# Regular graphs with many triangles are structured

Pim van der Hoorn[*]

Department of Mathematics and Computer Science
TU Eindhoven
Eindhoven, Netherlands

`w.l.f.v.d.hoorn@tue.nl`

Gabor Lippner[†]

Department of Mathematics
Northeastern University
Boston, MA, U.S.A.

`g.lippner@northeastern.edu`

Elchanan Mossel[‡]

Department of Mathematics, and
MIT

Cambridge, MA, U.S.A.

`elmos@mit.edu`

## Abstract

We compute the leading asymptotics of the logarithm of the number of $d$-regular graphs having at least a fixed positive fraction $c$ of the maximum possible number of triangles, and provide a strong structural description of almost all such graphs.

When $d$ is constant, we show that such graphs typically consist of many disjoint $(d+1)$-cliques and an almost triangle-free part. When $d$ is allowed to grow with $n$, we show that such graphs typically consist of very dense sets of size $d + o(d)$ together with an almost triangle-free part.

This confirms a conjecture of Collet and Eckmann from 2002 and considerably strengthens their observation that the triangles cannot be totally scattered in typical instances of regular graphs with many triangles.

**Mathematics Subject Classifications:** 05C80, 05C30, 05C75

# 1 Introduction

It is easy to see that a $d$-regular graph on $n$ nodes cannot have more than $T_{\max} = T_{\max}(n, d) = \frac{n}{3}\binom{d}{2}$ triangles, and that this value is achieved exactly if the graph is a disjoint union of $(d+1)$-cliques. In this paper we prove a *strong, almost sure structural stability result* for this extremal problem: Let $0 < c \leqslant 1$ and let $\mathcal{G}_{d,c}(n)$ denote the set of $d$-regular graphs on $n$ labeled nodes that contain at least $c \cdot T_{\max}$ triangles. We show that, for constant $d$ and large $n$, almost all graphs in $\mathcal{G}_{d,c}(n)$ consist of a disjoint union of $(d+1)$-cliques and an almost triangle-free part.

This result may not seem surprising at first, and especially for $c$ close to 1 it seems quite natural to expect. However we would like to emphasize that it also holds for very small positive $c$. In that regime the required number of triangles could easily be arranged in such a way that no vertex has more than one triangle in its neighborhood. However, as it turns out, such graphs only constitute a vanishing fraction of elements of $\mathcal{G}_{d,c}(n)$.

The study of $\mathcal{G}_{d,c}(n)$ has been initiated by Collet and Eckmann [5], who proved

$$1 - c\frac{2d}{d+1} \leqslant \liminf_{n \to \infty} \frac{\log|\mathcal{G}_{d,c}(n)|}{\log|\mathcal{G}_d(n)|} \leqslant \limsup_{n \to \infty} \frac{\log|\mathcal{G}_{d,c}(n)|}{\log|\mathcal{G}_d(n)|} \leqslant 1 - \frac{c}{3}, \tag{1}$$

where $\mathcal{G}_d(n)$ is the set of all $d$-regular graphs on $n$ labeled nodes. They conjectured that this limit exists. Using a heuristic counting argument they concluded that the triangles should be clustered in a "typical" element of $\mathcal{G}_{d,c}(n)$.

In this paper we prove their conjecture and find the correct limiting value

$$\lim_{n \to \infty} \frac{\log|\mathcal{G}_{d,c}(n)|}{\log|\mathcal{G}_d(n)|} = 1 - c \cdot \frac{d-1}{d+1}. \tag{2}$$

Furthermore we prove that the observation that triangles are clustered in typical elements of $\mathcal{G}_{d,c}(n)$ holds in a very strong sense: for almost all graphs in the class, almost all triangles are contained in $(d+1)$-cliques.

## 1.1 Probabilistic context

While our methods are purely combinatorial, the results are more conveniently stated in the language of random graphs. 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 i.i.d. variables $X \sim X_1, X_2, \ldots$ there exists a "rate function" $I(x)$ depending on the distribution of $X$ such that

$$\mathbb{P}\left(\sum_1^N X_i \geqslant Nx\right) \approx e^{-N \cdot I(x)}.$$

### 1.1.1 Upper tail interpretation

Let $\mathbb{G}_d(n)$ be a random $d$-regular graph. Our results can be related to the logarithmic scaling of the upper tail probability $\mathbb{P}\left(T(\mathbb{G}_d(n)) > cT_{\max}\right)$, where $T(G)$ denotes the number of triangles in the graph $G$, by noting that

$$\mathbb{P}\left(T(\mathbb{G}_d(n)) > cT_{\max}\right) = \frac{|\mathcal{G}_{d,c}(n)|}{|\mathcal{G}_d(n)|}.$$

In random graphs, the question about the upper tail for triangles in $\mathbb{G}(n,p)$ has been long studied for a constant factor of deviation from the mean [10]. More precisely, let $t(\mathbb{G}(n,p))$ denote the triangle density in the Erdős-Rényi random graph, normalized so that $\mathbb{E}\left[t(\mathbb{G}(n,p))\right] = p^3$. One would like to understand the asymptotic behavior of

$$r(n,p,\delta) = -\log \mathbb{P}\left(t(\mathbb{G}(n,p) > (1+\delta)p^3\right).$$

The dense case ($p$ a constant) has been reduced to an analytic variational problem by Chatterjee and Varadhan [4] using methods from graph limits. However, the solution of this variational problem is only known in certain parameter ranges (see [17] for details). In the sparse (case $p = o(1)$) regime the asymptotics $r(n,p,\delta) \approx n^2 p^2 \log(1/p)$ have been determined in a long series of papers by many authors [2, 7, 9, 11, 15, 22]. The variational methods were extended to (part of) the sparse regime in [3] and using this, Lubetzky and Zhao [18] found the exact asymptotics of $r(n,p,\delta)$ in the $n^{-1/42}\log n \leqslant p \ll 1$ range. Recently, Cook and Dembo [6] and Augeri [1] extended it to the range $n^{-1/2} \ll p \ll 1$, and Harel, Mousset and Samotij [8] to all $n^{-1}\log n \ll p \ll 1$.

