Regular graphs with linearly many triangles

Pim van der Hoorn\(^1\), Gabor Lippner\(^2\), and Elchanan Mossel\(^3,4\)

\(^1\)Northeastern University, Department of Physics
\(^2\)Northeastern University, Department of Mathematics
\(^3\)MIT, Department of Mathematics
\(^4\)MIT Institute for Data Systems and Society

April 5, 2019

Abstract

We compute the leading asymptotics of the probability that a random \(d\)-regular graph has linearly many triangles. We also show that such graphs typically consist of many disjoint \(d + 1\)-cliques and an almost triangle-free part.

1 Introduction

What is the probability that a random graph has a lot more triangles than expected? This is a typical question in the field of large deviations, the theory that studies the tail behavior of random variables or, stated differently, the behavior of random objects conditioned on a parameter being far from its expectation. For example, one of the earliest results of this flavor, Cramér’s Theorem states that for iid variables \(X_1, X_2, \ldots\) there exists a “rate function” \(I(x)\) depending on the distribution of \(X\) such that

\[ P \left( \sum_{i=1}^{N} X_i \geq Nx \right) \approx e^{-N \cdot I(x)}. \]

In random graphs, the question about the upper tail for triangles in \(G(n, p)\) has been long studied for a constant factor of deviation from the mean \([8]\). More precisely, let \(t(G(n, p))\) denote the triangle density in the Erdős-Rényi random graph, normalized so that \(E[t(G(n, p))] = p^3\). One would like to understand the asymptotic behavior of

\[ r(n, p, \delta) = -\log P \left( t(G(n, p)) > (1 + \delta)p^3 \right) \]

The dense case \((p = o(1))\) has been reduced to an analytic variational problem by Chatterjee and Varadhan \([5]\) using methods from graph limits. However, the solution of this variational problem is only known in certain parameter ranges (see \([15]\) for details). In the sparse \((p = o(1))\) regime the asymptotics \(r(n, p, \delta) \approx n^{3/2} p^2 \log(1/p)\) have been determined in a long series of papers by many authors. \([20, 13, 9, 7, 3, 6]\). Finally, the variational methods were extended to (part of) the sparse regime in \([4]\) and using this, Lubetzky and Zhao \([16]\) found the exact asymptotics of \(r(n, p, \delta)\) in the \(n^{-1/42} \log n \leq p \ll 1\) range.

In the case of random regular graphs \(G_d(n)\), much less is known. Kim, Sudakov, and Vu \([12]\) obtained that the distribution of small subgraphs of \(G_d(n)\) is asymptotically Poisson in the sparse case, implying an asymptotic formula for the tail probability \(P \left( T(G_d(n)) > (1 + \delta)E[T(G_d(n))] \right)\), where \(T(G)\) denotes the number of triangles in the graph \(G\).
1.1 Maximum entropy random graphs with triangles

In this note we are interested in the more extreme tail probability \( P(T(G_d(n)) > tn/3) \). The reason for analyzing this tail probability stems from a related problem of finding random graph models that maximize entropy under specific constraints.

Let \( P_n \) be some probability distribution on the set \( G(n) \) of graphs on \( n \) labeled nodes. Then the entropy of \( P_n \) is defined as

\[
\mathcal{E}[P_n] = \sum_{G \in G(n)} -P_n(G) \log (P_n(G_n)) .
\]

(1)

Now let \( G^*(n) \) denote the set of graphs on \( n \) labeled nodes with some additional properties, for instance specified edge or triangle densities. Then, in order to study the structure of “typical” graphs with these constraints, one wants to find the uniform distribution on \( G^*(n) \). This corresponds to finding the distribution \( P_n^* \) that maximizes the entropy \( \mathcal{E}[P_n]\) subject to the constraint that \( P_n^* = 0 \) outside \( G^*(n) \).

