CAMBRIDGE
UNIVERSITY PRESS

ARTICLE

Approximation of subgraph counts in the uniform attachment model

Johan Björklund[1,2], Cecilia Holmgren[1,2], Svante Janson[1,2] (iD) and Tiffany Y. Y. Lo[2,3]* (iD)

[1]Department of Electrical Engineering, Mathematics and Science, University of Gävle, Gävle, Sweden, [2]Department of Mathematics, Uppsala University, Uppsala, Sweden, and [3]Department of Mathematics, Stockholm University, Stockholm, Sweden
Corresponding author: Tiffany Y. Y. Lo; Email: tiffany.y.y.lo@math.su.se

Abstract

We use Stein's method to obtain distributional approximations of subgraph counts in the uniform attachment model or random directed acyclic graph; we provide also estimates of rates of convergence. In particular, we give uni- and multi-variate Poisson approximations to the counts of cycles and normal approximations to the counts of unicyclic subgraphs; we also give a partial result for the counts of trees. We further find a class of multicyclic graphs whose subgraph counts are a.s. bounded as $n \to \infty$.

1. Introduction

A random uniform attachment graph, denoted G_n^m and also known as a uniform random recursive *dag* (directed acyclic graph), can be constructed recursively as follows. Fix $m \geqslant 1$. The initial graph G_1^m is a single isolated vertex. To construct G_n^m from G_{n-1}^m, we add vertex n to G_{n-1}^m, with vertex n born with m edges, which are labelled by $1, \ldots, m$. For $i \in [m] := \{1, \ldots, m\}$, the other endpoint of $n^{(i)}$, the ith edge of vertex n, is then uniformly chosen among the existing vertices of G_{n-1}^m, i.e., in $[n-1]$. Thus, G_n^m has n vertices and $(n-1)m$ edges, and each edge has a label in $[m]$. Observe that we allow multiple edges when $m \geqslant 2$. An edge in G_n^m can be thought of as always pointing towards the vertex with the smaller label, and so there is no real distinction between the undirected and directed versions of the uniform attachment graph.

When $m = 1$, the model is the random recursive tree first studied in [16]; and when $m \geqslant 2$, the model was first introduced in [6]. There is an abundance of literature on random recursive trees (see e.g. [7] for an overview), but here we mention [9–11, 13], which provide Poisson and normal approximations to the counts of subtree copies. For $m \geqslant 2$, results on vertex degrees can be found in [6, 18], and results on depths and path lengths are available in [1, 3, 5, 19]. The recent paper [14] studies the number of vertices that can be reached from vertex n via a directed path, where the edge is thought as pointing from the larger vertex label towards the smaller one.

In this article, we consider the counts of subgraphs of G_n^m isomorphic to a given fixed graph as $n \to \infty$, where the parameter m is fixed with $m \geq 2$ (Figure 1). We provide uni- and multivariate
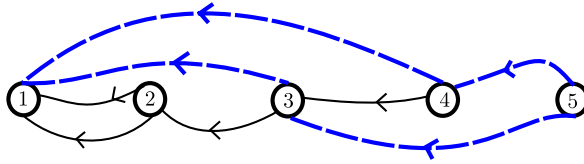
Figure 1. A 4-cycle (with blue and dotted edges) as a subgraph of G_5^2.

Poisson approximations to the counts of cycles. We also prove that the counts of unicyclic sub-graphs are asymptotically normal (after suitable renormalisation). We conjecture that the counts of more general trees are approximately normal and show that this conjecture holds for stars. All these approximation results are accompanied with convergence rates. We also consider multi-cyclic subgraphs. In particular, we show that if in addition to having more than two cycles, the subgraph is 'leaf-free' (i.e. has no vertices of degree one), then the number of copies in the uniform attachment graph is bounded a.s. For multicyclic subgraphs that are not leaf-free, we identify the exact rate of growth of the expectation and an upper bound on the variance. We conjecture that the count of a subgraph of this type converges a.s. after suitable rescaling, and prove a special case.

The distributional approximation tool that we use here is Stein's method, using in particular the size-bias coupling for both types of approximations; see [2, 4, 17] for an overview. This is in contrast to the analytic and contraction methods employed in [9–11]. In [13], Stein's method was applied in proving the Poisson and normal approximation results for the counts of subgraphs and their functionals in random recursive trees. However, we note that, for normal approximation, the approach that we use here is different from that in [13], which does not provide a convergence rate.

A major challenge in the distributional approximation problem of subgraph counts in the uniform attachment graph is the computation of the variance of the counts. In particular, our methods require a lower bound on the variance. In the case $m = 1$, the variance can be computed explicitly by using an elegant bijection between the random binary tree and the random recursive tree [[15], Section 2.3.2], as done in [13]. When $m \geqslant 2$, the order of the variance of the count of certain non-tree subgraphs can be obtained by analysing the covariances of the indicators that a certain subgraph exists in G_n^m, and finding the pairs of subgraphs that contribute the dominant term in the variance, but this task becomes difficult in the case of trees. In contrast to the case of $m = 1$, it is also much harder to obtain an explicit expression for the mean and variances in the case of $m \geqslant 2$. This is due to the fact for most subgraphs, the same set of vertices in the uniform attachment graph can form a copy in several ways. Consequently, we are only able to give the order of the mean and variance (i.e., within constant factors), and in some cases only an upper bound, and we only consider fixed subgraphs and a fixed m.

Remark 1.1. It would be interesting to extend the results of this paper to subgraphs H_n depending on n, or to $m = m(n)$ growing with n. Our methods still work in principle, but it becomes more complicated to estimate the various expressions, and we have not pursued this. We leave these as open problems.

2. Notation

The objective of this paper is thus to study the number of copies of a given graph H in G_n^m. We will only consider connected H. We emphasise that we generally see both G_n^m and the H as *undirected*. However, each edge in G_n^m has a canonical direction where an edge between the vertices i and j is directed towards the smaller of i and j. Thus every copy of H in G_n^m also has an induced direction on its edges. Note, however, that different copies may induce different directions in H.

Let K_n^m be the *m-fold complete multigraph on n*, defined to be the multigraph on $[n]$ such that for every pair (i, j) with $1 \leqslant i < j \leqslant n$, there are m edges $j \to i$, and these edges are labelled $1, \ldots, m$. We denote the edge $j \to i$ with label a by $ji^{(a)}$. Note that we can regard G_n^m as a subgraph of K_n^m in the obvious way (preserving edge labels); in fact, we may define G_n^m as the subgraph of K_n^m obtained by choosing, for every $j \in \{2, \ldots, n\}$ and every $a \in [m]$, exactly one of the edges $ji^{(a)}$ with label a and $i \in [j-1]$, all choices uniformly at random and independent.

Let Γ be the set of all copies of H in K_n^m that do not contain two different edges $ji_1^{(a)}$ and $ji_2^{(a)}$ with $i_1, i_2 < j$ and the same label a. (If there are two such edges $ji_1^{(a)}$ and $ji_2^{(a)}$, then the copy can never be a subgraph of G_n^m.) We say that Γ is the set of *potential copies* of H in G_n^m. For $\alpha \in \Gamma$, let $\mathbf{1}_\alpha$ be the indicator variable that takes value 1 if $\alpha \subseteq G_n^m$, and 0 otherwise. The number W_n of copies of H in G_n^m is thus

$$W_n = \sum_{\alpha \in \Gamma} \mathbf{1}_\alpha, \tag{2.1}$$

and we are therefore interested in approximating the distribution of this sum.

Example 2.1. *For a simple example, let H be a triangle. Then, a potential copy of H in G_n^m is described by its vertex set $\{i, j, k\} \subseteq [n]$, where we may assume $i < j < k$, together with its edges $ji^{(a)}$, $ki^{(b)}$, and $kj^{(c)}$, where $a, b, c \in [m]$ and $b \neq c$.*

Remark 2.2. Let m_H be the largest integer k such that the k-core of H is non-empty, i.e., there exists a non-empty subgraph of H where each node has degree at least k (in the subgraph). For example, $m_H = 1$ if H is a tree, $m_H = 2$ if H is a cycle, and $m_H = k - 1$ if H is the complete graph on k vertices. Then it is easy to see that if $m < m_H$, then there are no potential copies of H in G_n^m for any n (since in a copy of H, the vertex with largest label in the m_H-core has at least m_H edges to vertices with smaller labels); hence $\Gamma = \emptyset$ and thus $W_n = 0$ deterministically. On the other hand, if $m \geqslant m_H$ and n is large enough, then there are potential copies (since we may create a copy by assigning labels in decreasing order to the vertices of H, each time choosing a vertex that has at most m_H edges to the remaining vertices).

Remark 2.3. Recall that G_n^m is a loop-less multigraph. Similarly, H can be a loop-less multigraph in the results below, although we for simplicity write "graph" and "subgraph".

2.1. Probability distances

To precisely state our results, we also need to define the metrics in consideration here. The *total variation distance* d_{TV} between two probability measures ν_1 and ν_2 supported on $\mathbb{Z}^+$ is defined as

$$d_{\mathrm{TV}}(\nu_1, \nu_2) = \sup_{A \subset \mathbb{Z}^+} |\nu_1(A) - \nu_2(A)|; \tag{2.2}$$

and the *Wasserstein distance* d_{W} between two probability measures ν_1, ν_2 supported on $\mathbb{R}$ is defined as

$$d_{\mathrm{W}}(\mu, \nu) = \sup_f \left| \int f(x) d\mu(x) - \int f(x) d\nu(x) \right|, \tag{2.3}$$

where the supremum is over all $f : \mathbb{R} \to \mathbb{R}$ such that $|f(x) - f(y)| \leqslant |x - y|$ for any $x, y \in \mathbb{R}$. Note that if, for example, Y is a random variable with the standard normal distribution $\mathcal{N}(0, 1)$, then for any random variable X, the usual Kolmogorov distance can be bounded by

$$d_{\mathrm{K}}\big(\mathcal{L}(X), \mathcal{L}(Y)\big) := \sup_{a \in \mathbb{R}}\big|\mathbb{P}[X \leqslant a] - \mathbb{P}[Y \leqslant a]\big| = O\big(\sqrt{d_{\mathrm{W}}\left(\mathcal{L}(X), \mathcal{L}(Y)\right)}\big). \tag{2.4}$$

(see e.g. [[17], Proposition 2.1]).

2.2. Further notation

We sometimes tacitly assume that n is not too small.

$\xrightarrow{\text{a.s.}}$ and $\xrightarrow{\text{P}}$ denote convergence almost surely (a.s.) and in probability, respectively.

$\mathrm{Po}(\mu)$ denotes the Poisson distribution with mean μ, and $\mathcal{N}(0,1)$ is the standard normal distribution.

C denotes constants that may vary from one occurrence to the next. They do not depend on n, but they may depend on m, H, and other parameters.

3. Main results

Consider first subgraphs that are cycles. The first two theorems show that when n is large, the count of any cycle with a fixed number of edges is approximately Poisson, and that the joint distribution of the counts of cycles of different numbers of edges can be approximated by independent Poisson variables. For convenience, we refer to a cycle with ℓ edges and vertices simply as an ℓ-cycle. Note also we view a pair of parallel edges as a 2-cycle.

Theorem 3.1. *Fixing the positive integers $m \geq 2$ and $\ell \geqslant 2$, let W_n be the number of ℓ-cycles in G_n^m, and let $\mu_n := \mathbb{E}W_n$. Then,*

$$\mu_n = \Theta(\log n). \tag{3.1}$$

and there is a positive constant $C = C(m, \ell)$ such that

$$d_{\mathrm{TV}}\left(\mathcal{L}(W_n), \mathrm{Po}(\mu_n)\right) \leqslant C \log^{-1} n. \tag{3.2}$$

Remark 3.2. Let W_n be as above and $Y_n' \sim \mathrm{Po}(\mu_n)$ and define

$$Z_n := \frac{W_n - \mu_n}{\sqrt{\mu_n}}, \qquad Y_n := \frac{Y_n' - \mu_n}{\sqrt{\mu_n}}. \tag{3.3}$$

It follows from Theorem 3.1 that

$$d_{\mathrm{K}}\big(\mathcal{L}(Z_n), \mathcal{L}(Y_n)\big) \leqslant d_{\mathrm{TV}}\big(\mathcal{L}(Z_n), \mathcal{L}(Y_n)\big) \leqslant C \log^{-1} n. \tag{3.4}$$

On the other hand, the classical Berry–Esseen theorem implies that

$$d_{\mathrm{K}}\big(\mathcal{L}(Y_n), \mathcal{N}(0,1)\big) \leqslant C\mu_n^{-\frac{1}{2}}, \tag{3.5}$$

Combining (3.4) and (3.5) with the triangle inequality, and using (3.1), we obtain

$$d_{\mathrm{K}}\big(\mathcal{L}(Z_n), \mathcal{N}(0,1)\big) \leqslant C \log^{-\frac{1}{2}} n. \tag{3.6}$$

In particular, the cycle count W_n is asymptotically normal. In fact, the estimate (3.6) is sharp. Since $\mathcal{N}(0,1)$ has a continuous distribution function, while the distribution function of Z_n has a jump $\mathbb{P}(W_n = k)$ (if this is non-zero) at $(k - \mu_n)/\sqrt{\mu_n}$, it follows by choosing $k = \lfloor \mu_n \rfloor$ and using (3.2) and (3.1) that

$$d_{\mathrm{K}}\big(\mathcal{L}(Z_n), \mathcal{N}(0,1)\big) \geq \tfrac{1}{2}\mathbb{P}\big(W_n = \lfloor \mu_n \rfloor\big) = \Theta\big(\mu_n^{-\frac{1}{2}}\big) = \Theta\big(\log^{-\frac{1}{2}} n\big). \tag{3.7}$$

Consequently, combining (3.5) and (3.7),

$$d_{\mathrm{K}}\big(\mathcal{L}(Z_n), \mathcal{N}(0,1)\big) = \Theta\big(\log^{-\frac{1}{2}} n\big). \tag{3.8}$$

In the following corresponding theorem on multi-variate Poisson approximation, the error bound that we obtain is slightly inferior to the one in Theorem 3.1. As in Remark 3.2, this theorem implies also a multivariate normal approximation (we omit the details).

