ASYMPTOTIC NORMALITY IN RANDOM GRAPHS
WITH GIVEN VERTEX DEGREES

SVANTE JANSON

ABSTRACT. We consider random graphs with a given degree sequence and show, under weak technical conditions, asymptotic normality of the number of components isomorphic to a given tree, first for the random multigraph given by the configuration model and then, by a conditioning argument, for the simple uniform random graph with the given degree sequence. Such conditioning is standard for convergence in probability, but much less straightforward for convergence in distribution as here. The proof uses the method of moments, and is based on a new estimate of mixed cumulants in a case of weakly dependent variables.

The result on small components is applied to give a new proof of a recent result by Barbour and Röllin on asymptotic normality of the size of the giant component in the random multigraph; moreover, we extend this to the random simple graph.

1. INTRODUCTION

Let $G(n, \mathbf{d})$ be a random (simple) graph with n labelled vertices and a given degree sequence $\mathbf{d} = (d_1, \ldots, d_n)$, chosen uniformly at random among all such graphs. (We assume tacitly that $\mathbf{d}$ is such that some such graph exists.) We will denote the vertices by $v_1, \ldots, v_n$; thus v_i has by definition degree d_i.

The standard way to constuct a random graph $G(n, \mathbf{d})$ is by the *configuration model*, which was introduced by Bollobás [11]. As is well-known, this method constructs first a random multigraph, which we denote by $G^*(n, \mathbf{d})$, and then obtains $G(n, \mathbf{d})$ by conditioning on the event that $G^*(n, \mathbf{d})$ is simple; see Section 4.

We are, as most papers in this field, interested in asymptotic results as $n \to \infty$, where the degree sequence $\mathbf{d}_n = (d_i^{(n)})_1^n$ depends on n and satisfies suitable conditions. The standard method is to first prove results for the random multigraph $G^*(n, \mathbf{d}_n)$ and then obtain corresponding results for $G(n, \mathbf{d}_n)$ by conditioning as above. In the present paper, we make the common assumption that the (asymptotic) degree distribution has a finite second moment; see Section 2 for precise assumptions. Then, it is well-known that $\mathbb{P}\big(G^*(n, \mathbf{d}_n) \text{ is simple}\big) \geqslant c$ for some $c > 0$ (at least for large n), see Remark 2.3, and as a consequence, any property that holds w.h.p. (i.e., with probability tending to 1) for $G^*(n, \mathbf{d}_n)$ holds w.h.p. also after conditioning on $G^*(n, \mathbf{d}_n)$ being simple. Hence, the transfer of such results from

the multigraph $G^*(n, \mathbf{d}_n)$ to the simple graph $G(n, \mathbf{d}_n)$ is trivial. (Cf. Bollobás and Riordan [13], where transfer is made possible by far from trivial arguments also when $\mathbb{P}\big(G^*(n, \mathbf{d}_n)$ is simple$\big) \to 0$.)

However, it has repeatedly been remarked that this simple method fails for distributional results, for example that some random variable X_n defined by the graph is asymptotically normal, since probabilities like $\mathbb{P}(X_n \leqslant x)$ may be changed by the conditioning. Hence, although in many cases it seems intuitively clear that a few loops or multiple edges should not affect the asymptotic results, and we still expect the same asymptotic distribution for $G(n, \mathbf{d}_n)$ as for $G^*(n, \mathbf{d}_n)$, it is typically difficult to prove this rigorously. We know only two papers where this has been done for asymptotic normality of some variables, in both cases by first proving a result for $G^*(n, \mathbf{d}_n)$ and then showing that the proof can be modified to work for $G(n, \mathbf{d}_n)$: Janson and Łuczak [24] showing asymptotic normality of the size of the k-core (using a rather complicated extra argument for $G(n, \mathbf{d}_n)$) and Riordan [36] showing asymptotic normality of the size of the giant component in the weakly supercritical case (using a simple extra argument for $G(n, \mathbf{d}_n)$ noting that the proof only uses local explorations involving $o(n)$ vertices).

The purpose of present paper is to do this in another case. Barbour and Röllin [8] recently proved, for $G^*(n, \mathbf{d}_n)$, a theorem on asymptotic normality of a class of "local" statistics that include, for example, the number of small components of a given type. They further used this to show asymptotic normality of the size of the giant component in the supercritical case. Our main results show that the same results on small components and the size of the giant hold for the random simple graph $G(n, \mathbf{d}_n)$. (We also weaken somewhat the technical conditions for these results in [8].) Precise statements are given in Section 3 below.

We achieve these results by the time-honoured method of moments. (Barbour and Röllin [8] use Stein's method.) Using the method of moments, we show joint convergence of, e.g., the number of components of a given type and the numbers of loops and pairs of parallel edges in $G^*(n, \mathbf{d}_n)$; we may then obtain the result for $G(n, \mathbf{d}_n)$ by conditioning on the latter numbers being 0. In order to do this, we thus show convergence of mixed moments. Calculations of means and (co)variances are rather straightforward, and the central part of the proof is to obtain bounds on higher-order cumulants. This is similar to results by Féray [17, 18].

To describe the idea, consider as a simple case the covariance $\mathrm{Cov}(I_1, I_2)$ of two indicators, each indicating that a particular set of vertices (and half-edges) form a copy of a given graph H. If the two sets of vertices are disjoint, then I_1 and I_2 are only weakly dependent; we exploit this by constructing a modification I_2' of I_2 that is independent of I_1, and such that $I_2' \overset{\mathrm{d}}{=} I_2$ and $\mathbb{P}(I_2' \neq I_2 \mid I_1) = O(\mathbb{E}\, I_2 / N)$. This implies $\mathrm{Cov}(I_1, I_2) = \mathbb{E}(I_1 I_2) - \mathbb{E}(I_1 I_2') = \mathbb{E}\big(I_1(I_2 - I_2')\big) = O\big(\mathbb{E}\, I_1 \, \mathbb{E}\, I_2 / N\big)$. The general case is an extension of this, although the details are quite technical, see Section 6. (The idea to construct a suitable independent modification is used also in the Stein coupling constructed by [8], although their modification is both constructed and used differently from our modifications.)

One feature of the construction that may be of independent interest is that we in the proof use the bipartite version of the configuration model to construct a random bipartite graph with the vertices $[n]$ of $G^*(n, \mathbf{d}_n)$ as vertices on one side, and the edges of $G^*(n, \mathbf{d}_n)$ as vertices on the other side; equivalently, we construct first the random bipartite graph obtained by bisecting each edge in $G^*(n, \mathbf{d}_n)$; see Section 4.

Remark 1.1. As remarked by Barbour and Röllin [8], there are not many papers at all proving asymptotic normality for statistics of the multigraph $G^*(n, \mathbf{d}_n)$; apart from [24], [36] and [8] just mentioned, we know of Angel, van der Hofstad and Holmgren [1] (number of loops and multiple edges when the second moment of the degree distribution is infinite; this is obviously not relevant for simple graphs), KhudaBukhsh, Woroszylo, Rempała and Koeppl [30] (an epidemic on the graph), Ball [4] (a more general epidemic model, and the giant component in site or bond percolation), and Athreya and Yogeshwaran [3] (certain statistics in a subcritical case).

Remark 1.2. The method of moments is a very old method. Applications of it are typically messy and lead to long calculations using combinatorial estimates of multiple sums, while other methods may give shorter and more elegant proofs. Nevertheless, it is a powerful method that often works in combinatorial problems. I have seen several cases where results first have been proved by the method of moments and later reproved using other methods. It seems likely that the results here will be another example of this in the future. For example, perhaps a combination of Stein's method for normal approximation and Stein–Chen's method for Poisson approximation might be used instead of the method of moments to show the joint convergence used in our proofs below.

2. Assumptions and notation

2.1. **Some notation.** $[n] := \{1, \ldots, n\}$. $\mathbb{N} := \{0, 1, \ldots\}$. $\mathfrak{S}_n$ is the set of all permutations of $[n]$.

$(n)_r := n!/(n-r)! = n(n-1)\cdots(n-r+1)$ is the descending factorial. Similarly,

$$((n))_r := \frac{n!!}{(n-2r)!!} = n(n-2)\cdots(n-2(r-1)) = 2^r (n/2)_r. \qquad (2.1)$$

$(x)_+ := \max(x, 0)$. We interpret $0/0 := 0$ and $0 \cdot \infty := 0$.

Unspecified limits are as $n \to \infty$; w.h.p. (with high probability) means with probability tending to 1 as $n \to \infty$. $\xrightarrow{\mathrm{d}}$ and $\xrightarrow{\mathrm{p}}$ denote convergence in distribution and probability, respectively. C and c denote positive constants that may be different at each occurrence.

If G is a (multi)graph, we let $V(G)$ denote its vertex set and $E(G)$ its edge set; furthermore, $v(G) := |V(G)|$ and $e(G) := |E(G)|$ are the numbers of vertices and edges of G, and $q(G)$ is its number of components. The number of vertices of degree k is denoted by $n_k(G)$. For convenience, we often write $v \in G$ for $v \in V(G)$ and $|G|$ for $|V(G)|$.

The degree of a vertex $v \in G$ is denoted $d_G(v)$. (A loop in a multigraph contributes 2 to the degree of its endpoint.) Thus $\sum_{v \in G} d_G(v) = 2e(G)$.

The component containing a vertex $v \in G$ is denoted $\mathcal{C}(v) = \mathcal{C}_G(v)$. It may be regarded as a rooted multigraph, i.e., a multigraph with a distinguished vertex, viz. v. We denote the components of G arranged in decreasing order by $\mathcal{C}_1(G), \mathcal{C}_2(G), \ldots, \mathcal{C}_{q(G)}(G)$ (with ties resolved by any fixed rule, e.g. lexicographically).

If X is a random variable and $\mathcal{E}$ an event, then $\mathbb{E}(X; \mathcal{E}) := \mathbb{E}(X\mathbf{1}\{\mathcal{E}\})$.

We may ignore obvious roundings and write e.g. $n^{1/2}$ or εn when we mean the nearest larger or smaller integer.

2.2. **Basic assumptions.** As in the introduction, we assume that for each $n \geqslant 1$, we are given a degree sequence $\mathbf{d}_n = (d_i)_1^n$. Thus $d_i = d_i^{(n)}$ depends on n, as do many other quantities introduced below, but often we omit n from the notation for convenience. Also as above, $G(n, \mathbf{d}_n)$ is the random simple graph with degree sequence $\mathbf{d}_n$, and $G^*(n, \mathbf{d}_n)$ is the random multigraph with degree sequence $\mathbf{d}_n$ given by the configuration model. (We assume tacitly that $\mathbf{d}_n$ is such that a graph with these vertex degrees exists; in particular, $\sum_i d_i$ is even.) Let

$$N := \sum_{i=1}^n d_i. \tag{2.2}$$

Thus the random graph $G^*(n, \mathbf{d}_n)$ has n vertices and $N/2$ edges.

Let $n_k = n_k(\mathbf{d}_n) := |\{i \in [n] : d_i = k\}|$, the number of vertices of degree k in $G^*(n, \mathbf{d}_n)$. Further, let D_n denote the degree of a uniformly random vertex, i.e., D_n is a random variable with the distribution

$$\mathbb{P}(D_n = k) = n_k/n, \qquad k \geqslant 0. \tag{2.3}$$

We will always assume the following.

(A1) D_n, the degree of a randomly chosen vertex, converges in distribution to a random variable D with a finite and positive mean $\mu := \mathbb{E} D$. In other words, there exists a probability distribution $(p_k)_{k=0}^\infty$ such that

$$\frac{n_k}{n} \to p_k = \mathbb{P}(D = k), \qquad k \geqslant 0, \tag{2.4}$$

and $\mu = \sum_{k=0}^\infty k p_k \in (0, \infty)$.

(A2) $\mathbb{E} D_n \to \mathbb{E} D = \mu$. Assuming (A1), this is equivalent to D_n being uniformly integrable.

(See e.g. [19, Theorem 5.5.9] for the equivalence with uniform integrability.) Sometimes we will also use one or several of the following assumptions.

(A3) $\mathbb{E} D_n^2 \to \mathbb{E} D^2 < \infty$. Assuming (A1), this is equivalent to D_n^2 being uniformly integrable. (This will always be assumed when studying $G(n, \mathbf{d}_n)$.)

(A4) $\sup_n \mathbb{E} D_n^m < \infty$ for every $m < \infty$. This implies uniform integrability of every powers D_n^m, and thus, assuming (A1), $\mathbb{E} D_n^m \to \mathbb{E} D^m$ for every $m \geqslant 0$; in particular (A2) and (A3). (We use this strong condition only in a few results.)

(A5) $\mathbb{E} D(D - 2) > 0$. This is the *supercritical* case, when $G^*(n, \mathbf{d}_n)$ w.h.p. has a giant component of order $\Theta(n)$, see Molloy and Reed [32, 33] with refinements in, e.g., [25], [13], [28].

(A6) $p_1 > 0$ and $p_0 + p_1 < 1$. This excludes some less interesting cases; see Remark 3.5.

The degree sequences $\mathbf{d}_n$ will be fixed throughout the paper, and unspecified constants may depend on them through (explicit or implicit) constants in these assumptions, for example the distribution (p_i).

Note that by (2.2), $\mathbb{E}\, D_n = N/n$. Hence, (A2) is equivalent to

$$N/n \to \mu. \tag{2.5}$$

In particular, since $0 < \mu < \infty$ by (A1), $N = \Theta(n)$. Hence, in estimates with unspecified constants, it does not matter whether we use e.g. $O(n)$ or $O(N)$.

Remark 2.1. We assume that $G(n, \mathbf{d}_n)$ has n vertices. This is customary, but as always just for notational convenience. See [13], where the results are formulated with the given $\mathbf{d}_n$ allowed to have arbitrary lengths $\to \infty$.

Remark 2.2. One common version of the configuration model uses random vertex degrees d_i that are i.i.d. copies of a random variable D (e.g. ignoring one half-edge when the sum of degrees is odd). Then, (A1) holds a.s., and thus (with suitable assumptions on D), the results below apply conditioned on the vertex degrees. However, unconditioned results are somewhat different, see Section 12.

Let $d_{\max} := \max_i d_i$. The uniform integrability of D_n^2 in (A3) means that for any (deterministic) sequence $\omega(n) \to \infty$, $\mathbb{E}\, D_n^2 \mathbf{1}\{D_n \geqslant \omega(n)\} = \frac{1}{n} \sum_i d_i^2 \mathbf{1}\{d_i \geqslant \omega(n)\} \to 0$; this implies (choosing e.g. $\omega(n) = n^{1/3}$) that

$$d_{\max} = o\big(n^{1/2}\big). \tag{2.6}$$

(A3) obviously also implies

$$\mathbb{E}\, D_n^2 = O(1). \tag{2.7}$$

Remark 2.3. The condition (2.7) (together with e.g. (A1)) implies

$$\liminf_{n \to \infty} \mathbb{P}\big(G^*(n, \mathbf{d}_n) \text{ is simple}\big) > 0, \tag{2.8}$$

see [21; 23]; hence (A1)–(A3) imply (2.8). The lower bound (2.8) is the basis for most applications of the configuration models to the random simple graph $G(n, \mathbf{d}_n)$, since if often allows simple conditioning as said in the introduction.

When also (2.6) holds, (2.8) can easily be shown using the method of moments; this has been the standard method since the introduction of the configuration model [11], see [21] for a proof with no further assumptions. This method is also the basis of our proofs below.

Remark 2.4. We will assume (A3) for our results for the simple graph $G(n, \mathbf{d}_n)$. In fact, the main results really require only (2.6) and (2.7); by considering subsequences we may then assume that $\mathbb{E}\, D_n^2 \to \mu_2$ for some $\mu_2 < \infty$, and the proofs then hold with $\mathbb{E}\, D^2$ replaced by μ_2 in e.g. (7.7). However, we prefer to state the results for the more natural condition (A3). (An example satisfying (2.6) and (2.7) but not (A3) is obtained from any example satisfying (A1)–(A3) by changing the degrees of $n^{1/3}$ vertices to $n^{1/3}$.)

As said above, it is shown in [21; 23] that (2.8) holds also without assuming (2.6), but the proofs are then somewhat more complicated, and we do not know whether they can be used to replace the assumption (A3) by (2.7) in the theorems below. In fact, it seems likely that the results extend also to cases without (2.7), but that would require completely different methods.

2.3. **The corresponding branching process.** The usual exploration process reveals the edges (i.e., pairings of half-edges) in $G^*(n, \mathbf{d}_n)$ one by one as they are needed, starting with the half-edges at a vertex v and then continuing with the unpaired half-edges at the neighbours of v, and so on, until the component $\mathcal{C}(v)$ is fully explored. Let V be a uniformly random vertex in $G^*(n, \mathbf{d}_n)$. As is well-known, see e.g. [13, Lemma 4] for a formal statement, assuming (A1)–(A2), the exploration process of $\mathcal{C}(V)$ may be approximated by a branching process $\mathcal{T}$ (regarded as a rooted tree, finite or infinte), in the sense that for any fixed K, the first K generations of the two processes may be coupled such that they are isomorphic w.h.p. (In particular, if one of the processes has only $k < K$ non-empty generations, then so has the other w.h.p.) The branching process $\mathcal{T}$ is a Galton–Watson process where the root has offspring distribution D as in (A1), and all other vertices have offspring distribution $\widehat{D} - 1$ where $\widehat{D}$ has the size-biased distribution

$$\mathbb{P}(\widehat{D} = k) := \hat{p}_k := \frac{k p_k}{\mu}, \qquad k \geqslant 1. \tag{2.9}$$

For a rooted unlabelled tree T, we define

$$p_T := \mathcal{P}(\mathcal{T} \cong T), \tag{2.10}$$

with $\cong$ meaning isomorphism (i.e., equality) as unlabelled rooted graphs. Similarly, if T is an unrooted unlabelled tree, we define p_T by (2.10), now interpreting $\cong$ as isomorphism (equality) of unlabelled unrooted graphs.

For random variables that are functionals $g(\mathcal{T})$ of $\mathcal{T}$, we write

$$\mathbb{E}^\dagger g(\mathcal{T}) := \mathbb{E}\big(g(\mathcal{T}); |\mathcal{T}| < \infty\big) = \sum_T p_T g(T), \tag{2.11}$$

summing over finite unlabelled (rooted or unrooted) trees T. Obviously, it here suffices that g is defined for finite trees.

We denote the probability generating function of D by

$$f(z) := \mathbb{E}\, z^D = \sum_k p_k z^k. \tag{2.12}$$

In the supercritical case (A5), let ζ be the unique root in $[0, 1)$ of

$$f'(\zeta) = \mu\zeta. \tag{2.13}$$

(If (A5) does not hold, we may take $\zeta := 1$, which always satisfies (2.13).) As is well-known, then $\mathbb{P}\big(|\mathcal{T}| < \infty\big) = f(\zeta)$. See further Sections 3.3 and 11.

3. Main results

The main results are stated below. Proofs are given in later sections. We begin with some notation for subgraph counts.

Let G and H be (multi)graphs. (Think of G as big and H small.) We often regard the small graph H as unlabelled; in particular, when we talk about

several distinct graphs H, we mean non-isomorphic. However, for formal definitions it is convenient to regard both G and H as labelled, meaning that both vertices and edges, and also half-edges, are labelled. (For simple graphs it suffices to label vertices, since the edges are identified by their endpoints. Similarly, half-edges need to be labelled only for loops.) We count subgraphs of G isomorphic to H in four different ways:

- $Z_H(G)$ is the number of *labelled* copies of H in G.
- $Z_H^{\mathsf{u}}(G)$ is the number of *unlabelled* copies of H in G.
- $\overline{Z}_H(G)$ is the number of *labelled isolated* copies of H in G.
- $\overline{Z}_H^{\mathsf{u}}(G)$ is the number of *unlabelled isolated* copies of H in G.

Thus, $Z_H(G)$ can be defined as the number of injective maps $H \to G$, mapping vertices to vertices, edges to edges and half-edges to half-edges such that the relations between them are preserved. $\overline{Z}_H(G)$ is the number of such maps such that the vertices in the image H' of H have no other edges than those in H', i.e., the degrees are preserved. Furthermore,

$$Z_H^{\mathsf{u}}(G) = \frac{1}{\operatorname{aut}(H)} Z_H(G), \qquad \overline{Z}_H^{\mathsf{u}}(G) = \frac{1}{\operatorname{aut}(H)} \overline{Z}_H(G), \qquad (3.1)$$

where $\operatorname{aut}(H) = Z_H(H)$ is the number of automorphisms of H. If H is connected, then $\overline{Z}_H^{\mathsf{u}}(G)$ is the number of components of G isomorphic to H.

The simple relations (3.1) shows that it is equivalent to study labelled or unlabelled copies. Nevertheless we will consider both since it often is natural to count unlabelled copies (for example when counting components), but labelled copies often (but not always) are more convenient in our proofs.

We simplify the notation and write $Z_H := Z_H\big(G(n, \mathbf{d}_n)\big)$ and $Z_H^* := Z_H\big(G^*(n, \mathbf{d}_n)\big)$, and similarly for the other versions of subgraph counts.

Example 3.1. Let C_1 be a loop and C_2 a double edge. Then a multigraph G is simple if and only if $Z_{\mathsf{C}_1}^{\mathsf{u}}(G) = Z_{\mathsf{C}_2}^{\mathsf{u}}(G) = 0$. If particular, we obtain $G(n, \mathbf{d}_n)$ from $G^*(n, \mathbf{d}_n)$ by conditioning on the event $Z_{\mathsf{C}_1}^{\mathsf{u}*} = Z_{\mathsf{C}_2}^{\mathsf{u}*} = 0$.

3.1. **Small components.** Our first theorem gives the asymptotic distributions for the number of small components of different types in $G^*(n, \mathbf{d}_n)$. It is well-known that most small components are trees. More precisely, for a tree T, there is typically a linear number of components T, and asymptotic normality was recently shown by Barbour and Röllin [8] using Stein's method. We give a new proof of this (under somewhat weaker conditions), since this is the basis of our work below. We complement this with the easy results that the number of components isomorphic to a given conneted unicyclic graph has an asymptotic Poisson distribution, and that there w.h.p. is no small (i.e. fixed size) component with more than one cycle. The latter results follow by standard moment calculations and are presumably known, although we do not know any specific reference.

Theorem 3.2 (Mainly [8]). *Consider the random multigraph $G^*(n, \mathbf{d}_n)$ and assume (A1)–(A2) and (A6).*

(i) *If H is a tree, then*

$$\frac{\overline{Z}_H^{\mathsf{u}*} - \mathbb{E}\,\overline{Z}_H^{\mathsf{u}*}}{\sqrt{n}} \xrightarrow{\mathrm{d}} N\big(0, \sigma_H^2\big) \qquad (3.2)$$

for some $\sigma_H^2 = \sigma_{H,H} \geqslant 0$ given by (3.7) below. Furthermore,

$$\mathbb{E}\,\overline{Z}_H^{\mathsf{u}*}/n \to \lambda_H := \frac{\mu^{-e(H)}}{\mathrm{aut}(H)} \prod_{u \in H} p_{d_H(u)} d_H(u)! \tag{3.3}$$

and, if $v(H) > 1$,

$$\sigma_H^2 > 0 \iff \lambda_H > 0 \iff d_H(u) \in \{k : p_k > 0\}\ \forall u \in H. \tag{3.4}$$

Moreover, for every tree H,

$$\lambda_H = p_H/|H|, \tag{3.5}$$

where p_H is given by (2.10) (for unrooted H).

