g(n))$ if $f(n) \geq cg(n)$ for sufficiently large n. If both statements are true, we write $f(n) = \Theta(g(n))$. Regarding our study of the clustering coefficient, some results make a statement about the asymptotic behavior of a function with respect to a sufficiently large d. These are marked by $\mathcal{O}_d(\cdot), \Omega_d(\cdot), \Theta_d(\cdot)$, respectively.

2.1. Spherical random geometric graphs (SRGGs). In this model, n vertices are distributed uniformly on the d-dimensional unit sphere $\mathcal{S}^{d-1}$ and vertices u, v connected whenever their L_2-distance is below the connection threshold t_{uv}, which is again chosen such that the connection probability of u, v is fixed. This model thus differs from the GIRG model in its ground space (sphere instead of torus) and the fact that it uses homogeneous weights; i.e., the marginal connection probability between each pair of vertices is the same. We mainly use this model as a comparison, since its behavior in high dimensions was extensively studied previously [4, 10, 17].

2.2. Useful bounds and concentration inequalities. Throughout this paper, we use the following approximation of the binomial coefficient.

LEMMA 2.1. *For all $n \geq 1$ and all $1 \leq k \leq \frac{1}{2}n$, we have*

$$\binom{n}{k} = n^k \Theta(k)^{-k}.$$

That is, there are constants $c_1, c_2 > 0$ such that for all $n \geq 1$,

$$n^k \left(c_1 k\right)^{-k} \leq \binom{n}{k} \leq n^k \left(c_2 k\right)^{-k}.$$

Proof. We start with the upper bound and immediately get that for all n, k,

$$\binom{n}{k} \leq \frac{n^k}{k!}.$$

From Stirling's approximation, we get for all $k \geq 1$ that

$$\sqrt{2\pi k} \left(\frac{k}{e}\right)^k e^{\frac{1}{12k+1}} \leq k!.$$

Because $k \geq 1$, the left side is lower bounded by $\left(\frac{k}{e}\right)^k$ and hence

$$\binom{n}{k} \leq n^k \left(e^{-1} k\right)^{-k}.$$

For the lower bound, we observe that

$$\binom{n}{k} \geq \left(\frac{(n-k)}{k}\right)^k = (n-k)^k k^{-k}.$$

We claim that there is a constant $c > 0$ such that $(n-k)^k \geq (cn)^k$, which is equivalent to $1 - \frac{k}{n} \geq c$. As $k \leq \frac{1}{2}n$, this inequality is true for all $c < \frac{1}{2}$, which finishes the proof. ∎

We use the following well-known concentration bounds.

THEOREM 2.2 (Theorem 2.2 in [29], Chernoff–Hoeffding bound). *For $1 \leq i \leq k$, let X_i be independent random variables taking values in $[0,1]$, and let $X := \sum_{i=1}^{k} X_i$. Then, for all $0 < \varepsilon < 1$,*

(i) $\Pr[X > (1+\varepsilon)\mathbb{E}[X]] \leq \exp(-\frac{\varepsilon^2}{3}\mathbb{E}[X])$,

(ii) $\Pr[X < (1-\varepsilon)\mathbb{E}[X]] \leq \exp(-\frac{\varepsilon^2}{2}\mathbb{E}[X])$, *and*

(iii) $\Pr[X \geq t] \leq 2^{-t}$ *for all* $t \geq 2e\mathbb{E}[X]$.

3. Cliques in the low-dimensional regime.

We start by proving the results from Tables 1, 2, and 3 for the case $d = o(\log(n))$. We remind the reader that the results in this section hold for the standard GIRG model but remark that our bounds up to (and including) section 4 are also applicable if norms other than L_∞ are used.

3.1. Bounds on the clique probability.

Recall that we denote by K_k the random variable that represents the number of cliques of size k in G and that q_k is the probability that a set of k vertices chosen uniformly at random forms a clique. Then the expectation of K_k is

$$\mathbb{E}[K_k] = \binom{n}{k} q_k.$$

In the following, we derive upper and lower bounds on q_k. Our bounds here are very general and remain valid regardless of how the dimension scales with n and which L_p-norm is used. One may also easily extend them to the nonthreshold version of the weight sampling model. Although our bounds are asymptotically tight for constant d, they become less meaningful if d scales with n. We therefore derive sharper bounds in section 4 for the case $d = \omega(\log(n))$.

An upper bound on q_k. In this section, we derive an upper bound on q_k by considering the event that a set of k random vertices forms a star centered around the vertex of minimal weight. As this is necessary to form a clique, it gives us an upper bound on q_k that is very general and independent of d. To get sharper upper bounds, we combine this technique with Theorem 1.6 in section 3.4.

Recall that $U_k = \{v_1, \ldots, v_k\}$ is a set of k random vertices with (random) weights $w_1, \ldots, w_k$. In the following, we assume without loss of generality that v_1 is of minimal weight among all vertices in U_k. We start by analyzing how the minimal weight w_1 is distributed.

LEMMA 3.1. *Let G be any GIRG with a power-law weight distribution with exponent $\beta > 2$. Furthermore, let $U_k = \{v_1, \ldots, v_k\}$ be a set of k random vertices and assume that v_1 is of minimal weight w_1 among U_k. Then, w_1 is distributed according to the density function*

$$\rho_{w_1}(x) = \frac{(\beta-1)k}{w_0^{(1-\beta)k}} \cdot x^{(1-\beta)k-1}$$

in the interval $[w_0,\infty]$. Conditioned on the weight w_1, the weight w_i for all $2 \le i \le k$ is distributed independently as

$$\rho_{w_i|w_1}(x) = \frac{\beta-1}{w_1^{1-\beta}} x^{-\beta}.$$

Proof. Recall that the weight w of each vertex is independently sampled from the Pareto distribution such that

$$\Pr[w \le x] = 1 - \left(\frac{x}{w_0}\right)^{1-\beta}. \tag{3.1}$$

Accordingly, the probability that the minimal weight w_1 is at most x is

$$\Pr[w_1 \le x] = 1 - \Pr[w \ge x]^k = 1 - \left(\frac{x}{w_0}\right)^{k(1-\beta)}.$$

To find the density function of w_1, we differentiate this term and get

$$\rho_{w_1}(x) = \frac{\mathrm{d}\Pr[w_v \le x]}{\mathrm{d}x} = \frac{\mathrm{d}}{\mathrm{d}x}\left(1 - \left(\frac{x}{w_0}\right)^{(1-\beta)k}\right) = \frac{(\beta-1)k}{w_0^{(1-\beta)k}} \cdot x^{(1-\beta)k-1}.$$

The conditional density function $\rho_{w_i|w_1}(x)$ of w_i is then

$$\rho_{w_i|w_1}(x) = \frac{\rho_w(x)}{\int_{w_1}^{\infty} \rho_w(x)\mathrm{d}x} = \frac{x^{-\beta}}{\int_{w_1}^{\infty} x^{-\beta}\mathrm{d}x} = \frac{\beta-1}{w_1^{1-\beta}} x^{-\beta},$$

where $\rho_w(x) = \frac{\beta-1}{w_0^{1-\beta}} x^{-\beta}$ is the (unconditional) density function of a single weight. $\square$

We proceed by bounding the probability of the event $\mathbf{E}_{\text{star}}$ that U_k is a star with center v_1. We start with the following lemma.

Lemma 3.2. *Let G be any GIRG with a power-law weight distribution with exponent $\beta > 2$, let $U_k = \{v_1,\dots,v_k\}$ be a set of k random vertices, and assume that v_1 is of minimal weight w_1 among U_k. Furthermore, let $\mathbf{E}_{star}$ be the event that U_k is a star with center v_1 and let $w_-, w_+ \ge w_0$. Then,*

$$\Pr[\mathbf{E}_{star} \cap w_- \le w_1 \le w_+] \le \mathcal{C}\left(\frac{\lambda}{n}\right)^{k-1}\left(\frac{\beta-1}{\beta-2}\right)^{k-1}\left(w_+^{k(3-\beta)-2} - w_-^{k(3-\beta)-2}\right)$$

with

$$\mathcal{C} := \frac{(\beta-1)k w_0^{-(1-\beta)k}}{(3-\beta)k-2}.$$

Proof. Define $P := \Pr[\mathbf{E}_{\text{star}} \cap w_- \le w_1 \le w_+]$. As the marginal connection probability of two vertices u,v with weights w_u, w_v is $\min\{\lambda w_u w_v, 1\}$, we get

$$\begin{aligned} P &\le \int_{w_-}^{w_+}\int_{w_1}^{\infty}\dots\int_{w_1}^{\infty} \frac{\lambda^{k-1} w_1^{k-1} w_2 \dots w_k}{n^{k-1}} \rho_{w_1}(w_1)\rho_{w_2|w_1}(w_2)\dots\rho_{w_k|w_1}(w_k)\mathrm{d}w_k \dots \mathrm{d}w_1 \\ &= \left(\frac{\lambda}{n}\right)^{k-1}\int_{w_-}^{w_+} w_1^{k-1}\rho_{w_1}(w_1)\left(\int_{w_1}^{\infty} w_2 \rho_{w_2|w_1}(w_2)\mathrm{d}w_2\right)^{k-1}\mathrm{d}w_1. \end{aligned}$$

By Lemma 3.1, we have

$$\int_{w_1}^{\infty} w_2 \rho_{w_2|w_1}(w_2) \mathrm{d}w_2 = \frac{\beta-1}{w_1^{1-\beta}} \int_{w_1}^{\infty} w_2^{1-\beta} \mathrm{d}w_2 = \frac{\beta-1}{\beta-2} w_1, \tag{3.2}$$

and therefore, our expression for P simplifies to

$$\begin{aligned} P &\leq \left(\frac{\lambda}{n}\right)^{k-1} \int_{w_-}^{w_+} w_1^{k-1} \rho_{w_1}(w_1) \left(\frac{\beta-1}{\beta-2} w_1\right)^{k-1} \mathrm{d}w_1 \\ &= \left(\frac{\lambda}{n}\right)^{k-1} \left(\frac{\beta-1}{\beta-2}\right)^{k-1} \frac{(\beta-1)k}{w_0^{(1-\beta)k}} \int_{w_-}^{w_+} w_1^{k(3-\beta)-3} \mathrm{d}w_1 \\ &= \mathcal{C} \left(\frac{\lambda}{n}\right)^{k-1} \left(\frac{\beta-1}{\beta-2}\right)^{k-1} \left(w_+^{k(3-\beta)-2} - w_-^{k(3-\beta)-2}\right), \end{aligned}$$

as desired. ∎

Corollary 3.3. *Let G be any GIRG with a power-law weight distribution with exponent $\beta > 2$, let $U_k = \{v_1, \ldots, v_k\}$ be a set of k random vertices, and assume that v_1 is of minimal weight w_1 among U_k. Furthermore, let $\mathbf{E}_{star}$ be the event that U_k is a star with center v_1. Then,*

$$q_k \leq \Pr[\mathbf{E}_{star}] \leq \begin{cases} \Theta(1)^k n^{\frac{k}{2}(1-\beta)} & \text{if } k > \frac{2}{3-\beta} \text{ and } 2 < \beta < 3, \\ \Theta(1)^k n^{1-k} & \text{otherwise.} \end{cases}$$

Proof. Set $w_+ = \sqrt{n/\lambda}$ and observe that

$$\Pr[\mathbf{E}_{\text{star}}] = \Pr[\mathbf{E}_{\text{star}} \cap w_0 \leq w_1 \leq w_+] + \Pr[\mathbf{E}_{\text{star}} \cap w_1 \geq w_+].$$

Note that, if $w_1 \geq w_+ = \sqrt{n/\lambda}$, the formation of a clique (and thus a star) is guaranteed and, hence,

$$\Pr[\mathbf{E}_{\text{star}} \cap w_1 \geq w_+] = \Pr[w_1 \geq w_+] = \left(\frac{w_+}{w_0}\right)^{k(1-\beta)} = \Theta(1)^k n^{\frac{k}{2}(1-\beta)}.$$

To bound $\Pr[\mathbf{E}_{\text{star}} \cap w_0 \leq w_1 \leq w_+]$, we use Lemma 3.2 and observe that

$$\mathcal{C} = \frac{(\beta-1)k w_0^{-(1-\beta)k}}{(3-\beta)k-2}$$

is positive if and only if $k > \frac{2}{3-\beta}$ and $2 < \beta < 3$. In this case Lemma 3.2 implies that $\Pr[\mathbf{E}_{\text{star}}] \leq \Theta(1)^k n^{\frac{k}{2}(1-\beta)}$, as desired. Otherwise, we get $\Pr[\mathbf{E}_{\text{star}} \cap w_0 \leq w_1 \leq w_+] \leq \Theta(1)^k n^{1-k}$, which dominates $\Pr[\mathbf{E}_{\text{star}} \cap w_1 \geq w_+] = \Theta(1)^k n^{\frac{k}{2}(1-\beta)}$. ∎

The above lemma shows that there is a phase transition at $k = \frac{2}{3-\beta}$ if $2 < \beta < 3$, as previously observed by Michielan and Stegehuis [35]. We remark that our bound is independent of the geometry and also works in the temperate variant of the model, i.e., in the nonthreshold case.

A lower bound on q_k. To obtain a matching lower bound, we employ a similar strategy that yields the following lemma.

LEMMA 3.4. *Let $G = \mathrm{GIRG}(n, \beta, w_0, d)$ be a standard GIRG with any L_p-norm and let $w_0 \le w_+ \le \sqrt{n/\lambda}$. Then,*

$$\Pr\left[(U_k \text{ is clique}) \cap (w_0 \le w_1 \le w_+)\right] \ge 2^{-d(k-1)} \mathcal{C} \left(\frac{\lambda}{n}\right)^{k-1} \left(w_+^{k(3-\beta)-2} - w_0^{k(3-\beta)-2}\right)$$

with

$$\mathcal{C} := \frac{(\beta-1)k w_0^{-(1-\beta)k}}{(3-\beta)k-2}.$$

Proof. To get a lower bound on $\Pr[U_k \text{ is clique} \cap w_0 \le w_1 \le w_+]$, we consider the event that every $u \in U_k \setminus \{v_1\}$ is placed at a distance of at most $t_{uv_1}/2$ from v_1. Then, by the triangle inequality, for any $u, v \in U_k \setminus \{v_1\}$, we may bound the distance $d(u,v)$ as

$$d(u,v) \le \frac{1}{2} t_{uv_1} + \frac{1}{2} t_{vv_1} \le t_{uv}$$

because $w_1 \le w_v, w_u$. Hence, u and v are adjacent. The probability that a random vertex u is placed at distance of at most $t_{uv}/2$ from v is equal to the volume $\nu(r)$ of the ball of radius $r = t_{uv}/2$ (but at most 1), i.e.,

$$\min\{1, \nu(t_{uv}/2)\} = \min\left\{1, 2^{-d}\nu(t_{uv})\right\} = \min\left\{1, 2^{-d}\lambda \frac{w_u w_v}{n}\right\}.$$

We remark that it is easy to verify that the above term is also a valid lower bound for the probability of the described event if we are working with some L_p-norm where $p < \infty$. Conditioned on a value of w_1 smaller than $\sqrt{n/\lambda}$, the probability that a vertex $u \in U_k \setminus \{v_1\}$ is placed within distance $t_{uv_1}/2$ from v_1 is thus at least $2^{-d}\lambda w_1^2/n$. Thus,

$$\begin{aligned}
\Pr\left[(U_k \text{ is clique}) \cap (w_0 \le w_1 \le w_+)\right] &\ge \int_{w_0}^{w_+} \frac{2^{-d(k-1)}\lambda^{k-1} w_1^{2(k-1)}}{n^{k-1}} \rho_{w_1}(w_1) \mathrm{d}w_1 \\
&= \left(\frac{2^{-d}\lambda}{n}\right)^{k-1} \frac{(\beta-1)k}{w_0^{(1-\beta)k}} \int_{w_0}^{w_+} w_1^{k(3-\beta)-3} \mathrm{d}w_1 \\
&= 2^{-d(k-1)} \mathcal{C} \left(\frac{\lambda}{n}\right)^{k-1} \left(w_+^{k(3-\beta)-2} - w_0^{k(3-\beta)-2}\right),
\end{aligned}$$

where the first equality is due to Lemma 3.1. ∎

COROLLARY 3.5. *Let $G = \mathrm{GIRG}(n, \beta, w_0, d)$ be a standard GIRG with any L_p-norm. Then,*

$$q_k \ge \begin{cases} \Theta(1)^k n^{\frac{k}{2}(1-\beta)} & \text{if } k > \frac{2}{3-\beta} \text{ and } 2 < \beta < 3, \\ \Theta(1)^k 2^{-dk} n^{1-k} & \text{otherwise.} \end{cases}$$

Proof. The proof is identical to that of Corollary 3.3 with Lemma 3.4 instead of Lemma 3.2. ∎

Hence, the asymptotic behavior of our lower bound for q_k is the same as that of the upper bound up to a factor of $2^{-d(k-1)}\Theta(1)^k$. We remark that our bounds are easily adaptable to the nonthreshold version of the GIRG model as here it is still guaranteed that a pair of vertices placed within its respective connection threshold is adjacent with a constant probability.

3.1.1. A sharper upper bound on q_k for bounded weights. Our upper and lower bounds for the cases $k < \frac{2}{3-\beta}$ and $\beta \geq 3$ still differ by a factor that is exponential in d. In this section, we prove Theorem 1.6, which we restate for the sake of readability, and thereby narrow this gap down under the assumption that the weights of the vertices in U_k are bounded. While this condition is not always met w.h.p. if we choose U_k at random from all vertices, we show how to leverage it to obtain a better bound on the number of cliques in the entire graph in the subsection 3.4.

THEOREM 1.6. *Let G be a standard GIRG and consider $k \geq 3$. Furthermore, let $U_k = \{v_1, v_2, \ldots, v_k\}$ be a set of vertices chosen uniformly at random and assume without loss of generality that $w_1 \leq \cdots \leq w_k$. Let $\mathbf{E}^c_{star}$ be the event that v_1 connects to all vertices in $U_k \setminus \{v_1\}$ and that $w_k \leq c^d w_1$ for some constant $c \geq 1$ with $c^2 \left(w_1^2/(\tau n)\right)^{1/d} \leq 1/4$. Then, the probability that U_k is a clique conditioned on $\mathbf{E}^c_{star}$ fulfills*

$$\left(\frac{1}{2}\right)^{d(k-1)} k^d \leq \Pr\left[U_k \text{ is clique} \mid \mathbf{E}^c_{star}\right] \leq c^{d(k-2)} \left(\frac{1}{2}\right)^{d(k-1)} k^d.$$

Recall that $\tau = 2^d/\lambda$ is a parameter controlling the average degree by influencing the connection threshold. We require the condition $c^2(w_1^2/(\tau n))^{1/d} \leq 1/4$ to ensure that the maximal connection threshold of any pair of vertices in U_k is so small that we can measure the distance between two neighbors of a given vertex as we would in $\mathbb{R}^d$, i.e., without having to take the topology of the torus into account. We remark that this condition is asymptotically fulfilled as long as $d = o(\log(n))$ and $w_1 = \mathcal{O}(n^{1/2-\varepsilon})$ for any $\varepsilon > 0$.

In the following, we let $\{w_i\}_i^k = \{w_1, \ldots, w_k\}$ be the sequence of weights of the vertices in $U_k = \{v_1, \ldots, v_k\}$ whereby we assume without loss of generality that $w_1 \leq \cdots \leq w_k$. We denote by "U_k is star" the event that U_k is a star centered at the vertex of minimum weight (which is v_1). In order to prove Theorem 1.6, we start by showing that $\Pr\left[U_k \text{ is clique} \mid U_k \text{ is star}, \{w_i\}_i^k\right]$ is monotonically increasing in w_i for all $2 \leq i \leq k$. We remark that this property only holds if we condition on having a star centered at the vertex of minimal weight in U_k. With this statement, we may subsequently assume all $u \in U_k \setminus \{v_1\}$ to have a weight of w_1 and $c^d w_1$ for deriving a lower and upper bound on $\Pr\left[U_k \text{ is clique} \mid \mathbf{E}^c_{star}\right]$, respectively.

LEMMA 3.6. *Let $G = \mathrm{GIRG}(n, \beta, w_0, d)$ be a standard GIRG and denote by "U_k is star" the event that U_k forms a star centered at the vertex of minimal weight in U_k. Let further $U_k = \{v_1, \ldots, v_k\}$ be a set of k random vertices with $w_1 \leq \cdots \leq w_k$. Then, the conditional probability $\Pr\left[U_k \text{ is clique} \mid U_k \text{ is star}, \{w_i\}_i^k\right]$ is monotonically increasing in $w_2, \ldots, w_k$.*

Proof. For any $1 \leq i, j \leq k$, denote by t_{ij} the connection threshold $t_{v_i v_j} = \left(\frac{w_i w_j}{\tau n}\right)^{1/d}$. In the following, we abbreviate t_{i1} by t_i for $2 \leq i \leq k$. Note that, since we assume the use of L_∞-norm and condition on $\mathbf{E}^c_{\text{star}}$, the vertex v_i is uniformly distributed in the cube of radius t_i around v_1 for all $2 \leq i \leq k$. Thus, all components of $\mathbf{x}_{v_i}$ are independent and uniformly distributed random variables in the interval $[-t_i, t_i]$ (we choose our coordinate system such that $\mathbf{x}_{v+1}$ is the origin). The probability that U_k is a clique conditioned on $\mathbf{E}^c_{\text{star}}$ is hence equal to the probability that the distance between each pair of points is below their respective connection threshold in every dimension. If we denote by p the probability that this event occurs in one fixed dimension, we get that $\Pr\left[U_k \text{ is clique} \mid \mathbf{E}^c_{\text{star}}, \{w_i\}_i^k\right] = p^d$ because all dimensions are independent. Therefore, it suffices to show the desired monotonicity only for p. In the following, we therefore only consider one fixed dimension and denote the coordinate of the vertex v_i in this dimension with x_i.

Note that the probability p is equal to

$$p = \int_{-t_2}^{t_2} \int_{-t_3}^{t_3} \cdots \int_{-t_k}^{t_k} \rho(x_2) \dots \rho(x_k) \mathbb{1}(x_2, \dots, x_k) \mathrm{d}x_k \dots \mathrm{d}x_2,$$

where $\rho(x_i) = \frac{1}{2t_i}$ is the density function of x_i as x_i is uniformly distributed in the interval $[-t_i, t_i]$, and $\mathbb{1}(x_2, \dots, x_k)$ is an indicator function that is 1 if and only if for all $2 \leq i, j \leq k$, we have $|x_i - x_j| \leq t_{ij}$. We show that p is monotonically increasing in w_i for all $2 \leq i \leq k$. For this, assume without loss of generality that we increase the weight of v_2 by a factor $\xi > 1$. This weight change increases the threshold t_2 by a factor of $\xi^{1/d}$ and we denote the connection threshold between v_i and v_j after the weight change by $\tilde{t}_{ij}$. The connection probability $\tilde{p}$ after this weight increases is

$$\tilde{p} = \int_{-\xi^{1/d} t_2}^{\xi^{1/d} t_2} \int_{-t_3}^{t_3} \cdots \int_{-t_k}^{t_k} \tilde{\rho}(x_2) \rho(x_3) \dots \rho(x_k) \tilde{\mathbb{1}}(x_2, \dots, x_k) \mathrm{d}x_k \dots \mathrm{d}x_2,$$

where $\tilde{\rho}(x_2) = \frac{1}{2\xi^{1/d} t_2} = \rho(x_2)/\xi^{1/d}$, and where $\tilde{\mathbb{1}}$ is defined like $\mathbb{1}$ with the only difference that it uses the new weight of v_2, i.e., $\tilde{\mathbb{1}}$ is 1 if and only if $|x_i - x_j| \leq \tilde{t}_{ij}$ for all $2 \leq i, j \leq k$. Substituting $x_2 = \xi^{1/d} y$, we get

$$\begin{aligned} \tilde{p} &= \xi^{1/d} \int_{-t_2}^{t_2} \int_{-t_3}^{t_3} \cdots \int_{-t_k}^{t_k} \tilde{\rho}(\xi^{1/d} y) \rho(x_3) \dots \rho(x_k) \tilde{\mathbb{1}}(\xi^{1/d} y, \dots, x_k) \mathrm{d}y \dots \mathrm{d}x_k \\ &= \int_{-t_2}^{t_2} \int_{-t_3}^{t_3} \cdots \int_{-t_k}^{t_k} \rho(y) \rho(x_3) \dots \rho(x_k) \tilde{\mathbb{1}}(\xi^{1/d} y, \dots, x_k) \mathrm{d}y \dots \mathrm{d}x_k. \end{aligned}$$

We claim that $\tilde{p} \geq p$, which we show by proving that $\mathbb{1}(y, \dots, x_k) = 1$ implies $\tilde{\mathbb{1}}(\xi^{1/d} y, \dots, x_k) = 1$. For this, assume that $y, x_3, \dots, x_k$ are such that $\mathbb{1}(y, \dots, x_k) = 1$. Note that it suffices to show that for all $3 \leq i \leq k$, if $|x_i - y| \leq t_{2i}$, then $|x_i - \xi^{1/d} y| \leq \xi^{1/d} t_{2i}$. More formally, we have to show that $d_i := |x_i - y| \leq t_{2i}$ implies ${d_i}' := |x_i - \xi^{1/d} y| \leq \xi^{1/d} t_{2i}$.