Theorem 3.3. *Fix the positive integers $m, r \geq 2$ and $\ell(i) \in [2, \infty)$, $i \in [r]$, with $\ell(i) \neq \ell(j)$ for $i \neq j$. Let $W_n^{(i)}$ be the number of cycles of length $\ell(i)$ in G_n^m, and let $\mu_{n,i} := \mathbb{E} W_n^{(i)}$. Then, there is a positive constant $C = C(m, (\ell(i))_{i \in [r]})$ such that*

$$d_{\mathrm{TV}}\left(\mathcal{L}\left(\left\{W_n^{(i)}\right\}_{i=1}^r\right), \prod_{i=1}^r \mathrm{Po}(\mu_{n,i})\right) \leqslant \frac{C \log \log n}{\log n}. \tag{3.9}$$

Example 3.4. *The case $\ell = 2$ is simple in Theorems 3.1 and 3.3, since the numbers of pairs of parallel edges with larger endpoint j are independent for different j; moreover, we have the exact formula $\mu_n = \binom{m}{2} \sum_{j=2}^n \frac{1}{j-1}$. For $\ell = 3$, Example 2.1 yields*

$$\mu_n = m^2(m-1) \sum_{1 \leqslant i < j < k \leqslant n} \frac{1}{(j-1)(k-1)^2} = m^2(m-1)\big(\log n + O(1)\big). \tag{3.10}$$

Next, we state the normal approximation for the count of any *unicyclic* graph, i.e., a graph that contains exactly one cycle. For completeness, we include cycles in the theorem below. Let $\mathcal{C}_\ell$ be an ℓ-cycle, let $s \geq 0$, and let, for $i = 1, \ldots, s$, $\mathcal{T}_i$ be a tree with t_i edges and a distinguished root, so that $\mathcal{T}_i$ has $t_i + 1$ vertices. We consider a graph $\Lambda := \Lambda_{\ell, t_1, \ldots, t_s}$ which can be constructed by attaching each $\mathcal{T}_i$ to $\mathcal{C}_\ell$, using a vertex of $\mathcal{C}_\ell$ as the distinguished root of $\mathcal{T}_i$. (This does not specify Λ uniquely, but we choose one possibility.) By combining the trees $\mathcal{T}_i$ that are attached at the same vertex in the cycle, we may assume that $s \leqslant \ell$ and that each $\mathcal{T}_i$ is attached to $\mathcal{C}_\ell$ at a distinct vertex. Denote by $\Gamma = \Gamma_\Lambda^{(n,m)}$ the set of all potential copies of Λ; and saving notation, let

$$W_n := \sum_{\alpha \in \Gamma} \mathbf{1}_\alpha, \quad \mu_n := \mathbb{E} W_n, \quad \sigma_n^2 := \mathrm{Var}(W_n). \tag{3.11}$$

Theorem 3.5. *Fixing the integers $m \geq 2$, $\ell \geq 2$, $s \geq 0$, $t_1, \ldots, t_s \geq 1$, and a unicyclic subgraph Λ as above, let W_n, μ_n and σ_n be as in (3.11), and $Y_n := (W_n - \mu_n)/\sigma_n$. Let $t = \sum_{i=1}^s t_i \geq 0$; then*

$$\mu_n = \Theta\big(\log^{t+1} n\big). \tag{3.12}$$

$$\sigma_n^2 = \Theta\big(\log^{2t+1} n\big), \tag{3.13}$$

and there is a positive constant $C = C(m, \ell, t)$ such that

$$d_{\mathrm{W}}\big(\mathcal{L}(Y_n), \mathcal{N}(0,1)\big) \leqslant C \log^{-\frac{1}{2}} n. \tag{3.14}$$

Remark 3.6. In view of (2.4) and (3.14), Theorem 3.5 implies an error bound

$$d_{\mathrm{K}}\big(\mathcal{L}(Y_n), \mathcal{N}(0,1)\big) = O\big(\log^{-\frac{1}{4}} n\big) \tag{3.15}$$

for the Kolmogorov distance, which in the case of a cycle is clearly not as sharp as the error bound in (3.6).

The next theorem concerns trees, and the normal approximation result relies on the assumption that the variance of the counts of the tree of choice is of the exact order $\Theta(n)$. The precise order of the variance in this case is much harder to establish, essentially due to the fact that the total covariance of the positively correlated pairs of copies and that of the negatively correlated pairs of copies are both of order $O(n)$. However, we are able to prove that the variance of the

count of a star is precisely $\Theta(n)$. Below, let Λ be a fixed tree on t vertices, let Γ be the set of potential copies of Λ, and define again W_n, μ_n, and σ_n^2 by (3.11).

Theorem 3.7. *Fix the positive integers m, t and the tree Λ with t vertices. Let W_n, μ_n and σ_n be as above, and $Y_n := (W_n - \mu_n)/\sigma_n$. Then*

$$\mu_n = \Theta(n), \tag{3.16}$$

$$\sigma_n^2 = O(n). \tag{3.17}$$

If $\sigma_n^2 = \Theta(n)$, then there is a constant $C = C(m, t)$ such that

$$d_W \left(\mathcal{L}(Y_n), \mathcal{N}(0, 1) \right) \leqslant C n^{-\frac{1}{2}}. \tag{3.18}$$

In the trivial cases $t = 1$ and $t = 2$ (a single vertex and a single edge, respectively), W_n is deterministic and $\sigma_n^2 = 0$; thus Y_n is not even defined. We conjecture that these are the only cases where (3.18) does not hold.

Conjecture 3.8. *If Λ is tree with at least 2 edges, then $\sigma_n^2 = \Theta(n)$, and thus the normal approximation (3.18) holds.*

We show in Section 9.2 that this holds at least for stars.

Theorem 3.9. *Conjecture 3.8 holds when Λ is a star S_ℓ with $\ell \geq 2$ edges.*

The remaining subgraphs are the multicyclic ones. We consider in particular the ones that have the following properties.

Definition 3.10. *We say that a graph is*

- *multicyclic if it has at least two (not necessarily edge- or vertex-disjoint) cycles;*
- *leaf-free if it has no node of degree 1.*

As examples of connected graphs that are both multicyclic and leaf-free, we have two edge-disjoint cycles joined by an edge, or two cycles that share precisely an edge; see Figure 2 for another example.

Theorem 3.11. *For any connected graph H that is both multicyclic and leaf-free, and any $m \geqslant 2$, the expected number of copies of H in G_n^m is bounded as $n \to \infty$.*

Remark 3.12. We may define the infinite random graph G_∞^m with vertex set $\{1, 2, \ldots\}$ as the union of G_n^m over all $n \geq 1$. Let H be any fixed graph and let as above W_n be the number of copies of H in G_n^m; also, let $W_\infty \leq \infty$ be the number of copies of H in G_∞^m. Then, as $n \to \infty$, $W_n \nearrow W_\infty$. In particular, $W_n \xrightarrow{\text{p}} W_\infty$, and since both (3.12)–(3.13) and (3.16)–(3.17) imply $W_n \xrightarrow{\text{p}} +\infty$, we see that if H is unicyclic or a tree, then $W_\infty = \infty$ a.s. On the other hand, Theorem 3.11 implies by monotone convergence that if H is multicyclic and leaf-free, then $\mathbb{E} W_\infty < \infty$ and thus $W_\infty < \infty$ a.s.

Note further that since $W_n \xrightarrow{\text{a.s.}} W_\infty$ as $n \to \infty$, in particular, W_n converges in distribution to W_∞ (without any normalisation). However, we do not expect that this limiting distribution has any nice form such as Poisson, since W_∞ is mainly determined by the random wirings of the first few edges in G_n^m.

We also consider the expected counts of a graph that are multicyclic but not leaf-free. Note that every such graph H can be constructed in the following way. Start with a graph H' that is both multicyclic and leaf-free, and let $\mathcal{T}_i$, $i = 1, \ldots, s$ be trees with t_i edges and a distinguished root. The graph H is obtained by attaching every $\mathcal{T}_i$ to H', with one of the vertices of H' as the
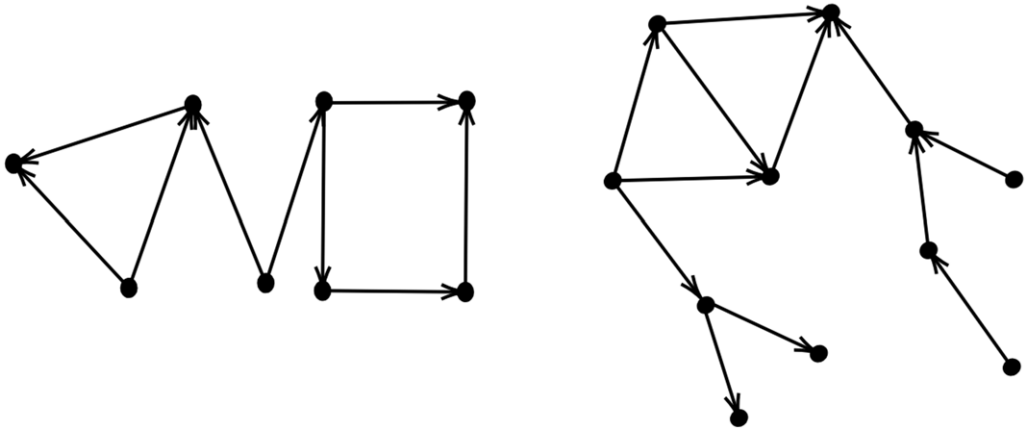
Figure 2. *Left*: a graph that is both multicyclic and leaf-free. *Right*: a graph that is multicyclic but not leaf-free.

distinguished root of $\mathcal{T}_i$. As before, we may assume that any pair of $\mathcal{T}_i$ do not share the same vertex of H'. The graph H' is known as the 2-*core* of H. See Figure 2 for an example.

In the following theorem, we for convenience include the case $t = 0$, which is just Theorem 3.11. (Then $s = 0$ and $H' = H$.) Recall the definition of m_H in Remark 2.2, and that $W_n = 0$ if $m < m_H$.

Theorem 3.13. *Fixing a connected multicyclic graph H and a positive integer $m \geqslant m_H$, let W_n be the number of copies of H in G_n^m, and let $t := \sum_i t_i \geq 0$ be the number of vertices in $H \setminus H'$, where H' as above is the 2-core of H. Then,*

$$\mathbb{E}W_n = \Theta\big(\log^t n\big), \tag{3.19}$$

$$\mathrm{Var}(W_n) = O\big(\log^{2t} n\big). \tag{3.20}$$

The precise order of $\mathrm{Var}(W_n)$ is more difficult to establish for the same reason as for trees. For this family of graphs, we do not expect the count to be approximately Poisson or normal (after renormalisation). Instead, we make the following conjecture, where we, as commented above, do not expect the distribution of the limit to have a nice form.

Conjecture 3.14. *Let H and H' be as in Theorem 3.13, and let W_n and W_n' be the numbers of copies of H and H' in G_n^m, respectively. Then there exists a constant $c > 0$ such that $(\log n)^{-t} W_n - cW_n' \xrightarrow{\text{a.s.}} 0$ as $n \to \infty$, and thus*

$$\frac{W_n}{\log^t n} \xrightarrow{\text{a.s.}} cW_\infty'. \tag{3.21}$$

We can prove a special case. It appears likely that the general case can be shown by similar arguments, but the details seem complicated and we leave this as an open problem.

Theorem 3.15. *Conjecture 3.14 holds when $t = 1$.*

We note also that at least for some multicyclic graphs H', $W_\infty' = 0$ with positive probability, and thus $\mathbb{P}(W_\infty = 0) \geqslant \mathbb{P}(W_\infty' = 0) > 0$; hence although $\mathbb{E}W_n \to \infty$ by (3.19), $\mathbb{P}(W_n = 0) \geqslant c > 0$, which in particular shows that we cannot have Poisson or normal convergence. We do not know whether this holds for every simple multigraph H', and we give just one example.