(ii) *If H is a connected unicyclic multigraph, so $e(H) = v(H)$, then*

$$\overline{Z}_H^{\mathsf{u}*} \xrightarrow{\ \mathrm{d}\ } \mathrm{Po}(\lambda_H), \tag{3.6}$$

with λ_H as in (3.3).

(iii) *If H is a connected multigraph with more than one cycle, i.e., $e(H) > v(H)$, then $\overline{Z}_H^{\mathsf{u}*} = 0$ w.h.p.*

Moreover, the limits in (i)–(iii) *hold jointly, for any finite number of distinct connected multigraphs H_i, with a joint limit $N(0, \Sigma)$ for the trees, for a covariance matrix $\Sigma = (\sigma_{H_i, H_j})$ given by*

$$\sigma_{H_1, H_2} := \delta_{H_1, H_2} \lambda_{H_1} + \lambda_{H_1} \lambda_{H_2} \Big(\frac{2e(H_1)e(H_2)}{\mu} - \sum_{k \geqslant 0} \frac{n_k(H_1)n_k(H_2)}{p_k} \Big), \tag{3.7}$$

and with the Poisson limits in (ii) *independent of each other and of the joint normal limit in* (i). *Furthermore, this holds with convergence of all mixed moments. In particular, if H_1 and H_2 are trees, then*

$$\mathrm{Cov}\big(\overline{Z}_{H_1}^{\mathsf{u}*}, \overline{Z}_{H_2}^{\mathsf{u}*}\big) = \sigma_{H_1, H_2} n + o(n). \tag{3.8}$$

Moreover, if each tree H_i has $v(H_i) > 1$ and satisfies the condition in (3.4), then the covariance matrix Σ is non-singular.

In (3.7), δ_{H_1, H_2} is the Kronecker delta, equal to 1 when $H_1 = H_2$ (as unlabelled graphs); furthermore, recall that $0/0 = 0$ and $0 \cdot \infty = 0$.

Remark 3.3. The case $v(H) = 1$, i.e. $v = \mathsf{K}_1$, has to be excepted in (3.4), since trivially $Z_{\mathsf{K}_1}^{\mathsf{u}*} = n_0$ is deterministic, so $\sigma_{\mathsf{K}_1}^2 = 0$.

Remark 3.4. The moment convergence in (3.2) implies that if H is a tree such that $\sigma_H^2 > 0$, then (3.2) can be written

$$\frac{\overline{Z}_H^{\mathsf{u}*} - \mathbb{E}\,\overline{Z}_H^{\mathsf{u}*}}{\big(\mathrm{Var}\,\overline{Z}_H^{\mathsf{u}*}\big)^{1/2}} \xrightarrow{\ \mathrm{d}\ } N(0, 1), \tag{3.9}$$

so $\overline{Z}_H^{\mathsf{u}*}$ really is asymptotically normal. The same applies in other theorems below, e.g. Theorem 3.9.

Remark 3.5. The assumption (A6) excludes the two extremal cases $p_1 = 0$ and $p_0 + p_1 = 1$, which are less interesting in the present context; in these cases, there are only a few $(o(n))$ small components except (possibly) isolated vertices and edges. Theorem 3.2 hold also when (A6) fails, except that then

$\sigma_H^2 = 0$ for every tree H (and (3.4) may fail) and thus e.g. (3.2) says only that the variable converges to 0 in probability; cf. Theorem 3.10 and Remark 7.11, where (A6) is not assumed.

Remark 3.6. Cf. the similar results for Erdős–Rényi graphs $G(n, p)$ and $G(n, m)$, shown by e.g. [16], [6], [7]; see Example 12.3.

Remark 3.7. In spite of (3.3), we cannot in general replace $\mathbb{E} \overline{Z}_H^{\mathsf{u}*}$ by its asymptotic value $n\lambda_H$ in (3.2); the reason is that this would require $\mathbb{E} \overline{Z}_H^{\mathsf{u}*} = n\lambda_H + o(n^{1/2})$, and this rate of convergence in (3.3) does not hold without further assumptions on the rate of convergence in (2.4). In particular, it does not hold in the case of random vertex degrees mentioned in Remark 2.2, see Section 12.

Remark 3.8. As said above, we reprove the result by Barbour and Röllin [8] on tree components. We do not treat the more general class of local statistics studied in [8], and leave it as an open problem whether our methods apply to them in full generality. On the other hand, we include below results on subgraph counts not covered by [8].

One of our main results is that Theorem 3.2 transfers to the simple random graph $G(n, \mathbf{d}_n)$, under weak assumptions. Obviously, we have to consider only simple graphs H, since $Z_H^{\mathsf{u}} = 0$ if H contains a loop or a multiple edge.

Theorem 3.9. *Assume* (A1)–(A3) *and* (A6). *Then, for simple graphs* H, *the results in Theorem 3.2 hold also for the random simple graph* $G(n, \mathbf{d}_n)$ *and variables* $\overline{Z}_H^{\mathsf{u}}$, *with the same* λ_H, σ_H^2 *and* Σ. *Furthermore, if* H *is a tree, then*

$$\mathbb{E} \overline{Z}_H^{\mathsf{u}} - \mathbb{E} \overline{Z}_H^{\mathsf{u}*} = o(n^{1/2}), \tag{3.10}$$

and thus it does not matter whether we normalize $\overline{Z}_H^{\mathsf{u}}$ *as in (3.2) using* $\mathbb{E} \overline{Z}_H^{\mathsf{u}}$ *or* $\mathbb{E} \overline{Z}_H^{\mathsf{u}*}$.

3.2. **Small subgraphs.** With a stronger moment assumption on the degrees, we have similar results for the number of copies (not necessarily isolated) of a given tree. We assume (A4), i.e., that every moment $\mathbb{E} D_n^m$ is bounded, and thus in particular that D has finite moments of all orders. (This implies (A2)–(A3) as said above.)

Theorem 3.10. *Assume* (A1) *and* (A4).
  (i) *For every tree* T *there exists* $\tilde{\sigma}_T^2 \geqslant 0$ *such that*

$$\frac{Z_T^{\mathsf{u}*} - \mathbb{E} Z_T^{\mathsf{u}*}}{\sqrt{n}} \xrightarrow{\mathrm{d}} N(0, \tilde{\sigma}_T^2). \tag{3.11}$$

  Moreover, (3.11) holds jointly for any number of such trees T, *with a joint limit* $N(0, \tilde{\Sigma})$ *for some covariance matrix* $\tilde{\Sigma}$. *Furthermore, the limits hold with convergence of all moments.*
  (ii) *The results in* (i) *hold also for the random simple graph* $G(n, \mathbf{d}_n)$ *and variables* Z_T^{u}, *with the same* $\tilde{\sigma}_T^2$ *and* $\tilde{\Sigma}$. *Furthermore,*

$$\mathbb{E} Z_T^{\mathsf{u}} - \mathbb{E} Z_T^{\mathsf{u}*} = o(n^{1/2}), \tag{3.12}$$

  and thus it does not matter whether we normalize Z_T^{u} *as in (3.11) using* $\mathbb{E} Z_T^{\mathsf{u}}$ *or* $\mathbb{E} Z_T^{\mathsf{u}*}$.

Remark 3.11. We consider only trees in Theorem 3.10. If H is a cycle, then the proofs below show that $Z_H^{\mathsf{u}*}$ and Z_H^{u} are asymptotically Poisson distributed. However, for general unicyclic H this is not true; this is similar to the case for $G(n,p)$, see e.g. [14] and [27, Example 3.21]. It can be shown that for a general connected unicyclic H, $Z_H^{\mathsf{u}*}$ and Z_H^{u} converge to a compund Poisson distribution. This is besides the point of the present paper, so we leave the details to the reader.

Remark 3.12. We leave explicit calculations of asymptotics of means and variances in Theorem 3.10 to the reader. Note that (for the simple graph $G(n,\mathbf{d}_n)$) there are further trivially deterministic cases: $Z_{\mathsf{K}_2}^{\mathsf{u}} = N/2$, the number of edges in $G(n,\mathbf{d}_n)$; similarly, for any $r \geqslant 1$, $Z_{\mathsf{K}_{1,r}}^{\mathsf{u}} = \sum_{i=1}^{n} \binom{d_i}{r}$. We do not know whether there also are further cases where the variance (perhaps of some linear combination) is $o(n)$ and thus vanishes in the limit taken in (3.11).

Remark 3.13. The condition (A4) assumes that all moments of the degree distribution are bounded. This is presumably stronger than necessary; it seems likely that it is enough that $\sup_n \mathbb{E}\, D_n^M < \infty$ for some M depending on T for convergence in distribution (3.11), although (A4) presumably is required for moment convergence. Some condition of this type is necessary, since otherwise $Z_T^{\mathsf{u}*}$ may be dominated by a few vertices of high degrees, as in the following example.

Example 3.14. Let $\mathbf{d}_n$ have one vertex v_1 of degree $n^{0.4}$, $n^{0.5}$ vertices u_i of degree $n^{0.1}$, $n/2$ vertices of degree 3 and the rest, about $n/2$, of degree 1. Then (A1)–(A3) are satisfied, with $p_1 = p_3 = 1/2$ and $\mu = 2$. Let H be the tree with $v(H) = 22$ obtained by taking two disjoint stars $\mathsf{K}_{1,10}$ and joining their central vertices by an edge.

Let X_n be the number of edges in $G^*(n,\mathbf{d}_n)$ between v_1 and a vertex of degree $n^{0.1}$. Each such edge is the central edge in $\approx 2n^{10\cdot 0.4+10\cdot 0.1} = 2n^5$ labelled copies of H, while the $O(n)$ other edges are central edges in at most $O(n^2)$ copies of H each. Hence, $Z_H^* = \big(2+o(1)\big)n^5 + O\big(n^3\big)$. Furthermore, it is easy to see (e.g. by the method of moments) that $X_n \xrightarrow{\mathrm{d}} \mathrm{Po}(1/2)$. Hence,

$$Z_H^* / \big(2n^5\big) \xrightarrow{\mathrm{d}} \mathrm{Po}\big(1/2\big). \tag{3.13}$$

Thus, after suitable norming, Z_H^* and $Z_H^{\mathsf{u}*}$ are asymptotically Poisson distributed and not normal.

3.3. **Giant component and susceptibility.** It is well-known that in the supercritical case when (A5) holds, there is w.h.p. a unique component $\mathcal{C}_1$ of order $\Theta(n)$ in $G^*(n,\mathbf{d}_n)$ or $G(n,\mathbf{d}_n)$, known as the giant component; see Molloy and Reed [32, 33] and, e.g., [25], [13], [28]; moreover, $\mathbb{E}\,|\mathcal{C}_1|/n \to 1 - f(\zeta)$, which we recall from Section 2.3 is the survival probability of the branching process $\mathcal{T}$. Ball and Neal [5] proved (under some extra technical conditions, in particular the existence of a third moment $\mathbb{E}\, D^3 < \infty$) that the variance satisfies $\mathrm{Var}\,|\mathcal{C}_1|/n \to \sigma^2$, with σ^2 given by (3.15) below. Moreover, Barbour and Röllin [8] proved (under the same techical conditions) that for the random multigraph $G^*(n,\mathbf{d}_n)$, the size $|\mathcal{C}_1|$ is asymptotically normal.

We reprove this here with our methods (removing some unnecessary conditions); moreover, we show that the result holds for the simple random graph $G(n, \mathbf{d}_n)$ too. (Similar results for $G(n, p)$ and $G(n, m)$ have been known for a long time, see e.g. [37], [34], [35], [12].)

Remark 3.15. In the weakly supercritical case, where $\mathbb{E}\, D_n(D_n - 2) \to 0$ but $\mathbb{E}\, D_n(D_n - 2) \gg n^{-1/3}$, Riordan [36] showed asymptotic normality of the size (and nullity) of the giant component, both for $G^*(n, \mathbf{d}_n)$ and $G(n, \mathbf{d}_n)$ (assuming that the degrees are bounded), using methods different from ours. Note that in this case, the giant is smaller: $\mathbb{E}\, |\mathcal{C}_1| \ll n$, but $\mathrm{Var}\, |\mathcal{C}_1| \gg n$. It does not seem to be possible to prove these results by the method in the present paper.

Theorem 3.16 (Partly [5] and [8])**.** (i) *Assume* (A1)–(A2) *and* (A5)–(A6). *Then the size* $|\mathcal{C}_1|$ *of the giant component in* $G^*(n, \mathbf{d}_n)$ *has an asymptotically normal distribution:*

$$\frac{|\mathcal{C}_1| - \mathbb{E}\, |\mathcal{C}_1|}{\sqrt{n}} \xrightarrow{\mathrm{d}} N(0, \sigma^2), \qquad (3.14)$$

where $\sigma^2 > 0$ *is given by, with* f *and* ζ *as in* (2.12)–(2.13),

$$\sigma^2 = f(\zeta) + \frac{\mu^2 \zeta^2}{\mu - f''(\zeta)} + 2\frac{\mu^3 \zeta^4}{(\mu - f''(\zeta))^2} - f(\zeta^2) - 2\frac{\mu \zeta^2}{\mu - f''(\zeta)} f'(\zeta^2)$$

$$- \frac{\mu^2}{(\mu - f''(\zeta))^2}\left(\zeta^4 f''(\zeta^2) + \zeta^2 f'(\zeta^2)\right). \qquad (3.15)$$

Furthermore, $\mathrm{Var}\, |\mathcal{C}_1|/n \to \sigma^2$.

(ii) *Assume* (A1)–(A3) *and* (A5)–(A6). *Then the results of* (i) *hold for* $G(n, \mathbf{d}_n)$ *too, with the same* σ^2.

The formula (3.15) is given (in an equivalent form) by Ball and Neal [5]. They also state that $\sigma^2 > 0$, but the proof seems omitted; we give a proof in Section 10 under our (weaker) assumptions.

The proof of Theorem 3.16 is given in Sections 9–11; it is based on the results above and a truncation argument; as in [5] and [8], we obtain results for the giant by subtracting all small components, and the main ideas in this part of the proof are similar. As part of the proof, we show a more general result in Theorem 3.17; we first introduce more notation.

A *graph functional* ψ is a real-valued functional $\psi(H)$ defined for unlabelled multigraphs H. Let ψ be a graph functional, and define for a multigraph G

$$\Psi(G) := \sum_{v \in G} \psi(\mathcal{C}(v)) = \sum_{j=1}^{q(G)} |\mathcal{C}_j| \psi(\mathcal{C}_j) = \sum_{H} |H| \psi(H) \overline{Z}_H^{\mathsf{u}}(G), \qquad (3.16)$$

summing over all unlabelled connected H; the middle equality holds because each component $\mathcal{C} \cong H$ is counted $|\mathcal{C}| = |H|$ times in the sum over v.

In the supercritical case, the second sum in (3.16) may be dominated by the single term coming from the giant component, and it is sometimes more

interesting to exclude it and consider only small components. We define

$$\Psi^\dagger(G) := \sum_{j=2}^{q(G)} |\mathcal{C}_j| \psi(\mathcal{C}_j). \qquad (3.17)$$

Recall also the notation $\mathbb{E}^\dagger$ defined in (2.11).

Theorem 3.17. (i) *Assume* (A1)–(A2) *and* (A5). *Let* ψ *be a graph functional such that*

$$\psi(H) = O\big(e(H)^m + 1\big) \qquad (3.18)$$

for some constant m and all connected H. Define $\Psi^\dagger$ by (3.17). *Then*

$$\mathbb{E}\,\Psi^\dagger\big(G^*(n, \mathbf{d}_n)\big) = n\,\mathbb{E}^\dagger\,\psi(\mathcal{T}) + o(n), \qquad (3.19)$$

$$\mathrm{Var}\,\Psi^\dagger\big(G^*(n, \mathbf{d}_n)\big) = n\sigma_\psi^2 + o(n), \qquad (3.20)$$

and

$$\frac{\Psi^\dagger\big(G^*(n, \mathbf{d}_n)\big) - \mathbb{E}\,\Psi^\dagger\big(G^*(n, \mathbf{d}_n)\big)}{\sqrt{n}} \xrightarrow{\ \mathrm{d}\ } N(0, \sigma_\psi^2), \qquad (3.21)$$

where, summing over unlabelled, unrooted trees T_1, T_2,

$$\sigma_\psi^2 := \sum_{T_1, T_2} |T_1||T_2|\psi(T_1)\psi(T_2)\sigma_{T_1, T_2}$$

$$= \mathbb{E}^\dagger\big(|\mathcal{T}|\psi(\mathcal{T})^2\big) + \frac{2}{\mu}\big(\mathbb{E}^\dagger(e(\mathcal{T})\psi(\mathcal{T}))\big)^2 - \sum_{k \geqslant 0} \frac{1}{p_k}\big(\mathbb{E}^\dagger(n_k(\mathcal{T})\psi(\mathcal{T}))\big)^2 \qquad (3.22)$$

with all sums and expectations absolutely convergent.

   (ii) *Assume* (A1)–(A3) *and* (A5). *Then the results in* (i) *hold for $G(n, \mathbf{d}_n)$ too.*

Remark 3.18. If $p_k = 0$ for some k, then $\mathcal{T}$ cannot contain any vertices of degree k, so $n_k(\mathcal{T}) = 0$. Hence the terms in the last sum in (3.22) with $p_k = 0$ vanish and are not a problem (recall that we interpret $0/0$ as 0); we could write the sum as a sum over $\{k : p_k > 0\}$. The same applies to other sums below with p_k in the denominator.

Remark 3.19. We do not assume (A6) in Theorem 3.17, since the results hold also when (A6) fails (so $p_1 = 0$ by (A5)), but this case is less interesting since then $\sigma_\psi^2 = 0$, because $|\mathcal{T}| = \infty$ a.s. and thus $\mathbb{E}^\dagger = 0$.

Example 3.20. Let $\psi(H) := 1$. Then (3.17) yields

$$\Psi^\dagger(G) = \sum_{j=2}^{q(G)} |\mathcal{C}_j| = |G| - |\mathcal{C}_1|. \qquad (3.23)$$

Hence the asymptotic normality of $\mathcal{C}_1$ in (3.14) is equivalent to the asymptotic normality of $\Psi^\dagger(G^*(n, \mathbf{d}_n))$ and $\Psi^\dagger(G(n, \mathbf{d}_n))$, which follows from Theorem 3.17, with $\sigma^2 = \sigma_\psi^2$. We will show in Lemma 11.1 that this yields (3.15), and in Lemma 10.5 that $\sigma^2 > 0$, assuming (A5) and (A6).

Similarly, we may take $\psi_1(H) := e(H)/|H|$, so that (3.17) yields, for $G = G^*(n, \mathbf{d}_n)$ or $G(n, \mathbf{d}_n)$,

$$\Psi_1^\dagger(G) = \sum_{j=2}^{q(G)} e(\mathcal{C}_j) = N/2 - e(\mathcal{C}_1). \qquad (3.24)$$

Theorem 3.17 thus shows asymptotic normality of $e(\mathcal{C}_1)$; more generally, taking a linear combinaton $a\psi(H) + b\psi_1(H) = a + be(H)/|H|$, we obtain joint asymptotic normality of $|\mathcal{C}_1|$ and $e(\mathcal{C}_1)$. (Cf. [36] for the weakly supercritical case not studied here.) We conjecture that the limit distribution has a non-singular covariance matrix; however, we have not verified this; see Remark 10.6.

Example 3.21. Let $\psi(H) = |H|$; then (3.17) yields

$$n^{-1}\Psi^\dagger(G) = n^{-1} \sum_{j=2}^{q(G)} |\mathcal{C}_j|^2, \qquad (3.25)$$

which is the modified susceptibility $\widehat{\chi}(G)$ studied in [22]. It was shown there that, assuming (A1)–(A3) and (A5)–(A6), $\widehat{\chi}(G(n, \mathbf{d}_n))$ converges in probability to $\mathbb{E}^\dagger |\mathcal{T}|$. Theorem 3.17 yields convergence of $\mathbb{E}\,\widehat{\chi}(G(n, \mathbf{d}_n))$ to the same limit and, moreover, asymptotic normality

$$n^{1/2}\big(\widehat{\chi}(G(n, \mathbf{d}_n)) - \mathbb{E}\,\widehat{\chi}(G(n, \mathbf{d}_n))\big) \xrightarrow{\text{d}} N\big(0, \sigma^2\big), \qquad (3.26)$$

for some σ^2 given by (3.22). This σ^2 could be evaluated explicitly similarly to Lemma 11.1, but we have not done so. We show in Example 10.7 that $\sigma^2 > 0$.

Remark 3.22. We assume (A5) (supercriticality) in Theorems 3.16 and 3.17. This is for the truncation argument allowing us to ignore large components. (For example Lemma 9.4.)

In the critical case $\mathbb{E}\,D(D-2) = 0$, there are typically a few large components of order $\Theta(n^{2/3})$, see [36, Theorem 1.3], and Theorem 3.16 does not hold; thus Theorem 3.17 does not hold.

In the subcritical case, Theorem 3.16 is not interesting, but other functionals are, for example the susceptibility in Example 3.21. We conjecture that Theorem 3.17 holds under suitable conditions, now for Ψ rather than $\Psi^\dagger$ (the difference should be negligible), but that stronger conditions than above (on the degree sequences or on the graph functional, or both) are required to keep the contribution from large components small. Perhaps (A4) will do; alternatively, (3.18) may be replaced by a stronger assumption. The paper is long as it is, and we do not consider this case further.

4. The configuration model

The standard way to constuct a random multigraph $G^*(n, \mathbf{d}_n)$ is by the *configuration model*, which was introduced by Bollobás [11]. (See [9; 38] for related models and arguments.) As is well-known, we then assign a set of d_i *half-edges* to each vertex v_i; this gives a total of N half-edges, and we choose a perfect matching of them uniformly at random. This defines $G^*(n, \mathbf{d}_n)$ by regarding each pair of half-edges in the matching as an edge.

We will use this standard version in Section 8, but in the main part of our arguments (Section 7), it will be convenient to use a variation of this construction that yields the same result in a somewhat more circuitous way. To see it, we may start with the standard construction above, but also assume that we label the edges by putting a *cuff* on each edge, with the cuffs labelled $1, \ldots, N/2$ (uniformly at random). Furthermore, each cuff i has two half-edges, labelled $2i-1$ and $2i$ (randomly), joined to one each of the half-edges making the edge. (The half-edges are now really quarter-edges, but we keep the name.)

Our version of the configuration model can now be described as follows. Let $\Upsilon_i = \{\omega_{i1}, \ldots, \omega_{id_i}\}$ be the set of half-edges assigned to vertex v_i in the standard model above. Let $\Upsilon := \bigcup_i \Upsilon_i$ be set of all half-edges and label them (arbitrarily) as $\alpha_1, \ldots, \alpha_N$. Take also a second set of half-edges $\beta_1, \ldots, \beta_N$ representing the half-edges at the cuffs, with β_{2j-1} and β_{2j} at cuff j, which we now for emphasis denote χ_j. Let $\pi \in \mathfrak{S}_N$ be a uniformly random permutation, and join each α_i to $\beta_{\pi(i)}$. We denote the result by $\widehat{G}(n, \mathbf{d}_n)$. We interpret $\widehat{G}(n, \mathbf{d}_n)$ as a bipartite graph by merging all half-edges in Υ_i into the vertex v_i, and the half-edges β_{2j-1} and β_{2j} into the cuff χ_j; thus the two vertex sets are $\{v_i : i \in [n]\}$ and $\{\chi_j : j \in [N/2]\}$.

