Triangles and subgraph probabilities in random regular graphs

Pu Gao

Submitted: Mar 3, 2021; Accepted: Jun 20, 2023; Published: Jan 12, 2024
© The author. Released under the CC BY-ND license (International 4.0).

Abstract

We improve the estimates of the subgraph probabilities in a random regular graph. Using the improved results, we further improve the limiting distribution of the number of triangles in random regular graphs.

Mathematics Subject Classifications: 05C80, 05C30

1 Introduction

Research in random graph theory started from the study of subgraphs [2]. The distributional results for small subgraphs and for large subgraphs are very different in nature. The distributions of small subgraphs in $\mathcal{G}(n, p)$ are well understood [18, 19]. Under some mild conditions, the number of subgraphs isomorphic to a fixed graph $H$, denoted by $Z_H$, is asymptotically normally distributed when $p$ exceeds some critical value $p_H$. Similar results hold for $\mathcal{G}(n, m)$ as well [18]. In particular, if $H$ is a balanced graph and $p$ is not too close to 1, then $Z_H$ is asymptotically normally distributed in $\mathcal{G}(n, p)$, and is thus highly concentrated around its expectation $\mathbb{E}Z_H$, when $\mathbb{E}Z_H \to \infty$ as $n \to \infty$. The picture for large subgraphs is very different. Take $Z_M$, the number of perfect matchings (assuming $n$ is even) for an example. The threshold of its appearance is at $p = \log n/n$. However, $\mathbb{E}Z_M \to \infty$ already when $p = C/n$ for $C > e$. This implies no concentration of $Z_M$ around its expectation for $p = O(1/n)$, and indeed no such concentration for larger $p$ either. It has been proved that $Z_M$ is asymptotically log-normally distributed for $p \gg n^{-1/2}$ [13].

For smaller values of $p$ its distribution is unknown. Similar results have been proved for the number of spanning trees and the number of Hamilton cycles [13], the number of $d$-factors [4], and the number of triangle factors [5]. What happens to the number of subgraphs whose size is between constant size and size of order $\Omega(n)$? For the number of matchings of size $\ell$, it turns out that its limiting distribution transitions from normal to log-normal at some critical point of $p$ [9], and this distributional transition may be a general phenomenon for other subgraphs as well.

University of Waterloo (pu.gao@uwaterloo.ca).
Another extensively studied random graph model is the model of random regular graphs. Let $\mathcal{G}(n, d)$ denote a graph chosen uniformly at random from the set of all $d$-regular graph on vertex set $[n]$. Much less is known about the distribution of $Z_H$ in $\mathcal{G}(n, d)$ where $H$ is of fixed size, or of size $\Omega(n)$. For $d = O(1)$, $\mathcal{G}(n, d)$ is locally tree-like and thus, the only nontrivial connected subgraphs of fixed sizes are cycles (see [21] for more discussions of properties of $\mathcal{G}(n, d)$ for fixed $d$). In $\mathcal{G}(n, d)$ the joint distribution of the number of cycles of any fixed length tends asymptotically to the distribution of independent Poisson random variables [1, 20]. This continues to hold for slowly growing $d$ [17]. For instance, the Poisson like distribution for the number of triangles is proved to hold for all $d = o(n^{1/5})$. Z. Gao and Wormald [11] proved normal distribution for strictly balanced graphs, which, in cases of short cycles, permits slightly larger $d$ than in [17]. For the number of triangles, their result holds for $d = o(n^{2/7})$. For large subgraphs, Janson [12] determined the limiting distribution of the number of perfect matchings and the number of Hamilton cycles in $\mathcal{G}(n, d)$ for constant $d$. For growing $d$, we are not aware of any results when the size of $H$ is beyond $\log n$. The recent development on the sandwich conjecture [6, 3] allows to translate properties and graph parameters from $\mathcal{G}(n, p)$ to $\mathcal{G}(n, d)$, but in general the distributional results do not translate.

There are two main obstacles in proving the limiting distribution of $Z_H$, even for $H$ of fixed size. The first obstacle lies in the computation of the moments of $Z_H$. To prove normality, one can compute its central $j$-th moment for each fixed $j$. However, due to the appearance of both positive and negative terms in the central moments computation, the leading terms cancel each other, and the errors in the subgraph probabilities quickly become the dominating terms when $d$ grows. (Take the triangle $H = C_3$ for an example. We only manage to compute the variance of $Z_{C_3}$ for $d = o(n^{2/5})$, and computing higher moments requires even smaller $d$.) The other standard method is computing the $k$-th raw moment for $k$ around $\mathbb{E}Z_H / \sigma(Z_H)$, where $\sigma(Z_H)$ is the standard deviation of $Z_H$. For such large $k$ (which typically grows quickly with $d$), it is difficult to bound the number of $k$-tuples of copies of $H$ sharing a specific number of edges. The second obstacle is to estimate the probabilities of subgraphs in $\mathcal{G}(n, d)$. Specifying a set of edges $H \subseteq \mathcal{G}(n, d)$ (i.e. $\mathcal{G}(n, d)$ contains all edges in $H$)? McKay [14, 15] gave an estimate of this probability which has a relative error $O(|H|d/n)$, and to our knowledge, this has been the best estimate to date. This immediately limits the study of $Z_H$ for large $H$, or the computation of high moments of $Z_H$ even for $H$ of constant size, which is required when proving the limiting distribution of $Z_H$. Take $H = C_3$ as an example. To determine the limiting distribution of $Z_{C_3}$, we need to estimate the $k$-th moment of $Z_{C_3}$ where $k \approx d^{3/2}$. This requires an asymptotic joint probability estimate of $\Theta(d^{3/2})$ edges. With errors in [14, 15], $d$ is required to be $o(n^{2/5})$.

For the second obstacle, we manage to improve the error in the subgraph probabilities in [14, 15]. See Theorem 6 in Section 2 for the precise statement. This new result allows us to estimate, for instance, the asymptotic probability that $\mathcal{G}(n, d)$ contains a specific perfect matching for $d$ up to $n^{1/2}$. As we only study small subgraphs in this paper, it turns out that a less precise form would be sufficient, which is stated in Theorem 2 as a corollary of Theorem 6.
For the first obstacle, Z. Gao and Wormald bypassed it by smartly considering a new variable $Z_H$, which counts isolated copies of $H$, i.e. copies of $H$ that do not share edges with other copies of $H$. In the range of $d$ where the number of non-isolated copies of $H$ is so small that it does not affect the standard deviation of $Z_H$, the distribution of $Z_H$ will immediately yield the distribution of $Z_H$. Working on $Z_H$ avoids having to deal with tuples of heavily intersecting copies of $H$. However, it also limits the range of $d$ for which the proof method can apply. For the case $H = C_3$, $d = o(n^{2/7})$ is required in [11], which did not even reach the natural stopping point $d = o(n^{2/5})$ (explained above). To relax the condition on $d$, it is possible to modify the definition of $Z_H$. Instead of counting isolated triangles, we may, for instance, count triangles that intersect only a bounded number of other triangles. With the new definition and with the new subgraph probabilities in Theorem 2, it is possible to get beyond $d = o(n^{2/7})$. However it is not sufficient to reach $d = n^{1/2-\alpha(1)}$. In this paper, we handle $Z_H$ directly, and prove the normality of the number of triangles for all $d = O(n^{1/2})$ where $d \to \infty$.

We also computed the variance of $Z_H$ when $H$ is a fixed strictly balanced graph. Surprisingly, although the variance seemingly carries less information than the limiting distribution and thus computing it should be easier than determining the limiting distribution, we are not able to determine the variance of the number of triangles for $d$ up to $n^{1/2}$. In fact, we can only compute the variance for $d = o(n^{2/5})$. Note that the limiting distribution does not imply the variance, as it is not sensitive to a lottery effect (events affecting $Z_H$ that occur with a tiny probability), whereas the variance is, although we do not expect such a lottery effect for the variance either. Another surprising discovery is that the number of triangles seems to behave more like the sum of a set of independent variables in $G(n,d)$ than in $G(n,p)$, whereas the intuition is the other way around. Let $X_{ijk}$ be the indicator variable that the three vertices $i, j, k$ induce a triangle. Then the number of triangles is $\sum X_{ijk}$ where the sum is over all 3-subsets of $[n]$. In $G(n,p)$, $X_{ijk}$ is independent of all but those $X_{uvw}$ such that $|\{i, j, k\} \cap \{u, v, w\}| = 2$. In $G(n,d)$, due to the dependency between edges, all these indicator variables are correlated. However, surprising cancellations of leading terms appear in the calculations for the variance of the number of triangles (see Section 4.2 for details). The variance of the number of triangles turns out smaller in $G(n,d)$ than in $G(n,d/n)$.

One of the challenges in computing the moments of the number of triangles is to bound the number of $k$-tuples of triangles where $k \approx d^{3/2}$, such that these $k$ triangles induce a large number of “holes” (see Definition 16 in Section 5.2). The bound and the arguments in this paper do not easily extend to other strictly balanced graphs in general. Extending the distributional result for the number of triangles to other strictly balanced graphs requires a smarter treatment for the holes.
2 Main results

2.1 Improved subgraph probabilities

Let \( \mathbf{d} = (d_1, \ldots, d_n) \) be a degree sequence and let \( G(n, \mathbf{d}) \) denote a uniformly random graph with degree sequence \( \mathbf{d} \). Let \( H_1 \) and \( H_2 \) be two disjoint graphs on \([n]\), i.e. \( E(H_1) \cap E(H_2) = \emptyset \). Let \( H^+_1 \) denote the event that \( H_1 \subseteq G(n, \mathbf{d}) \) and let \( H^-_2 \) denote the event that \( H_2 \cap G(n, \mathbf{d}) = \emptyset \). The conditional edge probability \( P(uv \in G(n, \mathbf{d}) \mid H^+_1, H^-_2) \) was first given by McKay [14, 15], which immediately applies to give an asymptotic estimate of \( P(H \subseteq G(n, \mathbf{d})) \) when \( H \) is not too large. This result was recently extended by Ohapkin and the author [8] to more general degree sequences, with the same order of the error term \( O(\Delta(\mathbf{d})/n) \) as in [8] (\( \Delta(\mathbf{d}) \) denotes the maximum degree in \( \mathbf{d} \)). Given a graph \( H \) on \([n]\), let \( \mathbf{d}^H = (d^H_1, \ldots, d^H_n) \) denote the degree sequence of \( H \) and let \( |H| \) denote the number of edges in \( H \). Let \( \Delta_H \) denote the maximum degree of \( H \). Given two degree sequences \( \mathbf{d} \) and \( \mathbf{d}' \), we say \( \mathbf{d} \preceq \mathbf{d}' \) if \( d_i \leq d'_i \) for every \( 1 \leq i \leq n \). Let \( \Delta = \Delta(\mathbf{d}) \) be the maximum component of \( \mathbf{d} \) and let \( M = \sum_{i=1}^n d_i \). The following result follows as a corollary of [8, Theorem 1].

**Theorem 1.** Let \( \mathbf{d} \) be such that \( \Delta^2 = o(M) \). Let \( H_1 \) and \( H_2 \) be two disjoint graphs on \([n]\) where \( \mathbf{d}^H_1 \preceq \mathbf{d} \). Suppose that \( \Delta_{H_2} = O(\Delta) \) and \( M - 2|H_1| = \Omega(M) \). Suppose further that \( uv \notin H_1 \cup H_2 \). Then,

\[
P(uv \in G(n, \mathbf{d}) \mid H^+_1, H^-_2) = \left( 1 + O\left( \frac{\Delta^2}{M} \right) \right) \frac{(d - d^H_u)(d - d^H_v)}{M - 2|H_1|}.
\]

One of the main results of this paper is an estimate of the above conditional probability with a relative error \( \Delta^2/n^3 \) (See Theorem 6). Due to the technical statement of the theorem, we start our discussion from a special and simpler case where \( \mathbf{d} = (d, \ldots, d) \) and \( H_2 = \emptyset \). Given graph \( H \) and \( u \in [n] \), let \( \mathcal{N}_H(u) \) denote the set of vertices that are adjacent to \( u \) in \( H \). The theorem below gives an approximation of \( P(uv \in G(n, \mathbf{d}) \mid H^+) \) with a less sharp error compared to Theorem 6 below. However, it is sufficient for studying \( Z_H \) in \( G(n, \mathbf{d}) \) when \( H \) is not too big.

**Theorem 2.** Let \( H \) be a graph on \([n]\) where \( d^H_i \leq d \) for every \( i \). Suppose that \( d = o(n) \), \( dn - |H| = \Omega(dn) \), and that \( uv \notin H \). Let \( \mathbf{d} = \mathbf{d} - \mathbf{d}^H \) where \( \mathbf{d} = (d, \ldots, d) \). Then,

\[
P(uv \in G(n, \mathbf{d}) \mid H^+) = \frac{d_u d_v}{dn} \left( 1 - \frac{\phi_H(uv)}{dn} \right) \left( 1 + O\left( \frac{|H|}{n^2} + \frac{|H|^2}{n^4} + \frac{d^2}{n^2} \right) \right), \tag{1}
\]

where

\[
\phi_H(uv) = -d - 2|H| - (d - 1)(d^H_u + d^H_v) + d^H_u d^H_v + \sum_{x \in \mathcal{N}_H(u)} d^H_x + \sum_{y \in \mathcal{N}_H(v)} d^H_y. \tag{2}
\]

**Remark 3.** By the definition of \( \phi_H(uv) \) it is immediate that if the maximum degree of \( H \) is \( O(1) \) then

\[
\phi_H(uv) = -2|H| - d(1 + d^H_u + d^H_v) + o(1), \tag{3}
\]
and this bound is sufficiently accurate in many applications. Given $H$, $\phi_H(uv)$ can be easily computed by examining the neighbourhood of $u$ and $v$ in $H$. In Figure 1 we give a few examples of the value of $\phi_H(uv)$. Edge $uv$ is coloured red, and edges in $H$ are coloured black. The value of $\phi_H(uv)$ is given next to the edge $uv$.