Example 3.16. *Let H' be the complete graph on 4 vertices, K_4, and construct H by adding a single edge to H'; thus we take $s = 1$ and $\mathcal{T}_1$ as an edge in the construction above. Let W_n and W'_n be as above. By Theorem 3.11 (or a simple calculation), we see that a.s. $W'_n \to W'_\infty < \infty$ as $n \to \infty$. Moreover, we claim that $\mathbb{P}(W'_\infty = 0) > 0$.*

To see this, let $H'' = K_4^-$ be K_4 minus one edge, and let W''_n be the number of copies of K_4^- in G_n^m. Fix a large N, and consider only $n > N$. Let $\mathcal{E}'_N$ be the event that there is a triangle in G_N^m. Let $\mathcal{E}''_{N,n}$ be the event that there is a copy of K_4 in $G_n^m \cup K_N^m$ with at most 2 vertices in $[N]$. In other words $\mathcal{E}''_{N,n}$ means that there is either a copy of K_4 in G_n^m, or a copy of K_4^- in G_n^m with the two non-adjacent vertices in $[N]$ and the two others in $[n] \setminus [N]$. Note that $\mathcal{E}''_{N,n}$ does not depend on the edges of G_N^m, and thus $\mathcal{E}'_N$ and $\mathcal{E}''_{N,n}$ are independent. If there is a copy of K_4 in G_n^m, then either $\mathcal{E}'_N$ or $\mathcal{E}''_{N,n}$ holds, and thus

$$\mathbb{P}(W'_n = 0) \geqslant \mathbb{P}(\text{not } \mathcal{E}'_N)\mathbb{P}(\text{not } \mathcal{E}''_{N,n}). \tag{3.22}$$

If $\mathcal{E}''_{N,n}$ holds, then there is a copy of $H'' = K_4^-$ in G_n^m with at least two vertices in $[n] \setminus [N]$, and thus $W''_\infty \geqslant W''_n \geqslant W''_N + 1$. Hence, by Markov's inequality,

$$\mathbb{P}(\mathcal{E}''_{N,n}) \leq \mathbb{E}(W''_\infty - W''_N) = \mathbb{E}W''_\infty - \mathbb{E}W''_N. \tag{3.23}$$

Since $\mathbb{E}W''_N \to \mathbb{E}W''_\infty$ as $N \to \infty$, we can choose N such that $\mathbb{E}W''_\infty - \mathbb{E}W''_N \leq \frac{1}{2}$; then (3.23) implies that $\mathbb{P}(\text{not } \mathcal{E}''_{N,n}) \geqslant \frac{1}{2}$ for all $n > N$. Moreover, $\mathbb{P}(\text{not } \mathcal{E}'_N) > 0$, since there is a positive probability that all edges in G_N^m lead to 1, and then there is no triangle. Consequently, (3.22) yields $\mathbb{P}(W'_n = 0) \geqslant c$, for some $c > 0$ that does not depend on $n > N$; thus also $\mathbb{P}(W'_\infty = 0) \geqslant c > 0$.

If there are no copies of H', then there can be no copies of H, and thus we conclude $\mathbb{P}(W_n = 0) \geqslant c > 0$ for all n. Hence, although $\mathbb{E}W_n \to \infty$ by (3.19), W_n does not converge in probability to $+\infty$, and in particular W_n cannot be asymptotically normal.

3.1. Discussion on possible future avenues

An important direction for future work is to verify if Conjecture 3.8 holds, or in other words, to prove that the variance of the count of any tree with at least 2 edges is of the precise order n. Another direction for future work is to compute the leading coefficients of the means and variances of the subgraph counts. More precise expressions will possibly also enable us to verify if the approximation results still hold if m and (or) the number of vertices of the subgraph are allowed to increase with n. It is also possible to prove a multivariate analogue for the normal approximation results. This can be done using, for instance, [[4], Theorem 12.1], which also uses Stein's method with the size-bias coupling.

3.2. Article outline

In the next section, we state the results from Stein's method that we apply in the approximation proofs. These results use a coupling that we construct in Section 5. In Section 6, we prove Theorem 3.11 and some additional lemmas that will be useful in the approximation proofs later. We prove the Poisson approximation results (Theorems 3.1 and 3.3) in Section 7 and the normal approximation for unicyclic graphs (Theorem 3.5) in Section 8, where we also prove Theorem 3.13. Section 9 contains the proofs of the normal approximation for trees (Theorems 3.7 and 3.9), and in the last section we give a proof of Theorem 3.15.

4. Preliminary: Stein's method

The error bounds in the Poisson and normal approximation results that we use are obtained from general results on Stein's method in terms of a coupling that we now describe. Let $(I_\alpha)_{\alpha \in \Gamma}$ be a

collection of 0-1 valued random variables. For each $\alpha \in \Gamma$, let the random variables $(J_{\beta\alpha})_{\beta\in\Gamma}$ be defined on the same space as $(I_\alpha)_{\alpha\in\Gamma}$, satisfying

$$\mathcal{L}(J_{\beta\alpha}; \beta \in \Gamma \setminus \{\alpha\}) = \mathcal{L}(I_\beta; \beta \in \Gamma \setminus \{\alpha\} \mid I_\alpha = 1). \tag{4.1}$$

Note that this is a special case of the size-bias coupling appearing in the literature of Stein's method; see [4, 17], and also [12]. The theorem below is a direct consequence of [[2], Chapter 2, equation (1.2)].

Theorem 4.1. *Let $(I_\alpha)_{\alpha\in\Gamma}$ be as above, where $\mathbb{E}I_\alpha = \pi_\alpha$. Suppose that for each $\alpha \in \Gamma$, there is a coupling of $(J_{\beta\alpha})_{\beta\in\Gamma}$ and $(I_\alpha)_{\alpha\in\Gamma}$ such that (4.1) holds. Let $W := \sum_{\alpha\in\Gamma} I_\alpha$, $\lambda := \sum_{\alpha\in\Gamma} \pi_\alpha$. Then,*

$$d_{\mathrm{TV}}\left(\mathcal{L}(W), \mathrm{Po}(\lambda)\right) \leqslant \min\left\{1, \lambda^{-1}\right\} \left(\sum_{\alpha\in\Gamma} \pi_\alpha^2 + \sum_{\alpha\in\Gamma}\sum_{\beta\neq\alpha} \pi_\alpha \mathbb{E}|J_{\beta\alpha} - I_\beta| \right). \tag{4.2}$$

Now, assume that there is a partition $\Gamma = \bigcup_{j=1}^{r} \Gamma_j$. The next theorem quantifies how well the random vector $(\sum_{\alpha\in\Gamma_j} I_\alpha)_{1\leqslant j\leqslant r}$ can be approximated by a vector of independent Poisson variables. In our application, Γ_j is to be taken as the set of potential cycles of length $\ell(j)$.

Theorem 4.2 ([[2], Theorem 10.K]). *Let $W_j = \sum_{\alpha\in\Gamma_j} I_\alpha$, $\mathbb{E}I_\alpha = \pi_\alpha$ and $\lambda_j = \mathbb{E}W_j$. Suppose that for each $\alpha \in \Gamma$, there is a coupling of $(J_{\beta\alpha})_{\beta\in\Gamma}$ and $(I_\alpha)_{\alpha\in\Gamma}$ such that (4.1) holds. Then,*

$$d_{\mathrm{TV}}\left(\mathcal{L}(\{W_j\}_{j=1}^{r}), \prod_{j=1}^{r} \mathrm{Po}(\lambda_j) \right) \leqslant \frac{1 + 2\log^+ (e\min_j \lambda_j)}{e\min_j \lambda_j} \sum_{\alpha\in\Gamma} \left(\pi_\alpha^2 + \sum_{\beta\neq\alpha} \pi_\alpha \mathbb{E}|J_{\beta\alpha} - I_\beta| \right). \tag{4.3}$$

We now state a normal approximation result for a collection of 0-1 variables, which follows from [[17], Theorem 3.20 and Corollary 3.24].

Theorem 4.3. *Let $(I_\alpha)_{\alpha\in\Gamma}$ be a collection of 0-1 variables and let $(J_{\beta\alpha})_{\beta\in\Gamma}$ be defined on the same space as $(I_\alpha)_{\alpha\in\Gamma}$, satisfying (4.1). Define $W := \sum_{\alpha\in\Gamma} I_\alpha$, $\mu := \mathbb{E}W$, $\sigma^2 := \mathrm{Var}(W)$ and $W^s := \sum_{\beta\in\Gamma\setminus\{K\}} J_{\beta K} + 1$, where the index $K \in \Gamma$ is chosen randomly with probabilities $\mathbb{P}(K=\alpha) = \mathbb{E}I_\alpha/\mu$. If $Z = (W - \mu)/\sigma$, then*

$$d_{\mathrm{W}}\left(\mathcal{L}(Z), \mathcal{N}(0,1)\right) \leqslant \frac{\mu}{\sigma^2}\sqrt{\frac{2}{\pi}}\sqrt{\mathrm{Var}\left(\mathbb{E}[W^s - W \mid W]\right)} + \frac{\mu}{\sigma^3}\mathbb{E}[(W^s - W)^2], \tag{4.4}$$

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

5. Construction of the size-bias couplings

As the edges of a uniform attachment graph are independent, we can construct the coupling appearing in Theorems 4.1–4.3 as follows. Fix the subgraph H, let $G := G_n^m$ be the uniform attachment graph and let Γ be the set of potential copies of H. For every $\alpha \in \Gamma$, we couple two graphs G and G^α by matching their attachment steps, except for the edges of α. In G^α, the edges of α are wired in a deterministic fashion to obtain the copy α of H; whereas in G, they are generated independently from the construction of G^α. For $\beta \in \Gamma$, let as above $\mathbf{1}_\beta$ be the indicator that the subgraph β is present in G, and let $\mathbf{1}_\beta^\alpha$ be the indicator that β is present in G^α. It follows that for every chosen α,

$$\mathcal{L}(\mathbf{1}_\beta^\alpha; \beta \in \Gamma \setminus \{\alpha\}) = \mathcal{L}(\mathbf{1}_\beta; \beta \in \Gamma \setminus \{\alpha\} \mid \mathbf{1}_\alpha = 1). \tag{5.1}$$

Let $\Gamma_\alpha^- \subset \Gamma$ be the set of copies β of H such that at least one edge of α has a different endpoint in β, and let $\Gamma_\alpha^+ \subset \Gamma \setminus \Gamma_\alpha^-$ be the set of copies of H that share at least one edge with α, excluding α itself. Observe that under the coupling, $\mathbf{1}_\beta^\alpha \leqslant \mathbf{1}_\beta$ for $\beta \in \Gamma_\alpha^-$ and $\mathbf{1}_\beta^\alpha \geqslant \mathbf{1}_\beta$ for $\beta \in \Gamma_\alpha^+$, while $\mathbf{1}_\beta^\alpha = \mathbf{1}_\beta$ for every $\beta \in \Gamma \setminus (\Gamma_\alpha^+ \cup \Gamma_\alpha^- \cup \{\alpha\})$. The error bound (4.2) in Theorem 4.1 therefore simplifies to

$$\min\{1, \lambda^{-1}\} \left(\sum_{\alpha \in \Gamma} \pi_\alpha^2 + \sum_{\alpha \in \Gamma} \sum_{\beta \in \Gamma_\alpha^+} \pi_\alpha \mathbb{E}[\mathbf{1}_\beta^\alpha - \mathbf{1}_\beta] + \sum_{\alpha \in \Gamma} \sum_{\beta \in \Gamma_\alpha^-} \pi_\alpha \mathbb{E}[\mathbf{1}_\beta - \mathbf{1}_\beta^\alpha] \right); \qquad (5.2)$$

noting that $\sum_{\beta \in \Gamma_\alpha^+} \mathbb{E}[\mathbf{1}_\beta^\alpha - \mathbf{1}_\beta]$ and $\sum_{\beta \in \Gamma_\alpha^-} \pi_\alpha \mathbb{E}[\mathbf{1}_\beta - \mathbf{1}_\beta^\alpha]$ are, respectively, the expected gain and loss in the copies of H after forcing α to be present in the graph G. We may simplify this further by omitting the negative terms in (5.2) and noting that, by (5.1),

$$\pi_\alpha \mathbb{E}\left[\mathbf{1}_\beta^\alpha\right] = \mathbb{P}(\mathbf{1}_\alpha = 1) \mathbb{E}\left[\mathbf{1}_\beta \mid \mathbf{1}_\alpha = 1\right] = \mathbb{E}\left[\mathbf{1}_\beta \mathbf{1}_\alpha\right], \qquad (5.3)$$

which yields the simpler (but somewhat larger) error bound

$$\min\{1, \lambda^{-1}\} \left(\sum_{\alpha \in \Gamma} \pi_\alpha^2 + \sum_{\alpha \in \Gamma} \sum_{\beta \in \Gamma_\alpha^+} \mathbb{E}\left[\mathbf{1}_\alpha \mathbf{1}_\beta\right] + \sum_{\alpha \in \Gamma} \sum_{\beta \in \Gamma_\alpha^-} \pi_\alpha \pi_\beta \right). \qquad (5.4)$$

As for Theorem 4.2, the simplified error bound is the same as (5.2) or (5.4), but with the factor $\min\{1, \lambda^{-1}\}$ replaced with

$$\frac{1 + 2\log^+ (e \min_j \lambda_j)}{e \min_j \lambda_j}. \qquad (5.5)$$

When applying Theorem 4.3, we first sample a copy $K \in \Gamma$ of subgraph H with probabilities $\mathbb{P}(K = \alpha)$ proportional to $\mathbb{E}\mathbf{1}_\alpha$, and construct the graphs G^K and G as above. The subgraph count W and its size-bias version W^s can then be found in G and G^K, respectively.