Finally, to obtain $G^*(n, \mathbf{d}_n)$, we merge the two edges at each cuff into one, and forget the cuffs. This yields evidently the same result as the standard configuration model.

The original vertices will sometimes be called *real* vertices.

Remark 4.1. The reader may recognize that our construction is just the standard configuration model construction of a random bipartite graph with the real vertices on one side and the $N/2$ cuffs (each of degree 2) on the other side, followed by contractions eliminating all cuffs.

5. CUMULANTS

Our proofs are based on the method of moments, in the form using cumulants, see e.g. [27, Section 6.1]. We denote the r-th cumulant of a random variable X by $\kappa_r(X)$, and the mixed cumulant of random variables $X_1, \ldots, X_r$ by $\kappa(X_1, \ldots, X_r)$. (The variables are assumed to have finite moments.) We recall the following properties of mixed cumulants, see e.g. [27, p. 147] or [31].

(κ1) $\kappa_r(X) = \kappa(X, \ldots, X)$ (r times).

(κ2) $\kappa(X_1, \ldots, X_r)$ is multilinear in $X_1, \ldots, X_r$.

(κ3) $\kappa(X_1, \ldots, X_r) = 0$ if $\{X_1, \ldots, X_r\}$ can be partitioned into two non-empty sets of random variables that are independent of each other.

(κ4) $\kappa(X_1, \ldots, X_r) = \sum_{I_1, \ldots, I_q} (-1)^{q-1} (q-1)! \prod_{p=1}^{q} \mathbb{E} \prod_{j \in I_p} X_j$, summing over all partitions of $\{1, \ldots, r\}$ into non-empty sets $\{I_1, \ldots, I_q\}$, $q \geqslant 1$.

(κ5) $\mathbb{E}(X_1 \cdots X_r) = \sum_{I_1, \ldots, I_q} \prod_{p=1}^{q} \kappa(\{X_i : i \in I_p\})$, summing as in ($\kappa$4).

6. THE MAIN LEMMA

In this section we state and prove Lemma 6.1 below, whch is the central part of the proof of the results in the present paper. Although the lemma is motivated by its application to the configuration model and random graphs,

we state it in a self-contained way. The lemma could be derived using the general theory by [18], and is very similar to [17, Theorem 1.4], but we give a complete proof, using the following notation.

Suppose that $r \geqslant 1$ and that for each $i \in [r]$ we are given $\ell_i \geqslant 1$ and two sequences $\alpha_{i1}, \ldots, \alpha_{i\ell_i} \in [N]$ and $\beta_{i1}, \ldots, \beta_{i\ell_i} \in [N]$. (These are fixed throughout this section.)

Let $\pi \in \mathfrak{S}_N$ be uniformly random and define the random indicator variables

$$Y_i := \mathbf{1}\{\pi(\alpha_{ij}) = \beta_{ij}, \forall j \in [\ell_i]\} = \prod_{j=1}^{\ell_i} \mathbf{1}\{\pi(\alpha_{ij}) = \beta_{ij}\}. \qquad (6.1)$$

Our goal in this section is to estimate the mixed cumulant $\kappa(Y_1, \ldots, Y_r)$.

Let $A_i := \{\alpha_{ij} : j \in [\ell_i]\}$ and $B_i := \{\beta_{ij} : j \in [\ell_i]\}$. Let Γ be the graph with vertex set $[r]$ and an edge ik if $A_i \cap A_k \neq \emptyset$ or $B_i \cap B_k \neq \emptyset$. (In other words, there is an edge ik when Y_i and Y_k use a common α or a common β in the definition (6.1).) The connected components of Γ are called *blocks*. Let b be the number of blocks, and denote the blocks by $\Gamma_1, \ldots, \Gamma_b$ (in some order, e.g. lexicographic); thus $1 \leqslant b \leqslant r$ and $\Gamma_1, \ldots, \Gamma_b$ form a partition of $[r]$. Furthermore, let

$$e := \Big| \bigcup_{i \in [r]} \{(\alpha_{ij}, \beta_{ij}) : j \in [\ell_i]\} \Big|, \qquad (6.2)$$

i.e., the number of distinct pairs $(\alpha_{ij}, \beta_{ij})$.

Lemma 6.1. *With notations as above,*

$$\big|\kappa(Y_1, \ldots, Y_r))\big| \leqslant C N^{-(b-1)-e}. \qquad (6.3)$$

where C is a constant that may depend on r and $\ell_1, \ldots, \ell_r$ but not on N.

Remark 6.2. The estimate $\big|\kappa(Y_1, \ldots, Y_r))\big| \leqslant C N^{-e}$ is straightforward; in particular, the case $b = 1$ of (6.3) is easy. If the indicators $\mathbf{1}\{\pi(\alpha) = \beta\}$ were independent for disjoint pairs (α, β), then the variables Y_i belonging to different blocks would be independent, and thus the mixed cumulant would vanish when $b \geqslant 2$. Of course, in our setting these indicators are not independent, but they are only weakly dependent and Lemma 6.1 yields a substitute with an estimate that becomes smaller when the number of blocks gets larger. We will see later that this is sufficient for our purposes.

Proof of Lemma 6.1. In this proof, C will denote various constants that may depend on r and $\ell_1, \ldots, \ell_r$, but not on N or α_{ij}, β_{ij}. The same holds for the implied constants in $O(\ldots)$.

For each set $R \subseteq [r]$, let $A(R) := \bigcup_{i \in R} A_i$, $B(R) := \bigcup_{i \in R} B_i$, and, generalizing (6.2),

$$e(R) := \Big| \bigcup_{i \in R} \{(\alpha_{ij}, \beta_{ij}) : j \in [\ell_i]\} \Big|. \qquad (6.4)$$

Thus, $e := e([r])$. Note that the sets $A(\Gamma_1), \ldots, A(\Gamma_b)$ are disjoint.

The idea of the proof is, hardly surprisingly, that the values $\pi(\alpha)$ for different α are almost independent, and thus indicators $\mathbf{1}\{\pi(\alpha) = \beta\}$ for different α are almost independent. However; this is only "almost", and

therefore we approximate using truly independent variables, constructed in a careful way.

We begin by constructing the random permutation π in a special way. Define recursively $L'_1, \ldots, L'_b$ and $L_1, \ldots, L_b$ by

$$L'_k := 1 + \sum_{j=1}^{k-1} L_j, \qquad L_k := \big|A(\Gamma_k)\big|L'_k, \qquad 1 \leqslant k \leqslant b. \qquad (6.5)$$

Assume $N \geqslant L_b$. (Otherwise (6.3) is trivial if C is large enough.)

Let $\hat{\tau}^1, \ldots \hat{\tau}^b$ be independent uniformly random permutations in $\mathfrak{S}_N$, and let $\tau^k := \hat{\tau}^k(1) \cdots \hat{\tau}^k(L_k)$ be the random string consisting of the first L_k values of $\hat{\tau}^k$. We write the elements of τ^k as $\tau^k(\ell)$, $\ell \in [L_k]$. Using a fixed bijection between $[L_k]$ and $A(\Gamma_k) \times [L'_k]$, we also regard τ^k as an array $\tau^k(\alpha, \ell)$, with $\alpha \in A(\Gamma_k)$ and $\ell \in [L'_k]$.

For $k \in [b]$ and a set of indices $J \subseteq [k-1]$, say that an element $\tau^k(\alpha, \ell)$ is *bad* if it also occurs in τ^j for some $j \in J$. (I.e., if it equals $\tau^j(\alpha', \ell')$ for some $(\alpha', \ell') \in A(\Gamma_j) \times [L'_j]$.) Finally, define $\pi_J^k(\alpha)$ for $\alpha \in A(\Gamma_k)$ as $\tau^k(\alpha, \ell)$ for the smallest $\ell \in [L'_k]$ such that this element is not bad. Note that we have defined L'_k in (6.5) so large that there is always at least one good element for each α. (We will not use $\pi_J^k(\alpha)$ for $\alpha \notin A(\Gamma_k)$; we may define these arbitrarily to make π_J^k a permutation if desired, but we may also just ignore them.)

By construction, for a given k and J, $\pi_J^k(\alpha)$ are distinct for $\alpha \in A(\Gamma_k)$, and these values are distinct from $\pi_I^j(\alpha')$ for all $j \in J$, $I \subseteq [j-1]$ and $\alpha' \in A(\Gamma_j)$.

In particular, consider the case $J_k = [k-1]$ for all $k \in [b]$. Then all $\pi_{[k-1]}^k(\alpha)$, for $k \in [b]$ and $\alpha \in A(\Gamma_k)$, are distinct, and by symmetry, they equal any sequence of $\sum_k |A(\Gamma_k)|$ distinct values in $[N]$ with the same probability. Hence, they have the same (joint) distribution as the values $\pi(\alpha)$ for $\alpha \in \bigcup_k A(\Gamma_k) = A([r])$. (Recall that the sets $A(\Gamma_1), \ldots, A(\Gamma_b)$ are disjoint.) Consequently, we may assume that

$$\pi(\alpha) = \pi_{[k-1]}^k(\alpha), \qquad \alpha \in A(\Gamma_k),\ k \in [b]. \qquad (6.6)$$

For $i \in [r]$, let $k(i)$ be the unique index such that $i \in \Gamma_{k(i)}$. Given $J \subseteq [k(i) - 1]$, define further a modification of Y_i in (6.1) by

$$Y_{i;J} := \prod_{j=1}^{\ell_i} \mathbf{1}\big\{\pi_J^{k(i)}(\alpha_{ij}) = \beta_{ij}\big\}. \qquad (6.7)$$

Note that by (6.1) and (6.6),

$$Y_i = Y_{i;[k(i)-1]}. \qquad (6.8)$$

Define also a random variable denoted $Y_{i;\Delta J}$ by

$$Y_{i;\Delta J} = \sum_{I \subseteq J} (-1)^{|J \setminus I|} Y_{i;I}. \qquad (6.9)$$

This can be regarded as a kind of inclusion–exclusion, or Möbius inversion, and (6.9) implies the inverse relation,

$$Y_{i;I} = \sum_{J \subseteq I} Y_{i;\Delta J}, \qquad I \subseteq [k(i)-1]. \tag{6.10}$$

We now conclude from (6.8), (6.10) and the multilinearity of mixed cumulants ($\kappa 2$) that

$$\kappa(Y_1, \dots, Y_r) = \kappa\big(Y_{1;[k(1)-1]}, \dots, Y_{r;[k(r)-1]}\big)$$

$$= \sum_{J_1 \subseteq [k(1)-1]} \cdots \sum_{J_r \subseteq [k(r)-1]} \kappa(Y_{1;\Delta J_1}, \dots, Y_{r;\Delta J_r}). \tag{6.11}$$

The number of terms in this sum is $O(1)$, so it suffices to estimate each term individually.

Thus, fix a sequence $J_1, \dots, J_r$ with $J_i \subseteq [k(i)-1]$. Let Λ be the graph with vertex set $[b]$ and an edge jk (with $j < k$) if $j \in \bigcup_{i \in \Gamma_k} J_i$.

Note first that a variable $Y_{i;\Delta J_i}$ by (6.9), (6.7) and the construction of π_J^k, depends only on $\tau^{k(i)}$ and τ^j for $j \in J_i$. Consequently, if Λ is disconnected, so we may divide Λ into two parts Λ_1 and Λ_2 with no edges between them, then the random variables $\{Y_{i;\Delta J_i} : i \in \Lambda_\ell\}$ depend only on τ^k for $k \in \Lambda_\ell$, $\ell \in \{1, 2\}$, and thus these two sets of variables are independent. Hence, ($\kappa 3$) yields the following.

Claim 1. *If Λ is disconnected, then $\kappa\big(Y_{1;\Delta J_1}, \dots, Y_{r;\Delta J_r}\big) = 0$.*

Consequently, it suffices to consider the case when Λ is connected. For a set $R \subseteq [r]$, let

$$Y_R^* := \prod_{i \in R} Y_{i;\Delta J_i}. \tag{6.12}$$

By ($\kappa 4$), the mixed cumulant $\kappa(Y_{1;\Delta J_1}, \dots, Y_{r;\Delta J_r})$ in (6.11) can be written as a linear combination of products

$$\prod_{p=1}^{q} \mathbb{E} Y_{R_p}^* \tag{6.13}$$

where $R_1, \dots, R_q$ is a partition of $[r]$. The number of terms and their coefficients are $O(1)$, and thus it suffices to estimate each product (6.13).

Define for $1 \leqslant j < k \leqslant b$, $\ell \in [L_j]$ and $m \in [L_k]$ the indicator

$$\xi_{jk\ell m} := \mathbf{1}\big\{\tau^j(\ell) = \tau^k(m)\big\} \tag{6.14}$$

and let $\Xi := \{\xi_{jk\ell m}\}$ be the array of all these indicators. Thus Ξ tells us exactly which coincidences there are among the values $\tau^k(i)$. (Recall that there are no such coincidences with the same k.) This also determines, for each k and $J \subseteq [k-1]$, the indices of the elements $\tau^k(\alpha, \ell)$ that are bad when constructing π_J, and thus exactly which element $\tau^k(\alpha, \ell)$ that $\pi_J^k(\alpha)$ equals, for each $\alpha \in A(\Gamma_k)$. This implies the following.

Claim 2. *For any given sequence of sets $I_1, \dots, I_r$, Ξ determines exactly what coincidences, if any, there are among all $\pi_{I_i}^{k(i)}(\alpha)$ for $i \in [r]$ and $\alpha \in A(\Gamma_k)$.*

Furthermore, by symmetry, conditioned on Ξ, any sequence of non-coinciding values $\pi_{I_i}^{k(i)}(\alpha)$ has the same distribution as a sequence drawn without replacement from $[N]$.

Let $R \subseteq [r]$, and consider $Z := \prod_{i \in R} Y_{i;I_i}$. This is, recalling (6.7), a product of factors of the type $\mathbf{1}\{\pi_{I_i}^{k(i)}(\alpha_{i\ell}) = \beta_{i\ell}\}$. There may be some pairs $(\alpha_{i\ell}, \beta_{i\ell})$ that are repeated; we let $\tilde{Z}$ be the product after deleting all factors such that $(\alpha_{i\ell}, \beta_{i\ell})$ repeats a previous pair. Note that the number of factors remaining in $\tilde{Z}$ is $e(R)$ by (6.4). Suppose that

$$\pi_{I_i}^{k(i)}(\alpha_{i\ell}) = \pi_{I_j}^{k(j)}(\alpha_{jm}) \tag{6.15}$$

for two of the remaining factors $\mathbf{1}\{\pi_{I_i}^{k(i)}(\alpha_{i\ell}) = \beta_{i\ell}\}$ and $\mathbf{1}\{\pi_{I_j}^{k(j)}(\alpha_{jm}) = \beta_{jm}\}$. If further $k(i) = k(j)$, then $\alpha_{i\ell} = \alpha_{jm}$, by our construction of $\pi_J^{k(i)}$, and then $\beta_{i\ell} \neq \beta_{jm}$ since we have eliminated repetitions. On the other hand, if $k(i) \neq k(j)$, then i and j belong to different blocks, so by the definition of blocks $B_i \cap B_j = \emptyset$ and thus $\beta_{im} \neq \beta_{pq}$. Hence, in any case, (6.15) implies $\beta_{i\ell} \neq \beta_{jm}$, and thus

$$\mathbf{1}\{\pi_{I_i}^{k(i)}(\alpha_{i\ell}) = \beta_{i\ell}\} \cdot \mathbf{1}\{\pi_{I_i}^{k(i)}(\alpha_{jm}) = \beta_{jm}\} = 0. \tag{6.16}$$

This and Claim 2 show that conditioned on Ξ, either $\prod_{i \in R} Y_{i;I_i}$ vanishes because the product contains conflicting indicators, or else all $\pi_{I_i}^{k(i)}(\alpha_{ij})$ occurring in the $e(R)$ factors remaining in $\tilde{Z}$ are distinct and obtained by drawing without replacement from $[N]$. Hence, in any case,

$$\mathbb{E}\Big(\prod_{i \in R} Y_{i;I_i} \,\Big|\, \Xi\Big) \leqslant \mathbb{E}\big(\tilde{Z} \mid \Xi\big) \leqslant \frac{1}{(N)_{e(R)}} \leqslant CN^{-e(R)}. \tag{6.17}$$

This is valid for any $I_i \subseteq [k(i) - 1]$. Hence, by the definitions (6.12) and (6.9)

$$\mathbb{E}\big(|Y_R^*| \mid \Xi\big) = \mathbb{E}\Big(\prod_{i \in R} |Y_{i;\Delta J_i}| \,\Big|\, \Xi\Big) \leqslant \sum_{I_i \subseteq J_i,\, i \in R} \mathbb{E}\Big(\prod_{i \in R} Y_{i;I_i} \,\Big|\, \Xi\Big) \leqslant CN^{-e(R)}. \tag{6.18}$$

Suppose that $j < k$, and suppose that τ^k and τ^j have no common element. Then, for every $J \subseteq [k-1]$, an element $\tau^k(\alpha, \ell)$ is bad for J if and only if it is bad for $J \cup \{j\}$, and thus $\pi_{J\cup\{j\}}^k(\alpha) = \pi_J^k(\alpha)$ for all $\alpha \in A(\Gamma_k)$; hence $Y_{i;J\cup\{j\}} = Y_{i;J}$ for $i \in \Gamma_k$. Consequently, (6.9) shows that then $Y_{i;\Delta J} = 0$ for every J such that $j \in J$. In contrapositive form, this shows the following, recalling (6.14),

Claim 3. *If $Y_{i;\Delta J} \neq 0$, $i \in \Gamma_k$ and $j \in J$, then τ^k and τ^j have at least one common element, i.e., $\xi_{jk\ell m} = 1$ for some $\ell \in [L_j]$ and $m \in [L_k]$.*

If F is a graph with vertex set $\subseteq [b]$, say that Ξ *covers* F if for every edge $jk \in F$ with $j < k$, there exist $\ell \in [L_j]$ and $m \in [L_k]$ such that $\xi_{jk\ell m} = 1$. For a set $R \subseteq [r]$, let Λ_R be the graph with edges

$$E(\Lambda_R) := \bigcup_{k \in [b]} \bigcup_{i \in R \cap \Gamma_k} \{jk : j \in J_i\} \tag{6.19}$$

Then, by (6.12) and Claim 3, if $Y_R^* \neq 0$, then Ξ covers Λ_R. This implies that we can improve (6.18) to

$$\mathbb{E}\big(|Y_R^*| \mid \Xi\big) \leqslant CN^{-e(R)}\mathbf{1}\{\Xi \text{ covers } \Lambda_R\}. \tag{6.20}$$

Hence,

$$\mathbb{E}\big|Y_R^*\big| \leqslant CN^{-e(R)}\,\mathbb{P}\big(\Xi \text{ covers } \Lambda_R\big). \tag{6.21}$$

Recall that we have fixed $J_1, \ldots, J_r$, so Λ_R is for each $R \subseteq [r]$ a fixed non-random graph on $[b]$. If F is any graph on $[b]$, let $\rho(F)$ be the number of edges in a spanning forest. (This equals b minus the number of components of F.) We claim that if F is any graph on $[b]$,

$$\mathbb{P}(\Xi \text{ covers } F) \leqslant CN^{-\rho(F)}. \tag{6.22}$$

To see this, we may replace F by a spanning forest, so it suffices to show (6.22) when F is a forest. We do this by induction on $\rho(F)$; the base case $\rho(F) = 0$ being trivial. If $\rho(F) > 0$, let k be a leaf in F, let $k\ell$ be the edge incident to k, and let $F' := F - k\ell$. Recall that Ξ covers $k\ell$ if τ^k and τ^ℓ have a common element. Thus, conditioning on τ^j for all $j \neq k$, the probability that Ξ covers $k\ell$ is at most $L_k L_\ell N^{-1} \leqslant CN^{-1}$. Hence,

$$\mathbb{P}\big(\Xi \text{ covers } F\big) \leqslant CN^{-1}\,\mathbb{P}\big(\Xi \text{ covers } F'\big), \tag{6.23}$$

and (6.22) follows by induction.

Combining (6.21) and (6.22), we find

$$\mathbb{E}\big|Y_R^*\big| \leqslant CN^{-e(R)-\rho(\Lambda_R)}. \tag{6.24}$$

Finally, consider as in (6.13) a partition $R_1, \ldots, R_q$ of $[r]$. Then (6.2) and (6.4) imply

$$e = e([r]) \leqslant \sum_p e(R_p). \tag{6.25}$$

Furthermore, $\bigcup_p \Lambda_{R_p} = \Lambda$ by (6.19). If we take a spanning subtree $\widehat{\Lambda}_{R_p} \subseteq \Lambda_{R_p}$ for each $p \leqslant q$, then the union of these subtrees has the same components as $\bigcup_p \Lambda_{R_p} = \Lambda$. Hence, if Λ is connected, $\bigcup_p \widehat{\Lambda}_{R_p}$ is connected and has thus at least $b - 1$ edges, which implies

$$\sum_p \rho(\Lambda_{R_p}) \geqslant b - 1. \tag{6.26}$$

Consequently, (6.24) implies by (6.25)–(6.26),

$$\left|\prod_{p=1}^q \mathbb{E}\,Y_{R_p}^*\right| \leqslant \prod_{p=1}^q \mathbb{E}\big|Y_{R_p}^*\big| \leqslant CN^{-\sum_p e(R_p)-\sum_p \rho(\Lambda_{R_p})} \leqslant CN^{-e-(b-1)}. \tag{6.27}$$

As said above, the mixed cumulant $\kappa(Y_{1;\Delta J_1}, \ldots, Y_{r;\Delta J_r})$ in (6.11) is by ($\kappa 4$) and (6.12) a linear combination of such products, and thus (6.27) yields

$$\big|\kappa(Y_{1;\Delta J_1}, \ldots, Y_{r;\Delta J_r})\big| \leqslant CN^{-e-(b-1)}. \tag{6.28}$$

We have here assumed that Λ is connected, but as said in Claim 1, the cumulant vanishes otherwise, so (6.28) holds in general.

Finally, also as said above, the result (6.3) follows from (6.11) and (6.28). $\qquad \square$

7. Proofs of Theorems 3.2, 3.9 and 3.10

We prove first Theorem 3.2 for $G^*(n, \mathbf{d}_n)$, and show then how the proof can be extended to Theorem 3.9 for the simple graph $G(n, \mathbf{d}_n)$. Theorem 3.10 follows by simple modifications.

As said earlier, we use the method of moments in the form with cumulants; we thus show convergence of all (mixed) cumulants. The proofs will use Lemma 6.1 to estimate higher cumulants, as detailed below. In addition, we estimate means and variances (and higher moments for unicyclic components) by standard methods; for the sake of focussing the presentation, we state these results as the following lemmas but postpone their proofs to Section 8. Recall that $Z_H^* := Z_H\big(G^*(n, \mathbf{d}_n)\big)$,

Lemma 7.1. *Assume* (A1)–(A2). *Then, for the random multigraph* $G^*(n, \mathbf{d}_n)$:

(i) *If* H *is a tree, then*

$$\mathbb{E}\,\overline{Z}_H^{\mathsf{u}*} = n\lambda_H + o(n) \tag{7.1}$$

with λ_H *given by* (3.3).

(ii) *If* H_1, H_2 *are trees, then*

$$\mathrm{Cov}\big(\overline{Z}_{H_1}^{\mathsf{u}*}, \overline{Z}_{H_2}^{\mathsf{u}*}\big) = n\sigma_{H_1,H_2} + o(n), \tag{7.2}$$

where σ_{H_1,H_2} *is given by* (3.7).