It turns out that in many cases, computing the rate function also comes down to solving an optimization problem involving entropy. For example, Chatterjee and Dembo [4] showed that, up to lower order terms, the rate function corresponding to the large deviation result for subgraph counting can be expressed as the solution to a specific optimization problem. For large deviations of triangles, let \( \mathcal{G}_n \) denote the set of undirected graphs on \( n \) nodes with edges weights \( g_{ij} \in [0,1] \), then the rate function is obtained, up to lower order terms, as

\[
r(n,p,\delta) = \inf \left\{ I_p(G) : G \in \mathcal{G}_n, t(G) > (1 + \delta)p^3 \right\} .
\]

where \( t(G) = n^{-3} \sum_{1 \leq i,j,k \leq n} g_{ij}g_{jk}g_{ki} \) and \( I_p(G) \) is the so-called relative entropy of the weighted graph \( G \)

\[
I_p(G) = \sum_{1 \leq i,j \leq n} g_{ij} \log \frac{g_{ij}}{p} + (1 - g_{ij}) \log \frac{1 - g_{ij}}{1 - p}.
\]

In the case of dense graphs, such optimizations problems can be used to establish structural results for constraint random graphs. In the case of edge and triangle densities, a collection of research by Kenyon, Radin and co-authors [10] showed that the limits of dense maximal entropy random graphs with given edge and triangle densities have a bipodal structure. This means that the graph is split into two components with specific inter- and intra-component connection probabilities.

Recently, some techniques have been extended to solve the problem of finding maximum entropy sparse graphs with a given power-law degree distribution [19]. However, the degree distribution is a relatively global characteristic and hence is not expected to influence graph structures that much. A natural extension of this problem is therefore to include a constraint related to triangles, try to find the corresponding maximum entropy solutions and see what this tells us about the structure of such graphs. A key motivation for this kind of question is the work by Krioukov [14], which hinted to the fact that triangle constraints might enforce the resulting maximum entropy solution to have some geometric component.

1.2 Results

Motivated by the question “can local triangle constraints induce global (geometric) behavior?”, we study the number of random regular graph \( G_d(n) \) conditioned on having linearly many triangles. As explained in the previous section, our setting is related to the entropy maximization problem with local and global constraints, i.e. where each node must have degree exactly \( d \) and must be incident to at least \( t \) triangles on average.

Let \( \mathcal{G}_{d,t}(n) \) denote the set of \( d \)-regular graphs on \( n \) labeled nodes that contain at least \( nt/3 \) triangles - or equivalently, where the nodes are incident to at least \( t \) triangles on average. It is clear that a node can be incident to no more than \( \binom{d}{3} \) triangles, so \( \mathcal{G}_{d,t}(n) \) is empty unless \( 0 \leq t \leq \binom{d}{2} \). We compute the leading asymptotics of \( |\mathcal{G}_{d,t}(n)| \) for fixed \( d, t \), as \( n \to \infty \), and provide a structural description of a “typical” element of \( \mathcal{G}_{d,t}(n) \). We then extend these results to case of \( k \)-cliques in \( d \)-regular graphs.
1.2.1 Number of \( d \)-regular graph with many triangles

**Theorem 1.1.** For fixed \( d \) and fixed \( 0 < t < \binom{d}{2} \) we have

\[
\left| \frac{\log |\mathcal{G}_{d,t}(n)|}{n \log n} - \left( \frac{d}{2} - \frac{t}{d+1} \right) \right| = O(1/\log n)
\]

The part \((d/2) n \log n\) is related to \(\log |\mathcal{G}_d(n)|\). In particular, using the results in [2], one can show that

\[
\lim_{n \to \infty} \frac{\log |\mathcal{G}_d(n)|}{n \log n} = \frac{d}{2}
\]

Since \(\mathbb{P}(T(\mathcal{G}_d(n)) > tn/3) = |\mathcal{G}_{d,1}(n)|/|\mathcal{G}_d(n)|\), we obtain the following result for this tail probability from Theorem 1.1:

**Corollary 1.2.** For fixed \( d \) and fixed \( 0 < t < \binom{d}{2} \) we have

\[
\lim_{n \to \infty} -\log \mathbb{P}(T(\mathcal{G}_d(n)) > tn/3) = \frac{t}{d+1}\]

1.2.2 Structure of \( d \)-regular graph with many triangles

It turns out, perhaps not so surprisingly, that in most elements of \(\mathcal{G}_{d,t}(n)\), most of the triangles cluster into (disjoint) \(d+1\)-cliques. To make this statement precise, let us call a node \( k \)-bad if it is not part of a \(d+1\)-clique but it is incident to at least one triangle.