6. Proof of Theorem 3.11 and some useful lemmas

Here we prove some lemmas that are useful for proving the main results later. We also prove Theorem 3.11 here, as some special cases of the result will be applied in the other proofs. To study the expected number of copies of a graph H, we use the following definition. Let h be the number of vertices in H.

Definition 6.1 (Vertex marks and mark sequence). *Let α be a potential copy of H and suppose that its vertices are $k_1 < \ldots < k_h$. We say that the mark of a vertex $k_i \in \alpha$ is the out-degree of k_i in α regarded as a directed graph (as always, with edges directed towards the smaller endpoint). The mark sequence is $(b_i)_{1 \leqslant i \leqslant h}$, with b_i being the mark of vertex k_i and for convenience, we occasionally refer to $b_i = k$ as a k-mark.*

In other words, for α to actually be a copy of H in G_n^m, there are for each $i \in [h]$ exactly b_i edges in G_n^m from k_i that have to have the endpoints determined by α. Note that the mark sequence does not entirely encode the configuration of the copy of H in G_n^m, but, together with the sequence of vertices k_i, the mark sequence gives the probability that the copy α is present in G_n^m. In fact, since the edges in G_n^m choose their endpoints independently,

$$\pi_\alpha := \mathbb{P}(\mathbf{1}_\alpha = 1) = \prod_{i=1}^h \frac{1}{(k_i - 1)^{b_i}} = \prod_{i=2}^h \frac{1}{(k_i - 1)^{b_i}}, \qquad (6.1)$$

where the final equality holds because always $b_1 = 0$, since edges are directed towards lower numbers, and thus α has no out-edge from k_1. Similarly, b_h equals the degree of k_h in α, since k_h has no in-edge.

Note that for a given unlabelled graph H, there is a finite number (at most $h!$) of non-isomorphic labelled versions of H. Hence, in order to obtain estimates of the expected number of copies, it suffices to consider each possible labelled version separately, and then consider only potential copies with vertex sequences $(k_i)_1^h$ where $k_1 < \ldots < k_h$ and k_i corresponds to vertex $i \in H$. The mark sequence of the potential copy depends only on the labelled version of H. Different labelled versions of H may yield different mark sequences $(b_i)_1^h$, but they all have the same length $h = v(H)$, the number of vertices in H, and the same sum

$$\sum_{i=1}^{h} b_i = e(H), \tag{6.2}$$

the number of edges in H.

For the proof of Theorem 3.11, we also use the following definition.

Definition 6.2 (*F*-number). *We define the F-number, $F(a_i)$ of a finite sequence (a_i) of natural numbers as $\sum_{a_i > 1} a_i + \sum_{a_i = 0} (-2)$, that is, we ignore all a_i with $a_i = 1$, count the a_i with $a_i = 0$ with weight -2 and count all other a_i with a_i as their weight.*

Proof of Theorem 3.11. Consider a connected graph H that is multicyclic and leaf-free, see Definition 3.10. Let $(b_i)_{1 \leqslant i \leqslant h}$ be the mark sequence of a potential copy of H. By (6.1), the expected number of copies of H with this mark sequence can be bounded by a constant times

$$S(b_1 \cdots b_h) := \sum_{1 \leqslant k_1 < \ldots < k_h < \infty} \prod_{i=2}^{h} \frac{1}{(k_i - 1)^{b_i}}. \tag{6.3}$$

Hence it suffices to show that this infinite sum is finite for every mark sequence $(b_i)_1^h$.

We will show this by modifying the mark sequence (b_i) in such a way that preserves the sum in (6.2) and increases the sum $S(b_1 \cdots b_h)$ in (6.3) until we reach a sequence where we can show that $S(b_1 \cdots b_h)$ is finite. Note that we do not claim that the modified sequences actually are mark sequences for some copies of H; we consider the value $S(b_1 \cdots b_h)$ as defined by (6.3), without worrying about a probabilistic interpretation in general.

Since H is leaf-free, each vertex in H has degree at least 2, and thus $b_h > 1$. We change the sequence $(b_i)_1^h$ as follows: In the first round, if $b_h > 2$ and $b_i = 0$ for some $i \neq 1$, then decrease b_h by 1 and increase the last such b_i (i.e., the one with maximal i) by 1. For instance, $0102023 \rightarrow 0102122$. We repeat this process until $b_h = 2$ or $b_i > 0$ for all $i \neq 1$. In the second round, if $b_h = 2$ and $b_{h-1} > 1$, then we repeat the same procedure with b_{h-1} until $b_{h-1} = 1$ or we have exhausted all b_i, $i \neq 1$, such that $b_i = 0$. Continue in the same way: in the pth round ($3 \leqslant p < h$), if $b_h = 2$, $b_{h-1} = \cdots = b_{h-p+2} = 1$, and $b_{h-p+1} > 1$, then repeat the procedure for b_{h-p+1} as we did for b_{h-1}. Stop when such moves are no longer possible. Denote the final sequence by $(b_i')_1^h$. We emphasise that these moves do not change the sum in (6.2) and never decrease $\prod_{i=1}^{h} (k_i - 1)^{-b_i}$, since we always shift mass in the mark sequence to the left; thus $S(b_1 \cdots b_h) \leqslant S(b_1' \cdots b_h')$.

All the possible final sequences $(b_i')_1^h$ are analysed below:

1. We have at least one intermediate zero left, say $b_q' = 0$ with $1 < q < h$ and q maximal among such indices. Then $b_{q+1}' = \cdots = b_{h-1}' = 1$ and $b_h' = 2$, since otherwise we would have modified the sequence further. In other words, we have a sequence of the form $0XXX \cdots XXX011 \cdots 112$. This is not possible, due to the *F*-number of the final part of

the sequence $b'_q \cdots b'_h = 0111 \cdots 1112$. Note that the moves that we made above cannot have affected b_i for $i < q$, so they were entirely in this subsequence, and it is easy to see that each move either preserved or increased the F-number of this subsequence. Thus $0 = F(0111 \cdots 1112) \geqslant F(b_q \cdots b_h)$. with $(b_i)_{1 \leqslant i \leqslant h}$ being the original mark sequence. A way of interpreting the F-number is as an upper bound of the number of edges directed from the last $h - q + 1$ nodes to the first $q - 1$ nodes, which equals the sum of the differences between outdegree and indegree for the last $h - q + 1$ nodes. In any copy of a connected graph H that is multicyclic and leaf-free, a vertex with mark 0 must receive at least 2 in-edges and thus contributes at most $a_i = -2$ to the count. A vertex i with mark 1 sends one edge but receives at least 1, and so contributes at most $a_i = 0$ to the count. Similarly, a vertex i with mark $k > 1$ may receive some in-edges and so contributes at most $a_i = k$ to the count. Since H is a connected graph, we must have some edge connecting the last $h - q + 1$ vertices and the first $q - 1$ vertices in H, and so $F(b_q \cdots b_h) > 0$, contradicting the above.

2. We have removed all the zeroes (except the first) and ended up with $0XX \cdots XXb'_h$ where $b'_h > 2$ and each X is at least 1. (In this case, we could not reduce the last number down to 2 before running out of zeroes.) The sequence can be compared to the sequence $0111 \cdots 1113$ which gives a larger or equal value of (6.3). For the latter sequence we have, if $h \geq 3$, by first summing over $k_h > k_{h-1}$,

$$S(011 \cdots 113) = \sum_{1 \leqslant k_1 < \ldots < k_h} \prod_{i=2}^{h-1} \frac{1}{k_i - 1} \cdot \frac{1}{(k_h - 1)^3}$$

$$\leqslant \sum_{1 \leqslant k_1 < \ldots < k_{h-1}} \prod_{i=2}^{h-2} \frac{1}{k_i - 1} \cdot \frac{1}{(k_{h-1} - 1)^3}, \tag{6.4}$$

which is the same sum with h replaced by $h - 1$; thus induction yields

$$S(011 \cdots 113) \leqslant S(03) = \sum_{1 \leqslant k_1 < k_2} \frac{1}{(k_2 - 1)^3} = \sum_{k_2 \geq 2} \frac{1}{(k_2 - 1)^2} < \infty. \tag{6.5}$$

3. We have $0XXX \cdots XXX2$ where the X's are at least 1. In this case, note that not all the X's can be 1, as H for topological reasons has more edges than vertices. Thus at least one X is at least 2. If $b'_2 = 1$, we can exchange b'_2 with b'_j for some $j < h$ such that $b'_j > 1$, again without decreasing S. We can then compare the sequence to $02111 \cdots 1112$. By arguing as in (6.4)–(6.5), we find

$$S(0211 \cdots 112) \leqslant S(022) = \sum_{1 \leqslant k_1 < k_2 < k_3} \frac{1}{(k_2 - 1)^2 (k_3 - 1)^2}$$

$$\leqslant \sum_{1 \leqslant k_1 < k_2} \frac{1}{(k_2 - 1)^3} = S(03) < \infty. \tag{6.6}$$

This completes the proof of the theorem. □

For the expected number of ℓ-cycles, which is not covered by Theorem 3.11, we use a simpler version of the argument above to prove the following lemma, which essentially says that it is enough to just consider one particular configuration of an ℓ-cycle. We first observe that in the mark sequence of a potential copy of a cycle, all marks are 0, 1, or 2; furthermore, there must be equal numbers of 0-marks and 2-marks by (6.2). An ℓ-cycle has ℓ marks. As noted above, we must have $b_1 = 0$ and $b_\ell = 2$, since edges are directed towards lower numbers. Note that an ℓ-cycle has at most $\lfloor \ell/2 \rfloor$ vertices of mark 0.

Lemma 6.3. *Among all potential copies of an ℓ-cycle on given vertices $k_1 < \ldots < k_\ell$, the configurations with the mark sequence $011 \cdots 112$ have the largest probability to occur in G_n^m.*

Proof. In a cycle, if a vertex sends j edges, it receives $2 - j$ edges. Consider the mark sequence $(b_i)_1^\ell$ of a potential copy of an ℓ-cycle. Since the F-number of the sequence of the last $k \leqslant \ell$ marks must be positive (by the earlier argument for proving Theorem 3.11), the number of 2-marks must always be larger than the number of 0-marks, except when $k = \ell$, in which case they are equal. Equivalently,

$$\sum_{\ell-k+1}^{\ell} (b_i - 1) \geqslant 0, \text{ with strict inequality when } 1 \leqslant k < \ell. \tag{6.7}$$

This holds since the cycle is connected, and so the total number of out-edges of the last k vertices must be larger than their total number of in-edges when $k < \ell$ ($k = \ell$ gives equality). The probability of the cycle appearing in G_n^m is given by (6.1). Consider all sequences $(b_i)_1^\ell$ with $b_i \in \{0, 1, 2\}$ such that (6.7) holds. We must have $b_1 = 0$ and $b_\ell = 2$. If there is another mark $b_j = 2$ with $j \neq \ell$ in the sequence, let j be the largest such index. Then there must also be an index $i \neq 1$ with $b_i = 0$; let i be the largest such index. We must have $i < j$, since otherwise (6.7) would be violated for $k = \ell - i + 1$. It is then easy to see that if we change both b_i and b_j to 1, then (6.7) still holds, and the value of (6.1) is increased. Consequently, the mark sequence with the largest probability is of the form $0111 \cdots 1112$. $\qquad \square$

The lemma below concerns trees and is used when we prove the normal approximation results.

Lemma 6.4. *Fix the positive integer t and let $\mathcal{T}$ be a rooted tree with t edges and thus $t + 1$ vertices. The following results concerning $\mathcal{T}$ hold.*

  (i) *Among all potential copies of $\mathcal{T}$ on a given vertex set, the configurations with the mark sequence $011 \cdots 11$ have the largest probability.*

  (ii) *Let N_x be the number of copies of $\mathcal{T}$ with vertex x as its distinguished root. Then $\mathbb{E}N_x = O(\log^t n)$, uniformly for $1 \leqslant x \leqslant n < \infty$.*

Proof. 1: Let $0b_2 \cdots b_{t+1}$ be the mark sequence of a potential copy of $\mathcal{T}$. Note first that if $2 \leqslant k \leqslant t + 1$, then we have, summing over the $t + 2 - k$ last vertices,

$$\sum_{i=k}^{t+1} b_i \geqslant t + 2 - k. \tag{6.8}$$

In fact, this sum is the number of edges with upper endpoint in $[k, t + 1]$, which equals the number of edges with at least one endpoint in $[k, t + 1]$, and in any connected graph, a proper subset of ℓ vertices is always adjacent to at least ℓ edges. (To see this, collapse all other vertices into one and pick a spanning tree with ℓ edges in the resulting graph.)

We argue similarly to the proof of Theorem 3.11, now modifing the mark sequence as follows. As long as some $b_i \geq 2$, we reduce the rightmost such b_i by 1 and increase the rightmost mark 0 to 1. We can never reach a sequence ending with a proper subsequence $011 \cdots 11$ with some $b_k = 0$ ($k \geq 2$) and $b_{k+1} = \cdots = b_{t+1} = 1$, because the first time this happens, all previous moves must have been inside $[k, \ldots, t + 1]$, so the sum in (6.8) has not changed and (6.8) still holds, a contradiction. Consequently, we do not stop until there is no mark 0 except b_1, but since the sum of all marks is t by (6.2) (and this sum is not changed), the final sequence is $011 \cdots 11$. As the procedure never decreases the probability (6.1), the mark sequence $011 \cdots 11$ indeed yields the largest probability.