(iii) *If* H *is a connected unicyclic multigraph, then*

$$\mathbb{E}\,\overline{Z}_H^{\mathsf{u}*} \to \lambda_H, \tag{7.3}$$

with λ_H *as in* (3.3). *Moreover, for any distinct such multigraphs* $H_1, \ldots, H_\ell$, *and integers* $r_1, \ldots, r_\ell \geqslant 0$, *the mixed factorial moments converge:*

$$\mathbb{E}\prod_{i=1}^{k}(\overline{Z}_{H_i}^{\mathsf{u}*})_{r_i} \to \prod_{i=1}^{k}\lambda_{H_i}^{r_i}. \tag{7.4}$$

(iv) *If* H *is a connected multigraph with more than one cycle, i.e.,* $e(H) > v(H)$, *then*

$$\mathbb{E}\,\overline{Z}_H^{\mathsf{u}*} \to 0. \tag{7.5}$$

Lemma 7.2. *Assume* (A1)–(A2). *For every tree* H, (3.5) *holds, i.e.,* $\lambda_H = p_H/|H|$.

Lemma 7.3. *Assume* (A1)–(A2) *and* (A6). *If* H *is a tree with* $v(H) > 1$, *then* $\sigma_{H,H} > 0 \iff \lambda_H > 0$. *More generally, if* $H_1, \ldots, H_\ell$ *are distinct trees with* $v(H_i) > 1$ *and* $\lambda_{H_i} > 0$ *for each* i, *then the matrix* $\big(\sigma_{H_i,H_j}\big)$ *is non-singular.*

The next lemma shows that, assuming the second moment condition (A3), Lemma 7.1(iii) extends to include also the numbers of loops $Z_{\mathsf{C}_1}^{\mathsf{u}*}$ and pair of multiple edges $Z_{\mathsf{C}_2}^{\mathsf{u}*}$, cf. Example 3.1. (We have to restrict to simple graphs H_i; if H is not simple, there is obviously a strong dependence between $\overline{Z}_H^{\mathsf{u}*}$ and $Z_{\mathsf{C}_1}^{\mathsf{u}*}$ or $Z_{\mathsf{C}_2}^{\mathsf{u}*}$.)

Lemma 7.4. *Assume* (A1)–(A3). *Let* $H_1, \ldots, H_k$ *be connected unicyclic simple graphs. Then, for any integers* $s_1, s_2, r_1, \ldots, r_k \geqslant 0$,

$$\mathbb{E}\Big(\big(Z^{\mathsf{u}*}_{\mathsf{C}_1} \big)_{s_1} \big(Z^{\mathsf{u}*}_{\mathsf{C}_2} \big)_{s_2} \prod_{i=1}^{k} (\overline{Z}^{\mathsf{u}*}_{H_i})_{r_i} \Big) \to \lambda^{s_1}_{*\mathsf{C}_1} \lambda^{s_2}_{*\mathsf{C}_2} \prod_{i=1}^{k} \lambda^{r_i}_{H_i}, \qquad (7.6)$$

where

$$\lambda_{*\mathsf{C}_1} := \frac{\mathbb{E}\, D(D-1)}{2}, \qquad\qquad \lambda_{*\mathsf{C}_2} := \Big(\frac{\mathbb{E}\, D(D-1)}{2} \Big)^2. \qquad (7.7)$$

As said above, we postpone the proofs of these lemmas to Section 8.

Fix a (multi)graph H. Let $h := v(H) \geqslant 1$, and assume that the vertices of H are labelled $1, \ldots, h$. Our aim is to estimate the cumulants $\kappa_r(Z^*_H)$ for (fixed) $r \geqslant 1$ and corresponding mixed cumulants. In this section, C denotes constants that may depend on the order r and the fixed (small) graph H (and corresponding graphs below), but not on n; the same holds for the implicit constants in $O(\ldots)$.

We use the version of the configuration model described in Section 4. A copy of H in $G^*(n, \mathbf{d}_n)$ then corresponds to a copy of $\widehat{H}$ in $\widehat{G}(n, \mathbf{d}_n)$, where $\widehat{H}$ is obtained from H by subdividing each edge into two, and regarding the new vertices as cuffs, and where we only count copies of $\widehat{H}$ that map real vertices to real vertices and cuffs to cuffs. We consider also the identity of the half-edges used in the construction, and see that an isolated labelled copy of $\widehat{H}$ in $\widehat{G}(n, \mathbf{d}_n)$ is described by the following data:

(ϕ1) For each vertex $i \in V(H) = [h]$: a real vertex $v_{\nu(i)}$ such that $d_{\nu(i)} = d_H(i)$. Furthermore, $v(1), \ldots, v(h)$ are distinct.

(ϕ2) For each edge ij in H: two half-edges $\alpha_{ij} \in \Upsilon_{\nu(i)}$ and $\alpha'_{ij} \in \Upsilon_{\nu(j)}$, and also a cuff χ_{ij} and a labelling of the two half-edges at χ_{ij} as β_{ij} and β'_{ij}. Furthermore, all these half-edges for $ij \in E(H)$ are distinct.

Each such family of data $\phi := \big(\nu(i), \alpha_{ij}, \alpha'_{ij}, \chi_{ij}, \beta_{ij}, \beta'_{ij} \big)$ defines a possible isolated labelled copy $\widehat{H}_\phi$ of $\widehat{H}$. Thus, if I_ϕ is the indicator that $\widehat{H}_\phi$ exists in $\widehat{G}(n, \mathbf{d}_n)$, and $\Phi(\widehat{H})$ is the set of all such data ϕ, then

$$\overline{Z}^*_H = \overline{Z}_H(G^*(n, \mathbf{d}_n)) = \sum_{\phi \in \Phi(\widehat{H})} I_\phi. \qquad (7.8)$$

Furthermore, $\widehat{H}_\phi$ exists in $G^*(n, \mathbf{d}_n)$ if and only if the construction of $\widehat{G}(n, \mathbf{d}_n)$ yields edges $\alpha_{ij}\beta_{ij}$ and $\alpha'_{ij}\beta'_{ij}$ for each $ij \in E(H)$, which is equivalent to $\pi(\alpha_{ij}) = \beta_{ij}$ and $\pi(\alpha'_{ij}) = \beta'_{ij}$. Consequently, each I_ϕ is the product of $2e(H)$ indicators of the type $\mathbf{1}\{\pi(\alpha) = \beta\}$. Each I_ϕ is thus a random variable of the type in (6.1).

Consider now a sequence $H_1, \ldots, H_r$ of multigraphs. By (7.8) and multilinearity (κ2), the mixed cumulant $\kappa\big(\overline{Z}^*_{H_1}, \ldots, \overline{Z}^*_{H_r} \big)$ can be expanded as

$$\kappa\big(\overline{Z}^*_{H_1}, \ldots, \overline{Z}^*_{H_r} \big) = \sum_{\phi_1 \in \Phi(\widehat{H}_1)} \cdots \sum_{\phi_r \in \Phi(\widehat{H}_r)} \kappa\big(I_{\phi_1}, \ldots, I_{\phi_r} \big), \qquad (7.9)$$

where the mixed cumulants may be estimated by Lemma 6.1.

It remains to estimate the parameters b and e in Lemma 6.1 and the corresponding number of terms in (7.9). This is done in the following lemmas,

using a standard type of argument that is common in applications of the method of moments in combinatorial problems.

Lemma 7.5. *Assume* (A1)–(A2). *Let* $H_1, \ldots, H_r$ *be a sequence of connected multigraphs. Then*

$$\left|\kappa\left(\overline{Z}^*_{H_1}, \ldots, \overline{Z}^*_{H_r}\right)\right| \leqslant C \sum_{\mathcal{F}} n^{v(\mathcal{F})-e(\mathcal{F})-q(\mathcal{F})+1}, \qquad (7.10)$$

where we sum over all unlabelled bipartite multigraphs $\mathcal{F}$ *that can be written as a union* $\bigcup_{i=1}^r \widehat{H}'_i$ *where* $\widehat{H}'_i \cong \widehat{H}_i$ *and, as above,* $\widehat{H}_i$ *is obtained from* H_i *by subdividing each edge into two.*

We consider only unions $\bigcup_{i=1}^r \widehat{H}'_i$ respecting the bipartition between real vertices and cuffs. Note that the set of $\mathcal{F}$ in (7.10) is finite and independent of n.

Proof. For notational simplicity we consider the case of a single multigraph H, i.e., $H_1 = \cdots = H_r = H$; the proof for mixed cumulants is the same.

Consider one term in the sum in (7.9); it is given by indicators $I_{\phi_1}, \ldots, I_{\phi_r}$ corresponding to r copies $\widehat{H}_{\phi_1}, \ldots, \widehat{H}_{\phi_r}$ of $\widehat{H}$. Let $\hat{F} := \bigcup_{i=1}^r \widehat{H}_{\phi_i}$. We apply Lemma 6.1 with $Y_i := I_{\phi_i}$. Note that e in (6.2) equals $e(\hat{F})$. If the graph Γ in Section 6 has an edge jk, then $\widehat{H}_{\phi_k}$ and $\widehat{H}_{\phi_j}$ have a common half-edge, and thus a common vertex (real or cuff); hence $\widehat{H}_{\phi_k}$ and $\widehat{H}_{\phi_j}$ are subgraphs of the same component of $\hat{F}$. This yields a surjective map from the blocks of Γ to the components of $\hat{F}$, and thus the number of blocks b in Lemma 6.1 satisfies $b \geqslant q(\hat{F})$. Hence, Lemma 6.1 yields

$$\left|\kappa_r\left(I_{\phi_1}, \ldots, I_{\phi_r}\right)\right| \leqslant C N^{1-q(\hat{F})-e(\hat{F})} \leqslant C n^{1-q(\hat{F})-e(\hat{F})}. \qquad (7.11)$$

Now let us count the number of $\phi_1, \ldots, \phi_r \in \Phi(\widehat{H})$ that yield a union $\hat{F}$ isomorphic to some given bipartite multigraph $\mathcal{F}$. Each cuff in $\hat{F}$ and its 2 half-edges may be chosen in $O(N) = O(n)$ ways. Let w be a real vertex in $\mathcal{F}$, and let $m := d_{\mathcal{F}}(w)$ be its degree. Also, let $K := \max_{i \in H} d_H(i) < \infty$. We consider by definition only copies of $\widehat{H}$ such that each vertex v has degree equal to the corresponding vertex in H, and thus $d_{\widehat{G}}(v) \leqslant K$. Hence, for each such choice of v corresponding to $w \in \mathcal{F}$, we have at most K^m choices of the m half-edges incident to it; hence the number of choices of v and its m half-edges is $O(nK^m) = O(n)$.

Consequently, each vertex (real or cuff) in $\mathcal{F}$ gives $O(n)$ choices of corresponding vertex and half-edges in $\hat{F}$. Hence, the number of $\hat{F}$ isomorphic to a given $\mathcal{F}$ is $O\left(n^{v(\mathcal{F})}\right)$. Finally, given $\hat{F}$, we can choose $\widehat{H}_1, \ldots, \widehat{H}_r$ in $O(1)$ ways. Hence, the total number of terms in (7.9) corresponding to a given $\mathcal{F}$ is $O\left(n^{v(\mathcal{F})}\right)$, and by (7.11), their total contribution is $O\left(n^{v(\mathcal{F})-e(\mathcal{F})-q(\mathcal{F})+1}\right)$. This yields (7.10). □

Lemma 7.6. *Let* $F := \bigcup_{i=1}^r H_i$, *where* $H_1, \ldots, H_r$ *is a sequence of connected multigraphs.*

(i) *Then* $v(F) \leqslant e(F) + q(F)$.

(ii) *If furthermore at least one* H_i *is not a tree, then* $v(F) \leqslant e(F)+q(F)-1$.

Proof. Consider first the case when F is connected. We may then reorder $H_1, \ldots, H_r$ such that each $\bigcup_{i=1}^{k} H_i$ is connected; furthermore, we may do this starting with any of the multigraphs as H_1, and in (ii) we may thus assume that H_1 is not a tree.

Now, choose for each $i \geqslant 2$ a spanning tree $T_k \subseteq H_k$, and define $F_k := H_1 \cup \bigcup_{i=2}^{k} T_i$, $k \geqslant 1$. Thus each F_k is connected.

Let $k > 1$ and let $T'_k := T_k \cap F_{k-1}$. Since T'_k is a subgraph of T_k, which is a tree, we see that T'_k is a forest, and since $F_k = T_k \cup F_{k-1}$ is connected, T'_k is not empty. Hence, $v(T'_k) \geqslant e(T'_k) + 1$. Consequently, since $F_k = T_k \cup F_{k-1}$,

$$
\begin{aligned}
v(F_k) - e(F_k) &= v(F_{k-1}) + v(T_k) - v(T'_k) - \big(e(F_{k-1}) + e(T_k) - e(T'_k)\big) \\
&\leqslant v(F_{k-1}) - e(F_{k-1}) + v(T_k) - e(T_k) - 1 \\
&= v(F_{k-1}) - e(F_{k-1})
\end{aligned}
\tag{7.12}
$$

when $k > 1$. Hence, by induction $v(F_k) - e(F_k) \leqslant v(H_1) - e(H_1)$. Furthermore, F_r is a spanning subgraph of F, and thus $v(F_r) = v(F)$ and $e(F_r) \leqslant e(F)$. Consequently,

$$
v(F) - e(F) \leqslant v(F_r) - e(F_r) \leqslant v(H_1) - e(H_1).
\tag{7.13}
$$

Moreover, $e(H_1) \geqslant v(H_1) - 1$ since H_1 is connected, and if H_1 is not a tree, then $e(H_1) \geqslant v(H_1)$. Consequently, (7.13) yields (i) and (ii) in the case $q(F) = 1$.

If $q(F) > 1$, i.e., F is disconnected, denote the components of F by F_i, $i = 1, \ldots, q(F)$. Then, by what just has been shown, $v(F_i) \leqslant e(F_i) + 1$, and if some $H_j \subseteq F_i$ is not a tree, then $v(F_i) \leqslant e(F_i)$. Hence, the result fullows by summing over all components F_i. $\qquad \square$

Lemma 7.7. *Assume* (A1)–(A2). *Let* $H_1, \ldots, H_r$ *be a sequence of connected multigraphs. Then*

$$
\kappa\big(\overline{Z}^*_{H_1}, \ldots, \overline{Z}^*_{H_r}\big) = O(n).
\tag{7.14}
$$

Furthermore, if at least one H_i *is not a tree, then*

$$
\kappa\big(\overline{Z}^*_{H_1}, \ldots, \overline{Z}^*_{H_r}\big) = O(1).
\tag{7.15}
$$

Proof. An immediate consequence of Lemmas 7.5 and 7.6, applying the latter to $\widehat{H}_1, \ldots, \widehat{H}_r$. $\qquad \square$

Proof of Theorem 3.2. Define

$$
X_H := \begin{cases} \big(\overline{Z}^{\mathsf{u}*}_H - \mathbb{E}\,\overline{Z}^{\mathsf{u}*}_H\big)/n^{1/2}, & H \text{ is a tree}, \\ X_H := \overline{Z}^{\mathsf{u}*}_H, & H \text{ has a cycle}. \end{cases}
\tag{7.16}
$$

First, (iii), the case when $e(H) > v(H)$, is easy. By Lemma 7.1(iv) and Markov's inequality, $\mathbb{P}(X_H \neq 0) \leqslant \mathbb{E}\,X_H \to 0$. (Recall that X_H is a non-negative integer.) In particular, $X_H \xrightarrow{\text{p}} 0$. Furthermore, (7.15) shows that every cumulant $\kappa_r(X_H) = O(1)$, and thus every moment is bounded by (κ5). This implies uniform integrability of every power, and thus $\mathbb{E}\,X_H^r \to 0$ for every r. Convergence (to 0) of joint moments with other X_H follows by the Cauchy–Schwarz inequality when we have shown convergence of moments also in (i) and (ii).

For (i) and (ii), and joint convergence of both, we use the method of moments, in the cumulant version. Let $\widetilde{X}_H$ be random variables defined for (unlabelled) trees and unicyclic multigraphs H to have the claimed joint limit distribution:

(i) For trees T, $\widetilde{X}_T$ have a joint normal distribution with $\mathbb{E}\,X_T = 0$ and $\mathrm{Cov}\big(\widetilde{X}_{T_1}, \widetilde{X}_{T_2}\big) = \sigma_{T_1, T_2}$.

(ii) For unicyclic F, $\widetilde{X}_F \sim \mathrm{Po}(\lambda_F)$ with $\widetilde{X}_F$ independent of all other $\widetilde{X}_H$.

We then claim that for any $r \geqslant 1$ and trees or connected unicyclic multigraphs $H_1, \ldots, H_r$,

$$\kappa\big(X_{H_1}, \ldots, X_{H_r}\big) \to \kappa\big(\widetilde{X}_{H_1}, \ldots, \widetilde{X}_{H_r}\big). \tag{7.17}$$

Indeed, this implies by $(\kappa 5)$ convergence of all mixed moments. Furthermore, normal and Poisson distributions have finite moment generating functions, and thus the joint distribution of any finite number of $\widetilde{X}_H$ is determined by its mixed moments. Hence, (7.17) implies convergence

$$\big(X_{H_1}, \ldots, X_{H_r}\big) \xrightarrow{\mathrm{d}} \big(\widetilde{X}_{H_1}, \ldots, \widetilde{X}_{H_r}\big) \tag{7.18}$$

as claimed in the theorem. Furthermore, all moments converge, as just seen.

Note that the right-hand side in (7.17) vanishes except in the two cases $r \geqslant 1$ and all H_i are the same unicyclic multigraph, or $r = 2$ and both H_1 and H_2 are trees. This follows by the independence assumption in (ii) above together with $(\kappa 3)$, and the fact that all mixed cumulants of order $r \geqslant 3$ vanish for joint normal distributions.

In order to show (7.17), we consider three cases.

Case 1: Every H_i is a tree. Consider three subcases. First, if $r = 1$, then

$$\kappa\big(X_{H_1}\big) = \mathbb{E}\,X_{H_1} = 0 = \kappa\big(\widetilde{X}_{H_1}\big). \tag{7.19}$$

Secondly, if $r = 2$, then (7.16) and (7.2) yield

$$\kappa\big(X_{H_1}, X_{H_2}\big) = \mathrm{Cov}\big(X_{H_1}, X_{H_2}\big) = n^{-1}\,\mathrm{Cov}\big(\overline{Z}^{\mathsf{u}*}_{H_1}, \overline{Z}^{\mathsf{u}*}_{H_2}\big)$$
$$\to \sigma_{H_1, H_2} = \kappa\big(\widetilde{X}_{H_1}, \widetilde{X}_{H_2}\big). \tag{7.20}$$

Thirdly, if $r \geqslant 3$, then (7.14) in Lemma 7.7 together with (3.1) and (7.16) yields

$$\kappa\big(X_{H_1}, \ldots, X_{H_r}\big) = C n^{-r/2} \kappa\big(\overline{Z}^*_{H_1}, \ldots, \overline{Z}^*_{H_r}\big) = O\big(n^{1-r/2}\big) \to 0. \tag{7.21}$$

This verifies (7.17) in each of the three subcases.

Case 2: Every H_i is unicyclic. This case is really nothing new. By Lemma 7.1(iii), we have convergence of all mixed factorial moments to the corresponding moments of $\widetilde{X}_{H_1}, \ldots, \widetilde{X}_{H_r}$, which by well-known algebraic identities is equivalent to convergence of all mixed moments, and thus to convergence of all mixed cumulants. (See [27, Section 6.1] and $(\kappa 4)$–$(\kappa 5)$.)

Case 3: At least one H_i is a tree and at least one is unicyclic. Suppose that there are $\ell \geqslant 1$ trees (not necessarily distinct) among $H_1, \ldots, H_r$. Then, by (7.15) in Lemma 7.7 together with (3.1) and (7.16),

$$\kappa\big(X_{H_1}, \ldots, X_{H_r}\big) = C n^{-\ell/2} \kappa\big(\overline{Z}^*_{H_1}, \ldots, \overline{Z}^*_{H_r}\big) = O\big(n^{-\ell/2}\big) \to 0. \tag{7.22}$$

Hence, (7.17) holds in this case too.

This completes the verification of (7.17) in all cases; as said above, this proves (7.18), with convergence of all moments.

Finally, Lemma 7.1 shows (3.3), Lemma 7.2 shows (3.5), Lemma 7.3 shows the first equivalence in (3.4), and the second is obvious from (3.3). Lemma 7.3 yields also the final claim on non-singularity of the covariance matrix $\Sigma = \left(\sigma_{H_i, H_j} \right)$. $\qquad\square$

To show Theorem 3.9 for the simple random graph $G(n, \mathbf{d}_n)$, we combine Theorem 3.2 with asymptotics of the counts $Z^{\mathsf{u}*}_{\mathsf{C}_1}$ and $Z^{\mathsf{u}*}_{\mathsf{C}_2}$ of loops and double edges. Note that $Z^{\mathsf{u}*}_{\mathsf{C}_j}$ counts all occurences of C_j, which in general differs from $\overline{Z}^{\mathsf{u}*}_{\mathsf{C}_j}$, the number of isolated occurences. The latter is already included in Theorem 3.2, but $Z^{\mathsf{u}*}_{\mathsf{C}_j}$ requires an extra argument.

$Z^{*}_{\mathsf{C}_1}$ and $Z^{*}_{\mathsf{C}_2}$ can be expressed as sums (7.8) for sets Φ of data ϕ as above, with the difference that in $(\phi 1)$ above, we omit the condition on the degree $d_{\nu(i)}$.

We may just as well consider a somewhat more general situation: we say that a *marked multigraph* $\widetilde{H} = \left(H, (\iota_w)_{w \in H} \right)$ is a multigraph H where each vertex v has a mark $\iota_v \in \{bound,\ free\}$. We then define $Z_{\widetilde{H}}(G)$ as the number of labelled copies of H in G such that each bound vertex $w \in H$ corresponds to a vertex v in G with the same degree $d_G(v) = d_H(w)$ (while there is no restriction for a free vertex). Hence, if all vertices in H are free, $Z_{\widetilde{H}} = Z_H$, and if all vertices are bound, $Z_{\widetilde{H}} = \overline{Z}_H$. We write, as in the unmarked case, $Z_{\widetilde{H}} := Z_{\widetilde{H}}(G(n, \mathbf{d}_n))$ and $Z^{*}_{\widetilde{H}} := Z_{\widetilde{H}}(G^*(n, \mathbf{d}_n))$.

Assume that $V(H) = [h]$ for convenience. Then $Z^{*}_{\widetilde{H}}$ is given by (7.8) for a set $\Phi(\widetilde{H})$ of ϕ defined by $(\phi 1)$–$(\phi 2)$, but with the degree condition in $(\phi 1)$ omitted for free vertices. Let cuffs in subdivided multigraphs be bound by default. Furthermore, define $\bigcup_i \widetilde{H}_i$, where $\widetilde{H}_i$ are marked multigraphs, by taking the union and marking a vertex as bound if it is bound in some $\widetilde{H}_i$, and free otherwise.

We extend Lemma 7.5 as follows.