Remark 4. We check that $\mathbb{P}(uv \in G(n, d))$ from Theorem 2 agrees with what we know about this probability. By Theorem 2,

$$\mathbb{P}(uv \in G(n, d)) = \frac{d}{n} \left(1 + \frac{d}{dn}\right) \left(1 + O\left(d^2/n^2\right)\right) = \frac{d}{n-1} \left(1 + O\left(d^2/n^2\right)\right),$$

whereas by symmetry, we know that $\mathbb{P}(uv \in G(n, d)) = d/(n-1)$. Hence the estimate from the theorem is accurate up to a relative error $O(d^2/n^2)$, as it is supposed to.

The following corollary on the upper bound of $\mathbb{P}(uv \in G(n, d) \mid H^+)$ will be useful in the study of the distribution of the number of triangles. The proof of the corollary is given in Section 5.1.

**Corollary 5.** Let $\mathbf{d} = (d, d, \ldots, d)$ where $dn$ is even and $d = o(n)$. Let $H$ be a graph on $[n]$ with degree sequence $\mathbf{d}^H \preceq \mathbf{d}$ such that $dn - 2|H| = \Omega(dn)$. Suppose $F \subseteq K_n \setminus H$ where $|F| = O(1)$. Let $\lambda_F = \mathbb{P}(F \subseteq G(n, d))$ and $\mathbf{d} = \mathbf{d} - \mathbf{d}^H$. Then,

$$\mathbb{P}(F \subseteq G(n, d) \mid H^+) \leq \lambda_F \left(1 + O\left(\frac{1}{n} + \frac{|H|}{dn} + \frac{d^2}{n^2}\right)\right). \quad (4)$$

Theorem 2 is a special and less precise version of the following more general result. Given $\mathbf{d}$, recall that $M = M(\mathbf{d}) = \sum_{i=1}^{n} d_i$. Let $M_j = M_j(\mathbf{d}) = \sum_{i=1}^{n} (d_i)_j$ for any integer $j \geq 2$.

**Theorem 6.** Let $\mathbf{d} = (d_1, \ldots, d_n)$ be such that $M$ is even and $\Delta^2 = o(M)$, where $\Delta = \Delta(\mathbf{d})$. Let $H_1$ and $H_2$ be disjoint graphs with degree sequences $\mathbf{d}^{H_1} \preceq \mathbf{d}$ and $\Delta_{H_2} = O(\Delta)$. Let $\tilde{\mathbf{d}} = \mathbf{d} - \mathbf{d}^{H_1}$, and let $\tilde{M} = M(\mathbf{d})$ and $\tilde{M}_j = M_j(\tilde{\mathbf{d}})$ for $j \geq 2$. Suppose $\tilde{M} = \Omega(M)$, and $uv \notin H_1 \cup H_2$. Then

$$\mathbb{P}(uv \in G(n, \mathbf{d}) \mid H_1^+, H_2^-)$$

$$= \left(1 + O\left(\frac{\Delta^2 n}{M^3}\right)\right) \frac{\tilde{d}_u \tilde{d}_v}{\tilde{M}} \left(1 - \frac{\tilde{M}_2}{M^2} - \frac{\tilde{M}_2}{M^2} - \sum_{(x,y) \in W} \frac{\tilde{d}_x \tilde{d}_y}{M^2}\right) \left(1 - \frac{\tilde{\phi}_{H_1, H_2}(uv)}{M}\right),$$

Figure 1: $\phi_H(uv)$
where
\[
\tilde{\phi}_{H_1,H_2}(uv) = -2(\tilde{d}_u + \tilde{d}_v) + 2 - \sum_{x \in N_{H_1 \cup H_2}(u)} \tilde{d}_x - \sum_{y \in N_{H_1 \cup H_2}(v)} \tilde{d}_y - \frac{(\tilde{d}_u + \tilde{d}_v - 2)\tilde{M}_2}{M} + \tilde{d}_u \tilde{d}_v,
\]
and \(W\) is the set of \((x, y)\) such that \(xy \in H_1 \cup H_2\), \(xu, yv \notin H_1 \cup H_2\), and \(\{x, y\} \cap \{u, v\} = \emptyset\).

Remark 7. To study \(Z_H\) for large \(H\), for instance, if \(H\) is a perfect matching, then the error \(O(|H|^2/d^2n^2)\) in Theorem 2 is too large, as the cumulative error becomes \(O(|H|^3/d^2n^2) = O(n/d^2)\) which is \(\Omega(1)\) for \(d = O(\sqrt{n})\). In such cases Theorem 6 needs to be applied.

Remark 8. A result by McKay [16, Theorem 4.6] applies to find the probability \(\mathbb{P}(H \subseteq \mathcal{G}(n, d))\) with a relative error \(d^2/n\). Theorem 6 applies to find this probability with a smaller relative error \(d^2|H|/n^2\).

### 2.2 Proof of Theorem 2

We prove that Theorem 2 follows as a corollary of Theorem 6. Apply Theorem 6 with \(M = dn\), \(H_1 = H\) and \(H_2 = \emptyset\), and note that \(M = M - 2|H|, \tilde{M}_2 = M_2 + O(d|H|)\), and thus,
\[
\frac{\tilde{M}_2}{M^3} = \frac{(d - 1)^2}{dn} \left(1 + O\left(\frac{|H|}{dn}\right)\right), \quad \frac{\tilde{M}_2}{M^2} = \frac{d - 1}{dn} \left(1 + O\left(\frac{|H|}{dn}\right)\right),
\]
and
\[
\frac{\tilde{M}_2}{M} = (d - 1) \left(1 - O\left(\frac{|H|}{dn}\right)\right)\quad \text{and} \quad \sum_{(x, y) \in W} \frac{\tilde{d}_x \tilde{d}_y}{M^2} = O\left(\frac{d^3|H|}{M^2}\right) = O\left(\frac{|H|}{n^2}\right).
\]

By letting \(\bar{\phi}_H = \tilde{\phi}_{H, \emptyset}\) and using the approximation of \(\tilde{M}_2/M\) above, it is straightforward to find that \(\bar{\phi}_H(uv) = \phi_H(uv) + d - d^2 + 2|H| + O(dn^2/|H|)\). Hence,
\[
\frac{1}{dn - 2|H|} \left(1 - \frac{\bar{\phi}_H(uv)}{dn - 2|H|}\right)
\]
\[
= \frac{1}{dn} \left(1 + \frac{2|H|}{dn} + O\left(\frac{|H|^2}{d^2n^2}\right)\right) \left(1 - \frac{\phi_H(uv)}{dn} + O\left(\frac{|H|}{dn}\right)\right)
\]
\[
= \frac{1}{dn} \left(1 - \frac{\phi_H(uv) - 2|H|}{dn} + O\left(\frac{|H|^2}{d^2n^2} + \frac{|H|}{n^2}\right)\right)
\]
\[
= \frac{1}{dn} \left(1 - \frac{\phi_H(uv) + d - d^2 + O(n^2/|H|)}{dn} + O\left(\frac{|H|^2}{d^2n^2} + \frac{|H|}{n^2}\right)\right)
\]
as \(\bar{\phi}_H(uv) = O(d^2)\). Moreover,
\[
1 - \frac{\tilde{M}_2^2}{M^3} = \frac{\tilde{M}_2}{M^2} - \sum_{(x, y) \in W} \frac{\tilde{d}_x \tilde{d}_y}{M^2} = \left(1 - \frac{d - 1}{n}\right) \left(1 + O\left(\frac{|H|}{n^2}\right)\right).
\]

Theorem 2 follows by noting further that \(d^2n/M^3 = O(d^2/n^2)\) and
\[
\left(1 - \frac{\phi_H(uv) + d - d^2}{dn}\right) \left(1 - \frac{d - 1}{n}\right) = \left(1 - \frac{\phi_H(uv)}{dn}\right) \left(1 + O\left(\frac{d^2}{n^2}\right)\right).
\]
2.3 Expectation and variance of strictly balanced subgraphs

Let $H$ be a fixed graph. Let

$$t = |V(H)| \quad \text{and} \quad h = |H|.$$  

Let $Z_H$ denote the number of subgraphs of $G(n, d)$ that are isomorphic to $H$; each such subgraph is called a copy of $H$. Let

$$\mu_H = \mathbb{E}Z_H, \quad \text{and} \quad \sigma_H^2 = \text{Var}Z_H.$$  

Define, for each $1 \leq j \leq h - 1$,

$$\rho_H(j) = \sup \left\{ \rho : |V(H')| \geq t \left( \frac{|H'|}{|H|} + \rho \right) \right\} \text{ for all subgraphs } H' \text{ of } H \text{ where } |H'| = j \right\}.$$  

See Example 10 below for values of $\rho_H$ where $H$ is a cycle or a clique.

Given a real number $x$ and a nonnegative integer $k$, let $(x)_k = \prod_{j=0}^{k-1}(x-j)$. Let $\text{aut}(H)$ denote the size of the automorphism group of $H$.

**Theorem 9.** Let $H$ be a fixed graph with $t$ vertices and $h$ edges. Then, the expectation of $Z_H$ is

$$\mu_H = \frac{(n)_t}{\text{aut}(H)} \varphi(d, H) \left( 1 - \frac{A_H d + B_H}{dn} + O(d^2/n^2) \right),$$

where

$$\varphi(d, H) = \prod_{v \in V(H)} (d)_{d_H(v)},$$

and $A_H$ and $B_H$ are constants depending only on $H$. Suppose further that $H$ is strictly balanced and

$$\frac{d^{h-1}}{n^{h-1}+1} = o(1), \quad \frac{d^{h+2}}{n^{h+2}} = o(1), \quad \frac{\mu_H^{1-j/h}}{n^{\rho_H(j)}} = o(1), \text{ for every integer } 1 \leq j \leq h - 1.$$  

Then,

$$\sigma_H^2 \sim \mu_H \sim \frac{(n)_t}{\text{aut}(H)} \varphi(d, H) \left( \frac{1}{(dn)^h} \right).$$

**Example 10.** (Cycles) Let $H = C_\ell$ where $\ell \geq 3$. Then, $\rho_{C_\ell}(j) = 1/\ell$ for every $1 \leq j \leq \ell - 1$. The assumptions of Theorem 9 are satisfied when $d^{\ell-1} = o(n)$ and $d^{\ell+2} = o(n^2)$. In particular, for $H = K_3$, the assumptions of Theorem 9 are satisfied when $d = o(n^{2/5})$.

(Cliques) Let $H = K_t$ where $t \geq 3$. Then, $\rho_{K_t}(j) = \frac{n(t-1)-2j}{t(t-1)}$ for every $(t-1)/2 < j \leq (t-1)$.

The assumptions of Theorem 9 are equivalent to

$$d \ll n^{1-2/(t+1)}, \quad d \ll n^{1-2t/(t^2-t+4)}, \quad d \ll n^{(t+s-3)/(t+s-1)},$$

which reduce to $d \ll n^{1-2t/(t^2-t+4)}$.

(Trees) Theorem 9 does not cover trees, as $d^{h-1} = o(n^{h-1})$ cannot be satisfied. Indeed the variance of $Z_H$ cannot be of order $\mu_H$ when $H$ is a tree. Consider $H = P_3$ to be a path of length three. Then $Z_H = d(d-1)^2n - 6Z_{C_3}$, and therefore $\sigma^2_{P_3} = 36\sigma^2_{C_3} \sim 36\mu_{C_3}$, which is much smaller than $\mu_{P_3}$.
2.4 Distribution of the number of triangles

Theorem 11. Suppose \( d = O(\sqrt{n}) \) and \( d = \omega(1) \). Let \( Z_{C_3} \) be the number of triangles in \( G(n, d) \). Then,

\[
\frac{Z_{C_3} - \mu_{C_3}}{\sqrt{\mu_{C_3}}} \xrightarrow{d} N(0, 1), \quad \text{as } n \to \infty.
\]

We will prove Theorem 6 in Section 3. The variance of \( Z_H \) for strictly balanced \( H \) will be estimated in Section 4. Finally, the limiting normal distribution of the number of triangles will be proved in Section 5. The proof for the case where \( d = o(\sqrt{n}) \) is given in Section 5.2 and the case where \( d = \Theta(\sqrt{n}) \) is treated in Section 5.3. The proof for Corollary 5 is presented in Section 5.1.

3 Proof of Theorem 6

The proof is obtained by expressing \( P(uv \in G(n, d) \mid H_1^{-1}, H_2^{-2}) \) by a function involving such conditional probabilities for other pairs \( u'v' \). Applying Theorem 1 for pairs \( u'v' \) yields a new estimate for \( P(uv \in G(n, d) \mid H_1^{-1}, H_2^{-2}) \) with improved error. By repeatedly applying this argument it is possible to further improve the relative error and to relax conditions on \( H_1^{-1} \) and \( H_2^{-2} \). However, the expression of the conditional probability would become more complicated after each iteration. We did not attempt this in the paper.