2: We use 1 to obtain

$$\mathbb{E}N_x \leqslant C \left(\sum_{x < k_1 < \ldots < k_t \leqslant n} \prod_{i=1}^{t} \frac{1}{k_i} + \sum_{r=1}^{t} \sum_{k_1 < \ldots < k_r < x < k_{r+1} < \ldots < k_t \leqslant n} \frac{1}{x} \prod_{i=2}^{t} \frac{1}{k_i} \right)$$

$$\leqslant C \left(\sum_{x < k_1 < \ldots < k_t \leqslant n} \prod_{i=1}^{t} \frac{1}{k_i} + \sum_{r=1}^{t} \sum_{k_1 < \ldots < k_r < x < k_{r+1} < \ldots < k_t \leqslant n} \prod_{i=1}^{t} \frac{1}{k_i} \right)$$

$$\leqslant C \prod_{i=1}^{t} \sum_{k_i=1}^{n} \frac{1}{k_i} \leqslant C (\log n + 1)^t, \tag{6.9}$$

which completes the proof. $\qquad\square$

7. Proof of the Poisson approximations for cycles

In this section we prove the Poisson approximation results in Theorems 3.1 and 3.3 using Theorems 3.11, 4.1 and 4.2. We fix in this section the integers $m \geq 1$ and $\ell \geq 2$ and denote by W_n the count of ℓ-cycles in G_n^m.

Proof of Theorems 3.1 and 3.3. We start by proving the result on the expected number of cycles μ_n in (3.1). By Lemma 6.3 and (6.1), μ_n can be bounded above by

$$\psi_{\ell,m} \sum_{k_1 < \ldots < k_\ell \leqslant n} \frac{1}{(k_\ell - 1)^2} \prod_{i=2}^{\ell-1} \frac{1}{k_i - 1}, \tag{7.1}$$

where $\psi_{\ell,m}$ is the total number of ways of forming a potential ℓ-cycle on ℓ given vertices. Summing over $k_1, \ldots, k_{\ell-1}$ (in this order), the above is at most $\psi_{\ell,m} \sum_{k_\ell=\ell}^{n} \frac{1}{k_\ell - 1}$, which can be bounded by $\psi_{\ell,m}(\log n + 1)$. The calculation for the lower bound is similar, with the factor $\psi_{\ell,m}$ replaced by 1; we consider just one configuration that has mark sequence $011 \cdots 11$.

Next, we construct the size-bias coupling of the cycle count described in Section 5. To prove the approximation result in Theorem 3.1, it then suffices to show that the sums in (5.4) are bounded as $n \to \infty$; the error bound then follows from (5.4) and (3.1).

Let Γ be the set of potential cycles and let as in (6.1) π_α be the probability of cycle $\alpha \in \Gamma$ being a subgraph of G_n^m. By (6.1) and Lemma 6.3, using the notation (6.3),

$$\sum_{\alpha \in \Gamma} \pi_\alpha^2 \leqslant C \sum_{k_1 < \ldots < k_\ell \leqslant n} k_\ell^{-4} \prod_{i=2}^{\ell-1} k_i^{-2} \leqslant C \cdot S(022 \cdots 224) \leqslant C \cdot S(011 \cdots 113), \tag{7.2}$$

which is finite by (6.5).

In the sum

$$\sum_{\alpha \in \Gamma} \sum_{\beta \in \Gamma_\alpha^+} \mathbb{E}\left[\mathbf{1}_\alpha \mathbf{1}_\beta\right], \tag{7.3}$$

each term $\mathbb{E}[\mathbf{1}_\alpha \mathbf{1}_\beta] = \mathbb{P}[\mathbf{1}_\alpha \mathbf{1}_\beta = 1]$ is the probability that G_n^m contains a specific copy of a graph H obtained by merging the cycles α and β. See Figure 3 for an illustration. There is only a finite number of such graphs H (up to isomorphism); each copy of a graph H arises for a bounded number of pairs (α, β); and each such graph H is connected, multicyclic and leaf-free (see Definition 3.10). Consequently, the sum (7.3) is bounded by a constant times the sum of the expected number of copies of such graphs H, and Theorem 3.11 shows that this sum is bounded as $n \to \infty$.

Finally, note that if $\beta \in \Gamma_\alpha^-$, then the cycles α and β have at least one common vertex. Consider a uniform attachment graph G_n^{2m} where the first m out-edges of any vertex $2 \leqslant i \leqslant n$ are coloured
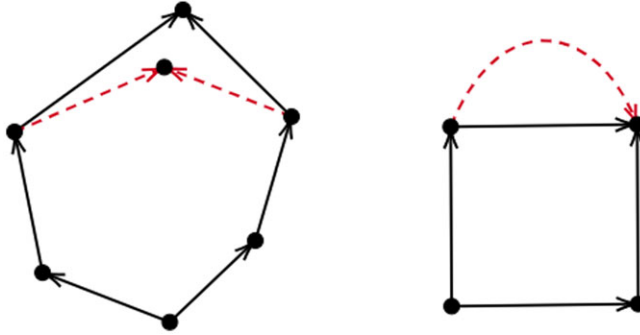
Figure 3. Two examples of positively correlated cycles. In each example, the red dashed edges form the additional segment that gives rise to the other cycle.

red, and the remaining m out-edges of these vertices are coloured blue. This graph can be thought of as overlaying two independent copies of G_n^m. Let $\alpha \cdot \beta$ be the subgraph of K_n^{2m} obtained by regarding α and β living in the red and blue parts, respectively, and taking the union of them. (Thus, the i-th edge of vertex j, $j^{(i)}$, that is part of β is now written as $j^{(i+m)}$.) With some slight abuse of notation, let $\mathbf{1}_{\alpha \cdot \beta}$ be the indicator of the subgraph $\alpha \cdot \beta$ in G_n^{2m}. By the edge independence, we have $\pi_\alpha \pi_\beta = \mathbb{E} \mathbf{1}_{\alpha \cdot \beta}$, which is the probability that G_n^{2m} contains a given copy of a graph H that, just as in the argument for Γ_α^+, is obtained by merging the cycles α and β. There is only a finite number of such graphs H, so it follows from Theorem 3.11 that the expected number of copies of any of them in G_n^{2m} is bounded as $n \to \infty$, and thus

$$\sum_{\alpha \in \Gamma} \sum_{\beta \in \Gamma_\alpha^-} \pi_\alpha \pi_\beta = \mathbb{E} \sum_{\alpha \in \Gamma} \sum_{\beta \in \Gamma_\alpha^-} \mathbf{1}_{\alpha \cdot \beta} = O(1). \tag{7.4}$$

This completes the proof of Theorem 3.1.

The proof of Theorem 3.3 is similar, using Theorem 4.2; we now have $\Gamma = \bigcup_{j=1}^r \Gamma_j$, where Γ_j is the set of potential $\ell(j)$-cycles. We estimate (7.2)–(7.4) as above, when necessary replacing ℓ by $\ell(j)$ and summing over $j = 1, \ldots, r$. □

8. Proof of the normal approximation for unicyclic graphs

8.1 Proofs of (3.12) and (3.19)

We first prove (3.12) in Theorem 3.5 and (3.19) in Theorem 3.13, which have similar proofs.

Proof of (3.12). Given a potential ℓ-cycle α, let $\Xi(\alpha)$ be the set of s-tuples $(\beta_1, \ldots, \beta_s)$ of potential copies of $\mathcal{T}_i$, $i = 1, \ldots, s$, that can be added to α to form a potential copy of Λ. Denote by Γ_{cycle} the set of potential ℓ-cycles. By independence of the edges, (3.1) and Lemma 6.4(ii), we have

$$\mu_n = \sum_{\alpha \in \Gamma_{\text{cycle}}} \sum_{\beta \in \Xi(\alpha)} \mathbb{E} \mathbf{1}_\alpha \mathbb{E} \mathbf{1}_\beta \leqslant O(\log^t n) \sum_{\alpha \in \Gamma_{\text{cycle}}} \mathbb{E} \mathbf{1}_\alpha = O(\log^{t+1} n), \tag{8.1}$$

noting that in the inequality we have used the assumption that ℓ and $\mathcal{T}_i$, $i = 1, \ldots, s$, are fixed.

For the lower bound on μ_n, we choose a suitable configuration on a set of vertices $k_1 < \ldots < k_{\ell+t}$. We use $k_1 < \ldots < k_\ell$ to form a cycle whose configuration has the mark sequence $011 \cdots 112$, and $k_{\ell+1} < \ldots < k_{\ell+t}$ to construct the trees $(\mathcal{T}_i)_{1 \leqslant i \leqslant s}$, taking their vertices in some order going from the roots outwards, so that each non-root vertex in $\mathcal{T}_i$ has mark 1; this gives a potential copy of Λ with all 1-marks except $b_1 = 0$ and $b_\ell = 2$. Hence μ_n is at least, for some constant $c > 0$, by

first summing iteratively over $k_1 < \ldots < k_{\ell-1}$,

$$\sum_{k_1 < \ldots < k_{\ell+t} \leqslant n} \frac{1}{(k_2-1)\cdots(k_{\ell-1}-1)(k_\ell-1)^2} \prod_{i=\ell+1}^{\ell+t} \frac{1}{k_i-1} \geqslant c \sum_{\ell \leqslant k_\ell < \ldots < k_{\ell+t} \leqslant n} \prod_{i=\ell}^{\ell+t} \frac{1}{k_i-1} \quad (8.2)$$
$$= \Theta\big(\log^{t+1} n\big),$$

as required. $\qquad\qquad\square$

Proof of (3.19) in Theorem 3.13. The upper bound follows as in (8.1), now letting α be a potential copy of H' and summing over all such α, using Theorem 3.11 and, as above, Lemma 6.4(ii). For the lower bound, it suffices to consider one fixed potential copy of H'. $\qquad\square$

8.2 Proofs of variance estimates

We next prove the variance estimates (3.13) in Theorem 3.5 and (3.20) in Theorem 3.13.

Proof of (3.13). Recall that a connected unicyclic graph Λ can be constructed as follows. Let $\mathcal{C}_\ell$ be an ℓ-cycle and let $\mathcal{T}_i$, $i = 1, \ldots, s$, be a tree with t_i edges and a distinguished root, so that $\mathcal{T}_i$ has $t_i + 1$ vertices. We construct $\Lambda = \Lambda_{\ell, t_1, \ldots, t_s}$ by attaching $\mathcal{T}_i$ to $\mathcal{C}_\ell$, using a vertex of $\mathcal{C}_\ell$ as the distinguished root of $\mathcal{T}_i$, and we may assume that $s \leqslant \ell$ and that at most one $\mathcal{T}_i$ attaches to each vertex in $\mathcal{C}_\ell$. Fixing Λ throughout this section, we write $\Gamma = \Gamma_\Lambda^{(n,m,\ell,t)}$ as the set of all potential copies of Λ in the uniform attachment graph $G := G_n^m$.

Given $\alpha \in \Gamma$, denote by Γ_α^- the set of potential copies of Λ that cannot coexist with α in the same instance of G_n^m (because at least one edge is incompatible with the edges of α), and denote by $\Gamma_\alpha^+ \subseteq \Gamma \setminus \Gamma_\alpha^-$ the set of potential copies of Λ that share at least one edge with α, and are compatible with α. (Note that now $\alpha \in \Gamma_\alpha^+$, unlike in Section 5.)

By edge independence, the covariance between α and any copy of Λ that does not belong to $\Gamma_\alpha^+ \cup \Gamma_\alpha^-$ is zero. Thus,

$$\sigma_n^2 = \sum_{\alpha \in \Gamma} \sum_{\beta \in \Gamma_\alpha^+} \big(\mathbb{E}\mathbf{1}_\alpha \mathbf{1}_\beta - \mathbb{E}\mathbf{1}_\alpha \mathbb{E}\mathbf{1}_\beta\big) - \sum_{\alpha \in \Gamma} \sum_{\beta \in \Gamma_\alpha^-} \mathbb{E}\mathbf{1}_\alpha \mathbb{E}\mathbf{1}_\beta. \quad (8.3)$$

Consider first the positive covariances, i.e., the first double sum in (8.3). We argue as in the estimate of (7.3). Let $\alpha\beta$ denote the union of the graphs α and β, and note that $\mathbf{1}_\alpha \mathbf{1}_\beta = \mathbf{1}_{\alpha\beta}$. Then $\alpha\beta$ is a connected graph, which either is unicyclic (an ℓ-cycle with attached trees) or multicyclic, in both cases with at most $2t$ edges outside the 2-core; see Figure 4 for examples. Each graph $\alpha\beta$ arises from a bounded number of pairs (α, β), and thus it follows by Theorem 3.13, summing over the finitely many isomorphism types of $\alpha\beta$ that can arise, that the expected number of pairs (α, β) in G_n^m with $\beta \in \Gamma_\alpha^+$ and $\alpha\beta$ multicyclic is $O(\log^{2t} n)$. Similarly, it follows from (3.12), again summing over a finite number of possible types of $\alpha\beta$, that the expected number of pairs with $\alpha\beta$ unicyclic is $\Theta(\log^{2t+1} n)$. Consequently, we have the exact order

$$\sum_{\alpha \in \Gamma} \sum_{\beta \in \Gamma_\alpha^+} \mathbb{E}\mathbf{1}_\alpha \mathbf{1}_\beta = \Theta(\log^{2t+1} n). \quad (8.4)$$