Lemma 7.8. *Assume* (A1)–(A2). *Let* $\widetilde{H}_1, \dots, \widetilde{H}_r$ *be a sequence of connected marked multigraphs. Then*

$$\left| \kappa \left(Z^{*}_{\widetilde{H}_1}, \dots, Z^{*}_{\widetilde{H}_r} \right) \right| \leqslant C \sum_{\mathcal{F}} n^{v(\mathcal{F}) - e(\mathcal{F}) - q(\mathcal{F}) + 1} \prod_{w \in \mathcal{F} \text{ is free}} \mathbb{E}\, D_n^{d_{\mathcal{F}}(w)} \quad (7.23)$$

where we sum over all unlabelled bipartite multigraphs $\mathcal{F}$ that can be written as a union $\bigcup_{i=1}^{r} \widehat{H}'_i$ where $\widehat{H}'_i \cong \widehat{H}_i$ and $\widehat{H}_i$ is obtained from $\widetilde{H}_i$ by subdividing each edge into two as above.

Proof. We follow the proof of Lemma 7.5. The only difference is when counting the number of ways we can choose a copy v of a vertex $w \in \mathcal{F}$ and its $m := d_{\mathcal{F}}(w)$ half-edges. If w is bound, then as in the proof of Lemma 7.5, $d_{\widehat{G}}(v) \leqslant K$ for some $K < \infty$, and then the m half-edges can be chosen in at most $K^m = O(1)$ ways for each v, giving as before $O(n)$ choices of vertex and half-edges.

On the other hand, if w is free, then we bound for each real vertex $v = v_i \in V(\widehat{G})$ the number of choices of the m half-edges by $d_{\widehat{G}}(v_i)^m = d_i^m$;

hence, the total number of choices of v and its half-edges is at most

$$\sum_{i=1}^{n} d_i^m = n \, \mathbb{E} \, D_n^m = n \, \mathbb{E} \, D_n^{d_{\mathcal{F}}(w)}. \tag{7.24}$$

This gives an additional factor $\mathbb{E} \, D_n^{d_{\mathcal{F}}(w)}$ for each free w, and thus (7.23) instead of (7.10). $\qquad\square$

Lemma 7.9. *Let* $F = F' \cup F''$ *with* $F' = \bigcup_{i=1}^{r} H_i$, *where* $r \geqslant 0$ *and* $H_1, \ldots, H_r$ *is a sequence of connected multigraphs, and* $F'' = \bigcup_{i=1}^{t} C_i$, *where* $t \geqslant 1$ *and each* C_i *is a cycle* C_ℓ, $\ell \geqslant 1$. *Say that a vertex* $w \in V(F'') \setminus V(F')$ *is free, and let*

$$s(F) := \frac{1}{2} \sum_{w \text{ free}} \left(d_F(w) - 2 \right)_+. \tag{7.25}$$

Then $v(F) + s(F) \leqslant e(F) + q(F) - 1$.

Proof. Let $S(F) := v(F) + s(F) - e(F) - q(F)$; thus the result is $S(F) \leqslant -1$.

We use induction on t. If $t = 1$, then each free vertex w has degree $d_F(w) = d_{C_1}(w) = 2$, and thus $s(F) = 0$. Hence, the result follows from Lemma 7.6(ii) applied to $\{H_i\} \cup \{C_1\}$.

Now suppose that the results hold for some t, and add another cycle C_{t+1}, of length ℓ, say. Denote the old F by F_t, so $F = F_t \cup C_{t+1}$. We consider the changes in the quantities $v(F), e(F), q(F), s(F), S(F)$, which we denote by $\Delta v, \Delta e$, and so on. We treat three cases separately.

(i). If the new cycle C_{t+1} is disjoint from F_t, then $\Delta v = \ell$, $\Delta e = \ell$, $\Delta q = 1$ and $\Delta s = 0$; the latter since all new vertices have degree 2 and thus do not contribute to s. Hence, $\Delta S = -1$.

(ii). Suppose that C_{t+1} is edge-disjoint but not vertex-disjoint from F_t. Then $\Delta e = \ell$ and $\Delta q = 0$. Each vertex in $C_{t+1} \setminus F_t$ contributes 1 to Δv and 0 to Δs, while each vertex in $C_{t+1} \cap F_t$ contributes 0 to Δv and at most 1 to Δs, since the new cycle increases the degree by 2. (If the vertex is not free in F_t, the contribution to Δs is 0.) Hence, each vertex contributes at most 1 to $\Delta(v + s)$, and thus $\Delta(v + s) \leqslant \ell = \Delta e$. Consequently, $\Delta S \leqslant 0$.

(iii). Suppose that C_{t+1} is has some edge in common with F_t. Then $\Delta q = 0$. The edges in $E(C_{t+1}) \setminus E(F_t)$ form $k \geqslant 0$ disjoint paths P_j. Suppose that P_j contains $\ell_j \geqslant 1$ edges. Then P_j contributes ℓ_j to Δe, the $\ell_j - 1$ internal vertices in P_j contribute 1 each to Δv and the two endpoints of P_j contribute at most $1/2$ each to Δs. There are no other contributions, and thus $\Delta(v + s) \leqslant \sum_j \ell_j = \Delta e$ and consequently $\Delta S \leqslant 0$.

We have shown that $\Delta S \leqslant 0$ in all cases, which completes the induction. $\qquad\square$

Lemma 7.10. *Assume* (A1)–(A3). *Let* $H_1, \ldots, H_r$ *be trees or connected unicyclic simple graphs. Then, the joint limits in distribution in Theorem 3.2(i)(ii) hold jointly with*

$$Z_{\mathsf{C}_1}^{\mathsf{u}*} \xrightarrow{\mathrm{d}} \mathrm{Po}\big(\lambda_{*\mathsf{C}_1}\big), \tag{7.26}$$

$$Z_{\mathsf{C}_2}^{\mathsf{u}*} \xrightarrow{\mathrm{d}} \mathrm{Po}\big(\lambda_{*\mathsf{C}_2}\big), \tag{7.27}$$

with $\lambda_{*\mathsf{C}_1}$ and $\lambda_{*\mathsf{C}_2}$ given by (7.7), and with the limits in (7.26) and (7.27) independent of each other and of the limits for $\overline{Z}_{H_i}^{\mathsf{u}*}$ in Theorem 3.2(i)(ii).

Proof. We extend the proof of Theorem 3.2, and show that the convergence (7.17) of joint cumulants holds also if we consider besides the variables X_{H_i} also $X_{*\mathsf{C}_1} := Z_{\mathsf{C}_1}^{\mathsf{u}*}$ and $X_{*\mathsf{C}_2} := Z_{\mathsf{C}_2}^{\mathsf{u}*}$ (possibly repeated several times), and the corresponding $\widetilde{X}_{*\mathsf{C}_1} \sim \mathrm{Po}(\lambda_{*\mathsf{C}_1})$ and $\widetilde{X}_{*\mathsf{C}_2} \sim \mathrm{Po}(\lambda_{*\mathsf{C}_2})$, independent of each other and all $\widetilde{X}_H$. We regard $*\mathsf{C}_j$, $j = 1, 2$, as symbols used only to denote these variables; they are not any graphs; for convenience we may say $H = *\mathsf{C}_j$, but it should be interpreted in this formal sense.

We have in the proof of Theorem 3.2 proved (7.17) when there is no $*\mathsf{C}_j$ among $H_1, \ldots, H_r$, by considering three different cases separately. We now consider two further cases.

Case 4: Some H_i is a $*\mathsf{C}_j$, and some H_j is a tree. Let $\ell \geqslant 1$ be the number of H_i that are trees. Regard each H_i as a marked multigraph $\widetilde{H}_i$; if $H_i = *\mathsf{C}_j$ we let every vertex be free, and otherwise we let every vertex be bound. By (3.1), applied also to $Z_{\mathsf{C}_j}^{\mathsf{u}*}$, we find similarly to (7.22)

$$\kappa\big(X_{H_1}, \ldots, X_{H_r}\big) = Cn^{-\ell/2}\kappa\big(Z_{\widetilde{H}_1}^*, \ldots, Z_{\widetilde{H}_r}^*\big), \qquad (7.28)$$

which we estimate by (7.23). If $d \geqslant 2$, then (A3) implies, through (2.6) and (2.7),

$$\mathbb{E}\, D_n^d \leqslant d_{\max}^{d-2}\, \mathbb{E}\, D_n^2 = O\big(n^{(d-2)/2}\big). \qquad (7.29)$$

Furthermore, if $\mathcal{F} = \bigcup \widehat{H}_i'$ is as in Lemma 7.8, then a vertex $w \in \mathcal{F}$ is free if and only if it belongs to $\widehat{H}_i'$ only for $H_i = *\mathsf{C}_j$. In particular, if $w \in \mathcal{F}$ is free then $d_{\mathcal{F}}(w) \geqslant 2$. Let $s(\mathcal{F})$ be as in (7.25) with $F = \mathcal{F}$, F'' the union of $\widehat{H}_i'$ for $H_i = *\mathsf{C}_j$ and F' the union of the other $\widehat{H}_i$. Then Lemma 7.8, (7.29), (7.25) and Lemma 7.9 yield

$$\big|\kappa\big(Z_{\widetilde{H}_1}^*, \ldots, Z_{\widetilde{H}_r}^*\big)\big| \leqslant C \sum_{\mathcal{F}} n^{v(\mathcal{F})-e(\mathcal{F})-q(\mathcal{F})+1+s(\mathcal{F})} \leqslant C. \qquad (7.30)$$

Consequently, (7.28) yields

$$\kappa\big(X_{H_1}, \ldots, X_{H_r}\big) = O\big(n^{-\ell/2}\big) \to 0. \qquad (7.31)$$

Hence, (7.17) holds in this case too.

Case 5: Some H_i is a $*\mathsf{C}_j$, but no H_j is a tree. This is similar to Case 2. Lemma 7.4 shows convergence of all mixed factorial moments, which is equivalent to convergence of all mixed moments and of all mixed cumulants.

This shows that (7.17) holds in all cases, which implies joint convergence in distribution (7.18) as above. $\qquad\square$

Proof of Theorem 3.9. Theorem 3.2(iii) (multicyclic H) transfers immediately by (2.8). Consider thus only trees and unicyclic H.

The joint convergence in distribution in (i) and (ii) for $G(n, \mathbf{d}_n)$ is an immediate consequence of Lemma 7.10 and conditioning on $Z_{\mathsf{C}_1}^{\mathsf{u}*} = Z_{\mathsf{C}_2}^{\mathsf{u}*} = 0$. (We keep the normalization by $\mathbb{E}\,\overline{Z}_H^{\mathsf{u}*}$ in (3.2).)

Furthermore, with X_H as in (7.16), we have by Theorem 3.2 convergence of every moment $\mathbb{E}\,X_H^r$, and thus $\mathbb{E}\,X_H^r = O(1)$. Furthermore, (2.8) holds

(as said in Remark 2.3, and a consequence of (7.26)–(7.27)); hence for every even integer r,

$$\mathbb{E}\big(X_H^r \mid Z_{\mathsf{C}_1}^{\mathsf{u}*} = Z_{\mathsf{C}_2}^{\mathsf{u}*} = 0\big) \leqslant \frac{\mathbb{E}\, X_H^r}{\mathbb{P}\big(Z_{\mathsf{C}_1}^{\mathsf{u}*} = Z_{\mathsf{C}_2}^{\mathsf{u}*} = 0\big)} = O(1). \qquad (7.32)$$

In other words, the conditioned random variables $\big(X_H \mid G^*(n, \mathbf{d}_n)$ is simple$\big)$ have bounded moments of arbitrary order; hence the convergence in distribution of these variables just shown implies that moment convergence holds also for the conditioned variables, i.e., for $G(n, \mathbf{d}_n)$. Convergence of mixed moments follows by the same argument.

In particular, this shows that for a tree H,

$$\mathbb{E}\big(X_H \mid G^*(n, \mathbf{d}_n)\text{ is simple}\big) = \frac{\mathbb{E}\, \overline{Z}_H^{\mathsf{u}} - \mathbb{E}\, \overline{Z}_H^{\mathsf{u}*}}{n^{1/2}} \to 0, \qquad (7.33)$$

which shows (3.10), and completes the proof. $\qquad\qquad\qquad\square$

Proof of Theorem 3.10. (i): As the proof of Theorem 3.2 (only Case 1 is relevant), but using Lemma 7.8 (with all vertices free) instead of Lemma 7.7; the extra factors in (7.23) are all $O(1)$ by the assumption (A4), so this does not affect the rest of the proof.

(ii): As the proof of Theorem 3.9, using again Lemma 7.8 and (A4) instead of Lemma 7.7. $\qquad\qquad\qquad\square$

Remark 7.11. Theorem 3.10 extends, with the same proof, to marked trees $\tilde{T}$, where as above each vertex is marked as either bound or free. This includes Theorems 3.2, 3.9 and 3.10, but also mixed cases. More generally, we may consider marked multigraphs where each vertex w is marked with a set $\mathcal{D}_w \subseteq \mathbb{N}$ of allowed degrees; we count only copies such that each w corresponds to a vertex $v \in G$ with $d_G(v) \in \mathcal{D}_w$. For example, we may count edges such that one endpoint has prime degree and the other is a leaf. The proofs above hold for this case too; we say that a vertex w is bound if $\mathcal{D}_w$ is finite, and free otherwise. If some vertex is free we assume (A4) as in Theorem 3.10; if all vertices are bound, the assumptions (A1)–(A2) or (A1)–(A3) in Theorems 3.2 and 3.9 are enough.

8. Means and variances

In this section we prove Lemmas 7.1–7.4 used in the proofs in Section 7. We will here use the standard version of the configuration model, recalled at the beginning of Section 4; we thus use the half-edges $\Upsilon = \bigcup_i \Upsilon_i = \{\alpha_j\}_1^N$ (see Section 4) and take a random perfect matching of them.

Proof of Lemma 7.1. Let H be a multigraph and let $h_k := n_k(H)$ be the number of vertices of degree k in H, $k \geqslant 0$. We argue as in Section 7; we denote the possible isolated labelled copies of H in $G^*(n, \mathbf{d}_n)$ by $\{H_\phi\}_{\phi \in \Phi(H)}$, and obtain in analogy with (7.8)

$$\overline{Z}_H^* = \sum_{\phi \in \Phi(H)} I_\phi, \qquad (8.1)$$

where now I_ϕ is the indicator that H_ϕ exists as an isolated subgraph in $G^*(n, \mathbf{d}_n)$. H_ϕ is specified by:

(ϕ'1) For each vertex $i \in V(H) = [h]$: a vertex $v_{\nu(i)}$ such that $d_{\nu(i)} = d_H(i)$. Furthermore, $v(1), \ldots, v(h)$ are distinct. (As (ϕ1).)

(ϕ'2) For each vertex $i \in V(H) = [h]$: also a bijection between the $d_H(i)$ half-edges at $i \in H$ and the half-edges $\Upsilon_{\nu(i)}$ at $v_{\nu(i)}$. These bijections define how the half-edges in $\bigcup_{i \in [h]} \Upsilon_{\nu(i)}$ are paired in H_ϕ.

For each $k \geqslant 0$, there are h_k vertices i such that $d_H(i) = k$, and we may choose the corresponding $\nu(i)$ in (ϕ'1) in $(n_k)_{h_k}$ ways; then for each of these vertices the bijection of half-edges in (ϕ'2) may be chosen in $k!$ ways. Hence, the number of H_ϕ is

$$|\Phi(H)| = \prod_{k \geqslant 0} (n_k)_{h_k} k!^{h_k}. \tag{8.2}$$

(The product in (8.2) is really finite, since $h_k \neq 0$ only for finitely many k.) Each I_ϕ is a product of $e(H)$ indicators of specific pairings of the type $\mathbf{1}\{\alpha_k$ and α_ℓ are paired in $G^*(n, \mathbf{d}_n)\}$. Hence,

$$\mathbb{E}\, I_\phi = \frac{1}{(N-1)(N-3)\cdots(n-2e(H)+1)} = \frac{1}{((N-1))_{e(H)}}. \tag{8.3}$$

Consequently, (8.1), (8.2) and (8.3) yield the exact formula

$$\mathbb{E}\, \overline{Z}_H^* = \frac{1}{((N-1))_{e(H)}} \prod_{k \geqslant 0} (n_k)_{h_k} k!^{h_k}. \tag{8.4}$$

As $n \to \infty$, we obtain from (8.4) using (A1) and λ_H defined in (3.3),

$$\mathbb{E}\, \overline{Z}_H^* = N^{-e(H)} \big(1 + O(N^{-1})\big) \prod_{k \geqslant 0} \big(np_k + o(n)\big)^{h_k} k!^{h_k}$$

$$= \frac{n^{v(H)}}{N^{e(H)}} \Big(\prod_{k \geqslant 0} \big(p_k k!\big)^{h_k} + o(1) \Big)$$

$$= n^{v(H)-e(H)} \mu^{-e(H)} \Big(\prod_{u \in H} p_{d_H(u)} d_H(u)! + o(1) \Big)$$

$$= n^{v(H)-e(H)} \operatorname{aut}(H) \big(\lambda_H + o(1)\big). \tag{8.5}$$

Using (3.1), this proves (7.1), (7.3) and (7.5).

Now consider two connected multigraphs H and H', and let $h_k' := n_k(H')$. By (8.1),

$$\overline{Z}_H^* \overline{Z}_{H'}^* = \sum_{\phi \in \Phi(H)} \sum_{\phi' \in \Phi(H')} I_\phi I_{\phi'}. \tag{8.6}$$

Let $\phi \in \Phi(H)$ and condition on $I_\phi = 1$. Then H_ϕ is a component of $G^*(n, \mathbf{d}_n)$, and the rest of the graph, i.e. $G^*(n, \mathbf{d}_n) \setminus H_\phi$, is given by another instance of the configuration model, with n_k replaced by $n_k - h_k$ and N by $N - 2e(H)$. Consequently, if we first only consider $\phi \in \Phi(H)$ and $\phi' \in \Phi(H')$ such that H_ϕ and $H_{\phi'}$ are disjoint, then, using (8.4) for $G^*(n, \mathbf{d}_n) \setminus H_\phi$,

$$\mathbb{E} \sum_{\phi, \phi': H_\phi \cap H_{\phi'} = \emptyset} I_\phi I_{\phi'} = \sum_\phi \mathbb{E}\, I_\phi \sum_{\phi': H_\phi \cap H_{\phi'} = \emptyset} \mathbb{E}\big(I_{\phi'} \mid I_\phi = 1\big)$$

$$= \sum_\phi \mathbb{E}\, I_\phi \frac{1}{((N - 2e(H) - 1))_{e(H')}} \prod_{k \geqslant 0} (n_k - h_k)_{h_k'} k!^{h_k'}$$

$$= \mathbb{E}\, \overline{Z}_H^* \frac{1}{((N-2e(H)-1))_{e(H')}} \prod_{k \geqslant 0} (n_k - h_k)_{h'_k}\, k!^{h'_k}$$

$$= \mathbb{E}\, \overline{Z}_H^* \, \mathbb{E}\, \overline{Z}_{H'}^* \frac{((N-1))_{e(H')}}{((N-2e(H)-1))_{e(H')}} \prod_{k \geqslant 0} \frac{(n_k - h_k)_{h'_k}}{(n_k)_{h'_k}}. \qquad (8.7)$$

If H_ϕ and $H_{\phi'}$ are not disjoint, then both can occur as components only if they coincide as unlabelled graphs. Hence, if $H \neq H'$, then $\mathbb{E}\big(\overline{Z}_H^* \overline{Z}_{H'}^*\big)$ is given by (8.7), while if $H = H'$, then there is an additional term $\mathrm{aut}(H)\, \mathbb{E}\, \overline{Z}_H^*$. We switch to counting unlabelled copies, using (3.1) as usual, and obtain

$$\mathbb{E}\big(\overline{Z}_H^{\mathsf{u}*} \overline{Z}_{H'}^{\mathsf{u}*}\big) = \delta_{H,H'}\, \mathbb{E}\, \overline{Z}_H^{\mathsf{u}*}$$

$$+ \mathbb{E}\, \overline{Z}_H^{\mathsf{u}*}\, \mathbb{E}\, \overline{Z}_{H'}^{\mathsf{u}*} \frac{((N-1))_{e(H')}}{((N-2e(H)-1))_{e(H')}} \prod_{k \geqslant 0} \frac{(n_k - h_k)_{h'_k}}{(n_k)_{h'_k}}. \qquad (8.8)$$

Consider first the case when H and H' are unicyclic. If $p_k > 0$ for every k such that $h_k h'_k > 0$, then all fractions in (8.8) tend to 1 as $n \to \infty$, and thus (8.8) and (7.3) yield

$$\mathbb{E}\big(\overline{Z}_H^{\mathsf{u}*} \overline{Z}_{H'}^{\mathsf{u}*}\big) \to \delta_{H,H'}\lambda_H + \lambda_H \lambda_{H'}. \qquad (8.9)$$

If $p_k = 0$ for some k with $h_k h'_k > 0$, then $\mathbb{E}\, \overline{Z}_H^{\mathsf{u}*} \to \lambda_H = 0$ by (7.3) and the definition (3.3) of λ_H (or directly by (8.5)), and (8.8) implies that (8.9) still holds. If $H \neq H'$, (8.9) yields (7.4) with $k = 2$ and $r_1 = r_2 = 1$; if $H = H'$, (8.9) yields $\mathbb{E}\big(\overline{Z}_H^{\mathsf{u}*}\big)_2 \to \lambda_H^2$, another instance of (7.4) ($k = 1$ and $r_1 = 2$).

This shows (7.4) when the degree $\sum_i r_i = 2$, The general case is proved in the same way. The factorial moments in (7.4) means that we only count copies that are distinct, and therefore disjoint, and we may condition on one indicator I_ϕ as in (8.7) and use induction; this is a standard argument and we omit the details.

Finally, consider the case of two trees H and H'. Then (8.8) yields

$$\mathrm{Cov}\big(\overline{Z}_H^{\mathsf{u}*}, \overline{Z}_{H'}^{\mathsf{u}*}\big) = \delta_{H,H'}\, \mathbb{E}\, \overline{Z}_H^{\mathsf{u}*}$$

$$+ \mathbb{E}\, \overline{Z}_H^{\mathsf{u}*}\, \mathbb{E}\, \overline{Z}_{H'}^{\mathsf{u}*} \Big(\frac{((N-1))_{e(H')}}{((N-2e(H)-1))_{e(H')}} \prod_{k \geqslant 0} \frac{(n_k - h_k)_{h'_k}}{(n_k)_{h'_k}} - 1 \Big). \qquad (8.10)$$

If $p_k > 0$ whenever $h_k h'_k > 0$, then

$$\frac{((N-1))_{e(H')}}{((N-2e(H)-1))_{e(H')}} \prod_{k \geqslant 0} \frac{(n_k - h_k)_{h'_k}}{(n_k)_{h'_k}}$$

$$= \Big(1 + \frac{2e(H)e(H')}{N} + O\big(N^{-2}\big)\Big) \prod_{k \geqslant 0} \Big(1 - \frac{h_k h'_k}{n_k} + O\big(n_k^{-2}\big)\Big)$$

$$= 1 + \frac{2e(H)e(H')}{\mu n} - \sum_{k \geqslant 0} \frac{h_k h'_k}{p_k n} + o\big(n^{-1}\big). \qquad (8.11)$$

Hence, (8.10) and (7.3) then yield

$$\mathrm{Cov}\big(\overline{Z}_H^{\mathsf{u}*}, \overline{Z}_{H'}^{\mathsf{u}*}\big) = \delta_{H,H'} n\lambda_H + n\lambda_H \lambda_{H'} \Big(\frac{2e(H)e(H')}{\mu} - \sum_{k \geqslant 0} \frac{h_k h_k'}{p_k}\Big) + o(n). \tag{8.12}$$

If $p_k = 0$ so $n_k = o(n)$ for some k with $h_k h_k' > 0$, then (8.4) yields $\mathbb{E}\,\overline{Z}_H^{\mathsf{u}*}, \mathbb{E}\,\overline{Z}_{H'}^{\mathsf{u}*} = O(n_k)$, and it is easily seen from (8.10) and an expansion as in (8.11) that $\mathrm{Cov}\big(\overline{Z}_H^{\mathsf{u}*}, \overline{Z}_{H'}^{\mathsf{u}*}\big) = o(n)$. Hence, (8.12) holds in this case too.

