# THE NUMBERS OF SPANNING TREES, HAMILTON CYCLES AND PERFECT MATCHINGS IN A RANDOM GRAPH

Svante Janson

Abstract. The numbers of spanning trees, Hamilton cycles and perfect matchings in a random graph $G_{nm}$ are shown to be asymptotically normal if $m$ is neither too large nor too small. At the lower limit $m \asymp n^{3/2}$, these numbers are asymptotically log-normal. For $G_{np}$, the numbers are asymptotically log-normal for a wide range of $p$, including $p$ constant.

The same results are obtained for random directed graphs and bipartite graphs.

The results are proved using decomposition and projection methods.

## Introduction and results

The number of small subgraphs of a given kind of a random graph has been studied by many authors. Typical results are that for both standard models $G_{np}$ and $G_{nm}$ of random graphs, for wide ranges of $p$ and $m$, the number of subgraphs isomorphic to a fixed graph is asymptotically normally distributed as $n \to \infty$, see for example [7], [4].

In this paper we will study some examples of large subgraphs. More precisely, we will study three examples of subgraph counts in $G_{np}$ and $G_{nm}$ where the subgraphs have $n$ vertices and $\asymp n$ edges. The results in these cases are rather different from the results for small subgraphs; the asymptotic distribution is still normal for $G_{nm}$ but log-normal for $G_{np}$, provided the edge density is neither too small nor too big. For a smaller edge density, $m \asymp n^{3/2}$, we find asymptotic log-normal distributions also for $G_{nm}$.

In order to state the results smoothly, we let $f(G)$, $g(G)$ and $h(G)$ denote the numbers of spanning subtrees, Hamilton cycles and perfect matchings in a graph $G$. We assume tacitly that $n$ is restricted to be even whenever we consider $h(G_{nm})$ or $h(G_{np})$, since $h(G) = 0$ when the order of $G$ is odd.

For a random variable $X$ (with positive, finite variance) let $X^* = (X - \operatorname{E} X)/(\operatorname{Var} X)^{1/2}$ denote its *standardization*. We write $a \ll b$ when $a$ and $b$ are positive and $a/b \to 0$.

**Theorem 1.** *Assume that $n \to \infty$, $m \gg n^{3/2}$ and $\binom{n}{2} - m \gg n$. Then the standardized variables $f(G_{nm})^*$, $g(G_{nm})^*$ and $h(G_{nm})^*$ converge in distribution to a standard normal distribution.*

*Moreover, with $p = m/\binom{n}{2}$, we have*

$$\operatorname{E} f(G_{nm}) = n^{n-2}p^{n-1} \exp\left(-\frac{1-p}{p} + O\Big((1-p)\frac{n^3}{m^2}\Big)\right), \tag{1.1}$$

$$\operatorname{Var} f(G_{nm}) \sim \frac{n^3}{8m^2}(1-p)^2\big(\operatorname{E} f(G_{nm})\big)^2, \tag{1.2}$$

$$\operatorname{E} g(G_{nm}) = \tfrac{1}{2}(n-1)!\, p^n \exp\left(-\frac{1-p}{p} + O\Big((1-p)\frac{n^3}{m^2}\Big)\right), \tag{1.3}$$

$$\operatorname{Var} g(G_{nm}) \sim \frac{n^3}{2m^2}(1-p)^2\big(\operatorname{E} g(G_{nm})\big)^2, \tag{1.4}$$

$$\mathrm{E}\,h(G_{nm}) = (n-1)!!\,p^{n/2}\exp\Bigl(-\frac{1-p}{4p}+O\Bigl((1-p)\frac{n^3}{m^2}\Bigr)\Bigr), \tag{1.5}$$

$$\operatorname{Var} h(G_{nm}) \sim \frac{n^3}{8m^2}(1-p)^2\bigl(\mathrm{E}\,h(G_{nm})\bigr)^2. \tag{1.6}$$

*If furthermore* $m/\binom{n}{2} \to \pi \geq 0$, *then*

$$\frac{m}{n^{3/2}}\bigl(f(G_{nm})/(n^{n-2}p^{n-1}e^{-(1-p)/p})-1\bigr) \xrightarrow{d} \mathrm{N}\bigl(0,\tfrac18(1-\pi)^2\bigr), \tag{1.7}$$

$$\frac{m}{n^{3/2}}\bigl(g(G_{nm})/(\tfrac12(n-1)!\,p^n e^{-(1-p)/p})-1\bigr) \xrightarrow{d} \mathrm{N}\bigl(0,\tfrac12(1-\pi)^2\bigr), \tag{1.8}$$

$$\frac{m}{n^{3/2}}\bigl(h(G_{nm})/((n-1)!!\,p^{n/2}e^{-(1-p)/4p})-1\bigr) \xrightarrow{d} \mathrm{N}\bigl(0,\tfrac18(1-\pi)^2\bigr). \tag{1.9}$$

We do not know whether the upper bound for $m$, namely $\binom{n}{2}-m \gg n$, is necessary for the conclusions of Theorem 1; in fact, it seems likely that asymptotic normality holds as soon as $\binom{n}{2}-m \gg n^{1/2}$, but we will not investigate this case any further here.

On the other hand, the lower bound $m \gg n^{3/2}$ is indeed necessary, and we have the following result for the limiting case, which shows that the asymptotic distribution then is log-normal. We write $X \sim \mathrm{LN}(\mu,\sigma^2)$ when $\log X \sim \mathrm{N}(\mu,\sigma^2)$.

**Theorem 2.** *Assume that* $n\to\infty$ *and* $m/n^{3/2}\to c>0$. *Then, with* $p=m/\binom{n}{2}$,

$$\mathrm{E}\,f(G_{nm}) \sim n^{n-2}p^{n-1}\exp\bigl(-\frac{1-p}{p}-\frac{1}{6c^2}\bigr), \tag{1.10}$$

$$\operatorname{Var} f(G_{nm}) \sim (e^{1/8c^2}-1)\bigl(\mathrm{E}\,f(G_{nm})\bigr)^2, \tag{1.11}$$

$$\mathrm{E}\,g(G_{nm}) \sim \tfrac12(n-1)!\,p^n\exp\bigl(-\frac{1-p}{p}-\frac{1}{6c^2}\bigr), \tag{1.12}$$

$$\operatorname{Var} g(G_{nm}) \sim (e^{1/2c^2}-1)\bigl(\mathrm{E}\,g(G_{nm})\bigr)^2, \tag{1.13}$$

$$\mathrm{E}\,h(G_{nm}) \sim (n-1)!!\,p^{n/2}\exp\bigl(-\frac{1-p}{4p}-\frac{1}{48c^2}\bigr), \tag{1.14}$$

$$\operatorname{Var} h(G_{nm}) \sim (e^{1/8c^2}-1)\bigl(\mathrm{E}\,h(G_{nm})\bigr)^2, \tag{1.15}$$

*and*

$$f(G_{nm})/\,\mathrm{E}\,f(G_{nm}) \xrightarrow{d} \mathrm{LN}\bigl(-\frac{1}{16c^2},\frac{1}{8c^2}\bigr), \tag{1.16}$$

$$g(G_{nm})/\,\mathrm{E}\,g(G_{nm}) \xrightarrow{d} \mathrm{LN}\bigl(-\frac{1}{4c^2},\frac{1}{2c^2}\bigr), \tag{1.17}$$

$$h(G_{nm})/\,\mathrm{E}\,h(G_{nm}) \xrightarrow{d} \mathrm{LN}\bigl(-\frac{1}{16c^2},\frac{1}{8c^2}\bigr). \tag{1.18}$$

For $G_{np}$ we obtain asymptotic normality only when $p\to1$; for smaller $p$ we have again a log-normal distribution.

**Theorem 3.** *Assume that* $n\to\infty$, $p\to1$ *and* $1-p\gg n^{-2}$. *Then the standardized variables* $f(G_{np})^*$, $g(G_{np})^*$ *and* $h(G_{np})^*$ *converge in distribution to the standard normal* *distribution. Moreover,*

$$\operatorname{E} f(G_{np}) = n^{n-2}p^{n-1}, \tag{1.19}$$

$$\operatorname{Var} f(G_{np}) \sim 2(1-p)\big(\operatorname{E} f(G_{np})\big)^2, \tag{1.20}$$

$$\operatorname{E} g(G_{np}) = \tfrac{1}{2}(n-1)!\,p^n, \tag{1.21}$$

$$\operatorname{Var} g(G_{np}) \sim 2(1-p)\big(\operatorname{E} g(G_{np})\big)^2, \tag{1.22}$$

$$\operatorname{E} h(G_{np}) = (n-1)!!\,p^{n/2}, \tag{1.23}$$

$$\operatorname{Var} h(G_{np}) \sim \tfrac{1}{2}(1-p)\big(\operatorname{E} h(G_{np})\big)^2. \tag{1.24}$$

**Theorem 4.** *Assume that* $n \to \infty$, $p \to \pi < 1$ *and* $\liminf pn^{1/2} > 0$. *Then (1.19), (1.21), (1.23) hold as above, and*

$$p^{1/2}\left(\log f(G_{np}) - \log\big(\operatorname{E} f(G_{np})\big) + \frac{1-p}{p}\right) \xrightarrow{d} \mathrm{N}\big(0, 2(1-\pi)\big), \tag{1.25}$$

$$p^{1/2}\left(\log g(G_{np}) - \log\big(\operatorname{E} g(G_{np})\big) + \frac{1-p}{p}\right) \xrightarrow{d} \mathrm{N}\big(0, 2(1-\pi)\big), \tag{1.26}$$

$$p^{1/2}\left(\log h(G_{np}) - \log\big(\operatorname{E} h(G_{np})\big) + \frac{1-p}{4p}\right) \xrightarrow{d} \mathrm{N}\big(0, \tfrac{1}{2}(1-\pi)\big). \tag{1.27}$$

**Remark 1.1.** The results in Theorem 2 and in Theorem 4 for constant $p$ may be written as $\log X - a \xrightarrow{d} \mathrm{N}(0, \sigma^2)$ without scaling, where $X$ is the random variable under consideration and $a = a(n)$ and $\sigma^2$ are suitable constants. Equivalently, $X/e^a \xrightarrow{d} \mathrm{LN}(0, \sigma^2)$. On the other hand, if $p \to 0$ in Theorem 4, it is necessary to scale $\log X$ to get convergence: $(\log X - a)/b \xrightarrow{d} \mathrm{N}(0, \sigma^2)$, which translates to $(X/e^a)^{1/b} \xrightarrow{d} \mathrm{LN}(0, \sigma^2)$, with $b \to \infty$. Note finally that the asymptotic normality in Theorems 1 and 3 also may be written

$$(\log X - a)/b \xrightarrow{d} \mathrm{N}(0, \sigma^2),$$

where now $b \to 0$.

**Remark 1.2.** Let $p \to 0$, with $p \gg n^{-1/2}$. The distribution of $\log f(G_{np})$ is concentrated at $\log \operatorname{E}(f(G_{np})) - (1-p)/p + O(p^{-1/2})$, which is below $\log \operatorname{E}(f(G_{np}))$. Hence the distribution of $f(G_{np})$ is concentrated way below its expectation; in particular, $f(G_{np})/\operatorname{E} f(G_{np}) \xrightarrow{p} 0$. This may look surprising at first sight, but it is actually a natural consequence of the large tail of a log-normal distribution. For example, by (1.25), the distribution of $f(G_{np})$ is well approximated by $\mathrm{LN}(\log \operatorname{E} f(G_{np}) - \frac{1-p}{p}, 2\frac{1-p}{p})$, which has the same expectation as $f(G_{np})$, but is concentrated at substantially lower values.

**Remark 1.3.** Since (1.25)–(1.27) hold if $pn^{1/2} \to c$, for every fixed $c > 0$, a simple compactness argument shows that they hold also if $pn^{1/2} \to 0$ sufficiently slowly. Similarly, if we rewrite (1.13)–(1.15) as

$$\frac{m}{n^{3/2}}\left(\log f(G_{nm}) - \log\big(\operatorname{E} f(G_{nm})\big) + \frac{n^3}{16m^2}\right) \to \mathrm{N}(0, \tfrac{1}{8}), \tag{1.28}$$

etc., they hold also when $m/n^{3/2} \to 0$ slowly. (By Theorem 1, (1.28) etc. hold also when $m/n^{3/2} \to \infty$ with $m/\binom{n}{2} \to 0$.)

We do not know how small $p$ and $m$ can be for these results to hold; it is possible that the asymptotic log-normality extends all the way down to the thresholds for the variables to be non-zero, which are at $p \sim \log n/n$ and $m \sim \frac{1}{2} n \log n$, see [1].

**Remark 1.4.** It follows from the proofs below that the standardized variables $f(G_{nm})^*$, $g(G_{nm})^*$ and $h(G_{nm})^*$ in Theorem 1 converge jointly to the same normal variable; thus $f(G_{nm})$, $g(G_{nm})$ and $h(G_{nm})$ are approximatively linear functions of each other. Similar results hold for joint convergence in Theorems 2–4.

After an informal discussion in Section 2, we prove generalizations of the results above in Sections 3 and 4, which together with some combinatorial estimates derived in Section 5 prove the theorems above.

In Section 6, finally, we give extensions to random directed graphs and bipartite graphs.

## 2. Some comments and heuristics

If we compare the results above for $G_{nm}$ and $G_{np}$ with $p \approx m/\binom{n}{2}$, we see that the variables vary on a larger scale for $G_{np}$. In other words, the variation in $f(G_{np})$, say, given the actual number of edges $e(G_{np})$, is negligible compared with the variation caused by the fluctutation in the number of edges. Hence $f(G_{np})$ is asymptotically like a function of $e(G_{np})$, which is asymptotically normal. (See the proof of Theorem 6 for details.) One might think that this would yield asymptotic normality for $f(G_{np})$, as it does for small subgraphs, but that is true only when $p \to 1$; in the situation of Theorem 4, it turns out that we get a log-normal distribution because $\mathrm{E}(f(G_{np}) \mid e(G_{np}) = m) = \mathrm{E} f\,(G_{nm})$ grows rapidly with the number of edges $m$, or, equivalently, that $\mathrm{E}\, f(G_{np}) = n^{n-2}p^n$ grows rapidly with $p$.