We note that $|y - \xi^{1/d} y| \leq \xi^{1/d} t_2 - t_2 = t_2(\xi^{1/d} - 1)$, as $|y|$ is at most t_2. Hence, the distance between v_i and v_2 increases by at most $t_2(\xi^{1/d} - 1)$ as well. Furthermore, recall that $t_2 \leq t_{2i}$ as we assume $w_i \geq w_1$. Accordingly,

$$\begin{aligned} d_i' &\leq d_i + (\xi^{1/d} - 1) t_2 \\ &\leq t_{2i} + (\xi^{1/d} - 1) t_2 \\ &\leq t_{2i} + (\xi^{1/d} - 1) t_{2i} \\ &= \xi^{1/d} t_{2i}, \end{aligned}$$

which finishes the proof. ∎

Before proceeding with the proof of Theorem 1.6, we remark that the above statement implies that the entire clique probability (conditional on a given weight sequence) is monotonically increasing in the involved weights. This will be useful in the next section.

COROLLARY 3.7. *Let $\{w_i\}_i^w = \{w_1, \dots, w_k\}$ be a weight sequence with $w_1 \leq \dots \leq w_k$. Then, the probability $\Pr[U_k \text{ is clique} \mid \{w_i\}_i^w]$ is monotonically increasing in $w_1, \dots, w_k$.*

Proof. Recall that "U_k is star" denotes the event that U_k is a star centered at the vertex of minimal weight in U_k and note that

$$\Pr[U_k \text{ is clique} \mid \{w_i\}_i^w] = \Pr[U_k \text{ is clique} \mid U_k \text{ is star}, \{w_i\}_i^w] \Pr[U_k \text{ is star} \mid \{w_i\}_i^w].$$

Hence, if we increase any of the weights $w_2, \dots, w_k$, then both of the above terms on the right-hand side are increasing. If we increase w_1, then note that

$$\begin{aligned}&\Pr[U_k \text{ is clique} \mid \{w_i\}_i^w]\\ &\quad = \Pr[U_k \text{ is clique} \mid U_k \setminus \{v_1\} \text{ is clique}, \{w_i\}_i^w] \Pr[U_k \setminus \{v_1\} \text{ is clique} \mid \{w_i\}_i^w].\end{aligned}$$

Here, the second factor remains the same if we change w_1, and the first factor can only increase if we increase w_1 as—no matter how $v_2, \dots, v_k$ arrange to form a clique—increasing w_1 only increases the probability that v_1 is adjacent to all of them. ∎

We go on with calculating $\Pr[U_k \text{ is clique} \mid \mathbf{E}^c_{\text{star}}]$ under the assumption of uniform weights, which afterwards implies Theorem 1.6.

LEMMA 3.8. *Let $G = \mathrm{GIRG}(n, \beta, w_0, d)$ be a standard GIRG, let U_k and $\mathbf{E}^c_{star}$ be defined as in Theorem* 1.6*, and assume that all vertices in $u \in U_k \setminus \{v_1\}$ have weight $w_u \geq w_1$. Then,*

$$\Pr[U_k \textit{ is clique} \mid \mathbf{E}^c_{star}] = \left(\frac{w_u}{w_1}\right)^{k-2} \left(2^{-(k-1)} \left((2-k)\left(\frac{w_u}{w_1}\right)^{1/d} + 2(k-1)\right)\right)^d.$$

Proof. Again, we only consider one fixed dimension and denote the coordinate of a vertex $u \in U_k$ in this dimension by x_u. Note that by Lemma 3.6, we may assume that the vertices $v_2, \dots, v_k$ all have the same weight. We further refer to the connection threshold between any $u \in U_k \setminus \{v_1\}$ and v_1 as t_0 and to the connection threshold between two vertices in $U_k \setminus \{v_1\}$ as t_u. This enables us to set the origin of our coordinate system such that x_u takes values in $[0, 2t_0]$ for all $u \in U_k$, i.e., such that $x_{v_1} = t_0$. Recall that for all $u \in U_k \setminus \{v_1\}$, x_u is a uniformly distributed random variable.

We refer to the event that the pairwise distance in the coordinates of all $u, v \in U_k \setminus \{v_1\}$ in our fixed dimension is below the connection threshold t_{uv} as U_k being a 1-*D clique.* We calculate $p := \Pr[U_k \text{ is 1-D clique} \mid \mathbf{E}^c_{\text{star}}]$ by integrating over the conditional probability $\Pr[U_k \text{ is 1-D clique} \mid \mathbf{E}^c_{\text{star}}, x_{max}]$ where x_{max} is the coordinate of the rightmost vertex (the one with the largest coordinate) in $U_k \setminus \{v_1\}$. Note that U_k is a 1-D clique if and only if we have $|x_u - x_{max}| \leq t_u$ for all $u \in U_k \setminus \{v_1\}$. We further note that

$$\Pr[U_k \text{ is 1-D clique} \mid \mathbf{E}^c_{\text{star}}, x_{max}] = \begin{cases} 1 & \text{if } x_{max} \leq t_u, \\ \left(\frac{t_u}{x_{max}}\right)^{k-2} & \text{otherwise.}\end{cases}$$

It remains to derive the distribution of x_{max}. For this, we derive its density function as

$$\begin{aligned}\rho(x_{max}) &= \frac{\mathrm{d}\Pr[x_{max} \leq x]}{\mathrm{d}x}\\ &= \frac{\mathrm{d}}{\mathrm{d}x}\left(\Pr[y \leq x]^{k-1}\right)\\ &= \frac{\mathrm{d}}{\mathrm{d}x}\left(\left(\frac{x}{2t_0}\right)^{k-1}\right)\\ &= \frac{k-1}{(2t_0)^{k-1}} x^{k-2}.\end{aligned}$$

With this, we may deduce

$$\begin{aligned}
\Pr\left[U_k \text{ is 1-D clique} \mid \mathbf{E}^c_{\text{star}}\right] &= \int_0^{2t_0} \rho(x_{max}) \Pr\left[U_k \text{ is 1-D clique} \mid x_{max}\right] \mathrm{d}x_{max} \\
&= \frac{k-1}{(2t_0)^{k-1}} \left(\int_0^{t_u} x^{k-2} \mathrm{d}x + t_u^{k-2} \int_{t_u}^{2t_0} 1 \mathrm{d}x \right) \\
&= 2^{-(k-1)} \left(\left(\frac{t_u}{t_0} \right)^{k-1} + (k-1) \frac{t_u^{k-2}}{t_0^{k-1}} (2t_0 - t_u) \right) \\
&= 2^{-(k-1)} \left(\left(\frac{t_u}{t_0} \right)^{k-1} + 2(k-1) \left(\frac{t_u}{t_0} \right)^{k-2} - (k-1) \left(\frac{t_u}{t_0} \right)^{k-1} \right) \\
&= 2^{-(k-1)} \left(\frac{w_u}{w_1} \right)^{\frac{k-2}{d}} \left((2-k) \left(\frac{w_u}{w_1} \right)^{1/d} + 2(k-1) \right).
\end{aligned}$$

Since $\Pr\left[U_k \text{ is clique} \mid \mathbf{E}^c_{\text{star}}\right] = \Pr\left[U_k \text{ is 1-D clique} \mid \mathbf{E}^c_{\text{star}}\right]^d$, this finishes the proof. ∎

Proof of Theorem 1.6. We use the monotonicity of $\Pr\left[U_k \text{ is clique} \mid \mathbf{E}^c_{\text{star}}\right]$ obtained by Lemma 3.8. Due to this monotonicity, it is sufficient to assume that $w_u = w_1$ for all $u \in U_k$ for deriving the lower bound. Hence, we have $w_u/w_1 = 1$ and the expression in Lemma 3.8 simplifies to

$$2^{-d(k-1)} k^d.$$

For the upper bound, we instead assume $w_u = c^d w_1$ implying that $(w_u/w_1)^{1/d} = c$. Lemma 3.8 implies that

$$\begin{aligned}
\Pr\left[U_k \text{ is clique} \mid \mathbf{E}^c_{\text{star}}\right] &\le c^{d(k-2)} 2^{-d(k-1)} \left((2-k)c + 2(k-1)\right)^d \\
&\le c^{d(k-2)} 2^{-d(k-1)} k^d = \frac{1}{c^d} \left(\frac{c}{2} \right)^{d(k-1)} k^d,
\end{aligned}$$

where we used that $(2-k)c + 2(k-1) \le k$ for all $c \ge 1$ and $k \ge 2$. ∎

3.2. Characterizing cliques by vertex weights. After establishing bounds on the clique probability in the whole graph, we now turn to characterizing the clique probability in specific parts of the graph in order to prove the statements in Table 4. Note that the proofs for the regime $k < \frac{2}{3-\beta}$ and $d = \omega(\log(n))$ are in subsection 4.2; all the rest is proven here.

Recall that $w_{\min}$ and $w_{\max}$ are the minimum and maximum weights among U_k. Furthermore we assume that $U_k = \{v_1, \ldots, v_k\}$ with associated weights $w_1 \le w_2 \le \cdots \le w_k$. Note that $w_1 = w_{\min}, w_k = w_{\max}$. We start by showing that cliques of size $k > \frac{2}{3-\beta}$ dominantly form within the heavy core.

LEMMA 3.9. *Let $k > \frac{2}{3-\beta}$, $\beta \in (2,3)$. Then for any $p \in (0,1)$, there is an $\varepsilon > 0$ such that*

$$\Pr\left[w_{\min} \in M_\varepsilon(\sqrt{n}) \mid U_k \textit{ is clique }\right] \ge p.$$

Proof. In the first part, we show that $\Pr\left[w_{\min} < \varepsilon\sqrt{n} \mid U_k \text{ is clique }\right] \le 1 - p$ for some $\varepsilon > 0$. To this end, recall that $U_k = \{v_1, v_2, \ldots, v_k\}$ is a set of k random vertices with weights $w_1, \ldots, w_k$ whereby we assume without loss of generality that $w_1 \le w_2 \le \cdots \le w_k$. We have

$$\Pr\left[w_{\min} < \varepsilon w \mid U_k \text{ is clique }\right] = \frac{\Pr\left[(w_{\min} < \varepsilon w) \cap (U_k \text{ is clique })\right]}{\Pr\left[U_k \text{ is clique }\right]}.$$

By Corollary 3.5, we get that there is a constant $c_1 > 0$ such that $\Pr[U_k \text{ is clique}] \geq c_1^k n^{\frac{k}{2}(1-\beta)}$. By Lemma 3.2, where we set $w_- = w_0$, we have

$$\begin{aligned}&\Pr[(w_{\min} < \varepsilon w) \cap (U_k \text{ is clique})]\\ &\quad\leq \frac{(\beta-1)k w_0^{-(1-\beta)}}{(3-\beta)k-2}\left(\frac{\lambda(\beta-1)w_0^{-(1-\beta)}}{\beta-2}\right)^{k-1} n^{1-k}(\varepsilon w)^{k(3-\beta)-2}\\ &\quad\leq c_2^k n^{1-k}(\varepsilon w)^{k(3-\beta)-2}\end{aligned}$$

for some constant $c_2 > 0$ as $\frac{(\beta-1)k}{(3-\beta)k-2}$ is at most a constant for all (potentially super-constant) $k > \frac{2}{3-\beta}$. Hence,

$$\begin{aligned}\Pr[w_{\min} < \varepsilon w \mid U_k \text{ is clique }] &\leq \frac{c_2^k n^{1-k}(\varepsilon w)^{k(3-\beta)-2}}{c_1^k n^{\frac{k}{2}(1-\beta)}}\\ &= \varepsilon^{k(3-\beta)-2}\left(\frac{c_2}{c_1}\right)^k \frac{n^{1-k} w^{k(3-\beta)-2}}{n^{\frac{k}{2}(1-\beta)}}.\end{aligned}$$

Setting $w = \sqrt{n}$ yields

$$\Pr\left[w_{\min} < \varepsilon\sqrt{n} \mid U_k \text{ is clique }\right] \leq \varepsilon^{k(3-\beta)-2}\left(\frac{c_2}{c_1}\right)^k = \varepsilon^{-2}\left(\frac{c_2\varepsilon^{3-\beta}}{c_1}\right)^k.$$

If k is a constant, we can see that (since $k(3-\beta)-2 > 0$ by our assumption on k), choosing an $\varepsilon > 0$ small enough, the above probability is at most $1-p$, as desired. For $k = \omega(1)$, choosing any $\varepsilon < (c_1/c_2)^{\frac{1}{3-\beta}}$ yields that the above probability is $o(1)$ and thus shows that our statement holds for sufficiently large n.

It remains to show that also $\Pr[w_{\min} > \sqrt{n}/\varepsilon \mid U_k \text{ is clique }] \leq 1-p$ for some $\varepsilon > 0$. Here, it suffices to observe that for $\varepsilon < \lambda$ a clique is formed if $w_{\min} > \sqrt{n}/\varepsilon$. Thus, if $\varepsilon < \lambda$,

$$\begin{aligned}\Pr\left[w_{\min} > \sqrt{n}/\varepsilon \mid U_k \text{ is clique }\right] &\leq \frac{\Pr[w_{\min} > \sqrt{n}/\varepsilon \cap U_k \text{ is clique }]}{\Pr[U_k \text{ is clique}]}\\ &\leq \frac{\Pr[w_{\min} \geq \sqrt{n}/\varepsilon]}{\Pr[w_{\min} \geq \sqrt{n}/\lambda]}\\ &= (\lambda/\varepsilon)^{k(1-\beta)},\end{aligned}$$

which approaches 0 as $\varepsilon \to 0$. ∎

The next lemma proves the claimed bounds for $w_{\max}$ based on the previous result. Note that this also proves what we want for the entire regime $d = \omega(\log(n)), \beta \in (2,3)$ after we establish suitable bounds for $w_{\min}$ in subsection 4.2.

LEMMA 3.10. *Assume that for every $p \in (0,1)$ there is some $\delta > 0$ such that*

$$\Pr\left[w_{\min} \in M_\delta(\sqrt{n}) \mid U_k \textit{ is clique }\right] \geq p.$$

Then, for every $p \in (0,1)$ there is also some $\varepsilon > 0$ such that

$$\Pr\left[w_{\max} \in M_\varepsilon(\sqrt{n}k^{\frac{1}{\beta-1}}) \mid U_k \textit{ is clique }\right] \geq p.$$

Proof. Bound

$$\begin{aligned}
&\Pr\left[w_{\max} > \sqrt{n}k^{\frac{1}{\beta-1}}/\varepsilon \mid U_k \text{ is clique }\right] \\
&\quad\le \Pr\left[w_{\max} > \sqrt{n}k^{\frac{1}{\beta-1}}/\varepsilon \mid U_k \text{ is clique } \cap w_{\min} \ge \delta\sqrt{n}\right] \\
&\qquad + \Pr\left[w_{\min} \le \delta\sqrt{n} \mid U_k \text{ is clique}\right].
\end{aligned}$$

From the assumption in our statement, we know that the second term is upper bounded by some function $f(\delta)$ that approaches 0 as $\delta \to 0$. We bound the first term as follows:

$$\begin{aligned}
&\Pr\left[w_{\max} > \sqrt{n}k^{\frac{1}{\beta-1}}/\varepsilon \mid U_k \text{ is clique } \cap w_{\min} \ge \delta\sqrt{n}\right] \\
&\quad\le \sum_{i=2}^{k} \Pr\left[w_i > \sqrt{n}k^{\frac{1}{\beta-1}}/\varepsilon \mid U_k \text{ is clique } \cap w_{\min} \ge \delta\sqrt{n}\right] \\
&\quad\le k \cdot \frac{\Pr\left[w_2 > \sqrt{n}k^{\frac{1}{\beta-1}}/\varepsilon \cap U_k \text{ is clique } \mid w_{\min} \ge \delta\sqrt{n}\right]}{\Pr\left[U_k \text{ is clique } \mid w_{\min} \ge \delta\sqrt{n}\right]}.
\end{aligned}$$

Conditional on $w_{\min} \ge \delta\sqrt{n}$, the U_k is a clique if $U_k \setminus \{v_2\}$ is a clique and if w_2 is at least $\sqrt{n}/(\lambda\delta)$ because then the connection probability of every pair of vertices is 1. Hence,

$$\begin{aligned}
&\Pr\left[w_{\max} > \sqrt{n}k^{\frac{1}{\beta-1}}/\varepsilon \mid U_k \text{ is clique } \cap w_{\min} \ge \delta\sqrt{n}\right] \\
&\quad\le k \cdot \frac{\Pr\left[U_k \setminus \{v_2\} \text{ is clique } \cap w_2 > \sqrt{n}k^{\frac{1}{\beta-1}}/\varepsilon \mid w_{\min} \ge \delta\sqrt{n}\right]}{\Pr\left[U_k \setminus \{v_2\} \text{ is clique } \cap w_2 \ge \sqrt{n}/(\lambda\delta) \mid w_{\min} \ge \delta\sqrt{n}\right]} \\
&\quad= k \cdot \frac{\Pr\left[w_2 > \sqrt{n}k^{\frac{1}{\beta-1}}/\varepsilon \mid w_{\min} \ge \delta\sqrt{n}\right]}{\Pr\left[w_2 \ge \sqrt{n}/(\lambda\delta) \mid w_{\min} \ge \delta\sqrt{n}\right]},
\end{aligned}$$

where the last step is due to independence. By the definition of the Pareto distribution, this is at most

$$k \cdot \left(\frac{\sqrt{n}k^{\frac{1}{\beta-1}}/\varepsilon}{\sqrt{n}/(\delta\lambda)}\right)^{1-\beta} = (\lambda\delta/\varepsilon)^{1-\beta},$$

so in total

$$\Pr\left[w_{\max} > \sqrt{n}k^{\frac{1}{\beta-1}}/\varepsilon \mid U_k \text{ is clique }\right] \le (\lambda\delta/\varepsilon)^{1-\beta} + f(\delta)$$

and setting $\delta = \sqrt{\varepsilon}$ yields that this function tends to zero as $\varepsilon \to 0$, and the proof of the first part is finished.

For the second part, bound

$$\begin{aligned}
&\Pr\left[w_{\max} < \varepsilon\sqrt{n}k^{\frac{1}{\beta-1}} \mid U_k \text{ is clique}\right] \\
&\quad\le \Pr\left[w_{\max} < \varepsilon\sqrt{n}k^{\frac{1}{\beta-1}} \mid U_k \text{ is clique} \cap w_{\min} \ge \delta\sqrt{n}\right] \\
&\qquad + \Pr\left[w_{\min} < \delta\sqrt{n} \mid U_k \text{ is clique}\right].
\end{aligned}$$

Again, by the assumption in our statement, we can make the second term arbitrarily small by adjusting δ. For the first term, we bound

$$\begin{aligned}&\Pr\left[w_{\max} < \varepsilon\sqrt{n}k^{\frac{1}{\beta-1}} \mid U_k \text{ is clique} \cap w_{\min} \geq \delta\sqrt{n}\right]\\ &\quad= \Pr\left[w_{\max} < \varepsilon\sqrt{n}k^{\frac{1}{\beta-1}} \mid w_{\min} \geq \delta\sqrt{n}\right]\\ &\qquad\cdot \frac{\Pr\left[U_k \text{ is clique} \mid (w_{\max} < \varepsilon\sqrt{n}k^{\frac{1}{\beta-1}}) \cap (w_{\min} \geq \delta\sqrt{n})\right]}{\Pr\left[U_k \text{ is clique} \mid w_{\min} \geq \delta\sqrt{n}\right]}.\end{aligned}$$

It follows by a coupling argument and the fact that the clique probability is monotonically increasing in the weights of the involved vertices that the fraction above is at most 1. Furthermore, from the definition of the Pareto definition we have

$$\begin{aligned}\Pr\left[w_{\max} < \varepsilon\sqrt{n}k^{\frac{1}{\beta-1}} \mid w_{\min} \geq \delta\sqrt{n}\right] &\leq \left(1 - \Omega\left(\frac{\varepsilon k^{\frac{1}{\beta-1}}}{\delta}\right)^{1-\beta}\right)^{k-1}\\ &= \exp\left(-\Omega(\varepsilon/\delta)^{1-\beta}\right).\end{aligned}$$

Hence, in total,

$$\Pr\left[w_{\max} < \varepsilon\sqrt{n}k^{\frac{1}{\beta-1}} \mid U_k \text{ is clique}\right] \leq \exp\left(-\Omega(\varepsilon/\delta)^{1-\beta}\right) + f(\delta),$$

where f is a function that tends to 0 as $\delta \to 0$. Setting $\delta = \sqrt{\varepsilon}$ yields that this holds for the entire right-hand side and finishes the proof. $\square$

We turn to the regime $d = o(\log(n))$ and $k \leq \frac{2}{3-\beta}$ or $\beta > 3$. We start with the following lemma, which tells us that here at least one vertex of small weight, i.e., weight on the order of $\exp(\mathcal{O}(1)d) = n^{o(1)}$, is involved in a clique. We afterwards extend this statement to the other vertices involved in a clique.

