Distribution of certain sparse spanning subgraphs in random graphs

Pu Gao[*]
Max-Planck-Institut für Informatik
janegao@mpi-inf.mpg.de

Abstract

We describe a general approach of determining the distribution of spanning subgraphs in the random graph $\mathcal{G}(n, p)$. In particular, we determine the distribution of spanning subgraphs of certain given degree sequences, which is a generalisation of the d-factors, of spanning triangle-free subgraphs, of (directed) Hamilton cycles and of spanning subgraphs that are isomorphic to a collection of vertex disjoint (directed) triangles.

1 Introduction

The distributions of subgraphs with fixed sizes in various random graph models have been investigated by many authors. A general approach by Ruciński [6, 7] showed that the numbers of subgraphs with fixed sizes in the binomial model $\mathcal{G}(n, p)$ are asymptotically normal for a large range of p. On the other hand, studies of distributions of subgraphs of sizes growing with n, for example, the spanning subgraphs, are much less common. The first breakthrough is perhaps due to Robinson and Wormald [8, 9] on proving that random regular graphs are a.a.s. Hamiltonian. Based on their work, Janson [3] deduced the limiting distribution of the number of Hamilton cycles in random regular graphs. The distributions of some types of spanning subgraphs (perfect matchings, Hamilton cycles, spanning trees) in random graphs $\mathcal{G}(n, p)$ and $\mathcal{G}(n, m)$ were determined by Janson [4]. These distributions behave significantly differently in $\mathcal{G}(n, m)$ and $\mathcal{G}(n, p)$. It was shown that within a big range of m, the numbers of these spanning subgraphs are asymptotically normally distributed in $\mathcal{G}(n, m)$, whereas in the corresponding $\mathcal{G}(n, p)$ with $p = m/\binom{n}{2}$, these random variables are asymptotically log-normally distributed. This is because the expectations of these variables in $\mathcal{G}(n, m)$ grow very fast as m grows. Therefore, even though the number of edges in $\mathcal{G}(n, p)$ has small deviation, the deviation of these random variables (e.g. the number of perfect matchings) can eventually be very large. This same phenomena was observed by the author [2] while studying the distribution of the number of d-factors in $\mathcal{G}(n, p)$.

In this paper, we extend and generalise the method in [2] and study additional types of spanning subgraphs. In Section 2, we describe the general method (Theorems 1 and 3) and give conditions under which the distribution of the random variable under investigation will follow a pattern

[*]Research supported by the Humboldt Foundation

of concentration in $\mathcal{G}(n,m)$ and log-normal distribution in $\mathcal{G}(n,p)$, which we call the *log-normal paradigm* in this paper. The method is also extended to cope with probability spaces of random directed graphs (See Theorem 4). In Section 3, we study the distribution of certain types of spanning subgraphs (including spanning subgraphs with certain degree sequences, triangle-free spanning subgraphs, undirected and directed Hamilton cycles, and spanning subgraphs isomorphic to a collection of vertex disjoint triangles). Their distributions are determined by verifying the conditions given in the theorems in Section 2. Note also that the method used by Janson in [4] is graph decomposition and projection whereas the approach used in [2] and in this paper proceeds via combinatorial counting, making extensive use of the switching method developed by McKay [5]. The proof of Theorem 1 is implicit in the proofs of [2, Theorem 2.3 and 2.4], which we abstract and generalise to a general approach for proving concentration in $\mathcal{G}(n,m)$ and log-normal distribution in $\mathcal{G}(n,p)$. The proof of Theorem 3 is essentially the same as the proof of [4, Theorem 6] with slight adaptation and generalisation. Both proofs of Theorems 1 and 3 are presented in Section 4. The specific problem on the number of Hamilton cycles has been studied in the past by a few authors. The first investigation was done by Wright for the directed Hamilton cycles in [11] and then the undirected Hamilton cycles in [10]. Even though both proofs in [11] and [10] are based on a similar counting trick, the analysis for the undirected version is much more complicated. The proof for the directed Hamilton cycles was redone by Frieze and Suen [1], probably unaware of the existing work of Wrignt, using basically the same approach. In [4], Janson reproved the same result for both the undirected and directed versions, using the method of graph decomposition and projection. In this paper, we present a completely new proof for both the undirected and directed version, which is indeed much simpler than the previous approaches.

2 A general approach

Let $\mathscr{S}$ denote a set of vertex-labelled graphs on a set $S=[n]$ of n vertices. For two graphs H_1 and H_2 both on vertex set S, let $H_1\cap H_2$ ($H_1\cup H_2$) denote the set of edges contained in both (either of) H_1 and H_2. For any integer $j\ge 0$, let $F_j(\mathscr{S})$ denote the set of ordered pairs $(H_1,H_2)\in\mathscr{S}\times\mathscr{S}$ such that $|H_1\cap H_2|=j$. Let $f_j=f_j(\mathscr{S})=|F_j(\mathscr{S})|$ and let $r_j=f_j/f_{j-1}$ for any $j\ge 1$, as long as $f_{j-1}\ne 0$. Let $X_n=X_n(\mathscr{S})$ denote the number of members of $\mathscr{S}$ that are contained in a random graph ($\mathcal{G}(n,p)$ or $\mathcal{G}(n,m)$, defined on the same vertex set S) as (spanning) subgraphs. Here S, p and m refer to sequences $(S(n))_{n\ge1}$, $(p(n))_{n\ge1}$ and $(m(n))_{n\ge1}$. Assume every graph in $\mathscr{S}$ has the same number $h(n)=\Omega(n)$ of edges. Let $N(n)=\binom{n}{2}$. We drop n from all these notations when there is no confusion. All asymptotics in this paper refer to $n\to\infty$. For any real x and any integer $\ell\ge 0$, define the ℓ-th falling factorial $[x]_\ell$ to be $\prod_{i=0}^{\ell-1}(x-i)$. Let

$$\mu_n \;=\; |\mathscr{S}|\binom{N-h}{m-h}\Big/\binom{N}{m},\quad \lambda_n=|\mathscr{S}|p^h. \tag{2.1}$$

Clearly,

$$\mathbf{E}_{\mathcal{G}(n,m)}X_n=\mu_n,\quad \mathbf{E}_{\mathcal{G}(n,p)}X_n=\lambda_n.$$

A simplification of μ_n (readers can also refer to Lemma 19 by taking $\ell=h$) gives

$$\mathbf{E}_{\mathcal{G}(n,m)}X_n \;=\; |\mathscr{S}|\cdot\frac{\binom{N-h}{m-h}}{\binom{N}{m}}=|\mathscr{S}|\cdot\frac{[m]_h}{[N]_h}=|\mathscr{S}|(m/N)^h\exp\left(-\frac{N-m}{mN}\frac{h^2}{2}+O(h^3/m^2)\right). \tag{2.2}$$

Theorem 1 *Let μ_n be defined as in (2.1). Assume that $h^3 = o(m^2)$, and for $\rho(n) = h^2/m$ and some function $\gamma(n)$, the following conditions hold:*

(a) for all $K > 0$ and for all $1 \le j \le K\rho(n)$,

$$r_j = \frac{h^2}{Nj}(1 + o(m/h^2));$$

(b) $r_j \le m/2N$ for all $4\rho(n) \le j \le \gamma(n)$;

(c) $t(n) := \sum_{j>\gamma(n)} f_j = o(\mu_n|\mathscr{S}|)$;

(d) $\mu_n \to \infty$, as $n \to \infty$.

Then in $\mathcal{G}(n,m)$,

$$X_n/\mathbf{E}_{\mathcal{G}(n,m)}(X_n) \xrightarrow{p} 1,$$

as $n \to \infty$.

Remark: The ratio r_j in condition (a) looks quite restrictive. However, as we will see in the next section, this ratio appears naturally if the edges in $\mathscr{S}$ are distributed randomly (see examples in Sections 3.1 and 3.2). In some cases, for instance, if we take $\mathscr{S}$ to be the set of graphs isomorphic to a given unlabelled graph on n vertices, the edges in $\mathscr{S}$ are likely to still distribute in some kind of "random-like" way and thus having r_j as expressed in condition (a) is expected. If we are lucky, we might have condition (b) satisfied for $\gamma(n) = h$. See the example in Section 3.4. But usually this is not the case, as the sequence r_j might decrease first and increase at its tail. Normally, in these cases, condition (c) is not difficult to check. See examples in Sections 3.3, 3.5 and 3.6.

Theorem 1 and its proof also gives the following proposition.

Proposition 2 *Assume all conditions (a)–(d) of Theorem 1 are satisfied. Then, for all $j = O(h^2/m)$,*

$$f_j(n) \sim |\mathscr{S}|^2 \exp(-h^2/N)(h^2/N)^j/j!.$$

The following theorem gives conditions under which X_n will be asymptotically log-normally distributed in $\mathcal{G}(n,p)$ if all conditions in Theorem 1 are satisfied by taking $m = pN$.

Theorem 3 *Assume $h^3 = o(p^2n^4)$. Let $\beta_n = h\sqrt{(1-p)/pN}$ and λ_n as defined in (2.1). Assume further that $\liminf_{n\to\infty} \beta_n > 0$. If for all $m = pN + O(\sqrt{pN})$, $X_n/\mathbf{E}_{\mathcal{G}(n,m)}(X_n) \xrightarrow{p} 1$, then*

$$\frac{\ln(e^{\beta_n^2/2}X_n/\lambda_n)}{\beta_n} \xrightarrow{d} \mathcal{N}(0,1), \quad \textit{as } n \to \infty,$$

where $\mathcal{N}(0,1)$ is the standard normal normal distribution.

By Theorems 1 and 3, to show that a random variable has a log-normal distribution in $\mathcal{G}(n,p)$, it is enough to check conditions (a)–(d) in Theorem 1 by taking $m = pN$. This method is particular powerful if we can estimate r_j without knowing f_j. This is the case in most examples established in Section 3. However, even in the case when r_j is obtained by estimating f_j first, Theorem 1 provides a guidance of which terms of f_j are non-negligible terms in the analysis, and by verifying the conditions in Theorem 1 we can make proofs very systematic. We will give one such example in Section 3.5 (the second proof for Theorem 12).