**Theorem 1.3.** Let \( d, t \) as before. With high probability a uniformly randomly chosen element of \(\mathcal{G}_{d,t}(n)\) has less than \(\frac{\log n}{\log \log n}\) bad nodes. Thus, the number of triangles that are not part of a \(d+1\)-clique is sublinear.

In Section 2.2 we prove a slightly more general result where we consider the case where a uniformly randomly chosen element of \(\mathcal{G}_{d,t}(n)\) has less than \(\varepsilon_n n\) bad nodes, with \(\varepsilon_n \to 0\), such that \(\varepsilon_n \log n \to \infty\).

Note that Theorem 1.3 hints at a graph structure similar to the bipodal case, where instead of two components, we now have a linear in \( n \) number of cliques and some remaining larger graph with a sub-linear number of triangles.

1.2.3 \( d \)-regular graph with many \( k \)-cliques

As a corollary to our methods, we also obtain similar results for regular graphs with many \( k \)-cliques. Let \(\mathcal{G}_{d,t_k,k}(n)\) denote the set of \( d \)-regular graphs on \( n \) nodes that contain at least \( t_k n/k \) subgraphs isomorphic to \(K_k\). As a natural extension of terminology, we call nodes \( k \)-bad if they are not part of a \( k \)-clique but are incident to a \( k-1 \) clique.

**Theorem 1.4.** For fixed \( d \), \( k \geq 3 \) and fixed \( 0 < t_k < \binom{k-1}{2} \) we have

\[
\left| \frac{\log |\mathcal{G}_{d,t_k,k}(n)|}{n \log n} - \left( \frac{d}{2} - \frac{t^*}{d+1} \right) \right| = O(1/\log n)
\]

where

\[
t^* = t_k \left( \frac{d-1}{k-3} \right)
\]

Furthermore, almost all elements of \(\mathcal{G}_{d,t_k,k}(n)\) will have at most \(\varepsilon n\) bad nodes.

2 Proofs

2.1 The number of regular graphs with a given number of triangles

The proof of Theorem 1.1 consist of establishing a lower and upper bound on \(\log |\mathcal{G}_{d,t}(n)|\). More precisely, we will show that

\[
\left( \frac{d}{2} - \frac{t}{d+1} \right) n \log n - O(n) \leq \log |\mathcal{G}_{d,t}(n)| \leq \left( \frac{d}{2} - \frac{t}{d+1} \right) n \log n + O(n).
\]

The theorem then follows after dividing by \(n \log n\) and letting \(n \to \infty\).
Proof of Theorem 1.2 (Lower bound). To establish a lower bound we construct a family of elements in $G_{d,t}(n)$ by letting $b = \left\lceil \frac{mn}{3} \left(\frac{d+1}{3}\right)^{-1} \right\rceil = \left\lceil \frac{2mn}{d(d+1)(d-1)} \right\rceil$ and taking $b$ disjoint $d+1$-cliques and an arbitrary $m = n - (d+1)b$ node $d$-regular graph. Clearly, these graphs will have at least $tn/3$ triangles. The asymptotic number of $d$-regular graphs on $m$ nodes is

$$\Theta\left(\frac{(md)!}{(md/2)!2^{md/2}(d!)^m}\right),$$

see [1]. So our family of elements has size asymptotically

$$\left(\frac{n}{d+1}\right)^{n-(d+1)} \cdots \left(\frac{n-(b-1)(d+1)}{d+1}\right)^{b} \frac{(md)!}{(md/2)!2^{md/2}(d!)^m}$$

Thus

$$\log |G_{d,t}(n)| \geq b(d+1) \log n - b \log b + \frac{dm}{2} \log(dm) - O(n)$$

$$= \left( bd + \frac{dm}{2} \right) \log n + O(n) = \left( \frac{dn}{2} - \frac{d(d+1)}{2} b + db \right) \log n - O(n)$$

$$= \left( \frac{dn}{2} - \frac{d(d-1)}{2} b \right) \log n - O(n) = \left( \frac{d}{2} - \frac{t}{d+1} \right) n \log n - O(n).$$

We now need to prove a matching upper bound on $|G_{d,t}(n)|$. We do this by uncovering the edges of such graphs in a suitably chosen order, and recording whether in each step a new triangle is created. We will define a function $\phi : G_{d,t}(n) \rightarrow \{0,1\}^{nd/2}$ that will record which edges of $G$ create triangles when added in this order.