The case of the random regular graph, $\mathbb{G}_d(n)$, has been studied much less. Kim, Sudakov, and Vu [14] obtained that the distribution of small subgraphs of $\mathbb{G}_d(n)$ is asymptotically Poisson in the sparse case, implying an asymptotic formula for the tail probability $\mathbb{P}\left(T(\mathbb{G}_d(n)) > (1+\delta)\mathbb{E}\left[T(\mathbb{G}_d(n))\right]\right)$. In particular, for fixed $d$ the expected number of triangles in $\mathbb{G}_d(n)$ is finite. This means that the "standard" regime for large deviations - exceeding the expected value by a constant factor - in this case is not interesting.

### 1.1.2 Maximum entropy random graphs with triangles

Our motivation for analyzing the logarithmic size of $\mathcal{G}_{d,c}(n)$ originates from a related problem of describing "typical" graphs with a specific set of constraints. In general, such constraints can be either local or global, eg. restricting the degree of each node or prescribing the total number of triangles. In addition, they can be "hard" or "soft", eg. each node must have degree $d$ or the expected degree of each node should be $d$.

To study typical graphs, one looks for the probability distribution $\mathbb{P}$ on $\mathcal{G}(n)$ satisfying these constraints that is as random as possible in the sense that it maximizes the Shannon entropy

$$\mathcal{E}[\mathbb{P}] = \sum_{G \in \mathcal{G}(n)} -\mathbb{P}(G)\log\left(\mathbb{P}(G)\right). \tag{3}$$

Results of Krioukov [16] suggest that certain constraints on the number of triangles might result in maximum entropy solutions with some geometric component. Our results

further point in this direction: we show that typical $d$-regular graphs with many triangles contain a large, highly structured, part.

We only consider hard constraints that are a mix between local ($d$-regular) and global (at least $T_{\max}$ triangles). Of course, the distributions satisfying such constraints are simply all probability distributions on the set $\mathcal{G}_{d,c}(n)$. In this case it is well-known that the uniform distribution $\mathbb{P}_{d,c}$ on $\mathcal{G}_{d,c}(n)$ maximizes the Shannon entropy, with value

$$\mathcal{E}[\mathbb{P}_{d,c}] = \sum_{G \in \mathcal{G}_{d,c}(n)} \frac{1}{|\mathcal{G}_{d,c}(n)|} \log |\mathcal{G}_{d,c}(n)| = \log |\mathcal{G}_{d,c}(n)|.$$

While in our setting the maximum entropy solution is clear, estimating its entropy - the logarithmic size of the set of $\mathcal{G}_{d,c}(n)$ - turns out to be an important first step describing the structure of typical graphs with the given constraints on degrees and triangles.

## 1.2 Results

Motivated by the question "can local constraints induce global (geometric) behavior?", we study the random $d$-regular graph $\mathbb{G}_d(n)$ conditioned on having at least a positive fraction of the maximum possible number of triangles. (For $d$ fixed this just means linearly many triangles, in $n$.) With respect to 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.

Recall that $T_{\max} = T_{\max}(n,d) = \binom{d}{2}n/3$ denotes the maximum number of triangles an $n$ vertex $d$-regular graph can have. Let $\mathcal{G}_{d,c}(n)$ denote the set of $d$-regular graphs on $n$ labeled nodes that contain at least $c \cdot T_{\max}$ triangles. We compute the leading asymptotics of $|\mathcal{G}_{d,c}(n)|$ for fixed $c$, as $n \to \infty$, where $d$ is either a constant or can grow with $n$ as long as $\log d = o(\log n)$. We provide a structural description of a "typical" element of $\mathcal{G}_{d,c}(n)$. We then extend these results to the case of $k$-cliques in $d$-regular graphs.

### 1.2.1 Number of $d$-regular graph with many triangles

The dependence of $d$ on $n$ will be suppressed from the notation. We always assume $d = o(\sqrt{n})$. We will emphasize when constant $d$ is assumed.

**Theorem 1.** *For any $0 < c < 1$ there is a $C > 0$ such that*

$$\frac{C}{\log \frac{n}{d}} \leqslant \frac{\log|\mathcal{G}_{d,c}(n)|}{\frac{dn}{2}\log\frac{n}{d}} - \left(1 - c \cdot \frac{d-1}{d+1}\right) \leqslant c \frac{\log d}{\log \frac{n}{d}} + \frac{C}{\log \frac{n}{d}}$$

*holds for any $n$ and any $d < \sqrt{n}/C$.*

The part $\frac{dn}{2}\log\frac{n}{d}$ is related to $\log|\mathcal{G}_d(n)|$, where $\mathcal{G}_d(n)$ denotes the set of $d$-regular graphs on $n$ nodes. In particular, results of [19] imply that for $d = o(\sqrt{n})$

$$\lim_{n \to \infty} \frac{\log|\mathcal{G}_d(n)|}{\frac{dn}{2}\log\frac{n}{d}} = 1.$$

The $C/\log(n/d))$ terms are $o(1)$ as long as $d = o(n)$. The $c\log d/\log(n/d)$ term on the right hand side is only $o(1)$ if $\log d = o(\log n)$. Unfortunately, for $d$ polynomial in $n$ we do not get a sharp logarithmic rate.

### 1.2.2 Structure of $d$-regular graph with many triangles

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

**Theorem 2.** *Let $d$ be fixed and $0 < c < 1$. With high probability a uniformly randomly chosen element of $\mathcal{G}_{d,c}(n)$ has less than $\frac{\log\log n}{\log n}n$ bad nodes. Thus, the number of triangles that are not part of a $(d + 1)$-clique is sublinear.*

Note that Theorem 2 implies a two-phase structure: the graph consist of many disjoint cliques and an almost triangle-free part.