We can generalise the results to random digraphs. Define $\mathcal{D}(n,m)$ to be the random digraph on n vertices with m directed edges chosen uniformly at random from the $2N$ ordered pairs of vertices. Define $\mathcal{D}(n,p)$ to be the random digraph on n vertices, which includes every directed edge independently with probability p. In this paper, we again define $\mathcal{D}(n,m)$ and $\mathcal{D}(n,p)$ on the vertex set S. With almost the same proofs of Theorems 1 and 3 we have the following theorem.

Theorem 4 *The same conclusions of Theorems 1 and 3 hold if we replace $\mathcal{G}(n,m)$, $\mathcal{G}(n,p)$, N by $\mathcal{D}(n,m)$, $\mathcal{D}(n,p)$ and $2N$.*

3 A few examples

3.1 A trivial example

Take $\mathscr{S}_1$ to be the set of all graphs on vertex set S with h edges. Then $|\mathscr{S}_1| = \binom{N}{h}$. The conclusion of Theorem 1 should hold trivially in this case as $X_n(\mathscr{S}_1)$ is constant (depending only on m and h). Nevertheless we verify conditions (a) and (b), also for later use in the next section. For all $0 \le j \le h$,

$$f_j = \binom{N}{j}\binom{N-j}{h-j}\binom{N-h}{h-j}.$$

Then for all $1 \le j \le h$,

$$r_j = \frac{(N-j+1)(h-j+1)^2}{j(N-h)(N-2h+j)} = \frac{h^2}{jN}(1 + O(j/h + h/n^2)).$$

This verifies conditions (a) and (b) (for $\gamma(n) = h$). With $|\mathscr{S}_1| = \binom{N}{h}$, we can easily check that condition (d) is satisfied.

3.2 Another trivial example

Let $0 < \hat{p} < 1$. Consider the set of graphs $\mathscr{S}_2$ that is obtained by including each element in $\mathscr{S}_1$ independently with probability $\hat{p}$. Then we have the following.

Theorem 5 *Assume $0 < \hat{p} \le 1$, $0 < p < 1$ are reals and $h = \Omega(n)$ is an integer that satisfy $m = p\binom{n}{2}$, $h^3 = o(m^2)$, $m^2\hat{p}^2N^h >> h^{3h+4}\ln n$. Let μ_n and λ_n be defined as in (2.1) and let $\beta_n = h\sqrt{(1-p)/pN}$. Then a.a.s. $X_n(\mathscr{S}_2)/\mu_n \xrightarrow{p} 1$ in $G(n,m)$, and*

$$\frac{\ln(e^{\beta_n^2/2}X_n(\mathscr{S}_2)/\lambda_n)}{\beta_n} \xrightarrow{d} \mathcal{N}(0,1), \quad \textit{in } \mathcal{G}(n,p),$$

provided $\liminf_{n\to\infty} \beta_n > 0$.

Proof. By the definition of $\mathscr{S}_2$, we have $s_2 = |\mathscr{S}_2| \sim \mathcal{B}(\binom{N}{h}, \hat{p})$ and $f_j \sim \mathcal{B}(M_j, \hat{p}^2)$, where $M_j = \binom{N}{j}\binom{N-j}{h-j}\binom{N-h}{h-j}$. The Chernoff bound gives that

$$\mathbf{P}(|f_j - \hat{p}^2 M_j| > 2\sqrt{3 \ln n \hat{p}^2 M_j}) < \exp(-3 \ln n) = n^{-3},$$

and

$$\mathbf{P}\left(s_2 > \binom{N}{h}\hat{p}/2\right) = 1 - o(1).$$

Therefore, with probability at least $1 - hn^{-3} - o(1) = 1 - o(1)$, for all $0 \le j \le h$,

$$f_j = \left(1 + O\left(\sqrt{\ln n/\hat{p}^2 M_j}\right)\right) M_j.$$

Note that for all j, $M_j > [n]_h/(h!)^3 > (N/h^3)^h$ and $\hat{p}$ satisfies

$$\hat{p}^2 >> \frac{h^{4+3h} \ln n}{m^2 N^h}.$$

Thus, a.a.s. for all $0 \le j \le h$,

$$f_j = \left(1 + o(m/h^2)\right) M_j.$$

By the calculations in Section 3.1, a.a.s. both conditions (a) and (b) (for $\gamma(n) = h$) are satisfied and a.a.s.

$$\mathbf{E}(X_n(\mathscr{S}_2)) \ge \mathbf{E}(X(\mathscr{S}_1))\hat{p}/2 \sim \frac{\hat{p}/2}{\sqrt{2\pi h}} \left(\frac{em}{h}\exp(-h/2m)\right)^h >> \frac{h^2\sqrt{\ln n}}{m} \left(em\sqrt{h/N}e^{-h/2m}\right)^h.$$

Since $h = \Omega(n)$ and $h^3 = o(m^2)$, the above tends to ∞ as $n \to \infty$ and so condition (d) is also satisfied. The theorem thereby follows. ∎

The following is a corollary of Theorem 5 by letting $\hat{p} = 1/2$.

Corollary 6 *Assume $0 < p < 1$ is a real and $h = \Omega(n)$ is an integer that satisfy $m = p\binom{n}{2}$, $h^3 = o(m^2)$, $m^2 N^h >> h^{3h+4} \ln n$. Then for almost all subsets $\mathscr{S}_2'$ of $\mathscr{S}_1$, the same conclusions of Theorem 5 (without "a.a.s.") hold when $\mathscr{S}_2$ is replaced by $\mathscr{S}_2'$.*

3.3 The number of spanning subgraphs with given degree sequences

In this section, we consider a non-trivial example where $\mathscr{S}$ is the set of graphs on S with a given degree sequence.

Let $\mathbf{d} = (d_1, \ldots, d_n)$ be a degree sequence and let $d_{\max} := \max\{d_i, 1 \le i \le n\}$. Let $\mathscr{S}_3$ denote the set of graphs on S with degree sequence $\mathbf{d}$. Thus, $X_n(\mathscr{S}_3)$ counts all spanning subgraphs with degree sequence $\mathbf{d}$. The sequence $\mathbf{d}$ refers to $(\mathbf{d}(n))_{n \ge 1}$. We again drop n from the notation when there is no confusion.

A special case when $\mathbf{d}$ is constant was studied by the author in [2]. The distribution of the number of d-factors in $\mathcal{G}(n, p)$ was shown to follow the log-normal paradigm. The core part of the proof in [2] is to estimate r_j using the switching method. We will generalise this proof to cope with general degree sequences $\mathbf{d}$. Let $h = \sum_{i=1}^n d_i/2$, $\bar{d}_1 = 2h/n$ and $\bar{d}_2 = \sum_{i=1}^n d_i^2/n$. Let $M_i = \bar{d}_i n$ for $i = 1, 2$.

Assume $d_{\max}^4 = o(h)$. The following estimate of $|\mathscr{S}_3|$ was first obtained by McKay [5].

$$|\mathscr{S}_3| = \frac{M_1!}{(M_1/2)!2^{M_1/2}\prod_{i=1}^n d_i!}\exp\left(-\frac{M_2-M_1}{2M_1} - \frac{(M_2-M_1)^2}{4M_1^2} + O(d_{\max}^4/h)\right). \tag{3.1}$$

The main theorem is as follows.

Theorem 7 *Let $0 < p < 1$ be a real and $0 < m < N$ an integer and $\mathbf{d}$ a degree sequence satisfying $m = pN$, $d_{\max}^3 = o(p^2 n)$, $h^3 = o(m^2)$ and $d_{\max}^4 = o(h)$. Assume further that $\bar{d}_2 = \bar{d}_1^2(1 + o(m/h^2))$. Let $X_{n,\mathbf{d}}$ denote the number of spanning subgraphs with degree sequence $\mathbf{d}$. Let $\mu_{n,\mathbf{d}}$ and $\lambda_{n,\mathbf{d}}$ be defined as μ_n and λ_n in (2.1). Let $\beta_n = h\sqrt{(1-p)/pN}$. Then $X_{n,\mathbf{d}}/\mu_{n,\mathbf{d}} \xrightarrow{p} 1$ in $G(n,m)$, and*

$$\frac{\ln(e^{\beta_n^2/2}X_{n,\mathbf{d}}/\lambda_{n,\mathbf{d}})}{\beta_n} \xrightarrow{d} \mathcal{N}(0,1), \quad \text{in } \mathcal{G}(n,p),$$

provided $\liminf_{n\to\infty} \beta_n > 0$.

Remark: The condition $\bar{d}_2 = \bar{d}_1^2(1 + o(m/h^2))$ is rather restrictive. The degree sequences are restricted to those that are very concentrated around their average. So the graphs under consideration are "almost-regular". The condition $d_{\max}^4 = o(h)$ is probably not needed as we only need a lower bound of $|\mathscr{S}_3|$ to verify Theorem 1 (d). However, to avoid complication, we include $d_{\max}^4 = o(h)$ in the assumptions.

The following theorem, proved in [2], is a direct corollary of Theorem 7.

Theorem 8 *Let $0 < p < 1$ be a real and $0 < m < N$ and $d > 0$ be integers satisfying $m = pN$, $d^3 = o(p^2 n)$. Let $X_{n,d}$ denote the number of d-factors in a random graph ($\mathcal{G}(n,m)$ or $\mathcal{G}(n,p)$). Let $\mu_{n,d} = \mathbf{E}_{\mathcal{G}(n,m)}X_{n,d}$, $\lambda_{n,d} = \mathbf{E}_{\mathcal{G}(n,p)}X_{n,d}$ and let $\beta_n = d\sqrt{(1-p)/2p}$. Then $X_{n,d}/\mu_{n,d} \xrightarrow{p} 1$ in $G(n,m)$, and*

$$\frac{\ln(e^{\beta_n^2/2}X_{n,d}/\lambda_{n,d})}{\beta_n} \xrightarrow{d} \mathcal{N}(0,1), \quad \text{in } \mathcal{G}(n,p),$$

provided $\liminf_{n\to\infty} \beta_n > 0$.