More precisely, given $G \in G_{d,t}(n)$, define the lexicographic ordering $\prec$ on the set of edges of $G$ as follows. Let $e = (i_1j_1)$ and $f = (i_2j_2)$ be two edges of $G$ with $i_1 < j_1$ and $i_2 < j_2$. Let us declare $e \prec f$ if $i_1 \leq i_2$, or if $i_1 = i_2$ and $j_1 < j_2$. Let $e_1 \prec e_2 \cdots \prec e_{nd/2}$ denote the edges of $G$ in increasing lexicographic order. Let $G[k]$ denote the subgraph of $G$ consisting of $e_1, \ldots, e_k$.

Finally define $\phi(G)(k) = 1$ if $e_k$ is incident to a triangle in $G[k]$ and 0 otherwise. In other words, denoting $e_k = (ij)$, we have $\phi(G)(k) = 1$ if and only if there is a triangle $(hij)$ in $G$ such that $h < \min(i,j)$.

For any $x \in \{0,1\}^{nd/2}$ let us denote $|x| = \sum_{i=1}^{nd/2} x(j)$. Then $|\phi(G)|$ denotes the total number of edges $e_{k+1}$ that upon adding to the graph $G[k]$ have created at least one new triangle. Moreover, any vector $x \in \{0,1\}^{nd/2}$ describes a profile of which edges revealed a new triangle. The next lemma gives an upper bound on the number of graphs in $G_{d,t}(n)$ with a given triangle reveal profile.

Lemma 2.1.

$$|\phi^{-1}(x)| \leq n^{nd/2-|x|}d^{d|x|}$$

Proof. The idea is to reconstruct a $G \in \phi^{-1}(x)$ by starting from the empty graph and adding edges 1-by-1, according to the lexicographic order. The vector $x$ dictates whether the next edge added has to create a triangle with previously added edges. By the definition of the lexicographic order, and since all nodes in $G$ have to have exactly degree $d$, given $G[k-1]$, the smaller index endpoint of the next edge $e_k$ is pre-determined. It has to start at the smallest index node, $j$, that does not yet have $d$ edges incident to it. There are never more than $n$ choices for the other endpoint of this edge $e_k$. However when $x(k) = 1$, the number of choices is limited to current 2nd neighbors of $j$. There are never more than $d^2$ such nodes.

Thus, adding the edges 1-by-1, we never have more than $n$ options to choose from when $x(k) = 0$, and at most $d^2$ options when $x(k) = 1$. This amounts to the upper bound claimed.
The main idea for the upper bound is now to consider a specific set of triangle reveal profiles \( x \in \{0,1\}^{nd/2} \), in which at least \( nt/(d+1) \) edges have revealed triangles.

**Proof of Theorem 1.1 (Upper bound).** Define

\[
L = \left\{ x \in \{0,1\}^{nd/2} : |x| \geq \frac{nt}{d+1} - 1 \right\}.
\]

Then, by Lemma 2.1 we see that

\[
|\phi^{-1}(L)| \leq |L| \cdot n^{1+n\left(4 - \frac{4}{d+1}\right)} \cdot d^{nd/2} \leq n^{n\left(4 - \frac{4}{d+1}\right)} \cdot n \cdot (2d)^{nd/2}.
\]  

(2)

To finish the proof, we will show that \( |\mathcal{G}_{d,t}(n)| \leq \frac{dn}{2} \phi^{-1}(L) \) as claimed above. Combined with (2) we get

\[
\log |\mathcal{G}_{d,t}(n)| \leq \log \frac{dn}{2} + \left( \frac{d}{2} - \frac{t}{d+1} \right) n \log n + \log n + \frac{nd}{2} \log(2d)
\]

\[
= \left( \frac{d}{2} - \frac{t}{d+1} \right) n \log n + O(n).
\]

□

We are thus left to prove Lemma 2.2. For this we first show that the expected value of \( |\phi(G_{n})| \) is at least \( \frac{tn}{d+1} \). Then the lemma will follow from a standard Markov-inequality argument.