LEMMA 3.11. *Let $d = o(\log(n))$ and U_k be a set of k random vertices. Let $w_{\min}$ be the minimum weight among U_k. If $\beta > 3$ or $k < \frac{2}{3-\beta}$, there is a constant $c > 0$ (independent of k) such that for all $p \in (0,1)$ there is an $\varepsilon > 0$ such that*

$$\Pr\left[w_{\min} \leq e^{cd}/\varepsilon \mid U_k \textit{ is clique }\right] \geq p.$$

Proof. Similarly as in the proof of Lemma 3.9, we use Lemma 3.2 to obtain

$$\begin{aligned}&\Pr\left[(w_1 \geq w/\varepsilon) \cap (U_k \text{ is clique })\right]\\ &\quad\leq \frac{(1-\beta)kw_0^{-(1-\beta)}}{(3-\beta)k-2}\left(\frac{\lambda(\beta-1)w_0^{-(1-\beta)}}{\beta-2}\right)^{k-1} n^{1-k}(w/\varepsilon)^{k(3-\beta)-2}\\ &\quad\leq c_2^k n^{1-k}(w/\varepsilon)^{k(3-\beta)-2}\end{aligned}$$

for some constant $c_2 > 0$. By Corollary 3.5, we get that there is a constant c_1 such that $\Pr[U_k \text{ is clique}] \geq (c_1 2^{-d})^k n^{1-k}$. Define $\alpha = k(3-\beta)-2$ and note how this implies

$$\Pr\left[w_1 \geq w/\varepsilon \mid U_k \text{ is clique }\right] \leq (w/\varepsilon)^{\alpha}\left(\frac{c_2}{c_1 2^{-d}}\right)^k = (w/\varepsilon)^{-2}\left(\frac{c_2(w/\varepsilon)^{3-\beta}}{c_1 2^{-d}}\right)^k.$$

Note that due to our assumptions on β and k, we have $\alpha < 0$. If $\beta \in (2,3)$, we only have to consider the case $k < \frac{2}{3-\beta} = \text{const}$. Hence $\Pr[w_1 \geq w/\varepsilon \mid U_k \text{ is clique }] = c_3(w/\varepsilon)^{\alpha}2^{dk}$ for some constant $c_3 > 0$ and setting $w \geq 2^{\frac{dk}{-\alpha}}$ yields that the above probability is at most c_3/ε^α as desired (recall that $\alpha < 0$). Note that c_3 depends on k; however, since we only consider a constant number of different values of k (namely all integers between 3 and $\frac{2}{3-\beta}$), we may as well choose it independent of k by taking the maximum over all these k.

If $\beta > 3$, let us choose $w = 2^{\max\{1, \frac{1}{\beta-3}\}d}$. Then, as $w \geq 1$, for any $\varepsilon < (c_2/c_1)^{\frac{1}{3-\beta}}$, we have $\Pr[w_1 \geq \varepsilon w \mid U_k \text{ is clique }] \leq \varepsilon^{-2} c_4^k$ for some constant $c_4 < 1$, and the proof is finished. ∎

We now proceed by bounding the maximum weight associated to a clique. To this end, we relate the probability that U_k is a clique (assuming $w_{\min}$ is small) to the probability that U_{k-1} is a clique in the following two lemmas.

LEMMA 3.12. *If $d = o(\log(n))$ and $\beta > 3$, there are constants $a, t_0 > 0$ such that for all $t \geq t_0, w \geq w_0$ and all $k \geq 3$,*

$$\Pr[U_k \textit{ is clique} \cap w_{max} \geq t \mid w_{\min} \leq w] \leq \Pr[U_{k-1} \textit{ is clique} \mid w_{\min} \leq w] \cdot \frac{awkt^{2-\beta}}{n}.$$

Proof. Note that if we condition on any $w_{\min}$, all vertices in $U_k \setminus \{v_{\min}\}$ are distributed as independent Pareto random variables with parameters $\beta, w_{\min}$. By a union bound,

$$\Pr[U_k \text{ is clique} \cap w_{\max} \geq t \mid w_{\min}] \leq \sum_{v \in U_k \setminus \{v_{\min}\}} \Pr[U_k \text{ is clique} \cap w_v \geq t \mid w_{\min}].$$

Now for any $v \in U_k \setminus \{v_{\min}\}$,

$$\begin{aligned}
&\Pr[U_k \text{ is clique} \cap w_v \geq t \mid w_{\min}] \\
&\quad = \Pr[\Pr[U_k \text{ is clique} \cap w_v \geq t \mid U_k \setminus \{v\} \text{ is clique}, w_{\min}]]\Pr[U_k \setminus \{v\} \text{ is clique} \mid w_{\min}] \\
&\quad = \Pr[U_k \text{ is clique} \cap w_v \geq t \mid U_k \setminus \{v\} \text{ is clique}, w_{\min}]\Pr[U_{k-1} \text{ is clique} \mid w_{\min}].
\end{aligned}$$

Hence, it remains to bound the first factor above. To this end, we consider the necessary event that w_v is adjacent to $v_{\min}$ and obtain

$$\begin{aligned}
\Pr[U_k \text{ is clique} \cap w_v \geq t \mid U_k \setminus \{v\} \text{ is clique}, w_{\min}] &\leq \Pr[v \sim v_{\min} \cap w_v \geq t \mid w_{\min}] \\
&\leq \frac{\lambda w_{\min}}{n} \int_t^\infty c w^{1-\beta} \mathrm{d}w \\
&= \frac{a w_{\min} t^{2-\beta}}{n},
\end{aligned}$$

where c, a are constants. Summing over all v yields the desired statement. ∎

LEMMA 3.13. *Let $d = o(\log(n))$. Then there are constants $a, c > 0$ such that for any $w \geq w_0$,*

$$\Pr[U_k \textit{ is clique} \mid w_{\min} \leq w] \geq \Pr[U_{k-1} \textit{ is clique} \mid w_{\min} \leq w] \cdot \frac{a w^{1-\beta} e^{-cd}}{n}.$$

Proof. Fix any $v \in U_k \setminus \{v_{\min}\}$ and observe

$$\begin{aligned}
\Pr[U_k \text{ is clique} \mid w_{\min} \leq w] = &\Pr[U_k \text{ is clique} \mid U_k \setminus \{v\} \text{ is clique} \cap w_{\min} \leq w] \\
&\cdot \Pr[U_k \setminus \{v\} \text{ is clique } \mid w_{\min} \leq w],
\end{aligned}$$

so it remains to find a lower bound for the first factor. We consider the event that w_k is sufficiently large such that it is sufficiently likely that v_k is placed close enough to $v_{\min}$ to be connected to all the other vertices by the triangle inequality. We note that, if v is placed within distance

$$\phi := \left(\frac{w_v w_{\min}}{\tau n}\right)^{1/d} - \left(\frac{{w_{\min}}^2}{\tau n}\right)^{1/d} = \left(\frac{w_{\min}}{\tau n}\right)^{1/d}\left(w_v^{1/d} - w_{\min}^{1/d}\right)$$

of $v_{\min}$, then it must also be adjacent to all the other vertices. Assuming that $w_v \geq \alpha w_{\min}$ for some $\alpha > 0$, this event occurs with probability

$$\phi^d = \frac{\lambda w_{\min}}{n}\left(w_k^{1/d} - w_{\min}^{1/d}\right)^d \geq \frac{\lambda w_0^2}{n}\left(\alpha^{1/d} - 1\right)^d \geq \frac{\lambda w_0^2}{n}\left(\frac{\ln(\alpha)}{d}\right)^d,$$

where in the last step we used the inequality $e^x \geq 1 + x$. Choosing $\alpha = e^d$ yields that the above probability is at least $\lambda w_0^2/n$. Furthermore, the probability that $w_v \geq \alpha w_{\min} = e^d w_{\min}$ is

$$\left(\frac{\alpha w_{\min}}{w_0}\right)^{1-\beta} \geq \left(\frac{w e^d}{w_0}\right)^{1-\beta}.$$

In total,

$$\Pr\left[U_k \text{ is clique} \mid U_k \setminus \{v\} \text{ is clique} \cap w_{\min} \leq w\right] \geq \frac{\lambda w_0^2}{n}\left(\frac{w e^d}{w_0}\right)^{1-\beta} = \frac{a w^{1-\beta} e^{-cd}}{n}$$

for some constants $a, c > 0$. ∎

Using the previous two lemmas, we can bound the maximum weight associated to a clique as follows.

LEMMA 3.14. *If $d = o(\log(n))$ and $\beta > 3$, there is a constant $c > 0$ such that for any $p \in (0,1)$, there is an $\varepsilon > 0$ such that*

$$\Pr\left[w_{\max} \leq e^{cd} k^{\frac{1}{\beta-2}}/\varepsilon \mid U_k \text{ is clique}\right] \geq p.$$

Proof. Observe that for any δ, α,

$$\begin{aligned}
&\Pr\left[w_{\max} \geq e^{cd} k^{\frac{1}{\beta-2}}/\varepsilon \mid U_k \text{ is clique}\right] \\
&\quad \leq \Pr\left[w_{\max} \geq e^{cd} k^{\frac{1}{\beta-2}}/\varepsilon \mid U_k \text{ is clique} \cap w_{\min} \leq e^{\alpha d}/\delta\right] \\
&\qquad + \Pr\left[w_{\min} \geq e^{\alpha d}/\delta \mid U_k \text{ is clique}\right].
\end{aligned}$$

Now, by Lemma 3.11, if we choose a suitable constant α, we can make the second term arbitrarily small by choosing δ small enough. To bound the first term, observe further that

$$\begin{aligned}
&\Pr\left[w_{\max} \geq e^{cd} k^{\frac{1}{\beta-2}}/\varepsilon \mid U_k \text{ is clique} \cap w_{\min} \leq e^{\alpha d}/\delta\right] \\
&\quad = \frac{\Pr\left[U_k \text{ is clique } \cap w_{\max} \geq e^{cd} k^{\frac{1}{\beta-2}}/\varepsilon \mid w_{\min} \leq e^{\alpha d}/\delta\right]}{\Pr\left[U_k \text{ is clique } \mid w_{\min} \leq e^{\alpha d}/\delta\right]}.
\end{aligned}$$

By Lemma 3.12, there is a constant a_1 such that the numerator is bounded from above by

$$\frac{a_1}{n}\left(e^{\alpha d}/\delta\right) k \left(e^{cd} k^{\frac{1}{\beta-2}}/\varepsilon\right)^{2-\beta} = \frac{a_1}{n}\frac{e^{\alpha d + (2-\beta)cd}}{\delta \varepsilon^{2-\beta}}.$$

Similarly, by Lemma 3.13, the denominator is bounded from below by

$$\frac{a_2}{n} e^{-c'd} \left(e^{\alpha d}/\delta\right)^{1-\beta}$$

for some constants a_2, c'. Combining these, the fraction is bounded from above by

$$\frac{a_1}{a_2} \exp\left(\alpha d + c'd + (2-\beta)cd\right) \delta^{-\beta} \varepsilon^{\beta-2}.$$

In total, there are constants $s, a_3, a_4 > 0$ such that

$$\Pr\left[w_{\max} \geq e^{cd} k^{\frac{1}{\beta-2}}/\varepsilon \mid U_k \text{ is clique }\right] \leq a_3 \delta^{-\beta} \varepsilon^{\beta-2} + a_4 \delta^s.$$

Setting $\delta = \varepsilon^\gamma$ for any $0 < \gamma < \frac{\beta-2}{\beta}$ then yields that this term tends to 0 as ε tends to 0 and implies the desired statement. ∎

Finally, we show that the minimum weight associated to a clique is at least of order $k^{\frac{1}{\beta-2}}$. This essentially follows from the fact that this holds for stars centered at $v_{\min}$; we show additionally that conditioning on a clique only induces a bias towards even larger weights. This implies that our bounds are tight up to a factor $e^{\Theta(1)d} = n^{o(1)}$.

Lemma 3.15. *Let $d = o(\log(n))$. Assume further that $k < \frac{2}{3-\beta}$ or that $\beta > 3$. Then for every $p \in (0,1)$ there is an $\varepsilon > 0$ such that*

$$\Pr\left[w_{\max} \in M_\varepsilon^{(-)}(k^{\frac{1}{\beta-2}})\right] \geq p.$$

Proof. In the following—again—the event "U_k is star" denotes the event that U_k is a star centered at $v_{\min}$. Bound

$$\begin{aligned}
&\Pr\left[w_{\max} < \varepsilon k^{\frac{1}{\beta-2}} \mid U_k \text{ is clique}\right] \\
&\quad = \frac{\Pr\left[w_{\max} < \varepsilon k^{\frac{1}{\beta-2}} \mid U_k \text{ is star}\right] \Pr\left[U_k \text{ is clique} \mid U_k \text{ is star} \cap (w_{\max} < \varepsilon k^{\frac{1}{\beta-2}})\right]}{\Pr\left[U_k \text{ is clique} \mid U_k \text{ is star}\right]}.
\end{aligned}$$

It follows by a coupling argument and the fact that the clique probability is monotonically increasing in $w_2, \dots, w_k$ that

$$\frac{\Pr\left[U_k \text{ is clique} \mid U_k \text{ is star} \cap (w_{\max} < \varepsilon k^{\frac{1}{\beta-2}})\right]}{\Pr\left[U_k \text{ is clique} \mid U_k \text{ is star}\right]} \leq 1.$$

Moreover, by the Pareto distribution and the fact that the edges in a star are all independent, we get that there is a constant $c > 0$ such that

$$\begin{aligned}
\Pr\left[w_{\max} < w \mid U_k \text{ is star} \cap (w_{\min} = x)\right] &= \left(\frac{\int_x^w \frac{cxy}{n} y^{-\beta} \mathrm{d}y}{\int_x^\infty \frac{cxy}{n} y^{-\beta} \mathrm{d}y}\right)^{k-1} \\
&= \left(1 - \Omega(w/x)^{2-\beta}\right)^{k-1}.
\end{aligned}$$

Thus,

$$\begin{aligned}
\Pr\left[w_{\max} < \varepsilon k^{\frac{1}{\beta-2}} \mid U_k \text{ is star} \cap (w_{\min} = x)\right] &\leq \left(1 - \Omega\left(\frac{\varepsilon k^{\frac{1}{\beta-2}}}{x}\right)^{2-\beta}\right)^{k-1} \\
&= \exp(-\Omega((x/\varepsilon)^{\beta-2})).
\end{aligned}$$

As $w_{\min} \geq w_0$ deterministically, we get

$$\Pr\left[w_{\max} < \varepsilon k^{\frac{1}{\beta-2}} \mid U_k \text{ is star}\right] \leq \exp(-\Omega(1/\varepsilon)^{\beta-2}),$$

which approaches 0 as $\varepsilon \to 0$, as desired. ∎

3.3. Bounding the variance of typical cliques. After bounding the expected number of cliques and characterizing the clique probability by vertex weight, we use the gained insights for bounding the variance of cliques restricted to the dominant regimes identified previously. This, together with the results from Table 4, allows us further to derive concentration bounds on the total number of cliques later in section 5.

In the following, given a set of admissible weights M, we denote by $K_k(M)$ the number of k-cliques with vertex weights that are in M, and we derive bounds on the variance of $K_k(M_\varepsilon(w))$ for suitable w (recall the definition of M_ε from Definition 1.2).

LEMMA 3.16. *For all $k \geq 3$ and for any set of weights M,*

$$\operatorname{Var}[K_k(M)] \leq \mathbb{E}[K_k(M)] \sum_{\ell=1}^{k} \binom{k}{\ell} \mathbb{E}[K_{k-\ell}(M)].$$

Proof. For a set of vertices A, we write that "A is clique in M" if A's vertex weights are in M. Using this notation, we have

$$\mathbb{E}\left[K_k(M)^2\right] = \left(\sum_{A} \Pr[M_1 \text{ is clique in } M]\right),$$

where the sum goes over all k-element subsets of vertices. Accordingly,

$$\begin{aligned}
&\mathbb{E}\left[K_k(M)^2\right] \\
&\quad \leq \left(\sum_{A_1} \Pr[A_1 \text{ is clique in } M]\right)^2 \\
&\qquad + \sum_{A_1} \sum_{\substack{A_2 \\ |A_1 \cap A_2| \geq 1}} \Pr[(A_1 \text{ is clique in } M) \cap (A_2 \text{ is clique in } M)] \\
&\quad \leq \mathbb{E}[K_k(M)]^2 \\
&\qquad + \sum_{A_1} \sum_{\substack{A_2 \\ |A_1 \cap A_2| \geq 1}} \Pr[A_1 \text{ is clique in } M] \Pr[A_2 \setminus A_1 \text{ is clique in } M]. \\
&\quad = \mathbb{E}[K_k(M)]^2 \\
&\qquad + \binom{n}{k} \Pr[U_k \text{ is clique in } M] \sum_{\ell=1}^{k} \binom{k}{\ell} \binom{n}{k-\ell} \Pr[U_\ell \text{ is clique in } M] \\
&\quad = \mathbb{E}[K_k(M)]^2 + \mathbb{E}[K_k(M)] \sum_{\ell=1}^{k} \binom{k}{\ell} \mathbb{E}[K_{k-\ell}(M)].
\end{aligned}$$

By $\operatorname{Var}[X] = \mathbb{E}\left[X^2\right] - \mathbb{E}[X]^2$, the proof is finished. ∎

We further need the following auxiliary lemma.

LEMMA 3.17. *For $\beta \in (2,3)$ and all $\ell \geq 1, \ell = o(n)$,*

$$\mathbb{E}\left[K_\ell(M_\varepsilon^{(-)}(\sqrt{n}))\right] \leq \mathcal{O}\left(\frac{\ell \varepsilon^{1-\beta}}{n^{(3-\beta)/2}}\right) \mathbb{E}\left[K_{\ell+1}(M_\varepsilon^{(-)}(\sqrt{n}))\right].$$

Proof. Observe that

$$\begin{aligned}\mathbb{E}\left[K_{\ell+1}(M_\varepsilon^{(-)}(\sqrt{n}))\right] &= \binom{n}{\ell+1}\Pr\left[U_{\ell+1} \text{ is clique in } M_\varepsilon^{(-)}(\sqrt{n})\right] \\ &\geq \frac{n-\ell}{\ell+1}\binom{n}{\ell}\Pr\left[U_\ell \text{ is clique in } M_\varepsilon^{(-)}(\sqrt{n})\right]\Omega\left((\sqrt{n}/\varepsilon)^{1-\beta}\right) \\ &= \Omega\left(\frac{n^{(3-\beta)/2}}{\varepsilon^{1-\beta}\ell}\right)\mathbb{E}\left[K_\ell(M_\varepsilon^{(-)}(\sqrt{n}))\right].\end{aligned}$$

Here, in the penultimate step, we used the fact that $U_{\ell+1}$ is a clique if U_ℓ is a clique and the remaining $\ell+1$th vertex has a weight of $\Omega\left((\sqrt{n}/\varepsilon)\right)$. ∎

Combining these statements yields that the number of cliques in the dominant regimes established earlier is self-averaging. That is, restricting the number of cliques to either vertices of very low or of very large weight yields a random variable that concentrates well around its expectation.

LEMMA 3.18. *For all $k \geq 3$ and for any $w > 1$,*

$$\operatorname{Var}\left[K_k(M_\varepsilon^{(+)}(w))\right] = \mathcal{O}(w/\varepsilon)^{2k}\mathbb{E}\left[K_k(M_\varepsilon^{(+)}(w))\right].$$

Moreover,

$$\operatorname{Var}\left[K_k(M_\varepsilon^{(-)}(\sqrt{n}))\right] \leq \sum_{\ell=1}^{k}\mathcal{O}\left(\frac{\varepsilon^{1-\beta}k^2}{n^{(3-\beta)/2}}\right)^{\ell}\mathbb{E}\left[K_k(M_\varepsilon^{(-)}(\sqrt{n}))\right]^2.$$

Proof. As the maximum weight in $M_\varepsilon^{(+)}(w)$ is w/ε, the probability that $k-\ell$ vertices form a clique is $\mathcal{O}(w^2/(\varepsilon^2 n))^{k-\ell}$. Thus, the bound from Lemma 3.16 implies

$$\begin{aligned}\operatorname{Var}\left[K_k(M_\varepsilon^{(+)}(w))\right] &\leq \mathbb{E}\left[K_k(M_\varepsilon^{(+)}(w))\right]\sum_{\ell=1}^{k}\binom{k}{\ell}\binom{n}{k-\ell}\mathcal{O}((w^2/(\varepsilon^2 n))^{k-\ell} \\ &\leq \mathbb{E}\left[K_k(M_\varepsilon^{(+)}(w))\right]\mathcal{O}(w/\varepsilon)^{2k}\sum_{\ell=1}^{k}\binom{k}{\ell} \\ &\leq \mathbb{E}\left[K_k(M_\varepsilon^{(+)}(w))\right]\mathcal{O}(w/\varepsilon)^{2k}\end{aligned}$$

as $\sum_{\ell=1}^{k}\binom{k}{\ell} \leq 2^k$.

For the second part of the statement, note that by Lemma 3.17, we have

$$\mathbb{E}\left[K_{k-\ell}(M_\varepsilon^{(+)}(w))\right] \leq \mathcal{O}\left(\frac{\varepsilon^{1-\beta}k}{n^{(3-\beta)/2}}\right)^{\ell}\mathbb{E}\left[K_k(M_\varepsilon^{(+)}(w))\right].$$

Accordingly, by Lemma 3.16,

$$\operatorname{Var}\left[K_k(M_\varepsilon^{(-)}(\sqrt{n}))\right] \leq \mathbb{E}\left[K_k(M_\varepsilon^{(-)}(\sqrt{n}))\right]^2\sum_{\ell=1}^{k}\mathcal{O}\left(\frac{\varepsilon^{1-\beta}k^2}{n^{(3-\beta)/2}}\right)^{\ell}.$$ ∎

3.4. Bounds on the expected number of cliques. We proceed by turning the results from the previous sections into bounds on the expected number of cliques. Recall that

$$\mathbb{E}\left[K_k\right] = \binom{n}{k}q_k = n^k\Theta(k)^{-k}q_k. \tag{3.3}$$

Based on this relation and our bounds on q_k, we prove the following.

THEOREM 3.19. *Let $G = \mathrm{GIRG}(n, \beta, w_0, d)$ be a standard GIRG with dimensionality $d = o(\log(n))$. Then, we have*

$$\mathbb{E}[K_k] = \begin{cases} n\exp(-\Theta(1)dk)\,\Theta(k)^{-k} + o(1) & \text{if } k < \frac{2}{3-\beta} \text{ or } \beta \geq 3, \\ n^{\frac{k}{2}(3-\beta)}\Theta(k)^{-k} & \text{otherwise,} \end{cases}$$

whereby the statement holds for all (potentially superconstant) $k \geq 3$ if $\beta \neq 3$, and for $k = o(\log(n)/d)$ if $\beta = 3$.

Proof. Observe that Corollaries 3.3 and 3.5 directly imply $\mathbb{E}[K_k] = n^{\frac{k}{2}(3-\beta)}\Theta(k)^{-k}$ for the case $2 < \beta < 3$ and $k > \frac{2}{3-\beta}$. For the other case, note that Corollary 3.5 shows that $\mathbb{E}[K_k] \geq 2^{-dk} n\Theta(k)^{-k}$, and so it only remains to derive an upper bound, i.e., to show that $\mathbb{E}[K_k] \leq n\exp(-\Omega(1)dk)\Theta(k)^{-k} + o(1)$.

To this end, we define the set W_ℓ to be the set of all k-element subsets of vertices among which the minimum weight is at most ℓ, and we denote by $K_k(W_\ell)$ and $K_k(\overline{W_\ell})$ the random variables denoting the number of k-cliques in W_ℓ and $\overline{W_\ell}$, respectively. Clearly, for all ℓ and any $\varepsilon > 0$, we have $\mathbb{E}[K_k] = \mathbb{E}[K_k(W_\ell)] + \mathbb{E}\left[K_k(\overline{W_\ell})\right]$. We fix $\ell = n^{1/2-\varepsilon}$ and start by deriving a bound for $\mathbb{E}[K_k(W_\ell)]$.

Recall that we assume v_1 to be of minimal weight w_1 among U_k and that we denote by $\mathbf{E}_{\text{star}}$ the event that U_k is a star with center v_1. We let $\mathbf{E}^\ell_{\text{star}}$ be the event $\mathbf{E}_{\text{star}} \cap (U_k \in W_\ell)$ and observe that

$$\mathbb{E}[K_k(W_\ell)] \leq n^k \Theta(k)^{-k} \Pr\left[U_k \text{ is clique} \mid \mathbf{E}^\ell_{\text{star}}\right] \Pr\left[\mathbf{E}^\ell_{\text{star}}\right].$$

We note that $\Pr\left[\mathbf{E}^\ell_{\text{star}}\right] \leq \Pr[\mathbf{E}_{\text{star}}] = \Theta(1)^k n^{-(k-1)}$ as established in Corollary 3.3. It thus remains to show $\Pr\left[U_k \text{ is clique} \mid \mathbf{E}^\ell_{\text{star}}\right] \leq \mathcal{O}(1)^k \exp(-\Omega(1)dk)$.