Proof of Theorem 7. We generalise the proof in [2] and adapt it to our case of general $\mathbf{d}$. Recall that $F_j(\mathscr{S}_3)$ denotes the set of ordered pairs of graphs $(G_1, G_2) \in \mathscr{S}_3 \times \mathscr{S}_3$ such that $|G_1 \cap G_2| = j$. The following two switchings operating on elements of $\mathscr{S}_3 \times \mathscr{S}_3$ were first defined in [2].

s_1*-switching*: Take an edge $x \in G_1 \cap G_2$. Label the end vertices of x as u_2 and u_2'. Then take an edge $y \in G_1 \setminus G_2$ and label the end vertices of y as u_1 and u_1'. Then take an edge $z \in G_2 \setminus G_1$ and label its end vertices as u_3 and u_3'. An s-switching replaces x and y by $\{u_1, u_2\}$ and $\{u_1', u_2'\}$ in G_1 and replaces x and z by $\{u_2, u_3\}$ and $\{u_2', u_3'\}$ in G_2. An s-switching is applicable on the chosen triple $\{x, y, z\}$ with the given labeling, if and only if

(i) x and y are not adjacent and x and z are not adjacent;

(ii) all of $\{u_1, u_2\}, \{u_1', u_2'\}, \{u_2, u_3\}, \{u_2', u_3'\}$ are not in $G_1 \cup G_2$.

inverse s_1-switching: Choose a pair of 2-paths (u_1, u_2, u_3) and (u_1', u_2', u_3') such that $\{u_1, u_2\}, \{u_1', u_2'\} \in G_1 \setminus G_2$ and $\{u_2, u_3\}, \{u_2', u_3'\} \in G_2 \setminus G_1$. The inverse s-switching replaces $\{u_1, u_2\}$ and $\{u_1', u_2'\}$ by $\{u_1, u_1'\}$ and $\{u_2, u_2'\}$ in G_1 and replaces $\{u_2, u_3\}$ and $\{u_2', u_3'\}$ by $\{u_2, u_2'\}$ and $\{u_3, u_3'\}$ in G_2. The s-switching is applicable on the chosen pair of paths only if

(i') all vertices u_1, u_2, u_3, u_1', u_2', u_3' are distinct;

(ii') none of $\{u_1, u_1'\}$, $\{u_2, u_2'\}$ and $\{u_3, u_3'\}$ are contained in $G_1 \cup G_2$.

Figure 1 gives an example of the s-switching and its inverse, where the solid lines denote edges in G_1 and the dashed lines denote edges in G_2.

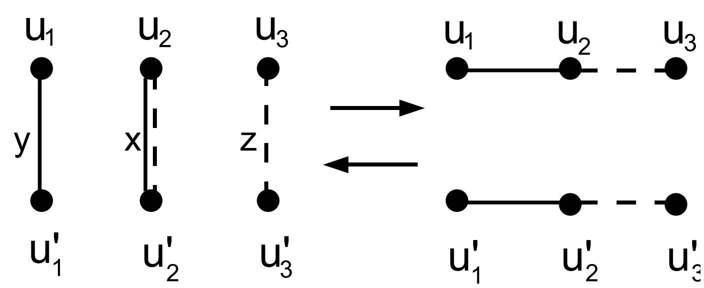

Figure 1: *s-switching and its inverse*

For any $j \geq 1$ and $g \in F_j(\mathscr{S}_3)$, an s_1-switching converts g into an element in $F_{j-1}(\mathscr{S}_3)$. For every g, let $N(g)$ denote the number of s_1-switchings applicable on g. There are j ways to choose x and for each chosen x there are two ways to label its end vertices. For any chosen x, the number of ways to choose y (or z) is $h - j + O(d_{\max}^2)$, where the error term $j + O(d_{\max}^2)$ counts all edges in $G_1 \cap G_2$ and all choices of y such that x and y are adjacent or u_1, u_2 are adjacent or u_1', u_2' are adjacent. For each chosen y (or z), there are two ways to label its end vertices. So $N(g) = 8j(h - j + O(d_{\max}^2))^2$. On the other hand, for any $g' \in F_{j-1}(\mathscr{S}_3)$, an inverse s-switching converts g' into an element in $F_j(\mathscr{S}_3)$. Let $N'(g')$ denote the number of inverse s_1-switchings applicable on g'. Recall that $M_2 = \bar{d}_2 n$. The number of 2-paths (u_1, u_2, u_3) with $\{u_1, u_2\} \in G_1$ and $\{u_2, u_3\} \in G_2$ is $M_2 + O(jd_{\max})$, where $O(jd_{\max})$ accounts for the miscount caused by edges in $G_1 \cap G_2$. Hence $N'(g') = (M_2 + O(jd_{\max}))^2 + O(M_2 d_{\max}^3 + jM_2 d_{\max})$, where the error term $O(M_2 d_{\max}^3)$ accounts for all miscounts that violate constraints (i') and (ii') while the error term $O(jM_2 d_{\max})$ accounts for the case that one of the two paths contains an edge in $G_1 \cap G_2$. Clearly, $\sum_{g \in F_j(\mathscr{S}_3)} N(g) = \sum_{g' \in F_{j-1}(\mathscr{S}_3)} N'(g')$. Thus,

$$r_j = \frac{M_2^2 + O(M_2 d_{\max}^3 + jM_2 d_{\max})}{8j(h - j + O(d^2))^2}.$$

Let $\alpha = 7/8$. For all $1 \leq j \leq \alpha h$, the above ratio is

$$r_j = \frac{M_2^2}{8h^2 j}(1 + O(d_{\max}^3/M_2 + jd_{\max}/M_2 + j/h + d_{\max}^2/h)). \tag{3.2}$$

Since $\bar{d}_2 = \bar{d}_1^2(1+o(m/h^2))$, $M_2 = h^2/n(1+o(m/h^2))$. Thus, we have $M_2^2/8h^2 = h^2/N(1+o(m/h^2))$. Now we verify that for all $j = O(h^2/m)$, all error terms in (3.2) are bounded by $o(m/h^2)$. Note that

$$
\begin{aligned}
&\frac{d_{\max}^3/M_2}{m/h^2} = O(d_{\max}^3 n/m) = O(d_{\max}^3/pn) = o(1);\\
&\frac{jd_{\max}/M_2}{m/h^2} = O(h^2 d_{\max} n/m^2) = o(d_{\max} n/m^{2/3}) = o(d/p^{2/3}n^{1/3}) = o(1);\\
&\frac{(j+d_{\max}^2)/h}{m/h^2} = \frac{(j+d_{\max}^2)h}{m} = O(h^3/m^2 + d_{\max}^2 h/m) = O(d_{\max}^2/m^{1/3}) + o(1)\\
&= O(d_{\max}^2/p^{1/3}n^{2/3}) + o(1) = o(1).
\end{aligned}
$$

Thus, Theorem 1 (a) is verified. Next we verify that condition (b) holds by taking $\gamma(n) = \alpha h$. By (3.2) and the above calculation we have

$$
\begin{aligned}
r_j &= \frac{h^2}{Nj}\Big(1 + o(m/h^2) + O(jd_{\max}/M_2 + j/h)\Big)\\
&= \frac{h^2}{Nj}\Big(1 + o(m/h^2)\Big) + O(d_{\max}/n + h/N) = \frac{h^2}{Nj} + o(m/n^2).
\end{aligned}
$$

Thus, $r_j \leq m/2N$ for all $j \geq 4h^2/m$. This verifies condition (b) by taking $\gamma(n) = \alpha h$. Next, we verify condition (d). By (3.1) and (2.2),

$$
\mathbf{E}_{\mathcal{G}(n,m)}(X_{n,\mathbf{d}}) \sim \frac{(2h)!p^h}{h!2^h\prod_{i=1}^n d_i!}\exp\left(-\frac{M_2-M_1}{2M_1} - \frac{(M_2-M_1)^2}{4M_1^2} - \frac{(1-p)h^2}{2m}\right).
$$

Since $M_2 = O(h^2/n)$, we have

$$
\exp\left(-\frac{M_2-M_1}{2M_1} - \frac{(M_2-M_1)^2}{4M_1^2} - \frac{(1-p)h^2}{2m}\right) = \exp(O(h^2/m)).
$$

We also have

$$
\prod_{i=1}^n d_i! \leq (d_{\max}!)^{2h/d_{\max}}.
$$

Hence

$$
\begin{aligned}
\ln \mathbf{E}_{\mathcal{G}(n,m)}(X_{n,\mathbf{d}}) &\geq h\ln p + h\ln(2h/e) - \frac{2h}{d_{\max}}\ln d_{\max}! - O(h^2/m)\\
&\geq h\left(\ln\left(2hd_{\max}^{3/2}/e\sqrt{n}\right) - \frac{2\ln d_{\max}!}{d_{\max}} + O(h/m)\right) \quad (\text{since } d_{\max}^3 = o(p^2 n))\\
&\geq h\left(\ln\left(2hd_{\max}^{3/2}/e\sqrt{n}\right) - 2\ln d_{\max} + O(h/m)\right).
\end{aligned}
$$

Since $h = \Omega(n)$, we further have $2hd_{\max}^{3/2}/e\sqrt{n} = O(\sqrt{n})$ and so

$$
\ln \mathbf{E}_{\mathcal{G}(n,m)}(X_{n,\mathbf{d}}) \geq h\left(\frac{1}{2}\ln n - \frac{1}{2}\ln d_{\max} + O(1)\right) \to \infty, \tag{3.3}
$$

which verifies condition (d). Lastly, we verify condition (c). Let G be a graph in $\mathscr{S}_3$, and let $\kappa_j(G)$ denote the number of graphs in $\mathscr{S}_3$ that share at least j edges with G. We estimate a uniform upper bound of $\kappa_j(G)$ for all G.