Let \( G \) be the class of graphs \( G \) in \( G(n, d) \) such that \( H_1^{-1} \subseteq G \) and \( H_2^{-2} \cap G = \emptyset \). Let \( G^+ \subseteq G \) be the set of graphs containing edge \( uv \) and let \( G^- = G \setminus G^+ \). We will estimate \( |G^+|/|G^-| \) by defining switchings that relate graphs in \( G^+ \) to graphs in \( G^- \) and counting the number of switchings that can be applied to a graph \( G \in G^+ \), and the number of switchings that can produce a graph \( G' \in G^- \).

Define the switching as follows. Given \( G \in G^+ \), a forward switching specifies an ordered pair of vertices \((x, y)\) such that

(a) \( xy \in G \setminus H_1 \) and \( u, v, x \) and \( y \) are all distinct;

(b) none of \( ux \) and \( vy \) is in \( G \cup H_2 \).

Then the forward switching replaces edges \( uv \) and \( xy \) by \( ux \) and \( vy \). The resulting graph \( G' \) is obviously in \( G^- \). The inverse operation which converts \( G' \) to \( G \) is called a backward switching. See Figure 2 for an illustration.

Let \( f(G) \) be the number of forward switchings that can be applied to \( G \). The number of ways to choose \((x, y)\) such that \( xy \in G \setminus H_1 \) is \( \widehat{M} \). Among these, the number of choices where \( x \in \{u, v\} \) is \( \widehat{d}_u + \widehat{d}_v \), and there are the same number of choices where \( y \in \{u, v\} \).

The number of choices where \( \{x, y\} = \{u, v\} \) is \( 2 \). Thus, by inclusion-exclusion, the number of choices for \((x, y)\) satisfying (a) is \( \widehat{M} - 2(\widehat{d}_u + \widehat{d}_v) + 2 \). To bound \( f(G) \) we will subtract the number of choices for \((x, y)\) satisfying (a) but not (b).

Let \( X_u \) be the number of choices of \((x, y)\) satisfying (a) where \( ux \) is an edge in \( G \cup H_2 \), and let \( X_v \) be the number of choices of \((x, y)\) satisfying (a) where \( vy \) is an edge in \( G \cup H_2 \).
Let $X_{uv}$ be the number of choices of $(x, y)$ satisfying (a) where both $ux$ and $vy$ are edges in $G \cup H_2$ and $xy \in G \setminus H_1$. Then,

$$f(G) = \tilde{M} - 2(\tilde{d}_u + \tilde{d}_v) + 2 - X_u - X_v + X_{uv}.$$ 

Claim 12. $\mathbb{E}X_{uv} = O(\Delta^4/M)$ and

$$\mathbb{E}X_u = \sum_{x \in \mathbb{N}_{H_1 \cup H_2}(u)} \tilde{d}_x + \frac{(\tilde{d}_u - 1)\tilde{M}_2}{\tilde{M}} + O(\Delta^4/M),$$

$$\mathbb{E}X_v = \sum_{y \in \mathbb{N}_{H_1 \cup H_2}(v)} \tilde{d}_y + \frac{(\tilde{d}_v - 1)\tilde{M}_2}{\tilde{M}} + O(\Delta^4/M).$$

By the claim,

$$\mathbb{E}f(G) = \tilde{M} - 2(\tilde{d}_u + \tilde{d}_v) + 2 - \sum_{x \in \mathbb{N}_{H_1 \cup H_2}(u)} \tilde{d}_x - \sum_{y \in \mathbb{N}_{H_1 \cup H_2}(v)} \tilde{d}_y$$

$$- \frac{(\tilde{d}_u + \tilde{d}_v - 2)\tilde{M}_2}{\tilde{M}} + O\left(\frac{\Delta^4}{\tilde{M}}\right)$$

$$= \left(1 + O\left(\frac{\Delta^4}{\tilde{M}^2}\right)\right) \cdot \Sigma,$$

where

$$\Sigma = \tilde{M} - 2(\tilde{d}_u + \tilde{d}_v) + 2 - \sum_{x \in \mathbb{N}_{H_1 \cup H_2}(u)} \tilde{d}_x - \sum_{y \in \mathbb{N}_{H_1 \cup H_2}(v)} \tilde{d}_y - \frac{(\tilde{d}_u + \tilde{d}_v - 2)\tilde{M}_2}{\tilde{M}}.$$

Next, we estimate $b(G')$, the number of backward switchings that can be applied to $G'$. To perform a backward switching, we choose $x$ and $y$ such that $xu \in G' \setminus H_1$ and $yv \in G' \setminus H_1$ and $xy \notin G' \cup H_2$. Then replace $ux, yv$ by $uv$ and $xy$. The number of ways to choose $(x, y)$ such that $xu \in G' \setminus H_1$ and $yv \in G' \setminus H_1$ is $\tilde{d}_u \tilde{d}_v$. Thus, $b(G') = \tilde{d}_u \tilde{d}_v - Y_1 - Y_2$ where $Y_1$ is the number of pairs $(x, y)$ such that $xu \in G' \setminus H_1$, $yv \in G' \setminus H_1$ and $xy \in G' \cup H_2$, and $Y_2$ is the number of vertices $x$ such that $xu \in G' \setminus H_1$ and $xv \in G' \setminus H_1$. To obtain upper and lower bounds for $\mathbb{E}b(G')$ it is sufficient to estimate $\mathbb{E}Y_1$ and $\mathbb{E}Y_2$. 

![Figure 2: Forward switching](image-url)
Claim 13.

\[ \mathbb{E}Y_1 = \frac{d_u d_v M_2}{M^3} + \sum_{(x, y) \in \mathcal{W}} \frac{d_x d_y d_u d_v}{M^2} + O\left(\frac{d_u d_v \Delta^5 n}{M^3}\right) \]

\[ \mathbb{E}Y_2 = \frac{\tilde{d}_u \tilde{d}_v \tilde{M}_2}{M^2} + O\left(\frac{\tilde{d}_u \tilde{d}_v \Delta^4}{M^3}\right). \]

By the claim and noting that \( \Delta^3/M^2 = O(\Delta^5 n/M^3) \) we have

\[ \mathbb{E}b(G') = \frac{\tilde{d}_u \tilde{d}_v}{\tilde{M}^2} \left(1 - \frac{M_2}{\tilde{M}^2} - \frac{\tilde{M}_2}{\tilde{M}^2} - \sum_{(x, y) \in \mathcal{W}} \frac{\tilde{d}_x \tilde{d}_y}{M^2}\right) \left(1 + O\left(\frac{\Delta^5 n}{M^3}\right)\right). \]

Note that \( \sum_{G \in \mathcal{G}^+} f(G) = \sum_{G \in \mathcal{G}^-} b(G') \) and thus \( |\mathcal{G}^+| \cdot \mathbb{E}f(G) = |\mathcal{G}^-| \cdot \mathbb{E}b(G') \). Therefore, noting that \( \Delta^4/M^2 = O(\Delta^5 n/M^3) \) we have

\[ \mathbb{P}(uv \in \mathcal{G}(n, d) \mid H_1, H_2) = \frac{|\mathcal{G}^+|}{|\mathcal{G}^+| + |\mathcal{G}^-|} = \frac{\mathbb{E}b}{\mathbb{E}f + \mathbb{E}b} \]

\[ = \frac{\tilde{d}_u \tilde{d}_v}{\tilde{M} - 2(\tilde{d}_u + \tilde{d}_v) + 2 - \sum_{x \in \mathcal{N}_{H_1 \cup H_2}(u)} \tilde{d}_x - \sum_{y \in \mathcal{N}_{H_1 \cup H_2}(v)} \tilde{d}_y - \left(\frac{(\tilde{d}_u + \tilde{d}_v - 2)M_2}{M}\right)} \left(1 + O\left(\frac{\Delta^5 n}{M^3}\right)\right). \]

Now we prove Claims 12 and 13.

**Proof of Claim 12.** First we estimate \( \mathbb{E}X_u \). Let \( X_{u, 1} \) be the number of \((x, y)\) such that \( ux \in H_1 \cup H_2 \), \( xy \in G \setminus H_1 \) and \( y \notin \{u, v\} \), \( X_{u, 2} \) be the number of \((x, y)\) such that \( ux \in G \setminus H_1 \) and \( xy \in G \setminus H_1 \) and \( y \neq u \), and \( X_{u, 3} \) be the number of \((x, y)\) such that \( ux \in G \setminus H_1 \) and \( xy \in G \setminus H_1 \) and \( y = v \). Then \( X_u = X_{u, 1} + X_{u, 2} - X_{u, 3} \). Obviously,

\[ X_{u, 1} = \sum_{x \in \mathcal{N}_{H_1 \cup H_2}(u)} \tilde{d}_x - X', \]

where \( X' \) is the number of \((x, y)\) such that \( ux \in H_1 \cup H_2 \), \( xy \in G \setminus H_1 \) and \( y = v \). Then, by Theorem 1,

\[ \mathbb{E}X' = \sum_{x \in \mathcal{N}_{H_1 \cup H_2}(u)} \mathbb{P}(xy \mid (H_1 \cup \{uv\})^+, H_2^-) = O(\Delta_{H_1 \cup H_2} \Delta^2/M) = O(\Delta^3/M). \]
Thus, $E_{x,u} = \sum_{x \in \mathcal{N}_{H_1 \cup H_2}(u)} \tilde{d}_x + O(\Delta^3/M)$. On the other hand,

$$E_{X,u_2} = \sum_{x \in [n] \setminus \{(u,v)\} \cup \mathcal{N}_{H_1 \cup H_2}(u)} \mathbb{P}(ux \mid (H_1 \cup \{uv\})^+, H_2^-) \cdot (\tilde{d}_x - 1).$$

Again by Theorem 1,

$$E_{X,u_2} = \sum_{x \in [n] \setminus \{(u,v)\} \cup \mathcal{N}_{H_1 \cup H_2}(u)} \frac{M}{\tilde{d}_x} (1 + O(\Delta^2/M)) \cdot (\tilde{d}_x - 1)$$

$$= (\tilde{d}_u - 1) \sum_{x \in [n] \setminus \{(u,v)\} \cup \mathcal{N}_{H_1 \cup H_2}(u)} \frac{M}{\tilde{d}_x^2} + O(\Delta^4/M)$$

$$= \frac{\tilde{d}_u - 1}{M} \left( \tilde{M}_2 - \sum_{x \in \mathcal{N}_{H_1 \cup H_2}(u)} \frac{M}{\tilde{d}_x^2} \right) + O(\Delta^4/M)$$

$$= \frac{\tilde{d}_u - 1}{M} \tilde{M}_2 + O(\Delta^4/M).$$

Finally, by Theorem 1

$$E_{X,u_3} = \sum_{x \in [n] \setminus \{(u,v)\} \cup \mathcal{N}_{H_1 \cup H_2}(u)} \mathbb{P}(vx \mid (H_1 \cup \{uv\})^+, H_2^-)$$

$$\leq (1 + O(\Delta^2/M)) \sum_{x \in [n]} \frac{d_u d_v (d_x^2)}{M^2} = O(\Delta^3/M).$$

Combining all together, we have

$$E_{X,u} = \sum_{x \in \mathcal{N}_{H_1 \cup H_2}(u)} \tilde{d}_x + \frac{\tilde{d}_u - 1}{M} \tilde{M}_2 + O\left(\frac{\Delta^4}{M}\right),$$

as $\Delta_{H_1} \leq \Delta$ and $\Delta_{H_2} = O(\Delta)$ by assumption. By symmetry,

$$E_{X,v} = \sum_{y \in \mathcal{N}_{H_1 \cup H_2}(v)} \tilde{d}_y + \frac{\tilde{d}_v - 1}{M} \tilde{M}_2 + O\left(\frac{\Delta^4}{M}\right).$$

Finally, we bound $E_{X_{uv}}$ by $O(\Delta^4/M)$. First we expose the neighbours of $u$ and $v$ in $G$. Next, for each $x \in \mathcal{N}_{H_1 \cup H_2}(u)$ and $y \in \mathcal{N}_{H_1 \cup H_2}(v)$, if $x \neq y$ and $xy \notin H_1$, the probability that $xy \in \mathcal{G}(n, d)$, conditional on $H_1^+, H_2^-$, and the presence of the exposed edges incident with $u$ and $v$, is at most $(1 + O(\Delta^2/M))d^2/(\tilde{M} - 4\Delta) = O(\Delta^2/M)$. There are at most $O(\Delta^2)$ such pairs of $(x, y)$, as $\Delta_{H_2} = O(\Delta)$. Thus,

$$E_{X_{uv}} = O(\Delta^2) \cdot O(\Delta^2/M) = O(\Delta^4/M).$$
Proof of Claim 13. Let \( Y_{1,1} \) be the number of pairs \((x, y)\) such that \( xu \in G' \setminus H_1 \) and \( yv \in G' \setminus H_1 \) and \( xy \in H_1 \cup H_2 \) and \( Y_{1,2} \) be the number of pairs \((x, y)\) such that \( xu \in G' \setminus H_1 \) and \( yv \in G' \setminus H_1 \) and \( xy \in G' \setminus H_1 \). Then \( Y_1 = Y_{1,1} + Y_{1,2} \). We first estimate \( EY_{1,1} \). Recall that \( W \) is the set of \((x, y)\) such that \( xy \in H_1 \cup H_2 \), \( xu, yv \notin H_1 \cup H_2 \), and \( \{x, y\} \cap \{u, v\} = \emptyset \).