Furthermore, we have $0 \leq \mathbb{E}\mathbf{1}_\alpha \mathbb{E}\mathbf{1}_\beta \leq \mathbb{E}\mathbf{1}_\alpha \mathbf{1}_\beta$ for all α and $\beta \in \Gamma_\alpha^+$, with $\mathbb{E}\mathbf{1}_\alpha \mathbb{E}\mathbf{1}_\beta \leq \frac{1}{2}\mathbb{E}\mathbf{1}_\alpha \mathbf{1}_\beta$ unless both α and β contain vertex 1, and it is easy to see that also

$$\sum_{\alpha \in \Gamma} \sum_{\beta \in \Gamma_\alpha^+} (\mathbb{E}\mathbf{1}_\alpha \mathbf{1}_\beta - \mathbb{E}\mathbf{1}_\alpha \mathbb{E}\mathbf{1}_\beta) = \Theta(\log^{2t+1} n). \quad (8.5)$$

Alternatively, this follows from (8.4) and the argument for Γ_α^- below.
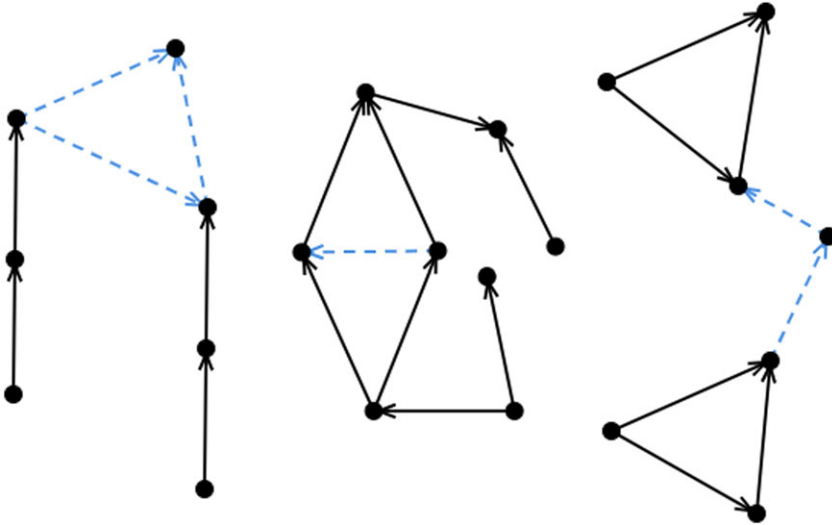
Figure 4. Examples of positively correlated copies of Λ (with $\ell = 3$ and $\mathcal{T}_1$ being a path with 2 edges). The blue dashed edges are the shared edges.

We see also that the dominant term is contributed by pairs $\alpha \in \Gamma$ and $\beta \in \Gamma_\alpha^+$ such that α and β have the same cycle but no other common vertices.

For the negative covariances, we show that the last double sum in (8.3) is of order $O(\log^{2t} n)$. To do so, we modify the argument for showing the sum in (7.4) is bounded. We construct a uniform attachment graph G_n^{2m} where the first m out-edges of any vertex $2 \leqslant i \leqslant n$ are coloured red, and the remaining m out-edges are coloured blue. If α and β are two potential copies of Γ, let $\alpha \cdot \beta$ be the subgraph of K_n^{2m} obtained by regarding α as living in the red part and β in the blue part, and then taking the union of them (see Figure 5 for examples). Note that if $\beta \in \Gamma_\alpha^-$, then α and β must share at least one vertex. Thus, $\alpha \cdot \beta$ is connected; furthermore, it contains at least two cycles, and it has at most $2t$ edges in the tree parts outside its 2-core. Once again, with some slight abuse of notation, let $\mathbf{1}_{\alpha \cdot \beta}$ be the indicator that $\alpha \cdot \beta$ is present in G_n^{2m}. By independence of the edges, we have $\mathbb{E}\mathbf{1}_\alpha \mathbb{E}\mathbf{1}_\beta = \mathbb{E}\mathbf{1}_{\alpha \cdot \beta}$. Since Λ is fixed, the total negative covariance in the last sum in (8.3) can thus be bounded by some constant times the expected number of copies of subgraphs of the possible types of $\alpha \cdot \beta$ in G_n^{2m}. Again, Theorem 3.13 implies that this is $O(\log^{2t} n)$.

Combining the estimates above of the sums in (8.3) yields the result $\sigma_n^2 = \Theta(\log^{2t+1} n)$. □

Proof of Theorem 3.13. We have proved (3.19) in Section 8.1, and (3.20) follows from Theorem 3.11 and an argument similar to the proof above of (3.13) (where we may simplify and ignore the negative terms), noting that a subgraph obtained by merging two multicyclic and leaf-free subgraphs at one or more vertices and edges is still multicyclic and leaf-free. □

8.3 Proof of Theorem 3.5

We now complete the proof of Theorem 3.5 by proving the normal approximation result (3.14).

***Proof of* (3.14).** We bound the error terms appearing in (4.4) using the coupling described in Section 5. Let W_n^s be the size-bias version of W_n defined there. For $\alpha \in \Gamma$, let Γ_α^- and Γ_α^+ be the subsets of Γ defined at the beginning of this subsection, and let as in Section 5, G^α be the graph G forced to contain all edges of α. Denote by $\mathbf{1}_\beta^\alpha$ the indicator that a copy β of Λ is in G^α, and let
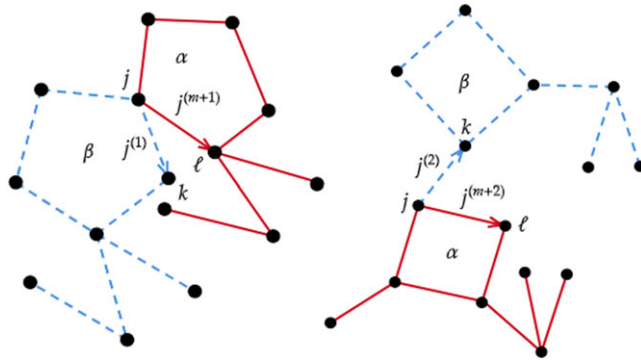
Figure 5. Two examples of two negative correlated copies α and β that are embedded in the graph G_n^{2m}, where vertex j is the common vertex and $j^{(1)}$ (left) and $j^{(2)}$ (right) are the directed edges that have different recipients in α and β. The red solid (resp. blue dashed) edges are the edges of α (resp. β.).

$\tilde{\mathbf{i}}_\beta^\alpha := \mathbf{1}_\beta^\alpha - \mathbf{1}_\beta$. Also, if $\beta \in \Gamma_\alpha^+$, let $\beta \setminus \alpha$ be the graph obtained from β by deleting the edges that also are in α; then $\mathbf{1}_\beta^\alpha = \mathbf{1}_{\beta \setminus \alpha}$.

If $\beta \in \Gamma_\alpha^+$, then $\tilde{\mathbf{i}}_\beta^\alpha \geq 0$ since if β exists in G, then it will also exist in G^α; furthermore, if $\beta \in \Gamma_\alpha^-$, then $\mathbf{1}_\beta^\alpha = 0$ and thus $\tilde{\mathbf{i}}_\beta^\alpha = -\mathbf{1}_\beta$, and if $\beta \notin \Gamma_\alpha^+ \cup \Gamma_\alpha^-$, then $\tilde{\mathbf{i}}_\beta^\alpha = 0$. Hence, the construction of W^s in Section 5 yields

$$\mathbb{E}[W_n^s - W_n \mid G] = \sum_{\alpha \in \Gamma} \frac{\mathbb{E}\mathbf{1}_\alpha}{\mu_n} \sum_{\beta \in \Gamma} \tilde{\mathbf{i}}_\beta^\alpha = \sum_{\alpha \in \Gamma} \frac{\mathbb{E}\mathbf{1}_\alpha}{\mu_n} \left(\sum_{\beta \in \Gamma_\alpha^+} \tilde{\mathbf{i}}_\beta^\alpha - \sum_{\beta \in \Gamma_\alpha^-} \mathbf{1}_\beta \right)$$

$$=: \sum_{\alpha \in \Gamma} \frac{\mathbb{E}\mathbf{1}_\alpha}{\mu_n} \left(R_\alpha^+ - R_\alpha^- \right). \tag{8.6}$$

Taking the variance yields, using the (Cauchy–Schwarz) inequality $(a + b)^2 \leqslant 2(a^2 + b^2)$ for all real a and b,

$$\mu_n^2 \mathrm{Var}\big(\mathbb{E}[W_n^s - W_n \mid G]\big) \leqslant 2 \, \mathrm{Var}\left(\sum_{\alpha \in \Gamma} \mathbb{E}\mathbf{1}_\alpha R_\alpha^+ \right) + 2 \, \mathrm{Var}\left(\sum_{\alpha \in \Gamma} \mathbb{E}\mathbf{1}_\alpha R_\alpha^- \right). \tag{8.7}$$

We proceed to bound the variances on the right-hand side. The first equals

$$A_1 := \sum_{\alpha \in \Gamma} \sum_{\beta \in \Gamma} \mathbb{E}\mathbf{1}_\alpha \mathbb{E}\mathbf{1}_\beta \mathrm{Cov}(R_\alpha^+, R_\beta^+) = \sum_{\alpha \in \Gamma} \sum_{\beta \in \Gamma} \sum_{\gamma_1 \in \Gamma_\alpha^+} \sum_{\gamma_2 \in \Gamma_\beta^+} \mathbb{E}\mathbf{1}_\alpha \mathbb{E}\mathbf{1}_\beta \, \mathrm{Cov}\big(\tilde{\mathbf{i}}_{\gamma_1}^\alpha, \tilde{\mathbf{i}}_{\gamma_2}^\beta\big). \tag{8.8}$$

In (8.8), we only need to consider pairs of γ_1, γ_2 that have at least one edge in common, since otherwise $\tilde{\mathbf{i}}_{\gamma_1}^\alpha$ and $\tilde{\mathbf{i}}_{\gamma_2}^\beta$ are independent. Moreover, since $\gamma_1 \in \Gamma_\alpha^+$ and $\gamma_2 \in \Gamma_\beta^+$, γ_1 has an edge (and thus a vertex) in common with α and γ_2 has an edge (and thus a vertex) in common with β. We argue similarly to the proof of (3.13) above, this time considering the uniform attachment graph G_n^{3m}, which we regard as three independent copies of G that are coloured red, blue, and green. Let $\alpha \cdot \beta * \gamma_1 \gamma_2$ be the graph formed by a red copy of α, a blue copy of β, and a green copy of $(\gamma_1 \setminus \alpha) \cup (\gamma_2 \setminus \beta)$. Then

$$\mathrm{Cov}\big(\tilde{\mathbf{i}}_{\gamma_1}^\alpha, \tilde{\mathbf{i}}_{\gamma_2}^\beta\big) \leq \mathbb{E}\big[\tilde{\mathbf{i}}_{\gamma_1}^\alpha \tilde{\mathbf{i}}_{\gamma_2}^\beta\big] \leq \mathbb{E}\big[\mathbf{1}_{\gamma_1}^\alpha \mathbf{1}_{\gamma_2}^\beta\big] = \mathbb{E}\big[\mathbf{1}_{\gamma_1 \setminus \alpha} \mathbf{1}_{\gamma_2 \setminus \beta}\big]. \tag{8.9}$$

and thus

$$\mathbb{E}\mathbf{1}_\alpha \mathbb{E}\mathbf{1}_\beta \, \mathrm{Cov}\big(\tilde{\mathbf{i}}_{\gamma_1}^\alpha, \tilde{\mathbf{i}}_{\gamma_2}^\beta\big) \leq \mathbb{E}\mathbf{1}_\alpha \mathbb{E}\mathbf{1}_\beta \mathbb{E}\big[\mathbf{1}_{\gamma_1 \setminus \alpha} \mathbf{1}_{\gamma_2 \setminus \beta}\big] = \mathbb{E}\big[\mathbf{1}_{\alpha \cdot \beta * \gamma_1 \gamma_2}\big]. \tag{8.10}$$

Note that $\alpha\cdot\beta*\gamma_1\gamma_2$ is a connected graph, which has at least two cycles. The edges in the trees attached to the 2-core of $\alpha\cdot\beta*\gamma_1\gamma_2$ come from the attached trees in α, β, γ_1, and γ_2, and thus the total number of them is at most $4t$. Consequently, Theorem 3.13 applies. Together with (8.8)–(8.10), and summing over the finite number of possible types of the graph $\alpha\cdot\beta*\gamma_1\gamma_2$, we obtain

$$A_1 = O(\log^{4t} n). \tag{8.11}$$

Similarly, the second variance on the right-hand side of (8.7) equals

$$A_2 := \sum_{\alpha\in\Gamma}\sum_{\beta\in\Gamma}\mathbb{E}\mathbf{1}_\alpha\mathbb{E}\mathbf{1}_\beta\mathrm{Cov}(R_\alpha^-, R_\beta^-) = \sum_{\alpha\in\Gamma}\sum_{\beta\in\Gamma}\sum_{\gamma_1\in\Gamma_\alpha^-}\sum_{\gamma_2\in\Gamma_\beta^-}\mathbb{E}\mathbf{1}_\alpha\mathbb{E}\mathbf{1}_\beta\,\mathrm{Cov}(\mathbf{1}_{\gamma_1}, \mathbf{1}_{\gamma_2}). \tag{8.12}$$