To show this, we apply Theorem 1.6. However, for this, we need to ensure that the weights $w_2, \ldots, w_k$ are at most $w_1 c^d$ for some constant $c > 1$, and that the maximal connection threshold $c^2(\frac{w_1^2}{\tau n})^{1/d}$ is at most $1/4$. As $w_1 \leq \ell = n^{1/2-\varepsilon}$ and $d = o(\log(n))$, the second condition is fulfilled for large enough n and every $\varepsilon > 0$. Regarding the first condition, however, we observe that some of the vertices in U_k might have a weight larger than $w_1 c^d$. Yet, it is very unlikely that this occurs for many vertices of U_k. We split U_k into the two parts $U_k^{(1)}$ and $\overline{U_k^{(1)}}$, such that $U_k^{(1)}$ is the set of vertices in U_k with weight at most $w_1 c^d$ and $\overline{U_k^{(1)}} = U_k \setminus U_k^{(1)}$. We set $s = \max\{3, \lfloor k/2 \rfloor\}$ and bound

$$\begin{aligned} &\Pr\left[U_k \text{ is clique} \mid \mathbf{E}^\ell_{\text{star}}\right] \\ &\quad\leq \Pr\left[U_k^{(1)} \text{ is clique} \mid \mathbf{E}^\ell_{\text{star}} \cap (|U_k^{(1)}| \geq s)\right] \Pr\left[|U_k^{(1)}| \geq s \mid \mathbf{E}^\ell_{\text{star}}\right] \\ &\qquad+ \Pr\left[U_k^{(1)} \text{ is clique} \mid \mathbf{E}^\ell_{\text{star}} \cap (|U_k^{(1)}| < s)\right] \Pr\left[|U_k^{(1)}| < s \mid \mathbf{E}^\ell_{\text{star}}\right]. \end{aligned}$$

As probabilities are at most one, we may bound

$$\begin{aligned} &\Pr\left[U_k \text{ is clique} \mid \mathbf{E}^\ell_{\text{star}}\right] \\ &\quad\leq \Pr\left[U_k^{(1)} \text{ is clique} \mid \mathbf{E}^\ell_{\text{star}} \cap |U_k^{(1)}| \geq s\right] + \Pr\left[|U_k^{(1)}| < s \mid \mathbf{E}^\ell_{\text{star}}\right]. \end{aligned} \tag{3.4}$$

By Theorem 1.6, the former probability is $\exp(-\Theta(1)ds)$ as there is a constant $\xi > 0$ such that $s \geq \xi k$ for all $k \geq 3$. We thus proceed by bounding the latter term in the above expression.

We observe that, conditioned on $\mathbf{E}^\ell_{\text{star}}$, the weights $w_2, \ldots, w_k$ are i.i.d. random variables, and so $|\overline{U_k^{(1)}}|$ is distributed according to the binomial distribution $\mathrm{Bin}(k-1, p)$, where $p = \Pr[w_2 \geq w_1 c^d \mid \mathbf{E}^\ell_{\text{star}}]$.

CLAIM 3.20. *There is a constant $b > 1$ such that $p \leq b^{-d+\mathcal{O}(1)}$.*

We defer the proof of this claim and proceed by showing how it helps in our main proof. Due to the binomial nature of $|\overline{U_k^{(1)}}|$, we get

$$\begin{aligned}
\Pr\left[|U_k^{(1)}| < s \mid \mathbf{E}_{\text{star}}^{\ell}\right] &= \Pr\left[|\overline{U_k^{(1)}}| > k - s \mid \mathbf{E}_{\text{star}}^{\ell}\right] \\
&= \sum_{i=k-s+1}^{k-1} \binom{k-1}{i} p^i (1-p)^{k-1-i} \\
&\leq \sum_{i=k-s+1}^{k-1} \binom{k-1}{i} b^{-di+\mathcal{O}(1)i} \\
&\leq b^{-d(k-s+1)+\mathcal{O}(1)k} \sum_{i=k-s+1}^{k-1} \binom{k-1}{i} \\
&\leq b^{-d(k-s+1)+\mathcal{O}(1)k},
\end{aligned}$$

where the last step is due to the fact that $\sum_{i=k-s+1}^{k-1} \binom{k-1}{i} \leq 2^k$. Now, observe that there is a constant $\xi > 0$ such that $k - s + 1 \geq \xi k$ for all $k \geq 3$. Thus, the above expression is $\exp(-\Omega(1)dk)\mathcal{O}(1)^k$, and so $\mathbb{E}[K_k(W_\ell)] = n \exp(-\Omega(1)dk)\mathcal{O}(k)^{-k}$. It remains to prove Claim 3.20.

Proof of Claim 3.20. Fix a vertex $u \in U_k \setminus \{v_1\}$ and observe that

$$p = \Pr\left[w_u \geq w_1 c^d \mid (u \sim v_1) \cap (U_k \in W_\ell)\right]$$

because the weight of each $u \in U_k \setminus \{v_1\}$ is an i.i.d. random variable conditioned on $\mathbf{E}_{\text{star}}^{\ell} = \mathbf{E}_{\text{star}} \cap U_k \in W_\ell$ and is, in particular, independent of whether other vertices in U_k are adjacent to v_1. Now, assume that $w_1 = x \leq \ell = n^{1/2-\varepsilon}$ and observe that

$$\Pr\left[w_u \geq w_1 c^d \mid (u \sim v_1) \cap (w_1 = x)\right] = \frac{\Pr\left[(w_u \geq w_1 c^d) \cap (u \sim v_1) \mid w_1 = x\right]}{\Pr\left[u \sim v_1 \mid w_1 = x\right]}. \tag{3.5}$$

Note that for all $x \leq n^{1/2-\varepsilon}$,

$$\begin{aligned}
\Pr\left[u \sim v_1 \mid w_1 = x\right] \geq \int_x^{\frac{n}{\lambda x}} \frac{\lambda x w_u}{n} \rho(w_u) \mathrm{d}w_u &= \frac{\lambda x}{n} \Theta(1) \left(x^{2-\beta} - \left(\frac{n}{\lambda x}\right)^{2-\beta}\right) \\
&= \frac{\lambda x^{3-\beta}}{n} \Theta(1) \left(1 - \left(\frac{n}{\lambda x^2}\right)^{2-\beta}\right) \\
&\geq (1 - o(1)) \frac{\lambda x^{3-\beta}}{n} \Theta(1),
\end{aligned}$$

where the last step follows from the fact that $x \leq n^{1/2-\varepsilon}$. On the other hand, we get

$$\begin{aligned}
\Pr\left[(w_u \geq w_1 c^d) \cap (u \sim v_1) \mid w_1 = x\right] \leq \int_{xc^d}^{\infty} \frac{\lambda x w_u}{n} \rho(w_u) \mathrm{d}w_u &= \frac{\lambda x}{n} \Theta(1) (xc^d)^{2-\beta} \\
&= \frac{\lambda x^{3-\beta}}{n} \Theta(1) c^{(2-\beta)d}.
\end{aligned}$$

Thus, we get from (3.5) that

$$\Pr\left[w_u \geq w_1 c^d \mid (u \sim v_1) \cap (w_1 = x)\right] \leq (1 + o(1))\Theta(1) c^{(2-\beta)d} = b^{-d+\mathcal{O}(1)},$$

if we choose $b = c^{\beta-2}$, which is greater than 1, since $c > 1$ and $\beta > 2$. ∎

It remains to bound $\mathbb{E}\left[K_k(\overline{W_\ell})\right]$. We observe that

$$\mathbb{E}\left[K_k(\overline{W_\ell})\right] \leq n^k \Theta(k)^{-k}\Big(\Pr\left[(U_k \text{ is clique}) \cap (\ell \leq w_1 \leq \sqrt{n/\lambda})\right] + \Pr\left[w_1 \geq \sqrt{n/\lambda}\right]\Big).$$

As in the proof of Corollary 3.3, we have

$$\Pr\left[w_1 \geq \sqrt{n/\lambda}\right] \leq \Theta(1)^k n^{\frac{k}{2}(3-\beta)-k},$$

and from Lemma 3.2, we get

$$\begin{aligned}\Pr\left[(U_k \text{ is clique}) \cap (\ell \leq w_1 \leq \sqrt{n/\lambda})\right] &\leq \Theta(1)^k n^{-(k-1)} n^{(1/2-\varepsilon)(k(3-\beta)-2)} \\ &\leq \Theta(1)^k n^{-(k-1)-\delta}\end{aligned}$$

for some $\delta \geq (1/2-\varepsilon)(3(\beta-3)+2) > 0$ since we have $k \geq 3$ and $k(\beta-3)+2>0$ for $k < \frac{2}{3-\beta}$ or $\beta \geq 3$. This implies

$$\mathbb{E}\left[K_k(\overline{W_\ell})\right] \leq n^{1-\delta}\Theta(k)^{-k}.$$

Now, we distinguish three cases. In the first case, $2<\beta<3$ and $k<\frac{2}{3-\beta}$. Here, as k is at most a constant, $n^{-\delta}$ is asymptotically smaller than $\exp(-\Theta(1)dk)$ (recall that $d = o(\log(n))$), which finishes the proof for this case. If $\beta = 3$, recall that we only have to show the statement for $k = o(\log(n)/d)$ and under this assumption, again, $n^{-\delta}$ is asymptotically smaller than $\exp(-\Theta(1)dk)$. For the case $\beta>3$, recall that δ is at least a constant since $\delta \geq (1/2-\varepsilon)(3(\beta-3)+2)$. As $3(\beta-3)+2$ is strictly greater than 2, we may choose $\varepsilon>0$ sufficiently small to get $\delta>1$. Then, $\mathbb{E}\left[K_k(\overline{W_\ell})\right] = o(1)$ for all $k \geq 3$, which finishes the proof. ∎

Summarizing our results on $\mathbb{E}[K_k]$, we see that in the case $2<\beta<3$, there is a phase transition at $k = \frac{2}{3-\beta}$, which was previously observed (in the case of constant d) by Michielan and Stegehuis [35] and also by Bläsius et al. [7] for HRGs. Regarding the influence of d, we find that the number of cliques in the regime $k<\frac{2}{3-\beta}$ or $\beta>3$ decreases exponentially in dk.

3.5. Bounds on the clique number. Now, we turn our bounds on the expected number of cliques into bounds on the clique number. The results of this section constitute the first two columns of Table 3.

Upper bounds. We start with the upper bounds stated in the first row of Table 3.

THEOREM 3.21. *Let $G = \mathrm{GIRG}(n,\beta,w_0,d)$ be a standard GIRG with $\beta<3$. Then, $\omega(G) = \mathcal{O}(n^{(3-\beta)/2})$ a.a.s.*

Proof. We may upper bound the clique number with Markov's inequality, which tells us that

$$\Pr[\omega(G) \geq k] = \Pr[K_k \geq 1] \leq \mathbb{E}[K_k].$$

The goal now is to choose k such that $\mathbb{E}[K_k] \leq n^{-\epsilon}$ for some $\epsilon>0$, as then there is no clique larger than k a.a.s. For $\beta<3$, which is a prerequisite of this theorem, we have $\mathbb{E}[K_k] \leq n^{\frac{k}{2}(3-\beta)}(c_1 k)^{-k}$ for some constant $c_1>1$ and large enough k, according to Theorem 3.19. If we set $k = n^{\frac{1}{2}(3-\beta)}$, we get

$$\begin{aligned}\mathbb{E}[K_k] &\leq n^{\frac{1}{2}(3-\beta)n^{\frac{1}{2}(3-\beta)}}(c_1 n^{\frac{1}{2}(3-\beta)})^{-n^{\frac{1}{2}(3-\beta)}} \\ &= c_1^{-n^{\frac{1}{2}(3-\beta)}},\end{aligned}$$

which is asymptotically smaller than $n^{-\epsilon}$ because

$$c_1^{-n^{\frac{1}{2}(3-\beta)}} \leq n^{-\epsilon} \Leftrightarrow \exp\left(-\log(c_1)n^{\frac{1}{2}(3-\beta)}\right) \leq \exp\left(-\epsilon\log(n)\right),$$

which holds for $\beta < 3$. ∎

Regarding our contribution in the second row of Table 3, we obtain the following theorem.

THEOREM 3.22. *Let $G = \mathrm{GIRG}(n, \beta, w_0, d)$ be a standard GIRG with $\beta = 3$ and $d = \mathcal{O}(\log\log(n))$. Then, $\omega(G) = \mathcal{O}(\log(n)/\log\log(n))$ a.a.s.*

Proof. Analogous to the proof of Theorem 3.21, we upper bound the clique number using Markov's inequality, i.e., $\Pr[\omega(G) \geq k] \leq \mathbb{E}[K_k]$, and choose k such that $\mathbb{E}[K_k] \leq n^{-\epsilon}$ for some $\epsilon > 0$. Now for $\beta = 3$, we have $\mathbb{E}[K_k] \leq n \cdot (c_1 k)^{-k}$ for some constant $c_1 > 1$ and sufficiently large k, which follows by Corollary 3.3 and Lemma 2.1. It remains to determine the value of k for which our desired upper bound on $\mathbb{E}[K_k]$ is valid, which can be done by solving the following equation for k:

$$\begin{aligned} & n \cdot (c_1 k)^{-k} = n^{-\epsilon} \\ \Leftrightarrow \qquad & (c_1 k)^{-k} = n^{-1-\epsilon}. \end{aligned}$$

This yields

$$k = \frac{1}{c_1}\exp\left(\mathcal{W}(c_1\log(n^{1+\epsilon}))\right),$$

where $\mathcal{W}$ is the Lambert $\mathcal{W}$ function defined by the identity $\mathcal{W}(z)e^{\mathcal{W}(z)} = z$. With this choice of k, indeed

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

In order to simplify the expression for k, note that for growing z, we have that $\mathcal{W}(z) = \log(z) - \log\log(z) + o(1)$ and thus

$$\begin{aligned} k &= \frac{1}{c_1}\exp\left(\log(c_1(1+\epsilon)\log(n)) - \log\log(c_1(1+\epsilon)\ln(n)) + o(1)\right) \\ &= (1+o(1))\frac{(1+\epsilon)\log(n)}{\log(c_1(1+\epsilon)\log(n))} = \mathcal{O}\left(\frac{\log(n)}{\log\log(n)}\right). \end{aligned}$$ ∎

Finally, using a similar argumentation, we obtain the following for the upper bounds we contribute to the last row of Table 3.

THEOREM 3.23. *Let $G = \mathrm{GIRG}(n, \beta, w_0, d)$ be a standard GIRG with $\beta > 3$ and $d = o(\log(n))$. Then, $\omega(G) = \mathcal{O}(\log(n)/(d + \log\log(n)))$ a.a.s.*

Proof. The proof is analogous to the one of Theorem 3.22. However, if $\beta > 3$, we may use the stronger upper bound from Theorem 3.19, i.e.,

$$\mathbb{E}[K_k] = n\exp(-\Theta(1)dk)\Theta(k)^{-k} + o(1).$$

That is, there are constants $c_1, c_2 > 0$ such that $\mathbb{E}[K_k] = n\left(c_1\exp(c_2 d)k\right)^{-k} + o(1)$. Just like before, setting

$$k = \frac{1}{c_1 \exp(c_2 d)} \exp\left(\mathcal{W}(c_1 \exp(c_2 d) \log(n^{1+\epsilon}))\right)$$

yields $\mathbb{E}[K_k] \leq n^{-\epsilon} + o(1)$. Using the asymptotic expansion of the Lambert $\mathcal{W}$ function, we then obtain

$$\begin{aligned} k &= (1+o(1)) \frac{(1+\epsilon)\log(n)}{\log(c_1 \exp(c_2 d)(1+\epsilon)\log(n))} \\ &= (1+o(1)) \frac{(1+\epsilon)\log(n)}{\log(c_1(1+\epsilon)) + c_2 d + \log\log(n)} \\ &= \mathcal{O}\left(\frac{\log(n)}{d + \log\log(n)}\right). \end{aligned}$$ ∎

Lower bounds. To get matching lower bounds, we distinguish once more the cases $\beta < 3$ and $\beta \geq 3$.

THEOREM 3.24. *Let $G = \mathrm{GIRG}(n, \beta, w_0, d)$ be a standard GIRG with $\beta < 3$. Then, $\omega(G) = \Omega(n^{(3-\beta)/2})$ a.a.s.*

Proof. We show that there are $\Omega(n^{\frac{1}{2}(3-\beta)})$ vertices with weight at least $\sqrt{n/\lambda}$ w.h.p. As all these vertices are connected with probability 1, this implies the existence of an equally sized clique.

Because the weight of each vertex is sampled independently, the number of vertices with weight above $\sqrt{n/\lambda}$, denoted by X, is the sum of n independent Bernoulli random variables with success probability

$$p = \left(\frac{n}{\lambda w_0^2}\right)^{\frac{1}{2}(1-\beta)} = \Theta\left(n^{\frac{1}{2}(1-\beta)}\right),$$

which we can infer from (3.1). Therefore, we get

$$\mathbb{E}[X] = np = \Theta\left(n^{\frac{1}{2}(3-\beta)}\right),$$

and by a Chernoff–Hoeffding bound (Theorem 2.2), we get $X \geq \mathbb{E}[X]/2$ w.h.p., which proves our lower bound. ∎

For the case $\beta \geq 3$, we use the concentration bounds obtained in the previous section applied to a subgraph of G of bounded weight.

THEOREM 3.25. *Let $G = \mathrm{GIRG}(n, \beta, w_0, d)$ be a standard GIRG with $\beta \geq 3$ and $d = o(\log(n))$. Then, $\omega(G) = \Omega(\log(n)/(d + \log\log(n)))$ a.a.s.*

Proof. Let $w > w_0$ be some fixed weight, and let $M = [w_0, w]$. Furthermore, by Lemma 3.4 there is a constant $c_1 > 0$ such that

$$\mathbb{E}[K_k(M)] \geq n\left(c_1 2^d k\right)^{-k}.$$

Now, setting

$$k = \frac{1}{2^d c_1} \exp\left(\mathcal{W}(2^d c_1 \log(n^{1-\varepsilon}))\right) = \Theta\left(\frac{\log(n)}{d + \log\log(n)}\right)$$

yields $\mathbb{E}[K_k(M)] \geq n^{\varepsilon}$. By Lemma 3.18 and as w is constant,

$$\mathrm{Var}[K_k(M)] \leq \mathbb{E}[K_k(M)]\, \mathcal{O}(1)^{2k}.$$

Thus, the inequality of Chebyshev yields

$$\Pr\left[|K_k(M) - \mathbb{E}[K_k(M)]| \geq \frac{1}{2}\mathbb{E}[K_k(G_{\leq w_c})]\right] \leq \frac{\mathrm{Var}[K_k(M)]}{\frac{1}{4}\mathbb{E}[K_k(M)]^2} = \frac{\mathcal{O}(1)^{2k}}{\mathbb{E}[K_k(M)]}.$$

As $k = \mathcal{O}(\log(n)/\log\log(n))$, we have $\mathcal{O}(1)^{2k} = n^{o(1)}$, and so the above term is at most $n^{o(1)-\varepsilon} = o(1)$. Accordingly, $\Pr[K_k(M) \geq 1] = 1 - o(1)$, as desired. □

4. Cliques in the high-dimensional regime. Now, we turn to the high-dimensional regime, where d grows faster than $\log(n)$. We shall see that, for constant k and $d = \omega(\log^2(n))$, the probability that U_k is a clique only differs from its counterpart in the IRG model by a factor of $(1 \pm o(1))$. However, as it turns out, the asymptotic behavior of cliques in the case $2 < \beta < 3$ is already the same as in the IRG model if $d = \omega(\log(n))$. For $\beta \geq 3$, we show that the number of triangles in the geometric case remains significantly larger than in the IRG model as long as $d = \log^{3/2}(n)$.

4.1. Bounding the clique probability for fixed weights. We consider the conditional probability that a set $U_k = \{v_1, \ldots, v_k\}$ of k independent random vertices with given weights $w_1, \ldots, w_k$ forms a clique. We derive bounds on this probability under the assumption that L_∞-norm is used, which afterwards allows bounding the expected number of cliques and the clique number. In fact, instead of only bounding the probability that U_k forms a clique, we bound the probability of the more general event that an arbitrary set of edges $\mathcal{A}$ is formed among the vertices of U_k. We denote by $E(U_k)$ the random variable representing the set of edges between the vertices in U_k and proceed by developing bounds on the probability of the event $E(U_k) \supseteq \mathcal{A}$.

The main difference from our previous bounds is that the connection threshold proportional to $(w_u w_v/n)^{1/d}$ now grows with n instead of shrinking, even for constant w_u, w_v. This requires us to pay closer attention to the topology of the torus. That is, we have to take into account that a single dimension of the torus is in fact a circle with a circumference of 1.

The bounds are formalized in Theorem 1.5, which we restate for the sake of readability.

THEOREM 1.5. *Let G be a standard GIRG and let $k \geq 3$ be a constant. Furthermore, let $U_k = \{v_1, \ldots, v_k\}$ be a set of vertices chosen uniformly at random and let $\{\kappa\}^{(k)} = \{\kappa_{ij} \mid 1 \leq i, j \leq k\}$ describe the pairwise product of weights of the vertices in U_k. Let $E(U_k)$ denote the (random) set of edges formed among the vertices in U_k. Then, for any set of edges $\mathcal{A} \subseteq \binom{U_k}{2}$,*

$$\Pr\left[E(U_k) \supseteq \mathcal{A} \mid \{\kappa\}^{(k)}\right] = \begin{cases} (1 \pm o(1)) \prod_{\{i,j\} \in \mathcal{A}} \frac{\kappa_{ij}}{n} & \text{if } d = \omega(\log^2(n)), \\ (1 \pm o(1)) \prod_{\{i,j\} \in \mathcal{A}} \left(\frac{\kappa_{ij}}{n}\right)^{1 \mp \mathcal{O}\left(\frac{\log(n)}{d}\right)} & \text{if } d = \omega(\log(n)). \end{cases}$$

This illustrates that the probability that U_k is a clique is at most $1 + o(1)$ times its counterpart in the IRG model if $d = \omega\left(\log^2(n)\right)$. For $d = \omega(\log(n))$, we get that these two probabilities only differ by a factor of $(1 + o(1))n^{o(1)}$, which is not much compared to the case $d = o(\log(n))$, where this factor is at least on the order of $n^{\binom{k-1}{2} - \frac{d(k-1)}{\log(n)}}$ among nodes of constant weight.

Before giving the proof, we derive some lemmas that make certain arguments easier to follow. We start with an upper bound on the probability that the set $U_k = \{v_1, \ldots, v_k\}$ forms a clique. Before deriving our bound, we need the following auxiliary lemma.