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

These results show similarities with studies on the structure of dense maximal entropy random graphs with given edge and triangle densities. In a collection of works by Kenyon, Radin, Ren and Sadun [12, 13, 20, 21] it was shown that the limits of these graph have a bipodal structure, at least in a narrow range just above the average. This means that the graph is split into two components with specific inter- and intra-component connection probabilities. In our setting we obtain a multipodal structure with a linear set of components, consisting of the $(d + 1)$-cliques and the triangle-free part.

We prove a similar result for the $1 \ll d \ll n$ case. Here, however, we cannot expect $(d+1)$-cliques to appear, as it is possible to construct families of examples with the correct leading logarithmic growth rate, that do not have any cliques. Instead, we introduce a notion of a pseudo-clique, which turns out to be a very dense subgraph of size $d+o(d)$ with the property that different pseudo-cliques must be disjoint. (See the explanation at the beginning of Section 2.2.2 for details.) It turns out that a typical element of the ensemble consists of a collection of these pseudo-cliques together with an almost triangle-free part.

**Theorem 3.** *Let $1 \ll d \ll \sqrt{n}$ and fix $0 < c < 1$. With high probability, almost all triangles of a uniformly randomly chosen element of $\mathcal{G}_{d,c}(n)$ are contained in pseudo-cliques.*

### 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 $\mathcal{G}_{d,c,k}(n)$ denote the set of $d$-regular graphs on $n$ nodes that contain at least $c \cdot T_{k,\max} = c\binom{d}{k-1}n/k$ subgraphs isomorphic to $K_k$. As a natural extension of terminology, we call nodes *bad* if they are not part of a $(d + 1)$-clique but are incident to a $k$-clique.

**Theorem 4.** *For $k \geqslant 3$ and fixed $0 < c < 1$ we have*

$$\left| \frac{\log|\mathcal{G}_{d,c,k}(n)|}{(d/2)n \log n} - \left( 1 - c \cdot \frac{d-1}{d+1} \right) \right| = O(\log d / \log n).$$

*Furthermore, for $d$ fixed, 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 consist of establishing a lower and upper bound on $\log |\mathcal{G}_{d,c}(n)|$. More precisely, we will show that

$$-Cdn \leqslant \log|\mathcal{G}_{d,c}(n)| - \left( 1 - c \cdot \frac{d-1}{d+1} \right) \frac{dn}{2} \log \frac{n}{d} \leqslant c \frac{dn}{2} \log d + Cdn.$$

The theorem then follows after dividing by $\frac{dn}{2} \log \frac{n}{d+1}$.

*Proof of Theorem 1 (Lower bound).* To establish a lower bound we construct a family of elements in $\mathcal{G}_{d,c}(n)$ by letting

$$b = c \cdot T_{\max} \cdot \binom{d+1}{3}^{-1} = \frac{c \cdot n}{d+1},$$

taking $b$ disjoint $(d+1)$-cliques and an arbitrary $m = n - (d+1)b = (1-c)n$ node $d$-regular graph. Clearly, these graphs will have at least $c \cdot T_{\max}$ triangles. Thus

$$|\mathcal{G}_{d,c}(n)| \geqslant \binom{n}{d+1}\binom{n-(d+1)}{d+1} \cdots \binom{n-(b-1)(d+1)}{d+1} \frac{|\mathcal{G}_d(m)|}{b!} = \binom{n}{m}\frac{(b(d+1))!}{(d+1)!^b}\frac{|\mathcal{G}_d(m)|}{b!}. \quad (4)$$

For $d = o(\sqrt{m})$, McKay and Wormald [19] show that the number of $d$-regular graphs on $m$ nodes is asymptotically

$$\frac{(dm)!}{(dm/2)! 2^{dm/2} (d!)^m} \exp\left( \frac{-d^2-1}{4} - \frac{d^3}{12m} + O(d^2/m) \right).$$

Stirling's approximation, $|\log k! - (k \log k - k + 1/2 \log k)| = O(1)$, then implies that

$$\log|\mathcal{G}_d(m)| \geqslant \frac{1}{2}dm \log \frac{m}{d} + \frac{dm}{2} - \frac{m}{2} \log d - O(m). \quad (5)$$

**Claim 5.** *Let $n = m + b(d+1) = (1-c)n + b(d+1)$ as before.*

  1.
$$\log \binom{n}{m} = nH(m/n) - \frac{\log n}{2} + O(1) = nH(1-c) - \frac{\log n}{2} + O(1) = O(n)$$

  *where $H(x) = -x \log x - (1-x)\log(1-x)$.*

*2.*

$$\log \frac{(b(d+1))!}{(d+1)!^b} = b(d+1)\log b - \frac{b}{2}\log(d+1) + O(b) = cn\log\frac{cn}{d+1} + O(dn).$$

*3.*

$$\log b! = b\log b - b + 1/2\log b + O(1) = \frac{cn}{d+1}\log\frac{cn}{d+1} + O(dn).$$

*Proof.* All three parts follow easily from Stirling's approximation. □

Combining these estimates with (4) and (5) we get

$$\log|\mathcal{G}_{d,c}(n)| \geqslant \log\binom{n}{m} + \log\frac{(b(d+1))!}{(d+1)!^b} - \log b! + \log|\mathcal{G}_d(m)|$$

$$\geqslant \frac{d(1-c)n}{2}\log\frac{(1-c)n}{d} + \frac{cnd}{d+1}\log\frac{cn}{d+1} + O(dn)$$

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

In the last step we used that $\log\frac{d}{d+1}, \log c, \log(1-c)$ are all $O(1)$. □

We now need to prove a matching upper bound on $|\mathcal{G}_{d,c}(n)|$. 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 use an approach inspired by the configuration model. Let us denote by $\mathcal{G}_d^*(n)$ (respectively, $\mathcal{G}_{d,c}^*(n)$) the set of $d$-regular graphs (respectively, $d$-regular graphs with at least $c \cdot T_{\max}$ triangles) on $n$ labeled nodes, where additionally the edges leaving each node are assigned labels 1 through $d$. This means that each edge gets two labels, one from each end.