**Lemma 2.3.**

\[
E[|\phi(G_{n})|] \geq \frac{tn}{d+1}
\]

**Proof.** Let \( X_{e}(\sigma) \) be the indicator variable of the edge \( e \) of \( G \) creating a triangle when it is added in the lexicographic order of \( G_{n} \). Then \( |\phi(G_{n})| = \sum_{e} X_{e}(\sigma) \) and so

\[
E[|\phi(G_{n})|] = \sum_{e} E[X_{e}(\sigma)] = \sum_{e} P(X_{e}(\sigma) = 1)
\]

Let \( e = (ij) \) and let \( e \) be incident to exactly \( t_{e} \) triangles in \( G \). Let \( v_{1}, v_{2}, \ldots, v_{t_{e}} \) denote the third nodes of these triangles. \( X_{e}(\sigma) \) is 1 if at least one of these triangles are formed at the moment when adding \( e \), which is equivalent to at least one of these nodes preceding both \( i \) and \( j \) in the \( \sigma \)-order. That is, \( \min(\sigma(v_{1}), \sigma(v_{2}), \ldots, \sigma(v_{t_{e}})) < \min(\sigma(i), \sigma(j)) \). Then \( X_{e}(\sigma) = 0 \) if and only if either \( i \) or \( j \) has the smallest \( \sigma \) value among \( i, j, v_{1}, v_{2}, \ldots, v_{t_{e}} \). Since the \( \sigma \)-order of these nodes is a uniformly random permutation on \( t_{e} + 2 \) elements, we get

\[
P(X_{e}(\sigma) = 0) = 2/(t_{e} + 2)
\]

and hence \( P(X_{e}(\sigma) = 1) = 1 - 2/(t_{e} + 2) \).

Thus, since \( t_{e} \leq d - 1 \), we get

\[
E[|\phi(G_{n})|] = \sum_{e} P(X_{e}(\sigma) = 1) = \sum_{e} \left( 1 - \frac{2}{t_{e} + 2} \right) = \sum_{e} \frac{t_{e}}{t_{e} + 2} \geq \sum_{e} \frac{t_{e}}{d+1} \geq \frac{tn}{d+1}.
\]  

(3)

where the last inequality follows from \( \sum_{e} t_{e} \) being 3 times the total number of triangles in \( G \), which is in turn at least \( tn/3 \). This finishes the proof of the lemma. □
Proof of Lemma 2.2. By simple algebraic considerations
\[
\frac{|S_n G \cap \phi^{-1}(L)|}{|S_n G|} = \frac{|\{\sigma \in S_n : \phi(G_\sigma) \in L\}|}{|S_n|}.
\] (4)
This is obvious when $G$ has no automorphisms (that is, when $S_n G$ is in bijection with $S_n$), but it also holds in the general case since the stabilizers of different elements of the orbit $S_n G$ are conjugate and hence have the same cardinality.

Consider a uniformly random permutation $\sigma \in S_n$. By (4) it is enough to show that with probability at least $\frac nk$ we have $\phi(G_\sigma) \in L$, which is equivalent to $|\phi(G_\sigma)| \geq \frac{tn}{d+1}$. Hence, using Lemma 2.2,
\[
\frac{tn}{d+1} \leq \mathbb{E} |\phi(G_\sigma)| \leq \left(\frac{tn}{d+1} - 1\right) P\left(|\phi(G_\sigma)| < \frac{tn}{d+1} - 1\right) + \frac{dn}{2} P\left(|\phi(G_\sigma)| \geq \frac{tn}{d+1} - 1\right),
\] from which
\[
P\left(|\phi(G_\sigma)| \geq \frac{tn}{d+1} - 1\right) \geq \frac{2}{dn}.
\]

We end this section with the proof of Corollary 1.2.