We have no similar, simple explanation for the emergence of asymptotic log-normal distributions for $G_{nm}$ in Theorem 2, but note that in this case we have $\operatorname{Var} X \sim (\mathrm{E}\, X)^2$ for our variables. This means that we have a natural end-point for the normal phase, since then $X^* \geq -\mathrm{E}\, X/(\operatorname{Var} X)^{1/2}$ is uniformly bounded below, and thus $X^* \xrightarrow{d} \mathrm{N}(0,1)$ cannot hold.

It is instructive to study the asymptotics of, say, $f(G_{np})$ and $f(G_{nm})$ using the decomposition in [4]. Using the notation there, the decomposition may be written as

$$f(G_{np}) = \sum_H \widehat{f}(H;p) S_n(H;p) \tag{2.1}$$

and it is easily seen (cf. the calculations in Section 4 below or [4, Example 12.2]), that $\widehat{f}(H;p)$ vanishes unless $H$ is a forest and that if $H$ is a forest with components of orders $v_1, \dots, v_r$, then $\widehat{f}(H;p) \asymp (\mathrm{E} f\,) \prod (np)^{1-v_i}$ and thus the normalized versions satisft $\widehat{f}^*(H;p) \asymp \mathrm{E} f \prod n^{1-v_i/2} p^{(1-v_i)/2}$.

If $p$ is constant, then $\widehat{f}^*(H;p) \asymp \mathrm{E} f$ for all $H = kK_2$, i.e. when $H$ consists of isolated edges, while all other $\widehat{f}^*(H;p)$ are smaller. It follows that $f(G_{np})$ can be approximated by $\sum_{k=0}^\infty \widehat{f}(kK_2;p) S_n(kK_2;p)$; moreover, $S_n(kK_2;p)$ can be approximated by a Hermite polynomial in $S_n(K_2;p)$ and the sum may be approximated by

$$\sum_0^\infty \mathrm{E} f \,\frac{1}{k!} \Bigl(2\frac{1-p}{p}\Bigr)^{k/2} h_k(S_n(K_2;p)^*) = \mathrm{E} f\, \exp\biggl(\Bigl(2\frac{1-p}{p}\Bigr)^{1/2} S_n(K_2;p)^* - \frac{1-p}{p}\biggr), \tag{2.2}$$

which gives (1.25), cf. [4].

If $p \to 0$, then $f^*(kK_2;p) \asymp \mathrm{E} f\, p^{-k/2}$ is (asymptotically) larger the larger $k$ is. In this case, each single term in the expansion (2.1) is negligible compared with the others, and we do not know how to make the argument above rigorous. It is, nevertheless, tempting to use the same approximations and arrive again at the approximation (2.2), which would imply (1.25) if the error could be controlled.

This is not only an heuristical motivation for the log-normal limits in Theorem 4 (with the proper scaling); it also suggests a method for proving them. Recalling that $S_n(K_2;p)$ and $S_n(K_2;p)^*$ are linear functions of $e(G_{np})$, consider instead of $f(G)$ the modified variable

$$\psi(G) = e^{-ae(G)} f(G), \tag{2.3}$$

where $a$ is a suitable constant (depending on $n$ and $p$) such that the approximation (2.2) of $f(G_{np})$ is equivalent to $\psi(G_{np}) \approx C_n$, for some constants $C_n$. As we shall see later (Remark 4.1), this can be verified by estimating the variance of $\psi(G_{np})$, at least when $p \gg n^{-1/2}$; this is perhaps the simplest proof of (1.25).

Turning to $G_{nm}$, we use the heuristics from [4] that the asymptotics for $G_{nm}$ usually are as for $G_{np}$ with $p = m/\binom{n}{2}$, if we ignore all terms in (2.1) such that $H$ contains an isolated edge. In our case, if $n^{-1/2} \ll p \le 1$, $f^*(P_2;p) \asymp \mathrm{E} f\,(np^2)^{-1/2}$, while all other $f^*(H;p)$, for $H$ such that every component of $H$ has at least three vertices, are smaller. This suggests that $f(G_{nm})$ has the same asymptotics as a linear function of $S_n(P_2;p)$, and thus is asymptotically normal. Moreover, if $pn^{1/2} \to c$, we have $f^*(kP_2;p) \asymp \mathrm{E} f$ for every $k$, and the heuristics suggests an approximating exponential sum similar to (2.2) (but with $S_n(P_2,p)^*$ instead), which would give a log-normal limit as in (1.16). We warm the reader, however, that these conclusions are not completely correct; they happen to give the right qualitivative behaviour of $f(G_{nm})$, but the asymptotic variances they give are wrong, e.g. by a factor 9 for (1.7). (A closer examination shows that the error comes from replacing $\tau$ i Theorem 5 below by $\lim n^2\gamma_2$.)

Again, we may make this argument rigorous (and obtain the correct variances) by considering the modified variable in (2.3). Note that the effect of the modification is quite different for $G_{nm}$ compared to $G_{np}$ which was studied above; since $e(G_{nm}) = m$ is constant, $\psi(G_{nm})$ is just a constant times $f(G_{nm})$ so limit results for one of them trivially transfers to the other. On the other hand, as we will see later, the expansion (2.1) of $\psi(G_{np})$ is dominated by the $P_2$ term (when $p \gg n^{-1/2}$), which gives asymptotic normality of $\psi(G_{np})$. Moreover, the modifying factor in (2.3) cancels essentially the strong dependence of $f(G_{np})$ on the number of edges present, which enables us, by methods of [4], to conclude that $\psi(G_{nm})$ and $\psi(G_{np})$ have the same asymptotics. We will prove Theorems 1 and 2 by this method in Section 4, leaving some combinatorial estimates to Section 5.

The heuristics above suggest, however, a short-cut where Theorem 1 is proved without any of this machinery. Noting that $S_n(P_2;p)$ is a linear function of the number of copies of $P_2$ in $G_{np}$ and $e(G_{np})$, we see that ignoring all terms but the $P_2$ term in (2.1), as suggested above, is equivalent to approximating $\psi(G_{nm})$ by a linear function of the number of copies of $P_2$ in $G_{nm}$. In the next section, we shall prove that this can be done, with a negligible error, by simple moment estimates. This is very similar to the "first projection method" to prove (normal) limits for $G_{np}$, but has to our knowledge not been used before for $G_{nm}$.

We finally remark that these considerations also suggest a method to treat the case $p \ll n^{-1/2}$. The argument above suggests an approximation $\psi(G_{np}) \approx C \exp(bS_n(P_2;p))$, for some constants $b$ and $C$, which conceivably could be proved by computing the variance of the modified modification

$$e^{-aS_n(K_2;p) - bS_n(P_2;p)} f(G_{np})$$

for suitable $a$ and $b$. (For small $p$ one might add more terms in the exponent.) We have not tried this approach.

## 3. A general result

Since most of the argument is the same for the three variables that we consider, we shall state and prove a more general result which will be used to prove Theorem 1.

Suppose that we are given, for each $n$, a set $\widetilde{\mathcal{A}}$ of unlabelled graphs with $\leq n$ vertices, and, for a graph $G$ with $n$ vertices, let $\varphi(G)$ be the number of subgraphs of $G$ that are isomorphic to some member in $\widetilde{\mathcal{A}}$. We assume that all graphs in $\widetilde{\mathcal{A}}$ have the same number $\mu$ of edges. We let $\mathcal{A}$ be the set of subgraphs of $K_n$ (the complete graph on the set of vertices where our random graphs live) that are isomorphic to some member of $\widetilde{\mathcal{A}}$, and let $N = \varphi(K_n)$ be the number of elements of $\mathcal{A}$. (Note that $\widetilde{\mathcal{A}}, \varphi, \mu, \mathcal{A}, N$ as well as $m$, $p$ and other quantities introduced below depend on $n$. A more careful notation would be $\widetilde{\mathcal{A}_n}, \varphi_n, \mu_n, \dots$, but for simplicity we will omit the subscripts.)

The three variables in Theorems 1–4 are evidently examples of such $\varphi$, and we have:

$f$ : $\widetilde{\mathcal{A}} = \{\text{trees on } n \text{ vertices}\}$, $\mu = n-1$, $N = n^{n-2}$;

$g$ : $\widetilde{\mathcal{A}} = \{C_n\}$, $\mu = n$, $N = \frac{1}{2n}n!\quad (n \geq 3)$;

$h$ : $\widetilde{\mathcal{A}} = \{\frac{n}{2}K_2\}$ (a graph consisting of $n/2$ disjoint edges), $\mu = \frac{n}{2}$, $N = (n-1)!!$.

We further define

$$\Lambda(x) = \frac{1}{N^2} \sum_{A_1, A_2 \in \mathcal{A}} (1+x)^{e(A_1 \cap A_2)}, \tag{3.1}$$

where $e(G)$ denotes the number of edges in $G$.

Let $N(H)$ denote the number of elements of $\mathcal{A}$ that contain a given subgraph $H$ of $K_n$ and define $\gamma(H) = N(H)/N$. Then, summing over all subgraphs $H$ of $K_n$ without isolated vertices,

$$\Lambda(x) = \frac{1}{N^2} \sum_{A_1, A_2} \sum_{H \subseteq A_1 \cap A_2} x^{e(H)} = \frac{1}{N^2} \sum_H N(H)^2 x^{e(H)} = \sum_H \gamma(H)^2 x^{e(H)}. \tag{3.2}$$

Since $N(H)$ and $\gamma(H)$ depend on $H$ only up to isomorphism, (3.2) yields, for small $x$ and fixed $n$,

$$\Lambda(x) = 1 + \binom{n}{2}\gamma(K_2)^2 x + \left(\frac{(n)_3}{2}\gamma(P_2)^2 + \frac{(n)_4}{8}\gamma(2K_2)^2\right)x^2 + O(x^3), \tag{3.3}$$

where $K_2$ is an edge, $P_2$ a path of length 2 and $2K_2$ consists of two independent edges. Thus, using the shorthand $\gamma_1 = \gamma(K_2)$, $\gamma_2 = \gamma(P_2)$ and $\gamma_3 = \gamma(2K_2)$,

$$\log \Lambda(x) = 1 + \binom{n}{2}\gamma_1^2 x + \left(\frac{(n)_3}{2}\gamma_2^2 + \frac{(n)_4}{8}\gamma_3^2 - \frac{1}{2}\binom{n}{2}^2\gamma_1^4\right)x^2 + O(x^3). \tag{3.4}$$

By counting the number of edges and pairs of edges in element of $\mathcal{A}$, we have $N\mu = \binom{n}{2}N(K_2)$, and $N\frac{\mu(\mu-1)}{2} = \frac{(n)_3}{2}N(P_2) + \frac{(n)_4}{8}N(2K_2)$, and thus

$$\gamma_1 = \mu/\tbinom{n}{2}, \tag{3.5}$$

$$\gamma_3 = \frac{4}{(n)_4}(\mu(\mu-1) - (n)_3\gamma_2). \tag{3.6}$$

Substitution of this into (3.4) yields, after simplifications, the following.

**Lemma 3.1.** *With the notations above,* $\log \Lambda(x)$ *has the Taylor expansion*

$$\log \Lambda(x) = \lambda_1 x + \lambda_2 x^2 + O(x^3), \quad |x| < 1, \tag{3.7}$$

*with*

$$\lambda_1 = \frac{\mu^2}{\binom{n}{2}} \tag{3.8}$$

$$\begin{aligned} \lambda_2 &= \frac{n+1}{n-3} \frac{(n)_3}{2} \gamma_2^2 - \frac{4}{n-3} \mu(\mu-1)\gamma_2 + \frac{2\mu^2}{n(n-1)} \left( \frac{(\mu-1)^2}{(n-2)(n-3)} - \frac{\mu^2}{n(n-1)} \right) \\ &= \frac{n^3}{2} (\gamma_2 - \gamma_1^2)^2 - \frac{\mu^3}{\binom{n}{2}^2} + O(n^2\gamma_2^2 + \frac{\mu^2}{n^4} + \frac{\mu^4}{n^6}). \end{aligned} \tag{3.9}$$

□

After these preliminaries, we state our result.

**Theorem 5.** *With the notations above, suppose that* $n \to \infty$ *and* $\mu = O(n)$*, that* $n^2(\gamma_2 - \gamma_1^2) \to \tau \in (-\infty, \infty)$ *(or, equivalently,* $n^2 \frac{N(P_2)}{N} - \frac{4\mu^2}{n^2} \to \tau$*), and that for every sequence* $x = x_n$ *with* $n^{-1} \ll x \ll n^{1/2}$*,*

$$\Lambda(x) \le \exp\left( \lambda_1 x + \lambda_2 x^2 + o\left(\frac{x^2}{n}\right) \right). \tag{3.10}$$

*If* $m \gg n^{3/2}$ *and* $\binom{n}{2} - m \gg n$*, then, with* $p = m/\binom{n}{2}$*,*

$$\mathrm{E}\,\varphi(G_{nm}) = Np^{\mu} \exp\left( -\frac{\mu^2}{2m}(1-p) + O\left((1-p)\frac{n^3}{m^2}\right) \right), \tag{3.11}$$

$$\mathrm{Var}\,\varphi(G_{nm}) = \left( \frac{\tau^2}{8} + o(1) \right) (1-p)^2 \frac{n^3}{m^2} \left( \mathrm{E}\,\varphi(G_{nm}) \right)^2 \tag{3.12}$$

*and*

$$(1-p)^{-1} \frac{m}{n^{3/2}} \left( \varphi(G_{nm}) \Big/ Np^{\mu} \exp\left( -\frac{\mu^2}{2m}(1-p) \right) - 1 \right) \xrightarrow{d} \mathrm{N}\left(0, \frac{\tau^2}{8}\right). \tag{3.13}$$

*If furthermore* $\tau \neq 0$*, the standardized variable* $\varphi(G_{nm})^*$ *converges in distribution to the standard normal distribution.*

*Proof.* Let $X = \varphi(G_{nm})$ and let $Y$ denote the number of copies of $P_2$ in $G_{nm}$. We shall prove (3.11) and

$$\mathrm{Var}\,X \le \left( \frac{\tau^2}{2} + o(1) \right) \frac{(1-p)^2}{np^2} (\mathrm{E}X)^2, \tag{3.14}$$

$$\mathrm{Cov}(X, Y) = (\tau + o(1)) \frac{(1-p)^2}{n^2p^2} \,\mathrm{E}X \,\mathrm{E}\,Y, \tag{3.15}$$

$$\mathrm{Var}\,Y = (2 + o(1)) \frac{(1-p)^2}{n^3p^2} (\mathrm{E}\,Y)^2. \tag{3.16}$$

The result then follows easily. First we must have equality in (3.14) by (3.15), (3.16) and the Cauchy–Schwarz inequality; this is equivalent to (3.12). If $\tau \neq 0$, then (3.14)–(3.16) yield

$$\mathrm{E}(X^* - \mathrm{sign}(\tau)Y^*)^2 = \mathrm{Var}(X^* - \mathrm{sign}(\tau)Y^*) \to 0. \tag{3.17}$$

(This is another way of expressing the asymptotic equality in the Cauchy–Schwarz inequality.) Since $Y^* \xrightarrow{d} \mathrm{N}(0,1)$ by [4,Theorem 19], $X^* \xrightarrow{d} \mathrm{N}(0,1)$ now follows. Finally, this, (3.11) and (3.12) yield (3.13). The case $\tau = 0$ is simpler, with (3.12) and (3.13) following directly from (3.14).