There are $\binom{h}{j}$ ways to choose $h-j$ edges from G. Removing these $h-j$ edges generates a deficiency degree sequence $\mathbf{d}'$, where $d'_i = d_i - a_i$, where a_i is the number of edges incident with G that are removed. Hence $\sum_{i=1}^n d'_i = 2(h-j)$ and for any G,

$$\kappa_j(G) \le \binom{h}{j} \max\left\{ g(\mathbf{d}') : \ \mathbf{d}' \text{ with } \sum_{i=1}^n d'_i = 2(h-j) \right\},$$

where $g(\mathbf{d}')$ denotes the number of graphs with degree sequence $\mathbf{d}'$. By (3.1), $g(\mathbf{d}') < M!/2^{M/2}(M/2)! < (M/2)^{M/2}$, where $M = 2(h-j)$. Therefore,

$$\sum_{j\ge\gamma(n)} f_j = \sum_G \kappa_{(1-\alpha)h}(G) \le |\mathscr{S}_3| \binom{h}{(1-\alpha)h} (h\alpha)^{h\alpha}. \tag{3.4}$$

Recall that $\alpha = 1/8$. By (3.3), (3.4) and the assumption $d_{\max} = o(n^{1/3})$, it is straightforward to check that

$$\sum_{j\ge\gamma(n)} f_j = o(|\mathscr{S}_3|\mu_n),$$

which completes the proof of the theorem. ∎

3.4 triangle-free subgraphs

In this section, we consider another example where $\mathscr{S}_4$ is the set of all triangle-free graphs on S with h edges. Then $X_n(\mathscr{S}_4)$ counts the number of triangle-free subgraphs with h edges.

Theorem 9 *Let $0 < p < 1$ be a real and $0 < m < N$ an integer satisfying $m = pN$, $h^3 = o(m^2)$ (or equivalently $h^3 = o(p^2n^4)$) and $h^4 = o(pn^5)$. Let X_n denote the number of triangle-free subgraphs with h edges. Let μ_n and λ_n be defined as in (2.1) and let $\beta_n = h\sqrt{(1-p)/pN}$. Then $X_n/\mu_n \xrightarrow{p} 1$ in $G(n,m)$, and*

$$\frac{\ln(e^{\beta_n^2/2}X_n/\lambda_n)}{\beta_n} \xrightarrow{d} \mathcal{N}(0,1), \quad \text{in } \mathcal{G}(n,p),$$

provided $\liminf_{n\to\infty} \beta_n > 0$.

Proof. Recall that $F_j(\mathscr{S}_4) = \{(G_1,G_2) \in \mathscr{S}_4\times\mathscr{S}_4 : |G_1\cap G_2| = j\}$. Consider $j \ge 1$ and the classes $F_j(\mathscr{S}_4)$ and $F_{j-1}(\mathscr{S}_4)$. Let K_n denote the complete graph on S. We define two other switchings operating on $\mathscr{S}_4 \times \mathscr{S}_4$ as follows.

s_2-switching: Let x be an edge in $G_1 \cap G_2$. Choose y and z from $K_n \setminus G_1 \cup G_2$, such that $G_1 \cup y$ and $G_2 \cup z$ are triangle-free. Replace x by y in G_1 and replace x by z in G_2.

inverse s_2-switching: Let x be an edge in $K_n \setminus G_1 \cup G_2$ such that $G_1 \cup x$ and $G_2 \cup x$ are triangle-free. Let $y \in G_1 \setminus G_2$ and $z \in G_2 \setminus G_1$. Replace y by x in G_1 and replace z by x in G_2.

Clearly, an s_2-switching converts an element $g \in F_j(\mathscr{S}_4)$ to an element $g' \in F_{j-1}(\mathscr{S}_4)$ and an inverse s_2-switching converts an element $g' \in F_{j-1}(\mathscr{S}_4)$ to an element $g \in F_j(\mathscr{S}_4)$ for some $j \ge 1$.

For any $g \in F_j(\mathscr{S}_4)$, let $N(g)$ denote the number of s-switchings that are applicable on g. There are j ways to choose x. Given x, the number of ways to choose y and z is $N - O(h + T_1(g))$ and $N - O(h + T_2(g))$ respectively, where $T_i(g)$ denotes the number of 2-paths in G_i. Let $T(g) = \max\{T_1(g), T_2(g)\}$. So $N(g) = j(N - O(h + T(g)))^2$. We have the following claim.

Claim 10 *$T(g) = O(h^2/n)$.*

Then $N(g) = jN^2(1 + O(h^2/n^3))$. For any $g' \in F_{j-1}(\mathscr{S}_4)$, let $N'(g')$ denote the number of inverse s_2'-switchings applicable on g'. Then $N'(g') = (N - O((2h - j + 1) + T(g')))(h - j + 1)^2 = Nh^2(1 + O(h^2/n^3 + j/h))$. Since $\sum_{g\in F_j(\mathscr{S}_4)} N(g) = \sum_{g'\in F_{j-1}(\mathscr{S}_4)} N'(g')$, we have that for all $j \geq 1$,

$$r_j = \frac{Nh^2}{jN^2}(1 + O(h^2/n^3 + j/h)) = \frac{h^2}{jN}(1 + o(m/h^2) + O(j/h)). \tag{3.5}$$

Note that $O(h^2/n^3) = o(m/h^2)$ because $h^4 = o(pn^5)$. Next we verify conditions (a) and (b) of Theorem 1. For all $j = O(h^2/m)$, $j/h = O(h/m) = o(m/h^2)$ since $h^3 = o(m^2)$. Thus

$$r_j = \frac{Nh^2}{jN^2}(1 + O(h^2/n^3 + j/h)),$$

which verifies condition (a). By (3.5), for all $j \geq 3h^2/m$,

$$r_j = \frac{h^2}{jN}(1 + o(1)) + O(h/N) \leq \frac{m}{2N},$$

which verifies condition (b) (for $\gamma(n) = h$). Next, we verify condition (d). Obviously $\mathscr{S}_4$ is larger than the set of bipartite graphs with h edges and with vertex-bipartition $([n/2], S - [n/2])$. The latter has size $\binom{n^2/4}{h}$. Thus

$$\begin{aligned}
\mathbf{E}_{\mathcal{G}(n,m)} X_n(\mathscr{S}_4) &\geq \binom{n^2/4}{h} p^h \exp\left(-\frac{(1-p)h^2}{2m} + o(1)\right) \\
&= \frac{(n^2p/4)^h}{h!} \exp(-2h^2/n^2 + O(h^3/n^4) - (1-p)h^2/2m + o(1)) \\
&\sim \frac{1}{\sqrt{2\pi h}} \left(\frac{en^2p}{4h}\right)^h \exp(-2h^2/n^2 - (1-p)h^2/2m) \\
&\geq \frac{1}{\sqrt{2\pi h}} \left(\frac{emp}{4h} \exp(-3h/m)\right)^h \quad (\text{as } n^2 > m),
\end{aligned}$$

where the second equality holds because $[n^2/4]_h = (n^2/4)^h \exp(-2h^2/n^2 + O(h^3/n^4))$ and the third asymptotics holds because $h^3 = o(n^4)$ as $h^3 = o(m^2)$ by the assumption. Since $h^3 = o(m^2)$, $\exp(-3h/m) \to 1$. Since $h = \Omega(n)$, we have $m >> n^{3/2}$. We also have $p = m/N = \Theta(m/n^2)$. Thus,

$$\frac{mp}{h} = \Theta\left(\frac{m^2}{n^2 h}\right) >> \frac{m^{4/3}}{n^2} >> 1.$$

This implies that

$$\mathbf{E}_{\mathcal{G}(n,m)} X_n(\mathscr{S}_4) \to \infty, \quad \text{as } n \to \infty. \blacksquare$$

It only remains to prove Claim 10.

Proof of Claim 10. It is sufficient to prove that for any graph G with h edges and n vertices, the number of 2-paths it contains is bounded by $O(h^2/n)$. Let $\mathbf{d} = (d_1, \ldots, d_n)$ denote the degree sequence of G. Then G contains exactly $\sum_{i=1}^{n} d_i(d_i - 1)/2 = \sum_{i=1}^{n} d_i^2/2 - h$ 2-paths. On the other hand, by the Cauchy-Schwarz inequality,

$$\sum_{i=1}^{n} d_i^2 \geq \frac{(\sum_{i=1}^{n} d_i)^2}{n} = \frac{4h^2}{n},$$

which completes the proof of the claim. ∎

3.5 Hamilton cycles

The most interesting examples of $\mathscr{S}$ are perhaps taking $\mathscr{S}$ as the set of graphs that are isomorphic to a given unlabelled graph H on a set of n vertices. However, counting $f_j(\mathscr{S})$ or estimating r_j is normally difficult. In this and the next sections, we consider two such examples. In Section 3.3 we have shown that the number of 2-factors follows the log-normal paradigm. In what follows, we pick two extreme cases from the set of 2-regular graphs on n vertices, as candidates for H. One is the longest possible cycle, the cycle with length n, whereas the other is a collection of shortest possible cycles, i.e. the union of vertex disjoint triangles.

In this section we consider H (H') to be a cycle (directed cycle) with length n and $\mathscr{S}_5$ ($\mathscr{S}'_5$) to be the set of graphs (directed graphs) on S that are isomorphic to H (H'). Thus, $X_n(\mathscr{S}_5)$ and $X_n(\mathscr{S}'_5)$ count the numbers of undirected and directed Hamilton cycles respectively. It is well known that

$$|\mathscr{S}_5| = (n-1)!/2, \text{ and } |\mathscr{S}'_5| = (n-1)!. \tag{3.6}$$

We have the following theorem for the undirected version.

Theorem 11 *Let $0 < p < 1$ be a real and $0 < m < N$ an integer satisfying $m = pN$ and $p >> n^{-1/2}$. Let X_n denote the number of Hamilton cycles in $\mathcal{G}(n, m)$ (or $\mathcal{G}(n, p)$). Let $\mu_n = \mathbf{E}_{\mathcal{G}(n,m)} X_n$ and let $\lambda_n = \mathbf{E}_{\mathcal{G}(n,p)} X_n$. Then $X_n/\mu_n \xrightarrow{p} 1$ in $G(n, m)$. Assume further that* $\limsup_{n\to\infty} p(n) < 1$, *then*

$$\frac{\ln(e^{\beta_n^2/2} X_n/\lambda_n)}{\beta_n} \xrightarrow{d} \mathcal{N}(0, 1), \quad \textit{in } \mathcal{G}(n, p),$$

where $\beta_n = \sqrt{2(1-p)/p}$.

Using almost the same proof as in Theorem 11, we immediately obtain the following paralelling theorem.

