SUSCEPTIBILITY IN SUBCRITICAL RANDOM GRAPHS

SVANTE JANSON AND MALWINA J. LUCZAK

Abstract. We study the evolution of the susceptibility in the subcritical random graph $G(n, p)$ as n tends to infinity. We obtain precise asymptotics of its expectation and variance, and show it obeys a law of large numbers. We also prove that the scaled fluctuations of the susceptibility around its deterministic limit converge to a Gaussian law. We further extend our results to higher moments of the component size of a random vertex, and prove that they are jointly asymptotically normal.

1. Introduction

The *susceptibility* $\chi(G)$ of a graph G (deterministic or random) is defined as the mean size of the component containing a random vertex. (As is well known, for random graphs of the random-cluster model, this, or rather its expectation, corresponds to the magnetic susceptibility in Ising and Potts models.) If G has n vertices and components $\mathcal{C}_1, \dots, \mathcal{C}_K$, where K is the number of components, then thus

$$\chi(G) = \sum_{i=1}^{K} \frac{|\mathcal{C}_i|}{n} |\mathcal{C}_i| = \frac{1}{n} \sum_{i=1}^{K} |\mathcal{C}_i|^2. \tag{1.1}$$

We define, for integers $k \geq 1$,

$$S_k(G) := \sum_{i=1}^{K} |\mathcal{C}_i|^k. \tag{1.2}$$

Thus $\chi(G) = n^{-1} S_2(G)$, and similarly $n^{-1} S_{m+1}(G)$ is the mth moment of the size of the component containing a random vertex. (Note that, by choosing a uniform random vertex, we bias the components by their sizes. The mean size of a uniformly chosen random component is n/K, which is different and which will not be treated here.)

The purpose of this paper is to study $\chi(G(n, p))$, or equivalently $S_2(G(n, p))$ for the standard Erdős–Rényi random graph $G(n, p)$ with n vertices where each possible edge appears with probability p, independently of all other edges; we will also give extensions to $S_k(G(n, p))$ for larger k.

We consider asymptotics as $n \to \infty$, with $p = p(n)$ a function of n. (All unspecified limits are as $n \to \infty$.)

Date: 31 May, 2008; revised 27 August, 2008.

2000 *Mathematics Subject Classification.* 05C80, 60C05, 60F05.

Key words and phrases. susceptibility, law of large numbers, central limit theorem.

It is well-known, see e.g. Bollobás [2] and Janson, Łuczak and Ruciński [13], that, if np is a little larger than 1, $np - 1 \gg n^{-1/3}$ to be precise, then $G(n,p)$ has w.h.p. a giant component which is much larger than the others (the *supercritical case*). It is easily seen that then the giant component will dominate all other terms in the sum (1.2); hence, if the largest component is $\mathcal{C}_1$, then $S_k(G(n,p)) = (1+o_{\mathrm{p}}(1))|\mathcal{C}_1|^k$ and $\chi(G(n,p)) = (1+o_{\mathrm{p}}(1))|\mathcal{C}_1|^2/n$. See Appendix A for a more precise statement (and proof).

Similarly, if $np = 1 + O(n^{-1/3})$ (the *critical case*), then there are several components of the order $n^{2/3}$; in this case S_k will be of order $n^{2k/3}$, and thus χ of order $n^{1/3}$, and it follows from Aldous [1] that these quantities, properly normalized, converge in distribution to some random variables but not to constants. See Appendix B for details.

In this paper we therefore concentrate on the case $np < 1$, and in particular $1 - np \gg n^{-1/3}$ (the *subcritical case*). We will prove the following results for $\chi(G(n,p))$, together with similar results for $S_k(G(n,p))$ stated later.

We use O_{p} and o_{p} in the standard sense, see e.g. [13, pp. 10–11], and write $X_n \sim_{\mathrm{p}} a_n$ for $X_n = a_n + o_{\mathrm{p}}(a_n)$ or, equivalently, $X_n/a_n \xrightarrow{\mathrm{p}} 1$. We will also write $X_n = O_{L^p}(a_n)$ if $\|X_n\|_{L^p} := (\mathbb{E}\,|X_n|^p)^{1/p} = O(a_n)$, and, similarly, $X_n = o_{L^p}(a_n)$ if $\|X_n\|_{L^p} = o(a_n)$. (Here, X_n and a_n are sequences of random variables and positive numbers respectively.)

Theorem 1.1. *Uniformly, for all $n \geq 1$ and $0 \leq p < n^{-1}$,*

$$\mathbb{E}\,\chi(G(n,p)) = \frac{1}{1-np}\Big(1 + O\Big(\frac{1}{n(1-np)^3}\Big)\Big), \tag{1.3}$$

$$\operatorname{Var}\chi(G(n,p)) = O\Big(\frac{1}{n(1-np)^5}\Big), \tag{1.4}$$

and

$$\chi(G(n,p)) = \frac{1}{1-np}\Big(1 + O_{\mathrm{p}}\Big(\big(n(1-np)^3\big)^{-1/2}\Big)\Big). \tag{1.5}$$

In particular, if $1 - np \gg n^{-1/3}$, then $\chi(G(n,p)) \sim_{\mathrm{p}} 1/(1-np)$.

One way to handle the explosion at $p = 1/n$ is to consider $1/\mathbb{E}\,\chi$ or $1/\chi$. In this form we can obtain uniform estimates for all p.