Proof of Corollary 1.2. It follows from [2] Theorem 1] that $|G_d(n)|$, the number of $d$-regular graphs on $n$ nodes, satisfies, as $n \to \infty$,
\[
|G_d(n)| \sim e^{-\frac{(d-1)/2 - (d-1)^2/4}{2} - nd/2} \left(\frac{nd}{d!}\right)^n,
\]
where $(a)_b = a(a-1) \ldots (a-b+1)$ and we note that $nd$ is even so that $nd/2$ is an integer. It then follows that
\[
\log(|G_d(n)|) \sim \log \left(e^{-\frac{(d-1)/2 - (d-1)^2/4}{2} - nd/2} \left(\frac{nd}{d!}\right)^n\right)
\] (5)\[
= \log ((nd)_{nd/2}) - n \log(d) - \frac{nd \log(2)}{2} - \frac{d - 1}{2} - \frac{(d-1)^2}{4}.
\]
Clearly the leading order comes from the first term $\log ((nd)_{nd/2})$. Using Stirling’s approximation we have that, as $n \to \infty$,
\[
\log ((nd)_{nd/2}) = \log((nd)!) - \log((nd/2)!) \sim \frac{nd}{2} \log(n) \sim \frac{dn \log(n)}{2},
\]
and hence we conclude that
\[
\lim_{n \to \infty} \frac{\log(|G_d(n)|)}{n \log(n)} = \lim_{n \to \infty} \frac{\log((nd)_{nd/2})}{n \log(n)} = \frac{d}{2}.
\]
Since $P(T(G_d(n)) > tn/3) = |G_{d,t}(n)|/|G_d(n)|$, Theorem 1.1 now implies that
\[
- \log P(T(G_d(n)) > tn/3) = \frac{\log |G_d(n)|}{n \log n} - \frac{\log |G_{d,t}(n)|}{n \log n} = \frac{t}{d + 1} + O\left(\frac{1}{\log n}\right),
\]
from which the result follows. \qed
2.2 The structure of regular graphs with a given number of triangles

A simple extension of the methods of the proof of Theorem 1.3 yields a strong structural description of a typical graph with \(tn/3\) triangles: it is a collection of disjoint \(d+1\)-cliques and an almost triangle-free graph. The number of cliques, as well as the size of the triangle-free part, can be expressed as a linear-in-\(n\) function of \(t\).

Let us say that a node in \(G\) is bad if it’s not in a \(d+1\)-clique, but it is in a triangle. The following statement is a (very) slight strengthening of Theorem 1.3.

**Theorem 2.4.** Let \(\varepsilon > 0\) fixed. Among all \(d\)-regular graphs with at least \(tn/3\) triangles, the proportion of those where more than \(\varepsilon n\) nodes are bad goes to 0 as \(n \to \infty\). This remains true even if \(\varepsilon \to 0\), as long as \(\varepsilon \log n \to \infty\).

We will make use of the following simple observation.

**Lemma 2.5.** Let \(G\) be a \(d\)-regular graph. If all edges incident to a node \(v\) are incident to exactly \(d-1\) triangles, then \(v\) is part of a \(d+1\)-clique.

**Proof of Theorem 2.4.** Let \(G_{d,t}^\varepsilon(n) \subset G_{d,t}(n)\) denote the subset of those graphs where at least \(\varepsilon n\) nodes are bad. Our goal is to show

\[
\lim_{n \to \infty} \frac{|G_{d,t}^\varepsilon(n)|}{|G_{d,t}(n)|} = 0
\]

Let \(t_e\) denote, as earlier, the number of triangles the edge \(e\) is incident to. If \(1 \leq t_e < d - 2\) then

\[
\frac{t_e}{t_e + 2} = \frac{t_e}{t_e + 3} + \left(\frac{t_e}{t_e + 3} - \frac{t_e}{t_e + 3}\right) = \frac{t_e}{t_e + 3} + \frac{1}{(1 + \frac{1}{t_e})(t_e + 3)} > \frac{t_e}{d + 1} + \frac{1}{3(d + 1)}.
\]