Again, we only need to consider pairs of γ_1, γ_2 that have at least one edge in common. Moreover, since $\gamma_1\in\Gamma_\alpha^-$ and $\gamma_2\in\Gamma_\beta^-$, γ_1 has a vertex in common with α and γ_2 has a vertex in common with β. We consider again G_n^{3m} and now let $\alpha\cdot\beta\cdot\gamma_1\gamma_2$ be the graph consisting of a red copy of α, a blue copy of β, and green copies of γ_1 and γ_2. Then

$$\mathbb{E}\mathbf{1}_\alpha\mathbb{E}\mathbf{1}_\beta\,\mathrm{Cov}(\mathbf{1}_{\gamma_1}, \mathbf{1}_{\gamma_2}) \leq \mathbb{E}\mathbf{1}_\alpha\mathbb{E}\mathbf{1}_\beta\mathbb{E}[\mathbf{1}_{\gamma_1}\mathbf{1}_{\gamma_2}] = \mathbb{E}[\mathbf{1}_{\alpha\cdot\beta\cdot\gamma_1\gamma_2}], \tag{8.13}$$

and $\alpha\cdot\beta\cdot\gamma_1\gamma_2$ is a connected graph with at least three cycles and at most $4t$ edges outside its 2-core. Consequently, we now obtain from (3.12) and Theorem 3.13

$$A_2 = O(\log^{4t} n). \tag{8.14}$$

By (8.7) and the bounds (8.11) and (8.14) above, we deduce that

$$\mu_n^2\mathrm{Var}(\mathbb{E}[W_n^s - W_n \mid G]) = O(\log^{4t} n). \tag{8.15}$$

Moreover, W_n is determined by G, and thus the conditional Jensen's inequality yields

$$\mathrm{Var}(\mathbb{E}[W_n^s - W_n \mid W_n]) = \mathrm{Var}(\mathbb{E}[\mathbb{E}[W_n^s - W_n \mid G] \mid W_n]) \leqslant \mathrm{Var}(\mathbb{E}[W_n^s - W_n \mid G]). \tag{8.16}$$

Noting also that $\sigma_n^2 = \Theta(\log^{2t+1} n)$ by (3.13), the first error term in (4.4) is therefore

$$\frac{\mu_n}{\sigma_n^2}\sqrt{\frac{2}{\pi}}\sqrt{\mathrm{Var}(\mathbb{E}[W_n^s - W_n \mid W_n])} = O(\log^{4t/2 - (2t+1)} n) = O(\log^{-1} n). \tag{8.17}$$

We now turn to the second error term in (4.4). By first conditioning on the graph G, we have in analogy with (8.6),

$$\mathbb{E}[(W_n^s - W_n)^2 \mid G] = \sum_{\alpha\in\Gamma}\frac{\mathbb{E}\mathbf{1}_\alpha}{\mu_n}(R_\alpha^+ - R_\alpha^-)^2 \leq 2\sum_{\alpha\in\Gamma}\frac{\mathbb{E}\mathbf{1}_\alpha}{\mu_n}((R_\alpha^+)^2 + (R_\alpha^-)^2), \tag{8.18}$$

and thus

$$\mu_n\mathbb{E}[(W_n^s - W_n)^2] \leq 2\sum_{\alpha\in\Gamma}\mathbb{E}\mathbf{1}_\alpha(R_\alpha^+)^2 + 2\sum_{\alpha\in\Gamma}\mathbb{E}\mathbf{1}_\alpha(R_\alpha^-)^2. \tag{8.19}$$

We argue as above. The first sum on the right-hand side equals

$$B_1 := \sum_{\alpha\in\Gamma}\mathbb{E}\mathbf{1}_\alpha\sum_{\gamma_1\in\Gamma_\alpha^+}\sum_{\gamma_2\in\Gamma_\alpha^+}\mathrm{Cov}(\tilde{\mathbf{1}}_{\gamma_1}^\alpha, \tilde{\mathbf{1}}_{\gamma_2}^\alpha). \tag{8.20}$$

Hence, using (8.9),

$$B_1 \leqslant \sum_{\alpha\in\Gamma}\sum_{\gamma_1\in\Gamma_\alpha^+}\sum_{\gamma_2\in\Gamma_\alpha^+}\mathbb{E}\mathbf{1}_\alpha\mathbb{E}[\mathbf{1}_{\gamma_1\setminus\alpha}\mathbf{1}_{\gamma_2\setminus\alpha}] = \sum_{\alpha\in\Gamma}\sum_{\gamma_1\in\Gamma_\alpha^+}\sum_{\gamma_2\in\Gamma_\alpha^+}\mathbb{E}\mathbf{1}_{\alpha*\gamma_1\gamma_2}, \tag{8.21}$$

where $\alpha*\gamma_1\gamma_2$ is the subgraph formed by a red copy of α and a blue copy of $(\gamma_1\setminus\alpha)\cup(\gamma_2\setminus\alpha)$, regarded as a subgraph of K_n^{2m} coloured red and blue as above, and $\mathbf{1}_{\alpha*\gamma_1\gamma_2}$ is the indicator that

this subgraph exists in G_n^{2m}. Since $\gamma_1, \gamma_2 \in \Gamma_\alpha^+$, they have at least one vertex each in common with α, and thus $\alpha * \gamma_1 \gamma_2$ is connected. Moreover, $\alpha * \gamma_1 \gamma_2$ has at least one cycle, and at most $3t$ edges outside its 2-core. Consequently, we now obtain from (3.12) and Theorem 3.13

$$B_1 = O\big(\log^{3t+1} n\big). \tag{8.22}$$

Similarly, the second sum in (8.19) equals

$$
\begin{aligned}
B_2 &:= \sum_{\alpha \in \Gamma} \mathbb{E}\mathbf{1}_\alpha \sum_{\gamma_1 \in \Gamma_\alpha^-} \sum_{\gamma_2 \in \Gamma_\alpha^-} \mathrm{Cov}(\mathbf{1}_{\gamma_1}, \mathbf{1}_{\gamma_2}) \leqslant \sum_{\alpha \in \Gamma} \sum_{\gamma_1 \in \Gamma_\alpha^-} \sum_{\gamma_2 \in \Gamma_\alpha^-} \mathbb{E}\mathbf{1}_\alpha \mathbb{E}\big[\mathbf{1}_{\gamma_1} \mathbf{1}_{\gamma_2}\big] \\
&= \sum_{\alpha \in \Gamma} \sum_{\gamma_1 \in \Gamma_\alpha^-} \sum_{\gamma_2 \in \Gamma_\alpha^-} \mathbb{E}\mathbf{1}_{\alpha \cdot \gamma_1 \gamma_2},
\end{aligned}
\tag{8.23}
$$

where $\alpha \cdot \gamma_1 \gamma_2$ is the subgraph formed by a red copy of α and a blue copy of $\gamma_1 \cup \gamma_2$. We obtain from (3.12) and Theorem 3.13

$$B_2 = O\big(\log^{3t+1} n\big). \tag{8.24}$$

We obtain from (8.19), (8.22), and (8.24) that

$$\mu_n \mathbb{E}\big[(W_n^s - W_n)^2\big] = O\big(\log^{3t+1} n\big). \tag{8.25}$$

By this and (3.13), we conclude that the second error term in (4.4) is

$$\frac{\mu_n}{\sigma_n^3} \mathbb{E}[(W_n^s - W_n)^2] = O\big(\log^{3t+1-3(2t+1)/2} n\big) = O\big(\log^{-\frac{1}{2}} n\big). \tag{8.26}$$

Finally, we use (8.17) and (8.26) in (4.4) and obtain (3.14), which completes the proof of Theorem 3.5. $\qquad\square$

9. Proof of the normal approximations for trees

In this section we prove Theorem 3.7 and Theorem 3.9.

9.1 Proof of Theorem 3.7.

We start by proving (3.16): By Lemma 6.4 (i) and (6.1), an upper bound is, by summing in the order $k_1, \ldots, k_t$,

$$\mu_n \leqslant C \sum_{1 \leqslant k_1 < \ldots < k_t \leqslant n} \prod_{i=2}^{t} \frac{1}{k_i - 1} \leqslant C \sum_{k_t \leqslant n} 1 = Cn, \tag{9.1}$$

The lower bound follows similarly, or simply because there are $\Theta(n^t)$ potential copies of Λ, and each has probability $\geqslant n^{-(t-1)}$.

(3.17): To obtain an upper bound on σ_n^2, we argue as in the proof of (3.13). We use again (8.3), and the only difference from the argument in Section 8 is that for $\alpha \in \Gamma$ and $\beta \in \Gamma_\alpha^+$, their union $\alpha\beta$ does not have to contain a cycle; however, it is always a connected graph with at most $2t$ edges. There is still a finite number of types of $\alpha\beta$, and we obtain by using (depending on the number of cycles in $\alpha\beta$) (3.16), (3.12) and Theorem 3.13

$$\sum_{\alpha \in \Gamma} \sum_{\beta \in \Gamma_\alpha^+} \mathbb{E}\mathbf{1}_{\alpha\beta} \leqslant O(n) + O\big(\log^{2t} n\big) = O(n), \tag{9.2}$$

which implies (3.17).

If $\sigma_n^2 = \Theta(n)$, then we can use Theorem 4.3 to prove the asymptotic normality of W_n in a similar vein as before. Again, we argue as in the proof of Theorem 3.5 in Section 8, and the only difference

is that the graphs $\alpha \cdot \beta * \gamma_1 \gamma_2$, $\alpha \cdot \beta \cdot \gamma_1 \gamma_2$, $\alpha * \gamma_1 \gamma_2$, $\alpha \cdot \gamma_1 \gamma_2$ do not have to contain a cycle. As in the estimation of variance above, we therefore use not only (3.12) and Theorem 3.13 but also (3.16) in the estimates; as a result we obtain the bound $O(n)$ in (8.15) and (8.25). Hence, if $\sigma_n^2 = \Theta(n)$, the terms in (4.4) are

$$\frac{\mu_n}{\sigma_n^2} \sqrt{\mathrm{Var}(\mathbb{E}[W_n^s - W_n \mid G])} = O(n^{-1/2}), \quad \frac{\mu_n}{\sigma_n^3} \mathbb{E}[(W_n^s - W_n)^2] = O(n^{-1/2}), \tag{9.3}$$

implying (3.18).

9.2 *Proof of Theorem 3.9.*

Recall that the upper bound $O(n)$ for the variance σ_n^2 is already proved in (3.17). To obtain a lower bound, we state and prove the following general lemma. In our application of it, X and ξ_v will be taken as the count of trees W_n and the recipient of a directed edge v in the uniform attachment graph G_n^m.

Lemma 9.1. *Let $(\xi_v)_{v \in \mathcal{I}}$ be a family of independent random variables (where $\mathcal{I}$ is an arbitrary index set), and let X be any random variable such that $\mathbb{E}X^2 < \infty$. For each $v \in \mathcal{I}$, let*

$$X_v := \mathbb{E}(X \mid \xi_v) - \mathbb{E}X. \tag{9.4}$$

Then

$$\mathrm{Var}(X) \geqslant \sum_{v \in \mathcal{I}} \mathrm{Var}(X_v) = \sum_{v \in \mathcal{I}} \mathrm{Var}\big[\mathbb{E}(X \mid \xi_v)\big]. \tag{9.5}$$

Proof. Assume first that $\mathcal{I}$ is finite and define

$$Y := \sum_{v \in \mathcal{I}} X_v. \tag{9.6}$$

Note that

$$\mathbb{E}Y = \sum_{v \in \mathcal{I}} \mathbb{E}X_v = 0. \tag{9.7}$$

It follows from (9.4) that the random variable X_v is a function of ξ_v. Consequently, the variables X_v are independent, and thus

$$\mathbb{E}(Y^2) = \mathrm{Var}(Y) = \sum_{v \in \mathcal{I}} \mathrm{Var}(X_v). \tag{9.8}$$

Moreover, (9.4) implies also

$$\mathbb{E}\left[(X - \mathbb{E}X)X_v\right] = \mathbb{E}\left[\mathbb{E}(X - \mathbb{E}X \mid \xi_v)X_v\right] = \mathbb{E}\left(X_v^2\right) = \mathrm{Var}(X_v). \tag{9.9}$$

Consequently,

$$\mathbb{E}\left[(X - \mathbb{E}X)Y\right] = \sum_{v \in \mathcal{I}} \mathbb{E}\left[(X - \mathbb{E}X)X_v\right] = \sum_{v \in \mathcal{I}} \mathrm{Var}(X_v) = \mathbb{E}(Y^2). \tag{9.10}$$

Hence,

$$\mathbb{E}\left[(X - \mathbb{E}X - Y)Y\right] = 0. \tag{9.11}$$

and thus

$$\begin{aligned}
\mathbb{E}\left[(X - \mathbb{E}X)^2\right] &= \mathbb{E}\left[(X - \mathbb{E}X - Y)^2\right] + 2\mathbb{E}\left[(X - \mathbb{E}X - Y)Y\right] + \mathbb{E}\left[Y^2\right] \\
&= \mathbb{E}\left[(X - \mathbb{E}X - Y)^2\right] + \mathbb{E}\left[Y^2\right] \\
&\geq \mathbb{E}\left[Y^2\right].
\end{aligned} \tag{9.12}$$

The result follows from (9.12) and (9.8), which completes the proof for finite $\mathcal{I}$.