Hence we only have to prove the moment estimates. In order to do so, we use the well-known estimate that, for $0 \le l \le k$ we have

$$\begin{aligned}(k)_l &= k^l \exp \sum_0^{l-1} \log(1 - \frac{i}{k}) \\ &= k^l \exp\Big(-\frac{l(l-1)}{2k} - \frac{2l^3 - 3l^2 + l}{12k^2} - \frac{l^4}{12k^3} - \frac{l^5}{20k^4} + O\Big(\frac{l^3}{k^3} + \frac{l^6}{k^5}\Big)\Big). \end{aligned} \tag{3.18}$$

Thus, using that $\mu = O(n)$, $p \gg n^{-1/2}$ and $1 - p \gg n^{-1}$,

$$\mathrm{E}\, X = N\frac{(m)_\mu}{\binom{n}{2}_\mu} = Np^\mu \exp\Big(-\frac{\mu(\mu-1)}{2m}(1-p) - \frac{\mu^3}{6m^2}(1-p^2) + O\Big(\frac{1-p}{n^2p^3}\Big)\Big); \tag{3.19}$$

in particular,

$$\mathrm{E}\, X = Np^\mu \exp\Big(-\frac{\mu^2}{2m}(1-p) + O\Big(\frac{1-p}{np^2}\Big)\Big), \tag{3.20}$$

which is (3.11).

Moreover,

$$\mathrm{E}\, X^2 = \sum_{A_1, A_2 \in \mathcal{A}} \frac{(m)_{e(A_1 \cup A_2)}}{(\binom{n}{2})_{e(A_1 \cup A_2)}} = \sum_{A_1, A_2 \in \mathcal{A}} \frac{(m)_{2\mu - e(A_1 \cap A_2)}}{(\binom{n}{2})_{2\mu - e(A_1 \cap A_2)}}, \tag{3.21}$$

and, for $0 \le e \le \mu$, using (3.18) as for (3.19),

$$\begin{aligned}
&\frac{(m)_{2\mu-e}}{(\binom{n}{2})_{2\mu-e}} \\
&= p^{2\mu-e} \exp\Big(-\frac{(2\mu-e)(2\mu-e-1)}{2m}(1-p) - \frac{(2\mu-e)^3}{6m^2}(1-p^2) + O(\frac{1-p}{n^2p^3})\Big) \\
&= p^{2\mu-e} \exp\Big(-\mu(2\mu-1)\frac{1-p}{m} + e(2\mu - \tfrac12)\frac{1-p}{m} - \frac{e^2}{2}\frac{1-p}{m} - \frac{8\mu^3}{6m^2}(1-p^2) \\
&\qquad + 2e\frac{\mu^2}{m^2}(1-p^2) + O\Big(\frac{\mu}{m^2}(1-p)e^2\Big) + O\Big(\frac{1-p}{n^2p^3}\Big)\Big) \\
&= p^{2\mu} \exp\Big(-\mu(2\mu-1)\frac{1-p}{m} - \frac{4\mu^3}{3m^2}(1-p^2)\Big) y^e \Big(1 + O\Big(e^2\frac{1-p}{n^2p}\Big) + O\Big(\frac{1-p}{n^2p^3}\Big)\Big),
\end{aligned} \tag{3.22}$$

where

$$y = p^{-1} \exp\Big((2\mu - \tfrac12)\frac{1-p}{m} + 2\frac{\mu^2}{m^2}(1-p^2)\Big) = p^{-1}\Big(1 + 2\mu\frac{1-p}{m} + O\Big(\frac{1-p}{n^2p^2}\Big)\Big). \tag{3.23}$$

Since $e^2 = O\big(\frac{1}{p^2}(1+p)^e\big)$, we obtain from (3.21), (3.22), (3.19) and (3.1), that

$$\begin{aligned}
&\mathrm{E}\, X^2 \\
&= N^2p^{2\mu} \exp\Big(-\mu(2\mu-1)\frac{1-p}{m} - \frac{4\mu^3}{3m^2}(1-p^2)\Big)\Big(\Lambda(y-1) + O\Big(\frac{1-p}{n^2p^3}\Lambda((1+p)y-1)\Big)\Big) \\
&= (\mathrm{E}\, X)^2 \exp\Big(-\mu^2\frac{1-p}{m} - \frac{\mu^3}{m^2}(1-p^2)\Big)\Big(\Lambda(y-1) + O\Big(\frac{1-p}{n^2p^3}\Lambda\big(y + O(1)\big)\Big)\Big)
\end{aligned} \tag{3.24}$$

Hence, using (3.10) and Lemma 3.1, and noting that our assumptions yield $\lambda_1 = O(1)$, $\lambda_2 = O(n^{-1})$ and $y - 1 \sim \frac{1-p}{p}$,

$$\begin{aligned}
\frac{\mathrm{E}\,X^2}{(\mathrm{E}\,X)^2} &\leq \Bigl(1 + O\Bigl(\frac{1-p}{n^2p^3}\Bigr)\Bigr) \exp\Bigl(-\mu^2\frac{1-p}{m} - \frac{\mu^3}{m^2}(1-p^2) + \lambda_1(y-1) \\
&\qquad\qquad + \lambda_2(y-1)^2 + o\Bigl(\frac{(1-p)^2}{np^2}\Bigr)\Bigr) \\
&= \exp\Bigl(-\mu^2\frac{1-p}{m} + \lambda_1\Bigl(\frac{1-p}{p} + \frac{2\mu}{m}\frac{1-p}{p}\Bigr) - \frac{\mu^3}{m^2}(1-p^2) \\
&\qquad\qquad + \lambda_2\Bigl(\frac{1-p}{p}\Bigr)^2 + o\Bigl(\frac{(1-p)^2}{np^2}\Bigr)\Bigr) \\
&= \exp\Bigl(\frac{n^3}{2}(\gamma_2 - \gamma_1^2)^2\Bigl(\frac{1-p}{p}\Bigr)^2 + o\Bigl(\frac{(1-p)^2}{np^2}\Bigr)\Bigr) \\
&= 1 + \Bigl(\frac{\tau^2}{2} + o(1)\Bigr)\frac{(1-p)^2}{np^2},
\end{aligned} \tag{3.25}$$

which yields (3.14).

Let $\mathcal{B}$ be the set of the $N_Y = \frac{1}{2}(n)_3$ copies of $P_2$ in $K_n$. Then

$$\mathrm{E}\,Y = N_Y\frac{(m)_2}{(\binom{n}{2})_2} = N_Y p^2 \frac{1-1/m}{1-1/\binom{n}{2}}, \tag{3.26}$$

$$\mathrm{E}\,XY = \sum_{\substack{A\in\mathcal{A}\\ B\in\mathcal{B}}} \frac{(m)_{e(A\cup B)}}{(\binom{n}{2})_{e(A\cup B)}}. \tag{3.27}$$

Now there are $N(P_2) = \gamma_2 N$ elements of $A$ containing a given element of $\mathcal{B}$, and thus $\gamma_2 N N_Y$ pairs $(A,B)$ with $B \subset A$ and $e(A\cup B) = \mu$. Further, given $B \in \mathcal{B}$, there are $N(K_2) = \gamma_1 N$ elements of $\mathcal{A}$ containing a given edge in $B$, and thus $\gamma_1 N - \gamma_2 N$ elements of $\mathcal{A}$ whose intersection with $B$ equals that edge. This gives $2(\gamma_1 - \gamma_2)NN_Y$ pairs $(A,B)$ with $e(A\cap B) = 1$ and $e(A\cup B) = \mu+1$. There remain $(1 - 2\gamma_1 + \gamma_2)NN_Y$ pairs $(A,B)$ with $e(A\cap B) = 0$ and $e(A\cup B) = \mu+2$. Hence

$$\begin{aligned}
\mathrm{E}\,XY &= NN_Y\Bigl((1-2\gamma_1+\gamma_2)\frac{(m)_{\mu+2}}{(\binom{n}{2})_{\mu+2}} + 2(\gamma_1-\gamma_2)\frac{(m)_{\mu+1}}{(\binom{n}{2})_{\mu+1}} + \gamma_2\frac{(m)_\mu}{(\binom{n}{2})_\mu}\Bigr) \\
&= \mathrm{E}X N_Y\Bigl((1-2\gamma_1+\gamma_2)\frac{(m-\mu)(m-\mu-1)}{(\binom{n}{2}-\mu)(\binom{n}{2}-\mu-1)} + 2(\gamma_1-\gamma_2)\frac{m-\mu}{\binom{n}{2}-\mu} + \gamma_2\Bigr) \\
&= \mathrm{E}X\,\mathrm{E}Y\Bigl(\frac{1-1/m}{1-1/\binom{n}{2}}\Bigr)^{-1}\Bigl((1-2\gamma_1+\gamma_2)\frac{(1-\mu/m)(1-(\mu+1)/m)}{(1-\mu/\binom{n}{2})(1-(\mu+1)/\binom{n}{2})} \\
&\qquad\qquad + 2(\gamma_1-\gamma_2)p^{-1}\frac{1-\mu/m}{1-\mu/\binom{n}{2}} + \gamma_2 p^{-2}\Bigr).
\end{aligned} \tag{3.28}$$

We use the expansion

$$\frac{1-\nu/m}{1-\nu/\binom{n}{2}} = 1 - \frac{(1-p)\nu/m}{1-\nu/\binom{n}{2}} = 1 - \nu\frac{1-p}{m} - \nu^2\frac{p(1-p)}{m^2} + O\Bigl(\frac{\nu^3(1-p)}{n^6 p}\Bigr),$$

valid for $\nu = O(n)$, and the relations $\gamma_1 = \mu/\binom{n}{2} = p\mu/m$, $\gamma_2 = O(n^{-2} + \gamma_1^2) = O(n^{-2})$, and obtain after a straightforward but lengthy calculation, that

$$\frac{\operatorname{Cov}(X,Y)}{\mathrm{E}X\,\mathrm{E}Y} = \frac{\mathrm{E}\,XY}{\mathrm{E}X\,\mathrm{E}Y} - 1 = (\gamma_2 - \gamma_1^2)\Big(\frac{1-p}{p}\Big)^2 + O\Big(\frac{\mu(1-p)}{n^4p^3}\Big), \tag{3.29}$$

which yields (3.15).

Finally, we note that $\mathrm{E}\,Y^2$ is given by the same formulas as $\mathrm{E}\,XY$, if we replace $\mathcal{A}$ by $\mathcal{B}$, $N$ by $N_Y$, $\mu$ by 2, $\gamma_1$ by $2/\binom{n}{2}$ and $\gamma_2$ by $1/N_Y = 2/(n)_3$. With these substitutions, (3.29) becomes

$$\frac{\operatorname{Var} Y}{(\mathrm{E}\,Y)^2} = \frac{\mathrm{E}\,Y^2}{(\mathrm{E}\,Y)^2} - 1 = \frac{2}{n^3}\Big(\frac{1-p}{p}\Big)^2 + O\Big(\frac{1-p}{n^4p^3}\Big), \tag{3.30}$$

which yields (3.16) and completes the proof. □

**Remark 3.1.** It follows from the proof that equality holds in (3.10). It is, however, convenient to have to verify only the inequality.

**Remark 3.2.** The proof yields, implicitly, the approximation

$$n^{1/2}p(1-p)^{-1}(\mathrm{E}X)^{-1}(X - \mathrm{E}X) \approx \tfrac{1}{\sqrt{2}}\tau Y^*, \tag{3.31}$$

where the difference between the two sides tends to 0 in probability (and in $L^2$) as $n \to \infty$. Here $Y$ may be replaced by the sum of the square of the degrees of the vertices.

It is likely that this method can be used also in the case $m \asymp n^{3/2}$ to prove a generalization of Theorem 2, but we have not attempted this and will instead use another method in the next section.

We now turn briefly to $G_{np}$, and obtain the following general version of Theorem 4 as a corollary of Theorem 5. A special case was given as [3, Theorem 6].

**Theorem 6.** *With assumptions as in Theorem 5, suppose further that $\mu/n \to \kappa \ge 0$. If $p \to \pi < 1$ and $p \gg n^{-1/2}$, then*

$$p^{1/2}\Big(\log\varphi(G_{np}) - \log \mathrm{E}\,\varphi(G_{np}) + \frac{\mu^2(1-p)}{n^2p}\Big) \xrightarrow{d} \mathrm{N}\big(0, 2\kappa^2(1-\pi)\big), \tag{3.32}$$

*with* $\mathrm{E}\,\varphi(G_{np}) = Np^\mu$.