Theorem 12 *If all assumptions with N, $\mathcal{G}(n, p)$ and $\mathcal{G}(n, m)$ replaced by $2N$, $\mathcal{D}(n, p)$ and $\mathcal{D}(n, m)$ in Theorem 11 hold, then the same conclusion of Theorem 11 holds (for $\beta_n = \sqrt{(1-p)/p}$ by the definition of β_n in Theorem 3).*

The second moment of the number of directed Hamilton cycles was originally estimated by Wright [11], which was later redone by Frieze and Suen [1] using basically the same approach.

However, extending the proof to the undirected version, done in [10], is not trivial. Indeed, the proof for the undirected version uses much more complicated counting and analysis. In this paper, we give a completely new and much simpler proof for the undirected Hamilton cycles (Theorem 11), using again the switching method. The same proof, with only slight modification of the switchings that cope with directed edges, works for the directed version (Theorem 12). However, for the directed version, we present a second proof instead, following the recursive functions obtained in [11, 1]. We will verify the conditions in Theorem 1 by analysing these recursive functions, which is eventually equivalent to the analysis in [11, 1]. We do so as this is an example to show that with the guidance of Theorem 1, the analysis can become cleaner and more systematic.

We first prove Theorem 11 by defining another two switchings.

Proof of Theorem 11. We define another two switchings as follows.
h-switching: Choose an edge $xy \in G_1 \cap G_2$. Then choose edges $x_1y_1 \in G_1 \setminus G_2$, $x_2y_2 \in G_2 \setminus G_1$ such that xyx_1y_1 and xyx_2y_2 are in a cyclic order in G_1 and G_2 respectively. Replace xy and x_1y_1 by xx_1 and yy_1 in G_1, and replace xy and x_2y_2 by xx_2 and yy_2 in G_2. The h-switching is applicable if and only if

(a) the six vertices x, y, x_i and y_i for $i = 1, 2$ are all distinct;

(b) the edges xx_1 and yy_1 are not in G_2 and the edges xx_2 and yy_2 are not in G_1.

inverse h-switching: Choose a pair of vertices $\{x, y\}$ such that $xy \notin G_1 \cup G_2$. For $i = 1, 2$, choose x_i and y_i such that $xx_i \in G_i$ and $yy_i \in G_i$ and xx_iyy_i is in a cyclic order in G_i. The inverse h-switching replaces xx_i and yy_i by xy and x_iy_i in G_i for $i = 1, 2$. The operation is applicable if and only if

(a') the six vertices x, y, x_i and y_i for $i = 1, 2$ are all distinct;

(b') the edges xx_i and yy_i are not in $G_1 \cap G_2$ for $i = 1, 2$;

(c') $x_1y_1 \notin G_2$ and $x_2y_2 \notin G_1$.

For $g \in F_j$, let $N(g)$ be the number of h-switchings applicable on g. There are $2j$ ways to choose and label the end vertices of the edge $xy \in G_1 \cap G_2$. For any chosen xy, there are $n - j + O(1)$ ways to choose and label the end vertices of the edge $x_iy_i \in G_i$, where $j + O(1)$ accounts for the case that $x_iy_i \in G_1 \cap G_2$ and the case that condition (a) is violated. Thus, a rough estimation of $N(g)$ is $2j(n - j + O(1))^2$. The only miscounts are those xy and x_iy_i such that condition (b) is violated. Clearly, the miscount due to the violation of condition (b) is $O(jn)$ because for any chosen xy, there are exactly two choices for x_1y_1 (equivalently x_2y_2), such that either xx_1 or yy_1 is in G_2 (equivalently, either xx_2 or yy_2 is in G_1). Thus, $N(g) = 2jn^2(1 - j/n + O(n^{-1}))^2$.

On the other hand, for $g' \in F_{j-1}$, let $N'(g')$ denote the number of inverse h-switchings applicable on g'. There are $n^2 - O(n)$ ways to choose and label vertices x and y such that $xy \notin G_1 \cup G_2$. For any chosen xy, there are two ways to choose x_i and y_i from G_i for $i = 1, 2$ respectively, such that xx_i, $yy_i \in G_i$ and xx_iyy_i is in a cyclic order in G_i. Thus, $N'(g')$ is approximately $4(n^2 - O(n))$. The only miscounts are those choices that violate conditions (a') or (b') or (c'). There are only $O(n)$ choices of xy so that (a') or (c') can possibly be violated, and there are only $O(jn)$ choices of xy so that (b') can possibly be violated. Therefore, $N'(g') = 4n^2(1 + O(j/n))$.

Hence for all $1 \le j \le n/2$,

$$r_j = \frac{4n^2}{2jn^2}(1 + O(j/n)) = \frac{2}{j}(1 + O(j/n)),$$

from which we can easily verify Theorem 1 (a), (b) (for $\gamma(n)=n/2$) and condition (d) is trivially true. The proof will be completed by verifying condition (c). Let G be a Hamilton cycle, and let $\kappa_j(G)$ denote the number of Hamilton cycles that share at least j edges with G. There are $\binom{n}{j}$ ways to choose j edges from G. These chosen edges form $r\le j$ disjoint paths. Contract each path into a special vertex. The total number of vertices including these special vertices is then $n-j$. There are $(n-j-1)!/2$ Hamilton cycles on these vertices. For every such Hamilton cycles, expand each special vertex by its corresponding path (there are two ways to expand each special vertex). Then each expanded Hamilton cycle corresponds to a Hamilton cycle that shares at least j edges with G. Thus, for every G,

$$\kappa_j(G)\le\binom{n}{j}\frac{(n-j-1)!}{2}\cdot 2^j<n!2^j/j!.$$

It is then straightforward to verify that

$$\sum_{j\ge n/2} f_j\le|\mathscr{S}_5|n!2^j/j!=o(|\mathscr{S}_5|\mu_n).\ \blacksquare$$

A second proof of Theorem 12. For a given directed cycle H of length n, let $f'_j(n)$ denote the number of directed Hamilton cycles on the same vertex set, which shares exactly j edges with H. Then $f_j=|\mathscr{S}'_5|f'_j(n)$ for all j. Thus, $r_j=f'_j(n)/f'_{j-1}(n)$. It was proved in [11, 1] that

$$\begin{aligned}
f'_0(n) &= \sum_{k=0}^{n-1}\binom{n}{k}(-1)^k(n-k-1)!+(-1)^n,\quad\text{for all } n\ge 1; && (3.7)\\
f'_j(n) &= \binom{n}{j}f_0(n-j),\quad\text{for all } j\le n-1; && (3.8)\\
f'_n(n) &= 1,\quad\text{for all } n\ge 0. && (3.9)
\end{aligned}$$

We give a short sketch of (3.7)–(3.9). The last equation is trivial. The equation (3.8) is obtained by contracting paths formed by edges contained in $G_1\cap G_2$ as described in the proof of Theorem 11. The nice property for the directed version is that after contracting these paths, the resulting two Hamilton cycles are edge disjoint, which is not the case for the undirected version. The equation (3.8) follows by observing that there is a unique way to expand each path to obtain the original directed Hamilton cycles. The equation (3.7) follows from an inclusion-exclusion argument.

Thus, by (3.7) and (3.8), for all $j\le n-1$,

$$\begin{aligned}
r_j &= \frac{\binom{n}{j}f'_0(n-j)}{\binom{n}{j-1}f'_0(n-j+1)}\\
&= \frac{n-j+1}{j}\cdot\frac{\sum_{k=0}^{n-j-1}\binom{n-j}{k}(-1)^k(n-j-k-1)!+(-1)^{n-j}}{\sum_{k=0}^{n-j}\binom{n-j+1}{k}(-1)^k(n-j-k)!+(-1)^{n-j+1}}\\
&= \frac{n-j+1}{j}\cdot\frac{(n-j)!\sum_{k=0}^{n-j-1}(-1)^k/k!(n-j-k)+(-1)^{n-j}}{(n-j)!(n-j+1)\sum_{k=0}^{n-j}(-1)^k/k!(n-j-k+1)+(-1)^{n-j+1}}\\
&= \frac{1}{j}\cdot\frac{\sum_{k=0}^{n-j-1}(-1)^k/k!(n-j-k)+(-1)^{n-j}/(n-j)!}{\sum_{k=0}^{n-j}(-1)^k/k!(n-j-k+1)+(-1)^{n-j+1}/(n-j)!(n-j+1)}.
\end{aligned}$$

Let

$$H(n,j)=\sum_{k=0}^{n-j-1}\frac{(-1)^k}{k!(n-j-k)}.$$

Next we estimate $H(n,j)$. First consider j such that $j\le n-2\ln n$. Let $k^*=\max\{\lceil \ln n\rceil, \lceil m(n-j)/n^2\ln n\rceil\}$.

$$H(n,j)=\frac{1+O(k^*/(n-j))}{n-j}\sum_{k=0}^{k^*}\frac{(-1)^k}{k!}+\sum_{k=k^*+1}^{n-j-1}\frac{(-1)^k}{k!(n-j-k)}.$$

By the choice of k^*, $k^*/(n-1)=o(m/n^2)$. We also have

$$\sum_{k=k^*+1}^{\infty}\frac{(-1)^k}{k!}=O((k^*!)^{-1})=O((e/k^*)^{k^*})=O(n^{-3}),$$

as $k^*\ge\ln n$. Thus,

$$H(n,j)=\frac{1+o(m/n^2)}{n-j}(e^{-1}+O(n^{-3}))+O(n^{-3})=(1+o(m/n^2))\frac{e^{-1}}{n-j}.$$

Hence, for all $j\le n-2\ln n$,

$$r_j=\frac{1}{j}\frac{H(n,j)+(-1)^{n-j}/(n-j)!}{H(n+1,j)+(-1)^{n-j+1}/(n-j+1)!}=\frac{1}{j}(1+o(m/n^2)).$$

This verifies conditions (a) and (b) (by taking $\gamma(n)=n-2\ln n$) of Theorem 1. Condition (c) follows in an analogous argument as in the proof of Theorem 11 and condition (d) holds trivially. ∎

3.6 Collection of disjoint triangles

In this section, we assume $n\equiv 0 \pmod 3$ and consider H (H') to be the unlabelled graph on n vertices consist of $n/3$ vertex disjoint triangles (directed triangles). Let $\mathscr{S}_6$ ($\mathscr{S}_6'$) denote the set of graphs on S that are isomorphic to H (H'). Then

$$|\mathscr{S}_6|=\frac{n!}{6^{n/3}(n/3)!},\quad |\mathscr{S}_6'|=\frac{n!}{3^{n/3}(n/3)!}. \tag{3.10}$$

The following theorem determines the limiting distribution of $X_n=X_n(\mathscr{S}_6)$.