LEMMA 4.1. *Let $\ell \in \mathbb{N}$, $\ell \geq 1$. There is a constant $x_0 < 1$ such that for all $x_0 \leq x \leq 1$, we have*

$$x^{\ell+1} \leq \ell x - \ell + 1.$$

Proof. We substitute $x = 1 - y$ and instead show that there is some $y_0 > 0$ such that for all $0 \leq y \leq y_0$,

$$(1-y)^{\ell+1} \leq 1 - \ell y.$$

We get from a Taylor series that there is a constant $c \geq 0$ such that for all $0 \leq y \leq 1$,

$$(1-y)^{\ell+1} \leq 1 - (\ell+1)y + cy^2 = 1 - \ell y + cy^2 - y.$$

Now, for all $0 \leq y \leq 1/c$, the term $cy^2 - y$ is negative, and our statement follows. $\square$

In the remainder of this section, we frequently analyze events occurring in a single fixed dimension on the torus and use the following notation. Recall that a single dimension of the torus is a circle of circumference 1, which we denote by $\mathbb{S}^1$. We define the set of points that are within a distance of at most r around a fixed point $\mathbf{x}$ on this circle as $A(r, \mathbf{x})$, and we denote by $\overline{A}(r, \mathbf{x})$ the complement of $A(r, \mathbf{x})$, i.e., the set $\mathbb{S}^1 \setminus A(r, \mathbf{x})$. Observe that $A(r, \mathbf{x})$ and $\overline{A}(r, \mathbf{x})$ are coherent circular arcs. Assume that the position of v_i in our fixed dimension is $\mathbf{x}_{v_i}$. For any pair of vertices v_i, v_j, we define the sets $A_{ij} := A(t_{v_i v_j}, \mathbf{x}_{v_i})$ and $\overline{A}_{ij} = \overline{A}(t_{v_i v_j}, \mathbf{x}_{v_i})$. We further define $A_i = A(t_0, \mathbf{x}_{v_i})$ and $\overline{A}_i = \overline{A}(t_0, \mathbf{x}_{v_i})$, whereby we note that $\overline{A}_{ij} \subseteq \overline{A}_i$ for all i, j because t_0 is the minimal connection threshold.

In the following, we derive upper and lower bounds on the probability that U_k is a clique.

THEOREM 4.2 (upper bound). *Let $G = \mathrm{GIRG}(n, \beta, w_0, d)$ be a standard GIRG with $d = \omega(\log(n))$, and let $U_k = \{v_1, \ldots, v_k\}$ be a set of k random vertices. For any constant $k \in \mathbb{N}_{\geq 3}$, any $\mathcal{A} \subseteq \binom{U_k}{2}$, and sufficiently large n, we have*

$$\Pr\left[E(U_k) \supseteq \mathcal{A} \mid \{\kappa\}^{(k)}\right] \leq \left(1 - \left(\frac{\kappa_0}{n}\right)^{r/d} \sum_{\{i,j\} \in \mathcal{A}} \left(1 - \left(\frac{\kappa_{ij}}{n}\right)^{1/d}\right)\right)^d,$$

where $r = (3(k-2)+1)(k-1)$.

Proof. In the following, we denote by t_0 the minimal connection threshold of any two vertices, i.e., $t_0 = (\frac{w_0^2}{\tau n})^{1/d}$. Note that $2t_{uv} = (\frac{\kappa_{uv}}{n})^{1/d}$ for any pair of vertices u, v. We again consider only one fixed dimension of the torus as they are all independent due to our use of L_∞-norm.

To get an upper bound on the desired probability, we derive a lower bound on $1 - p$. We define the event $\mathbf{E}_i$ as the event that v_i falls into $\overline{A}_{ji}$ for some $\{i, j\} \in \mathcal{A}$ with $i < j$, and the event $\mathbf{E}_i^{\mathrm{dis}}$ as the event that $v_i \notin \bigcup_{j=1}^{i-1} \overline{A}_j$ and the sets $\overline{A}_j$ are disjoint for all $1 \leq j \leq i$. Note that $\mathbf{E}_i$ and $\mathbf{E}_i^{\mathrm{dis}}$ are disjoint as $\overline{A}_{ji} \subseteq \overline{A}_j$. Then

$$\begin{aligned} 1 - p &\geq \Pr[\mathbf{E}_2] + \Pr\left[\mathbf{E}_3 \cap \mathbf{E}_2^{\mathrm{dis}}\right] + \Pr\left[\mathbf{E}_4 \cap \mathbf{E}_3^{\mathrm{dis}} \cap \mathbf{E}_2^{\mathrm{dis}}\right] + \ldots \\ &= \sum_{i=2}^{k} \Pr\left[\mathbf{E}_i \,\middle|\, \bigcap_{j=1}^{i-1} \mathbf{E}_j^{\mathrm{dis}}\right] \Pr\left[\bigcap_{j=1}^{i-1} \mathbf{E}_j^{\mathrm{dis}}\right]. \end{aligned} \tag{4.1}$$

Note that this is a valid bound because all the events we sum over are disjoint.

Now, let $\mathcal{A}_i^- = \{\{i,j\} \in \mathcal{A} \mid j \le i\}$ be the set of edges from vertex i to a lower-indexed vertex. If we condition on $\bigcap_{j=1}^{i-1} \mathbf{E}_j^{\mathrm{dis}}$, then the probability of $\mathbf{E}_i$ is simply

$$\Pr\left[\mathbf{E}_i \,\middle|\, \bigcap_{j=1}^{i-1} \mathbf{E}_j^{\mathrm{dis}}\right] = \sum_{\{i,j\}\in\mathcal{A}_i^-} (1-2t_{ij}) = \sum_{\{i,j\}\in\mathcal{A}_i^-} \left(1 - \left(\frac{\kappa_{ij}}{n}\right)^{1/d}\right)$$

because all the sets $\overline{A}_{ji}$ are disjoint. It remains to bound $\Pr[\bigcap_{j=1}^{i-1} \mathbf{E}_j^{\mathrm{dis}}]$. We obtain

$$\Pr\left[\bigcap_{j=1}^{i-1} \mathbf{E}_j^{\mathrm{dis}}\right] = \prod_{j=1}^{i-1} \Pr\left[\mathbf{E}_j^{\mathrm{dis}} \,\middle|\, \bigcap_{\ell=1}^{j-1} \mathbf{E}_\ell^{\mathrm{dis}}\right].$$

Now, the probability $\Pr[\mathbf{E}_j^{\mathrm{dis}} \mid \bigcap_{\ell=1}^{j-1} \mathbf{E}_\ell^{\mathrm{dis}}]$ is equal to the probability that v_j is placed outside of $\overline{A}_\ell$ for all $1 \le \ell < j$ while, at the same time, $\overline{A}_\ell \cap \overline{A}_j = \emptyset$. If we consider one fixed set $\overline{A}_\ell$, we note that this requires v_j to be of distance at least $1-2t_0$ from v_ℓ as, otherwise, $\overline{A}_\ell$ and $\overline{A}_j$ overlap. Hence, we may define a "forbidden" region around v_ℓ which includes $\overline{A}_\ell$ and all points within distance $1-2t_0$ of v_ℓ. This region has volume $3(1-2t_0)$ and so the probability that v_j falls outside the forbidden region is at least $1-3(1-2t_0)$. We refer the reader to Figure 1 for an illustration. Now considering the forbidden region of all v_ℓ with $1 \le \ell < j$, the combined volume of these forbidden regions is at most $3(j-1)(1-2t_0)$, and hence

$$\begin{aligned}\Pr\left[\bigcap_{j=1}^{i-1} \mathbf{E}_j^{\mathrm{dis}}\right] = \prod_{j=1}^{i-1} \Pr\left[\mathbf{E}_j^{\mathrm{dis}} \,\middle|\, \bigcap_{\ell=1}^{j-1} \mathbf{E}_\ell^{\mathrm{dis}}\right] &\ge \prod_{j=1}^{i-1} (1-3(j-1)(1-2t_0)) \\ &\ge (1-3(i-2)(1-2t_0))^{i-1}.\end{aligned}$$

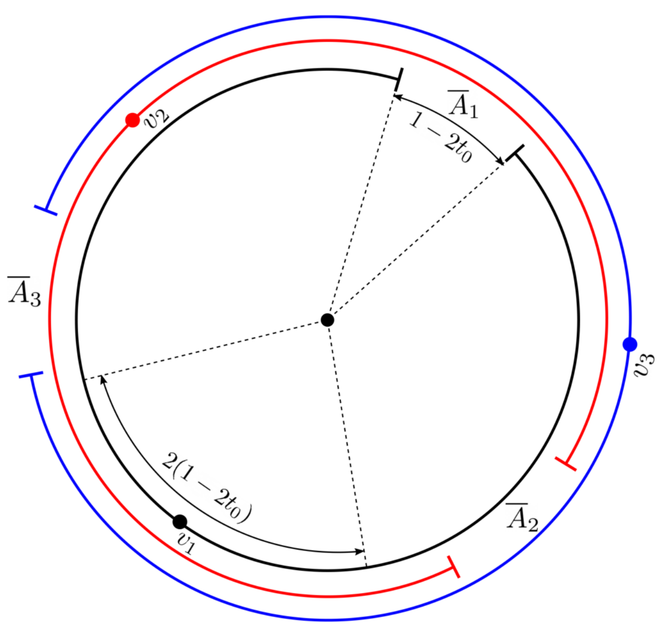

FIG. 1. *Illustration for the proof of Theorem 4.2. The colored circles represent the sets A_i for v_1, v_2, v_3. The probability that v_2 is placed in the region indicated by the black circular arc such that $\overline{A}_1 \cap \overline{A}_2 = \emptyset$ is at least $1-3(1-2t_0)$ as indicated by the black arrows. (Color available online.)*

We get from (4.1)

$$\begin{aligned}
1-p &\geq \sum_{i=2}^{k} \Pr\left[\mathbf{E}_i \,\middle|\, \bigcap_{j=1}^{i-1} \mathbf{E}_j^{\mathrm{dis}}\right] \Pr\left[\bigcap_{j=1}^{i-1} \mathbf{E}_j^{\mathrm{dis}}\right] \\
&\geq \sum_{i=2}^{k} (1-3(i-2)(1-2t_0))^{i-1} \sum_{\{i,j\}\in\mathcal{A}_i^-} \left(1-\left(\frac{\kappa_{ij}}{n}\right)^{1/d}\right) \\
&\geq (1-3(k-2)(1-2t_0))^{k-1} \sum_{\{i,j\}\in\mathcal{A}} \left(1-\left(\frac{\kappa_{ij}}{n}\right)^{1/d}\right) \\
&= \left(3(k-2)\left(\frac{\kappa_0}{n}\right)^{1/d} - 3(k-2)+1\right)^{k-1} \sum_{\{i,j\}\in\mathcal{A}} \left(1-\left(\frac{\kappa_{ij}}{n}\right)^{1/d}\right).
\end{aligned}$$

It now remains to show

$$3(k-2)\left(\frac{\kappa_0}{n}\right)^{1/d} - 3(k-2)+1 \geq \left(\frac{\kappa_0}{n}\right)^{\frac{3(k-2)+1}{d}}$$

for sufficiently large n. Recalling that $\left(\frac{\kappa_0}{n}\right)^{1/d}$ tends to 1 as n grows, this is equivalent to showing that there is some $x_0 < 1$ such that for all $x_0 \leq x \leq 1$ and $\ell = 3(k-2)$, we have

$$\ell x - \ell + 1 \geq x^{\ell+1}.$$

This follows by Lemma 4.1, and the proof is finished. ∎

THEOREM 4.3 (lower bound). *Let $G = \mathrm{GIRG}(n,\beta,w_0,d)$ be a standard GIRG with $d = \omega(\log(n))$, and let $U_k = \{v_1,\ldots,v_k\}$ be a set of k random vertices. Then, for every $\mathcal{A} \subseteq \binom{U_k}{2}$,*

$$\Pr\left[E(U_k) \supseteq \mathcal{A} \mid \{\kappa\}^{(k)}\right] \geq \prod_{i=1}^{k} \left(1 - \sum_{\{i,j\}\in\mathcal{A}_i^-} \left(1-\left(\frac{\kappa_{ij}}{n}\right)^{1/d}\right)\right)^d,$$

where $\mathcal{A}_i^- = \{\{i,j\} \in \mathcal{A} \mid j \leq i\}$ is the set of edges in $\mathcal{A}$ between vertex i and a previous vertex.

Proof. We sample the position of $v_2,\ldots,v_k$ one after another and again only consider the probability p that $E(U_k) \supset \mathcal{A}$ in one fixed dimension. When the position of v_i is sampled, the probability that v_i falls into one of A_{ji} for $\{i,j\} \in \mathcal{A}_i$ is at least

$$1 - \sum_{\{i,j\}\in\mathcal{A}_i^-} \left(1-\left(\frac{\kappa_{ij}}{n}\right)^{1/d}\right),$$

where equality holds if the first $i-1$ vertices are placed such that all the sets $\overline{A}_{ji}$ for $1 \leq j < i$ are disjoint. Accordingly, we have

$$p \geq \prod_{i=1}^{k} \left(1 - \sum_{\{i,j\}\in\mathcal{A}_i^-} \left(1-\left(\frac{\kappa_{ij}}{n}\right)^{1/d}\right)\right).$$

To have $E(U_k) \supset \mathcal{A}$, it is sufficient that this event occurs in all dimensions, so the final bound is p^d. ∎

To see how these bounds behave as $n \to \infty$, we need the following lemma.

LEMMA 4.4. *Let Ψ, d, ℓ be nonnegative functions of n. Assume that $\frac{-\ell \ln(\Psi)}{d} = o(1)$ and $\Psi < 1$ for all n. Then,*

$$\frac{-\ell \ln(\Psi)}{d} - e\left(\frac{-\ell \ln(\Psi)}{d}\right)^2 \leq 1 - \Psi^{\ell/d} \leq \frac{-\ell \ln(\Psi)}{d}.$$

Proof. Observe that

$$1 - \Psi^{\ell/d} = 1 - \exp\left(\frac{\ell \ln(\Psi)}{d}\right).$$

For the upper bound, we use the well-known fact that, for all x, we have $\exp(x) \geq 1 + x$, which directly implies our statement. For the lower bound, we use the Taylor series expansion of exp and bound

$$\begin{aligned}
1 - \Psi^{\ell/d} = 1 - \exp\left(\frac{\ell \ln(\Psi)}{d}\right) &= -\sum_{i=1}^{\infty} \frac{1}{i!}\left(\frac{\ell \ln(\Psi)}{d}\right)^i \\
&= \sum_{i=1}^{\infty} \frac{(-1)^{i-1}}{i!}\left(\frac{-\ell \ln(\Psi)}{d}\right)^i \\
&= \frac{-\ell \ln(\Psi)}{d} - \left(\frac{-\ell \ln(\Psi)}{d}\right)^2 \sum_{i=2}^{\infty} \frac{(-1)^i}{i!}\left(\frac{-\ell \ln(\Psi)}{d}\right)^{i-2}. \qquad (4.2)
\end{aligned}$$

Because $\frac{-\ell \ln(\Psi)}{d} = o(1)$, we get that $(\frac{-\ell \ln(\Psi)}{d})^{i-2} \leq 1$ for all $i \geq 2$ and sufficiently large n. Assuming that all the terms in the above sum are positive yields

$$\begin{aligned}
1 - \Psi^{\ell/d} &\geq \frac{-\ell \ln(\Psi)}{d} - \left(\frac{-\ell \ln(\Psi)}{d}\right)^2 \sum_{i=0}^{\infty} \frac{1}{i!} \\
&= \frac{-\ell \ln(\Psi)}{d} - e\left(\frac{-\ell \ln(\Psi)}{d}\right)^2,
\end{aligned}$$

as desired. ∎

LEMMA 4.5. *Let*

$$P := \left(1 - \left(\frac{\kappa_0}{n}\right)^{r/d} \sum_{\{i,j\} \in \mathcal{A}} \left(1 - \left(\frac{\kappa_{ij}}{n}\right)^{1/d}\right)\right)^d.$$

Then there is a constant $\delta \geq 0$ such that

$$\ln(P) \leq \sum_{\{i,j\} \in \mathcal{A}} \ln\left(\frac{\kappa_{ij}}{n}\right)\left(1 - \frac{\delta \ln(n)}{d}\right).$$

Proof. We have

$$\ln(P) = d \ln\left(1 - \left(\frac{\kappa_0}{n}\right)^{r/d} \sum_{\{i,j\} \in \mathcal{A}} \left(1 - \left(\frac{\kappa_{ij}}{n}\right)^{1/d}\right)\right).$$

We get from a Taylor series that

$$\ln(x) = \sum_{i=1}^{\infty} (-1)^{i-1} \frac{(x-1)^i}{i} = -\sum_{i=1}^{\infty} \frac{(1-x)^i}{i}$$

and thus

$$\begin{aligned}\ln(P) &= -d \sum_{i=1}^{\infty} \frac{1}{i} \left(\left(\frac{\kappa_0}{n}\right)^{r/d} \sum_{\{i,j\} \in \mathcal{A}} \left(1 - \left(\frac{\kappa_0}{n}\right)^{1/d}\right)\right)^i \\ &\leq -d \left(\frac{\kappa_0}{n}\right)^{r/d} \sum_{\{i,j\} \in \mathcal{A}} \left(1 - \left(\frac{\kappa_{ij}}{n}\right)^{1/d}\right).\end{aligned}$$

We now apply the lower bound from Lemma 4.4 with $\Psi = \kappa_{ij}/n$. Note that this fulfills the condition $-\ln(\Psi)/d = o(1)$ as $d = \omega(\log(n))$ and $\kappa_{ij}/n = \Omega(1/n)$. We obtain

$$\begin{aligned}\ln(P) &\leq d \left(\frac{\kappa_0}{n}\right)^{r/d} \sum_{\{i,j\} \in \mathcal{A}} \left(\frac{\ln\left(\frac{\kappa_{ij}}{n}\right)}{d} + \frac{e \ln^2\left(\frac{\kappa_{ij}}{n}\right)}{d^2}\right) \\ &= \left(\frac{\kappa_0}{n}\right)^{r/d} \sum_{\{i,j\} \in \mathcal{A}} \ln\left(\frac{\kappa_{ij}}{n}\right) \left(1 - \frac{e}{d} \ln\left(\frac{n}{\kappa_0}\right)\right).\end{aligned}$$

Note that the above term is negative, so we need to lower bound the term $\left(\frac{\kappa_0}{n}\right)^{r/d}$ to proceed. We get from Lemma 4.4

$$\left(\frac{\kappa_0}{n}\right)^{r/d} \geq 1 - \frac{r}{d} \ln\left(\frac{n}{\kappa_0}\right) - \frac{er^2}{d^2} \ln\left(\frac{n}{\kappa_0}\right)^2 = 1 - \frac{c \ln(n)}{d}$$

for some constant $c > 0$ and sufficiently large n. With this,

$$\begin{aligned}\ln(P) &\leq \sum_{\{i,j\} \in \mathcal{A}} \ln\left(\frac{\kappa_{ij}}{n}\right) \left(1 - \frac{e}{d} \ln\left(\frac{n}{\kappa_0}\right)\right) \left(1 - \frac{c \ln(n)}{d}\right) \\ &\leq \sum_{\{i,j\} \in \mathcal{A}} \ln\left(\frac{\kappa_{ij}}{n}\right) \left(1 - \frac{\delta \ln(n)}{d}\right)\end{aligned}$$

for some $\delta > 0$ and sufficiently large n. ∎

LEMMA 4.6. *Let*

$$Q := \prod_{i=1}^{k} \left(1 - \sum_{\{i,j\} \in \mathcal{A}_i^-} \left(1 - \left(\frac{\kappa_{ij}}{n}\right)^{1/d}\right)\right)^d.$$

Then there is a constant δ *such that*

$$\ln(Q) \geq \sum_{\{i,j\} \in \mathcal{A}} \ln\left(\frac{\kappa_{ij}}{n}\right) \left(1 + \frac{\delta \ln(n)}{d}\right).$$

Proof. By the upper bound from Lemma 4.4 and the Taylor series of ln, we obtain

$$
\begin{aligned}
\ln(Q) := d\sum_{i=1}^{k}\ln\left(1-\sum_{\{i,j\}\in\mathcal{A}_i^-}\left(1-\left(\frac{\kappa_{ij}}{n}\right)^{1/d}\right)\right)
&\geq d\sum_{i=1}^{k}\ln\left(1-\sum_{\{i,j\}\in\mathcal{A}_i^-}\left(\frac{\ln\left(\frac{n}{\kappa_{ij}}\right)}{d}\right)\right)\\
&= -d\sum_{i=1}^{k}\sum_{\ell=1}^{\infty}\frac{1}{\ell}\left(\sum_{\{i,j\}\in\mathcal{A}_i^-}\left(\frac{\ln\left(\frac{n}{\kappa_{ij}}\right)}{d}\right)\right)^{\ell}\\
&= -d\sum_{i=1}^{k}\sum_{\{i,j\}\in\mathcal{A}_i^-}\left(\frac{\ln\left(\frac{n}{\kappa_{ij}}\right)}{d}\right)\sum_{\ell=1}^{\infty}\frac{1}{\ell}\left(\sum_{\{i,j\}\in\mathcal{A}_i^-}\left(\frac{\ln\left(\frac{n}{\kappa_{ij}}\right)}{d}\right)\right)^{\ell-1}\\
&= \sum_{\{i,j\}\in\mathcal{A}}\ln\left(\frac{\kappa_{ij}}{n}\right)\sum_{\ell=1}^{\infty}\frac{1}{\ell}\left(\sum_{\{i,j\}\in\mathcal{A}_i^-}\left(\frac{\ln\left(\frac{n}{\kappa_{ij}}\right)}{d}\right)\right)^{\ell-1}\\
&\geq \sum_{\{i,j\}\in\mathcal{A}}\ln\left(\frac{\kappa_{ij}}{n}\right)\sum_{\ell=1}^{\infty}\frac{1}{\ell}\left(\frac{k\ln\left(\frac{n}{\kappa_0}\right)}{d}\right)^{\ell-1}.
\end{aligned}
$$

Now, we have

$$
\sum_{\ell=1}^{\infty}\frac{1}{\ell}\left(\frac{k\ln\left(\frac{n}{\kappa_0}\right)}{d}\right)^{\ell-1} \leq 1+\frac{k\ln\left(\frac{n}{\kappa_0}\right)}{d}\sum_{\ell=0}^{\infty}\left(\frac{k\ln\left(\frac{n}{\kappa_0}\right)}{d}\right)^{\ell} \leq 1+\frac{2k\ln\left(\frac{n}{\kappa_0}\right)}{d}
$$

because $\frac{k\ln\left(\frac{n}{\kappa_0}\right)}{d}=o(1)$ and so the above geometric sum converges to $1+o(1)$. ∎

Proof of Theorem 1.5. Combining Theorem 4.2, Theorem 4.3, Lemma 4.5, and Lemma 4.6, we get that there are constants $\delta,\delta'>0$ such that

$$
\prod_{\{i,j\}\in\mathcal{A}}\left(\frac{\kappa_{ij}}{n}\right)^{1+\frac{\delta\ln(n)}{d}} \leq \Pr\left[E(U_k)\supseteq\mathcal{A}\mid\{\kappa\}^{(k)}\right] \leq \prod_{\{i,j\}\in\mathcal{A}}\left(\frac{\kappa_{ij}}{n}\right)^{1-\frac{\delta'\ln(n)}{d}}.
$$

If $d=\omega(\log(n))$, $\frac{\delta\ln(n)}{d}=o(1)$ and the proof of this case is finished. If $d=\omega(\log^2(n))$, observe that

$$
\prod_{\{i,j\}\in\mathcal{A}}\left(\frac{\kappa_{ij}}{n}\right)^{1+\frac{\delta\ln(n)}{d}} \geq \left(\frac{n}{\kappa_0}\right)^{-\binom{k-1}{2}\frac{\delta\ln(n)}{d}}\prod_{\{i,j\}\in\mathcal{A}}\left(\frac{\kappa_{ij}}{n}\right).
$$

Similarly,

$$
\prod_{\{i,j\}\in\mathcal{A}}\left(\frac{\kappa_{ij}}{n}\right)^{1-\frac{\delta'\ln(n)}{d}} \leq \left(\frac{n}{\kappa_0}\right)^{\binom{k-1}{2}\frac{\delta'\ln(n)}{d}}\prod_{\{i,j\}\in\mathcal{A}}\left(\frac{\kappa_{ij}}{n}\right).
$$

Since $d = \omega(\log^2(n))$, we get that

$$\left(\frac{n}{\kappa_0}\right)^{\binom{k-1}{2}\frac{\delta \ln(n)}{d}} = (1+o(1))n^{1/\omega(\log(n))} = (1+o(1))\exp\left(\frac{\log(n)}{\omega(\log(n))}\right) = 1+o(1)$$

as we assume k to be constant. $\square$

Bounds for triangles. We proceed by deriving a lower bound for the probability that three vertices form a triangle.