*Proof.* Note that (3.13) implies

$$\varphi(G_{nm})\Big/N\Big(\frac{m}{\binom{n}{2}}\Big)^\mu \exp\Big(-\frac{1}{2}\mu^2\Big(\frac{1}{m} - \frac{1}{\binom{n}{2}}\Big)\Big) \xrightarrow{p} 1 \tag{3.33}$$

and thus

$$\log\varphi(G_{nm}) - \Big(\log N + \mu\log\frac{m}{\binom{n}{2}} - \frac{1}{2}\mu^2\Big(\frac{1}{m} - \frac{1}{\binom{n}{2}}\Big)\Big) \xrightarrow{p} 0. \tag{3.34}$$

Let $M = e(G_{np})$ be the number of edges in $G_{np}$. Then $M/n^{3/2} \xrightarrow{p} \infty$ and $(\binom{n}{2} - M)/n \xrightarrow{p} \infty$, and it follows by conditioning on $M$ that

$$\log\varphi(G_{np}) - \Big(\log N + \mu\log\frac{M}{\binom{n}{2}} - \frac{1}{2}\mu^2\Big(\frac{1}{M} - \frac{1}{\binom{n}{2}}\Big)\Big) \xrightarrow{p} 0. \tag{3.35}$$

Moreover, $M \sim \mathrm{Bi}(\binom{n}{2}, p)$ and thus $M^* \xrightarrow{d} \mathrm{N}(0,1)$, which implies

$$\log \frac{M}{\binom{n}{2}} = \log p + \log\Bigl(1 + \Bigl(\frac{1-p}{\binom{n}{2}p}\Bigr)^{1/2} M^*\Bigr) = \log p + \Bigl(\frac{1-p}{\binom{n}{2}p}\Bigr)^{1/2} M^* + O_p\Bigl(\frac{1-p}{n^2p}\Bigr) \quad (3.36)$$

and $\mu^2/M - \mu^2/\binom{n}{2}p \xrightarrow{p} 0$. Now (3.32) follows by Cramér's theorem. The formula for $\mathrm{E}\,\varphi(G_{np})$ is evident (for any $p$). □

The same proof yields asymptotic normality of $\varphi(G_{np})$ when $p \to 1$ with $1 - p \gg n^{-1}$ and $\kappa > 0$, but the following simpler proof, using the first projection method, yields a more general result.

**Theorem 7.** *Suppose that $n \to \infty$, $\mu/n \to \kappa > 0$, $p \to 1$ and $1 - p \gg n^{-2}$, and that if $x = \frac{1-p}{p}$, then*

$$\Lambda(x) \le \exp\bigl(\lambda_1 x + o(x)\bigr). \quad (3.37)$$

*Then the standardized variable $\varphi(G_{np})^*$ tends in distribution to $\mathrm{N}(0,1)$, with $\mathrm{E}\,\varphi(G_{np}) = Np^{\mu}$ and*

$$\operatorname{Var}\varphi(G_{np}) \sim 2\kappa^2(1-p)\bigl(\mathrm{E}\,\varphi(G_{np})\bigr)^2. \quad (3.38)$$

*Proof.* As in the proof of Theorem 5, it suffices to prove that if $X = \varphi(G_{np})$ and $Y = e(G_{np}) \sim \mathrm{Bi}(\binom{n}{2}, p)$, then

$$\operatorname{Var} X \le (2\kappa^2 + o(1))(1-p)(\mathrm{E}\,X)^2, \quad (3.39)$$

$$\operatorname{Cov}(X, Y) = \frac{2\kappa + o(1)}{n}(1-p)\,\mathrm{E}\,X\,\mathrm{E}\,Y, \quad (3.40)$$

$$\operatorname{Var} Y = \frac{2 + o(1)}{n^2}(1-p)(\mathrm{E}\,Y)^2. \quad (3.41)$$

These are easily verified. First, by (3.37),

$$\begin{aligned} \mathrm{E}\,X^2 &= \sum_{A_1, A_2 \in \mathcal{A}} p^{2\mu - e(A_1 \cap A_2)} = N^2 p^{2\mu} \Lambda\Bigl(\frac{1}{p} - 1\Bigr) \\ &\le (\mathrm{E}\,X)^2\Bigl(1 + \lambda_1 \frac{1-p}{p} + o(1-p)\Bigr), \end{aligned}$$

which gives (3.39) since $\lambda_1 = \mu^2/\binom{n}{2} \sim 2\kappa^2$ by Lemma 3.1. Similarly,

$$\mathrm{E}\,XY = N\mu p^{\mu} + N\left(\tbinom{n}{2} - \mu\right)p^{\mu+1} = \mathrm{E}\,X\left(\mu(1-p) + \tbinom{n}{2}p\right),$$

and thus

$$\frac{\operatorname{Cov}(X, Y)}{\mathrm{E}X\,\mathrm{E}\,Y} = \frac{\mu(1-p)}{\mathrm{E}\,Y} \sim \frac{2\kappa n(1-p)}{n^2 p},$$

which is (3.40). Finally, $\operatorname{Var} Y = \binom{n}{2}p(1-p)$, which implies (3.41). □

**Remark 3.3.** In Theorem 7, we only have to asssume an estimate of $\Lambda(x)$ at $x = 1/p - 1$. Similarly, in Theorems 5 and 6 it suffices that (3.10) holds for $x \asymp \binom{n}{2}/m - 1$ and $x \asymp 1/p - 1$, respectively.

## 4. A second proof, and the case $m \asymp n^{3/2}$

In this section we use the decomposition methods of [4] to give a second proof of Theorem 5, at the same time proving the following result for the limiting case.

**Theorem 8.** *Suppose that the conditions of Theorem 5 are fulfilled, with (3.10) holding for $x \asymp n^{1/2}$, that $\mu/n \to \kappa$ and that*

$$\gamma(iP_2 + jK_2) = \bigl(1 + o(1)\bigr)\gamma(P_2)^i \gamma(K_2)^j \tag{4.1}$$

*for any $i, j \geq 0$.*

(i) *If $m/n^{3/2} \to c > 0$, then*

$$\operatorname{E}\varphi(G_{nm}) \sim N\Bigl(\frac{m}{\binom{n}{2}}\Bigr)^{\mu} \exp\Bigl(-\frac{\mu^2}{2m} + \kappa^2 - \frac{\kappa^3}{6c^2}\Bigr), \tag{4.2}$$

$$\operatorname{Var}\varphi(G_{nm}) \sim (e^{\tau^2/8c^2} - 1)\bigl(\operatorname{E}\varphi(G_{nm})\bigr)^2, \tag{4.3}$$

$$\varphi(G_{nm})/\operatorname{E}\varphi(G_{nm}) \xrightarrow{d} \operatorname{LN}\Bigl(-\frac{\tau^2}{16c^2}, \frac{\tau^2}{8c^2}\Bigr). \tag{4.4}$$

(ii) *If $pn^{1/2} \to c > 0$, then*

$$p^{1/2}\biggl(\log\varphi(G_{np}) - \log\operatorname{E}\varphi(G_{np}) + \frac{\mu^2}{n^2 p}\biggr) \xrightarrow{d} \operatorname{N}(0, 2\kappa^2). \tag{4.5}$$

*Proof of Theorems 5 (again) and 8.*
We first observe that (4.2) follows by (3.19), and that part (ii) of Theorem 8 follows from part (i) by the proof of Theorem 6 with only minor modifications. Moreover, the argument in (3.24)–(3.25) yields

$$\frac{\operatorname{Var}\varphi(G_{nm})}{(\operatorname{E}\varphi(G_{nm}))^2} = \frac{\operatorname{E}(\varphi(G_{nm}))^2}{(\operatorname{E}\varphi(G_{nm}))^2} - 1 \leq \exp\Bigl(\frac{\tau^2}{8c^2} + o(1)\Bigr) - 1,$$

while (4.4) and Fatou's lemma yield, with $Z \sim \operatorname{LN}(-\frac{\tau^2}{16c^2}, \frac{\tau^2}{8c^2})$,

$$\liminf \frac{\operatorname{Var}\varphi(G_{nm})}{(\operatorname{E}\varphi(G_{nm}))^2} \geq \operatorname{Var}(Z) = \exp(\frac{\tau^2}{8c^2}) - 1,$$

and (4.3) follows. Thus we only have to prove (3.13) and (4.4). We let, as usual, $p = m/\binom{n}{2}$, and restrict ourselves, for simplicity, to the case $p \to \pi$ with $0 \leq \pi < 1$.

As explained in Section 2, we will replace $\varphi$ by another function, which on $G_{nm}$ differs from $\varphi$ by only a constant factor, before applying the decomposition. It is convenient to change (2.2) a little, and we define, for any graph $G$ with $n$ vertices,

$$\psi(G) = \varphi(G)(1 - a)^{e(G) - \mu}, \tag{4.6}$$

for some constant $a = a_n \in (0, 1)$ to be chosen later.

Our method requires us to compare $G_{nm}$ not only with a fixed $G_{np}$, but also to vary the edge probability. We therefore define a random graph process $G_n(t)$ as follows. Let, for each edge $e \in K_n$, $T_e$ be independent, identically distributed random variables with a uniform distribution on (0,1), and let $G_n(t)$ be the subgraph of $K_n$ with edge set $\{e : T_e \le t\}$. Thus $G_n(p) \cong G_{np}$. We also define

$$I_e(t) = I(e \in G_n(t)) = I(T_e \le t).$$

After these preliminaries, we begin by computing the expectation and variance of $X(t) = \psi(G_n(t))$.

Since

$$X(t) = \psi(G_n(t)) = \sum_{A\in\mathcal{A}} \prod_{e\in A} I_e(t) \prod_{e\notin A} (1 - aI_e(t)) = \sum_{A\in\mathcal{A}} Y_A(t), \tag{4.7}$$

say, we have

$$\mathrm{E}\, X(t) = \sum_{A\in\mathcal{A}} \mathrm{E}\, Y_A(t) = Nt^{\mu}(1 - at)^{\binom{n}{2}-\mu} \tag{4.8}$$

and, using $\mathrm{E}\, I_e(t) = t$, $\mathrm{E}\big(I_e(t)(1-aI_e(t))\big) = t(1-a)$ and $\mathrm{E}(1-aI_e(t))^2 = 1-t+t(1-a)^2$,

$$\begin{aligned}
\mathrm{E}\, X(t)^2 &= \sum_{A_1,A_2\in\mathcal{A}} \mathrm{E}\, Y_{A_1}(t) Y_{A_2}(t) \\
&= \sum_{A_1,A_2} t^{e(A_1\cap A_2)}(t(1-a))^{2(\mu - e(A_1\cap A_2))}(1-t+t(1-a)^2)^{\binom{n}{2}-2\mu+e(A_1\cap A_2)} \\
&= t^{2\mu}(1-a)^{2\mu}(1-t+t(1-a)^2)^{\binom{n}{2}-2\mu} N^2 \Lambda\Big(\frac{t(1-t+t(1-a)^2)}{t^2(1-a)^2} - 1\Big) \\
&= (\mathrm{E}\, X(t))^2 (1-a)^{2\mu}\Big(\frac{1-2at+a^2t}{(1-at)^2}\Big)^{\binom{n}{2}-2\mu}(1-at)^{-2\mu}\Lambda\left(\frac{1-t}{t(1-a)^2}\right).
\end{aligned} \tag{4.9}$$

Consequently we have, assuming $a < 1/2$ and letting

$$x = \frac{1-t}{t(1-a)^2} = O\Big(\frac{1}{t}\Big) \quad \text{and} \quad \lambda(x) = \log \Lambda(x),$$

$$\begin{aligned}
&\log \frac{\mathrm{E}\, X(t)^2}{(\mathrm{E}\, X(t))^2} \\
&= 2\mu\big(\log(1-a) - \log(1-at)\big) + \big(\tbinom{n}{2} - 2\mu\big)\log\Big(1 + \frac{a^2(t-t^2)}{(1-at)^2}\Big) + \lambda(x) \\
&= -2\mu(a - at) - \mu\big(a^2 - (at)^2\big) + O(\mu a^3) + \big(\tbinom{n}{2} - 2\mu\big)\Big(\frac{a^2t(1-t)}{(1-at)^2} + O\big((a^2t)^2\big)\Big) + \lambda(x) \\
&= -2\mu a(1-t) + \tbinom{n}{2}a^2t(1-t) + 2\tbinom{n}{2}a^3t^2(1-t) - \mu a^2\big(1 - t^2 + 2t(1-t)\big) \\
&\qquad\qquad + O(\mu a^3 + n^2a^4t^2) + \lambda(x) \\
&= (1+2at)\big(\tbinom{n}{2}a^2t(1-t) - 2\mu a(1-t)\big) - \mu a^2(1-t)^2 + O(\mu a^3 + n^2a^4t^2) + \lambda(x) \\
&= (1+2at)\frac{1-t}{\binom{n}{2}t}\big(\tbinom{n}{2}at - \mu\big)^2 - (1+2at)\frac{1-t}{\binom{n}{2}t}\mu^2 - \mu a^2(1-t)^2 + O(\mu a^3 + n^2a^4t^2) + \lambda(x) \\
&= -(1+2at)(1-a)^2\frac{\mu^2}{\binom{n}{2}}x - \mu a^2(1-t)^2 + O\Big(\frac{1}{\binom{n}{2}t}\big(\tbinom{n}{2}at - \mu\big)^2 + \mu a^3 + n^2a^4t^2\Big) + \lambda(x).
\end{aligned} \tag{4.10}$$

We choose

$$a = \frac{\mu}{m} = \frac{\mu}{\binom{n}{2}p} \tag{4.11}$$

and consider in the sequel of the proof only $t$ with

$$t = p + O(p^{1/2}/n); \tag{4.12}$$

this implies that $t \sim p$ and $\binom{n}{2}at - \mu = \mu(\frac{t}{p} - 1) = O(\frac{\mu}{np^{1/2}}) = O(\frac{\mu}{m^{1/2}})$. For such $t$, (4.10) yields after simplifications

$$\log \frac{\mathrm{E}\, X(t)^2}{(\mathrm{E}\, X(t))^2} = \lambda(x) - \frac{\mu^2}{\binom{n}{2}}x + \frac{\mu^3}{\binom{n}{2}^2}x^2 + O\Big(\frac{\mu^2}{n^4t^2} + \frac{\mu^4}{n^6t^3}\Big). \tag{4.13}$$

We now use our assumption (3.10) and the relations (3.8) and (3.9), noting that the latter gives $\lambda_2 = \frac{1}{2n}\tau^2 - \frac{\mu^3}{\binom{n}{2}^2} + o(\frac{1}{n})$, and obtain,

$$\log \frac{\mathrm{E}\, X(t)^2}{(\mathrm{E}\, X(t))^2} \le \frac{\tau^2}{2n}x^2 + o\Big(\frac{x^2}{n}\Big) + O\Big(\frac{1}{n^2t^3}\Big). \tag{4.14}$$

Since $x \asymp 1/t \asymp 1/p$, (4.14) implies if $p \gg n^{-1/2}$,

$$\frac{\operatorname{Var} X(t)}{(\mathrm{E}\, X(t))^2} = \frac{\mathrm{E}\, X(t)^2}{(\mathrm{E}\, X(t))^2} - 1 \le \tau^2\frac{x^2}{2n} + o\Big(\frac{x^2}{n}\Big) = (\tau^2 + o(1))\frac{(1-p)^2}{2np^2}, \tag{4.15}$$

and if $m/n^{3/2} \to c > 0$ and thus $p \sim 2cn^{-1/2}$,

$$\frac{\operatorname{Var} X(t)}{(\mathrm{E}\, X(t))^2} \le \exp(\frac{\tau^2}{8c^2}) - 1 + o(1). \tag{4.16}$$

We next consider the orthogonal decomposition

$$\psi(G(t)) = \sum_H \widehat{\psi}(H;t)S_n(H;t) \tag{4.17}$$

studied in [4]. Here $H$ ranges over all unlabelled graphs without isolated vertices and $S_n(H;t) = \sum_{H_1}\prod_{e\in H_1}(I_e(t) - t)$, with the summation over the subgraphs $H_1$ of $K_n$ with $H_1 \cong H$ (each repeated $\operatorname{aut}(H)$ times).