If $\mathcal{I}$ is infinite, the case just proved shows that for every finite subset $\mathcal{I}_1 \subset \mathcal{I}$, we have

$$\mathrm{Var}(X) \geq \sum_{v \in \mathcal{I}_1} \mathrm{Var}(X_v), \tag{9.13}$$

which implies (9.5). $\qquad\square$

Remark 9.2. From a more abstract point of view, a minor modification of the proof above shows that Y is the orthogonal projection of $X - \mathbb{E}X$ onto the linear subspace of L^2 consisting of sums of the type $\sum_{v \in \mathcal{I}} f_v(\xi_v)$, and the inequality (9.5) is thus an instance of the fact that orthogonal projections in L^2 never increase the variance (or the norm).

Proof of Theorem 3.9. Fix $j \in [n]$ and $a \in [m]$. Recall that the edge $j^{(a)}$ in G_n^m is randomly chosen as one of the edges $ji^{(a)}$ ($i \in [j-1]$) in K_n^m; we denote the recipient of the edge $j^{(a)}$ by $\xi = \xi_{j,a}$; thus ξ is uniformly distributed on $[j-1]$, and $j^{(a)} = j\xi^{(a)}$. We condition on ξ, and decompose W_n as

$$W_n = W_n^{(0)} + W_n^{(1)} + W_n^{(2)}, \tag{9.14}$$

where

- $W_n^{(0)}$ is the number of copies of S_ℓ that do not contain the edge $ji^{(a)}$ for any $i \in [j-1]$.
- $W_n^{(1)}$ is the number of copies α of S_ℓ that contain $j^{(a)} = j\xi^{(a)}$ and such that the center of α is j.
- $W_n^{(2)}$ is the number of copies α of S_ℓ that contain $j^{(a)} = j\xi^{(a)}$ and such that the center of α is ξ.

Then $W_n^{(0)}$ is clearly independent of ξ. Moreover, for every $i < j$, $\mathbb{E}(W_n^{(1)} \mid \xi = i)$ is the expected number of stars $S_{\ell-1}$ with center j and $\ell - 1$ leaves in $[n] \setminus \{i, j\}$. Since the edges from j have their endpoints uniformly distributed in $[j-1]$, this number does not depend on i. In other words, $\mathbb{E}(W_n^{(1)} \mid \xi)$ does not depend on ξ.

Similarly, $\mathbb{E}(W_n^{(2)} \mid \xi = i)$ is the expected number of stars $S_{\ell-1}$ with center i and $\ell - 1$ leaves in $[n] \setminus \{i, j\}$. Hence,

$$\mathbb{E}\big(W_n^{(2)} \mid \xi = i\big) = \sum_{q=0}^{\ell-1} W'_{j,i,q}, \tag{9.15}$$

where $W'_{j,i,q}$ is the number of such copies of $S_{\ell-1}$ with q leaves in $(i, n]$ and $r := \ell - 1 - q$ in $[1, i)$. Assume for simplicity that $i \geqslant n/10$. Then, counting the number of ways to choose first the vertices and then the edges of such a copy, and multiplying with the probability that it exists in G_n^m,

$$\mathbb{E}W'_{j,i,q} = \binom{i-1}{r} \sum_{\substack{i < i_1 < \ldots < i_q \leqslant n \\ i_1, \ldots, i_q \neq j}} \binom{m}{r} r! \, m^q (i-1)^{-r} \prod_{k=1}^{q} (i_k - 1)^{-1}$$

$$= \left(1 + O\left(\tfrac{1}{n}\right)\right) \binom{m}{r} m^q \frac{1}{q!} \left(\sum_{p=i+1}^{n} \frac{1}{p-1} + O\left(\tfrac{1}{n}\right)\right)^q$$

$$= \binom{m}{r} \frac{m^q}{q!} \log^q\left(\frac{n}{i}\right) + O\left(\frac{1}{n}\right). \tag{9.16}$$

Since $\log(n/i)$ is positive and monotonically decreasing in $i \in [n/10, n)$, it follows that if $\frac{n}{10} \leqslant i \leqslant k < j \leqslant n$, then

$$
\begin{aligned}
\mathbb{E}\big(W_n \mid \xi = i\big) - \mathbb{E}\big(W_n \mid \xi = k\big) &= \mathbb{E}\big(W_n^{(2)} \mid \xi = i\big) - \mathbb{E}\big(W_n^{(2)} \mid \xi = k\big) \\
&= \sum_{q=0}^{\ell-1} \big(W'_{j,i,q} - W'_{j,k,q}\big) \\
&= \sum_{q=0}^{\ell-1} \binom{m}{r} \frac{m^q}{q!} \Big(\log^q\Big(\frac{n}{i}\Big) - \log^q\Big(\frac{n}{k}\Big) + O\Big(\frac{1}{n}\Big)\Big) \\
&\geqslant \frac{m^{\ell-1}}{(\ell-1)!} \Big(\log^{\ell-1}\Big(\frac{n}{i}\Big) - \log^{\ell-1}\Big(\frac{n}{k}\Big)\Big) + O\Big(\frac{1}{n}\Big). \quad (9.17)
\end{aligned}
$$

In particular, if $i \in [n/10, n/4]$ and $k \in [n/3, n/2]$, then

$$
\mathbb{E}\big(W_n \mid \xi = i\big) - \mathbb{E}\big(W_n \mid \xi = k\big) \geqslant \frac{m^{\ell-1}}{(\ell-1)!} \big(\log^{\ell-1} 4 - \log^{\ell-1} 3\big) + O\Big(\frac{1}{n}\Big) \geqslant c. \quad (9.18)
$$

for some $c > 0$ if n is large enough, say $n \geqslant n_1$.

Let ξ' be an independent copy of $\xi = \xi_{j,a}$ and write $g(\xi) := \mathbb{E}(W_n \mid \xi)$. Assume $n \geqslant n_1$. Then (9.18) shows that $g(i) - g(k) \geqslant c$ if $i \in [n/10, n/4]$ and $k \in [n/3, n/2]$; hence it follows that if $j > n/2$, then

$$
\begin{aligned}
\mathrm{Var}\big[\mathbb{E}(W_n \mid \xi)\big] = \mathrm{Var}\, g(\xi) &= \tfrac{1}{2} \mathbb{E}\,\big(g(\xi) - g(\xi')\big)^2 \\
&\geqslant \tfrac{1}{2} \mathbb{P}\big(\xi \in [n/10, n/4]\big) \mathbb{P}\big(\xi' \in [n/3, n/2]\big) c^2 \geqslant c_1, \quad (9.19)
\end{aligned}
$$

for some constant $c_1 > 0$.

We now apply Lemma 9.1 with the family of all random variables $\xi_{j,a}$, thus letting $\mathcal{I} = \{2, \ldots, n\} \times [m]$. Then (9.5) and the lower bound in (9.19) (for $j \geqslant n/2$) show that for $n \geqslant n_1$,

$$
\mathrm{Var}(W_n) \geqslant \sum_{j=\lceil n/2 \rceil}^{n} \sum_{a=1}^{m} c_1 = \Omega(n). \quad (9.20)
$$

Since $\mathrm{Var}\, W_n = \sigma_n^2 = O(n)$ by (3.17), this completes the proof. $\qquad \square$

10. Proof of Theorem 3.15

Proof of Theorem 3.15. Let h' be the number of vertices in H', and let $r \in [h]$ be the number of them such that attaching an edge at that vertex yields a copy of H.

Fix a copy H'_0 of H' in G_∞^m, and let its vertices be $k_1 < \cdots < k_{h'}$. For every $j \in \mathbb{N}$ and $a \in [m]$, let I_{ja} be the indicator of the event that the edge $j^{(a)}$ in G_m^n has exactly one endpoint in H'_0 and moreover that attaching $j^{(a)}$ to H'_0 yields a copy of H. Then, for $n \geqslant k_{h'}$, the number of copies of H in G_n^m that contain the given subgraph H'_0 is

$$
W_n^{H'_0} := \sum_{j=1}^{n} \sum_{a=1}^{m} I_{ja} = \sum_{j=k_{h'}+1}^{n} \sum_{a=1}^{m} I_{ja} + O(1). \quad (10.1)
$$

Condition on the existence of H'_0. Then the variables I_{ja} in the final sum are independent and have the Bernoulli distributions $\mathrm{Be}\,(r/(j-1))$. It follows by the lemma below that as $n \to \infty$, $W_n^{H'_0}/\log n \xrightarrow{\text{a.s.}} rm$. Thus, every copy of H' in G_∞^m has asymptotically $rm \log n$ attached edges that

make copies of H. Summing over the a.s. finite number of copies H'_0 of H' in G^m_∞, we obtain

$$\frac{W_n}{\log n} \xrightarrow{\text{a.s.}} rmW'_\infty, \tag{10.2}$$

which verifies Conjecture 3.14 in this case. □

The proof used the following simple law of large numbers; it is certainly known but for completeness we include a proof.

Lemma 10.1. *Let $I_i \in \mathrm{Be}\,(p_i)$ be independent random variables with $p_i \in [0, 1]$, for $i = 1, 2, \ldots$. If $\sum_{i=1}^\infty p_i = \infty$, then*

$$\frac{\sum_{i=1}^n I_i}{\sum_{i=1}^n p_i} \xrightarrow{\text{a.s.}} 1. \tag{10.3}$$

Proof. Let $b_n := \sum_{i=1}^n p_i$. We have

$$\sum_{n=1}^\infty \frac{\mathbb{E}(I_n - p_n)^2}{b_n^2} \leqslant \sum_{n=1}^\infty \frac{p_n}{b_n^2} = \frac{1}{p_1} + \sum_{n=2}^\infty \frac{b_n - b_{n-1}}{b_n^2} \leqslant \frac{1}{p_1} + \sum_{n=2}^\infty \int_{b_{n-1}}^{b_n} \frac{\mathrm{d}s}{s^2}$$

$$= \frac{1}{p_1} + \int_{b_1}^\infty \frac{\mathrm{d}s}{s^2} < \infty. \tag{10.4}$$

Since the I_n are independent, it follows that $b_n^{-1} \sum_{i=1}^n (I_i - p_i) \xrightarrow{\text{a.s.}} 0$, see [[8], Theorem VII.8.3], which is equivalent to (10.3). □

Acknowledgement

References

[1] Arya, S., Golin, M. J. and Mehlhorn, K. (1999) On the expected depth of random circuits. *Comb. Prob. Comp.* **8** 209–228.

[2] Barbour, A. D., Holst, L. and Janson, S. (1992) *Poisson Approximation*. Oxford University Press.

[3] Broutin, N. and Fawzi, O. (2012) Longest path distance in random circuits. *Comb. Prob. Comp.* **21** 856–881.

[4] Chen, L. H. Y., Goldstein, L. and Shao, Q-M. (2011) *Normal Approximation by Stein's Method*. Springer.

[5] Devroye, L. and Janson, S. (2011) Long and short paths in uniform random recursive dags. *Arkiv Matematik.* **49** 61–77.

[6] Devroye, L. and Lu, J. (1995) The strong convergence of maximal degrees in uniform random recursive trees and dags. *Random Struct. Algor.* **7** 1–14.

[7] Drmota, M. (2009) *Random Trees: An Interplay between Combinatorics and Probability*. SpringerWienNewYork.

[8] Feller, W. (1971) *An Introduction to Probability Theory and its Applications*. Volume II. 2nd ed.John Wiley & Sons, Inc.

[9] Feng, Q. and Mahmoud, H. M. (2010) On the variety of shapes on the fringe of a random recursive tree. *J. Appl. Prob.* **47** 191–200.

[10] Feng, Q., Mahmoud, H. M. and Panholzer, A. (2008) Phase changes in subtree varieties in random recursive and binary search trees. *SIDMA.* **22** 160–184.

[11] Fuchs, M. (2008) Subtree sizes in recursive trees and binary search trees: Berry–Esseen bounds and Poisson approximations. *Comb. Prob. Comp.* **17** 661–680.

[12] Goldstein, L. and Rinott, Y. (1996) Multivariate normal approximations by Stein's method and size bias couplings. *J. Appl. Probab.* **33** 1–17.

[13] Holmgren, C. and Janson, S. (2015) Limit laws for functions of fringe trees for binary search trees and random recursive trees. *Electron. J. Probab.* **20** 1–51.

[14] Janson, S. (2024) The number of descendants in a random directed acyclic graph. *Random Struc. Algor.* **64** 768–803.

[15] Knuth, D. E. (1997) *The Art of Computer Programming*. Volume 1. 3rd ed. Addison-Wesley.

[16] Na, H. S. and Rapoport, A. (1970) Distribution of nodes of a tree by degree. *Math. Biosci.* **6** 313–329.

[17] Ross, N. (2011) Fundamentals of Stein's method. *Probab. Surv.* **8** 210–293.

[18] Tsukiji, T. and Mahmoud, H. M. (2001) A limit law for outputs in random recursive circuits. *Algorithmica.* **31** 403–412.
[19] Tsukiji, T. and Xhafa, F. (1996) On the depth of randomly generated circuits. In *Algorithms–ESA'96: Fourth Annual European Symposium Barcelona, Spain, September 25-27, 1996,* Vol. 1136. Lecture Notes in Comput. Sci. Springer, 208–220.