Theorem 13 *Let $0<p<1$ be a real and $0<m<N$ an integer satisfying $m=pN$ and $\liminf_{n\to\infty}p(n)>0$. Let X_n denote the number of subgraphs that are isomorphic to a set of $n/3$ vertex disjoint triangles. Let $\mu_n=\mathbf{E}_{\mathcal{G}(n,m)}X_n$ and let $\lambda_n=\mathbf{E}_{\mathcal{G}(n,p)}X_n$. Then $X_n/\mu_n\xrightarrow{p}1$ in $\mathcal{G}(n,m)$. Assume further that $\limsup_{n\to\infty}p(n)<1$, then*

$$\frac{\ln(e^{\beta_n^2/2}X_n/\lambda_n)}{\beta_n}\xrightarrow{d}\mathcal{N}(0,1),\quad in\ \mathcal{G}(n,p),$$

where $\beta_n=\sqrt{2(1-p)/p}$.

Remark: Indeed, the condition of $\liminf_{n\to\infty} p(n) > 0$ can be replaced by $p(n) \geq n^{-\delta}$, for some small constant δ. For instance, we checked that $\delta = 1/16$ works and there is still room for further improvement. However, $p >> n^{-1/2}$ does not seem to be sufficient. For the purpose of a cleaner presentation, we only consider $\liminf_{n\to\infty} p(n) > 0$ in the proof. For readers who are interested in improving the condition of p, we give quite tight bounds in Lemmas 15 and 16, and we also point out here that there is plenty of room in the proofs of Lemma 18 and Theorem 13 to improve the range of p.

Almost the same proof of the previous theorem, with slight modifications of the switchings defined in the proof of Theorem 13, concerning the directions of edges, yields the following corresponding theorem for the directed version.

Theorem 14 *If all assumptions with N, $\mathcal{G}(n,p)$ and $\mathcal{G}(n,m)$ replaced by $2N$, $\mathcal{D}(n,p)$ and $\mathcal{D}(n,m)$ in Theorem 13hold, then the same conclusion of Theorem 13 holds (for $\beta_n = \sqrt{(1-p)/p}$ by the definition of β_n in Theorem 3).*

For any $(G_1, G_2) \in \mathscr{S}_6 \times \mathscr{S}_6$, the edges in G_1 and G_2 can intersect in two ways. We say $e \in G_1 \cap G_2$ is of type 1 if the triangles $T_i \in G_i$ with $e \in T_i$ for $i = 1, 2$ are distinct. We say e is of type 2 if T_1 and T_2 are on the same vertex set.

Let $F_{\ell,t}$ denote the set of $(G_1, G_2) \in \mathscr{S}_6 \times \mathscr{S}_6$ such that number of edges in $G_1 \cap G_2$ of type 1 and 2 is ℓ and t respectively. Clearly $F_{\ell,t}$ is non-empty only if t is a multiple of 3. Clearly $F_j(\mathscr{S}_6) = \cup_k F_{j-3k,3k}$. Let $f_{\ell,t} = |F_{\ell,t}|$. Then $f_j = \sum_{k=0}^{\lfloor j/3 \rfloor} f_{j-3k,3k}$.

Lemma 15 *For any $t \geq 0$ and $\ell \geq 1$ such that $n - 4\ell - 3t - 1 > 0$ and $n - 3\ell - 3t - 12 > 0$,*

$$\frac{2}{\ell} \frac{(n-4\ell-3t-1)^2}{(n-3\ell-3t)^2} \leq \frac{f_{\ell,3t}}{f_{\ell-1,3t}} \leq \frac{2}{\ell} \frac{(n-4\ell-3t+4)^2}{(n-3\ell-3t-12)^2}.$$

Proof. We define two switchings operating on $\mathscr{S}_6 \times \mathscr{S}_6$ as shown in Figure 2.
t_1-switching: Take an edge of type 1 in $G_1 \cap G_2$ and label the end vertices x and y. Let u (v) be the vertex that is adjacent to both x and y in G_1 (G_2). Take a triangle T_1 (T_2) in G_1 (G_2) that is distinct from xyu (xyv) which does not contain any edge in $G_1 \cap G_2$. Label the vertices of T_1 (T_2) as $u_1u_2u_3$ ($v_1v_2v_3$). Replace these four triangles in $G_1 \cup G_2$ by xuu_1, $yu_2u_3 \in G_1$ and xvv_1, $yv_2v_3 \in G_2$. The t_1-switching is applicable only if $v \notin T_1$, $u \notin T_2$ and $T_1 \cap T_2 = \emptyset$. See Figure 2.
inverse t_1-switching: A vertex x is pure if both triangles containing x in G_1 and G_2 do not contain any edge in $G_1 \cap G_2$. Choose a pure vertex x and label its neighbours in G_1 (G_2) as u and u_1 (v and v_1). Then choose another pure vertex y that is distinct from x, u_i and v_i for $i = 1, 2$. Label the neighbours of y in G_1 (G_2) as u_2 and u_3 (v_2 and v_3). Replace these four triangles under consideration by xyu, $u_1u_2u_3 \in G_1$ and xyv, $v_1v_2v_3 \in G_2$.

For any $g = (G_1, G_2) \in F_{\ell,3t}$, let $N(g)$ be the number of t_1-switchings that are applicable on g. Clearly $N(g) \leq 2\ell(6(n/3 - (\ell+t)))^2$, as there are 2 ways to label x and y for a chosen edge from $G_1 \cap G_2$, and in G_1 (G_2) there are at most $n/3 - (\ell+t)$ choices for the triangle $u_1u_2u_3$ ($v_1v_2v_3$) and for each choice there are 6 ways to label the vertices. We also have

$$N(g) \geq 2\ell \cdot 6(n/3 - (\ell+t) - 1) \cdot 6(n/3 - (\ell+t) - 4),$$

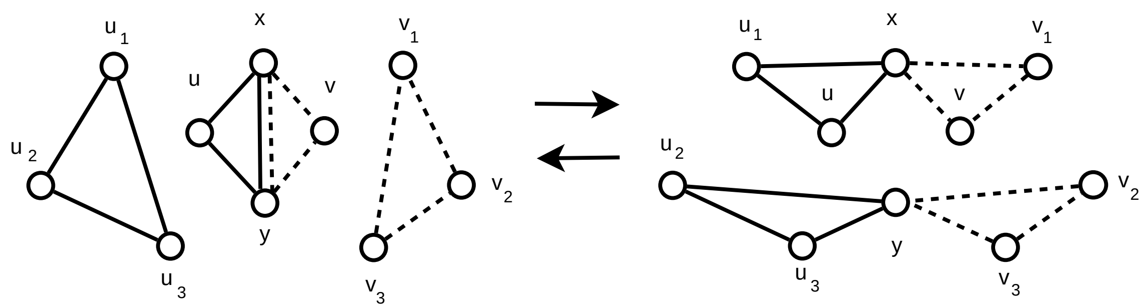

Figure 2: t_1*-switching and its inverse*

because for any chosen xy, the number of triangles in G_1 which contain no edges in G_2 and do not contain v is at least $n/3-(\ell+t)-1$, whereas given the triangle $u_1u_2u_3$, the number of triangles in G_1 which contain no edges in G_1 and do not contain any of u, u_i, $i=1,2,3$ is at least $n/3-(\ell+t)-4$. On the other hand, for any $g'=(G_1,G_2)\in F_{\ell-1,3t}$, let $N'(g')$ be the number of inverse t_1-switchings applicable on g'. The number of pure vertices is exactly $n-4(\ell-1)-3t$. Hence the number of ways to choose x is $n-4(\ell-1)-3t$ and for any chosen x, the number of ways to label u, u_1, v, v_1 is 4. The number of ways to choose y is $n-4(\ell-1)-3t-\delta$, where δ counts the number of pure vertices among x, u, u_1, v and v_1. Therefore, $1\le\delta\le 5$ always. Hence,

$$\frac{16(n-4(\ell-1)-3t-5)^2}{2\ell\cdot(6(n/3-(\ell+t)))^2}\le\frac{f_{\ell,3t}}{f_{\ell-1,3t}}\le\frac{16(n-4(\ell-1)-3t)^2}{2\ell\cdot 36(n/3-(\ell+t)-4)^2}. \quad \blacksquare$$

Lemma 16 *For any* $\ell\ge 0$ *and* $t\ge 1$,

$$\frac{f_{\ell,3t}}{f_{\ell,3(t-1)}}=\frac{32(n-4\ell-3t)^3}{3(n-3\ell-3t)^4}(1+O(1/(n-4\ell-3t))).$$

Proof. We define another two switching operations on $\mathscr{S}_6\times\mathscr{S}_6$ as shown in Figure 3.
t_2-switching: Let xyz be a triangle that is contained in both G_1 and G_2. Take two distinct triangles from G_1 (G_2) which do not contain any edge in $G_1\cap G_2$ and label the end vertices as $x_1y_1z_1$ and $x_2y_2z_2$ ($x_1'y_1'z_1'$ and $x_2'y_2'z_2'$) respectively. Replace the six triangles under consideration by $aa_1a_2\in G_1$ and $aa_1'a_2'\in G_2$, where $a\in\{x,y,z\}$. This switching is applicable only if all these fifteen vertices a, a_i, a_i' for $a\in\{x,y,z\}$ and $i=1,2$ are distinct.

inverse t_2-switching: Recall from the definition of inverse t_1-switching that a vertex x is pure if both triangles containing x in G_1 and G_2 do not contain any edge in $G_1\cap G_2$. Choose three pure vertices a, $a\in\{x,y,z\}$ and label the neighbours of a in G_1 (G_2) by a_1 and a_2 (a_1' and a_2'). The inverse t_2-switching replaces the six triangles under consideration by xyz, $x_iy_iz_i\in G_1$ for $i=1,2$ and xyz, $x_i'y_i'z_i'\in G_2$ for $i=1,2$. This switching is applicable only if all these fifteen vertices a, a_i, a_i' for $a\in\{x,y,z\}$ and $i=1,2$ are distinct.