Given $G^* \in \mathcal{G}_{d,c}^*(n)$, we define the *configuration ordering* $\prec$ on the set of edges of $G^*$ as follows. Let $e = \{i_1, j_1\}$ and $f = \{i_2, j_2\}$ be two edges of $G^*$ with $i_1 < j_1$ and $i_2 < j_2$. Let us declare $e \prec f$ if $i_1 < i_2$, or if $i_1 = i_2$ and the label of $e$ is smaller than the label of $f$ at their common node. Let $e_1 \prec e_2 \prec \cdots \prec e_{nd/2}$ denote the edges of $G^*$ in increasing configuration order. Let $G^*[k]$ denote the subgraph of $G^*$ consisting of $e_1, \ldots, e_k$.

Finally we define the "triangle reveal profile" function

$$\phi : \mathcal{G}_{d,c}^*(n) \to \{0,1\}^{nd/2}$$

that records the moments new triangles appear in the graph when edges are added 1-by-1 according to the configuration ordering. So we let $\phi(G^*)(k) = 1$ if $e_k$ is incident to a triangle in $G^*[k]$ and 0 otherwise. Denoting $e_k = \{i, j\}$, it is clear that we have $\phi(G^*)(k) = 1$ if and only if there is a triangle $\{h, i, j\}$ in $G^*$ such that $h < \min(i, j)$.

For any $x \in \{0,1\}^{nd/2}$ let us denote $\|x\| = \sum_{j=1}^{nd/2} x(j)$. Then $\|\phi(G^*)\|$ denotes the total number of edges $e_k$ that upon adding to the graph $G^*[k-1]$ have created at least one new triangle. The next lemma gives an upper bound on the number of graphs in $\mathcal{G}_{d,c}^*(n)$ that have a fixed triangle reveal profile.

**Lemma 6.** *For every $x \in \{0,1\}^{dn/2}$*

$$\left|\phi^{-1}(x)\right| \leqslant (dn)^{dn/2} \cdot \left(\frac{d^2}{n}\right)^{\|x\|}.$$

*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 configuration order. Just like in the configuration model, each node starts with $d$ half-edges, labeled 1 through $d$. First we take the half-edge with label 1 at node 1, and join it to any other half-edge. We can do this in $dn - 1$ ways. Then, in each subsequent step, we take the smallest node that still has half-edges, pick the one with the smallest label, and match it to any another half-edge. If we did not have constraints on triangles, the total number of possible (multi-)graphs we could create this way would be $(dn - 1)(dn - 3) \cdots 3 \cdot 1$, which is an upper bound on $|\mathcal{G}_d^*(n)|$.

In our case, the vector $x$ dictates whether the next edge added has to create a triangle with previously added edges. By the definition of the configuration order, the number of possible choices for the $k$th edge is $dn - (2k - 1)$, as the starting half-edge is fixed and there are exactly $dn - (2k - 1)$ available half-edges at this step. However, when $x(k) = 1$, the number of choices for the ending half-edge is limited. Suppose the starting half-edge is incident to node $j$. Then, in order for this edge to create a triangle, the ending half-edge most be incident to one of the current 2nd neighbors of $j$. There are never more than $d^2$ second neighbors, and thus never more than $d^3$ possible half-edges to choose from.

Thus we get the upper bound

$$\begin{aligned} |\phi^{-1}(x)| &\leqslant \prod_{j:x(j)=0} (dn - (2j - 1)) \cdot \prod_{j:x(j)=1} d^3 \\ &\leqslant d^{3\|x\|} \cdot (dn)^{dn/2 - \|x\|} \\ &= (dn)^{dn/2} \cdot \left(\frac{d^2}{n}\right)^{\|x\|} \end{aligned}$$

which proves the lemma. $\qquad\square$

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 a $c\frac{d-1}{d+1}$ fraction of edges have revealed triangles.

*Proof of Theorem 1 (Upper bound).* Let us introduce the following short hand notation,

$$T_c = c \cdot \frac{dn}{2} \frac{d-1}{d+1}, \tag{6}$$

as it will come up frequently. Define

$$L = \left\{ x \in \{0,1\}^{\frac{nd}{2}} : \|x\| \geqslant T_c - 1 \right\}.$$

Then, by Lemma 6, and using $d^2 \leqslant n$, we see that

$$\left|\phi^{-1}(L)\right| \leqslant |L|(dn)^{dn/2} \left(\frac{d^2}{n}\right)^{T_c-1} \leqslant 2^{dn/2}(dn)^{dn/2} \left(\frac{d^2}{n}\right)^{T_c-1} \tag{7}$$

To finish the proof, we will show that $\left|\mathcal{G}_{d,c}^*(n)\right| \leqslant \frac{dn}{2}|\phi^{-1}(L)|$. For this, consider the symmetric group $S_n$, which acts on $\mathcal{G}_{d,c}^*(n)$ by permuting the node labels. For $\sigma \in S_n$ and $G^* \in \mathcal{G}_{d,c}^*(n)$, let us denote by $G_\sigma^*$ the graph obtained by applying $\sigma$ to the node labels. Furthermore let $S_n G^* = \{G_\sigma^* : \sigma \in S_n\} \subset \mathcal{G}_{d,c}^*(n)$ denote the orbit of $G^*$ under the action of $S_n$. We finish the proof modulo the following result, which we establish at the end of this section.