Now,

\[
EY_{1,1} = \sum_{(x, y) \in W} \Pr(ux, vy \in G(n, d) \mid H_1^+, (H_2 \cup \{uv\})^-).
\]

By Theorem 1,

\[
EY_{1,1} = (1 + O(\Delta^2/M)) \sum_{(x, y) \in W} \frac{\bar{d}_u \bar{d}_x \bar{d}_y \bar{d}_v}{M^3} = \sum_{(x, y) \in W} \frac{\bar{d}_u \bar{d}_x \bar{d}_y \bar{d}_v}{M^3} + O \left( \frac{\bar{d}_u \Delta^5 n}{M^3} \right), \tag{5}
\]

since \(|W| = O(\Delta n)\). Similarly,

\[
EY_{1,2} = \sum_{(x, y) \in W} \frac{\bar{d}_u (\bar{d}_x)_2 (\bar{d}_y)_2 \bar{d}_v}{M^3} \left( 1 + O \left( \frac{\Delta^2}{M} \right) \right) = \sum_{(x, y) \in W} \frac{\bar{d}_u (\bar{d}_x)_2 (\bar{d}_y)_2 \bar{d}_v}{M^3} + O \left( \frac{\bar{d}_u \Delta^5 n}{M^3} \right),
\]

where the summation is over all pairs \((x, y)\) such that \( x \in [n] \setminus (\{u, v\} \cup \mathcal{N}_{H_1 \cup H_2}(u)) \) and \( y \in [n] \setminus (\{u, v, x\} \cup \mathcal{N}_{H_1 \cup H_2}(v) \cup \mathcal{N}_{H_1 \cup H_2}(x)) \). Thus,

\[
EY_{1,2} = O \left( \frac{\bar{d}_u \Delta^5 n}{M^3} \right) + \frac{\bar{d}_u \bar{d}_v \bar{d}_x \bar{d}_y}{M^3} \left( M_2^2 - a_1 - a_2 \right),
\]

where

\[
a_1 = \bar{M}_2 \left( \sum_{x \in \{u, v\} \cup \mathcal{N}_{H_1 \cup H_2}(u)} (\bar{d}_x)_2 + \sum_{j \in \{u, v\} \cup \mathcal{N}_{H_1 \cup H_2}(v)} (\bar{d}_y)_2 \right) - \sum_{x \in \{u, v\} \cup \mathcal{N}_{H_1 \cup H_2}(u)} (\bar{d}_x)_2 \sum_{j \in \{u, v\} \cup \mathcal{N}_{H_1 \cup H_2}(v)} (\bar{d}_y)_2 = O(\Delta^5 n),
\]

and

\[
a_2 = \sum (\bar{d}_x)_2 (\bar{d}_y)_2,
\]

where the sum is over all pairs \((x, y)\) such that \( y \in \{x\} \cup \mathcal{N}_{H_1 \cup H_2}(x) \) in addition to \( x \in [n] \setminus (\{u, v\} \cup \mathcal{N}_{H_1 \cup H_2}(u)) \) and \( y \in [n] \setminus (\{u, v\} \cup \mathcal{N}_{H_1 \cup H_2}(v)) \). Trivially, \( a_2 = O(\Delta^4 |H_1 \cup H_2|) = O(\Delta^5 n) \). Thus,

\[
EY_{1,2} = \frac{\bar{d}_u \bar{d}_v \bar{d}_x \bar{d}_y}{M^3} + O \left( \frac{\bar{d}_u \Delta^5 n}{M^3} \right).
\]

It follows now that

\[
EY_1 = \frac{\bar{d}_u \bar{d}_v \bar{d}_x \bar{d}_y}{M^3} + \sum_{(x, y) \in W} \frac{\bar{d}_u \bar{d}_x \bar{d}_y}{M^2} + O \left( \frac{\bar{d}_u \Delta^5 n}{M^3} \right).
\]
With a similar but simpler argument, we have

\[ \mathbb{E}Y_2 = (1 + O(\Delta^2/M)) \sum_{x \in [n] \setminus \langle \{u, v\} \cup N_{H_1 \cup H_2}(u) \cup N_{H_1 \cup H_2}(v) \rangle} \frac{\tilde{d}_u \tilde{d}_v (\tilde{d}_x)^2}{M^2} \]

\[ = (1 + O(\Delta^2/M)) \frac{\tilde{d}_u \tilde{d}_v (M_2 - O(\Delta^3))}{M^2} + O\left(\frac{\tilde{d}_u \tilde{d}_v \Delta^3}{M^2}\right). \]

4 Variance of the number of strictly balanced subgraphs

Let \( H \) be a fixed graph with \( t \) vertices and \( h \) edges. Recall that \( Z_H \) denotes the number of subgraphs of \( G(n, d) \) that are isomorphic to \( H \), and \( \mu_H = \mathbb{E}Z_H \), \( \sigma^2_H = \text{Var}Z_H \). The goal of this section is to estimate \( \sigma^2_H \) for strictly balanced \( H \). We start by estimating \( \mu_H \).

4.1 Expectation

Given a graph \( H \), let

\[ \varphi(d, H) = \prod_{v \in V(H)} (d)_{\text{aut}(H)}. \]

Let \( J \) be a copy of \( H \) in \( K_n \), and let

\[ \lambda_H = \mathbb{P}(J \subseteq G(n, d)). \]

By Theorem 2,

\[ \lambda_H = \frac{\varphi(d, H)}{(dn)^h} \left(1 - \frac{A_H d + B_H}{dn} + O(d^2/n^2)\right), \tag{6} \]

where \( A_H \) and \( B_H \) are constants that depend only on \( H \). The values of \( A_H \) and \( B_H \) are given for \( H \) being a cycle in Example 14 below. By linearity of expectation,

\[ \mu_H = \frac{(n)_t}{\text{aut}(H)} \lambda_H, \tag{7} \]

where \( \text{aut}(H) \) denotes the size of the automorphism group of \( H \). Hence,

\[ \mu_H = \Theta(n^t(d/n)^h) = \Theta(d^h n^{t-h}). \]

**Example 14.** Let \( H = C_\ell \), a cycle of length \( \ell \geq 3 \). By Theorem 2,

\[ \lambda_{C_\ell} = \left(\frac{d-1}{n}\right)^{\ell} \left(1 - \sum_{i=1}^{\ell} \frac{\phi_i}{dn} + O(d^2/n^2)\right), \]

where

\[ \phi_1 = -d, \]
\[ \phi_2 = -2d, \]
\[ \phi_i = -2d - 2(i-3) - 1 \text{ for all } 3 \leq i \leq \ell - 1, \]
\[ \phi_\ell = -3d + 3 - 2(\ell - 3). \]
In particular,
\[
\lambda_{C_3} = \left(\frac{d-1}{n}\right)^3 \left(1 + \frac{6d-3}{dn} + O(d^2/n^2)\right). \tag{8}
\]

### 4.2 Variance of $Z_{C_3}$

In this section, we calculate $\sigma^2_{C_3}$ in detail. In the next section, we will briefly sketch how to extend the argument for $C_3$ to other strictly balanced graphs. Let $N = \binom{n}{3}$, and let $H_1, \ldots, H_N$ be an enumeration of all copies of triangles in $K_n$. Let $X_i = 1_{\{H_i \subseteq G(n,d)\}}$. Then $Z_{C_3} = \sum_{i=1}^{N} X_i$. Moreover, $\mathbb{E}X_i = \lambda$ for every $i$ where $\lambda = \lambda_{C_3}$ has been estimated in (8). Hence,

\[
\sigma^2_{C_3} = \mathbb{E}(Z_{C_3} - \mu_{C_3})^2 = \mathbb{E} \left(\sum_{i=1}^{N} (X_i - \lambda)\right)^2 = \sum_{(i,j) \in [N]^2} (\mathbb{E}X_iX_j - \lambda^2). \tag{9}
\]

We split the above sum into four parts according to how $H_i$ and $H_j$ intersect.

(a) $H_i$ and $H_j$ are vertex disjoint. For such pairs $(i, j)$, by Theorem 2 and (2),

\[
\mathbb{E}X_iX_j = \left(\frac{d-1}{n}\right)^6 \left(1 + \frac{12d+12}{dn} + O(d^2/n^2)\right).
\]

Thus,

\[
\mathbb{E}X_iX_j - \lambda^2 = \left(\frac{d-1}{n}\right)^6 \frac{18}{dn} + O(d^8/n^8).
\]

The number of vertex disjoint pairs $(H_i, H_j)$ is $\binom{n}{3}\binom{n-3}{3} = (n)_6/36$. Hence, the contribution to (9) from part (a) is

\[
\frac{(n)_6}{36} \left(\left(\frac{d-1}{n}\right)^6 \frac{18}{dn} + O(d^8/n^8)\right) = \left(\frac{d-1}{n}\right)^6 \frac{(n)_6}{2dn} + O(d^8/n^2). \tag{10}
\]

(b) $H_i$ and $H_j$ share exactly one vertex. For such pairs,

\[
\mathbb{E}X_iX_j = \frac{(d(d-1))^4(d)_4}{(dn)^6} \left(1 + \frac{16d-4}{dn} + O\left(d^2/n^2\right)\right) = \frac{(d-1)^5(d-2)(d-3)}{dn^6} \left(1 + \frac{16d-4}{dn} + O\left(d^2/n^2\right)\right).
\]

Thus,

\[
\mathbb{E}X_iX_j - \lambda^2 = -\frac{2(2d-3)(d-1)^5}{dn^6} + O(d^8/n^7 + d^8/n^8).
\]
The number of pairs \((H_i, H_j)\) sharing exactly one vertex is \(\binom{n}{3} 3(n-3) = (n)_5/4\). Thus, the contribution to (9) from part (b) is

\[
\frac{(n)_5}{4} \left( \frac{2(2d - 3)(d - 1)^5}{dn^6} + O\left(\frac{d^6}{n^7} + \frac{d^8}{n^8}\right) \right) = \frac{(n)_5(2d - 3)(d - 1)^5}{2dn^6} + O\left(\frac{d^6}{n^2} + \frac{d^8}{n^3}\right).
\]  

(11)

(c) \(H_i\) and \(H_j\) share exactly two vertices. For such pairs,

\[
\mathbb{E}X_iX_j = \frac{(d)_2^2(d)_2^2}{(dn)^5} \left( 1 + \frac{13d - 11}{dn} \right) = \frac{(d - 1)^4(d - 2)^2}{dn^5} \left( 1 + \frac{13d - 11}{dn} \right).
\]

Thus,

\[
\mathbb{E}X_iX_j - \lambda^2 = \frac{(d - 1)^4(d - 2)^2}{dn^5} + O(d^6/n^6).
\]

The number of such pairs \((H_i, H_j)\) is \((n)_3 3(n-3) = (n)_4/2\). Thus, the contribution to (9) from part (c) is

\[
\frac{(n)_4}{2} \left( \frac{(d - 1)^4(d - 2)^2}{dn^5} + O(d^6/n^6) \right) = \frac{(n)_4(d - 1)^4(d - 2)^2}{2dn^5} + O(d^6/n^7). \tag{12}
\]

(d) \(H_i = H_j\). For such pairs,

\[
\mathbb{E}X_iX_j = \mathbb{E}X_i = \lambda.
\]

Thus,

\[
\mathbb{E}X_iX_j - \lambda^2 = \lambda - \lambda^2 = \left( \frac{d - 1}{n} \right)^3 \left( 1 + \frac{6d - 3}{dn} + O(d^2/n^2) \right) + O(d^6/n^6).
\]

There are \((n)_3/6\) pairs \((H_i, H_j)\) where \(H_i = H_j\). Thus, the contribution to (9) from part (d) is

\[
\frac{(n)_3}{6} \left( \left( \frac{d - 1}{n} \right)^3 \left( 1 + \frac{6d - 3}{dn} \right) + O(d^5/n^5) \right) + O\left(\frac{d^5}{n^2}\right).
\]  

(13)

By (10)–(12), the total contribution to (9) from parts (a)–(c) is

\[
-\frac{(d - 1)^4(d - 2)(n - 1)}{2dn^4} + O(d^8/n^2).
\]
Combining with (13) and by (9) we have
\[
\sigma^2_{C_3} = \frac{(n)_3}{6} \left( \frac{d - 1}{n} \right)^3 \left( 1 + \frac{6d - 3}{dn} \right) - \frac{(d - 1)^4(d - 2)(n - 1)_3}{2dn^2} + O(d^8/n^2)
\]
\[
= (1 + o(1)) \frac{(d - 1)^3}{6} + O(d^8/n^2).
\]

Since \( d = o(n^{2/5}) \), we have
\[
d^8/n^2 = o(d^3).
\]

It follows now that
\[
\sigma^2_{C_3} \sim \mu_{C_3} \sim \frac{(d - 1)^3}{6}.
\]

### 4.3 Variance of \( Z_H \): proof of Theorem 9

Recall that \( \rho_H(j) = \sup \{ \rho : |V(H')| \geq t \left( \frac{|H'|}{|H|} + \rho \right) \forall H' \subseteq H \text{ s.t. } H' \neq H, |H'| = j \} \)

and \( \eta_H(j) \) is positive for every \( 1 \leq j \leq h - 1 \) as \( H \) is strictly balanced.