For any $g\in F_{\ell,3t}$ and $g'\in F_{\ell,3t-3}$, define $N(g)$ and $N'(g')$ the same way as in the proof of

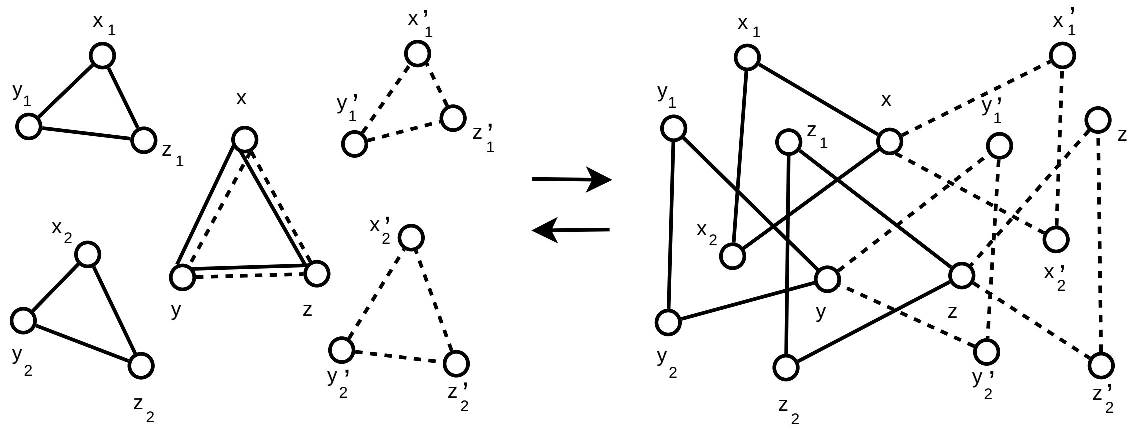

Figure 3: t_2-*switching and its inverse*

Lemma 15. Following an analogous argument of Lemma 15, it is not hard to show that

$$6t \cdot 6^2 \binom{n/3 - (\ell + t) - 6}{2}^2 \leq N(g) \leq 6t \cdot 6^2 \binom{n/3 - (\ell + t)}{2}^2$$
$$(4(n - 4(\ell - 1) - 3t - 10))^3 \leq N'(g') \leq (4(n - 4(\ell - 1) - 3t))^3.$$

Thus,

$$\frac{32(n - 4\ell - 3t - 6)^3}{3(n - 3\ell - 3t)^4} \leq \frac{f_{\ell,3t}}{f_{\ell,3(t-1)}} \leq \frac{32(n - 4\ell - 3t + 4)^3}{3(n - 3\ell - 3t - 21)^4}. \quad \blacksquare$$

Corollary 17 *For all $j = o(n)$,*

$$\frac{f_{j-3k-3,3k+3}}{f_{j-3k,3k}} \sim \frac{4[j - 3k - 1]_3}{3n}.$$

Proof. This follows by Lemmas 15 and 16 and

$$\frac{f_{j-3k-3,3k+3}}{f_{j-3k,3k}} = \frac{f_{j-3k-3,3k+3}}{f_{j-3k-3,3k}} \prod_{i=0}^{2} \frac{f_{j-3k-i-1,3k}}{f_{j-3k-i,3k}}. \quad \blacksquare$$

Lemma 18 *Assume* $\liminf_{n\to\infty} p(n) > 0$. *Let* $\gamma(n) = n/\ln\ln n$. *Then*

$$\sum_{j \geq \gamma(n)} f_j = o(|\mathscr{S}_6|\mu_n).$$

Proof. Let $G \in \mathscr{S}_6$ and let $\kappa_j(G)$ be the number of graphs in $\mathscr{S}_6$ which shares at least j edges with G. We estimate an upper bound of $\kappa_j(G)$. Let $j = \ell + 3t$ and we consider the number of graphs G' in $\mathscr{S}_6$ that shares at least ℓ and $3t$ edges of type 1 and 2 respectively with G. Then there are $\binom{n/3}{t}$ ways to choose the t triangles contained both in G and G'. Then there are $\binom{n/3-t}{\ell}3^\ell$ ways to choose the ℓ triangles in G that contain the ℓ edges of type 1 and to locate these ℓ edges. Given these ℓ edges in G', there are at most $[n-3t-2\ell]_\ell$ ways to choose another ℓ vertices to form the ℓ triangles in G'. Then there are at most

$$\frac{(n-3t-3\ell)!}{6^{n/3-t-\ell}(n/3-t-\ell)!} \le 9^n n^{2(n/3-t-\ell)}$$

ways to partition the remaining $n-3t-3\ell$ vertices into vertex disjoint triangles in G'. Hence

$$\kappa_j(G) \le \sum_\ell \binom{n/3}{t}\binom{n/3-t}{\ell} 3^\ell [n-3t-2\ell]_\ell 9^n n^{2(n/3-t-\ell)} \le n \cdot \max_\ell \{n^t n^{2\ell} \ell^{-\ell} 9^n n^{2(n/3-t-\ell)}\},$$

where $t=(j-\ell)/3$. Thus,

$$\ln(\kappa_j(G)) \le \max_\ell \{(2n/3-t)\ln n - \ell \ln(\ell)\} + O(n).$$

We consider only $j \ge \gamma(n)$. So the maximum is achieved at $\ell = n^{1/3}$. Thus

$$\ln(\kappa_j(G)) \le \frac{2n}{3}\ln n - \frac{j}{3}\ln n + O(n),$$

We also have

$$\ln \mu_n = n \ln p + \frac{2n}{3}\ln n + O(n).$$

So

$$\ln(\kappa_j(G)) - \ln \mu_n \le -\frac{j}{3}\ln n - n\ln p + O(n) \to -\infty,$$

as $n \to \infty$ since $\liminf_{n\to\infty} p(n) > 0$, which completes the proof of the lemma. ∎

Proof of Theorem 13. For any $j \ge 0$,

$$r_j = \sum_{k=0}^{\lfloor j/3 \rfloor} f_{j-3k,3k} \Big/ \sum_{k=0}^{\lfloor (j-1)/3 \rfloor} f_{j-1-3k,3k}. \tag{3.11}$$

By Corollary 17, for all $j = o(n^{1/3})$, $r_j \sim f_{j,0}/f_{j-1,0}$. By Lemma 15, this ratio is asymptotic to $2/j$. This verifies Theorem 1 (a). Let $\gamma(n) = n/\ln\ln n$. Lemma 18 verifies condition (c) whereas condition (d) is trivially true. The proof is completed by verifying condition (b). Since $r_j \sim 2/j$ for all $j = o(n^{1/3})$, we only need to show that for all $n^{1/3}/\ln n \le j \le \gamma(n)$, $r_j \le m/2N$. It follows directly from the following two facts.

(a) Let $\widehat{k} = \min\{k : j-3k \le \ln n\}$. By Corollary 17,

$$\sum_{k=0}^{\lfloor j/3 \rfloor} f_{j-3k,3k} \sim \sum_{k=0}^{\widehat{k}} f_{j-3k,3k}, \qquad \sum_{k=0}^{\lfloor (j-1)/3 \rfloor} f_{j-1-3k,3k} \sim \sum_{k=0}^{\widehat{k}} f_{j-1-3k,3k}.$$

(b) By Lemma 15, for all $0 \le k \le \widehat{k}$, $f_{j-3k,3k}/f_{j-1-3k,3k} = o(1)$. ∎

4 Proofs of Theorems 1 and 3

Before approaching Theorems 1 and 3, we first prove a technical lemma.

Lemma 19 *Let $N = \binom{n}{2}$ and let $p = m(n)/N$, where $0 < m(n) < N$. Then for any integer $\ell = \ell(n) \geq 0$ such that $\limsup_{n\to\infty} \ell(n)/m(n) < 1$,*

$$\binom{N-\ell}{m-\ell}/\binom{N}{m} = p^\ell \exp\left(-\frac{1-p}{pN}\frac{\ell^2-\ell}{2} + O(\ell^3/m^2)\right).$$

Moreover, if $\ell = \Omega(\sqrt{m})$, then

$$\binom{N-\ell}{m-\ell}/\binom{N}{m} = p^\ell \exp\left(-\frac{1-p}{pN}\frac{\ell^2}{2} + O(\ell^3/m^2)\right).$$

Proof.

$$\begin{aligned}
\binom{N-\ell}{m-\ell}/\binom{N}{m} &= \frac{[m]_\ell}{[N]_\ell} = \prod_{i=0}^{\ell-1}\frac{m-i}{N-i} \\
&= \prod_{i=0}^{\ell-1}\frac{m}{N}\exp\left(-\frac{i}{m}+\frac{i}{N}+O(i^2/m^2)\right) \quad (\text{since } \limsup_{n\to\infty} \ell(n)/m(n) < 1) \\
&= p^\ell \exp\left(-\frac{1-p}{pN}\frac{\ell^2-\ell}{2} + O(\ell^3/m^2)\right).
\end{aligned}$$

If we have further that $\ell = \Omega(\sqrt{m})$, then $\ell/pN = O(\ell^3/m^2)$. ∎

Proof of Theorem 1. In this proof, the probability space refers to the random graph $\mathcal{G}(n,m)$ only. Let $s = |\mathscr{S}|$. By (2.1) and (2.2),

$$\mathbf{E}X_n = s(m/N)^h \exp\left(-\frac{N-m}{mN}\frac{h^2}{2} + O(h^3/m^2)\right).$$

We also have

$$\mathbf{E}X_n^2 = \sum_{j=0}^{h} f_j \binom{N-(2h-j)}{m-(2h-j)}/\binom{N}{m}.$$