LEMMA 4.7. *Let $G = \mathrm{GIRG}(n, \beta, w_0, d)$ be a standard GIRG and let v_1, v_2, v_3 be three random vertices. Then,*

$$\Pr\left[v_2 \sim v_3 \mid v_1 \sim v_2, v_3\right] \geq \left(3 - 3\left(\frac{\kappa_0}{n}\right)^{-1/d} + \left(\frac{\kappa_0}{n}\right)^{-2/d}\right)^d.$$

Proof. We consider one fixed dimension and give an upper bound on the probability that $v_2 \not\sim v_3$ conditioned on $v_1 \sim v_2, v_3$. In the following, we abbreviate $p := \frac{\kappa_0}{n}$. Note that we may assume that all three vertices are of weight w_0 as it is easy to verify that larger weights only increase the probability of forming a triangle. Conditioned on the event $v_1 \sim v_2, v_3$, the vertices v_2, v_3 are uniformly distributed within a circular arc of length $p^{1/d}$ around v_1. In order for $v_2 \not\sim v_3$ to occur, v_3 needs to be placed within a circular arc of length $1 - p^{1/d}$ opposite to v_2. Hence, the probability that $v_2 \not\sim v_3$ conditioned on the event that v_2 is placed at distance x of v_1 is

$$\Pr\left[v_2 \not\sim v_3 \mid |x_{v_1} - x_{v_2}|_C = x\right] = \begin{cases} x/p^{1/d} & \text{if } x \leq 1 - p^{1/d}, \\ (1-p^{1/d})/p^{1/d} & \text{otherwise,} \end{cases}$$

where $|\cdot|_C$ denotes the distance on the circle. Since v_2 is distributed uniformly within distance $\frac{1}{2}p^{1/d}$ around v_1, we obtain

$$\begin{aligned} \Pr\left[v_2 \not\sim v_3 \mid v_1 \sim v_2, v_3\right] &\leq \frac{1}{\frac{1}{2}p^{1/d}} \int_0^{\frac{1}{2}p^{1/d}} \Pr\left[v_2 \not\sim v_3 \mid |x_{v_1} - x_{v_2}|_C = x\right] \mathrm{d}x \\ &= \frac{1}{\frac{1}{2}p^{1/d}} \left(\int_0^{1-p^{1/d}} \frac{x}{p^{1/d}} \mathrm{d}x + \int_{1-p^{1/d}}^{\frac{1}{2}p^{1/d}} \frac{1-p^{1/d}}{p^{1/d}} \mathrm{d}x\right). \end{aligned}$$

Solving the integrals then yields

$$\Pr\left[v_2 \not\sim v_3 \mid v_1 \sim v_2, v_3\right] \leq 3p^{-1/d} - 2 - p^{-2/d},$$

implying the desired bound. $\square$

We use this finding to show in the following lemma that, although cliques of size at least 4 already behave like in the IRG model if $d = \omega(\log(n))$, the number of triangles in the geometric case is still larger than that in the nongeometric case.

LEMMA 4.8. *Let $G = \mathrm{GIRG}(n, \beta, w_0, d)$ be a standard GIRG with $d = \omega(\log(n))$ and let v_1, v_2, v_3 be three random vertices. Then,*

$$\Pr\left[v_2 \sim v_3 \mid v_1 \sim v_2, v_3\right] \geq (1+o(1))\left(\frac{\kappa_0}{n}\right)^{1 - \frac{\ln(n)^2}{d^2} \pm \mathcal{O}\left(\frac{\ln(n)^3}{d^3}\right)}.$$

Proof. By Lemma 4.7, we have $\Pr[v_2 \sim v_3 \mid v_1 \sim v_2, v_3] = q^d$ with

$$q := 3\left(1 - \left(\frac{\kappa_0}{n}\right)^{-1/d}\right) + \left(\frac{\kappa_0}{n}\right)^{-2/d}.$$

In the following, we once again abbreviate $p = \frac{\kappa_0}{n}$ and start by decomposing q into a sum.

$$\begin{aligned} q &= 3\left(1 - p^{-1/d}\right) + p^{-2/d} \\ &= 3\left(1 - \exp\left(\frac{-\ln(p)}{d}\right)\right) + \exp\left(\frac{-2\ln(p)}{d}\right) \\ &= -3\sum_{i=1}^{\infty}\frac{1}{i!}\left(\frac{-\ln(p)}{d}\right)^i + \sum_{i=0}^{\infty}\frac{1}{i!}\left(\frac{-2\ln(p)}{d}\right)^i \\ &= 1 + \sum_{i=1}^{\infty}\frac{1}{i!}\left(\left(\frac{-2\ln(p)}{d}\right)^i - 3\left(\left(\frac{-\ln(p)}{d}\right)^i\right)\right) \\ &= 1 + \sum_{i=1}^{\infty}\frac{1}{i!}\left(\frac{-\ln(p)}{d}\right)^i (2^i - 3). \end{aligned}$$

We now proceed by bounding $d\ln(q)$ and apply the Taylor series of ln to get

$$d\ln(q) = -d\sum_{i=1}\frac{(1-q)^i}{i} = \sum_{i=1}(-1)^{i+1} d\frac{(q-1)^i}{i}. \tag{4.3}$$

By abbreviating $x := \ln(p)/d$ and our above sum, we get

$$\begin{aligned} q - 1 &= \sum_{i=1}^{\infty}\frac{1}{i!}\left(\frac{-\ln(p)}{d}\right)^i (2^i - 3) = x + \frac{1}{2}x^2 - \frac{5}{6}x^3 + \dots \\ &= x\underbrace{\left(1 + \frac{1}{2}x - \frac{5}{6}x^2 + R\right)}_{Z}, \end{aligned} \tag{4.4}$$

where $|R| = \mathcal{O}(|x|^3)$ because $|x| = o(1)$. Inserting this into the first terms of the sum in (4.3) yields

$$\begin{aligned} d\ln(q) &= dxZ - d\frac{1}{2}x^2Z^2 + d\frac{1}{3}x^3Z^3 + \dots \\ &= dx\left(Z - \frac{1}{2}xZ^2 + \frac{1}{3}x^2Z^3 + \sum_{i=4}(-1)^{i+1}x^{i-1}Z^i\right). \end{aligned}$$

Using the definition of Z in (4.4), we may carefully compute the leading terms of the above sum to obtain

$$d\ln(q) = dx\left(1 - x^2 + R'\right),$$

with $|R'| = \mathcal{O}(|x|^3)$. Recalling that $x = \ln(p)/d$ gives

$$d\ln(q) = \ln(p)\left(1 - \left(\frac{\ln(p)}{d}\right)^2 \pm \mathcal{O}\left(\left|\frac{\ln(p)}{d}\right|^3\right)\right).$$

As $p = \kappa_0/n$, this shows

$$q^d = \left(\frac{\kappa_0}{n}\right)^{1-\frac{\ln(n/\kappa_0)^2}{d^2}\pm\mathcal{O}\left(\frac{\ln(n)^3}{d^3}\right)} = (1+o(1))\left(\frac{\kappa_0}{n}\right)^{1-\frac{\ln(n)^2}{d^2}\pm\mathcal{O}\left(\frac{\ln(n)^3}{d^3}\right)}.$$

Here, the factor of $1+o(1)$ derives from the fact that

$$\frac{\ln(n/\kappa_0)^2}{d^2} = \frac{\ln(n)^2 - 2\ln(n)\ln(\kappa_0)+\ln(\kappa_0)^2}{d^2} = \frac{\ln(n)^2}{d^2} - \frac{1}{\omega(\log(n))}. \qquad \square$$

The above bound implies that the probability that a triangle forms among vertices of constant weight is still significantly larger than in the nongeometric models as long as $d = o(\log^{3/2}(n))$. We summarize this behavior in the following theorem.

THEOREM 4.9. *Let $G = \mathrm{GIRG}(n,\beta,w_0,d)$ be a standard GIRG with $\beta > 3$ and $d = \omega(\log(n))$. Then, there is a constant $\delta \geq 0$ such that the expected number of triangles fulfills*

$$\mathbb{E}[K_3] = \begin{cases} \Omega\left(\exp\left(\frac{\ln^3(n)}{d^2}\right)\right) & \text{if } 3<\beta<\infty, \\ \Theta\left(\exp\left(\frac{\ln^3(n)}{d^2}\right)\right) & \text{if } \beta=\infty, \end{cases}$$

where $\beta = \infty$ refers to the case where every vertex has the same weight.

Proof. We note that by Lemma 4.8

$$\Pr[U_3 \text{ is clique}] \geq (1+o(1))\left(\frac{\kappa_0}{n}\right)^2\left(\frac{\kappa_0}{n}\right)^{1-\frac{\ln(n)^2}{d^2}\pm\mathcal{O}\left(\frac{\ln(n)^3}{d^3}\right)}$$

and thus

$$\mathbb{E}[K_3] = \binom{n}{3}\Pr[U_3 \text{ is clique}] \geq (1+o(1))\kappa_0^3 n^{\Omega\left(\frac{\ln^2(n)}{d^2}\right)} = \Omega\left(\exp\left(\frac{\ln^3(n)}{d^2}\right)\right),$$

and the first part of the statement is shown. For the second part, we note that in the case of constant weights, the bound from Lemma 4.7 is the exact (conditional) probability that a triangle is formed and the transformations from Lemma 4.8 still apply for obtaining an upper bound. $\square$

4.2. Characterizing cliques by vertex weights. In this section, we extend the bounds on q_k obtained above to the entire graph and characterize it by the weights of the associated vertices assuming $d = \omega(\log(n))$ and $\beta \in (2,3)$ (the other parameter regimes are discussed in the subsequent section). To this end, we will mainly be concerned with bounding the following integral.

LEMMA 4.10. *Let ρ be the density function of the Pareto distribution, let $\varepsilon = \varepsilon(n) = o(1)$, and define*

$$\Lambda(k) := \int_0^\infty \cdots \int_0^\infty \rho(w_1)\cdots\rho(w_k) \prod_{1\leq i<j\leq k}\left(\frac{\kappa_{ij}}{n}\right)^{1-\varepsilon} \mathrm{d}w_k \ldots \mathrm{d}w_1.$$

For every constant k, we have

$$\Lambda(k) = \mathcal{O}\left(n^{\frac{k}{2}(1-\beta)}\right).$$

Remark 4.11. Clearly, $\Lambda(k)$ is an upper bound on the clique probability if we used w_0 as the lower integration limit instead of 0. However, we show the more general statement where the integration limit is 0 as it will be useful later in the proof of Lemma 4.12.

Proof of Lemma 4.10. We use a technique similar to that of Daly et al. [14], who bound

$$\Lambda(k) \leq \sum_{m=0}^{k} \binom{k}{m} \Lambda(k,m),$$

where

$$\Lambda(k,m) := \underbrace{\int_0^{\sqrt{n/\lambda}} \cdots \int_0^{\sqrt{n/\lambda}}}_{m \text{ times}} \underbrace{\int_{\sqrt{n/\lambda}}^{\infty} \cdots \int_{\sqrt{n/\lambda}}^{\infty}}_{k-m \text{ times}} \rho(w_1)\cdots\rho(w_k) \prod_{1\leq i<j\leq k} \left(\frac{\kappa_{ij}}{n}\right)^{1-\varepsilon} \mathrm{d}w_k \ldots \mathrm{d}w_1.$$

We start with the first extreme case:

$$\Lambda(k,k) = \int_0^{\sqrt{n/\lambda}} \cdots \int_0^{\sqrt{n/\lambda}} \rho(w_1)\cdots\rho(w_k) \prod_{1\leq i<j\leq k} \left(\frac{\kappa_{ij}}{n}\right)^{1-\varepsilon} \mathrm{d}w_1 \ldots \mathrm{d}w_k.$$

Since $\rho(w) = cw^{-\beta}$ for some constant c and as k is constant,

$$\begin{aligned}
&\Lambda(k,k)\\
&\quad = \Theta(1) \int_0^{\sqrt{n/\lambda}} \cdots \int_0^{\sqrt{n/\lambda}} \prod_{1\leq i<j\leq k} \left(\frac{\kappa_{ij}}{n}\right)^{1-\varepsilon} w_1^{-\beta} \cdots w_k^{-\beta} \mathrm{d}w_1 \ldots \mathrm{d}w_k\\
&\quad = \Theta(1) n^{-\binom{k}{2}(1-\varepsilon)} \left(\int_0^{\sqrt{n/\lambda}} w^{(k-1)(1-\varepsilon)-\beta} \mathrm{d}w \right)^k\\
&\quad = \Theta(1) n^{-\binom{k}{2}(1-\varepsilon)} n^{\frac{1}{2}k(k-1)(1-\varepsilon)+\frac{k}{2}(1-\beta)}\\
&\quad = \Theta(1) n^{\frac{k}{2}(1-\beta)}
\end{aligned}$$

as $\binom{k}{2} = \frac{1}{2}k(k-1)$. Note that the second step holds for sufficiently large n since $k \geq 3$ and $\beta < 3$ as this leads to an exponent in the integral that is strictly greater than -1. The above bounds only hold for sufficiently large n as $\varepsilon = o(1)$. For the other extreme case, let $w_{\min} = \min\{w_1, \ldots, w_k\}$; then Corollary 3.3 yields that

$$\Lambda(k,0) = \Pr\left[w_{\min} \geq \sqrt{n/\lambda}\right] = \Theta(1) n^{\frac{k}{2}(1-\beta)}$$

because $\kappa_{ij} = 1$ if $w_i, w_j \geq \sqrt{n/\lambda}$.

If $m \geq 3$, we bound

$$\begin{aligned}
\Lambda(k,m) &\leq \Lambda(m,m) \cdot \Lambda(k-m,0)\\
&= \Theta(1) n^{\frac{m}{2}(1-\beta)} n^{\frac{k-m}{2}(1-\beta)}\\
&= \Theta(1) n^{\frac{k}{2}(1-\beta)},
\end{aligned}$$

as desired. The case $m = \{1,2\}$ thus remains. To this end, we use that

$$\begin{aligned}
\Lambda(k,m) &\leq \Lambda(3,m)\Lambda(k-3,0)\\
&= \Theta(1)\Lambda(3,m) n^{\frac{k-3}{2}(1-\beta)}
\end{aligned}$$

and so it suffices to show that $\Lambda(3,m) = \Theta(1) n^{\frac{3}{2}(1-\beta)}$ for $m \in \{1,2\}$. For $m = 1$, we bound

$$
\begin{aligned}
&\Lambda(3,1)\\
&\quad\le \Theta(1)\int_0^{\sqrt{n/\lambda}}\int_{\sqrt{n/\lambda}}^{\infty}\int_{\sqrt{n/\lambda}}^{\infty}\left(\frac{w_1^2w_2w_3}{n}\right)^{2(1-\varepsilon)} w_1^{-\beta}w_2^{-\beta}w_3^{-\beta}\mathrm{d}w_1\mathrm{d}w_2\mathrm{d}w_3\\
&\quad= \Theta(1)n^{-2(1-\varepsilon)}\int_0^{\sqrt{n/\lambda}} w^{2(1-\varepsilon)-\beta}\mathrm{d}w\left(\int_{\sqrt{n/\lambda}}^{\infty} w^{1-\varepsilon-\beta}\mathrm{d}w\right)^2\\
&\quad= \Theta(1)n^{-2(1-\varepsilon)+(1-\varepsilon)+\frac{1}{2}(1-\beta)+(1-\varepsilon)+(1-\beta)}\\
&\quad= \Theta(1)n^{\frac{3}{2}(1-\beta)}
\end{aligned}
$$

as desired. Again, this works because for sufficiently large n, the exponent in the integral starting at 0 is greater than -1. For the case $m=2$, we use

$$
\begin{aligned}
\Lambda(3,2)\le \Theta(1)&\underbrace{\int_0^{\sqrt{n/\lambda}}\int_0^{w_1}\int_{\sqrt{n/\lambda}}^{n/w_1}\prod_{1\le i<j\le 3}\left(\frac{w_iw_j}{n}\right)^{1-\varepsilon}(w_1w_2w_3)^{-\beta}\mathrm{d}w_3\mathrm{d}w_2\mathrm{d}w_1}_{I_1}\\
&+\Theta(1)\underbrace{\int_0^{\sqrt{n/\lambda}}\int_0^{w_1}\int_{n/w_1}^{\infty}\left(\frac{w_1w_2^2w_3}{n^2}\right)^{1-\varepsilon}(w_3w_2w_1)^{-\beta}\mathrm{d}w_3\mathrm{d}w_2\mathrm{d}w_1}_{I_2}.
\end{aligned}
$$

Note that this is a valid bound since the vertices v_1, v_2 are interchangeable and thus P_2 is at most twice as large as the two integrals above, which is captured by the $\Theta(1)$ terms. Calculations now show that

$$
\begin{aligned}
I_1 &= n^{-3(1-\varepsilon)}\int_0^{\sqrt{n/\lambda}}\int_0^{w_1}\int_{\sqrt{n/\lambda}}^{n/w_1}\prod_{1\le i<j\le 3}(w_iw_j)^{1-\varepsilon}(w_1w_2w_3)^{-\beta}\mathrm{d}w_3\mathrm{d}w_2\mathrm{d}w_1\\
&= \Theta(1)n^{-3(1-\varepsilon)}\int_0^{\sqrt{n/\lambda}} w_1^{2(1-\varepsilon)-\beta}\int_0^{w_1} w_2^{2(1-\varepsilon)-\beta}\int_{\sqrt{n/\lambda}}^{n/w_1} w_3^{2(1-\varepsilon)-\beta}\mathrm{d}w_3\mathrm{d}w_2\mathrm{d}w_1\\
&\le \Theta(1)n^{-3(1-\varepsilon)}\int_0^{\sqrt{n/\lambda}} w_1^{2(1-\varepsilon)-\beta}\int_0^{w_1} w_2^{2(1-\varepsilon)-\beta}\left(\frac{n}{w_1}\right)^{2(1-\varepsilon)-\beta+1}\mathrm{d}w_2\mathrm{d}w_1\\
&= \Theta(1)n^{-(1-\varepsilon)-\beta+1}\int_0^{\sqrt{n/\lambda}} w_1^{-1}\int_0^{w_1} w_2^{2(1-\varepsilon)-\beta}\mathrm{d}w_2\mathrm{d}w_1\\
&= \Theta(1)n^{\varepsilon-\beta}\int_0^{\sqrt{n/\lambda}} w_1^{2(1-\varepsilon)-\beta}\mathrm{d}w_1 = \Theta(1)n^{\varepsilon-\beta+1-\varepsilon-\frac{1}{2}\beta+\frac{1}{2}} = \Theta(1)n^{\frac{3}{2}(1-\beta)}
\end{aligned}
$$

and

$$
\begin{aligned}
I_2 &= \int_0^{\sqrt{n/\lambda}}\int_0^{w_1}\int_{n/w_1}^{\infty}\left(\frac{w_1w_2^2w_3}{n^2}\right)^{1-\varepsilon}(w_3w_2w_1)^{-\beta}\mathrm{d}w_3\mathrm{d}w_2\mathrm{d}w_1\\
&\le \Theta(1)n^{-2(1-\varepsilon)}\int_0^{\sqrt{n/\lambda}} w_1^{1-\varepsilon-\beta}\int_0^{w_1} w_2^{2(1-\varepsilon)-\beta}\left(\frac{n}{w_1}\right)^{1-\varepsilon-\beta+1}\mathrm{d}w_2\mathrm{d}w_1\\
&= \Theta(1)n^{-(1-\varepsilon)-\beta+1}\int_0^{\sqrt{n/\lambda}} w_1^{-1}\int_0^{w_1} w_2^{2(1-\varepsilon)-\beta}\mathrm{d}w_2\mathrm{d}w_1\\
&= \Theta(1)n^{\frac{3}{2}(1-\beta)}
\end{aligned}
$$

as desired.

The above lemma not only implies our claimed bounds on the expected number of cliques but also allows us to show that cliques of all sizes form dominantly among vertices of weight on the order of $\sqrt{n}$.

LEMMA 4.12. *Let $d = \omega(\log(n))$ and $\beta \in (2,3)$. Then for all (potentially super-constant) k and all $p \in (0,1)$, there is an $\varepsilon > 0$ such that*

$$\Pr\left[w_{\min} \geq \varepsilon\sqrt{n} \mid U_k \text{ is clique}\right] \geq p.$$

For $k > \frac{2}{3-\beta}$, we already showed the statement in Lemma 3.9. For $k \leq \frac{2}{3-\beta}$, we use an argument inspired by the techniques introduced in [27]. Note that in the following, we assume k to be constant since we are in the case $k \leq \frac{2}{3-\beta}$. We write

$$\Pr\left[w_{\min} < \varepsilon\sqrt{n} \mid U_k \text{ is clique}\right] = \frac{\Pr\left[w_{\min} < \varepsilon\sqrt{n} \cap U_k \text{ is clique}\right]}{\Pr\left[U_k \text{ is clique}\right]}$$

and proceed by showing that this probability can be made arbitrarily small by choosing ε large enough. By considering the event that $w_{\min} \geq \sqrt{n/\lambda}$, we immediately get $\Pr[U_k \text{ is clique}] = \Omega(n^{\frac{k}{2}(1-\beta)})$. To show our statement, we proceed by showing that

$$\Pr\left[w_{\min} < \varepsilon\sqrt{n} \cap U_k \text{ is clique}\right] \leq f(\varepsilon) n^{\frac{k}{2}(1-\beta)}$$

for a function f that tends to 0 as $\varepsilon \to 0$. To this end, note that by Theorem 1.5, there is a function $\varepsilon = \varepsilon(n) = o(1)$ such that

$$\begin{aligned}\Pr\left[U_k \text{ is clique}\right] &\leq \Lambda(k)\\ &= \int_0^\infty \cdots \int_0^\infty \rho(w_1)\cdots\rho(w_k) \prod_{1\leq i<j\leq k} \left(\frac{\kappa_{ij}}{n}\right)^{1-\varepsilon} \mathrm{d}w_k \ldots \mathrm{d}w_1.\end{aligned}$$

Substituting $w_i = y_i\sqrt{n/\lambda}$ and recalling that $\rho(w) = cw^{-\beta}$ for some constant c yields

$$\begin{aligned}&\Lambda(k)\\ &= \int_0^\infty \cdots \int_0^\infty c^k (y_1\cdots y_k)^{-\beta} \sqrt{\frac{n}{\lambda}}^{-k\beta} \prod_{1\leq i<j\leq k} \left(\frac{\min\{n, \lambda n y_i y_j\}}{n}\right)^{1-\varepsilon} \sqrt{\frac{n}{\lambda}}^{k} \mathrm{d}y_k \ldots \mathrm{d}y_1\\ &= \left(\frac{n}{\lambda}\right)^{\frac{k}{2}(1-\beta)} \underbrace{\int_0^\infty \cdots \int_0^\infty c^k (y_1\cdots y_k)^{-\beta} \prod_{1\leq i<j\leq k} \left(\min\{1, \lambda y_i y_j\}\right)^{1-\varepsilon} \mathrm{d}y_k \ldots \mathrm{d}y_1}_{:=I}.\end{aligned}$$

By Lemma 4.10, we have $\Lambda(k) = \mathcal{O}(n^{\frac{k}{2}(1-\beta)})$ and thus $I < \infty$. Now

$$\begin{aligned}&\Pr\left[U_k \text{ is clique} \cap w_{\min} < \varepsilon\sqrt{n}\right]\\ &\quad\leq k \int_0^{\varepsilon\sqrt{n}} \int_0^\infty \cdots \int_0^\infty \rho(w_1)\cdots\rho(w_k) \prod_{1\leq i<j\leq k} \left(\frac{\kappa_{ij}}{n}\right)^{1-\varepsilon} \mathrm{d}w_k \ldots \mathrm{d}w_1.\end{aligned}$$

Using the same substitution as above, this yields

$$\begin{aligned}&\Pr\left[U_k \text{ is clique} \cap w_{\min} < \varepsilon\sqrt{n}\right]\\ &\leq k \left(\frac{n}{\lambda}\right)^{\frac{k}{2}(1-\beta)} \underbrace{\int_0^{\varepsilon\sqrt{\lambda}} \cdots \int_0^\infty c^k (y_1\cdots y_k)^{-\beta} \prod_{1\leq i<j\leq k} \left(\min\{1, \lambda y_i y_j\}\right)^{1-\varepsilon} \mathrm{d}y_k \ldots \mathrm{d}y_1}_{:=J(\varepsilon)}.\end{aligned}$$

As $\lim_{\varepsilon\to\infty} J(\varepsilon) = I < \infty$, $J(\varepsilon)$ can be made arbitrarily small by choosing $\varepsilon > 0$ small enough, which yields the desired statement. ∎

4.3. Bounds on $\mathbb{E}[K_k]$ and $\omega(G)$. We use the findings from subsection 4.1 to prove the third column of Table 3 and Table 1, respectively. We treat the cases $2 < \beta < 3$, $\beta = 3$, and $\beta > 3$ separately.