We may assume that $H$ is a subgraph of $K_n$. Then, using (4.7),

$$\begin{aligned}\widehat{\psi}(H;t) &= \frac{\mathrm{E}\,\psi(G(t))S_n(H;t)}{\mathrm{E}(S_n(H;t))^2} = \frac{\mathrm{E}\big(\psi(G(t))\prod_{e\in H}(I_e - t)\big)}{\operatorname{aut}(H)(t(1-t))^{e(H)}} = \\ &= \frac{1}{\operatorname{aut}(H)}t^{-e(H)}(1-t)^{-e(H)}\sum_A \mathrm{E}\Big(Y_A(t)\prod_{e\in H}(I_e - t)\Big),\end{aligned} \tag{4.18}$$

and, summing over $F$ without isolated vertices,

$$\begin{aligned}\mathrm{E}\Big(Y_A(t)\prod_{e\in H}(I_e - t)\Big) &= \mathrm{E}\, Y_A(t)\left(\frac{\mathrm{E}\, I_e(I_e - t)}{\mathrm{E}\, I_e}\right)^{e(H\cap A)}\left(\frac{\mathrm{E}(1 - aI_e)(I_e - t))}{E(1 - aI_e)}\right)^{e(H\setminus A)} \\ &= \mathrm{E}\, Y_A(t)(1-t)^{e(H\cap A)}\Big(\frac{-at(1-t)}{1 - at}\Big)^{e(H) - e(H\cap A)} \\ &= \mathrm{E}\, Y_A(t)(1-t)^{e(H)}\Big(\frac{1 - at}{-at}\Big)^{e(H\cap A) - e(H)} \\ &= \mathrm{E}\, Y_A(t)(1-t)^{e(H)}\Big(\frac{1 - at}{-at}\Big)^{-e(H)}\sum_{F\subseteq A\cap H}\Big(\frac{1}{-at}\Big)^{e(F)}.\end{aligned}$$

Consequently, with $F$ ranging over the $2^{e(H)}$ subgraphs of $H$ without isolated vertices,

$$\begin{aligned}\widehat{\psi}(H;t) &= \frac{1}{\operatorname{aut} H}\frac{\operatorname{E} X(t)}{N}(-a)^{e(H)}(1-at)^{-e(H)}\sum_{F\subseteq H}\Bigl(-\frac{1}{at}\Bigr)^{e(F)}N(F)\\ &= \frac{1}{\operatorname{aut} H}\operatorname{E} X(t)(1-at)^{-e(H)}(-a)^{e(H)}\sum_{F\subseteq H}(-at)^{-e(F)}\gamma(F). \end{aligned} \tag{4.19}$$

In particular, choosing $H = P_2$ we have 4 choices of $F$, namely $P_2$, $K_2$ (twice) and the empty graph. Hence, using $ap = \gamma_1 = \gamma(K_2)$ and our assumptions on $\gamma_2 = \gamma(P_2)$ and $t$,

$$\begin{aligned}\widehat{\psi}(P_2;t) &= \frac{1}{2}\operatorname{E} X(t)(1-at)^{-2}a^2\Bigl(1-\frac{2}{at}\gamma(K_2)+\frac{1}{(at)^2}\gamma(P_2)\Bigr).\\ &= \frac{1}{2}\operatorname{E} X(t)(1-at)^{-2}p^{-2}\gamma_1^2\Bigl(1-2\frac{p}{t}+\frac{p^2}{t^2}\frac{\gamma_2}{\gamma_1^2}\Bigr)\\ &= \frac{1}{2}\operatorname{E} X(t)\frac{1}{n^2p^2}(\tau+o(1)). \end{aligned} \tag{4.20}$$

The variance of the $P_2$ term in (4.17) is thus

$$\begin{aligned}\widehat{\psi}(P_2;t)^2 \operatorname{Var} S_n(P_2;t) &= \frac{1}{4}(\operatorname{E} X(t))^2\frac{1}{n^4p^4}(\tau^2+o(1))\cdot 2(n)_3t^2(1-t)^2\\ &= (\operatorname{E} X(t))^2(\tau^2+o(1))\frac{(1-p)^2}{2np^2}. \end{aligned} \tag{4.21}$$

Moreover, it follows easily from (4.8) that $\operatorname{E} X(t) \sim \operatorname{E} X(p)$ for the $t$ that we consider.

Consider first the case $p \gg n^{-1/2}$ (Theorem 5), and define

$$\beta = n^{-1/2}p^{-1}\operatorname{E} X(p). \tag{4.22}$$

Comparing (4.21) with (4.15), we then see, since the decomposition (4.17) is orthogonal and thus $\operatorname{Var} X(t) \geq \operatorname{Var}(\widehat{\psi}(P_2;t)S_n(P_2;t))$, that equality holds in (4.15), which now may be written

$$\operatorname{Var} X(t) = \bigl(\tfrac{1}{2}\tau^2(1-\pi)^2+o(1)\bigr)\,\beta^2. \tag{4.23}$$

Moreover, the variance of the remaining terms in (4.17) is $o(\beta^2)$. Thus, using the terminology of [4], and excepting the degenerate case $\tau = 0$, $X(t)$ is dominated by $P_2$. In particular $\widehat{\psi}^*(H;t) = o(\beta)$ for every $H \neq P_2$.

It follows by [4, Theorem 2 or 3], and trivially when $\tau = 0$, that

$$\bigl(X(p)-\operatorname{E} X(p)\bigr)/\beta \xrightarrow{d} \operatorname{N}\bigl(0,\tfrac{1}{2}\tau^2(1-\pi)^2\bigr). \tag{4.24}$$

In the case $p \sim 2cn^{-1/2}$ (Theorem 8), we, more generally, have to consider $H = kP_2$, $k \geq 1$.

For such $H$, if $F$ is a subgraph of $H$ without isolated vertices, then $F \cong iP_2 + jK_2$ for some $i$ and $j$, and by the assumption (4.1),

$$\begin{aligned}(-a)^{e(H)}(-at)^{-e(F)}\gamma(F) &= \Bigl(\frac{p}{t}\Bigr)^{e(F)}p^{-e(H)}(-\gamma_1)^{e(H)-e(F)}\gamma(F)\\ &= (1+o(1))p^{-2k}(-\gamma_1)^{2k-2i-j}\gamma_2^i\gamma_1^j. \end{aligned} \tag{4.25}$$

It follows by (4.19), since each of the $k$ components in $H$ may contribute a $P_2$, a $K_2$ (in two ways) or nothing to $F$, that

$$\begin{aligned}\widehat{\psi}(kP_2;t)/\operatorname{E}X(t) &= \frac{1}{2^k k!}\sum_F p^{-2k}(-1)^j\gamma_1^{2k-2i}\gamma_2^i + o(n^{-2k}p^{-2k})\\ &= \frac{p^{-2k}}{2^k k!}(\gamma_2 - 2\gamma_1^2+\gamma_1^2)^k + o(n^{-2k}p^{-2k})\\ &= n^{-2k}p^{-2k}\Big(\frac{\tau^k}{2^k k!}+o(1)\Big). \end{aligned}\tag{4.26}$$

In this case we define

$$\beta = \operatorname{E}X(p)$$

and obtain

$$\widehat{\psi}^*(kP_2;t)/\beta = n^{3k/2}p^k\widehat{\psi}(kP_2;t)/\beta \to \frac{1}{k!}\Big(\frac{\tau}{4c}\Big)^k. \tag{4.27}$$

By computing the contribution of the $kP_2$ terms in (4.17) to the variance of $X(t)$, we obtain

$$\begin{aligned}\liminf \operatorname{Var}(X(t))/\beta^2 &\geq \sum_{k=1}^{\infty}\lim|\widehat{\psi}^*(kP_2;t)/\beta|^2 \operatorname{aut}(kP_2)\\ &= \sum_{k=1}^{\infty}(k!)^{-2}\Big(\frac{\tau}{4c}\Big)^{2k}2^k k! = \exp\Big(\frac{\tau^2}{8c^2}\Big)-1. \end{aligned}\tag{4.28}$$

Thus equality holds in (4.16). Moreover, if $\tau \neq 0$, [4, Proposition 4.7] shows that $X(p)$ is almost finitely dominated by $\{kP_2\}_{k=1}^{\infty}$, and [4, Theorem 2] then yields

$$\begin{aligned}\big(X(p)-\operatorname{E}X(p)\big)/\beta &\xrightarrow{d} \sum_{k=1}^{\infty}\frac{1}{k!}\Big(\frac{\tau}{4c}\Big)^k :U(P_2;1)^k:\\ &= \exp\Big(\frac{\tau}{4c}U(P_2;1)-\frac{1}{2}\Big(\frac{\tau}{4c}\Big)^2\operatorname{Var}U(P_2;1)\Big)-1,\end{aligned}$$

with $U(P_2;1)\sim \mathrm{N}(0,2)$, and thus

$$X(p)/\beta \xrightarrow{d} \mathrm{LN}(-\frac{\tau^2}{16c^2},\frac{\tau^2}{8c^2}). \tag{4.29}$$

(When $\tau = 0$, this is trivial by (4.16).)

We now apply [4, Theorem 9(i) (or (iii), (iv))], which shows that $\psi(G_{nm})$ has the same asymptotic distribution (4.24) or (4.29) as $X(p)$, provided that furthermore

$$\operatorname{Var}(\xi(t)) = o(n^2\beta^2/p), \tag{4.30}$$

where $\xi(t)$ is the drift of $X(t)$, see [4]. Since

$$\psi(G_{nm})/\operatorname{E}X(p) = \varphi(G_{nm})(1-a)^{m-\mu}/Np^{\mu}(1-ap)^{\binom{n}{2}-\mu}, \tag{4.31}$$

easy calculations then yield (3.13) and (4.4).

It remains to prove (4.30). By [4, Proposition 2.8],

$$\xi(t)=(1-t)^{-1}\sum_e\frac{\partial X(t)}{\partial I_e}(1-I_e(t))=(1-t)^{-1}\sum_e\sum_A\frac{\partial Y_A(t)}{\partial I_e}(1-I_e(t)), \tag{4.32}$$

where $Y_A(t)$ is as in (4.7) and the derivatives are interpreted formally. In other words,

$$\xi(t)=(1-t)^{-1}\sum_{A\in\mathcal{A}}\sum_e Y_{A,e}(t), \tag{4.33}$$

where $Y_{A,e}(t)$ is obtained from $Y_A(t)$ by replacing the factor $I_e(t)$ by $1-I_e(t)$ if $e\in A$, and replacing $1-aI_e(t)$ by $-a(1-I_e(t))$ if $e\notin A$. Hence,

$$\mathrm{E}\,\xi^2(t)=(1-t)^{-2}\sum_{A_1,A_2}\sum_{e_1,e_2}E(Y_{A_1,e_1}Y_{A_2,e_2}). \tag{4.34}$$

By the independence of $I_e(t)$ in $G_n(t)$, we easily obtain

$$\frac{\mathrm{E}(Y_{A_1e_1}Y_{A_2e_2})}{\mathrm{E}\,Y_{A_1}Y_{A_2}}=\begin{cases}\frac{1-t}{t} & e_1=e_2\in A_1\cap A_2\\ \frac{-a(1-t)}{t(1-a)} & e_1=e_2\in(A_1\setminus A_2)\cup(A_2\setminus A_1)\\ \frac{a^2(1-t)}{1-t+t(1-a)^2} & e_1=e_2\in A_1^c\cap A_2^c\\ 0 & e_1\neq e_2\quad\text{and}\quad e_1\in A_2\quad\text{or}\quad e_2\in A_1\\ (\frac{1-t}{t(1-a)})^2 & e_1\in A_1\setminus A_2,\ e_2\in A_2\setminus A_1\\ \frac{1-t}{t(1-a)}\cdot\frac{-a(1-t)}{1-t+t(1-a)^2} & e_1\in A_1\setminus A_2,\ e_2\in A_1^c\cap A_2^c\quad\text{or conversely}\\ (\frac{-a(1-t)}{1-t+t(1-a)^2})^2 & e_1,e_2\in A_1^c\cap A_2^c,\ e_1\neq e_2\end{cases}$$

and thus, counting the number of different terms and completing a square,

$$\begin{aligned}&\sum_{e_1,e_2}\mathrm{E}(Y_{A_1e_1}Y_{A_2e_2})/\,\mathrm{E}(Y_{A_1}Y_{A_2})\\ &\quad=e(A_1\cap A_2)\frac{1-t}{t}-2e(A_1\setminus A_2)\frac{a(1-t)}{t(1-a)}\\ &\qquad+e(A_1^c\cap A_2^c)\Big(\frac{a^2(1-t)}{1-t+t(1-a)^2}-\Big(\frac{a(1-t)}{1-t+t(1-a)^2}\Big)^2\Big)\\ &\qquad+\Big(e(A_1\setminus A_2)\frac{1-t}{t(1-a)}-e(A_1^c\cap A_2^c)\frac{a(1-t)}{1-t+t(1-a)^2}\Big)^2\\ &\quad\le e(A_1\cap A_2)\frac{1}{t}+\tbinom{n}{2}a^2+\Big(\frac{\mu}{t(1-a)}+O\Big(\frac{e(A_1\cap A_2)}{t}\Big)-\tbinom{n}{2}\frac{a}{1-2ta+ta^2}+O(\mu a)\Big)^2\\ &\quad=O\Big(1+\Big(\frac{e(A_1\cap A_2)}{t}\Big)^2+n^2a^2+\mu^2a^2+\Big(\frac{\mu}{t}(1-2ta+ta^2)-\tbinom{n}{2}(a-a^2)\Big)^2\Big)\\ &\quad=O\Big(1+\Big(\frac{e(A_1\cap A_2)}{t}\Big)^2+\mu^2a^2+n^4a^4+\Big(\frac{\mu}{t}-\binom{n}{2}\frac{\mu}{m}\Big)^2\Big)\\ &\quad=O\Big(1+\Big(\frac{e(A_1\cap A_2)}{t}\Big)^2+\frac{n^4\mu^4}{m^4}+\mu^2\Big(\frac{1}{t}-\frac{1}{p}\Big)^2\Big).\end{aligned} \tag{4.35}$$