Let $g(j) = f_j\binom{N-(2h-j)}{m-(2h-j)}/\binom{N}{m}$. By condition (a), for every $K > 0$ and any $1 \leq j \leq Kh^2/m$,

$$\frac{g(j)}{g(j-1)} = r_j \cdot \frac{N}{m}(1+O(h/m)) = \frac{h^2}{mj}(1+O(h/m)+o(m/h^2)) = \frac{h^2}{mj}(1+o(m/h^2)), \tag{4.1}$$

where the last equality holds because $h^3 = o(m^2)$. By condition (c) and the fact that for any integer $0 \leq j \leq h$, $\binom{N-(2h-j)}{m-(2h-j)} \leq \binom{N-h}{m-h}$, we also have that

$$\sum_{j>\gamma(n)} g(j) \leq t(n)\binom{N-h}{m-h}/\binom{N}{m} = t(n)\mu_n/s.$$

Then for all sufficiently large $K > 0$,

$$\begin{aligned}\mathbf{E}X_n^2 &= \sum_{j=0}^{h} g(j) = \sum_{j=0}^{Kh^2/m} g(j) + O(g(Kh^2/m)) + O(t(n)\mu_n/s) \\ &= \left(1 + O\left(K^{-1}\right)\right) \sum_{j=0}^{Kh^2/m} g(j) + O(t(n)\mu_n/s),\end{aligned} \tag{4.2}$$

where the second equality holds because of condition (b) and the last equality holds by (4.1). Next, we estimate $\sum_{j=0}^{Kh^2/m} g(j)$. By (4.1) and Lemma 19,

$$\begin{aligned}&\sum_{j=0}^{Kh^2/m} g(j) = f_0 \frac{\binom{N-2h}{m-2h}}{\binom{N}{m}} \sum_{j=0}^{Kh^2/m} \frac{(h^2/m)^j}{j!}(1 + o(jm/h^2)) \\ &= f_0 \cdot (m/N)^{2h} \exp\left(-\frac{N-m}{mN}\frac{(2h)^2}{2} + O(h^3/m^2)\right)\Big(\exp(h^2/m + o(K)) + \Gamma(K)\Big), \\ &= f_0 \cdot (m/N)^{2h} \exp\left(-\frac{N-m}{mN}2h^2\right)\exp(h^2/m)\Big(1 + o(K) + O(\Gamma(K)\exp(-h^2/m))\Big),\end{aligned} \tag{4.3}$$

where

$$\Gamma(K) = O\left(\frac{(h^2/m)^{Kh^2/m}}{(Kh^2/m)!}\right) = O\left(\left(\frac{(eh^2/m)}{(Kh^2/m)}\right)^{Kh^2/m}\right).$$

Letting $K \to \infty$ in both (4.2) and (4.3), we have $\Gamma(K) \to 0$, since $h^2/m = \Omega(1)$. Thus,

$$\mathbf{E}X_n^2 = (1 + o(1))f_0 \cdot (m/N)^{2h} \exp\left(-\frac{N-m}{mN}2h^2\right)\exp(h^2/m) + O(t(n)\mu_n/s). \tag{4.4}$$

We also have

$$s^2 = \sum_{j=0}^{h} f_j = f_0 \sum_{j=0}^{h}\prod_{i=1}^{j} r_i.$$

With the same reasoning as before, it is enough to sum over the first Kh^2/N terms, leaving a negligible tail plus an error term $O(t(n))$, and then let $K \to \infty$. This yields

$$s^2 = (1 + o(1))f_0 \exp(h^2/N) + O(t(n)).$$

Since $t(n) = o(\mu_n s) = o(s^2)$ by condition (c), we obtain

$$f_0 \sim s^2 \exp(-h^2/N).$$

Combining with (4.4) and again by condition (c), we obtain

$$\begin{aligned}\mathbf{E}X_n^2 &= (1 + o(1))s^2(m/N)^{2h} \exp\left(-\frac{N-m}{mN}2h^2\right)\exp(h^2/m - h^2/N) + O(t(n)\mu_n/s) \\ &= (1 + o(1))s^2(m/N)^{2h} \exp\left(-\frac{N-m}{mN}h^2\right) + o(\mu_n^2) \sim (\mathbf{E}X_n)^2.\end{aligned}$$

By condition (d), $\mathbf{E}X_n \to \infty$ as $n \to \infty$. Then for every $\epsilon > 0$,

$$\mathbf{P}(|X_n/\mathbf{E}X_n - 1| > \epsilon) \to 0, \quad \text{as } n \to \infty,$$

by Chebyshev's inequality. ∎

Proof of Theorem 3. Let Y_n denote the number of edges in $\mathcal{G}(n,p)$, then $Y_n \sim Bin(N,p)$. Hence we have

$$Y_n - pN = O_p(\sqrt{p(1-p)N}), \tag{4.5}$$

where $f(n) = O_p(g(n))$ for some $g(n) \geq 0$ means $\mathbf{P}(|f(n)| > Kg(n)) \to 0$ as $K \to \infty$ and $n \to \infty$. Similarly we use the notation $f(n) = o_p(g(n))$ meaning that for every $\epsilon > 0$, $\mathbf{P}(|f(n)| > \epsilon g(n)) \to 0$ as $n \to \infty$. Since $X_n/\mathbf{E}_{\mathcal{G}(n,m)X_n} \xrightarrow{p} 1$ in $\mathcal{G}(n,m)$ for all $m = pN + O(\sqrt{p(1-p)N})$ by assumption and $\ln(\mathbf{E}_{\mathcal{G}(n,m)}X_n) = \ln|\mathscr{S}| + h\ln(m/N) + (N-m)h^2/2mN + o(1)$ by (2.2), by conditioning on Y_n, we have

$$\ln X_n - \ln|\mathscr{S}| - h\ln(Y_n/N) + \frac{1-Y_n/N}{Y_n}\frac{h^2}{2} \xrightarrow{p} 0. \tag{4.6}$$

By (4.5),

$$\frac{1-Y_n/N}{Y_n}\frac{h^2}{2} = \frac{h^2(1-p)}{2Np}\left(1 + O_p\left(\sqrt{\frac{p}{(1-p)N}} + \sqrt{\frac{1-p}{pN}}\right)\right) = \frac{h^2(1-p)}{2Np} + o_p(1), \tag{4.7}$$

where the equality above holds because $h^3 = o(p^2n^4)$. We also have

$$\ln(Y_n/N) = \ln p(1 + Y_n^*\sqrt{(1-p)/pN}) = \ln p + \sqrt{(1-p)/pN}Y_n^* + O_p((1-p)/pN), \tag{4.8}$$

where

$$Y_n^* = \frac{Y_n - pN}{\sqrt{p(1-p)N}}$$

is the normalised variable of Y_n. Recall that $\lambda_n = |\mathscr{S}|p^h$ from (2.1) and $\mathbf{E}X_n = \lambda_n$. Combining with (4.6)–(4.8), we have

$$\ln(X_n/\lambda_n) + \frac{\beta_n^2}{2} = \beta_n Y_n^* + o_p(1). \tag{4.9}$$

Since $\beta_n = \Omega(1)$, (4.9) immediately yields

$$\frac{\ln(e^{\beta_n^2/2}X_n/\lambda_n)}{\beta_n} = Y_n^* + o_p(1).$$

Since $Y_n^* \xrightarrow{d} \mathcal{N}(0,1)$, the theorem follows. ∎

5 Concluding remarks

It was proved in [4] that $m >> n^{3/2}$ is required for the concentration of X_n in $\mathcal{G}(n,m)$, where X_n denotes the number of Hamilton cycles or perfect matchings or spanning trees, as the variable will become asymptotically log-normally distributed when $m = \Theta(n^{3/2})$. We believe that most of the ranges of p that we presented in the paper are tight, except for sets of vertex disjoint triangles. It is also a little surprising that the critical point of m when X_n changes from small deviation ($\mathbf{E}X_n^2 \sim (\mathbf{E}X_n)^2$) to large deviation ($\limsup_{n\to\infty} \mathbf{E}X_n^2/(\mathbf{E}X_n)^2 > 1$) in $\mathcal{G}(n,m)$ seems to be different for Hamilton cycles and for sets of vertex disjoint triangles. We guess $m = n^{5/3}$ might be the critical point for the latter case.

As explained in Section 3.5, the most interesting set $\mathscr{S}$ to be studied is perhaps the one containing graphs isomorphic to an unlabelled graph H on n vertices. Unfortunately, for a general H, both f_j and r_j seem hard to compute. It will be interesting to know whether for all such graphs H, the corresponding random variables X_n follow the log-normal paradigm. If not, is it possible to characterise the class of H, for which the distribution of X_n follows this pattern?

References

[1] A. Frieze and S. Suen, Counting the number of Hamilton cycles in random digraphs, *Random Structures Algorithms* 3 (1992), no. 3, 235-241.

[2] P. Gao, Distribution of spanning regular subgraphs in random graphs, preprint.

[3] S. Janson, Random regular graphs: asymptotic distributions and contiguity, *Combin. Probab. Comput.* 4 (1995), no. 4, 369–405.

[4] S. Janson, The numbers of spanning trees, Hamilton cycles and perfect matchings in a random graph, *Combin. Probab. Comput.* 3 (1994), 97–126.

[5] B.D. McKay, Asymptotics for symmetric 0-1 matrices with prescribed row sums, *Ars Combinatoria* 19A (1985), 15–25.

[6] A. Rućinski, Subgraphs of random graphs: a general approach, *Random graphs* ’83 (Poznan, 1983), pp. 221–229, North-Holland Math. Stud. 118, North-Holland, Amsterdam, 1985.

[7] A. Ruciński, When are small subgraphs of a random graph normally distributed? *Probab. Theory Related Fields*, 78 (1988), no. 1, 1–10.

[8] R. W. Robinson and N.C. Wormald, Almost all cubic graphs are Hamiltonian, *Random Structures Algorithms* 3 (1992), no. 2, 117–125.

[9] R.W. Robinson and N. C. Wormald, Almost all regular graphs are Hamiltonian, *Random Structures Algorithms* 5 (1994), no. 2, 363–374.

[10] E. M. Wright, For how many edges is a graph almost certainly Hamiltonian? *J. London Math. Soc.* (2) 8 (1974), 44-48.

[11] E. M. Wright, For how many edges is a digraph almost certainly Hamiltonian? *Proc. Amer. Math. Soc.* 41 (1973), 384–388.