For $\beta \in (2,3)$, we show that if $d = \omega(\log(n))$, then the expected number of cliques is $n^{\frac{k}{2}(3-\beta)}\Theta(k)^{-k}$ for all $k \geq 3$ just like in the IRG model. This is despite the fact that the probability that a clique forms among vertices of constant weight is still significantly higher than in the IRG model if $\log(n) \ll d \ll \log^{3/2}(n)$. The reason for this is that the probability of forming a clique among vertices of weight close to $\sqrt{n}$ behaves like that of the IRG model if $d = \omega(\log(n))$ and because cliques forming among these high-weight vertices dominate all the others.

THEOREM 4.13. *Let $G = \mathrm{GIRG}(n, \beta, w_0, d)$ be a standard GIRG with $\beta < 3$ and $d = \omega(\log(n))$. Then, $\mathbb{E}[K_k] = n^{\frac{k}{2}(3-\beta)}\Theta(k)^{-k}$ for $k \geq 3$ and $\omega(G) = \Theta(n^{(3-\beta)/2})$ a.a.s.*

Proof. Observe that Corollary 3.3 and Corollary 3.5 imply our desired bounds if $k > \frac{2}{3-\beta}$. Otherwise, just considering the event that $w_{\min} \geq \sqrt{n/\lambda}$ gives us the desired lower bound on q_k and thus on $\mathbb{E}[K_k]$. To get an upper bound, note that $q_k \leq \Lambda(k)$ by Theorem 1.5, and that $\Lambda(k) = \mathcal{O}(n^{\frac{k}{2}(3-\beta)})$ by Lemma 4.10, which directly implies our claimed bounds on $\mathbb{E}[K_k]$. To bound $\omega(G)$, now the same argumentation as in Theorem 3.21 applies. ◻

We note that the phase transition at $k = \frac{2}{3-\beta}$ is no longer present. We continue with the case where $\beta = 3$.

THEOREM 4.14. *Let $G = \mathrm{GIRG}(n, \beta, w_0, d)$ be a standard GIRG with $\beta = 3$ and $d = \omega(\log(n))$. Then, $\omega(G) = \mathcal{O}(1)$.*

Proof. We show that the number of cliques of size $k = 4$ is such that for every $\varepsilon > 0$ there is a constant $c > 0$ such that

$$\Pr[K_4 \geq c] \leq \varepsilon.$$

That is, the probability that $K_4 \geq c$ can be made arbitrarily small by choosing c large enough. This fact is sufficient to show that the clique number is $\Theta(1)$.

We start by observing that $\beta = 3$ implies that the maximum weight w_{max} is a.a.s. on the order of $\sqrt{n}$. More precisely, denoting by X_w the number of vertices with weight at least w, we get by Markov's inequality that for every w,

$$\Pr[w_{max} \geq w] = \Pr[X_w \geq 1] \leq \mathbb{E}[X_w] = \Theta(1)nw^{1-\beta} = \Theta(1)nw^{-2}. \tag{4.5}$$

Thus, for every $c \geq 0$, we have

$$\Pr\left[w_{max} \geq c\sqrt{n}\right] = \mathcal{O}(1)c^{-2}. \tag{4.6}$$

With this and Markov's inequality, we may bound for every $t, c \geq 0$

$$\begin{aligned} \Pr[K_4 \geq t] &\leq \Pr\left[K_4 \geq t \mid w_{max} \leq c\sqrt{n}\right] + \Pr\left[w_{max} \geq c\sqrt{n}\right] \\ &\leq \frac{\mathbb{E}\left[K_4 \mid w_{max} \leq c\sqrt{n}\right]}{t} + \Pr\left[w_{max} \geq c\sqrt{n}\right]. \end{aligned} \tag{4.7}$$

To bound $\mathbb{E}[K_4 \mid w_{max} \leq c\sqrt{n}]$, we note that a random weight w_i fulfills

$$\Pr\left[w_i \le x \mid w_{max} \le c\sqrt{n}\right] = \frac{\Pr\left[\bigcap_j w_j \le c\sqrt{n} \cap w_i \le x\right]}{\Pr\left[\bigcap_j w_j \le c\sqrt{n}\right]}$$
$$= \frac{\Pr\left[w_i \le x\right]}{\Pr\left[w_i \le c\sqrt{n}\right]}$$

and since $\Pr\left[w_i \le c\sqrt{n}\right] = 1 - o(1)$, we get that the density of w_i is $\rho_{w_i}(x) = c' x^{-\beta}$ for some $c' \ge 0$ independent of x. Using that, we use Theorem 1.5 to bound

$$\begin{aligned}
&\Pr\left[U_k \text{ is clique } \mid w_k \le c\sqrt{n}\right] \\
&\quad \le \Theta(1) \int_{w_0}^{c\sqrt{n}} \cdots \int_{w_0}^{c\sqrt{n}} \prod_{1 \le i < j \le k} \left(\frac{\kappa_{ij}}{n}\right)^{1-\varepsilon} w_1^{-\beta} \cdots w_k^{-\beta} \mathrm{d}w_1 \ldots \mathrm{d}w_k \\
&\quad = \Theta(1) n^{-\binom{k}{2}(1-\varepsilon)} \left(\int_{w_0}^{c\sqrt{n}} w^{(k-1)(1-\varepsilon)-\beta} \mathrm{d}w\right)^k
\end{aligned}$$

for some function $\varepsilon(n) = o(1)$. Because $k \ge 4$ and $\beta = 3$, we observe that the exponent $(k-1)(1-\varepsilon) - \beta$ is greater than -1 for sufficiently large n, and hence, the above integral evaluates to

$$\Pr\left[U_k \text{ is clique } \mid w_k \le c\sqrt{n}\right] = c^{k((k-1)(1-\varepsilon)+1-\beta)} \Theta(1) n^{\frac{k}{2}(1-\beta)}.$$

For $k = 4$, we obtain

$$\Pr\left[U_k \text{ is clique } \mid w_k \le c\sqrt{n}\right] \le \Theta(1) c^4 n^{-4}$$

and accordingly

$$\mathbb{E}\left[K_4 \mid w_{max} \le c\sqrt{n}\right] = \binom{n}{4} \Pr\left[U_k \text{ is clique } \mid w_k \le c\sqrt{n}\right] \le \Theta(1) c^4.$$

By (4.7) and (4.6), this implies

$$\Pr\left[K_4 \ge t\right] = \Theta(1) c^4 t^{-1} + \mathcal{O}(1) c^{-2}.$$

Setting $c = t^{1/5}$ yields $\Pr[K_4 \ge t] = \mathcal{O}(1)(t^{-1/5} + t^{-2/5})$ and shows that the probability that $K_4 \ge t$ can be made arbitrarily small by increasing t. To bound the clique number, we note that the existence of a clique of size k implies the existence of $\binom{k}{4}$ cliques of size 4, and so

$$\Pr\left[\omega(G) \ge k\right] \le \Pr\left[K_4 \ge \binom{k}{4}\right],$$

which can be made arbitrarily small by choosing k large enough. Hence the probability that the clique number grows as any superconstant function $f(n) = \omega(1)$ is in $o(1)$, which shows that the clique number is in $\mathcal{O}(1)$ a.a.s. ∎

Finally, we deal with the case where $\beta > 3$, where we show that, in this case, there are no cliques of size 4 or larger a.a.s.

THEOREM 4.15. *Let $G = \mathrm{GIRG}(n, \beta, w_0, d)$ be a standard GIRG with $\beta > 3$ and $d = \omega(\log(n))$. Then, $\mathbb{E}[K_k] = o(1)$ for $k \ge 4$ and, thus, $\omega(G) \le 3$ a.a.s.*

Proof. We use a similar strategy as in the above paragraph. In analogy to (4.5), we now have

$$\Pr[w_{max} \geq n^{\alpha}] \leq \Theta(1) n^{1-(1-\beta)\alpha},$$

which is $o(1)$ if $\alpha > \frac{1}{\beta-1}$. For some α in the range $1/2 > \alpha > \frac{1}{\beta-1}$, we get

$$\begin{aligned}
&\Pr[U_k \text{ is clique } \mid w_k \leq n^{\alpha}] \\
&\quad \leq \Theta(1) \int_{w_0}^{n^{\alpha}} \cdots \int_{w_0}^{n^{\alpha}} \prod_{1 \leq i < j \leq k} \left(\frac{\kappa_{ij}}{n}\right)^{1-\varepsilon} w_1^{-\beta} \cdots w_k^{-\beta} \mathrm{d}w_1 \ldots \mathrm{d}w_k \\
&\quad = \Theta(1) n^{-\binom{k}{2}(1-\varepsilon)} \left(\int_{w_0}^{n^{\alpha}} w^{(k-1)(1-\varepsilon)-\beta} \mathrm{d}w\right)^k \\
&\quad = \Theta(1) n^{\left(\alpha-\frac{1}{2}\right)k(k-1)(1-\varepsilon)+\alpha k(1-\beta)} = o\left(n^{\alpha k(1-\beta)}\right).
\end{aligned}$$

Accordingly,

$$\mathbb{E}\left[K_4 \mid w_{max} \leq c\sqrt{n}\right] = \binom{n}{4} \Pr\left[U_k \text{ is clique } \mid w_k \leq c\sqrt{n}\right] = \Theta(1) n^{4+4\alpha(1-\beta)} = o(1)$$

as $\alpha(1-\beta) < -1$. By Markov's inequality this implies

$$\Pr[K_4 \geq 1] \leq \mathbb{E}[K_4 \mid w_{max} \leq n^{\alpha}] + \Pr[w_{max} \geq n^{\alpha}] = o(1),$$

as desired. □

5. Concentration bounds.

We use the insights gained so far to obtain the following concentration bounds on the total number of cliques that hold for almost all parameter regimes and clique sizes we consider.

THEOREM 1.4. *We have $K_k/\mathbb{E}[K_k] \to_p 1$; that is, for all $\delta > 0$,*

$$\Pr\left[\left|\frac{K_k}{\mathbb{E}[K_k]} - 1\right| \geq \delta\right] = o(1)$$

if one of the following conditions holds.

(i) $d = o(\log(n))$, $\beta \in (2,3)$, $k \neq \frac{2}{3-\beta}$, *and* $k = o(n^{(3-\beta)/4})$.
(ii) $d = \omega(\log(n))$, $\beta \in (2,3)$, $k = o(n^{(3-\beta)/4})$.
(iii) $d = o(\log(n))$, $\beta > 3$ *and* $k = o\left(\log(n)/(\log\log(n) + d)\right)$.

Proof. We start with the regimes where cliques dominantly form among vertices of weight $\sqrt{n}$. Recall that this covers case (ii), and case (i) if we additionally assume $k < \frac{2}{3-\beta}$. We write

$$K_k = K_k(M_{\varepsilon}^{(-)}(\sqrt{n})) + K_k(\overline{M_{\varepsilon}^{(-)}(\sqrt{n})}).$$

Since we are in the regime where cliques form dominantly in $M_{\varepsilon}(\sqrt{n})$, as shown in subsection 3.2 and subsection 4.2, there is some function $\varepsilon = o(1)$ such that

$$\mathbb{E}\left[K_k(M_{\varepsilon}^{(-)}(\sqrt{n}))\right] = (1 - o(1))\mathbb{E}[K_k] \text{ and } \mathbb{E}\left[K_k(\overline{M_{\varepsilon}^{(-)}(\sqrt{n})})\right] = o(1)\mathbb{E}[K_k].$$

Furthermore, by Lemma 3.18 and Chebyshev's inequality, we have

$$\Pr\left[\left|\frac{K_k(M_\varepsilon^{(-)}(\sqrt{n}))}{\mathbb{E}[K_k]}-1\right|\geq\delta\right]\leq(1+o(1))\delta^{-2}\sum_{\ell=1}^{k}\mathcal{O}\left(\frac{\varepsilon^{1-\beta}k^2}{n^{(3-\beta)/2}}\right)^{\ell}$$
$$=\mathcal{O}\left(\frac{\delta^{-2}\varepsilon^{1-\beta}k^2}{n^{(3-\beta)/2}}\right).$$

If we choose an $\varepsilon=o(1)$ that decays sufficiently slowly, this tends to 0 for every $\delta>0$ due to our assumption on k. Furthermore,

$$K_k(\overline{M_\varepsilon^{(-)}(\sqrt{n})})/\mathbb{E}[K_k]\to_p 0$$

by Markov's inequality, and the proof of this case is finished.

On the other hand, if cliques dominantly form among low-weight vertices (this is the case for case (iii), and case (1) for $k>\frac{2}{3-\beta}$), write

$$K_k=K_k(M_\varepsilon^{(+)}(w))+K_k(\overline{M_\varepsilon^{(+)}(w)})$$

for some $w=e^{\Theta(1)d}k^{\frac{1}{2-\beta}}$ and note that this covers the cases (iii) and (iv). By Lemma 3.16,

$$\Pr\left[\left|\frac{K_k(M_\varepsilon^{(+)}(w))}{\mathbb{E}[K_k]}-1\right|\geq\delta\right]\leq(1+o(1))\delta^{-2}\frac{\mathcal{O}(w/\varepsilon)^k}{\mathbb{E}[K_k]}.$$

Note that due to $k=o(\frac{\log(n)}{\log\log(n)+d})=n^{o(1)}$ and $w=n^{\Theta(1)d/\log(n)}k^{\frac{1}{2-\beta}}$, we have $\mathcal{O}(w)^k=n^{o(1)}$, and further $\mathbb{E}[K_k]\geq n^c$ for some constant $c>0$ (cf. Theorem 3.25). Accordingly,

$$\Pr\left[\left|\frac{K_k(M_\varepsilon^{(+)}(w))}{\mathbb{E}[K_k]}-1\right|\geq\delta\right]\leq(1+o(1))\delta^{-2}\varepsilon^{-k}n^{-c+o(1)}.$$

If we choose an $\varepsilon=o(1)$ that decays sufficiently slowly, the above term is $o(1)$. Moreover, as above, for $\varepsilon=o(1)$, we have

$$\mathbb{E}\left[K_k(M_\varepsilon^{(+)}(w))\right]=(1-o(1))\mathbb{E}[K_k]\text{ and }\mathbb{E}\left[K_k(\overline{M_\varepsilon^{(+)}(w)})\right]=o(1)\mathbb{E}[K_k],$$

and the rest of the proof is analogous to the previous case. ∎

6. Asymptotic equivalence with IRGs. We continue by studying the infinite-dimensional limit of our model, i.e., the case where n is fixed and d goes to infinity. We prove that in this situation, the GIRG model becomes in a strong sense equivalent to the nongeometric IRG model. That is, we prove that the total variation distance of the distribution over all possible graphs with n vertices goes to 0 as $d\to\infty$. We prove the following theorem.

THEOREM 1.1. *Let $\mathcal{G}(n)$ be the set of all graphs with n vertices, let $\{w\}_1^n$ be a weight sequence, and consider $G_{IRG}=\mathrm{IRG}(\{w\}_1^n)\in\mathcal{G}(n)$ and a standard GIRG $G_{GIRG}=\mathrm{GIRG}(\{w\}_1^n,d)\in\mathcal{G}(n)$ with any L_p-norm. Then,*

$$\lim_{d\to\infty}\|G_{GIRG},G_{IRG}\|_{TV}=0.$$

We split the proof of this theorem into two parts by considering the case of an L_p-norm for $1\leq p<\infty$ and the case of L_∞-norm separately. Our investigations in subsection 1.4 further show why RGGs on the torus become equivalent to nongeometric models as d tends to infinity and why this is not the case if we use the hypercube instead as previously observed in [18, 13].

6.1. Equivalence for L_p-norms with $1 \le p < \infty$. Our argument builds upon a multivariate central-limit theorem similar to the one used by Devroye et al., who establish a similar statement for SRGGs [17] .

Before starting the proof, we introduce some necessary auxiliary statements. Our argumentation builds upon the following Berry–Esseen theorem introduced in [38].

THEOREM 6.1 (Theorem 1.1 in [38]). *Let $Z_1, \dots, Z_d$ be independent zero-mean random variables taking values in $\mathbb{R}^m$. Let further $Z := \sum_{i=1}^d Z_i$ and assume that the covariance matrix of Z is the identity matrix. Let $X \in \mathbb{R}^m$ be a random variable following the standard m-variate normal distribution $\mathcal{N}(0, I)$. Then for any convex set $A \subseteq \mathbb{R}^m$, we have*

$$|\Pr[Z \in A] - \Pr[X \in A]| \le (42d^{1/4} + 16) \sum_{i=1}^d \mathbb{E}\left[\|Z_i\|^3\right],$$

where $\|Z_i\|$ is the L_2-norm of Z_i.

This illustrates that for $d \to \infty$, the (random) distance between two vertices behaves like a Gaussian random variable. Throughout this section, we use the following notation. For any $u, v \in V$, we define $\Delta_{(u,v)} \in \mathbb{R}^d$ as the componentwise distance of u and v, i.e., $\Delta_{(u,v),i} := |\mathbf{x}_{ui} - \mathbf{x}_{vi}|_C = \min\{|\mathbf{x}_{ui} - \mathbf{x}_{vi}|, 1 - |\mathbf{x}_{ui} - \mathbf{x}_{vi}|\}$. Recall that t_{uv} is the connection threshold of the vertices u, v, i.e., u, v are adjacent if and only if their distance is at most t_{uv}. We may express

$$\Pr[u \sim v] = \Pr\left[\left\|\Delta_{(u,v)}\right\|_p \le t_{uv}\right] = \Pr\left[\sum_{i=1}^d \Delta_{(u,v),i}^p \le t_{uv}^p\right]$$

and we further note that $\Delta_{(u,v),i}$ and $\Delta_{(u,v),j}$ are i.i.d. random variables. Define

$$\mu := \mathbb{E}\left[\Delta_{(u,v),i}^p\right] \qquad \text{and} \qquad \sigma^2 := \mathrm{Var}\left[\Delta_{(u,v),i}^p\right],$$

and let the random variable $Z_{(u,v),i}$ be defined as

$$Z_{(u,v),i} := \frac{\Delta_{(u,v),i}^p - \mu}{\sqrt{d}\sigma}.$$

Now define $Z_{(u,v)} := \sum_{i=1}^d Z_{(u,v),i}$ and observe that this allows us to express

$$\Pr[u \sim v] = \Pr\left[\sum_{i=1}^d \Delta_{(u,v),i}^p \le t_{uv}^p\right] = \Pr\left[Z_{(u,v)} \le \frac{t_{uv}^p - d\mu}{\sqrt{d}\sigma}\right].$$

Working with $Z_{(u,v)}$ instead of $\Delta_{(u,v)}$ has the advantage that we have $\mathbb{E}\left[Z_{(u,v)}\right] = 0$ and

$$\mathrm{Var}\left[Z_{(u,v)}\right] = \sum_{i=1}^d \mathrm{Var}\left[Z_{(u,v),i}\right] = \frac{1}{d\sigma^2} \sum_{i=1}^d \mathrm{Var}\left[\Delta_{(u,v),i}^p\right] = 1.$$

These properties are useful when applying Theorem 6.1. Now recall that t_{uv} is defined so that the marginal connection probability $\Pr[u \sim v]$ is equal to $\min\{1, \frac{\lambda w_u w_v}{n}\}$, which is required in order to ensure that $\mathbb{E}[\deg(v)] \propto w_v$ for all v. We use this to establish the following lemma describing the asymptotic behavior of the threshold t_{uv}.

LEMMA 6.2. *Let* $G = \mathrm{GIRG}(n, \beta, w_0, d) = (V, E)$ *be a standard GIRG with* L_p*-norm for* $p \in [1, \infty)$. *Denote by* Φ *the cumulative density function of the standard Gaussian distribution, i.e.,* $\Phi(x) = \sqrt{2\pi}\int_{-\infty}^{x} e^{-t^2/2}dt$. *Then for any* $u, v \in V$ *with* $u \neq v$ *and* $\frac{\lambda w_u w_v}{n} < 1$, *we have*

$$\lim_{d\to\infty} \frac{t_{uv}^p - d\mu}{\sqrt{d}\sigma} = \Phi^{-1}\left(\frac{\lambda w_u w_v}{n}\right).$$

Proof. Let u, v be fixed. In the remainder of this proof, we abbreviate $Z_{(u,v)}$ with Z, and $Z_{(u,v),i}$ with Z_i. For every $c \in \mathbb{R}$, define the set $A_c = \{x \in \mathbb{R} \mid x \leq c\}$. Let further $X \sim \mathcal{N}(0, 1)$ be a standard Gaussian random variable. We get from Theorem 6.1 that

$$\begin{aligned} |\Pr[Z \leq c] - \Pr[X \leq c]| &\leq (42d^{1/4} + 16)\sum_{i=1}^{d} \mathbb{E}\left[|Z_i|^3\right] \\ &= \frac{42d^{1/4} + 16}{d^{3/2}\sigma^3}\sum_{i=1}^{d} \mathbb{E}\left[\left|\Delta_{(u,v),i}^p - \mu\right|^3\right] \\ &\leq \frac{42d^{1/4} + 16}{d^{1/2}\sigma^3} && (6.1) \\ &= o_d(1) && (6.2) \end{aligned}$$

because $\Delta_{(u,v),i}^p - \mu \in [-1, 1]$, which shows that Z converges to a standard Gaussian random variable as $d \to \infty$. In particular, this statement is true for $c = \Phi^{-1}(\frac{\lambda w_u w_v}{n})$. At the same time,

$$\Pr\left[X \leq \Phi^{-1}\left(\frac{\lambda w_u w_v}{n}\right)\right] = \frac{\lambda w_u w_v}{n} = \Pr\left[Z \leq \frac{t_{uv}^p - d\mu}{\sqrt{d}\sigma}\right],$$

where the second step follows by the definition of t_{uv} and Z. Hence, by (6.1),

$$\lim_{d\to\infty}\left|\Pr\left[Z \leq \Phi^{-1}\left(\frac{\lambda w_u w_v}{n}\right)\right] - \Pr\left[Z \leq \frac{t_{uv}^p - d\mu}{\sqrt{d}\sigma}\right]\right| = 0.$$

Since the function $f(c) = \Pr[Z \leq c]$ converges to the cumulative density function Φ of the standard Gaussian distribution and since this function is continuous and strictly monotonically increasing, we infer that

$$\lim_{d\to\infty} \frac{t_{uv}^p - d\mu}{\sqrt{d}\sigma} = \Phi^{-1}\left(\frac{\lambda w_u w_v}{n}\right). \qquad \square$$

With this, we prove the main theorem of this section.