**Lemma 7.** *For any $G^* \in \mathcal{G}_{d,c}^*(n)$ we have*

$$\left|S_n G^* \cap \phi^{-1}(L)\right| \geqslant \frac{2}{dn}|S_n G^*|.$$

*In other words, randomly relabeling the nodes of $G^*$ yields, with not too small probability, a graph whose $\|\phi(G_\sigma^*)\| \geqslant T_c - 1$.*

Summing this inequality over all orbits of the $S_n$ actions yields $\left|\mathcal{G}_{d,c}^*(n)\right| \leqslant \frac{dn}{2}|\phi^{-1}(L)|$ as claimed above. Note that $|\mathcal{G}_{d,c}^*(n)| = |\mathcal{G}_{d,c}(n)| \cdot (d!)^n$. Combining this with (7) we get

$$
\begin{aligned}
|\mathcal{G}_{d,c}(n)| = \frac{|\mathcal{G}_{d,c}^*(n)|}{(d!)^n} &\leqslant \frac{dn}{2}\frac{|\phi^{-1}(L)|}{(d!)^n} \leqslant \frac{dn}{2}2^{dn/2}\frac{(dn)^{dn/2}}{(d/e)^{dn}}\left(\frac{d^2}{n}\right)^{T_c-1} \\
&= \frac{dn}{2}(\sqrt{2}e)^{dn}\left(\frac{n}{d}\right)^{\frac{dn}{2}-T_c+1}d^{T_c-1} \\
&= (\sqrt{2}e)^{dn}\left(\frac{n}{d}\right)^{\frac{dn}{2}-T_c}d^{T_c}\frac{n^2}{2d}
\end{aligned}
$$

Thus

$$\log|\mathcal{G}_{c,d}(n)| \leqslant \left(1 - c\cdot\frac{d-1}{d+1}\right)\frac{dn}{2}\log\frac{n}{d} + c\cdot\frac{dn}{2}\log d + O(dn).$$

$\square$

We are thus left to prove Lemma 7. For this we first show that for a uniform random permutation $\sigma$, the expected value of $\|\phi(G_\sigma^*)\|$ is at least $c\cdot\frac{dn(d-1)}{2(d+1)}$. Then the lemma will follow from a standard Markov-inequality argument.

**Lemma 8.** *Let $\sigma$ be a uniformly random permutation. Then*

$$\mathbb{E}\left[\|\phi(G_\sigma^*)\|\right] \geqslant T_c.$$

*Proof.* Let $X_e(\sigma)$ be the indicator variable that the edge $e$ of $G^*$ creates a triangle when it is added in the lexicographic order of $G_\sigma^*$. Then $\|\phi(G_\sigma^*)\| = \sum_e X_e(\sigma)$ and so

$$\mathbb{E}\left[\|\phi(G_\sigma^*)\|\right] = \sum_e \mathbb{E}\left[X_e(\sigma)\right] = \sum_e \mathbb{P}\left(X_e(\sigma) = 1\right).$$

Let $e = \{i, j\}$ 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. Then $X_e(\sigma)$ is 1 if at least one of these triangles is formed at the moment when $e$ is added, 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 $\mathbb{P}(X_e(\sigma) = 0) = 2/(t_e + 2)$ and hence $\mathbb{P}(X_e(\sigma) = 1) = 1 - 2/(t_e + 2)$.

Thus, since $t_e \leqslant d - 1$, we get

$$\mathbb{E}\left[\|\phi(G_\sigma^*)\|\right] = \sum_e \mathbb{P}(X_e(\sigma) = 1) = \sum_e \left(1 - \frac{2}{t_e + 2}\right) = \sum_e \frac{t_e}{t_e + 2} \geqslant \sum_e \frac{t_e}{d + 1} \geqslant T_c, \tag{8}$$

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 $c \cdot \frac{n}{3}\binom{d}{2}$. This finishes the proof of the lemma. $\qquad\square$

*Proof of Lemma 7.* By simple algebraic considerations

$$\frac{|S_n G^* \cap \phi^{-1}(L)|}{|S_n G^*|} = \frac{|\{\sigma \in S_n : \phi(G_\sigma^*) \in L\}|}{|S_n|}. \tag{9}$$

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 (9) it is enough to show that with probability at least $\frac{2}{dn}$ we have $\phi(G_\sigma^*) \in L$, which is equivalent to $\|\phi(G_\sigma^*)\| \geqslant T_c - 1$.

Observe that $\|\phi(G_\sigma^*)\|$ cannot be bigger than $\frac{dn}{2}$. Hence, using Lemma 8

$$T_c \leqslant \mathbb{E}\left[\|\phi(G_\sigma^*)\|\right] \leqslant (T_c - 1)\,\mathbb{P}\left(\|\phi(G_\sigma^*)\| < T_c - 1\right) + \frac{dn}{2}\mathbb{P}\left(\|\phi(G_\sigma^*)\| \geqslant T_c - 1\right)$$
$$\leqslant T_c - 1 + \frac{dn}{2}\mathbb{P}\left(\|\phi(G_\sigma^*)\| \geqslant T_c - 1\right), \tag{10}$$

from which we conclude that

$$\mathbb{P}\left(\|\phi(G_\sigma^*)\| \geqslant T_c - 1\right) \geqslant \frac{2}{dn}.$$

$\qquad\square$

## 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 yields a strong structural description of a typical graph with at least $c \cdot T_{\max}$ triangles:

**For $d$ fixed:** $1 - o(1)$ fraction of all triangles are contained in $(d+1)$-cliques.

**For $1 \ll d \ll \sqrt{n}$:** $1 - o(1)$ fraction of all triangles are contained in pseudo-cliques. Moreover, these pseudo-cliques are non-overlapping.

We will treat the two cases separately, but the following lemma will be useful for both. As before, we let $t_e$ denote the number of triangles the edge $e$ is incident to. We say the edge $e$ is $\delta$-*bad* if $1 \leqslant t_e \leqslant d - 1 - \delta d$.