Using the assumptions $\mu = O(n)$ and $t = p + O(p^{1/2}/n)$, this reduces to

$$O\Big(\Big(\frac{e(A_1\cap A_2)}{t}\Big)^2 + \frac{1}{t^4}\Big) = O\Big(\frac{1}{t^4}(1+t)^{e(A_1\cap A_2)}\Big).$$

Consequently,

$$\mathrm{E}(\xi(t))^2 = O\Big(t^{-4}\sum_{A_1,A_2}(1+t)^{e(A_1\cap A_2)}\,\mathrm{E}\,Y_{A_1}(t)Y_{A_2}(t)\Big). \tag{4.36}$$

This sum, $S$ say, is evaluated as in (4.9), the only difference being that the argument $\frac{1-t}{t(1-a)^2} = x$ of $\Lambda$ is replaced by $(1+t)(1+x)-1 = x+t+xt = x+O(1)$. The argument in (4.10)–(4.14) now yields

$$\begin{aligned}\log\frac{S}{(\mathrm{E}\,X(t))^2} &= \lambda(x+O(1)) - \frac{\mu^2}{\binom{n}{2}}x + \frac{\mu^3}{\binom{n}{2}^2}x^2 + O\Big(\frac{1}{n^2t^3}\Big)\\ &\le \frac{\tau^2}{2n}x^2 + o\Big(\frac{x^2}{n}\Big) + O(\lambda_1+\lambda_2 x)\\ &= O(1),\end{aligned} \tag{4.37}$$

and thus

$$S = O\big((\mathrm{E}\,X(t))^2\big) = O(np^2\beta^2), \tag{4.38}$$

$$\mathrm{E}(\xi(t))^2 = O(p^{-4}S) = O(np^{-2}\beta^2) = o(n^2\beta^2/p), \tag{4.39}$$

which proves (4.30) and completes the proof. □

**Remark 4.1.** It is easy to give a direct proof of Theorem 6 using (4.15) in the weak version $\frac{\operatorname{Var}X(p)}{(\mathrm{E}\,X(p))^2}\to 0$ which, by Chebyshev's inequality, yields $X(p)/\mathrm{E}\,X(p)\xrightarrow{p}1$ and thus

$$\log X(p) - \log\mathrm{E}\,X(p)\xrightarrow{p}0.$$

Since $X(p) = \psi(G_{np}) = \varphi(G_{np})(1-a)^{e(G_{np})-\mu}$, we have

$$\log\varphi(G_{np}) = \log X(p) - (e(G_{np})-\mu)\log(1-a),$$

and it follows easily that

$$p^{1/2}\big(\log\varphi(G_{np}) - \log\mathrm{E}\,X(p) + (\tbinom{n}{2}p-\mu)\log(1-a)\big)\xrightarrow{d}\mathrm{N}\big(0,2\kappa^2(1-\pi)\big),$$

which gives (3.32) by elementary calculations.

**Remark 4.2.** Note that this proof gives an asymptotic distribution with normalizing constants derived from the mean and variance of $\psi(G_{np})$, without caring about the actual mean and variance of $\varphi(G_{nm})$ which we have computed separately.

## 5. Proofs of Theorems 1–4

In this section we shall prove that the three variables considered in the introduction satisfy the conditions of Theorems 5–8, with the following parameters:

$f$, spanning subtrees: $\gamma_1 = 2/n$, $\gamma_2 = 3/n^2$ and thus $\tau = -1$; $\kappa = 1$.
$g$, Hamilton cycles: $\gamma_1 \sim 2/n$, $\gamma_2 \sim 2/n^2$ and thus $\tau = -2$; $\kappa = 1$.
$h$, perfect matchings: $\gamma_1 \sim 1/n$, $\gamma_2 = 0$ and thus $\tau = -1$; $\kappa = 1/2$.

Theorems 1–4 then follow immediately from Theorems 5, 8(i), 7, 6 and 8(ii), respectively.

The main problem is to verify (3.10), and we begin with a simple estimate.

**Lemma 5.1.** *If $x,\varepsilon>0$, and $\varepsilon^2(x+x^3)$ is sufficiently small, then*

$$\sum_{k=0}^{\infty}\frac{x^k}{k!}e^{-\varepsilon k^2}=\exp\Bigl(x-\varepsilon x-\varepsilon x^2+O\bigl(\varepsilon^2(x+x^3)\bigr)\Bigr). \tag{5.1}$$

*If furthermore $C<\infty$ is fixed, then also*

$$\sum_{k=0}^{n}\frac{x^k}{k!}e^{-\varepsilon k^2+Ck^3/n^2}\le\exp\Bigl(x-\varepsilon x-\varepsilon x^2+O\bigl((\varepsilon^2+n^{-2})(x+x^3)\bigr)\Bigr). \tag{5.2}$$

*Proof.* We have, with $y=xe^{-2\varepsilon x}=x-2\varepsilon x^2+O(\varepsilon^2x^3)$,

$$\begin{aligned}\sum_{k=0}^{\infty}\frac{x^k}{k!}e^{-\varepsilon k^2}&=\sum_k\frac{x^k}{k!}e^{-2\varepsilon kx-\varepsilon(k-x)^2+\varepsilon x^2}=e^{\varepsilon x^2}\sum_k\frac{y^k}{k!}e^{-\varepsilon(k-x)^2}\\&=e^{\varepsilon x^2}\sum_k\frac{y^k}{k!}\Bigl(1-\varepsilon(k-x)^2+O(\varepsilon^2(k-x)^4)\Bigr)\\&=e^{\varepsilon x^2+y}\Bigl(1-\varepsilon(y+(y-x)^2)+O\bigl(\varepsilon^2((y-x)^4+y^2+y)\bigr)\Bigr)\\&=e^{x-\varepsilon x^2+O(\varepsilon^2x^3)}\bigl(1-\varepsilon x+O(\varepsilon^2x^2+\varepsilon^2x+\varepsilon^3x^4+\varepsilon^6x^8)\bigr).\end{aligned} \tag{5.3}$$

Hence (5.1) follows by taking the logarithm, using $\varepsilon^2x^2\le\varepsilon^2(x+x^3)$ and $\varepsilon^6x^8\le\bigl(\varepsilon^2(x+x^3)\bigr)^3$.

For (5.2) we note that if $a_k=\frac{x^k}{k!}e^{-\varepsilon k^2+Ck^3/n^2}$ and $k<n$, then $a_{k+1}/a_k<\frac{x}{k}e^{3C}$. Hence $a_{k+1}/a_k<e^{-1}$ if $n\ge k\ge k_0=\lceil e^{3c+1}x\rceil$, and $\sum_{k_0+2\log n+1}^{n}a_k\le(1-e^{-1})^{-1}n^{-2}a_{k_0+1}=O(\frac{x}{n^2}a_{k_0})$. Consequently, it suffices to consider the sum with $k<k_0+2\log n+1$. If $x\ge\log n$, this implies $Ck^3/n^2=O(x^3/n^2)$, and (5.2) follows by (5.1). If $x<\log n$, we have $k=O(\log n)$ and thus

$$e^{Ck^3/n^2}\le 1+C_1\frac{k^3}{n^2}\le 1+C_2\frac{(k)_3+k}{n^2},$$

and again the result follows by (5.1) since

$$\sum_{k=1}^{\infty}\Bigl(1+C_2\frac{(k)_3+k}{n^2}\Bigr)\frac{x^k}{k!}e^{-\varepsilon k^2}\le\Bigl(1+C_2\frac{x^3+x}{n^2}\Bigr)\sum_{k=0}^{\infty}\frac{x^k}{k!}e^{-\varepsilon k^2}. \qquad\square$$

**Spanning subtrees.** In this case, the graphs $H$ that appear in (2.2) are the forests. If $H$ is a forest with $j_i$ components with $i$ vertices, $i=2,3,\dots$, then

$$\gamma(H)=\prod_i\Bigl(\frac{i}{n^{i-1}}\Bigr)^{j_i} \tag{5.4}$$

by a result of Moon [6]. There are, for each sequence $j_2,j_3,\dots$, with $\nu=\sum_2^\infty ij_i\le n$, exactly

$$\frac{(n)_\nu}{\prod_i j_i!\prod_i(i!)^{j_i}}\prod_i(i^{i-2})^{j_i}=\frac{(n)_\nu}{n^\nu}\prod_{i=2}^{\infty}\Bigl(\frac{i^{i-2}}{i!}n^i\Bigr)^{j_i}\frac{1}{j_i!} \tag{5.5}$$

such forest $H \subset K_n$, and thus

$$\Lambda(x) = \sum_H \gamma(H)^2 x^{e(H)} = \sum_{\{j_i\}} \frac{(n)_\nu}{n^\nu} \prod_{i=2}^{\infty} \frac{1}{j_i!} \Big( \frac{i^{i-2}}{i!} n^i \frac{i^2}{n^{2i-2}} \Big)^{j_i} x^{\sum j_i(i-1)}$$
$$= \sum_{\{j_i\}} \frac{(n)_\nu}{n^\nu} \prod_{i=2}^{\infty} \frac{1}{j_i!} \Big( \frac{i^i}{i!} n^{2-i} x^{i-1} \Big)^{j_i}. \tag{5.6}$$

We use the estimate

$$\frac{(n)_\nu}{n^\nu} = \prod_1^{\nu-1} (1 - \frac{k}{n}) \le \exp\big(-\sum_1^{\nu-1} \frac{k}{n}\big) = \exp\big(-\frac{\nu(\nu-1)}{2n}\big) \le \exp\big(-\frac{j_2(2j_2-1)}{n}\big) \tag{5.7}$$

and obtain finally, now summing over all sequences $\{j_i\}$ of non-negative integers and using Lemma 5.1,

$$\begin{aligned}
\Lambda(x) &\le \sum_{\{j_i\}} e^{-j_2(2j_2-1)/n} \prod_{i=2}^{\infty} \frac{1}{j_i!} \Big( \frac{i^i}{i!} n^{2-i} x^{i-1} \Big)^{j_i} \\
&= \sum_{j_2=0}^{\infty} e^{j_2/n - 2j_2^2/n} \frac{(2x)^{j_2}}{j_2!} \prod_{i=3}^{\infty} \frac{1}{j_i!} \Big( \frac{i^i}{i!} n^{2-i} x^{i-1} \Big)^{j_i} \\
&= \exp\Big( 2xe^{1/n} - \frac{2}{n} 2xe^{1/n} - \frac{2}{n} (2xe^{1/n})^2 + O\Big(\frac{1}{n^2}(x+x^3)\Big) + \sum_{i=3}^{\infty} \frac{i^i}{i!} n^{2-i} x^{i-1} \Big) \\
&= \exp\Big( 2x - \frac{2}{n}x - \frac{8}{n}x^2 + \frac{3^3}{3!}\frac{x^2}{n} + O\Big(\frac{x+x^3}{n^2}\Big)\Big)
\end{aligned} \tag{5.8}$$

for $x \ge 0$ with $x^3/n^2$ sufficiently small. Moreover, for the terms up to $x^2$ in the Taylor expanssion, the estimate above gives an error of at most $O(n^{-2})$ in each coefficient. Hence

$$\Lambda(x) \le \exp\Big( \lambda_1 x + \lambda_2 x^2 + O\Big(\frac{x+x^3}{n^2}\Big)\Big), \tag{5.9}$$

which implies (3.10) for $n^{-1} \ll x \le \delta n^{2/3}$ and (3.37) for $x = o(1)$. The remaining conditions in the theorems are immediately verified. Note, in particular, that

$$\gamma(iP_2 + jK_2) = \gamma(P_2)^i \gamma(K_2)^j$$

exactly, as soon as $n \ge 3i + 2j$.