Proof of Theorem 1.1 *for* $p \in [1, \infty)$. As n is fixed, the set $\mathcal{G}(n)$ is finite and so it suffices to show that for all $H \in \mathcal{G}(n)$, we have

$$\lim_{d\to\infty} \Pr[G_{\mathrm{GIRG}} = H] = \Pr[G_{\mathrm{IRG}} = H]. \tag{6.3}$$

First, we note that for any $u, v \in V$ with $\frac{\lambda w_u w_v}{n} \geq 1$, u and v are guaranteed to be connected in both G_{GIRG} and G_{IRG}. Hence, for every $H \in \mathcal{G}(n)$ in which u and v are not connected, we get $\Pr[G_{\mathrm{GIRG}} = H] = \Pr[G_{\mathrm{IRG}} = H] = 0$. For this reason, it suffices to show (6.3) for all $H \in \mathcal{G}(n)$ in which all $u, v \in V$ with $u \neq v$ and $\frac{\lambda w_u w_v}{n} \geq 1$ are connected.

Let H be an arbitrary but fixed such graph. We define the set $\mathcal{Q} = \{(u,v) \mid 1 \le u < v \le n, \frac{\lambda w_u w_v}{n} < 1\}$, which contains all pairs of vertices that are not connected with probability 1. For any event $\mathbf{E}$, we denote by $\mathbb{1}(\mathbf{E})$ the random variable that is 1 if $\mathbf{E}$ occurs and 0 otherwise. Similarly, we denote by $H_{(u,v)}$ an indicator variable that is 1 if the edge $\{u,v\}$ is present in H and 0 otherwise. Furthermore, for every $(u,v) \in \mathcal{Q}$, we define $\mathcal{N}_{(u,v)}$ to be an independent standard Gaussian random variable. Then,

$$\Pr\left[G_{\mathrm{IRG}} = H\right] = \Pr\left[\bigcap_{(u,v)\in\mathcal{Q}} \left(\mathbb{1}\left(\mathcal{N}_{(u,v)} \le \Phi^{-1}\left(\frac{\lambda w_u w_v}{n}\right)\right) = H_{(u,v)}\right)\right].$$

Furthermore, recall the definition of $Z_{(u,v)}$ and observe

$$\Pr\left[G_{\mathrm{GIRG}} = H\right] = \Pr\left[\bigcap_{(u,v)\in\mathcal{Q}} \left(\mathbb{1}\left(Z_{(u,v)} \le \frac{t_{uv}^p - d\mu}{\sqrt{d}\sigma}\right) = H_{(u,v)}\right)\right].$$

In addition, we define the random graph $\tilde{G}$ in which all $u \ne v \in V$ with $\frac{\lambda w_u w_v}{n} \ge 1$ are guaranteed to be connected, and in which for every $(u,v) \in \mathcal{Q}$, the edge $\{u,v\}$ is present if and only if $Z_{(u,v)} \le \Phi^{-1}(\frac{\lambda w_u w_v}{n})$. Accordingly,

$$\Pr\left[\tilde{G} = H\right] = \Pr\left[\bigcap_{(u,v)\in\mathcal{Q}} \left(\mathbb{1}\left(Z_{(u,v)} \le \Phi^{-1}\left(\frac{\lambda w_u w_v}{n}\right)\right) = H_{(u,v)}\right)\right].$$

From Lemma 6.2, we get that

$$\lim_{d\to\infty} \frac{t_{uv}^p - d\mu}{\sqrt{d}\sigma} = \Phi^{-1}\left(\frac{\lambda w_u w_v}{N}\right),$$

and so,

$$\lim_{d\to\infty} \left|\Pr\left[G_{\mathrm{GIRG}} = H\right] - \Pr\left[\tilde{G} = H\right]\right| = 0.$$

It therefore only remains to show $\lim_{d\to\infty} \Pr[\tilde{G} = H] = \Pr[G_{\mathrm{IRG}} = H]$.

For this, we let $m = |\mathcal{Q}|$ and we define the random vector $Z_i \in \mathbb{R}^m$ that has the random variables $Z_{(u,v),i}$ as its components for all $(u,v) \in \mathcal{Q}$. We further define $Z := \sum_{i=1}^d Z_i$. We use Theorem 6.1 and define the set $A \subseteq \mathbb{R}^m$ such that

$$x \in A \Leftrightarrow \forall (u,v) \in \mathcal{Q}: \begin{cases} x_{(u,v)} \le \Phi^{-1}\left(\frac{\lambda w_u w_v}{n}\right) & \text{if } H_{(u,v)} = 1, \\ x_{(u,v)} > \Phi^{-1}\left(\frac{\lambda w_u w_v}{n}\right) & \text{otherwise.} \end{cases}$$

It is easy to observe that A is convex. We get $\Pr[\tilde{G} = H] = \Pr[Z \in A]$ and $\Pr[G_{\mathrm{IRG}} = H] = \Pr[X \in A]$, where X is a random variable from the standard m-variate normal distribution.

We further note that for all $(u,v) \in \mathcal{Q}$, the random variables $\Delta_{(u,v),1}, \dots, \Delta_{(u,v),d}$ are independent, which implies that $Z_1, \dots, Z_d$ are independent as well. Furthermore, they have expectation 0 and for all $1 \le i,j \le d, (u,v),(s,t) \in \mathcal{Q}$ with $(u,v) \ne (s,t)$, the random variables $\Delta_{(u,v),i}$ and $\Delta_{st,j}$ are independent, even if $i = j$ and $\{u,v\} \cap \{s,t\} \ne \emptyset$ (because the torus is a homogeneous space). This implies that also $Z_{(u,v),i}$ and $Z_{(s,t),j}$ as well as $Z_{(u,v)}$ and $Z_{(s,t)}$ are independent. Hence, $\mathrm{Cov}\left[Z_{(u,v)}, Z_{(s,t)}\right] = 0$. Accordingly, the covariance matrix of Z is the identity matrix. Thus, Theorem 6.1 implies

$$
\begin{aligned}
|\Pr[Z \in A] - \Pr[X \in A]| &\le (42d^{1/4}+16)\sum_{i=1}^{d}\mathbb{E}\left[\|Z_i\|^3\right]\\
&= (42d^{1/4}+16)\sum_{i=1}^{d}\mathbb{E}\left[\left(\sum_{(u,v)\in\mathcal{Q}} Z_{(u,v),i}^2\right)^{3/2}\right]\\
&= (42d^{1/4}+16)\sum_{i=1}^{d}\frac{1}{d^{3/2}\sigma^3}\mathbb{E}\left[\left(\sum_{(u,v)\in\mathcal{Q}}\left(\Delta_{(u,v),i}^{p}-\mu\right)^2\right)^{3/2}\right]\\
&\le (42d^{1/4}+16)\sum_{i=1}^{d}\frac{m^{3/2}}{d^{3/2}\sigma^3}\\
&= (42d^{1/4}+16)\frac{dm^{3/2}}{d^{3/2}\sigma^3}\\
&= \frac{m^{3/2}}{\sigma^3}\left(\frac{42}{d^{1/4}}+\frac{16}{d^{1/2}}\right) = o_d(1),
\end{aligned}
$$

as $\Delta_{(u,v),i}^{p} - \mu \in [-1,1]$ and $m \le \binom{n}{2}$. This shows

$$
\lim_{d\to\infty}\Pr\left[\tilde{G} = H\right] = \Pr[G_{\mathrm{IRG}} = H],
$$

as desired. ∎

As mentioned before, the above result helps in getting an intuition for how the choice of the underlying ground space of geometric random graphs affects the impact of an increasing dimensionality. Recall that RGGs on the hypercube do *not* converge to Erdős–Rényi graphs as n is fixed and $d \to \infty$ [13, 18]. However, our results imply that they do when choosing the torus as ground space. These apparent disagreements are despite the fact that we apply the central limit theorem similarly.

As discussed before, the above proof relies on the fact that, for independent zero-mean variables $Z_1, \ldots, Z_d$, the covariance matrix of the random vector $Z = \sum_{i=1}^{d} Z_i$ is the identity matrix. This is due to the fact that the torus is a homogeneous space, implying that the probability measure of a ball of radius r is the same, regardless of where this ball is centered. This makes the random variables $Z_{(u,v)}$ and $Z_{(u,s)}$ independent. As a result their covariance is 0 although both "depend" on the position of u.

For the hypercube, this is not the case. Although the distance of two vertices can analogously be defined as a sum of independent, zero-mean random vectors over all dimensions just like we do above (the only difference being that $\Delta_{(u,v),i}$ is now the distance between u, v in dimension i in the hypercube, leading to different values of μ and σ^2), the random variables $Z_{(u,v)}$ and $Z_{(u,s)}$ do *not* have a covariance of 0.

In fact, one can verify that for every $1 \le i \le d$, there is a slightly positive covariance between $\Delta_{(u,v),i}$ and $\Delta_{(u,s),i}$ (equal to 1/180). This transfers to a covariance between $Z_{(u,v)}$ and $Z_{(u,s)}$, which stays constant as d grows, since

$$
\begin{aligned}
\mathrm{Cov}[Z_{(u,v)}, Z_{(u,s)}] &= \mathbb{E}\left[Z_{(u,v)} \cdot Z_{(u,s)}\right]\\
&= \sum_{i=1}^{d}\sum_{j=1}^{d}\mathbb{E}\left[Z_{(u,v),i} \cdot Z_{(u,s),j}\right]
\end{aligned}
$$

$$= \frac{1}{d\sigma^2} \sum_{i=1}^{d} \mathrm{Cov}[\Delta^p_{(u,v),i} \Delta^p_{(u,s),i}]$$
$$= \frac{\mathrm{Cov}[\Delta^p_{(u,v),1} \Delta^p_{(u,s),1}]}{\sigma^2},$$

where we used that $\mathbb{E}\left[Z_{(u,v),i} \cdot Z_{(u,s),j}\right] = 0$ if $i \neq j$. Accordingly, the covariance matrix Σ of Z is not the identity matrix. Nevertheless, our proof from the previous section still applies if we replace Z by $Y = \sum_{i=1}^{d} \Sigma^{-1} Z_i$. Now Y is the sum of independent random vectors and has the identity matrix as its covariance matrix, so Theorem 6.1 remains applicable. Furthermore, $\mathbb{E}\left[||\Sigma^{-1} Z_i||^3\right]$ is still proportional to $d^{-3/2}$ and thus remains bounded such that Y converges to a standard m-dimensional normal vector. This implies that Z converges to a random vector from $\mathcal{N}(0, \Sigma)$, showing that RGGs on the hypercube converge to a nongeometric model where the probability that any fixed graph is sampled can be described, like above, as the probability that $Z \sim \mathcal{N}(0, \Sigma)$ falls into the convex set A. In this model, however, the edges are not drawn independently, as Σ is not the identity matrix. In fact, for any three vertices s, u, v, the components $Z_{(u,v)}$ and $Z_{(u,s)}$ are slightly positively correlated, so there is a slightly higher probability that s, u, v form a triangle than in a corresponding Erdős–Rényi graph. This leads to a higher tendency to form cliques, which is in accordance with the observations from Erba et al. [18].

6.2. Asymptotic equivalence for L_∞-norm. In this section, we prove that our model also loses its geometry if L_∞-norm is used. We use a different technique to prove this theorem, as the L_∞-distance between two vertices is no longer a sum of independent random variables and central limit theorems no longer apply. Instead our argument builds upon the bounds we establish in subsection 4.1.

Proof of Theorem 1.1 *for L_∞-norm.* We show that for all $H \in \mathcal{G}(n)$,

$$\lim_{d\to\infty} \Pr[G_{\mathrm{GIRG}} = H] = \Pr[G_{\mathrm{IRG}} = H]. \tag{6.4}$$

We start by establishing a way to compute $\Pr[G = H]$ for any random variable G representing a distribution over all graphs in $\mathcal{G}(n)$. For this, we denote by $E(H)$ the set of edges of a graph $H = (V, E) \in \mathcal{G}(n)$. We further let $\binom{V}{2}$ be the set of all possible edges on the vertex set V. Now, for any $H \in \mathcal{G}$, we have

$$\Pr[G = H] = \Pr[E(G) \supseteq E(H)] - \sum_{E(H) \subset \mathcal{A} \subseteq \binom{V}{2}} \Pr[E(G) = \mathcal{A}].$$

That is, we may express the probability that G is sampled as the probability that a supergraph of G is sampled minus the probability that any proper supergraph of G is sampled. Now, for any $E(H) \subset \mathcal{A} \subseteq \binom{V}{2}$, the probability $\Pr[E(G) = \mathcal{A}]$ is the probability that G is a specific graph with at least $|E(H)| + 1$ edges. Now, we may repeatedly substitute terms of the form $\Pr[E(G) = \mathcal{A}]$ in the same way until we have an (alternating) sum consisting only of terms that have the form $\Pr[E(G) \supseteq \mathcal{A}]$ for some $E(H) \subset \mathcal{A} \subseteq \binom{V}{2}$. That is, we may calculate the probability $\Pr[G = H]$ even if we only know $\Pr[E(G) \supseteq \mathcal{A}]$ for any $\mathcal{A} \subseteq \binom{V}{2}$.

As n is fixed, in order to prove our statement in (6.4), it suffices to prove that, for each $\mathcal{A} \subseteq \binom{V}{2}$, we have

$$\lim_{d\to\infty} \Pr[E(G_{\mathrm{GIRG}}) \supseteq \mathcal{A}] = \Pr[E(G_{\mathrm{IRG}}) \supseteq \mathcal{A}] = \prod_{\{i,j\} \in \mathcal{A}} \frac{\kappa_{ij}}{n}.$$

Using Theorem 1.5, we get that

$$\Pr\left[E(G_{\mathrm{GIRG}})\supseteq\mathcal{A}\right]=\prod_{\{i,j\}\in\mathcal{A}}\left(\frac{\kappa_{ij}}{n}\right)^{1\mp\mathcal{O}_d\left(\frac{\ln(n)}{d}\right)}.$$

For $d\to\infty$, this clearly converges to $\Pr\left[E(G_{\mathrm{IRG}})\supseteq\mathcal{A}\right]$, and the proof is finished. ∎

REFERENCES

[1] W. Aiello, F. Chung, and L. Lu, *A random graph model for power law graphs*, Exp. Math., 10 (2001), pp. 53–66, https://doi.org/10.1080/10586458.2001.10504428.

[2] P. Almagro, M. Boguñá, and M. A. Serrano, *Detecting the ultra low dimensionality of real networks*, Nat. Commun., 13 (2022), 6096, https://doi.org/10.1038/s41467-022-33685-z.

[3] I. Artico, I. E. Smolyarenko, V. Vinciotti, and E. C. Wit, *How rare are power-law networks really?*, Proc. Roy. Soc. A, 476 (2020), 20190742, https://doi.org/10.1098/rspa.2019.0742.

[4] K. E. Avrachenkov and A. V. Bobu, *Cliques in high-dimensional random geometric graphs*, Appl. Network Sci., 5 (2020), pp. 1–24, https://doi.org/10.1007/s41109-020-00335-6.

[5] T. Bläsius and P. Fischbeck, *On the external validity of average-case analyses of graph algorithms*, in 30th Annual European Symposium on Algorithms (ESA 2022), 2022, 21, https://doi.org/10.4230/LIPIcs.ESA.2022.21.

[6] T. Bläsius, T. Friedrich, M. Katzmann, U. Meyer, M. Penschuck, and C. Weyand, *Efficiently generating geometric inhomogeneous and hyperbolic random graphs*, Network Sci., 10 (2022), pp. 361–380, https://doi.org/10.1017/nws.2022.32.

[7] T. Bläsius, T. Friedrich, and A. Krohmer, *Cliques in hyperbolic random graphs*, Algorithmica, 80 (2018), pp. 2324–2344, https://doi.org/10.1007/s00453-017-0323-3.

[8] A. Bonato, D. F. Gleich, M. Kim, D. Mitsche, P. Prałat, Y. Tian, and S. J. Young, *Dimensionality of social networks using motifs and eigenvalues*, PLOS ONE, 9 (2014), pp. 1–7, https://doi.org/10.1371/journal.pone.0106052.

[9] M. Brennan, G. Bresler, and D. Nagaraj, *Phase Transitions for Detecting Latent Geometry in Random Graphs*, https://arxiv.org/abs/1910.14167, 2020.

[10] S. Bubeck, J. Ding, R. Eldan, and M. Rácz, *Testing for High-Dimensional Geometry in Random Graphs*, https://arxiv.org/abs/1411.5713, 2015.

[11] F. Chung and L. Lu, *The average distances in random graphs with given expected degrees*, Proc. Natl. Acad. Sci. USA, 99 (2002), pp. 15879–15882, https://doi.org/10.1073/pnas.252631999.

[12] F. Chung and L. Lu, *Connected components in random graphs with given expected degree sequences*, Ann. Comb., 6 (2002), pp. 125–145, https://doi.org/10.1007/PL00012580.

[13] J. Dall and M. Christensen, *Random geometric graphs*, Phys. Rev. E, 66 (2002), 016121, https://doi.org/10.1103/PhysRevE.66.016121.

[14] F. Daly, A. Haig, and S. Shneer, *Asymptotics for Cliques in Scale-Free Random Graphs*, https://arxiv.org/abs/2008.11557, 2020.

[15] L. Daqing, K. Kosmidis, A. Bunde, and S. Havlin, *Dimension of spatially embedded networks*, Nat. Phys., 7 (2011), pp. 481–484, https://doi.org/10.1038/nphys1932.

[16] L. Decreusefond, E. Ferraz, H. Randriambololona, and A. Vergne, *Simplicial homology of random configurations*, Adv. Appl. Probab., 46 (2014), pp. 325–347, https://doi.org/10.1239/aap/1401369697.

[17] L. Devroye, A. György, G. Lugosi, and F. Udina, *High-dimensional random geometric graphs and their clique number*, Electron. J. Probab., 16 (2011), pp. 2481–2508, https://doi.org/10.1214/EJP.v16-967.

[18] V. Erba, S. Ariosto, M. Gherardi, and P. Rotondo, *Random geometric graphs in high dimension*, Phys. Rev. E, 102 (2020), 012306, https://doi.org/10.1103/PhysRevE.102.012306.

[19] P. Erdös and A. Rényi, *On random graphs* I, Publ. Math. Debrecen, 6 (1959), pp. 290–297.

[20] N. Fountoulakis and T. Müller, *Law of large numbers for the largest component in a hyperbolic model of complex networks*, Ann. Appl. Probab., 28 (2018), pp. 607–650, https://doi.org/10.1214/17-AAP1314.

[21] N. Fountoulakis, P. van der Hoorn, T. Müller, and M. Schepers, *Clustering in a hyperbolic model of complex networks*, Electron. J. Probab., 26 (2021), pp. 1–132, https://doi.org/10.1214/21-EJP583.

[22] T. Friedrich, A. Göbel, M. Katzmann, and L. Schiller, *Cliques in High-Dimensional Geometric Inhomogeneous Random Graphs*, https://doi.org/10.48550/arXiv.2302.04113, 2023.
[23] T. Friedrich and A. Krohmer, *On the diameter of hyperbolic random graphs*, SIAM J. Discrete Math., 32 (2018), pp. 1314–1334, https://doi.org/10.1137/17M1123961.
[24] E. N. Gilbert, *Random graphs*, Ann. Math. Statist., 30 (1959), pp. 1141–1144.
[25] W. Gu, A. Tandon, Y.-Y. Ahn, and F. Radicchi, *Principled approach to the selection of the embedding dimension of networks*, Nat. Commun., 12 (2021), 3772, https://doi.org/10.1038/s41467-021-23795-5.
[26] L. Gugelmann, K. Panagiotou, and U. Peter, *Random hyperbolic graphs: Degree sequence and clustering*, in 39th International Colloquium on Automata, Languages, and Programming (ICALP), 2012, pp. 573–585, https://doi.org/10.1007/978-3-642-31585-5_51.
[27] R. V. d. Hofstad, J. S. H. V. Leeuwaarden, and C. Stegehuis, *Optimal subgraph structures in scale-free configuration models*, Ann. Appl. Probab., 31 (2021), pp. 501–537, https://doi.org/10.1214/20-AAP1580.
[28] S. Janson, T. Łuczak, and I. Norros, *Large cliques in a power-law random graph*, J. Appl. Probab., 47 (2010), pp. 1124–1135, https://doi.org/10.1239/jap/1294170524.
[29] R. Keusch, *Geometric Inhomogeneous Random Graphs and Graph Coloring Games*, Doctoral thesis, ETH Zurich, 2018, https://doi.org/10.3929/ethz-b-000269658.
[30] C. Koch and J. Lengler, *Bootstrap percolation on geometric inhomogeneous random graphs*, in 43rd International Colloquium on Automata, Languages, and Programming (ICALP 2016), 2016, 147, https://doi.org/10.4230/LIPIcs.ICALP.2016.147.
[31] D. Krioukov, F. Papadopoulos, M. Kitsak, A. Vahdat, and M. Boguñá, *Hyperbolic geometry of complex networks*, Phys. Rev. E, 82 (2010), 036106, https://doi.org/10.1103/PhysRevE.82.036106.
[32] E. Levina and P. J. Bickel, *Maximum likelihood estimation of intrinsic dimension*, in Proceedings of the 17th International Conference on Neural Information Processing Systems, 2004, pp. 777–784.
[33] S. Liu, S. Mohanty, T. Schramm, and E. Yang, *Testing Thresholds for High-Dimensional Sparse Random Geometric Graphs*, https://arxiv.org/abs/2111.11316, 2021.
[34] S. Liu and M. Z. Racz, *Phase Transition in Noisy High-Dimensional Random Geometric Graphs*, https://arxiv.org/abs/2103.15249, 2021.
[35] R. Michielan and C. Stegehuis, *Cliques in geometric inhomogeneous random graphs*, J. Complex Netw., 10 (2022), cnac002, https://doi.org/10.1093/comnet/cnac002.
[36] M. Newman, *Clustering and preferential attachment in growing networks*, Phys. Rev. E, 64 (2001), 025102, https://doi.org/10.1103/PhysRevE.64.025102.
[37] M. Penrose, *Random Geometric Graphs*, Oxford University Press, 2003.
[38] M. Raič, *A multivariate Berry–Esseen theorem with explicit constants*, Bernoulli, 25 (2019), pp. 2824–2853, https://doi.org/10.3150/18-BEJ1072.
[39] M. Serafino, G. Cimini, A. Maritan, A. Rinaldo, S. Suweis, J. R. Banavar, and G. Caldarelli, *True scale-free networks hidden by finite size effects*, Proc. Natl. Acad. Sci. USA, 118 (2021), e2013825118, https://doi.org/10.1073/pnas.2013825118.
[40] I. Voitalov, P. van der Hoorn, R. van der Hofstad, and D. Krioukov, *Scale-free networks well done*, Phys. Rev. Res., 1 (2019), 033034, https://doi.org/10.1103/PhysRevResearch.1.033034.
[41] D. J. Watts and S. H. Strogatz, *Collective dynamics of 'small-world' networks*, Nature, 393 (1998), pp. 440–442, https://doi.org/10.1038/30918.
[42] W. Yang and D. Rideout, *High dimensional hyperbolic geometry of complex networks*, Mathematics, 8 (2020), 1861, https://doi.org/10.3390/math8111861.