We have shown (8.12) in general; this is the same as (7.2) with the definition (3.7), which completes the proof of Lemma 7.1. $\qquad\square$

Remark 8.1. We used labelled copies in the proof for convenience and transferred the results to unlabelled copies by (3.1). Alternatively, and essentially equivalently, we may define $\phi_1 \equiv \phi_2$ if $\phi_1, \phi_2 \in \Phi(H)$ and $H_{\phi_1} = H_{\phi_2}$ as unlabelled graphs, and note that then $I_{\phi_1} = I_{\phi_2}$. Hence $\Phi(H)$ splits into equivalence classes of $\mathrm{aut}(H)$ elements each. Define $\Phi_u(H)$ as a subset of $\Phi(H)$ consisting of one element from each equivalence class. Then, cf. (8.1) and (3.1), $|\Phi_u(H)| = \mathrm{aut}(H)^{-1}|\Phi(H)|$ and

$$\overline{Z}_H^{\mathsf{u}*} = \sum_{\phi \in \Phi_u(H)} I_\phi, \tag{8.13}$$

which can be used instead of (8.1), yielding the same results.

Proof of Lemma 7.4. This is similar to the proof of the special case (7.4) in Lemma 7.1.

The special case $k = 0$ is shown in the proof of (2.8) in [21, Section 7 including Remark 6]. In general, we take copies $H_{\phi_{ij}}$ of H_i for $i = 1, \ldots, k$ and $j = 1, \ldots, r_i$, and condition on $I_{\phi_{ij}} = 1$ for all such i and j. Then

$$\mathbb{E}\Big(\big(Z_{\mathsf{C}_1}^{\mathsf{u}*}\big)_{s_1} \big(Z_{\mathsf{C}_2}^{\mathsf{u}*}\big)_{s_2} \,\Big|\, I_{\phi_{ij}} = 1 \forall i, j\Big) \to \lambda_{*\mathsf{C}_1}^{s_1} \lambda_{*\mathsf{C}_2}^{s_2} \tag{8.14}$$

by the case $k = 0$ just discussed applied to $G^*(n, \mathbf{d}_n) \setminus \bigcup_{i,j} H_{\phi_{i,j}}$, and (7.6) follows by the argument in (8.7). (It is here convenient to count unlabelled copies as in Remark 8.1.) $\qquad\square$

Proof of Lemma 7.2. Let V be a uniformly random vertex in $G^*(n, \mathbf{d}_n)$. Then,

$$\mathbb{P}\big(\mathcal{C}(V) \cong H \mid G^*(n, \mathbf{d}_n)\big) = |H|\overline{Z}_H^{\mathsf{u}*}/n, \tag{8.15}$$

since each component isomorphic to H contains $|H|$ vertices that are possible choices of V. Hence, using (7.1),

$$\mathbb{P}\big(\mathcal{C}(V) \cong H\big) = |H|\,\mathbb{E}\,\overline{Z}_H^{\mathsf{u}*}/n \to |H|\lambda_H. \tag{8.16}$$

Furthermore, by the coupling of the exploration process and the branching process $\mathcal{T}$ discussed in Section 2.3,

$$\mathbb{P}\big(\mathcal{C}(V) \cong H\big) = \mathbb{P}\big(\mathcal{T} \cong H\big) + o(1) = p_H + o(1). \tag{8.17}$$

By (8.16) and (8.17), $|H|\lambda_H = p_H$. $\qquad\square$

Proof of Lemma 7.3. Suppose that this fails, and that $H_1, \ldots, H_k$ are distinct trees with $v(H_i) > 1$, $\lambda_{H_i} > 0$ and $\sum_{i,j} a_i a_j \sigma_{H_i, H_j} = 0$ for some real numbers $a_i \neq 0$. Hence, by (3.8),

$$\mathrm{Var}\Big(\sum_{i=1}^k a_i \overline{Z}_{H_k}^{\mathsf{u}*}\Big) = n \sum_{i,j=1}^k a_i a_j \sigma_{H_i, H_j} + o(n) = o(n). \qquad (8.18)$$

Hence, for any tree H, by the Cauchy–Schwarz inequality and (3.8),

$$\mathrm{Cov}\Big(\sum_{i=1}^k a_i \overline{Z}_{H_k}^{\mathsf{u}*}, \overline{Z}_H^{\mathsf{u}*}\Big) \leqslant \mathrm{Var}\Big(\sum_{i=1}^k a_i \overline{Z}_{H_k}^{\mathsf{u}*}\Big)^{1/2} \mathrm{Var}\big(\overline{Z}_H^{\mathsf{u}*}\big)^{1/2} = o(n), \quad (8.19)$$

and thus by (3.8) again,

$$\sum_{i=1}^k a_i \sigma_{H_i, H} = 0. \qquad (8.20)$$

For two trees T_1 and T_2, define

$$\langle T_1, T_2 \rangle := \frac{2e(T_1)e(T_2)}{\mu} - \sum_{k \geqslant 0} \frac{n_k(T_1) n_k(T_2)}{p_k}, \qquad (8.21)$$

so that (3.7) can be written

$$\sigma_{T_1, T_2} = \delta_{T_1, T_2} \lambda_{T_1} + \lambda_{T_1} \lambda_{T_2} \langle T_1, T_2 \rangle. \qquad (8.22)$$

Thus, (8.20) yields, for every tree H,

$$\sum_{i=1}^k a_i \lambda_{H_i} \delta_{H, H_i} + \lambda_H \sum_{i=1}^k a_i \lambda_{H_i} \langle H_i, H \rangle = 0. \qquad (8.23)$$

By assumption (A6), there exists an $r > 1$ such that $p_r > 0$. Given any tree T with $v(T) > 1$, we may replace a leaf by a vertex of degree r, joined to $r - 1$ new leaves. Denote the result by $T^{(1)}$; this is a tree with $n_k(T^{(1)}) = n_k(T) + (r-2)\delta_{k1} + \delta_{kr}$ and $e(T^{(1)}) = e(T) + r - 1$. Hence, for any other tree T',

$$\langle T^{(1)}, T' \rangle = \langle T, T' \rangle + \frac{2(r-1)e(T')}{\mu} - \frac{(r-2)n_1(T')}{p_1} - \frac{n_r(T')}{p_r}. \qquad (8.24)$$

Repeat this procedure and obtain a sequence of trees $T^{(j)}$, $j \geqslant 0$, with $T^{(0)} := T$ (replacing an arbitrary leaf each time). If j is large enough, then $v(T^{(j)}) > v(H_i)$ for $i = 1, \ldots, k$, and thus $T^{(j)} \neq H_i$ and (8.23) yields

$$\lambda_{T^{(j)}} \sum_{i=1}^k a_i \lambda_{H_i} \langle H_i, T^{(j)} \rangle = 0, \qquad j \text{ large}. \qquad (8.25)$$

If $\lambda_T > 0$, then also $\lambda_{T^{(j)}} > 0$ for every $j \geqslant 1$, and thus (8.25) yields $\sum_{i=1}^k a_i \lambda_{H_i} \langle H_i, T^{(j)} \rangle = 0$ for large j. Furthermore, it follows from (8.24) that this sum is a linear function of j, and thus it vanishes for all j, i.e.,

$$\sum_{i=1}^k a_i \lambda_{H_i} \langle H_i, T^{(j)} \rangle = 0, \qquad j \geqslant 0. \qquad (8.26)$$

In particular we can take $j = 0$ in (8.26), and see that $\sum_{i=1}^{k} a_i \lambda_{H_i} \langle H_i, T \rangle = 0$ for every tree T with $v(T) > 0$ and $\lambda_T > 0$. Hence, the second term in (8.23) vanishes for every tree H with $v(H) > 1$, and thus (8.23) and (8.26) imply

$$\sum_{i=1}^{k} a_i \lambda_{H_i} \delta_{H,H_i} = 0 \qquad (8.27)$$

for every such tree H. However, taking $H = H_1$, this yields $a_1 \lambda_{H_1} = 0$, a contradiction. $\qquad\square$

9. Proof of Theorem 3.17

We say that a graph functional ψ has *finite support* if $\psi(H) \neq 0$ for only finitely many unlabelled H, or equivalently, if there exists $K < \infty$ such that

$$\psi(H) = 0 \qquad \text{if } e(H) \geqslant K. \qquad (9.1)$$

Lemma 9.1. *Assume* (A1)–(A2). *Let ψ be a graph functional with finite support and define Ψ by* (3.16). *Then*

$$\mathbb{E}\,\Psi\big(G^*(n, \mathbf{d}_n)\big) = n\,\mathbb{E}\,\psi(\mathcal{T}) + o(n). \qquad (9.2)$$

Proof. Let V be a uniformly random vertex in $G^*(n, \mathbf{d}_n)$, and couple the exploration process of $\mathcal{C}(V)$ with $\mathcal{T}$ as in Section 2.3. Let K be as in (9.1). If the first K generations of the two processes are equal, then either $\mathcal{C}(V) = \mathcal{T}$ (as unlabelled rooted graphs), or both have at least K edges, and in both cases $\psi(\mathcal{C}(V)) = \psi(\mathcal{T})$. Hence, noting that (9.1) also implies that ψ is bounded,

$$\mathbb{E}\,\psi\big(\mathcal{C}(V)\big) = \mathbb{E}\,\psi(\mathcal{T}) + o(1). \qquad (9.3)$$

Furthermore, conditioning on $G^*(n, \mathbf{d}_n)$,

$$\mathbb{E}\big(\psi(\mathcal{C}(V)) \mid G^*(n, \mathbf{d}_n)\big) = \frac{1}{n} \sum_{i=1}^{n} \psi\big(\mathcal{C}(v_i)\big) = \frac{1}{n} \Psi\big(G^*(n, \mathbf{d}_n)\big). \qquad (9.4)$$

Thus, taking the expectation,

$$\mathbb{E}\,\psi(\mathcal{C}(V)) = \frac{1}{n}\,\mathbb{E}\,\Psi\big(G^*(n, \mathbf{d}_n)\big). \qquad (9.5)$$

The result (9.2) follows by combining (9.5) with (9.3). $\qquad\square$

Lemma 9.2. *Assume* (A1)–(A2). *Let ψ be a graph functional with finite support, let K be as in* (9.1) *and let $M := \sup_H |\psi(H)| < \infty$. Then, for n so large that $N \geqslant \mu n/2$ and $N \geqslant 4K$, with $c := 4 + 16/\mu$,*

$$\mathrm{Var}\big(\Psi\big(G^*(n, \mathbf{d}_n)\big)\big) \leqslant cMK^2\,\mathbb{E}\big|\psi(\mathcal{C}(V))\big|n. \qquad (9.6)$$

Proof. Let V_1 and V_2 be independent uniformly random vertices in $G^*(n, \mathbf{d}_n)$. Then

$$\mathbb{E}\big(\psi(\mathcal{C}(V_1))\psi(\mathcal{C}(V_2)) \mid G^*(n, \mathbf{d}_n)\big) = \frac{1}{n^2} \sum_{v_1, v_2 \in G^*(n, \mathbf{d}_n)} \psi(\mathcal{C}(v_1))\psi(\mathcal{C}(v_2))$$

$$= \frac{1}{n^2} \Psi\big(G^*(n, \mathbf{d}_n)\big)^2. \qquad (9.7)$$

We reveal the edges in $G^*(n, \mathbf{d}_n)$ in a special order. (Cf. the related argument in [5, Section 4.2].) First, let Π_1 and Π_2 denote independent

exploration processes of $\mathcal{C}(V_1)$ and $\mathcal{C}(V_2)$, starting at V_1 and V_2 as above. (These may thus conflict. Think of them as exploring different copies of $G^*(n, \mathbf{d}_n)$.) Let $Y_1 := \psi(\mathcal{C}(V_1))$ and $Y_2 := \psi(\mathcal{C}(V_2))$ be given by Π_1 and Π_2, respectively, and note that Y_1 and Y_2 are independent. Next, start from scratch and reveal first the edges of $\mathcal{C}(V_1)$ according to the process Π_1, but stop when K edges have been found, or when $\mathcal{C}(V_1)$ is exhausted (if it has less than K edges), where again K is as in (9.1); let $\mathcal{C}_K(V_1)$ denote the explored part of $\mathcal{C}(V_1)$. Then reveal edges according to Π_2 (starting from V_2) as long this does not involve any half-edge already paired. When the first conflict occurs, abandon Π_2 and pair the remaining half-edges uniformly at random in any order. This yields a copy of $G^*(n, \mathbf{d}_n)$, and we define $X_j := \psi(\mathcal{C}(Y_j))$, $j = 1, 2$, for it.

By the construction, and (9.1), $X_1 = Y_1$. Furthermore, $X_2 \neq Y_2$ only if a conflict has occurred when revealing one of the first K edges according to Π_2. This may happen either because V_2 belongs to $\mathcal{C}_K(V_1)$, or because one of the half-edges already found during the exploration of the first K edges of $\mathcal{C}(V_2)$ is paired by Π_2 with one of the half-edges in $\mathcal{C}_K(V_1)$. Condition on $\mathcal{C}_K(V_1)$. Since we reveal only at most K edges and thus $K + 1$ vertices in $\mathcal{C}_K(V_1)$, the first possibility has probability $\leqslant (K+1)/n$, and the second possibility has probability $\leqslant 2K/(N - 2K)$ for each of at most K pairings; hence, we obtain the union bound, using $N - 2K \geqslant N/2 \geqslant \mu n/4$,

$$\mathbb{P}\big(X_2 \neq Y_2 \mid \mathcal{C}_K(V_1)\big) \leqslant \frac{K+1}{n} + \frac{2K^2}{N - 2K} \leqslant \frac{2K^2}{n}\big(1 + 4\mu^{-1}\big). \qquad (9.8)$$

The construction shows that X_1 and X_2 have the correct joint distribution, while Y_1 and Y_2 have same individual distribution as these, but are independent of each other. Thus, by (9.7) and (9.4),

$$\mathbb{E}\big(X_1 X_2\big) = \mathbb{E}\big(\psi(\mathcal{C}(V_1))\psi(\mathcal{C}(V_2))\big) = n^{-2}\,\mathbb{E}\big(\Psi\big(G^*(n, \mathbf{d}_n)\big)^2\big), \qquad (9.9)$$

$$\mathbb{E}\big(Y_1 Y_2\big) = \big(\mathbb{E}\,Y_1\big)^2 = \big(\mathbb{E}\,\psi(\mathcal{C}(V_1))\big)^2 = n^{-2}\big(\mathbb{E}\,\Psi\big(G^*(n, \mathbf{d}_n)\big)\big)^2. \qquad (9.10)$$

Consequently, recalling $X_1 = Y_1$,

$$n^{-2}\,\mathrm{Var}\big(\Psi(G^*(n, \mathbf{d}_n))\big) = \mathbb{E}(X_1 X_2) - \mathbb{E}(Y_1 Y_2) = \mathbb{E}\big(X_1(X_2 - Y_2)\big). \qquad (9.11)$$

Furthermore, by (9.8), since X_1 is determined by $\mathcal{C}_K(V_1)$,

$$\mathbb{E}\big(|X_2 - Y_2| \mid X_1\big) \leqslant 2M\,\mathbb{P}\big(X_2 \neq Y_2 \mid X_1\big) \leqslant \frac{4MK^2}{n}\big(1 + 4\mu^{-1}\big). \qquad (9.12)$$

Hence, (9.11) yields

$$n^{-2}\,\mathrm{Var}\big(\Psi\big(G^*(n, \mathbf{d}_n)\big)\big) \leqslant \mathbb{E}\big(|X_1||X_2 - Y_2|\big) \leqslant \frac{cMK^2}{n}\,\mathbb{E}\,|X_1|, \qquad (9.13)$$

which yields (9.6). $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\square$

Lemma 9.3. *Assume* (A1)–(A2). *For every $\varepsilon > 0$, there exists $\varepsilon_1 > 0$ such that if $\mathcal{J} \subset [n]$ is any subset with $|\mathcal{J}| \leqslant \varepsilon_1 n$, then $\sum_{i \in \mathcal{J}} d_i < \varepsilon n$. In particular, every subgraph H of $G^*(n, \mathbf{d}_n)$ with $|H| \leqslant \varepsilon_1 n$ has $e(H) < \varepsilon n$ (deterministically).*

Proof. We have $D_n := d_I$, where I is a uniformly random index in $[n]$. The uniform integrability of D_n (see (A2)) means (see e.g. [19, Theorem 5.4.1]) that there exists ε_1 such that for any event $\mathcal{E}$ with $\mathbb{P}(\mathcal{E}) \leqslant \varepsilon_1$, we have

$\mathbb{E}(D_n; \mathcal{E}) < \varepsilon$. Let $\mathcal{E} := \{I \in \mathcal{J}\}$. Then $\mathbb{P}(\mathcal{E}) = |\mathcal{J}|/n$ and $\sum_{i \in \mathcal{J}} d_i = n\, \mathbb{E}(D_n; \mathcal{E})$. $\qquad\square$

We consider now only the supercritical case (A5).

Lemma 9.4. *Assume* (A1)–(A2) *and* (A5)*. Then there exist* $c > 0$, $\varepsilon > 0$ *and* $C < \infty$ *such that, for* $G^*(n, \mathbf{d}_n)$,

$$\mathbb{P}\big(e(\mathcal{C}(V)) = \ell\big) \leqslant C e^{-c\ell}, \qquad 0 \leqslant \ell \leqslant \varepsilon n, \tag{9.14}$$

$$\mathbb{P}\big(|\mathcal{C}(V)| = k\big) \leqslant C e^{-ck}, \qquad 1 \leqslant k \leqslant \varepsilon n. \tag{9.15}$$

Proof. Consider the exploration process, starting at a random vertex V and restarting at a new random vertex when a component is completely explored. Let Q_j be the number of unpaired half-edges in the explored part when j pairings have been made. In particular, $Q_0 = d(V) \overset{\mathrm{d}}{=} D_n$.

By (A5), $\mathbb{E}\,D^2 > 2\mu$, and thus there exists $K < \infty$ such that

$$s_K := \sum_{k=1}^{K} k^2 p_k > 2\mu. \tag{9.16}$$

Let $\varepsilon = (s_K - 2\mu)/(10K^3) > 0$. Consider until further notice only n that are so large that $n_k/n > p_k - \varepsilon$ for $k = 1, \dots, K$, and also $N/n < \mu + \varepsilon$, see (A1) and (2.5).

Let $k \leqslant K$ be such that $p_k \geqslant 3\varepsilon$. During the first εn steps of the exploration process, there is always at least $n_k - \varepsilon n > (p_k - 2\varepsilon)n$ unused vertices of degree k, and thus the probability that the next pairing is with a half-edge at an unused vertex of degree k is at least, noting that $p_k \leqslant \mu$,

$$\frac{k(n_k - \varepsilon n)}{N} > \frac{k(p_k - 2\varepsilon)n}{(\mu + \varepsilon)n} > \frac{k(p_k - 2\varepsilon)(1 - \varepsilon/\mu)}{\mu} > \frac{k(p_k - 3\varepsilon)}{\mu}. \tag{9.17}$$

Let $\xi_1, \xi_2, \dots$ be independent copies of a random variables ξ with the distribution

$$\mathbb{P}(\xi = k) = \begin{cases} k(p_k - 3\varepsilon)_+/\mu, & 1 \leqslant k \leqslant K, \\ 1 - \sum_{j=1}^{K} \mathbb{P}(\xi = j), & k = 0, \end{cases} \tag{9.18}$$

and define $S_m := \sum_{j=1}^{m}(\xi_j - 2)$.

By (9.17), we can couple the exploration process with the variables ξ_i such that for every $j \leqslant \varepsilon n$, $Q_{j+1} - Q_j \geqslant \xi_j - 2 = S_{j+1} - S_j$, and thus, by induction, $Q_j \geqslant S_j$.

Suppose now that the component $\mathcal{C}(V)$ has exactly ℓ edges. Then $Q_\ell = 0$, and hence $S_\ell \leqslant 0$. However, (9.18) implies that

$$\mu\, \mathbb{E}\,\xi = \sum_{k=1}^{K} k^2 (p_k - 3\varepsilon)_+ \geqslant \sum_{k=1}^{K} k^2 p_k - 3K^3\varepsilon > 2\mu \tag{9.19}$$

and thus $\mathbb{E}\,\xi > 2$. Furthermore, ξ is bounded, and thus $h(t) := \mathbb{E}\,e^{t(\xi-2)} < \infty$ for every real t. Hence, $h'(0) = \mathbb{E}(\xi - 2) > 0$, and thus there exist $t_0 < 0$ such that $h(t_0) < 1$. Consequently, using a Chernoff bound,

$$\mathbb{P}\big(e(\mathcal{C}(V)) = \ell\big) \leqslant \mathbb{P}(Q_\ell = 0) \leqslant \mathbb{P}(S_\ell \leqslant 0) \leqslant \mathbb{E}\,e^{t_0 S_\ell} = h(t_0)^\ell, \qquad \ell \leqslant \varepsilon n. \tag{9.20}$$

This proves (9.14), with $c := -\log h(t_0) > 0$, when n is large enough. The result (9.14) extends to all n, since it is trivial for small n if C is large enough.

We turn to (9.15). Let ε_1 be as in Lemma 9.3. Then, any component $\mathcal{C}$ with $|\mathcal{C}| = k \leqslant \varepsilon_1 n$ has at most εn edges, and at least $k - 1$. Consequently, by (9.14), if $k \leqslant \varepsilon_1 n$,

$$\mathbb{P}\big(|\mathcal{C}(V)| = k\big) \leqslant \sum_{\ell=k-1}^{\varepsilon n} \mathbb{P}\big(e(\mathcal{C}(V)) = \ell\big) \leqslant \sum_{\ell=k-1}^{\varepsilon n} C e^{-c\ell} \leqslant C' e^{-ck}. \quad (9.21)$$

Hence (9.15) too holds, if we redefine C and ε. $\square$

We have a similar estimate for the supercritical branching process $\mathcal{T}$.

Lemma 9.5. *Assume* (A1)–(A2) *and* (A5). *Then, for the branching process* $\mathcal{T}$ *in Section 2.3,*

$$\mathbb{P}\big(|\mathcal{T}| = k\big) \leqslant C e^{-ck}, \qquad 1 \leqslant k < \infty. \quad (9.22)$$

Proof. This follows by standard branching process theory.

Alternatively, by the coupling in Section 2.3, or by (9.3) with $\psi(T) := \mathbf{1}\{|T| = k\}$, we have for every fixed $k \geqslant 1$, $\mathbb{P}\big(|\mathcal{C}(V)| = k\big) \to \mathbb{P}\big(|\mathcal{T}| = k\big)$, and thus (9.22) follows from (9.15). $\square$

Proof of Theorem 3.17. The proofs of the two parts are essentially identical, and we give the details only for (i).

(i): Write $G_n^* := G^*(n, \mathbf{d}_n)$.