Recall \( \lambda_H \) and \( \mu_H \) from (6) and (7). Let \( N = \binom{\binom{n}{2}}{\text{aut}(H)} \), and let \( H_1, \ldots, H_N \) be an enumeration of all copies of \( H \) in \( K_n \). Let \( X_i = 1_{\{H_i \subseteq G(n,d)\}} \). Then, \( Z_H = \sum_{i=1}^N X_i \).

Moreover, \( \mathbb{E}X_i = \lambda \) for every \( i \) where \( \lambda = \lambda_H \). Hence,
\[
\sigma^2_H = \mathbb{E}(Z_H - \mu_H)^2 = \mathbb{E} \left( \sum_{i=1}^N (X_i - \lambda) \right)^2 = \sum_{(i,j) \in [N]^2} (\mathbb{E}X_iX_j - \lambda^2). \tag{14}
\]

Similarly to the case \( H = C_3 \), we split the above sum into four parts according to the intersection of \( H_i \) and \( H_j \). We will give only upper bounds for the contributions from the first three parts, and estimate asymptotically the contribution from the last part.

Under the hypotheses of the theorem, we verify that the contribution from the last part dominates. The calculations are similar to the case where \( H = C_3 \) and thus we only briefly sketch the proof.

(a) \( H_i \) and \( H_j \) are vertex disjoint. There are \( O(n^{2t}) \) such pairs. For each such pair,
\[
\mathbb{E}X_iX_j - \lambda^2 = O(1/dn + d^2/n^2) \left( \frac{d}{n} \right)^{2h}.
\]

Hence the total contribution to (14) from this part is \( \mu_H \cdot O(d^{h-1}n^{t-h-1} + d^{h+2}n^{t-h-2}) \).

(b) \( H_i \) and \( H_j \) are edge disjoint but share at least one vertex. There are \( O(n^{2t-1}) \) such pairs. For each such pair, by Theorem 2 and (3),
\[
\mathbb{E}X_iX_j = \lambda \cdot \frac{\prod_{v \in H}(d - O(1))_{d v}}{(dn)^h} \left( 1 - \frac{O(d)}{dn} + O \left( \frac{d^2}{n^2} \right) \right)
\]
\[
= \lambda^2 \left( 1 + O \left( d^{-1} + \frac{d^2}{n^2} \right) \right).
\]
Thus,

$$\mathbb{E}X_i X_j - \lambda^2 = \lambda^2 \cdot O(1/d + d^2/n^2) = \lambda \cdot O(d^{h-1}/n^h + d^{h+2}/n^{h+2}),$$

and the total contribution to (14) from this part is $\mu_H \cdot O(d^{h-1}n^{t-h-1} + d^{h+2}n^{t-h-3}).$

(c) $H_i$ and $H_j$ share $x \geq 2$ vertices and $1 \leq y \leq h - 1$ edges. There are $O(n^{2t-x})$ such pairs. For each such pair, we know $x \geq t(y/h + \rho_H(y))$ by definition of $\rho_H.$ Thus,

$$\mathbb{E}X_i X_j - \lambda^2 = O(\lambda^2 + \mathbb{E}X_i X_j) = O(\lambda^2 + (d/n)^{2h-y}).$$

The total contribution to (14) from part (c) is

$$\sum_{x,y} O(\lambda^2 n^{2t-x}) + O(n^{2t-x}(d/n)^{2h-y}) = \mu_H \sum_{x,y} O(n^{t-x}(d/n)^{h-y}) = \mu_H \sum_{x,y} O(n^{-t\rho_H(y)} (\mu_H^{1-y/h}).$$

(d) $H_i = H_j.$ The total contribution to (14) from this part is

$$\frac{(n_t)}{\text{aut}(H)} (\lambda - \lambda^2) \sim \mu_H \sim \frac{n^t}{\text{aut}(H)} \frac{\varphi(d,H)}{(dn)^h}.$$
5.1 Proof of Corollary 5

Let \( e_1, \ldots, e_{|F|} \) be an enumeration of the edges in \( F \). Let \( F_0 = \emptyset, H_0 = H, F_j = F_{j-1} \cup \{e_j\} \) and \( H_j = H \cup F_j \) for every \( 1 \leq j \leq |F| \). Fix \( 1 \leq j \leq |F| \), let \( u \) and \( v \) be the ends of \( e_j \) and let \( d = d - d_{H_j-1} \). By (1) and (3),

\[
\mathbb{P}\left( e_j \in G(n, d) \mid H_j^+ \right) = \frac{\tilde{d}_u \tilde{d}_v}{dn} \left( 1 - \frac{\phi_{H_j-1}(uv)}{dn} \right) \left( 1 + O\left( \frac{|H_j-1|}{n^2} + \frac{|H_j-1|^2}{dn^3} + \frac{d^2}{n^2} \right) \right)
\]

\[
= \frac{\tilde{d}_u \tilde{d}_v}{dn} + O\left( \frac{d|H_j-1|}{n^3} + \frac{|H_j-1|^2}{dn^3} + \frac{d^3}{n^3} \right) \times \left( 1 + \frac{d(1 + d_{u_j-1} + d_{v_j-1}) + O(|H_j-1|)}{dn} \right)
\]

\[
= \left( \frac{(d - d_{u_j-1})(d - d_{v_j-1})}{dn} \right) \left( 1 + \frac{d(1 + d_{u_j-1} + d_{v_j-1})}{dn} \right) + O\left( \frac{d^3}{n^3} + \frac{|H_j-1|}{n^2} \right),
\]

where the last step above holds since the errors \( d|H_j-1|/n^3 \) and \( |H_j-1|^2/dn^3 \) are absorbed by \( |H_j-1|/n^2 \) as \( |H_j-1| = O(dn) \). It is easy to verify using elementary calculus that

\[
\left( \frac{(d - d_{u_j-1})(d - d_{v_j-1})}{dn} \right) \left( 1 + \frac{d(1 + d_{u_j-1} + d_{v_j-1})}{dn} \right)
\]

is a decreasing function of \( d_{u_j-1} \) and \( d_{v_j-1} \), and thus the above product is at most

\[
\frac{(d - d_{u_j-1})(d - d_{v_j-1})}{dn} \left( 1 + O\left( \frac{1}{n} + \frac{d^2}{n^2} + \frac{|H_j-1|}{dn} \right) \right).
\]

It follows that

\[
\mathbb{P}\left( e_j \in G(n, d) \mid H_j^+ \right) \leq \frac{(d - d_{u_j-1})(d - d_{v_j-1})}{dn} \left( 1 + O\left( \frac{1}{n} + \frac{d^2}{n^2} + \frac{|H_j-1|}{dn} \right) \right)
\]

\[
= \frac{(d - d_{u_j-1})(d - d_{v_j-1})}{dn} \left( 1 + O\left( \frac{1}{n} + \frac{d^2}{n^2} + \frac{|H_j-1|}{dn} \right) \right).
\]

The assertion of the corollary follows by applying the above bound to

\[
\mathbb{P}(F \subseteq G(n, d) \mid H^+) = \prod_{j=1}^{|F|} \mathbb{P}\left( e_j \in G(n, d) \mid H_j^+ \right)
\]

and then comparing the resulting expression with (6).
5.2 Distribution of $Z_{C_3}$ when $d = o(\sqrt{n})$

Let $Z = Z_{C_3}$, $\lambda = \lambda_{C_3}$ and $\mu = \mu_{C_3}$. Then, $\mu = \Theta(d^2)$. Assume $d = o(\sqrt{n})$ and $d \to \infty$. We aim to show that

$$E(Z)_k \sim \mu^k,$$

for all $k = O(d/\sqrt{n})$. Then the distribution of $Z$ follows by Theorem 15 with $\mu_n = \mu$, $\sigma_n = \sqrt{\mu}$ and $s_n = 0$.

Let $N = \binom{n}{3}$ and let $H_1, \ldots, H_N$ be an enumeration of all copies of triangles in $K_n$. Let $\mathcal{F}$ be the set of ordered $k$-tuples $(j_1, \ldots, j_k) \in [N]^k$ that are pairwise distinct. Then, for any positive integer $k$,

$$E(Z)_k = \sum_{(j_1, \ldots, j_k) \in \mathcal{F}} \mathbb{P} \left( \bigcup_{i=1}^k H_{j_i} \subseteq G(n, d) \right).$$

(15)

**Definition 16.** Given a $k$-tuple $j$, we say a triple of vertices $\{x, y, z\}$ is a hole induced by $j_1, \ldots, j_k$, if there exist $j_i, j_{i+1}$ and $j_{i+2}$, such that $H_{j_i}$ contains $xy$, $H_{j_{i+1}}$ contains $yz$ and $H_{j_{i+2}}$ contains $xz$, and none of $H_{j_i}, H_{j_{i+1}}$ and $H_{j_{i+2}}$ is $xyz$.

We partition the summation into the following two parts.

(a) $\mathcal{F}_1 \subseteq \mathcal{F}$: the set of $(j_1, \ldots, j_k)$ such that $H_{j_1}, \ldots, H_{j_k}$ are pairwise edge disjoint.

(b) $\mathcal{F}_2 = \mathcal{F} \setminus \mathcal{F}_1$.

We will prove the following.

**Lemma 17.**

$$\sum_{(j_1, \ldots, j_k) \in \mathcal{F}_1} \mathbb{P} \left( \bigcup_{i=1}^k H_{j_i} \subseteq G(n, d) \right) \sim \mu^k.$$

**Lemma 18.**

$$\sum_{(j_1, \ldots, j_k) \in \mathcal{F}_2} \mathbb{P} \left( \bigcup_{i=1}^k H_{j_i} \subseteq G(n, d) \right) = o(\mu^k).$$

Now Theorem 11 follows by the above two lemmas and Theorem 15. It only remains to prove Lemmas 17 and 18.

**Proof of Lemma 17.** We say $H_{j_i}$ hits $v$ if $v$ is incident with an edge in $H_{j_i}$. Given a $k$-tuple $H_{j_1}, \ldots, H_{j_k}$, let $h(v)$ be the number of $H_j$ which hit $v$. Let $h_j$ be the number of vertices $v$ with $h(v) = j$. Then, given any $(j_1, \ldots, j_k) \in \mathcal{F}_1$, the number of vertices in $H^\oplus = \bigcup_{i=1}^k H_i$ is equal to $3k - \sum_{j \geq 2} (j - 1) h_j$ and the number of edges in $H^\oplus$ is $3k$.

Let $\mathcal{B}$ be the set of $(j_1, \ldots, j_k) \in [N]^k$ satisfying the following property.

$$h_j > 0 \text{ for some } j \geq 4; \text{ or } h_j \geq d^2/\sqrt{n} \text{ for some } j \in \{2, 3\}.$$

We further partition $\mathcal{F}_1$ into two parts: $\mathcal{F}_1 \setminus \mathcal{B}$ and $\mathcal{F}_1 \cap \mathcal{B}$.
Claim 19.
\[ |\mathcal{F}_1 \setminus \mathcal{B}| \sim \left( \frac{(n)_3}{6} \right)^k, \text{ and } |\mathcal{F}_1 \cap \mathcal{B}| = o \left( \left( \frac{(n)_3}{6} \right)^k \right). \]

Fix an arbitrary \((j_1, \ldots, j_k) \in \mathcal{F}_1 \setminus \mathcal{B}\). Let \(e_1, \ldots, e_{3k}\) be an enumeration of the edges in \(H^0 = \bigcup_{i=1}^k H_{j_i}\), where \(e_{3\ell-2}, e_{3\ell-1}, e_{3\ell}\) is the set of edges in \(H_{j_{\ell}}\). Let \(F_i = \bigcup_{j < i} e_j\). By (1) and (3), \(\phi_{F_i}(e_i) = O(i + d)\). Thus,
\[
\sum_{i=1}^k \frac{\phi_{F_i}(e_i)}{dn} = \frac{O(k^2 + dk)}{dn}.
\]

Since \(j \notin \mathcal{B}\), we have \(d_u^{H^0} = O(1)\) for every \(u\). Hence, by Theorem 2 (with \(\tilde{d}_u, \tilde{d}_v = d - O(1)\) and \(|H| = O(k)\))
\[
\mathbb{E} \left( \prod_{i=1}^k X_{H_{j_i}} \right) = \lambda^k \left( 1 + O \left( \frac{h_2 + h_3}{d} \right) \right) \left( 1 + O \left( \frac{k^2 + dk}{dn} + \frac{k^2 + d^2 k}{n^2} + \frac{k^3}{d^2 n^2} \right) \right) \sim \lambda^k,
\]
where the error \(O((h_2 + h_3)/d)\) accounts for the case that some \(H_{j_{\ell}}\) lands on a vertex which has been occupied by some \(H_{j_i}\) where \(i < \ell\) (note that for each such \(H_{j_{\ell}}\), an \(O(1/d)\) error is present as \(\tilde{d}_u, \tilde{d}_v = d - O(1)\) when Theorem 2 is applied), and the last step of asymptotic relation above holds as \(k = O(d^{3/2})\) and \(d = o(\sqrt{n})\). By Claim 19,
\[
\sum_{(j_1, \ldots, j_k) \in \mathcal{F}_1 \setminus \mathcal{B}} \mathbb{P} \left( \bigcup_{i=1}^k H_{j_i} \subseteq \mathcal{G}(n, d) \right) \sim \left( \frac{(n)_3}{6} \right)^k \lambda^k = \mu^k.
\]

On the other hand, by Corollary 5, for any \((j_1, \ldots, j_k) \in \mathcal{F}_1 \cap \mathcal{B}\), we have
\[
\mathbb{E} \left( \prod_{i=1}^k X_{H_{j_i}} \right) \leq \lambda^k \left( 1 + O \left( \frac{k}{n} + \frac{k^2}{dn} + \frac{d^2 k}{n^2} \right) \right) \sim \lambda^k,
\]
and again by Claim 19 we have
\[
\sum_{(j_1, \ldots, j_k) \in \mathcal{F}_1 \cap \mathcal{B}} \mathbb{P} \left( \bigcup_{i=1}^k H_{j_i} \subseteq \mathcal{G}(n, d) \right) = o \left( \left( \frac{(n)_3}{6} \right)^k \lambda^k = o(\mu^k). \right.
\]

Now Lemma 17 follows by (17) and (19).