**Hamilton cycles.** In this case, the graphs $H$ with $N(H) \ne 0$ are $C_n$ itself and all unions of disjoint paths. If $H$ is a disjoint union of $j_l$ paths of length $l$, $l = 1, 2, 3, \dots$, then

$$N(H) = \tfrac{1}{2}(n - \textstyle\sum_l l j_l - 1)!\, 2^{\sum j_l},$$

since by collapsing each component in $H$ to a single vertex, each Hamilton cycle (in $K_n$) containing $H$ may be obtained from a Hamilton cycle in the smaller set by choosing one of two possible orientations for each component in $H$; thus, with $\mu = \sum l j_l$,

$$\gamma(H) = 2^{\sum j_l} / (n-1)_\mu. \tag{5.10}$$

Given $j_1, j_2, \dots$ with $\nu = \sum(l+1)j_l \le n$, there are

$$\frac{(n)_\nu}{\prod_l j_l! \prod_l (l!)^{j_l}} \prod_l (\tfrac{1}{2} l!)^{j_l} = \frac{(n)_\nu}{\prod_l j_l!} 2^{-\sum j_l} \tag{5.11}$$

such graphs $H \subset K_n$. Finally, there are $N = \frac{1}{2}(n-1)!$ choices of $H \cong C_n$, each having $N(H) = 1$ and $\gamma(H) = 1/N$. Consequently,

$$\Lambda(x) = \sum_{\{j_l\}} \frac{(n)_\nu}{((n-1)_\mu)^2} \prod_l \frac{2^{j_l}}{j_l!} x^\mu + \frac{2}{(n-1)!} x^n. \tag{5.12}$$

We use the estimate

$$\begin{aligned}
\frac{(n)_\nu}{((n-1)_\mu)^2} &= \frac{n^2 (n)_\nu}{((n)_{\mu+1})^2} \\
&= n^{2+\nu-2(\mu+1)} \exp\Big(-\frac{\nu(\nu-1)}{2n} + 2\frac{(\mu+1)\mu}{2n} + O\Big(\frac{\nu^3}{n^2} + \frac{\mu^3}{n^2}\Big)\Big) \\
&= n^{\nu-2\mu} \exp\Big(-\frac{(2j_1+3j_2)(2j_1+3j_2-1)}{2n} + 2\frac{(j_1+2j_2+1)(j_1+2j_2)}{2n} \\
&\qquad\qquad + O\Big(\frac{j_1^3}{n^2} + \frac{j_2^3}{n^2} + \sum_3^n l j_l\Big)\Big) \\
&= n^{\nu-2\mu} \exp\Big(\frac{-2j_1^2 - 4j_1 j_2 - j_2^2}{2n} + \frac{2j_1}{n} + \frac{7j_2}{2n} + O\Big(\frac{j_1^3}{n^2} + \frac{j_2^3}{n^2} + \sum_3^n l j_l\Big)\Big)
\end{aligned} \tag{5.13}$$

and obtain, for some $C < \infty$, using Lemma 5.1, when $x^3/n^2$ is sufficiently small,

$$\begin{aligned}
\Lambda(x) &\le \sum_{\{j_l\}} \exp\Big(\frac{2}{n} j_1 - \frac{j_1^2}{n} + C\frac{j_1^3}{n^2} + \frac{7}{2n} j_2 + C\frac{j_2^3}{n^2} + \sum_3^n C l j_l\Big) \prod_l \frac{1}{j_l!} 2^{j_l} x^{l j_l} n^{(1-l)j_l} \\
&\le \sum_{j_1=0}^{n/2} \frac{1}{j_1!} (2xe^{2/n})^{j_1} e^{-j_1^2/n + Cj_1^3/n^2} \sum_{j_2=0}^{n/3} \frac{1}{j_2!} \Big(\frac{2x^2}{n} e^{7/2n}\Big)^{j_2} e^{Cj_2^3/n^2} \prod_{l=3}^\infty \sum_{j_l=0}^\infty \frac{1}{j_l!} \big(2e^{Cl} x^l n^{1-l}\big)^{j_l} \\
&\le \exp\Big(2e^{2/n} x(1 - \frac{1}{n}) - \frac{1}{n}(2e^{2/n}x)^2 + \frac{2}{n} e^{7/2n} x^2 + \sum_{i=3}^\infty 2n\Big(\frac{e^C x}{n}\Big)^l + O\big(n^{-2}(x + x^3)\big)\Big) \\
&= \exp\Big((2 + \frac{2}{n})x - \frac{2}{n} x^2 + O\Big(\frac{x + x^3}{n^2}\Big)\Big).
\end{aligned} \tag{5.14}$$

(The term $\frac{2}{(n-1)!} x^n$ is negligible and may be incorporated in e.g. the term with $j_1 = 1$, $j_2 = \cdots = 0$, without altering the estimates.)

Again, the estimates have errors at most $O(n^{-2})$ for the Taylor coefficients up to $x^2$, and the result may be written

$$\Lambda(x) \le \exp\Big(\lambda_1 x + \lambda_2 x^2 + O\Big(\frac{x + x^3}{n^2}\Big)\Big), \quad 0 \le x \le \delta n^{2/3}. \tag{5.15}$$

The conditions in Theorems 5–8 follow.

**Perfect Matchings.** We only have to consider $H = jK_2$, for which

$$\gamma(jK_2) = \frac{1}{n-1} \cdot \frac{1}{n-3} \cdot \dots \cdot \frac{1}{n-2j+1}. \tag{5.16}$$

Hence

$$\Lambda(x) = \sum_{j=0}^{n/2} \frac{(n)_{2j}}{2^j j!} \gamma(jK_2)^2 x^j = \sum_{j=0}^{n/2} \frac{1}{j!} \Big(\frac{x}{2}\Big)^j \prod_{i=0}^{j-1} \frac{n-2i}{n-2i-1}. \tag{5.17}$$

Denote this product by $b_j$. If $j \leq n/4$ we have, for $0 \leq i < j$,

$$\frac{n-2i}{n-2i-1} = \frac{n}{n-1}\Big(1 + \frac{2i}{n(n-2i-1)}\Big) \leq \frac{n}{n-1} \exp\Big(\frac{4i}{n^2}\Big)$$

and thus

$$b_j \leq \Big(\frac{n}{n-1}\Big)^j \exp\Big(\sum_{k=0}^{j-1} \frac{4i}{n^2}\Big) = \Big(\frac{n}{n-1}\Big)^j \exp\Big(\frac{2j(j-1)}{n^2}\Big) \leq \Big(\frac{n}{n-1}\Big)^j \Big(1 + 3\frac{j(j-1)}{n^2}\Big);$$

if $n/4 < j \leq n/2$ we have, for large $n$,

$$b_j = \frac{n}{n-2j+1} \prod_{1}^{j-1} \frac{n-2i}{n-2i+1} < n \leq 64\frac{j^3}{n^2} \leq 65\frac{(j)_3}{n^2}.$$

Consequently, for $x \geq 0$ and $n$ large,

$$\begin{aligned}
\Lambda(x) &\leq \sum_{j=0}^{\infty} \frac{1}{j!} \Big(\frac{x}{2}\Big)^j \Big(\frac{n}{n-1}\Big)^j \Big(1 + 3\frac{(j)_2}{n^2} + 65\frac{(j)_3}{n^2}\Big) \\
&= \exp\left(\frac{n}{2(n-1)}x\right) \left(1 + \frac{3}{(n-1)^2}\Big(\frac{x}{2}\Big)^2 + \frac{65n}{(n-1)^3}\Big(\frac{x}{2}\Big)^3\right) \\
&= \exp\left(\frac{n}{2(n-1)}x + O\Big(\frac{x^2+x^3}{n^2}\Big)\right),
\end{aligned} \tag{5.18}$$

or

$$\Lambda(x) \leq \exp\left(\lambda_1 x + \lambda_2 x^2 + O\Big(\frac{x^2+x^3}{n^2}\Big)\right), \tag{5.19}$$

where $\lambda_1 = n/2(n-1)$, $\lambda_2 = O(n^{-2})$. Hence the conditions of Theorems 5–8 are satisfied in this case as well.

## 6. Other random graphs

The results and proofs above hold with minor changes also for several other random graph models. We consider here two cases, omitting the detailed verifications that the arguments above and in the relevant parts of [4] still are valid.

**Directed graphs.** Let $D_{nm}$ be the random digraph without loops with $n$ (labelled) vertices and $m$ edges drawn without replacement from the $n(n-1)$ possible edges. Similarly, let $D_{np}$ be the random digraph without loops with $n$ vertices where each edge appears with probability $p$, independently of all others.

We now assume that $\widetilde{\mathcal{A}}$ is a set of unlabelled digraphs, and argue as above. There are two differences from the undirected case. First, $\binom{n}{2}$ has to be replaced by $n(n-1)$ everywhere, which leads to new factors $\frac{1}{2}$ in some constants in the asymptotic results. Secondly, in formulas as (3.2) and (4.17), we should sum over *digraphs* $H$. This makes no essential difference for the $K_2$-terms, if we interpret $K_2$ as a directed edge, but $P_2$ is replaced by three different digraphs $P_{2i}$, $i = 0, +, -$.

$$\underset{P_{20}}{\bullet \longrightarrow \bullet \longrightarrow \bullet} \qquad \underset{P_{2+}}{\bullet \longleftarrow \bullet \longrightarrow \bullet} \qquad \underset{P_{2-}}{\bullet \longrightarrow \bullet \longleftarrow \bullet}$$

There is also a fourth connected digraph with two edges, namely the cycle $C_2$ with two vertices.

Let $\gamma_1 = \gamma(K_2) = \mu/n(n-1)$ and define, as before, $\lambda_1$ and $\lambda_2$ by (3.7). (Note that (3.8) and (3.9) have to be modified.) Then the following analogue of Theorem 5 holds.

**Theorem 9.** *Suppose that $n \to \infty$ and $\mu = O(n)$, that $n^2(\gamma(P_{2i}) - \gamma_1^2) \to \tau_i \in (-\infty, \infty)$ for $i = 0, +, -$, that $\gamma(C_2) = O(n^{-2})$, and that for every sequence $x = x_n$ with $n^{-1} \ll x \ll n^{1/2}$,*

$$\Lambda(x) \le \exp\Bigl(\lambda_1 x + \lambda_2 x^2 + o\Bigl(\frac{x^2}{n}\Bigr)\Bigr). \tag{6.1}$$

*If $m \gg n^{3/2}$ and $n(n-1) - m \gg n$ then, with $p = m/n(n-1)$,*

$$\operatorname{E}\varphi(D_{nm}) = Np^{\mu}\exp\Bigl(-\frac{\mu^2}{2m}(1-p) + O\Bigl((1-p)\frac{n^3}{m^2}\Bigr)\Bigr), \tag{6.2}$$

$$\operatorname{Var}\varphi(D_{nm}) = \Bigl(\sigma^2 + o(1)\Bigr)(1-p)^2\frac{n^3}{m^2}\bigl(\operatorname{E}\varphi(D_{nm})\bigr)^2 \tag{6.3}$$

*and*

$$(1-p)^{-1}\frac{m}{n^{3/2}}\Bigl(\varphi(D_{nm})\Big/ Np^{\mu}\exp\Bigl(-\frac{\mu^2}{2m}(1-p)\Bigr) - 1\Bigr) \xrightarrow{d} \operatorname{N}\bigl(0, \sigma^2\bigr), \tag{6.4}$$

*where*

$$\sigma^2 = \tau_0^2 + \tfrac{1}{2}\tau_+^2 + \tfrac{1}{2}\tau_-^2. \tag{6.5}$$

*If furthermore $\sigma^2 > 0$, the standardized variable $\varphi(D_{nm})^*$ converges in distribution to the standard normal distribution.* $\square$

In the proof, we let $Y_i$ be the number of $P_{2i}$ in $D_{nm}$; then $Y_0^*$, $Y_+^*$ and $Y_-^*$ converge jointly to three independent standard normal distributions. The formula in Remark 3.2 is replaced by

$$n^{1/2}p(1-p)^{-1}(\operatorname{E}X)^{-1}(X - \operatorname{E}X) \approx \tau_0 Y_0^* + \tfrac{1}{\sqrt{2}}\tau_+ Y_+^* + \tfrac{1}{\sqrt{2}}\tau_- Y_-^*. \tag{6.6}$$

Theorems 6–8 are valid for digraphs if we replace $\mu^2/n^2$ by $\mu^2/2n^2$ in (3.32) and (4.5); $\kappa^2$ by $\frac{1}{2}\kappa^2$ in (3.32), (3.38), (4.2) and (4.5); $\tau^2/8$ by $\sigma^2$ in (4.3) and (4.4); and allow different $P_{2i}$ in (4.1).

For example, if $g(D)$ denotes the number of directed Hamilton cycles in $D$, we have $N = (n-1)!$, $\mu = n$, $\gamma_1 = 1/(n-1)$, $\gamma(P_{20}) = n/(n)_3 = 1/(n-1)(n-2)$, $\gamma(P_{2+}) = \gamma(P_{2-}) = 0$, $\gamma(C_2) = 0$, $\tau_0 = 1 - 1 = 0$, $\tau_+ = \tau_- = -1$ and $\sigma^2 = 1$. We can verify (6.1) as in Section 5 and obtain the following.

**Theorem 10.** *Assume that $n \to \infty$ and let $p = m/n(n-1)$.*

(i) *If $m \gg n^{3/2}$ and $n(n-1) - m \gg n$, then*

$$\mathrm{E}\, g(D_{nm}) = (n-1)!\, p^n \exp\Bigl(-\frac{1-p}{2p} + O\bigl((1-p)\frac{n^3}{m^2}\bigr)\Bigr), \tag{6.7}$$

$$\operatorname{Var} g(D_{nm}) \sim (1-p)^2 \frac{n^3}{m^2} \bigl(\mathrm{E}\, g(D_{nm})\bigr)^2 \tag{6.8}$$

*and*

$$g(D_{nm})^* \xrightarrow{d} \mathrm{N}(0,1). \tag{6.9}$$

(ii) *If $m/n^{3/2} \to c > 0$, then*

$$\mathrm{E}\, g(D_{nm}) \sim (n-1)!\, p^n \exp\bigl(-\frac{1-p}{2p} - \frac{1}{6c^2}\bigr), \tag{6.10}$$

$$\operatorname{Var} g(D_{nm}) \sim (e^{1/c^2} - 1)\bigl(\mathrm{E}\, g(D_{nm})\bigr)^2 \tag{6.11}$$

*and*

$$g(D_{nm}) / \mathrm{E}\, g(D_{nm}) \xrightarrow{d} \mathrm{LN}\bigl(-\frac{1}{2c^2}, \frac{1}{c^2}\bigr). \tag{6.12}$$

□

**Theorem 11.** *Assume that $n \to \infty$ and $p \to \pi \le 1$. Then*

$$\mathrm{E}\, g(D_{np}) = (n-1)!\, p^n. \tag{6.13}$$

(i) *If $\pi = 1$ and $1 - p \gg n^{-2}$, then*

$$\operatorname{Var} g(D_{np}) \sim (1-p)\bigl(\mathrm{E}\, g(D_{np})\bigr)^2, \tag{6.14}$$

$$g(D_{np})^* \to \mathrm{N}(0,1). \tag{6.15}$$

(ii) *If $0 \le \pi < 1$ and $\liminf p n^{1/2} > 0$, then*

$$p^{1/2}\Bigl(\log g(D_{np}) - \log\bigl(\mathrm{E}\, g(D_{np})\bigr) + \frac{1-p}{2p}\Bigr) \xrightarrow{d} \mathrm{N}(0, 1-\pi). \tag{6.16}$$

□

If we consider several variables, we can obtain multivariate normal limits of rank 1, 2 or 3. For example, let $h$ as before denote the number of perfect matchings (for even $n$), which has $\mu = n/2$, $\gamma_1 = 1/2(n-1)$, $\gamma(P_{20}) = \gamma(P_{2+}) = \gamma(P_{2-}) = \gamma(C_2) = 0$, $\tau_0 = \tau_+ = \tau_- = -\frac{1}{4}$ and $\sigma^2 = \frac{1}{8}$. We verify (6.1) as in Section 5; in fact $\Lambda(x) = \Lambda_{\mathrm{undirected}}(x/2)$, where the latter is given by (5.17). If $n \to \infty$, $m \gg n^{3/2}$ and $n(n-1) - m \gg n$, we have by (6.6) after standardization,

$$g(D_{nm})^* \approx -\tfrac{1}{\sqrt{2}} Y_+^* - \tfrac{1}{\sqrt{2}} Y_-^*, \tag{6.17}$$

$$h(D_{nm})^* \approx -\tfrac{1}{\sqrt{2}} Y_0^* - \tfrac{1}{2} Y_+^* - \tfrac{1}{2} Y_-^*; \tag{6.18}$$

hence these standardized variables converge jointly to two standard normal variables with correlation $0 \cdot \frac{1}{\sqrt{2}} + \frac{1}{\sqrt{2}} \cdot \frac{1}{2} + \frac{1}{\sqrt{2}} \cdot \frac{1}{2} = \frac{1}{\sqrt{2}}$.