Suppose more than \(\varepsilon n\) nodes of \(G\) are bad. Each bad node, by definition, is adjacent to at least two edges with \(1\) triangle. So the total number of edges for which \(1 \leq t_e \leq d - 2\) is at least \(\varepsilon n\). Combining Lemma 2.5 with (3) we get that for a uniformly random permutation \(\sigma \in S_n\)

\[
\mathbb{E} [\phi(G_\sigma)] = \sum_e \frac{t_e}{t_e + 2} \geq \varepsilon n \frac{1}{d + 1} + \sum_e \frac{t_e}{d + 1} \geq \frac{tn}{d + 1} + \frac{\varepsilon n}{3d + 3}.
\]

Hence, by the same computation as in (3) we get

\[
\mathbb{P} \left( \left| \phi(G_\sigma) \right| \geq \frac{tn}{d + 1} + \frac{\varepsilon n}{3d + 3} - 1 \right) \geq \frac{2}{dn}.
\]

Now let

\[L_\varepsilon = \left\{ x \in \{0, 1\}^n : \left| x \right| \geq \frac{nt}{d + 1} + \frac{\varepsilon n}{3d + 3} - 1 \right\}.
\]

By the previous considerations, for any \(G \in G_{d,t}^\varepsilon(n)\) we get that

\[
\frac{|S_n G \cap \phi^{-1}(L_\varepsilon)|}{|S_n G|} = \frac{|\{\sigma \in S_n : \phi(G_\sigma) \in L_\varepsilon\}|}{|S_n|} \geq \frac{2}{dn}.
\]

Summing the inequality \(|S_n G \cap \phi^{-1}(L_\varepsilon)| \geq \frac{2}{dn}|S_n|\) over the orbits of the \(S_n\) action in \(G_{d,t}^\varepsilon(n)\) we obtain the estimate

\[
|G_{d,t}^\varepsilon(n)| \leq \frac{dn}{2} |\phi^{-1}(L_\varepsilon)|,
\]

which, combined with Lemma 2.1 yields

\[
|G_{d,t}^\varepsilon(n)| \leq n^{1 + n \left(\frac{2}{d + 1} - \frac{\varepsilon}{3d + 3}\right) \cdot \frac{d}{d + 1}} \cdot \frac{2}{dn} \leq n^{n \left(\frac{2}{d + 1} - \frac{\varepsilon}{3d + 3}\right)} \cdot n \cdot (2d) \frac{d}{d + 1}.
\]

Finally, combined with Theorem 1.1, we obtain

\[
\log |G_{d,t}^\varepsilon(n)| \leq \left(\frac{d}{2} - \frac{t}{d + 1} - \frac{\varepsilon}{3d + 3}\right) n \log n + O(n) \leq \log |G_{d,t}(n)| - \frac{\varepsilon}{3d + 3} n \log n + O(n).
\]

As long as \(\varepsilon \log n \to \infty\), we get

\[
\limsup_{n \to \infty} \frac{|G_{d,t}^\varepsilon(n)|}{|G_{d,t}(n)|} = -\infty,
\]

finishing the proof.
2.3 $k$-cliques

We can easily extend the above results from triangles to $k$-cliques. Let $\mathcal{G}_{d,t,k}(n)$ denote the set of $d$-regular graphs on $n$ nodes that contain at least $t_{k,n}$ $k$-subgraphs isomorphic to $K_k$.

Proof of Theorem 1.4. The idea is a simple reduction the the $k=3$ case. Clearly, each $G \in \mathcal{G}_{d,t,k}(n)$ has at least
\[
\binom{n}{3} t_{k,n} \frac{k}{k(k-3)} = \frac{tn}{3}
\]
triangles, so $\mathcal{G}_{d,t,k}(n) \subset \mathcal{G}_{d,t}(n)$, which implies the upper bound of the theorem. On the other hand, the family of graphs constructed in Theorem 1.1 contain
\[
b \left( \frac{d+1}{k} \right) \geq \frac{tn}{3} \left( \frac{d+1}{k} \right) = \frac{tn}{k} \left( \frac{k}{k-3} \right) \left( \frac{d+1}{k} \right) = \frac{tk}{k}
\]
k-cliques, so this family is contained in $\mathcal{G}_{d,t,k}(n)$, implying the lower bound of the theorem. Finally, the structural statement follows directly from Theorem 2.4. □

Acknowledgements Pim van der Hoorn and Gabor Lippner was supported by ARO grant W911NF1610391, Gabor Lippner was also supported by NSF grant DMS 1800738, and Elchanan Mossel was supported by NSF grant DMS-1737944 and ONR grant N00014-17-1-2598.