**Lemma 9.** *Let $\varepsilon, \delta > 0$ fixed. Let $\mathcal{G}_{d,c}^{\varepsilon,\delta} \subset \mathcal{G}_{d,c}$ denote the subset of graphs where at least $\varepsilon(d/2)n$ edges are $\delta$-bad. Then*

$$\log|\mathcal{G}_{d,c}^{\varepsilon,\delta}(n)| \leqslant \left(1 - c\frac{d-1}{d+1} - \frac{\varepsilon\delta d}{3d+3}\right)\frac{dn}{2}\log\frac{n}{d} + \left(c + \frac{\varepsilon\delta d}{3d+3}\right)\frac{dn}{2}\log d + O(dn).$$

*Proof.* If $e$ is $\delta$-bad, then $1 \leqslant t_e \leqslant d - 1 - \delta d$, so

$$\frac{t_e}{t_e+2} = \frac{t_e}{d+1} + \left(\frac{t_e}{t_e+2} - \frac{t_e}{d+1}\right) = \frac{t_e}{d+1} + \frac{t_e}{t_e+2}\cdot\frac{d-1-t_e}{d+1} \geqslant \frac{t_e}{d+1} + \frac{1}{3}\cdot\frac{\delta d}{d+1}.$$

Suppose more than $\varepsilon(d/2)n$ edges of $G^* \in \mathcal{G}_{d,c}^*(n)$ are bad. Combining the above with (8) we get that for a uniformly random permutation $\sigma \in S_n$

$$\mathbb{E}\left[\|\phi(G_\sigma^*)\|\right] = \sum_e \frac{t_e}{t_e+2} \geqslant \sum_e \frac{t_e}{d+1} + \frac{\varepsilon\delta d^2 n}{6d+6} \geqslant T_c + \frac{\varepsilon\delta d^2 n}{6d+6}.$$

Hence, by the same computation as in (10) we get

$$\mathbb{P}\left(\|\phi(G_\sigma^*)\| \geqslant T_c + \frac{\varepsilon\delta d^2 n}{6d+6} - 1\right) \geqslant \frac{2}{dn}.$$

Now let

$$L_{\varepsilon,\delta} = \left\{x \in \{0,1\}^{\frac{nd}{2}} : |x| \geqslant T_c + \frac{\varepsilon\delta d^2 n}{6d+6} - 1\right\}.$$

By the previous considerations, for any $G^* \in \mathcal{G}_{d,c}^{*\,\varepsilon,\delta}(n)$ we get that

$$\frac{|S_n G^* \cap \phi^{-1}(L_{\varepsilon,\delta})|}{|S_n G^*|} = \frac{|\{\sigma \in S_n : \phi(G_\sigma^*) \in L_{\varepsilon,\delta}\}|}{|S_n|} \geqslant \frac{2}{dn}.$$

Summing the inequality $|S_n G^* \cap \phi^{-1}(L_{\varepsilon,\delta})| \geqslant \frac{2}{dn}|S_n|$ over the orbits of the $S_n$ action in $\mathcal{G}_{d,c}^{*\,\varepsilon,\delta}(n)$ we obtain the estimate

$$|\mathcal{G}_{d,c}^{*\,\varepsilon,\delta}(n)| \leqslant \frac{dn}{2}|\phi^{-1}(L_{\varepsilon,\delta})|,$$

which, combined with Lemma 6, yields

$$\begin{aligned}
|\mathcal{G}_{d,c}^{\varepsilon,\delta}(n)| = \frac{|\mathcal{G}_{c,d}^{*\,\varepsilon,\delta}(n)|}{(d!)^n} &\leqslant \frac{dn}{2}\frac{|\phi^{-1}(L_{\varepsilon,\delta})|}{(d!)^n}\\
&\leqslant \frac{dn}{2}2^{dn/2}\frac{(dn)^{dn/2}}{(d/e)^{dn}}\left(\frac{d^2}{n}\right)^{T_c + \frac{\varepsilon\delta d^2 n}{6d+6} - 1}\\
&= \frac{dn}{2}(\sqrt{2}e)^{dn}\left(\frac{n}{d}\right)^{\frac{dn}{2} - T_c + 1} d^{T_c - 1}\\
&= (\sqrt{2}e)^{dn}\left(\frac{n}{d}\right)^{\frac{dn}{2} - T_c - \frac{\varepsilon\delta d^2 n}{6d+6}} d^{T_c + \frac{\varepsilon\delta d^2 n}{6d+6}}\cdot\frac{n^2}{2d}.
\end{aligned}$$

Taking log of both sides finishes the proof. □

### 2.2.1 Fixed $d$

Let us say that a node in $G$ is bad if it is not in a $(d+1)$-clique, but it is in a triangle. The following statement is a (very) slight strengthening of Theorem 2.

**Theorem 10.** *Let $\varepsilon > 0$ and $d$ fixed. Among all d-regular graphs with at least $c \cdot T_{\max}$ 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, whose proof we omit.

**Lemma 11.** *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 10.* Let us set $\delta = 1/d$ and call $1/d$-bad edges simply "bad". Suppose now that more than $\varepsilon n$ nodes of $G$ are bad. Each bad node, by definition, is adjacent to at least two bad edges, so there are at least $\varepsilon n$ bad edges. Thus $G \in \mathcal{G}_{d,c}^{\frac{2\varepsilon}{d}, \frac{1}{d}}(n)$. Then, Lemma 9 combined with Theorem 1 and the fact that $d = O(1)$ gives

$$\log \frac{|\mathcal{G}_{d,c}^{\frac{2\varepsilon}{d}, \frac{1}{d}}(n)|}{|\mathcal{G}_{d,c}(n)|} = -\frac{\frac{2\varepsilon}{d}\frac{1}{d}d}{3d+3}\frac{dn}{2}\log n + O(dn \log d) = -\frac{\varepsilon}{3d+3}n \log n + O(n),$$

so indeed

$$\lim_{n \to \infty} \frac{|\mathcal{G}_{d,c}^{\frac{2\varepsilon}{d}, \frac{1}{d}}(n)|}{|\mathcal{G}_{d,c}(n)|} = 0,$$

as long as $\varepsilon \log n \to \infty$, proving that with high probability a graph conditioned on having at least $c \cdot T_{\max}$ triangles has $o(n)$ bad nodes, hence consists almost completely of $(d+1)$-cliques and a triangle-free part. $\qquad \square$