We truncate. Define, for $\ell, L \geqslant 1$,

$$\psi_\ell(H) := \psi(H) \cdot \mathbf{1}\{e(H) = \ell\}, \quad (9.23)$$

$$\psi_{\leqslant L}(H) := \psi(H) \cdot \mathbf{1}\{e(H) \leqslant L\}, \quad (9.24)$$

and consider the corresponding Ψ_ℓ and $\Psi_{\leqslant L}$.

Let $\varepsilon > 0$ be a fixed small number, chosen so small that Lemma 9.4 holds, as well as the following argument.

If $\Psi^\dagger(G_n^*) \neq \Psi_{\leqslant \varepsilon n}(G_n^*)$, then either $e(\mathcal{C}_1) \leqslant \varepsilon n$, and thus $|\mathcal{C}_1| \leqslant \varepsilon n + 1$, or $e(\mathcal{C}_2) > \varepsilon n$ and thus $|\mathcal{C}_2| > \varepsilon_1 n$ by Lemma 9.3. If ε is small enough, then both events have probability $O\big(e^{-cn}\big)$ by [13, Theorem 2]. Since (3.18) implies $\Psi^\dagger(G_n^*), \Psi_{\leqslant \varepsilon n}(G_n^*) = O\big(n^{m+1}\big)$, we thus have

$$\mathbb{E}\big(\Psi^\dagger(G_n^*) - \Psi_{\leqslant \varepsilon n}(G_n^*)\big)^2 = O\big(n^{2m+2} e^{-cn}\big) = O\big(e^{-cn}\big). \quad (9.25)$$

Hence, it suffices to prove the results for $\Psi_{\leqslant \varepsilon n}(G_n^*)$.

For every $\ell \leqslant \varepsilon n$, (3.18) and Lemma 9.4 yield

$$\mathbb{E}\,|\psi_\ell\big(\mathcal{C}(V)\big)| \leqslant C\ell^m \mathbb{P}\big(e(\mathcal{C}(V)) = \ell\big) \leqslant C\ell^m e^{-c\ell}, \quad (9.26)$$

and thus by (9.5) and Lemma 9.2 (with $M \leqslant C\ell^m$)

$$\mathbb{E}\,|\Psi_\ell(G_n^*)| = n\,\mathbb{E}\,|\psi_\ell(\mathcal{C}(V))| \leqslant C e^{-c\ell} n, \quad (9.27)$$

$$\mathrm{Var}\,\Psi_\ell(G_n^*) \leqslant C\ell^m \ell^2 \,\mathbb{E}\,|\psi_\ell(\mathcal{C}(V))|n \leqslant C\ell^{2m+2} e^{-c\ell} n \leqslant C e^{-c\ell} n. \quad (9.28)$$

Let, for $L \geqslant 1$,

$$X_n := n^{-1/2}\big(\Psi^\dagger(G_n^*) - \mathbb{E}\,\Psi^\dagger(G_n^*)\big), \quad (9.29)$$

$$X_{n;L} := n^{-1/2}\big(\Psi_{\leqslant L}(G_n^*) - \mathbb{E}\,\Psi_{\leqslant L}(G_n^*)\big). \quad (9.30)$$

Then, assuming $n \geqslant L/\varepsilon$,

$$X_n - X_{n;L} = X_n - X_{n;n\varepsilon} + \sum_{\ell=L+1}^{\varepsilon n} n^{-1/2}\big(\Psi_\ell(G_n^*) - \mathbb{E}\,\Psi_\ell(G_n^*)\big), \qquad (9.31)$$

and thus, using (9.25), (9.28) and Minkowski's inequality,

$$\big(\mathrm{Var}(X_n - X_{n;L})\big)^{1/2} \leqslant \big(\mathrm{Var}(X_n - X_{n;\varepsilon n})\big)^{1/2} + \sum_{\ell=L+1}^{\varepsilon n} n^{-1/2}\big(\mathrm{Var}\,\Psi_\ell(G_n^*)\big)^{1/2},$$

$$\leqslant Ce^{-cn} + \sum_{\ell=L+1}^{\varepsilon n} Ce^{-c\ell} \leqslant Ce^{-cL}. \qquad (9.32)$$

Furthermore, for every fixed $L \geqslant 1$, $\psi_{\leqslant L}$ is a functional with finite support, and (3.16) yields a finite linear combination

$$\Psi_{\leqslant L}(G_n^*) = \sum_{e(H)\leqslant L} |H|\psi(H)\overline{Z}_H^{\mathsf{u}*}. \qquad (9.33)$$

Terms where H has a cycle have variance $O(1)$, by Theorem 3.2 or directly by Lemma 7.7. Hence, they can be ignored, and Theorem 3.2 (with Remark 3.5) yields, by (i) and joint convergence for different trees,

$$X_{n;L} := \frac{\Psi_{\leqslant L}(G_n^*) - \mathbb{E}\,\Psi_{\leqslant L}(G_n^*)}{\sqrt{n}} \xrightarrow{\mathrm{d}} N\big(0,\sigma_L^2\big), \qquad (9.34)$$

where

$$\sigma_L^2 := \sum_{|T_1|,|T_2|\leqslant L} |T_1||T_2|\psi(T_1)\psi(T_2)\sigma_{T_1,T_2}, \qquad (9.35)$$

summing over pairs of trees of order less than L.

We will verify below that the sums in (3.22) converge absolutely. Thus, as $L \to \infty$, $\sigma_L^2 \to \sigma_\psi^2$ defined in (3.22). The estimate (9.32), which is uniform in $n \geqslant L/\varepsilon$, implies

$$\lim_{L\to\infty} \limsup_{n\to\infty} \mathbb{E}(X_n - X_{n;L})^2 = \lim_{L\to\infty} \limsup_{n\to\infty} \mathrm{Var}(X_n - X_{n;L}) = 0. \qquad (9.36)$$

This together with the limit (9.34) for each fixed L and $\sigma_L^2 \to \sigma_\psi^2$ imply, see e.g. [10, Theorem 4.2] or [29, Theorem 4.28], $X_n \xrightarrow{\mathrm{d}} N\big(0,\sigma_\psi^2\big)$, which is (3.21).

Furthermore, Theorem 3.2 shows also that $\mathrm{Var}\,X_{n;L} \to \sigma_L^2$ for every fixed L. Hence, Minkowski's inequality and (9.32) imply, for every fixed L,

$$\big|(\mathrm{Var}\,X_n)^{1/2} - \sigma_L\big| \leqslant \big|(\mathrm{Var}\,X_n)^{1/2} - (\mathrm{Var}\,X_{n;L})^{1/2}\big| + \big|(\mathrm{Var}\,X_{n,L})^{1/2} - \sigma_L\big|$$

$$\leqslant \big(\mathrm{Var}(X_n - X_{n;L})\big)^{1/2} + o(1) \leqslant Ce^{-cL} + o(1), \qquad (9.37)$$

and thus,

$$\big|(\mathrm{Var}\,X_n)^{1/2} - \sigma_\psi\big| \leqslant \big|(\mathrm{Var}\,X_n)^{1/2} - \sigma_L\big| + |\sigma_L - \sigma_\psi|$$

$$\leqslant Ce^{-cL} + |\sigma_L - \sigma_\psi| + o(1). \qquad (9.38)$$

Take $\limsup_{n\to\infty}$ in (9.38) and then let $L \to \infty$. This yields

$$\limsup_{n\to\infty}\big|(\mathrm{Var}\,X_n)^{1/2} - \sigma_\psi\big| = 0, \qquad (9.39)$$

and thus $\operatorname{Var} X_n \to \sigma_\psi^2$, which is equivalent to (3.20).

Similarly, if $n \geqslant L/\varepsilon$, then (9.25) and (9.27) yield

$$\mathbb{E}\,|\Psi^\dagger(G_n^*) - \Psi_{\leqslant L}(G_n^*)| \leqslant \mathbb{E}\,|\Psi^\dagger(G_n^*) - \Psi_{\leqslant \varepsilon n}(G_n^*)| + \sum_{\ell=L+1}^{\varepsilon n} \mathbb{E}\,|\Psi_\ell(G_n^*)|$$

$$\leqslant Ce^{-cn} + C\sum_{\ell=L+1}^{\varepsilon n} Ce^{-c\ell}n \leqslant Ce^{-cL}n. \qquad (9.40)$$

Moreover, $\mathbb{E}\,\Psi_{\leqslant L}(G_n^*)/n \to \mathbb{E}\,\psi_{\leqslant L}(\mathcal{T})$ by Lemma 9.1, and $\mathbb{E}\,\psi_{\leqslant L}(\mathcal{T}) \to \mathbb{E}^\dagger\,\psi(\mathcal{T})$ as $L \to \infty$ by (2.11), noting that $\mathbb{E}^\dagger\,|\psi(\mathcal{T})| < \infty$ by (3.18) and Lemma 9.5. Hence, for any fixed L, (9.40) implies

$$|\mathbb{E}\,\Psi^\dagger(G_n^*)/n - \mathbb{E}^\dagger\,\psi(\mathcal{T})| \leqslant \mathbb{E}\,|\Psi^\dagger(G_n^*)/n - \Psi_{\leqslant L}(G_n^*)/n|$$

$$+ |\mathbb{E}\,\Psi_{\leqslant L}(G_n^*)/n - \mathbb{E}\,\psi_{\leqslant L}(\mathcal{T})| + |\mathbb{E}\,\psi_{\leqslant L}(\mathcal{T}) - \mathbb{E}^\dagger\,\psi(\mathcal{T})|$$

$$\leqslant Ce^{-cL} + o(1) + |\mathbb{E}\,\psi_{\leqslant L}(\mathcal{T}) - \mathbb{E}^\dagger\,\psi(\mathcal{T})|. \qquad (9.41)$$

Taking $\limsup_{n\to\infty}$ and then letting $L \to \infty$ yields (3.19), similarly to (9.39).

Finally, to verify absolute convergence of the sums in (3.22), suppose temporarily that $\psi(T) \geqslant 0$, so that we may interchange order in the summations freely. Then, considering the terms in (3.7) separately and using (3.5) and (2.11),

$$\sum_{T_1,T_2} |T_1||T_2|\psi(T_1)\psi(T_2)\delta_{T_1,T_2}\lambda_{T_1} = \sum_T |T|^2\psi(T)^2\lambda_T = \sum_T p_T|T|\psi(T)^2$$

$$= \mathbb{E}^\dagger\big(|\mathcal{T}|\psi(\mathcal{T})^2\big), \qquad (9.42)$$

$$\sum_{T_1,T_2} |T_1||T_2|\psi(T_1)\psi(T_2)\lambda_{T_1}\lambda_{T_2}e(T_1)e(T_2) = \Big(\sum_T p_T\psi(T)e(T)\Big)^2$$

$$= \big(\mathbb{E}^\dagger\big(\psi(\mathcal{T})e(\mathcal{T})\big)\big)^2, \qquad (9.43)$$

$$\sum_{T_1,T_2} |T_1||T_2|\psi(T_1)\psi(T_2)\lambda_{T_1}\lambda_{T_2}\sum_{k\geqslant 0}\frac{n_k(T_1)n_k(T_2)}{p_k}$$

$$= \sum_{k\geqslant 0}\frac{1}{p_k}\Big(\sum_T p_T\psi(T)n_k(T)\Big)^2 = \sum_{k\geqslant 0}\frac{1}{p_k}\big(\mathbb{E}^\dagger\big(\psi(\mathcal{T})n_k(\mathcal{T})\big)\big)^2. \quad (9.44)$$

The expectations in (9.42) and (9.43) are finite by (3.18) and Lemma 9.5, and so is each expectation in (9.44). If also the sum in (9.44) converges, then (3.7) and (9.42)–(9.44) show the last equality in (3.22). In particular, if (9.44) is finite, then (3.22) yields, since $\sigma_\psi^2 \geqslant 0$,

$$\sum_{k\geqslant 0}\frac{1}{p_k}\big(\mathbb{E}^\dagger(n_k(\mathcal{T})\psi(\mathcal{T}))\big)^2 \leqslant \mathbb{E}^\dagger\big(|\mathcal{T}|\psi(\mathcal{T})^2\big) + \frac{2}{\mu}\big(\mathbb{E}^\dagger(e(\mathcal{T})\psi(\mathcal{T}))\big)^2. \quad (9.45)$$

For any $\psi \geqslant 0$, this applies to the truncation $\psi_{\leqslant L}$ for any L, since we have $\psi_{\leqslant L}(T)n_k(T) = 0$ for every T when $k > L$, and thus the sum over k in (9.44) converges. Hence (9.45) holds for $\psi_{\leqslant L}$, and letting $L \to \infty$ shows (by monotone convergence) that (9.45) holds for ψ too. We have already seen that the right-hand side in (9.45) is finite, and thus the sum on the left converges, for any $\psi \geqslant 0$ satisfying (3.18).

Consequently, all sums in (9.42)–(9.44) converge when $\psi \geqslant 0$. Hence, applying this to $|\psi|$, we see that for every ψ satisfying (3.18), all sums in (9.42)–(9.44) converge absolutely; hence the equalities in (9.42)–(9.44) hold in general, which verifies (3.22), with absolute convergence everywhere.

(ii): The main difference is that we use Theorem 3.9 instead of Theorem 3.2. Since we now assume (A3), (2.8) holds, and thus all estimates of means and variances in the proof of (i) hold automatically for $G(n, \mathbf{d}_n)$ too by conditioning. $\qquad\square$

Proof of Theorem 3.16. As said in Example 3.20, we apply Theorem 3.17 with $\psi(H) := 1$, together with Lemmas 10.5 and 11.1 below. $\qquad\square$

10. NON-ZERO VARIANCE?

The asymptotic variance σ_ψ^2 in Theorem 3.17 necessarily satifies $\sigma_\psi^2 \geqslant 0$; however, $\sigma_\psi^2 = 0$ is possible. We note first some trivial cases.

Example 10.1. (a) If (A6) does not hold, so (in the supercritical case) $p_1 = 0$, then $\sigma_\psi^2 = 0$ for every ψ, see Remark 3.19.
  (b) If $\psi(T) = \delta_{T,\mathsf{K}_1}$, then $\sigma_\psi^2 = 0$. In fact, then $\Psi(G) = n_0(G)$ counts isolated vertices, and thus $\Psi^\dagger(G(n, \mathbf{d}_n))$ is deterministic.
  (c) If $\psi(T) = 0$ for every tree T with $p_T > 0$, then $\psi(\mathcal{T}) = 0$ a.s. and thus $\sigma_\psi^2 = 0$. In this case, $\psi(\mathcal{C}_j) = 0$ for all but a few components $\mathcal{C}_j$.

Parts (b) and (c) of Example 10.1 show that the values of $\psi(H)$ for $H = \mathsf{K}_1$, trees H with $p_H = 0$, and non-trees H, do not affect σ_ψ^2.

We conjecture that the trivial cases in Example 10.1 (in combination) is the only way to get $\sigma_\psi^2 = 0$ (in the supercritical case). Formally:

Conjecture 10.2. *If* (A1)–(A3) *and* (A5)–(A6) *hold, then*

$$\sigma_\psi^2 = 0 \iff \psi(T) = 0 \text{ for every tree } T \text{ with } |T| > 1 \text{ and } p_T > 0. \quad (10.1)$$

Remark 10.3. As said in Section 3.3, we consider for simplicity only the supercritical case in Theorem 3.17, although we expect a similar result (for Ψ) also in the subcritical case under suitable conditions. However, note that then there are two further trivial cases with $\sigma_\psi^2 = 0$, viz. $\psi(H) = 1$ and $\psi(H) = e(H)/|H|$ as in Example 3.20 (and linear combinations of them), since then $\Psi(G) = |G|$ and $e(G)$, respectively, which are deterministic for $G(n, \mathbf{d}_n)$.

Although we have not been able to verify Conjecture 10.2, we can show $\sigma_\psi^2 > 0$ in many cases.

Lemma 10.4. *Assume* (A1)–(A2) *and* (A5)–(A6). *Suppose that the graph functional ψ satisfies* (3.18) *and that $\sigma_\psi^2 = 0$. Then ψ has the form, for some real constants a_k and every tree T with $p_T > 0$,*

$$\psi(T) = \frac{1}{|T|} \sum_{k=0}^\infty a_k n_k(T), \quad (10.2)$$

where, furthermore, for every $k \geqslant 0$,

$$p_k a_k = \mathbb{E}^\dagger\big(n_k(\mathcal{T})\psi(\mathcal{T})\big) - \frac{k p_k}{\mu} \mathbb{E}^\dagger\big(e(\mathcal{T})\psi(\mathcal{T})\big) \quad (10.3)$$

$$= \sum_{j=0}^{\infty} a_j \, \mathbb{E}^{\dagger} \Big(\frac{n_k(\mathcal{T})n_j(\mathcal{T})}{|\mathcal{T}|} - \frac{kp_k}{\mu} \frac{e(\mathcal{T})n_j(\mathcal{T})}{|\mathcal{T}|} \Big). \qquad (10.4)$$

Note that (10.4) is an eigenvalue equation for the vector $(a_k)_k$.

Proof. Let Φ be the linear space of all graph functionals satisfying (3.18). The right-hand side of (3.22) is a quadratic form in $\psi \in \Phi$. Denote the corresponding symmetric bilinear form by $\langle \psi_1, \psi_2 \rangle$; thus $\sigma_\psi^2 = \langle \psi, \psi \rangle$ and

$$\langle \psi_1, \psi_2 \rangle = \mathbb{E}^{\dagger}\big(|\mathcal{T}|\psi_1(\mathcal{T})\psi_2(\mathcal{T})\big) + \frac{2}{\mu} \, \mathbb{E}^{\dagger}(e(\mathcal{T})\psi_1(\mathcal{T})) \, \mathbb{E}^{\dagger}(e(\mathcal{T})\psi_2(\mathcal{T}))$$

$$- \sum_{k \geqslant 0} \frac{1}{p_k} \, \mathbb{E}^{\dagger}(n_k(\mathcal{T})\psi_1(\mathcal{T})) \, \mathbb{E}^{\dagger}(n_k(\mathcal{T})\psi_2(\mathcal{T})). \qquad (10.5)$$

Since (3.20) implies $\sigma_\psi^2 \geqslant 0$, the bilinear form (10.5) is positive semidefinite, and thus the Cauchy–Schwarz inequality holds for it. In particular, if $\langle \psi, \psi \rangle = \sigma_\psi^2 = 0$, then $\langle \psi, \psi' \rangle = 0$ for every graph functional $\psi' \in \Phi$. Hence, taking $\psi'(H) := \mathbf{1}\{H \cong T\}$ for a tree T,

$$0 = p_T |T|\psi(T) + \frac{2}{\mu} \, \mathbb{E}^{\dagger}(e(\mathcal{T})\psi(\mathcal{T}))p_T e(T) - \sum_{k \geqslant 0} \frac{1}{p_k} \, \mathbb{E}^{\dagger}(n_k(\mathcal{T})\psi(\mathcal{T}))p_T n_k(T).$$
$$(10.6)$$

Thus, for every tree T such that $p_T > 0$, using $2e(T) = \sum_k kn_k(T)$,

$$|T|\psi(T) = \sum_{k \geqslant 0} \frac{1}{p_k} \, \mathbb{E}^{\dagger}(n_k(\mathcal{T})\psi(\mathcal{T}))n_k(T) - \frac{2}{\mu} \, \mathbb{E}^{\dagger}(e(\mathcal{T})\psi(\mathcal{T}))e(T)$$

$$= \sum_{k \geqslant 0} \Big(\frac{1}{p_k} \, \mathbb{E}^{\dagger}(n_k(\mathcal{T})\psi(\mathcal{T})) - \frac{k}{\mu} \, \mathbb{E}^{\dagger}(e(\mathcal{T})\psi(\mathcal{T})) \Big) n_k(T), \qquad (10.7)$$

which yields (10.2)–(10.3); then (10.4) follows by substituting (10.2) in (10.3). $\qquad \square$

Lemma 10.5. *Assume* (A1)–(A2) *and* (A5)–(A6). *Let* $\psi(H) := 1$. *Then* $\sigma_\psi^2 > 0$.

Proof. Suppose that $\sigma_\psi^2 = 0$; thus (10.2)–(10.4) hold by Lemma 10.4. Furthermore, (10.2) trivially holds with $a_k = 1$, and since the coefficients a_k in (10.2) are uniquely determined for every k with $p_k > 0$, (10.3) yields

$$p_k = p_k a_k = \mathbb{E}^{\dagger} n_k(\mathcal{T}) - \frac{kp_k}{\mu} \, \mathbb{E}^{\dagger} e(\mathcal{T}), \qquad k \geqslant 0. \qquad (10.8)$$

Summing over k yields

$$1 = \sum_k \Big(\mathbb{E}^{\dagger} n_k(\mathcal{T}) - \frac{kp_k}{\mu} \, \mathbb{E}^{\dagger} e(\mathcal{T}) \Big) = \mathbb{E}^{\dagger} |\mathcal{T}| - \mathbb{E}^{\dagger} e(\mathcal{T}) = \mathbb{E}^{\dagger}\big(|\mathcal{T}| - e(\mathcal{T})\big)$$

$$= \mathbb{E}^{\dagger} 1 = \mathbb{P}(|\mathcal{T}| < \infty). \qquad (10.9)$$

This is a contradiction. $\qquad \square$

Remark 10.6. As said in Example 3.20 (and more generally in Conjecture 10.2), we conjecture that $\sigma_\psi^2 > 0$ also for $\psi(H) := e(H)/|H|$, and for

(non-zero) linear combinations of these two graph functionals, but we have failed to show this in general, and leave this as an open problem.

Frank Ball (personal communication) has noted that in the case of bounded maximum degree treated in his paper [4] (on epidemics, but with the giant component as a special case), his formulas [4, (5.25)–(5.27)] express the asymptotic variance as a sum of integrals of squares, making it easy to show that the asymptotic variance is strictly positive in these cases. It seems likely that this formula for the asymptotic variance holds more generally (perhaps by purely algebraic manipulations), which might lead to a general proof, but we have not checked the details.

Example 10.7. Let $\psi(H) := |H|$. (A6) shows both $p_1 > 0$ and the existence of $r > 1$ with $p_r > 0$. There exist arbitrarily large trees T with all degrees in $\{1, r\}$, and they have $p_T > 0$. However, (10.2) cannot hold for all such trees, since the right-hand side is bounded for them by $|a_1| + |a_r|$, while $\psi(T) = |T|$ is unbounded. Hence Lemma 10.4 shows that $\sigma_\psi^2 > 0$, as claimed in Example 3.21.

11. The variance of the giant

Lemma 11.1. *If* $\psi(H) = 1$, *then the variance* σ_ψ^2 *in* (3.22) *is given by* (3.15).

Proof. Let $\mathcal{T}_1$ be the Galton–Watson tree with offspring distribution $Y_1 := \widehat{D} - 1$, with $\widehat{D}$ given by (2.9). Recall that $\mathcal{T}$ has a root with a random number D of copies of $\mathcal{T}_1$ attached to it. The probability generating function of Y_1 is given by, using (2.9) and (2.12),

$$f_1(z) := \mathbb{E}\, z^{Y_1} = \mathbb{E}\, z^{\widehat{D}-1} = \sum_{k=1}^{\infty} \frac{k p_k}{\mu} z^{k-1} = \mu^{-1} f'(z). \tag{11.1}$$

The probability that the supercritical Galton–Watson process $\mathcal{T}_1$ is finite is the unique root $\zeta \in [0, 1)$ of $f_1(\zeta) = \zeta$, which by (11.1) is equivalent to (2.13).