References

[1] Edward A Bender and E Rodney Canfield. The asymptotic number of labeled graphs with given degree sequences. *Journal of Combinatorial Theory, Series A*, 24(3):296–307, 1978.

[2] Béla Bollobás. A probabilistic proof of an asymptotic formula for the number of labelled regular graphs. *European Journal of Combinatorics*, 1(4):311–316, 1980.

[3] Sourav Chatterjee. The missing log in large deviations for triangle counts. *Random Structures & Algorithms*, 40(4):437–451, 2012.

[4] Sourav Chatterjee and Amir Dembo. Nonlinear large deviations. *Advances in Mathematics*, 299:396–450, 2016.

[5] Sourav Chatterjee and S.R.S. Varadhan. The large deviation principle for the erdős-rényi random graph. *European Journal of Combinatorics*, 32(7):1000 – 1017, 2011. Homomorphisms and Limits.

[6] B. DeMarco and J. Kahn. Upper tails for triangles. *Random Structures & Algorithms*, 40(4):452–459, aug 2011.

[7] Svante Janson, Krzysztof Oleszkiewicz, and Andrzej Ruciński. Upper tails for subgraph counts in random graphs. *Israel Journal of Mathematics*, 142(1):61–92, 2004.

[8] Svante Janson and Andrzej Ruciński. The infamous upper tail. *Random Structures & Algorithms*, 20(3):317–342, 2002.

[9] Svante Janson and Andrzej Ruciński. The deletion method for upper tail estimates. *Combinatorica*, 24(3):615–640, 2004.

[10] Richard Kenyon, Charles Radin, Kui Ren, and Lorenzo Sadun. Bipodal structure in oversaturated random graphs. *International Mathematics Research Notices*, page rnw261, 2016.

[11] Richard Kenyon, Charles Radin, Kui Ren, and Lorenzo Sadun. The phases of large networks with edge and triangle constraints. *Journal of Physics A: Mathematical and Theoretical*, 50(43):435001, 2017.

[12] Jeong Han Kim, Benny Sudakov, and Van Vu. Small subgraphs of random regular graphs. *Discrete Mathematics*, 307(15):1961–1967, 2007.
[13] Jeong Han Kim and Van H Vu. Divide and conquer martingales and the number of triangles in a random graph. *Random Structures & Algorithms*, 24(2):166–174, 2004.

[14] Dmitri Krioukov. Clustering implies geometry in networks. *Physical review letters*, 116(20):208302, 2016.

[15] Eyal Lubetzky and Yufei Zhao. On replica symmetry of large deviations in random graphs. *Random Structures & Algorithms*, 47(1):109–146, 2015.

[16] Eyal Lubetzky and Yufei Zhao. On the variational problem for upper tails in sparse random graphs. *Random Structures & Algorithms*, 50(3):420–436, 2017.

[17] Charles Radin, Kui Ren, and Lorenzo Sadun. The asymptotics of large constrained graphs. *Journal of Physics A: Mathematical and Theoretical*, 47(17):175001, 2014.

[18] Charles Radin and Lorenzo Sadun. Phase transitions in a complex network. *Journal of Physics A: Mathematical and Theoretical*, 46(30):305002, 2013.

[19] Pim van der Hoorn, Gabor Lippner, and Dmitri Krioukov. Sparse maximum-entropy random graphs with a given power-law degree distribution. *Journal of Statistical Physics*, 173(3):806–844, Nov 2018.

[20] Van H Vu. A large deviation result on the number of small subgraphs of a random graph. *Combinatorics, Probability and Computing*, 10(1):79–94, 2001.