### 2.2.2 Growing $d$

An immediate generalization of Theorem 10 cannot hold for the $d \gg 1$ case, because one can exhibit a family of $d$-regular graphs with $c \cdot T_{\max}$ triangles that contain no cliques at all, yet have the optimal, $(1-c)(d/2)n \log \frac{n}{d}$, logarithmic growth rate. Such a family can be built, for example, by taking the disjoint union of many copies of $H$, together with a random $d$-regular graph, where $H$ is $K_{d+2}$ minus a perfect matching. Realizing the required $c \cdot T_{\max}$ triangles takes up only slightly more space this way than using copies of $K_{d+1}$, and the resulting decrease in the size of the random part is small enough that it does not affect the logarithmic growth rate. One can push this even further, and use disjoint $d + o(d)$ size components (these still contain roughly $\binom{d}{3}$ triangles each), and a large random $d$-regular part of the appropriate size.

We will show in this section, that a typical graph in the ensemble does, in fact, resemble an element of this last family. The main reason the previous argument fails for $d \gg 1$ is because now we cannot choose $\delta$ to be too small in Lemma 9, otherwise the gain will be less in magnitude than the error term $O(dn \log d)$. Nevertheless, if $\log d / \log n$ is small,

then the gap between the main term and the error term allows us to choose both $\varepsilon$ and $\delta$ to be small, which will be enough to learn something about the typical graphs in the ensemble. In particular, we can choose

$$\varepsilon = \delta^2 = (3c)^{2/3} \cdot \left( \frac{\log d}{\log \frac{n}{d^2}} \right)^{1/2}. \tag{11}$$

Then Lemma 9 implies that in a typical $d$-regular graph with at least $c \cdot T_{\max}$ triangles, most edges are incident to 0 or almost $d$ triangles. As it turns out, this implies a structural description similar to that of Theorem 10. Let us first informally explain the result. We call a subgraph $H \subset G$ a *dense spot* if $|H| \leqslant d + 1$ and $\deg_H(x) = d(1 - O(\delta))$ for all $x \in H$. Dense spots satisfy the following simple, combinatorial observations:

- Two dense spots are either disjoint, or they intersect in $d(1 - O(\delta))$ nodes.

- Intersection is transitive: if $H_1 \cap H_2 \neq \emptyset$ and $H_2 \cap H_3 \neq \emptyset$ then $H_1 \cap H_3 \neq \emptyset$.

- The union of a maximal, pairwise intersecting, family of dense spots has size $d(1 + O(\delta))$. We call these *pseudo-cliques*.

- It follows that any two pseudo-cliques must be disjoint.

The following is a restatement of Theorem 3.

**Theorem 12.** *Let $1 \ll d \ll n$. Let $\delta$ as in (11), and assume $\delta < 1/16$. With high probability, a random d-regular graph with at least $c \cdot T_{\max}$ triangles contains $(1 + O(\delta))cn/d$ pseudo-cliques. These pseudo-cliques contain $1 - O((\varepsilon + \delta)/c)$ fraction of all triangles.*

*Remark* 13. Theorem 12 is the strongest when $\log d = o(\log n)$, as in this case both $\varepsilon$ and $\delta$ are $o(1)$. However, when $d = n^\beta$ then $\delta = (3c)^{1/3} \left( \frac{\beta}{1-2\beta} \right)^{1/4}$, so we still get a non-trivial structural result when $\beta$ is small enough.

*Proof.* We set $\varepsilon$ and $\delta$ according to (11). Then, a careful calculation using Lemma 9 shows that we have

$$\lim_{n \to \infty} \frac{|\mathcal{G}_{d,c}^{\varepsilon,\delta}(n)|}{|\mathcal{G}_{d,c}(n)|} = 0,$$

so it is enough to consider a graph $G \in G_{d,c}(n) \setminus G_{d,c}^{\varepsilon,\delta}(n)$. The graph $G$ then has, by definition, less than $\varepsilon(d/2)n$ edges that are $\delta$-bad. Let us call a $\delta$-bad edge *bad* for brevity, and other edges *good*. Let us start by removing all edges with $t_e = 0$ from $G$, and denote the remaining graph by $G'$. Removing such edges does not change the $t_e$ value of the remaining edges. Let us call a node $v \in G'$ bad, if it is incident to at least $\delta d$ bad edges. Then, since $\varepsilon = \delta^2$, it follows that $G'$ cannot have more than $\delta n$ bad nodes.

The total number of triangles that are incident to either a bad edge or a bad node is at most $\varepsilon(d/2)n \cdot d + \delta n \binom{d}{2} = O((\varepsilon + \delta) \cdot T_{\max})$. We will show that the rest of the triangles are concentrated in pseudo-cliques.

**Definition 14.** A subgraph $H \subset G$ is a *dense spot* if $|V(H)| \leqslant d+1$ and each node $x \in H$ has $\deg_H(x) \geqslant (1-4\delta)d$.

**Claim 15.** *Let $H_1, H_2$ be dense spots. Then they are either disjoint, or $|H_1 \cap H_2| \geqslant (1-8\delta)d$. This follows from the fact the nodes in the intersection must have degree $\leqslant d$.*

**Claim 16.** *Let $H_1, H_2, H_3$ be dense spots. If $H_1 \cap H_2 \neq \emptyset$ and $H_2 \cap H_3 \neq \emptyset$ then $H_1 \cap H_3 \neq \emptyset$, since otherwise we would have $d+1 \geqslant |H_2| \geqslant |H_2 \cap H_1| + |H_2 \cap H_3| \geqslant 2d(1-8\delta)$ which contradicts $\delta < 1/16$.*

**Definition 17.** A subgraph $K \subset G$ is a *pseudo-clique* if there is a maximal family $\mathcal{H}$ of pairwise intersecting dense spots such that $K = \cup_{H \in \mathcal{H}} H$.