Finally, we prove Claim 19. First, we prove
\[
|\mathcal{F}_1 \cap \mathcal{B}| = o \left( \left( \frac{(n)_3}{6} \right)^k \right).
\]

For each \(j \geq 1\) let \(C_j(i)\) be the set of \(k\)-tuples with \(h_j = i\). Then, \([N]^k = \bigcup_{i \geq 0} C_j(i)\) for each \(j\). Define a switching from \(C_j(i)\) to \(C_j(i-1)\) as follows. Choose \(v\) and \(i_1 < i_2 < \cdots < i_j\)
such that \( h(v) = j \) and \( H_{i_1}, \ldots, H_{i_j} \) all use \( v \). There are exactly \( i \) ways to choose them. Then, we choose \( j - 1 \) distinct vertices \( v_1 < \cdots < v_{j-1} \), such that \( h(v_\ell) = 0 \) for all \( 1 \leq \ell \leq j - 1 \). Modify \( H_{i_\ell} \) for each \( 2 \leq \ell \leq j \) so that \( v \) is replaced by \( v_{j-1} \). The resulting \( k \)-tuple is in \( C_j(i-1) \). The number of ways to perform a switching is at least \( i(n-3k)j-1 \). On the other hand, the number of inverse switchings is at most \( (k)_j \). Hence, as \( k = O(d^{3/2}) \),

\[
\frac{|C_j(i)|}{|C_j(i-1)|} \leq \frac{(k)_j}{i(n-3k)j-1} \leq \frac{k^j}{i(n-3k)^{j-1}} = O \left( \frac{d^{3j/3}}{i n^{j-1}} \right).
\]

As \( d = o(n^{1/2}) \) it follows that

\[
|\bigcup_{j \geq 4} \bigcup_{i \geq 1} C_j(i)| \leq \sum_{j \geq 4} \sum_{i \geq 1} |C_j(i)| = O(d^6/in^3)\binom{n}{k} = o(\binom{n}{k}^3).
\]

For \( j = 2 \),

\[
\frac{|C_2(i)|}{|C_2(i-1)|} \leq \frac{(k)_2}{i(n-3k)} = O \left( \frac{d^3}{in} \right).
\]

It follows that the number of \( k \)-tuples where \( h_2 \geq d^2/\sqrt{n} \) is \( o(\binom{n}{k}) \) as \( d = o(\sqrt{n}) \). Similarly, the number of \( k \)-tuples where \( h_3 \geq d^2/\sqrt{n} \) is \( o(\binom{n}{k}) \). This confirms (20).

Next, we prove \( |F_1| \sim |\binom{n}{k}| \) which together with (20) confirms Claim 19. The total number of \( k \)-tuples in \( \binom{n}{k} \) is \( \left( \frac{n^3}{6} \right)^k \). It is thus sufficient to show that

\[
|\binom{n}{k} \setminus F_1| = o \left( \left( \frac{(n^3)}{6} \right)^k \right). \tag{21}
\]

Let \( F'_1 \subseteq \binom{n}{k} \) be the set of \((j_1, \ldots, j_k)\) where there exists \( j_u < j_v \) such that \( H_{j_u} \) and \( H_{j_v} \) share at least an edge. Then, \( F'_1 = \binom{n}{k} \setminus F_1 \). Obviously,

\[
|F'_1| \leq k^2 n^4 \left( \frac{(n^3)}{6} \right)^{k-2} = \left( \frac{(n^3)}{6} \right)^k \cdot O \left( \frac{k^2}{n^2} \right) = o \left( \left( \frac{(n^3)}{6} \right)^k \right),
\]

as \( k^2 = O(d^3) = o(n^2) \). This completes the proof for Claim 19. \( \Box \)

Proof of Lemma 18. Given \( j = (j_1, \ldots, j_k) \in F_2 \), let \( \mathcal{I}_j \) be the set of \( i \) where \( H_{j_i} \) intersects exactly \( j \) edges with \( \cup_{\ell \leq 3} H_{j_\ell} \), for \( j \in \{1, 2, 3\} \). By the definition of \( F_2 \), \( |\mathcal{I}_1(j)| + |\mathcal{I}_2(j)| + |\mathcal{I}_3(j)| \geq 1 \) for any \( j \in F_2 \). We will consider a set of parameters for \( j \) and count elements in \( F_2 \) according to these parameters. The first three parameters are

\[
i = |\mathcal{I}_1(j)|, \quad j = |\mathcal{I}_2(j)|, \quad \ell = |\mathcal{I}_3(j)|.
\]

Let \( u, v, w \) be the three vertices occupied by the triangle \( H_{j_i} \). We say that \( z, z \in \{u, v, w\} \), creates a collision, if \( z_i \) has been occupied by some triangle \( H_{j_c}, c < i \). \( H_{j_c} \) can create at most 3 collisions. Suppose \( i \in \mathcal{I}_1(j) \) and \( u, v, w \) is an edge that has been occupied
by some triangle $H_{j_{c}}$, $c < i$, then immediately $u_{i}$ and $v_{i}$ each creates a collision. These collisions are called inherent. Suppose $i \in I_{3}(j)$ then $H_{j_{i}}$ creates 3 collisions all of which are called inherent. All other collisions are called noninherent.

Let

$$t$$
denote the number of noninherent collisions.

Finally, let

$$F_{2}(i, j, \ell, t)$$
be the set of $j$ satisfying the set of parameters;

$$F_{2}(i, j, \ell) = \cup_{t < \log n \cdot k^{2}/n} F_{2}(i, j, \ell, t).$$

Now

$$F_{2} = (\cup_{t \geq \log n \cdot k^{2}/n} \cup_{(i, j, \ell; i + j + \ell \geq 1}} F_{2}(i, j, \ell, t)) \cup (\cup_{(i, j, \ell; i + j + \ell \geq 1}} F_{2}(i, j, \ell)).$$

Fix $j \in F_{2}$. We show that

$$\mathbb{P} \left( \cup_{i=1}^{k} H_{j_{i}} \subseteq \mathcal{G}(n, d) \right) \leq \lambda^{k-i-j-\ell} \left( \frac{d}{n} \right)^{2i+j} \left( 1 + O \left( \frac{k}{n} + \frac{k^{2}}{dn} + \frac{d^{2}k}{n^{2}} \right) \right).$$

(22)

Since

$$\mathbb{P} \left( \cup_{i=1}^{k} H_{j_{i}} \subseteq \mathcal{G}(n, d) \right) = \prod_{i=1}^{k} \mathbb{P} \left( H_{j_{i}} \subseteq \mathcal{G}(n, d) | (\cup_{c<i} H_{j_{c}})^{+} \right).$$

By Corollary 5, besides the relative error $O(1/n + k/dn + d^{2}/n^{2})$, if $i \notin (I_{1}(j) \cup I_{2}(j) \cup I_{3}(j))$ then the above conditional probability is at most $\lambda$; if $i \in I_{1}(j)$ then the conditional probability is at most $(d/n)^{2}$; if $i \in I_{2}(j)$ then the conditional probability is at most $d/n$; and if $i \in I_{3}(j)$ then the conditional probability is 1. Multiplying them together yields the main term in (22). There are $O(k)$ terms in the product, and multiply the relative errors in each term gives the error as in (22). As $k = O(d^{3/2})$ and $d = o(\sqrt{n})$ the error in (22) is $o(1)$.

Next, we prove that for some absolute constant $C > 0$:

$$|F_{2}(i, j, \ell, t)| \leq \frac{k^{i+j+\ell}}{i!j!\ell!} (3kn)^{i} (6kd)^{j} X^{\ell} \left( \frac{3k}{t} \right) \left( \frac{(n)_{i}}{6} \right)^{k-i-j-\ell} \left( \frac{Ck}{n} \right)^{t},$$

(23)

$$|F_{2}(i, j, \ell)| \leq \frac{k^{i+j+\ell}}{i!j!\ell!} (3kn)^{i} (6kd)^{j} Y^{\ell} \left( \frac{(n)_{i}}{6} \right)^{k-i-j-\ell},$$

(24)

where

$$X = C(i^{3/2} + j^{3/2} + t^{3/2}), \quad Y = C(i^{3/2} + j^{3/2} + (\log n \cdot k^{2}/n)^{3/2}).$$

There are at most $k^{i+j+\ell}/i!j!\ell!$ ways to fix $I_{1}$ for $I_{1}(j)$, $I_{2}$ for $I_{2}(j)$, and $I_{3}$ for $I_{3}(j)$. For every $h \in I_{1}$, there are at most 3k ways to choose the edge in $H_{j_{h}}$ which intersect with $\cup_{c<h} H_{j_{c}}$, and then at most $n$ ways to choose the other vertex in $H_{j_{h}}$. For every $h \in I_{2}$, there are at most 3k ways to choose an edge $x$ in $\cup_{c<h} H_{j_{c}}$, and then at most 2d ways to choose another edge incident with $x$ in $\cup_{c<h} H_{j_{c}}$. The upper bound 2d follows as we may assume that every vertex is incident with at most $d$ edges of $\cup_{c=1}^{k} H_{j_{c}}$ — otherwise, $\mathbb{P}(\cup_{i=1}^{k} H_{j_{i}} \subseteq \mathcal{G}(n, d)) = 0$. We claim the following bound on the number of holes.
Claim 20. The number of holes in $F_2(i, j, \ell, t)$ is at most $X$.

Proof of Claim 20. First we prove that any graph with $x$ edges has at most $3x^{3/2}$ triangles. Let $t_1, \ldots, t_n$ be the degree sequence of the graph, and we call a vertex big if its degree is at least $\sqrt{x}$ and small otherwise. Then the number of big vertices is at most $2x/\sqrt{x} = 2\sqrt{x}$. A triangle in the graph either uses only big vertices, or occupies a small vertex. Hence, the number of triangles in the graph is at most

$$\left(\frac{2\sqrt{x}}{3}\right) + \sum_{i: t_i < \sqrt{x}} \left(\frac{t_i}{2}\right) \leq \frac{4}{3} x^{3/2} + \frac{2x}{\sqrt{x}} \cdot \frac{2}{2} \leq 3x^{3/2},$$

where the maximum of the sum above appears when all non-zero degrees of small vertices are equal to $\sqrt{x}$.

Now, we bound the holes given parameters $i$, $j$, and $t$. Let $H = \bigcup_{h=1}^{k} H_{j_h}$ and we bound the largest possible number of holes that $H$ may contain. If $h \in I_1 \cup I_2$ then colour all three edges in $H_{j_h}$ red. Then the number of red edges is at most $3(i + j)$. Let $T_1$ be the number of holes where all their three edges are red, and let $T_2$ be the number of other holes. Immediately, $T_1 \leq 3(i + 2j)^{3/2} \leq C(i^{3/2} + j^{3/2})$ and $T_2 \leq C^{3/2}$. The claim follows.

First consider $F_2(i, j, \ell, t)$ for $t \geq \log n \cdot k^2/n$. By the above claim, there are at most $X^\ell$ ways to fix all $j_h$ for $h \in I_3$. Finally, there are at most $((n)_3/6)^{k-i-j-\ell}$ ways to fix $(j_i)_{i \in (I_1 \cup I_2 \cup I_3)}$. There are at most $3^k$ ways to choose the set of noninherent collisions, and each noninherent collision contributes an additional $Ck/n$ factor because there are at most $3k$ instead of $n$ choices to choose the vertex where a noninherent collision occurs. Multiplying all these upper bounds together yields (23).

For (24), the argument is similar except that we do not take union bound over all possible vertex collisions. Instead, we use the fact that $X \leq Y = C(i^3 + j^3 + (\log n \cdot k^2/n)^{3/2})$ when $t < \log n \cdot k^2/n$. Thus the number of ways to choose $j_h$ where $h \in I_3$ is at most $Y^\ell$ for all $j \in F_2(i, j, \ell, t)$, and (24) follows.