Let $\mathcal{T}_2$ be $\mathcal{T}_1$ conditioned on being finite. Then, see e.g. [2, Theorem 1.12.3], $\mathcal{T}_2$ is another Galton–Watson process, which is subcritical and has an offspring distribution Y_2 with probability generating function, using (11.1),

$$f_2(z) := \mathbb{E}\, z^{Y_2} = \frac{f_1(\zeta z)}{\zeta} = \frac{f'(\zeta z)}{\mu \zeta}. \tag{11.2}$$

In particular,

$$\mathbb{E}\, Y_2 = f_2'(1) = \frac{f''(\zeta)}{\mu}. \tag{11.3}$$

Since $\mathcal{T}_2$ is subcritical, thus $f''(\zeta)/\mu < 1$.

Let $n_k^{\mathsf{out}}(\mathcal{T}_2)$ be the number of vertices in $\mathcal{T}_2$ with outdegree k. Then, by a standard calculation for subcritical Galton–Watson trees, summing the expected number of such vertices in generation $j \geqslant 0$, and using (11.2)–(11.3) and (2.9),

$$\mathbb{E}\, n_k^{\mathsf{out}}(\mathcal{T}_2) = \sum_{j=0}^{\infty} (\mathbb{E}\, Y_2)^j \, \mathbb{P}(Y_2 = k) = \frac{1}{1 - \mathbb{E}\, Y_2} \zeta^{k-1} \, \mathbb{P}(Y_1 = k)$$

$$= \frac{\zeta^{k-1}\,\mathbb{P}(\widehat{D}=k+1)}{1-\mathbb{E}\,Y_2} = \frac{(k+1)p_{k+1}\zeta^{k-1}}{\mu-f''(\zeta)}. \tag{11.4}$$

If the root of $\mathcal{T}$ has degree ℓ, then the tree is finite with probability ζ^ℓ, and conditioned on this event, it has ℓ branches that are copies of $\mathcal{T}_2$. Hence, using (11.4) and (2.13),

$$\mathbb{E}^\dagger\,n_k(\mathcal{T}) = \sum_{\ell=0}^\infty p_l\zeta^\ell\big(\delta_{k\ell}+\ell\,\mathbb{E}\,n_{k-1}^{\mathsf{out}}(\mathcal{T}_2)\big) = p_k\zeta^k + \sum_{\ell=0}^\infty p_l\zeta^\ell\ell\frac{kp_k\zeta^{k-2}}{\mu-f''(\zeta)}$$

$$= p_k\zeta^k + f'(\zeta)\frac{kp_k\zeta^{k-1}}{\mu-f''(\zeta)} = p_k\zeta^k + \frac{kp_k\zeta^k\mu}{\mu-f''(\zeta)}. \tag{11.5}$$

Summing over k we find

$$\mathbb{E}^\dagger\,|\mathcal{T}| = \sum_{k=0}^\infty \mathbb{E}^\dagger\,n_k(\mathcal{T}) = f(\zeta) + \frac{\zeta f'(\zeta)\mu}{\mu-f''(\zeta)} = f(\zeta) + \frac{\mu^2\zeta^2}{\mu-f''(\zeta)}, \tag{11.6}$$

$$\mathbb{E}^\dagger\,e(\mathcal{T}) = \mathbb{E}^\dagger\big(|\mathcal{T}|-1\big) = \mathbb{E}^\dagger\,|\mathcal{T}| - \mathbb{P}(|\mathcal{T}|<\infty) = \mathbb{E}^\dagger\,|\mathcal{T}| - f(\zeta)$$

$$= \frac{\mu^2\zeta^2}{\mu-f''(\zeta)}. \tag{11.7}$$

Hence, (3.22) with $\psi=1$ yields

$$\sigma^2 = f(\zeta) + \frac{\mu^2\zeta^2}{\mu-f''(\zeta)} + 2\frac{\mu^3\zeta^4}{(\mu-f''(\zeta))^2} - \sum_{k=0}^\infty p_k\Big(1+k\frac{\mu}{\mu-f''(\zeta)}\Big)^2\zeta^{2k}, \tag{11.8}$$

which yields (3.15) by expanding the square and summing, using (2.12). $\quad\square$

12. Random degrees

One reason for the importance of the model $G(n,\mathbf{d}_n)$ is that for several models of random graphs, if we condition on the degree sequence, then we obtain a graph of the type $G(n,\mathbf{d}_n)$, for the observed degree sequence $\mathbf{d}_n$. This includes the Erdős–Rényi graphs $G(n,p)$ and $G(n,m)$, and several others, see e.g. [15]. Such random graphs can thus be regarded as $G(n,\mathbf{d}_n)$ based on a random degree sequence $\mathbf{d}_n$. We obtain easily results for such graphs too, by conditioning on $\mathbf{d}_n$. Note that, as pointed out e.g. by Ball and Neal [5], the randomness in the degree sequence will in general affect the asymptotic variance. We illustrate this by considering in some detail the counts of small isolated trees in the basic theorem Theorem 3.9. Similar versions of e.g. Theorems 3.16 and 3.17 follow similarly; see [5] for the variance of the size of the giant component in Theorem 3.16.

Theorem 12.1. *Let $(p_k)_0^\infty$ be a probability distribution satisfying* (A6). *Suppose that $\mathbf{d}_n$ is random, and such that, with $\mu := \sum_k kp_k$,*

$$\frac{n_k-p_kn}{\sqrt{n}} \xrightarrow{\mathrm{d}} \xi_k, \qquad k=0,1,\ldots, \tag{12.1}$$

$$\frac{\sum_k kn_k-\mu n}{\sqrt{n}} \xrightarrow{\mathrm{d}} \sum_{k=1}^\infty k\xi_k, \tag{12.2}$$

jointly, where ξ_k are jointly normal with $\mathbb{E}\,\xi_k = 0$ and some covariances $\mathrm{Cov}(\xi_k, \xi_\ell) = \gamma_{k\ell}$ *with* $\sum_{k,\ell} k\ell|\gamma_{k\ell}| < \infty$, *and furthermore that*

$$n^{-1} \sum_{k=1}^{\infty} k^2 n_k \overset{\mathrm{p}}{\longrightarrow} \sum_k p_k k^2 < \infty. \tag{12.3}$$

Then, for any tree H,

$$\frac{\overline{Z}_H^{\mathsf{u}} - n\lambda_H}{\sqrt{n}} \overset{\mathrm{d}}{\longrightarrow} N\big(0, \widehat{\sigma}_H^2\big), \tag{12.4}$$

where λ_H is as in (3.3); furthermore, joint convergence holds for several trees H, with limit $N(0, \widehat{\Sigma})$, where the covariance matrix $\widehat{\Sigma}$ is given by

$$\widehat{\sigma}_{H_1, H_2} := \sigma_{H_1, H_2} + \lambda_{H_1}\lambda_{H_2} \sum_{k,\ell=0}^{\infty} \Big(\frac{n_k(H_1)}{p_k} - \frac{k}{\mu}e(H_1)\Big)\Big(\frac{n_\ell(H_2)}{p_\ell} - \frac{\ell}{\mu}e(H_2)\Big)\gamma_{k\ell}. \tag{12.5}$$

Proof. For convenience, we use the Skorohod coupling theorem [29, Theorem 4.30], and may thus assume that the limits in (12.1)–(12.3) hold a.s. Then (A1)–(A3) hold a.s., and thus Theorem 3.9 applies, conditioned on the degree sequence. Hence, conditionally,

$$\frac{\overline{Z}_H^{\mathsf{u}} - \mathbb{E}\big(\overline{Z}_H^{\mathsf{u}*} \mid \mathbf{d}_n\big)}{\sqrt{n}} \overset{\mathrm{d}}{\longrightarrow} N\big(0, \sigma_H^2\big). \tag{12.6}$$

Consider a tree H and let $h_k := n_k(H)$. If $\lambda_H > 0$, then, a.s., using (8.4) and remembering that n_k and N now are random, together with the a.s. versions of (12.1) and (12.2),

$$\frac{\mathbb{E}\big(\overline{Z}_H^{\mathsf{u}*} \mid \mathbf{d}_n\big)}{n\lambda_H} = \frac{1 + O(n^{-1})}{n\,\mathrm{aut}(H)\lambda_H} N^{-e(H)} \prod_k n_k^{h_k} k!^{h_k}$$

$$= \big(1 + O(n^{-1})\big) \Big(\frac{N}{n\mu}\Big)^{-e(H)} \prod_k \Big(\frac{n_k}{np_k}\Big)^{h_k}$$

$$= 1 - \frac{e(H)}{\mu\sqrt{n}} \sum_{k=1}^{\infty} k\xi_k + \sum_{k=0}^{\infty} \frac{h_k}{p_k\sqrt{n}}\xi_k + o\big(n^{-1/2}\big)$$

$$= 1 + n^{-1/2} \sum_{k=0}^{\infty} \Big(\frac{h_k}{p_k} - \frac{k}{\mu}e(H)\Big)\xi_k + o\big(n^{-1/2}\big) \tag{12.7}$$

and thus

$$\frac{\mathbb{E}\big(\overline{Z}_H^{\mathsf{u}*} \mid \mathbf{d}_n\big) - n\lambda_H}{\sqrt{n}} \overset{\mathrm{d}}{\longrightarrow} \Xi_H := \lambda_H \sum_{k=0}^{\infty} \Big(\frac{h_k}{p_k} - \frac{k}{\mu}e(H)\Big)\xi_k. \tag{12.8}$$

On the other hand, if $\lambda_H = 0$, then $p_k = 0$ for some k with $h_k > 0$. Since $p_k = 0$, (12.1) implies $\xi_k \geqslant 0$ a.s., and thus $\xi_k = 0$ a.s., so $n_k = o(n^{1/2})$ a.s. Hence, (8.4) implies $\mathbb{E}\big(\overline{Z}_H^{\mathsf{u}*} \mid \mathbf{d}_n\big) = o\big(n^{1/2}\big)$ a.s., and thus (12.8) holds in this case too, with $\Xi_H = 0$.

Since (12.6) holds, with the same limit, conditioned on $\mathbf{d}_n$, the limits (12.6) and (12.8) hold jointly, with independent limits. Hence, we can take

their sum and obtain (12.4) with $\widehat{\sigma}^2_{H_1,H_2} := \sigma^2_{H_1,H_2} + \mathrm{Cov}\big(\Xi_{H_1}, \Xi_{H_2}\big)$, i.e., (12.5). $\qquad\qquad\square$

Example 12.2. One case studied also by Ball and Neal [5] is when the degrees d_i are i.i.d. random variables that are copies of a given D, which we assume satisfies $\mathbb{E}\, D^2 < \infty$. (We ignore one half-edge if the sum N otherwise becomes odd.) Then, by the central limit theorem and the law of large numbers, (12.1)–(12.3) hold with

$$\gamma_{k\ell} = \delta_{k\ell} p_k - p_k p_l. \tag{12.9}$$

Hence, Theorem 12.1 shows that (12.4) holds. Furthermore, the double sum in (12.5) is by (12.9)

$$\sum_{k=0}^{\infty} p_k \Big(\frac{n_k(H_1)}{p_k} - \frac{k}{\mu} e(H_1) \Big) \Big(\frac{n_k(H_2)}{p_k} - \frac{k}{\mu} e(H_2) \Big)$$
$$- \sum_{k=0}^{\infty} p_k \Big(\frac{n_k(H_1)}{p_k} - \frac{k}{\mu} e(H_1) \Big) \sum_{\ell=0}^{\infty} p_\ell \Big(\frac{n_\ell(H_2)}{p_\ell} - \frac{\ell}{\mu} e(H_2) \Big). \tag{12.10}$$

By simple algebra, using $\sum_k n_k(H) = |H|$ and $\sum_k k n_k(H) = 2e(H)$, the first sum in (12.10) equals

$$\sum_{k=0}^{\infty} \frac{n_k(H_1) n_k(H_2)}{p_k} - \frac{4e(H_1)e(H_2)}{\mu} + \frac{\mathbb{E}\, D^2}{\mu^2} e(H_1) e(H_2) \tag{12.11}$$

and the second and third are just

$$|H_j| - e(H_j) = 1. \tag{12.12}$$

Hence, (12.5) and (3.7) yield, after some interesting cancellations,

$$\widehat{\sigma}_{H_1,H_2} = \delta_{H_1,H_2} \lambda_{H_1} + \lambda_{H_1} \lambda_{H_2} \Big(\frac{\mathbb{E}\, D(D-2)}{\mu^2} e(H_1) e(H_2) - 1 \Big). \tag{12.13}$$

Example 12.3. Consider the Erdős–Rényi graphs $G(n,p)$ and $G(n,m)$, where we keep the average degree constant by choosing $p = \mu/n$ and $m = \mu n/2 + o(n^{1/2})$ for some fixed $\mu > 0$. The number of isolated trees of a given size in $G(n,m)$ was studied already by Erdős and Rényi [16], but in the range of m considered here, their result contains an error as was pointed out by Barbour [6], who proved asymptotic normality for $G(n,p)$ using Stein's method. The result was extended by Barbour, Karoński and Ruciński [7] to counts of isolated copied of individual trees, our $\overline{Z}^{\mathsf{u}}_T$; the method yields also joint convergence. (We are not aware of any similar result on asymptotic normality proved for $G(n,m)$, but such results might be in the literature.) Although this result thus can be proved directly in a rather simple way, at least for $G(n,p)$, we find it instructive to see how it follows, for both models, from the general results in the present paper.

Asymptotic normality of n_k was shown for $G(n,p)$ by [7], see also [27, Example 6.35]. This was extended to $G(n,m)$ in [20, Theorem 4.1], which implies (as a consequence of [20, (4.2)]) that (12.1)–(12.3) hold for both $G(n,p)$ and $G(n,m)$ (with $p = \mu/n$ and $m = \mu n/2 + o(n^{1/2})$ as above), with

$p_k = \mu^k e^{-\mu}/k!$, the Poisson $\mathrm{Po}(\mu)$ distribution, and asymptotic covariances

$$\gamma_{k\ell} = \begin{cases} \delta_{k\ell} p_k - p_k p_\ell + \frac{(k-\mu)(\ell-\mu)}{\mu} p_k p_\ell, & G(n,p), \\ \delta_{k\ell} p_k - p_k p_\ell - \frac{(k-\mu)(\ell-\mu)}{\mu} p_k p_\ell, & G(n,m). \end{cases} \tag{12.14}$$

Comparing with Example 12.2, we see that $\gamma_{k\ell}$ has an additional term, with different signs in the two cases. To calculate the contribution from that term to (12.5), we calculate, recalling $D \sim \mathrm{Po}(\mu)$ so $\mathbb{E}\, D^2 = \mu^2 + \mu$,

$$\sum_{k=0}^{\infty} \Big(\frac{n_k(H)}{p_k} - \frac{k}{\mu} e(H) \Big)(k-\mu) p_k = \sum_{k=0}^{\infty}(k-\mu) n_k(H) - \sum_{k=0}^{\infty} \frac{p_k k(k-\mu)}{\mu} e(H)$$

$$= 2e(H) - \mu|H| - e(H) = -(\mu-1)|H| - 1. \tag{12.15}$$

Define $\chi := +1$ for $G(n,p)$ and $\chi := -1$ for $G(n,m)$. Then, using (12.14) in (12.5) yields, using (12.13) and (12.15) in the calculations,

$$\widehat{\sigma}_{H_1,H_2} = \delta_{H_1,H_2} \lambda_{H_1}$$

$$+ \lambda_{H_1} \lambda_{H_2} \Big(\frac{\mu-1}{\mu} e(H_1)e(H_2) - 1 + \frac{\chi}{\mu} \big((\mu-1)|H_1| + 1\big)\big((\mu-1)|H_2| + 1\big) \Big). \tag{12.16}$$

Thus, Theorem 12.1 yields asymptotic normality of the counts of isolated trees, in both $G(n,p)$ and $G(n,m)$, with asymptotic covariances (12.16).

Remark 12.4. The model in Example 12.2 with $D \sim \mathrm{Po}(\mu)$ thus gives a result with a covariance matrix (12.13) that is half-way between the results for $G(n,p)$ and $G(n,m)$. As has been remarked before, the same is seen in the much simpler (but related) case $e(G)$ of the number of edges; elementary calculations yield $\mathrm{Var}\, e(G) = \mathrm{Var}(N/2) \sim \mu n/4$ for the model in Example 12.2, $\mathrm{Var}\, e(G) \sim \mu n/2$ for $G(n,p)$, and of course $\mathrm{Var}\, e(G) = 0$ for $G(n,m)$.

References

[1] Omer Angel, Remco van der Hofstad & Cecilia Holmgren: Limit laws for self-loops and multiple edges in the configuration model. Preprint, 2016. `arXiv:1603.07172`

[2] Krishna B. Athreya & Peter E. Ney: *Branching Processes*. Springer-Verlag, Berlin, 1972. MR 0373040

[3] Siva Athreya & D. Yogeshwaran: Central limit theorem for statistics of subcritical configuration models. Preprint, 2018. `arXiv:1808.06778`

[4] Frank Ball: Central limit theorems for SIR epidemics and percolation on configuration model random graphs. Preprint, 2018. `arXiv:1812.03105`

[5] Frank Ball & Peter Neal: The asymptotic variance of the giant component of configuration model random graphs. *Ann. Appl. Probab.* **27** (2017), no. 2, 1057–1092. MR 3655861

[6] A. D. Barbour: Poisson convergence and random graphs. *Math. Proc. Cambridge Philos. Soc.* **92** (1982), no. 2, 349–359. MR 0671189

[7] A. D. Barbour, Michał Karoński & Andrzej Ruciński: A central limit theorem for decomposable random variables with applications to random graphs. *J. Combin. Theory Ser. B* **47** (1989), no. 2, 125–145. MR 1047781

[8] A. D. Barbour & Adrian Röllin: Central limit theorems in the configuration model. Preprint, 2017. `arXiv:1710.02644v1`

[9] Edward A. Bender & E. Rodney Canfield: The asymptotic number of labeled graphs with given degree sequences. *J. Combinatorial Theory Ser. A* **24** (1978), no. 3, 296–307. MR 0505796

[10] Patrick Billingsley: *Convergence of Probability Measures.* Wiley, New York, 1968. MR 0233396

[11] Béla Bollobás: A probabilistic proof of an asymptotic formula for the number of labelled regular graphs. *European J. Combin.* **1** (1980), no. 4, 311–316. MR 0595929

[12] Béla Bollobás & Oliver Riordan: Asymptotic normality of the size of the giant component via a random walk. *J. Combin. Theory Ser. B* **102** (2012), no. 1, 53–61. MR 2871766

[13] Béla Bollobás & Oliver Riordan: An old approach to the giant component problem. *J. Combin. Theory Ser. B* **113** (2015), 236–260. MR 3343756

[14] Béla Bollobás & John C. Wierman: Subgraph counts and containment probabilities of balanced and unbalanced subgraphs in a large random graph. *Graph Theory and its Applications: East and West (Jinan, 1986)*, 63–70, Ann. New York Acad. Sci., 576, New York, 1989. MR 1110801

[15] Tom Britton, Mia Deijfen & Anders Martin-Löf: Generating simple random graphs with prescribed degree distribution. *J. Stat. Phys.*, **124** (2006), no. 6, 1377–1397. MR 2266448

[16] P. Erdős & A. Rényi: On the evolution of random graphs. *Magyar Tud. Akad. Mat. Kutató Int. Közl.* **5** (1960), 17–61. MR 0125031

[17] Valentin Féray: Asymptotic behavior of some statistics in Ewens random permutations. *Electron. J. Probab.* **18** (2013), paper no. 76, 32 pp. MR 3091722

[18] Valentin Féray: Weighted dependency graphs. *Electron. J. Probab.* **23** (2018), paper no. 93, 65 pp. MR 3858921

[19] Allan Gut: *Probability: A Graduate Course*, 2nd ed. Springer, New York, 2013. MR 2977961

[20] Svante Janson: Monotonicity, asymptotic normality and vertex degrees in random graphs. *Bernoulli* **13** (2007), no. 4, 952–965. MR 2364221

[21] Svante Janson: The probability that a random multigraph is simple. *Combin. Probab. Comput.* **18** (2009), no. 1-2, 205–225. MR 2497380

[22] Svante Janson: Susceptibility of random graphs with given vertex degrees. *J. Comb.* **1** (2010), no. 3-4, 357–387. MR 2799217

[23] Svante Janson: The probability that a random multigraph is simple. II. *J. Appl. Probab. 51A* (2014), 123–137. MR 3317354

[24] Svante Janson & Malwina Luczak: Asymptotic normality of the k-core in random graphs. *Ann. Appl. Probab.* **18** (2008), no. 3, 1085–1137. MR 2418239

[25] Svante Janson & Malwina Luczak: A new approach to the giant component problem. *Random Structures Algorithms* **34** (2009), no. 2, 197–216. MR 2490288

[26] Svante Janson, Malwina Luczak & Peter Windridge: Law of large numbers for the SIR epidemic on a random graph with given degrees. *Random Structures Algorithms* **45** (2014), no. 4, 726–763. MR 3275704

[27] Svante Janson, Tomasz Łuczak & Andrzej Ruciński: *Random Graphs.* Wiley, New York, 2000. MR 1782847

[28] Felix Joos, Guillem Perarnau, Dieter Rautenbach & Bruce Reed: How to determine if a random graph with a fixed degree sequence has a giant component. *Probab. Theory Related Fields* **170** (2018), no. 1-2, 263–310. MR 3748325

[29] Olav Kallenberg: *Foundations of Modern Probability*, 2nd ed. Springer, New York, 2002. MR 1876169

[30] Wasiur R. KhudaBukhsh, Casper Woroszylo, Grzegorz A. Rempała & Heinz Koeppl: Functional central limit theorem for susceptible-infected process on configuration model graphs. Preprint, 2017. `arXiv:1703.06328`

[31] V.P. Leonov & A.N. Shiryaev: On a method of semi-invariants. *Teor. Veroyatnost. i Primenen.* **4** (1959), 342–355. (Russian.) English transl.: *Theor. Probability Appl.* **4** (1959), 319–329. MR 0123345

[32] Michael Molloy & Bruce Reed: A critical point for random graphs with a given degree sequence. *Random Structures Algorithms* **6** (1995), no. 2-3, 161–179. MR 1370952

[33] Michael Molloy & Bruce Reed: The size of the giant component of a random graph with a given degree sequence. *Combin. Probab. Comput.* **7** (1998), no. 3, 295–305. MR 1664335

[34] Boris Pittel: On tree census and the giant component in sparse random graphs. *Random Structures Algorithms* **1** (1990), no. 3, 311–342. MR 1099795

[35] Boris Pittel & Nicholas C. Wormald: Counting connected graphs inside-out. *J. Combin. Theory Ser. B* **93** (2005), no. 2, 127–172. MR 2117934

[36] Oliver Riordan: The phase transition in the configuration model. *Combin. Probab. Comput.* **21** (2012), no. 1-2, 265–299. MR 2900063

[37] V. E. Stepanov: On the probability of connectedness of a random graph $\mathcal{G}_m(t)$. (Russian.) *Teor. Verojatnost. i Primenen* **15** (1970), 58–68. English transl.: *Theor. Probability Appl.* **15** (1970), 55–67. MR 0270406

[38] Nicholas C. Wormald: The asymptotic distribution of short cycles in random regular graphs. *J. Combin. Theory Ser. B* **31** (1981), no. 2, 168–182. MR 0630980

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

Email address: svante.janson@math.uu.se

URL: http://www.math.uu.se/svante-janson