**Claim 18.** *By definition, any dense spot $H$ is either disjoint from, or fully contained in, a pseudo clique $K$. Furthermore, any two distinct pseudo-cliques are disjoint.*

**Lemma 19.** *If $K$ is a pseudo-clique then $|K| \leqslant \frac{1-8\delta}{1-13\delta}(d+1) \leqslant (1+30\delta)d$.*

*Proof of Lemma 19.* Let $H \subset K$ be one of the dense spots in $K$. For any node $x \in H$ we have $\deg_H(x) \geqslant (1-4\delta)d$. But $\deg(x) = d$, thus the total number of edges going between $H$ and $K \setminus H$ is at most $|H| \cdot 4\delta d \leqslant 4\delta d(d+1)$. However, each node $y \in K \setminus H$ is contained in a dense spot $H'$, and thus $\deg_{H'}(y) \geqslant (1-4\delta)d$. Since $|H' \setminus H| \leqslant 8\delta d + 1 \leqslant 9\delta d$, we get that at least $(1-13\delta)d$ edges go from $y$ to $H$. Hence

$$|K \setminus H|(1-13\delta)d \leqslant 4\delta d(d+1),$$

from which

$$|K| \leqslant |H| + \frac{4\delta(d+1)}{1-13\delta} \leqslant \frac{1-8\delta}{1-13\delta}(d+1).$$

Finally

$$\frac{1-8\delta}{1-13\delta} = 1 + \frac{5\delta}{1-13\delta} \leqslant 1 + 5\frac{16}{16-13}\delta \leqslant 1 + 30\delta$$

when $\delta \leqslant 1/16$. $\qquad\square$

To finish the proof of Theorem 12, we need to show that any triangle that is only incident to good edges and good nodes is contained in a pseudo-clique. We will show slightly more: that a good edge connecting good nodes is in a pseudo-clique.

Consider a good edge $uv$ in $G'$, where both $u$ and $v$ are good nodes. Since we already removed the edges with no triangles, $t_{uv} \geqslant d - \delta d$. In particular $u$ and $v$ share at least $d - \delta d$ common neighbors. Each of $u$ and $v$ may be incident to at most $\delta d$ bad edges. That means that the subset $H_0$ of common neighbors of $u$ and $v$ that are connected to both of them via good edges has size $|H_0| \geqslant d - 3\delta d$. Let $H = H_0 \cup \{u, v\}$. We claim $H$ is a dense spot. Clearly $|H| \leqslant 1 + \deg(u) = d+1$, and by construction, $\deg_H(u), \deg_H(v) \geqslant (1-3\delta)d \geqslant (1-4\delta)d$. What remains to show is that for any node $x \in H_0$ we have $\deg_H(x) \geqslant (1-4\delta)d$. But $xu$ is a good edge, hence $x$ and $u$ have at least $(1-\delta)d$ common neighbors, or equivalently, at most $\delta d$ of $u$'s neighbors are not connected to $x$. Thus $x$ is connected to at least $(1-4\delta)d$ nodes in $H$, proving that indeed $H$ is a dense spot. So the $uv$ edge is contained in a dense spot, and thus in a pseudo-clique. $\quad\square$

### 2.3  $k$-cliques

We can easily extend the above results from triangles to $k$-cliques. Let $\mathcal{G}_{d,c,k}(n)$ denote the set of $d$-regular graphs on $n$ nodes that contain at least $c \cdot \binom{d}{k-1}\frac{n}{k}$ subgraphs isomorphic to $K_k$. (The maximum possible number of subgraphs isomorphic to $K_k$ is clearly $\binom{d}{k-1}\frac{n}{k}$.)

*Proof of Theorem 4.* The idea is a simple reduction the the $k = 3$ case. Clearly, each $G \in \mathcal{G}_{d,c,k}(n)$ has at least

$$c \cdot \binom{d}{k-1}\frac{n}{k} \frac{\binom{k}{3}}{\binom{d-2}{k-3}} = c \cdot \binom{d}{2}\frac{n}{3} = c \cdot T_{\max}$$

triangles, so $\mathcal{G}_{d,t_k,k}(n) \subset \mathcal{G}_{d,c}(n)$, which implies the upper bound of the theorem. On the other hand, the family of graphs constructed in Theorem 1 contain

$$b\binom{d+1}{k} = c \cdot \frac{n}{d+1}\binom{d+1}{k} = c \cdot \binom{d}{k-1}\frac{n}{k}$$

$k$-cliques, so this family is contained in $\mathcal{G}_{d,c,k}(n)$, implying the lower bound of the theorem. Finally, the structural statement follows directly from Theorem 10. $\square$

### Acknowledgements

## References

[1] Fanny Augeri. Nonlinear large deviation bounds with applications to traces of Wigner matrices and cycles counts in Erdős–Rényi graphs. *Preprint*, 2018, arXiv:1810.01558.

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

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

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

[5] Pierre Collet and Jean-Pierre Eckmann. The number of large graphs with a positive density of triangles. *Journal of Statistical Physics*, 109(5):923–943, 2002.

[6] Nicholas Cook and Amir Dembo. Large deviations of subgraph counts for sparse Erdős–Rényi graphs. *Advances in Mathematics*, 373:107289, 2020.

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

[8] Matan Harel, Frank Mousset, and Wojciech Samotij. Upper tails via high moments and entropic stability. *Preprint*, 2019, arXiv:1904.08212.

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

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

[11] Svante Janson and Andrzej Ruciński. The deletion method for upper tail estimates. *Combinatorica*, 24(4):615–640, 2004.

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

[13] 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.

[14] Jeong Han Kim, Benny Sudakov, and Van Vu. Small subgraphs of random regular graphs. *Discrete Mathematics*, 307(15):1961–1967, 2007.

[15] 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.

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

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

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

[19] Brendan D. McKay and Nicholas C. Wormald. Asymptotic enumeration by degree sequence of graphs with degrees o (n 1/2). *Combinatorica*, 11(4):369–382, 1991.

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

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

[22] 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.