By (22) and (23) we have

\[
\sum_{(j_1, \ldots, j_k) \in F_2(i, j, \ell, t)} \mathbb{P}(\bigcup_{i=1}^{k} H_{j_i} \subseteq G(n, d)) \leq (1 + o(1))|F_2(i, j, \ell, t)| \cdot X^{k-i-j-\ell} \left(\frac{d}{n}\right)^{2i+j} \\
\leq (1 + o(1))\frac{1}{i!j!\ell!} \left(\frac{n}{3}\right)^{k-i-j-\ell} \cdot (3kn)^{3} (6kd)^{3} X^{\ell} \left(\frac{Ck^2}{tn}\right)^{t} \cdot \left(\frac{d}{n}\right)^{2i+j} \\
= O(\mu^k) \cdot \left(\frac{Ck^2 \mu}{j\mu}\right)^{i} \left(\frac{Ck^2 \mu}{j\mu}\right)^{j} \left(\frac{CkX}{\ell\mu}\right)^{\ell} \left(\frac{Ck^2}{tn}\right)^{t} \\
= O(\mu^k) \cdot \left(\frac{Ck^2 \mu}{j\mu}\right)^{i} \left(\frac{Ck^2 \mu}{j\mu}\right)^{j} \exp\left(\frac{CkX}{\mu}\right) \left(\frac{Ck^2}{tn}\right)^{t}.
\]
by defining \((x/0)^0 = 1\), and noting that \((x/\ell)^t \leq e^x\). It follows then that
\[
\sum_{(j_1, \ldots, j_k) \in \mathcal{F}_2(i,j,\ell)} \mathbb{P}\left( \bigcup_{i=1}^k H_{j_i} \subseteq \mathcal{G}(n, d) \right)
= \frac{d^{2i+j}}{\mu} \cdot \left( \log n \right)^{2i+j} (6kn)^j (C(i^{3/2} + j^{3/2} + (\log n \cdot k^2/n)^{3/2}))^\ell
\]
\[
= \frac{d^{2i+j}}{\mu} \cdot \left( \log n \right)^{2i+j} (6kn)^j (C(i^{3/2} + j^{3/2} + (\log n \cdot k^2/n)^{3/2}))^\ell
\]

Since \(\mu = \Theta(d^3)\), \(k = \Theta(d^{3/2})\), \(d = o(n^{1/2})\), and \(t \geq \log n \cdot k^2/n\), it follows immediately that
\[
\sum_{i \geq \log n \cdot k^2/n} \sum_{(j_1, \ldots, j_k) \in \mathcal{F}_2(i,j,\ell)} \mathbb{P}\left( \bigcup_{i=1}^k H_{j_i} \subseteq \mathcal{G}(n, d) \right) = o(d^k).
\]

Now combining (22) and (24) we have
\[
\mathbb{P}\left( \bigcup_{i=1}^k H_{j_i} \subseteq \mathcal{G}(n, d) \right) \leq \left( 1 + o(1) \right) |\mathcal{F}_2(i,j,\ell)| \cdot \chi^{k-i-j-\ell} \left( \frac{d}{n} \right)^{2i+j}
\]
\[
\leq \frac{k^i+j+i}{i!j!} \left( \frac{d^2}{i!j!} \right)^{2i+j} \left( \log n \right)^{2i+j} \left( 6kn \right)^j \left( C(i^{3/2} + j^{3/2} + (\log n \cdot k^2/n)^{3/2}) \right)^\ell
\]
\[
= O(d^k) \cdot \left( \log n \right)^{2i+j} (6kn)^j (C(i^{3/2} + j^{3/2} + (\log n \cdot k^2/n)^{3/2}))^\ell
\]

Since \(\mu = \Theta(d^3)\), \(k = \Theta(d^{3/2})\), \(d = o(n^{1/2})\), summing the above over all \((i, j, \ell)\) where \(i + j + \ell \geq 1\) yields \(o(d^k)\).

5.3 Distribution of \(Z_{C_{\delta}}\) for \(d = \Theta(\sqrt{n})\)

Again let \(N = \binom{n}{3}\) and let \(H_1, \ldots, H_N\) be an enumeration of all copies of triangles in \(K_n\). Let \(\mathcal{F}\) be the set of ordered \(k\)-tuples \((j_1, \ldots, j_k) \in [N]^k\) that are pairwise distinct. We want to estimate the following sum for \(k = O(d^{3/2})\):

\[
\mathbb{E}(Z)_k = \sum_{(j_1, \ldots, j_k) \in \mathcal{F}} \mathbb{P}\left( \bigcup_{i=1}^k H_{j_i} \subseteq \mathcal{G}(n, d) \right).
\]

We partition the summation into parts \(\mathcal{F}_1\) and \(\mathcal{F}_2 = \mathcal{F} \setminus \mathcal{F}_1\), where \(\mathcal{F}_1 \subseteq \mathcal{F}\) is the set of \(j\) satisfying one the following:

\[
\mathcal{E}(Z)_k = \sum_{(j_1, \ldots, j_k) \in \mathcal{F}_1} \mathbb{P}\left( \bigcup_{i=1}^k H_{j_i} \subseteq \mathcal{G}(n, d) \right).
\]

\[
\mathcal{E}(Z)_k = \sum_{(j_1, \ldots, j_k) \in \mathcal{F}_2} \mathbb{P}\left( \bigcup_{i=1}^k H_{j_i} \subseteq \mathcal{G}(n, d) \right).
\]
(a) \(|\mathcal{I}_3(j)| \geq 1\), or \(h(v) > \log^2 n\) for some \(v\).

(b) there exists \(H_{j_i}\) which intersects at least two edges with \(\cup_{\ell < i} H_{j_\ell}\).

(c) there are more than \(\log n\) \(i\)'s such that \(H_{j_i}\) intersects exactly one edge with \(\cup_{\ell < i} H_{j_\ell}\).

(d) there exists an edge contained in at least three triangles.

(e) there is some \(H_{j_i}\) such that at least two of its edges are each contained in some other triangles.

We will prove the following.

**Lemma 21.**

\[
\sum_{j \in \mathcal{F}_1} \mathbb{P}\left( \bigcup_{i=1}^k H_{j_i} \subseteq \mathcal{G}(n, d) \right) = o(\mu^k).
\]

**Lemma 22.**

\[
\sum_{j \in \mathcal{F}_2} \mathbb{P}\left( \bigcup_{i=1}^k H_{j_i} \subseteq \mathcal{G}(n, d) \right) \sim \mu^k.
\]

**Proof of Lemma 22.** The structure of the union of the \(k\) triangles \(H_{j_1}, \ldots, H_{j_k}\) are simple, for \(j \in \mathcal{F}_2\). Except for at most \(\log n\) pairs of triangles, each pair sharing exactly one edge, all the other triangles are edge disjoint. Moreover, every vertex is contained in at most \(\log^2 n\) triangles. However, as we will show, there are \(\Omega(d)\) vertices \(v\) with \(h(v) \geq 2\). This means that the joint probability of a set of \(k\) edge disjoint triangles will not be asymptotic to \(\lambda^k\) any more, which was the case in the proof of Lemma 17.

We will partition \(\mathcal{F}_2\) according to two parameters. An edge is called a double edge, if it is contained in two triangles. For \(j \in \mathcal{F}_2\), the number of double edges is exactly \(|I_1(j)|\).

Recall that \(h_i\) denote the number of vertices \(v\) with \(h(v) = i\). Modify \(\mathcal{B}\) to state the set of \(j\) where

\[|h_2 - 9k^2/2n| > d^{2/3}, \text{ or } h_3 + h_4 > d \cdot d^{7/2} \log n / n^2, \text{ or } h_j > 0 \text{ for some } j \geq 5.\]

Define

\[
\mathcal{F}_2(i) = \{j \in \mathcal{F}_2 : |I_1(j)| = i\}, \quad \mathcal{F}_2^0(i) = \mathcal{F}_2(i) \setminus \mathcal{B}, \quad \mathcal{F}_2^1(i) = \mathcal{F}_2(i) \cap \mathcal{B}
\]

\[
\mathcal{F}_2^0 = \bigcup_{i=0}^{\log n} \mathcal{F}_2^0(i), \quad \mathcal{F}_2^1 = \bigcup_{i=0}^{\log n} \mathcal{F}_2^1(i) = \mathcal{F}_2 \setminus \mathcal{F}_2^0.
\]

Then for every \(i \leq \log n\),

\[
|\mathcal{F}_2(i)| = \binom{k}{2i} \frac{(2i)!}{2^i i!} \left(\frac{(n)_3}{6} - O(\mu n)\right)^{k-2i} \left(\frac{(n - O(i))}{2}\right)^i
\]

\[
= \frac{k^{2i}}{4^i i!} \left(\frac{(n)_3}{6}\right)^{k-2i} n^{4i} \exp\left(O\left(\frac{k^2}{n^2} + \frac{i^2}{k}\right)\right)
\]

\[
\sim \frac{k^{2i} n^{4i}}{4^i i!} \left(\frac{(n)_3}{6}\right)^{k-2i}.
\]
Claim 23.

\[ |\mathcal{F}_2(i)| \sim |\mathcal{F}_2(i)| \text{ for all } i \leq \log n, \text{ and } \sum_{j \in \mathcal{F}_2^0} \mathbb{P}(\bigcup_{i=1}^k H_{ji} \subseteq \mathcal{G}(n, d)) = o(\mu^k). \]

Next, fix \( j \in \mathcal{F}_2^0(i) \). We show that

\[ \mathbb{P}(\bigcup_{i=1}^k H_{ji} \subseteq \mathcal{G}(n, d)) \sim \lambda^{k-2i} \left( \frac{d}{n} \right)^{5i} \exp \left( -\frac{9k^2}{dn} \right). \]

The proof is almost the same as for (16). The accumulative error, except for that from \( \phi_F(uv)/dn \), is \( O((d^2k + k^2)/n^2 + k^3/d^2n^2) = o(1) \). The only two differences are listed below:

(a) The term \(-2|F|\) in \( \phi_F(uv) \) in (3) is no longer negligible. Its accumulative contribution after taking the product of the conditional probabilities is

\[ \prod_{i=1}^{3k-O(i)} \left( 1 + \frac{2i}{dn} \right) = \exp \left( \frac{9k^2}{dn} + O \left( \frac{ik}{dn} + \frac{k^3}{d^2n^2} \right) \right) \sim \exp \left( -\frac{9k^2}{dn} \right). \]

(b) For each \( v \) where \( h(v) = 2 \), the joint probability of the two triangles hitting \( v \) cannot be approximated by \( \lambda^2 \) with negligible error because \( h_2 = \Theta(d) \). Their accumulative contribution after taking the product of the conditional probabilities becomes

\[ \left( \frac{(d-2)(d-3)}{d(d-1)} \right)^{h_2} = \exp \left( \left( -\frac{4}{d} + O(1/d^2) \right) \left( \frac{9k^2}{2n} + O(d^2/3) \right) \right) \sim \exp \left( -\frac{18k^2}{dn} \right). \]

Multiplying the additional factors in (a) and (b) yields (29). Finally,

\[ \sum_{j \in \mathcal{F}_2^0} \mathbb{P}(\bigcup_{i=1}^k H_{ji} \subseteq \mathcal{G}(n, d)) = \sum_{i=0}^{\log n} \sum_{j \in \mathcal{F}_2^0(i)} \mathbb{P}(\bigcup_{i=1}^k H_{ji} \subseteq \mathcal{G}(n, d)) \]

\[ \sim \sum_{i=0}^{\log n} \binom{n}{3}^{k-2i} \frac{k^2n^{4i}}{4i!} \left( \frac{d}{n} \right)^{5i} \exp \left( -\frac{9k^2}{dn} \right) \]

\[ \sim \mu^k \exp \left( -\frac{9k^2}{dn} \right) \sum_{i=0}^{\log n} \frac{62i k^2n^{4i}}{n^{6i} 4i!} \left( \frac{n}{d} \right)^{i} \]

\[ \sim \mu^k \exp \left( -\frac{9k^2}{dn} \right) \exp \left( \frac{9k^2}{dn} \right) \]

\[ = \mu^k, \] (30)

as \( k = O(d^{3/2}) \), \( d = \Theta(\sqrt{n}) \) and \( \mu = \Theta(d^3) \).
It only remains to prove Claim 23.