Let us now consider digraphs with loops, and let $D'_{nm}$ and $D'_{np}$ denote the random digraphs defined as $D_{nm}$ and $D_{np}$ but allowing loops. The situation is now more complicated, since there is a further connected digraph with one edge, viz. the loop $C_1$, and several more digraphs with two edges. We therefore consider only the case when we count the number of some loopless subgraphs, i.e. when $\widetilde{\mathcal{A}}$ is a set of digraphs without loops. In this case, we can ignore all loops in the random graph and use the results above for digraphs without loops. Note in particular that $\Lambda(x)$, and thus condition (6.1), do not depend on whether we consider digraphs with loops or without.

For $D'_{np}$ we obtain $D_{np}$ by deleting all loops. Hence the results above for $D_{np}$ are valid for $D'_{np}$ as well.

For $D'_{nm}$, we have a random number $M$ of non-loops, but conditioned on $M$, the graph obtained by deleting all loops may be regarded as $D_{nM}$. Here $M$ has a hypergeometric distribution with parameters $n^2$, $n^2-n$ and $m$; thus $M$ is asymptotically normal and $\operatorname{Var} M \sim p(1-p)n$, where $p = m/n^2$. An argument similar to the one in the proof of Theorem 6 gives the asymptotic distribution, but it now depends on $p$ whether the extra variation caused by the variation in $M$ dominates the variation of $\varphi(D_{nM})$ for a fixed $M \approx m(1-\frac{1}{n})$ or not.

A simple calculation shows that if $p \to 0$, then the variation of $\varphi(D_{nM})$ dominates and we obtain the same results as before; if $p \to 1$, the variation in $M$ dominates (as was the case in Theorem 6); if $p \to \pi \in (0,1)$, both variations are of the same order and have to be combined in the final result. The expectation and variance of $\varphi(D'_{nm})$ are computed as in (3.19)–(3.25); the only significant difference is that in (3.25), the terms $\mu^2(1-p)/m = \mu^2(1-p)/pn^2$ and $\lambda_1(1-p)/p = \mu^2(1-p)/pn(n-1)$ no longer cancel exactly, which leads to an extra term in the variance estimate (which is negligible if $p \to 0$). We omit the details, but state the resulting version of Theorem 9.

**Theorem 12.** *Suppose that $n \to \infty$ and $\mu/n \to \kappa$, that $n^2(\gamma(P_{2i}) - \gamma_1^2) \to \tau_i \in (-\infty,\infty)$ for $i = 0,+,-$, that $\gamma(C_2) = O(n^{-2})$, and that for every sequence $x = x_n$ with $n^{-1} \ll x \ll n^{1/2}$, (6.1) holds. If $m \gg n^{3/2}$, $n^2 - m \gg n$ and $p = m/n^2 \to \pi \geq 0$, then*

$$\operatorname{E}\varphi(D'_{nm}) = Np^{\mu}\exp\Big(-\frac{\mu^2}{2m}(1-p) + O\Big((1-p)\frac{n^3}{m^2}\Big)\Big), \tag{6.19}$$

$$\operatorname{Var}\varphi(D'_{nm}) = \Big((1-\pi)\sigma^2 + \pi\kappa^2 + o(1)\Big)(1-p)\frac{n^3}{m^2}\big(\operatorname{E}\varphi(D'_{nm})\big)^2 \tag{6.20}$$

*and*

$$(1-p)^{-1/2}\frac{m}{n^{3/2}}\Big(\varphi(D'_{nm})\Big/Np^{\mu}\exp\Big(-\frac{\mu^2}{2m}(1-p)\Big) - 1\Big) \xrightarrow{d} \operatorname{N}\big(0,(1-\pi)\sigma^2 + \pi\kappa^2\big), \tag{6.21}$$

*where $\sigma^2$ is given by (6.5). If furthermore $\sigma^2 > 0$, or if $\sigma^2 = 0$, $\pi > 0$ and $\kappa > 0$, then the standardized variable $\varphi(D'_{nm})^*$ converges in distribution to a standard normal distribution.* $\square$

Theorems 6–8 are valid for random digraphs with loops with the same modifications as for digraphs without loops.

For example, for the number of directed Hamilton cycles we have $\sigma^2 = 1$ and $\kappa = 1$, which gives the following. (A weaker version of the variance estimate (6.23) is given in [2].)

**Theorem 13.** *Assume that $n \to \infty$, $m \gg n^{3/2}$ and $n^2 - m \gg n$, and let $p = m/n^2$.*

*Then*

$$\mathrm{E}\, g(D'_{nm}) = (n-1)!\, p^n \exp\left( -\frac{1-p}{2p} + O\Big((1-p)\frac{n^3}{m^2}\Big)\right), \tag{6.22}$$

$$\operatorname{Var} g(D'_{nm}) \sim (1-p)\frac{n^3}{m^2}\big(\mathrm{E}\, g(D'_{nm})\big)^2 \tag{6.23}$$

*and*

$$g(D'_{nm})^* \xrightarrow{d} \mathrm{N}(0,1). \tag{6.24}$$

□

The results in Theorem 10(ii) and Theorem 11 remain valid for $D'_{nm}$ and $D'_{np}$.

As a further example, we observe that Theorem 12 implies that, for $m$ as above and $n$ even, $\operatorname{Var} h(D'_{nm}) \sim \frac{1}{8}(1+p)(1-p)n^3m^{-2}$, since $\sigma^2 = \frac{1}{8}$ and $\kappa = \frac{1}{2}$. Moreover, (6.21) may be generalized to vector-valued variables. For example, if $m$ is as above, then, by (6.6),

$$(1-p)^{-1}\frac{m}{n^{3/2}}\Big(\frac{g(D_{nm})}{\mathrm{E}\, g(D_{nm})} - 1, \frac{h(D_{nm})}{\mathrm{E}\, h(D_{nm})} - 1\Big) \xrightarrow{d} \mathrm{N}(0,\Sigma), \tag{6.25}$$

with covariance matrix $\sigma_{11} = 1$, $\sigma_{12} = \frac{1}{4}$, $\sigma_{22} = \frac{1}{8}$. Hence, if furthermore $p \to \pi \le 1$, then

$$(1-p)^{-1/2}\frac{m}{n^{3/2}}\Big(\frac{g(D'_{nm})}{\mathrm{E}\, g(D'_{nm})} - 1, \frac{h(D'_{nm})}{\mathrm{E}\, h(D'_{nm})} - 1\Big) \xrightarrow{d} \mathrm{N}(0,\Sigma'), \tag{6.26}$$

with covariance matrix $\sigma'_{11} = (1-\pi) + \pi = 1$, $\sigma'_{12} = (1-\pi)\frac{1}{4} + \pi\frac{1}{2} = (1+\pi)/4$, $\sigma'_{22} = (1-\pi)\frac{1}{8} + \pi\frac{1}{4} = (1+\pi)/8$. In particular, $g(D'_{nm})^*$ and $h(D'_{nm})^*$ converge jointly to two normal variables with correlation $\frac{1+\pi}{4}\Big/\sqrt{\frac{1+\pi}{8}} = \sqrt{\frac{1+\pi}{2}}$.

**Bipartite graphs.** We consider bipartite graphs with an explicit bipartition, i.e. graphs whose vertex set is partitioned into two subsets, which we assume are coloured white and black, such that no edge joins two vertices of the same colour.

For simplicity we consider only random bipartite graphs with equally many vertices of each colour, although at least the case when the numbers are within a constant of each other presents no further difficulties.

Thus, let $B_{nm}$ be the random bipartite graph with $2n$ labelled vertices, $n$ white and $n$ black, and $m$ edges drawn without replacement from the $n^2$ possible edges. Similarly, let $B_{np}$ be the random bipartite graph with $n+n$ vertices and edges drawn independently with probability $p$.

We now assume that $\widetilde{\mathcal{A}}$ is a set of bipartite graphs. In the arguments above, $\binom{n}{2}$ has to be replaced by $n^2$ everywhere, which again leads to new factors $\frac{1}{2}$ in some constants. Moreover, we should sum over bipartite graphs $H$. Thus $P_2$ should be replaced by two different bipartite graphs, with the middle vertex white and black, respectively, but if we restrict ourselves to sets $\widetilde{\mathcal{A}}$ that are invariant under colour inversion, the two terms may be combined into one and we may proceed as before.

Hence, for colour symmetric $\widetilde{\mathcal{A}}$, Theorems 5–8 are valid for bipartite graphs if we replace $\binom{n}{2}$ by $n^2$, in particular $p = m/n^2$ in Theorem 5; $\tau^2/8$ by $\tau^2$ in (3.12), (3.13), (4.3) and (4.4); $\mu^2/n^2$ by $\mu^2/2n^2$ in (3.32) and (4.5); $\kappa^2$ by $\frac{1}{2}\kappa^2$ in (3.32), (3.38), (4.2) and (4.5); and allow different colourings of the $P_2$ in (4.1).

For example, for the number of perfect matchings we have $N = n!$, $\mu = n$, $\gamma_1 = 1/n$, $\gamma_2 = 0$, $\tau = -1$, $\kappa = 1$. Furthermore,

$$\Lambda(x) = \sum_{j=0}^{n}\binom{n}{j}\binom{n}{j} j!\,\gamma(jK_2)^2 x^j = \sum_{j=0}^{n}\frac{x^j}{j!} \le \exp(x), \tag{6.27}$$

which verifies (6.1) with $\lambda_1 = 1$ and $\lambda_2 = 0$. This gives the following results.

**Theorem 14.** *Assume that $n \to \infty$ and let $p = m/n^2$.*

(i) *If $m \gg n^{3/2}$ and $n^2 - m \gg n$, then*

$$\operatorname{E} h(B_{nm}) = n!\, p^n \exp\Bigl(-\frac{1-p}{2p} + O\Bigl((1-p)\frac{n^3}{m^2}\Bigr)\Bigr), \tag{6.28}$$

$$\operatorname{Var} h(B_{nm}) \sim (1-p)^2 \frac{n^3}{m^2} \bigl(\operatorname{E} h(B_{nm})\bigr)^2 \tag{6.29}$$

*and*

$$h(B_{nm})^* \xrightarrow{d} \mathrm{N}(0,1). \tag{6.30}$$

(ii) *If $m/n^{3/2} \to c > 0$, then*

$$\operatorname{E} h(B_{nm}) \sim n!\, p^n \exp\bigl(-\frac{1-p}{2p} - \frac{1}{6c^2}\bigr), \tag{6.31}$$

$$\operatorname{Var} h(B_{nm}) \sim (e^{1/c^2} - 1)\bigl(\operatorname{E} h(B_{nm})\bigr)^2 \tag{6.32}$$

*and*

$$h(B_{nm}) / \operatorname{E} h(B_{nm}) \xrightarrow{d} \mathrm{LN}\bigl(-\frac{1}{2c^2}, \frac{1}{c^2}\bigr). \tag{6.33}$$

□

**Theorem 15.** *Assume that $n \to \infty$ and $p \to \pi \le 1$. Then*

$$\operatorname{E} h(B_{np}) = n!\, p^n. \tag{6.34}$$

(i) *If $\pi = 1$ and $1 - p \gg n^{-2}$, then*

$$\operatorname{Var} h(B_{np}) \sim (1-p)\bigl(\operatorname{E} h(B_{np})\bigr)^2, \tag{6.35}$$

$$h(B_{np})^* \to \mathrm{N}(0,1). \tag{6.36}$$

(ii) *If $0 \le \pi < 1$ and $\liminf pn^{1/2} > 0$, then*

$$p^{1/2}\Bigl(\log h(B_{np}) - \log\bigl(\operatorname{E} h(B_{np})\bigr) + \frac{1-p}{2p}\Bigr) \xrightarrow{d} \mathrm{N}(0, 1-\pi). \tag{6.37}$$

□

These theorems can also be interpreted as asymptotic results for the permanent of a random 0–1 matrix. A weaker version of the variance estimate (6.29) is given in [5].

## References


1. B. Bollobás, *Random Graphs*, Academic Press, London, 1985.
2. A. Frieze and S. Suen, *Counting the number of Hamilton cycles in random digraphs*, Random Struct. Alg. **3** (1992), 235–241.
3. S. Janson, *Random trees in a graph and trees in a random graph*, Math. Proc. Camb. Phil. Soc. **100** (1986), 319–330.
4. S. Janson, *Orthogonal decompositions and functional limit theorems for random graph statistics*, Uppsala Univ. Dept. Math. Report 1992:3.
5. M. Jerrum, *An analysis of a Monte Carlo algorithm for estimating the permanent* (to appear).
6. J. W. Moon, *Enumerating labelled trees*, Graph theory and theoretical physics, ed. F. Harary, Academic Press, London, 1967, pp. 261–272.
7. A. Ruciński, *When are small subgraphs of a random graph normally distributed?*, Prob. Th. Rel. Fields. **78** (1988), 1–10.



Department of Mathematics, Uppsala University, PO Box 480, S-751 06 Uppsala, Sweden

*E-mail address*: svante.janson@math.uu.se