**Proof of Claim 23.** The proof is basically the same as the proof for Claim 19, except that in the switching we avoid switch away the double edges — there are at most \( \log n \) of them, and we estimate the number of switchings more carefully for \( h_2 \). The argument for bounds on \( h_j, j \geq 3 \) is much simpler and we skip the details. Now consider \( h_2 \), and assume that \( h_2 + h_4 = o(d) \) and \( h_j = 0 \) for all \( j \geq 5 \). Let \( \mathcal{C}(i, j) \) be the set of \( j \in F_2 \) with \( i \) double edges with \( h_2 = j \). Let \( \mathcal{C}(i) = \cup_{j \geq 4} \mathcal{C}(i, j) \). Define a switching from \( \mathcal{C}(i, j) \) to \( \mathcal{C}(i, j - 1) \) as follows. Choose \( v \) and \( i_1 < i_2 \) such that \( h(v) = 2 \), \( v \) is not incident with a double edge, and \( H_{i_1} \) and \( H_{i_2} \) both hit \( v \). There are exactly \( j - O(i) \) ways to choose them. Then, we choose a vertex \( u \) such that \( h(u) = 0 \). Modify \( H_{i_2} \) so that \( v \) is replaced by \( u \). The resulting \( k \)-tuple is in \( \mathcal{C}(i, j - 1) \). The number of ways to perform a switching is \((j - O(i))(n - 3k + O(\sum_j h_j))\). On the other hand, the number of inverse switchings is \(9(\binom{k - 2(j - O(i)) - \sum_{j \geq 3} h_j}{2}) \). Hence,

\[
\frac{|\mathcal{C}(i, j)|}{|\mathcal{C}(i, j - 1)|} = \frac{9(\binom{k - 2(j - O(i)) - \sum_{j \geq 3} h_j}{2})}{2j(n - 3k + O(\sum_j h_j))} = \frac{9(k - 2j)^2}{2j(n - 3k)} \left(1 + O\left(\frac{\log n}{\sqrt{n}}\right)\right).
\]

Let \( j^* \) be the root of \(9(k - 2j)^2 = 2(n - 3k)\). Then,

\[
j^* = \frac{5k}{12} + \frac{n}{36} - \sqrt{\frac{n^2 + 30kn - 99k^2}{36}} = \frac{9k^2}{2n} + O\left(\sqrt{d^3/n}\right).
\]

Now using (31) and following standard calculations we have

\[
\sum_{|j - j^*| > d^{2/3}} |\mathcal{C}(i, j)| = o(|\mathcal{C}(i)|)
\]

and the first assertion of the claim follows. The second assertions follows by the same calculations as for (29) and (30), except that we may use the bound in Corollary 5 which is of simpler form. For each \( j \in F_2 \cap B\),

\[
P\left(\bigcup_{i=1}^{k} H_{i,j} \subseteq G(n, d)\right) \leq \lambda^{k-2i}\left(\frac{d}{n}\right)^{5i}\left(1 + O\left(\frac{k}{n} + \frac{k^2}{dn} + \frac{d^2k}{n^2}\right)\right) = O\left(\lambda^{k-2i}\left(\frac{d}{n}\right)^{5i}\right).
\]

Since \( |F_2(i)| = o(|F_2(i)|) \) for all \( i \leq \log n \),

\[
\sum_{j \in F_2} P\left(\bigcup_{i=1}^{k} H_{i,j} \subseteq G(n, d)\right) = \sum_{i=0}^{\log n} o(|F_2(i)|) \cdot O\left(\lambda^{k-2i}\left(\frac{d}{n}\right)^{5i}\right) = o(1) \sum_{i=0}^{\log n} \frac{k^2i^4n^{4i}}{4^i i!} \left(\frac{(n)^3}{6}\right)^{k-2i} \lambda^{k-2i}\left(\frac{d}{n}\right)^{5i},
\]

\text{THE ELECTRONIC JOURNAL OF COMBINATORICS 31(1) (2024), #P1.2}
which is \( o(\mu^k) \) by following the same calculations as in (30).

**Proof of Lemma 21.** Let \( I_1(j) \) be as defined before. Let \( F_1^j \) be the set of \( j \in F_1 \) where \( j \) induces a hole. With almost the same proof as in Lemma 18, we consider \( t \): the number of noninherent collisions. For \( t \geq \log n \cdot k^2/n \), we get (25), and thus (26). Note that the change of the range of \( d \) from \( o(\sqrt{n}) \) to \( O(\sqrt{n}) \) does not affect the analysis, as \( t \geq \log n \cdot k^2/n \) and thus the term \((CK^2/tn)^j\) dominates the rate the whole product vanishes. For \( t < \log n \cdot k^2/n \), we get (27). Summing over all \((i,j,\ell)\) where \( \ell \geq 1 \) yields \( o(\mu^k) \).

Next, let \( F_2^j \subseteq F_1 \setminus F_1^j \) be the set of \( j \) where \( i + j > \log n \) where \( i = |I_1(j)| \) and \( j = |I_2(j)| \). As \( j \notin F_1 \) we know \( I_3(j) \) must be empty. Let

\[
F_2^j(i,j) = \{ j : I_1(j) = i, I_2(j) = j \}.
\]

We use the bound (24) (with \( \ell = 0 \)) for \( |F_1^2(i,j)| \):

\[
|F_2^1(i,j)| \leq \frac{k+i+j}{i!j!} \left( \frac{n}{3} \right)^{k-i-j} (3kn)^j (6kd)^j.
\]

Thus,

\[
\sum_{i,j: i+j > \log n} \sum_{j \in F_2^1(i,j)} \Pr \left( \bigcup_{i=1}^k H_{ji} \subseteq \mathcal{G}(n,d) \right) \leq O(1) \cdot \mu^k \sum_{i,j: i+j > \log n} \frac{(3k^2d^2/\mu n)^j}{i!} \left( \frac{6kd^2/\mu n}{j} \right) = o(\mu^k),
\]

as \( k = O(d^{3/2}) \) and thus \( 3k^2d^2/\mu n = O(1) \).

Next, let \( F_2^j \subseteq F_1 \setminus (F_1^1 \cup F_2^1) \) be the set of \( j \) where \( h(v) > \log^2 n \). The summation over \( F_2^j \) can be treated similarly as in \( F_1^1 \). Fix \( v \in [n] \). Since \( j \) induces no holes, \( I_3(j) = 0 \). Let \( L(j) \) be the set of \( i \) such that \( H_{ji} \) hits \( v \). Let

\[
F_2^j(i,j,r) = \{ j \in F_1^1 : I_1(j) = i, I_2(j) = j, L(j) = r \}.
\]

Since \( j \notin F_2^1 \), \( i \leq \log n \) and \( j \leq \log n \). Let \( r > \log^2 n \). There are at most \( k^{i+j+r}/i!j!r! \) ways to fix \( I_1, I_2 \) and \( L \) for \( I_1(j), I_2(j) \) and \( L(j) \) respectively. Given \( I_1, I_2 \) and \( L \), let \( L' = L \setminus (I_1 \cup I_2) \) and we must have \( |L'| \geq r - 2 \log n = r(1 - o(1)) \) as \( r > \log^2 n \). Thus,

\[
|F(i,j,r)| \leq \frac{k^{i+j+r}}{i!j!r!} \left( \frac{n}{3} \right)^{k-i-j} (3kn)^j (6kd)^j n^{2r},
\]

as for each \( r \in L' \), there are at most \( n^2 \) ways to choose the two vertices other than \( v \). Hence,

\[
\sum_{i,j,r: r > \log^2 n} \sum_{j \in F_2^1(i,j,\ell)} \Pr \left( \bigcup_{i=1}^k H_{ji} \subseteq \mathcal{G}(n,d) \right) = O(1) \cdot \sum_{i,j,r: r > \log^2 n} \frac{k^{i+j+r}}{i!j!r!} \left( \frac{n}{3} \right)^{k-i-j-r(1-o(1))} (3kn)^j (6kd)^j n^{2r} k^{i+j} \left( \frac{d}{n} \right)^{2i+j} = O(\mu^k)
\]

\[
= O(\mu^k) \sum_{r > \log^2 n} \frac{(k/n^{1-o(1)})^r}{r!} \sum_{i \geq 0} \frac{(3k^2d^2/\mu n)^i}{i!} \sum_{j \geq 0} \frac{(6kd^2/\mu n)^j}{j!} = o(\mu^k),
\]
as $k^2d^2/\mu n = O(1)$ and $k/n^{1-o(1)} = o(1)$.

Next, let $\mathcal{F}_1^1 \subset \mathcal{F}_1 \setminus (\mathcal{F}_1^1 \cup \mathcal{F}_1^2 \cup \mathcal{F}_1^3)$ be the set of $j$ where $\mathcal{I}_2(j)$ is nonempty. Let

$$\mathcal{F}_1^1(i, j) = \{ j \subseteq \mathcal{F}_1^1 : \mathcal{I}_1(j) = i, \mathcal{I}_2(j) = j \}.$$

As now $j \notin \mathcal{F}_1^1$, $h(v) \leq \log^2 n$ for every $n$ and thus we have the following better upper bound than (23) for $|\mathcal{F}_1^1(i, j)|$:

$$|\mathcal{F}_1^1(i, j)| \leq \left( \frac{(n/3)^i}{6} \right)^{k-i-j} (3kn)^i(12k^2d^2/\mu n)^j.$$

This is because every vertex in $\bigcup_{i=1}^{k} H_j$ is incident with at most $2\log^2 n$ instead of $d$ edges.

Using this improved upper bound, we have

$$\sum_{j \in \mathcal{F}_1^1} \mathbb{P}(\bigcup_{i=1}^{k} H_j \subseteq \mathcal{G}(n, d)) = O(\mu^k) \sum_{i \leq \log n} \frac{(3k^2d^2/\mu n)^i}{i!} \sum_{j \geq 1} \frac{(6k^2d\log^2 n/\mu n)^j}{j!} = o(\mu^k).$$

Finally, let $\mathcal{F}_1^0 = \mathcal{F}_1 \setminus (\bigcup_{i=1}^{k} \mathcal{F}_1^1)$, the set of $j$ where either an edge is contained in more than 2 triangles, or there is a triangle with two edges each contained in some other triangle. Summing over $j$ in the first case we have

$$\sum_{2 \leq i \leq \log n} \frac{k^i}{i!} \left( \frac{(n/3)^i}{6} \right)^{k-i} i \cdot \log n \cdot (3kn)^i-1 \lambda^{k-i} \left( \frac{d}{n} \right)^{2i} = O(\mu^k \cdot \log^2 n/k) = o(\mu^k).$$

Now we consider the latter case. If $H_{ja}$ has two edges each contained in some other triangles — call them $H_{ja}$ and $H_{jr}$. Obviously $\ell \neq r$ since otherwise $H_{ja}$ would be the same as $H_{ja}$, contradicting with them being distinct. There are two sub-cases: (a) $\ell < i$ and $r > i$; (b) $r > \ell > i$. In either case, there are at most $i^3$ ways to choose $u, \ell$ and $r$. Then, similar to the above, summing over $j$ in the second case yields

$$\sum_{2 \leq i \leq \log n} \frac{k^i}{i!} \left( \frac{(n/3)^i}{6} \right)^{k-i} i^3 \cdot O(n^2) \cdot (3kn)^i-2 \lambda^{k-i} \left( \frac{d}{n} \right)^{2i} = O(\mu^k \cdot \log^3 n/k) = o(\mu^k).$$

Combining the two cases we have $\sum_{j \in \mathcal{F}_1^0} \mathbb{P}(\bigcup_{i=1}^{k} H_j \subseteq \mathcal{G}(n, d)) = o(\mu^k)$, completing the proof of the lemma.

**Acknowledgements**

The author received support from NSERC grant RGPIN-04173-2019.

**References**

[1] 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.
[2] Paul Erdős and Alfréd Rényi. On the evolution of random graphs. *Publ. Math. Inst. Hung. Acad. Sci.*, 5(1):17–60, 1960.

[3] Pu Gao, Mikhail Isaev and Brendan McKay. Kim-vu’s sandwich conjecture is true for $d \geq \log^4 n$, *arXiv preprint arXiv:2011.09449*, 2020.

[4] Pu Gao. Distribution of the number of spanning regular subgraphs in random graphs. *Random Structures & Algorithms*, 43(3):338–353, 2013.

[5] Pu Gao. Distributions of sparse spanning subgraphs in random graphs. *SIAM Journal on Discrete Mathematics*, 27(1):386–401, 2013.

[6] Pu Gao, Mikhail Isaev, and Brendan D McKay. Sandwiching random regular graphs between binomial random graphs. In *Proceedings of the Fourteenth Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 690–701. SIAM, 2020.

[7] Pu Gao, Mikhail Isaev, and Brendan D McKay. Sandwiching random regular graphs between binomial random graphs. In *Proceedings of the Fourteenth Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 690–701. SIAM, 2020.

[8] Pu Gao and Yuval Ohapkin. Subgraph probability of random graphs with specified degrees and applications to chromatic number and connectivity. *arXiv preprint arXiv:2007.02216*, 2020.

[9] Pu Gao and Cristiane M Sato. A transition of limiting distributions of large matchings in random graphs. *Journal of Combinatorial Theory, Series B*, 116:57–86, 2016.

[10] Zhicheng Gao and Nicholas C Wormald. Asymptotic normality determined by high moments, and submap counts of random maps. *Probability theory and related fields*, 130(3):368–376, 2004.

[11] Zhicheng Gao and Nicholas C Wormald. Distribution of subgraphs of random regular graphs. *Random Structures & Algorithms*, 32(1):38–48, 2008.

[12] Svante Janson. Random regular graphs: asymptotic distributions and contiguity. *Combinatorics, Probability and Computing*, 4(4): 369 – 405, 1995.

[13] Svante Janson. The numbers of spanning trees, hamilton cycles and perfect matchings in a random graph. *Combinatorics, Probability and Computing*, 3(1):97–126, 1994.

[14] Brendan D McKay. Subgraphs of random graphs with specified degrees. *Congressus Numerantium*, 33: 213–223.1981.

[15] Brendan D McKay. Subgraphs of random graphs with specified degrees. In *Proceedings of the International Congress of Mathematicians 2010 (ICM 2010) (In 4 Volumes) Vol. I: Plenary Lectures and Ceremonies Vols. II–IV: Invited Lectures*, pages 2489–2501. World Scientific, 2010.

[16] Brendan D McKay. Asymptotics for symmetric 0-1 matrices with prescribed row sums. *Ars Combin* 19 (1985): 15–25.

[17] Brendan D McKay, Nicholas C Wormald, and Beata Wysocka. Short cycles in random regular graphs. *the electronic journal of combinatorics*, pages R66–R66, 2004.

[18] Andrzej Ruciński. Subgraphs of random graphs: a general approach. In *North-Holland Mathematics Studies*, volume 118, pages 221–229. Elsevier, 1985.
[19] Andrzej Ruciński. When are small subgraphs of a random graph normally distributed? *Probability Theory and Related Fields*, 78(1):1–10, 1988.

[20] Nicholas C Wormald. The asymptotic distribution of short cycles in random regular graphs. *Journal of Combinatorial Theory, Series B*, 31(2):168–182, 1981.

[21] Nicholas C Wormald. Models of random regular graphs. *London Mathematical Society Lecture Note Series*: 239–298, 1999.