Corollary 1.2. *Uniformly, for all $n \geq 1$ and $0 \leq p \leq 1$,*

$$\frac{1}{\mathbb{E}\,\chi(G(n,p))} = (1-np)_+ + O\big(n^{-1/3}\big), \tag{1.6}$$

$$\frac{1}{\chi(G(n,p))} = (1-np)_+ + O_{\mathrm{p}}\big(n^{-1/3}\big). \tag{1.7}$$

The last statement of Theorem 1.1 can be sharpened to asymptotic normality. We will also find the variance more precisely. We write $X_n \sim \mathrm{AsN}(\mu_n, \sigma_n^2)$ if (X_n) is a sequence of random variables and μ_n and $\sigma_n > 0$ are real numbers such that $(X_n - \mu_n)/\sigma_n \xrightarrow{\mathrm{d}} N(0,1)$.

Theorem 1.3. *If $p = p(n) < n^{-1}$ and further $p \gg n^{-2}$ and $1-np \gg n^{-1/3}$, then*

$$\chi(G(n,p)) \sim \operatorname{AsN}\Big(\frac{1}{1-np}, \frac{2p}{(1-np)^5}\Big)$$

and $\operatorname{Var}\chi(G(n,p)) \sim 2p/(1-np)^5$.

It follows easily from $\chi(G(n,p)) > 0$ that the asymptotic normality in Theorem 1.3 cannot hold for $1 - np = O(n^{-1/3})$. Similarly, it does not hold in the case $n^2p = O(1)$, when the number of edges is $O_{\mathrm{p}}(1)$, for example because $n\chi(G(n,p))$ is integer valued. (The variance asymptotics hold in this case too.)

The proof of Theorem 1.1 (given in Sections 3–4) is fairly simple and is based on studying how S_k evolves for the Erdős–Rényi random graph process $\mathcal{G}(n,t)$ (defined in Section 2). Heuristically, it is easy to see that (ignoring the difference between a random variable and its mean), S_k ought to be an approximative solution to the differential equation $f'(t) = f^2(t)$, which (with the initial value $f(0) = n$) is solved by $f(t) = n/(1 - nt)$. We make this precise and rigorous below. This simple idea has presumably been noticed by several people, and at least the leading terms in (1.3) and (1.5) are more or less known folk theorems. However, we do not know of any rigorous treatments, except [19] which uses the susceptibility to study a class of more complicated random graph process. Their processes include the Erdős–Rényi process studied here, so their results include the leading term asymptotics in (1.3) and (1.5) in the case where $p \le (1-\varepsilon)/n$ for some constant $\varepsilon > 0$. Their analysis involves branching processes approximation, as well as differential equations, and seems contingent on the fact that the component distribution (excluding the giant in the supercritical case) has exponentially decaying tails. Furthermore, Durrett [5] studies the expectation $\mathbb{E}\,\chi(G(n,p)))$ using the branching process approximation; the result is stated for fixed $np < 1$ but the argument yields (1.3) for all $np < 1$.

The proof of Theorem 1.3 is more involved; the asymptotic normality is based on using a martingale central limit theorem for a suitable modification of the process $S_k(\mathcal{G}(n,t))$ (Section 5), while the variance is estimated directly (Section 6).

In Section 7, the asymptotic results for S_k are interpreted using the Borel distribution and its moments.

Remark 1.4. It is seen from Theorem 1.1 that the susceptibility blows up at $p = 1/n$, which of course is another sign of the phase transition there, with the emergence of a giant component. In fact, our results give a new proof that there is no giant component for smaller p. In the opposite direction, the explosion of the susceptibility at (or close to) $p = 1/n$ shows that there are large components at that stage; it is tempting to conclude that a giant component emerges around this instance (as we know by other arguments), but a formal proof based on this seems to require some additional work. See

Spencer and Wormald [19] where this type of arguments is used for a class of more complicated random graph processes.

Remark 1.5. An alternative approach to at least some of our results is to use the standard branching process approximation of the neighbourhood exploration process; this will be treated elsewhere for a larger class of random graphs [3].

Remark 1.6. In this paper we study the random graph $G(n,p)$. Most or all of our results transfer easily to the random graph $G(n,m)$ with a fixed number of edges by monotonicity (Lemma 2.1) and the standard device of coupling $G(n,m)$ with $G(n,p)$ for a suitable p such that the expected number of edges is slightly smaller or larger than m. We leave the details to the reader.

Acknowledgements. This work was initiated while both SJ and MJL attended the programme "Combinatorics and Statistical Mechanics" at the Isaac Newton Institute, Cambridge, 2008, where SJ was supported by a Microsoft fellowship and MJL by the Newton Institute.

We thank a referee for a careful reading and correction of several minor errors.

2. Preliminaries

We first note a simple monotonicity.

Lemma 2.1. *If H is a subgraph of G, then $S_k(H) \le S_k(G)$ for every $k \ge 1$.*

Proof. It suffices to consider the case when G is obtained from H by either adding a single edge or adding a single vertex (and no edges); both cases are immediate. $\square$

The random graph process $\mathcal{G}(n,t)$ starts at $t = 0$ with n vertices and no edges, and where edges are added randomly and independently to every possible pair of vertices with rate 1, i.e., the time edge ij is added has an exponential distribution with mean 1. Hence, at a given time t, each possible edge is present with probability $1-e^{-t}$, so $\mathcal{G}(n,t)$ is a random graph $G(n, 1-e^{-t})$. We are interested in the subcritical case where $t < 1/n$; then the difference between $1 - e^{-t}$ and t is $O(t^2) = O(n^{-2})$ which is negligible, and we can see $\mathcal{G}(n,t)$ as a convenient version of $G(n,t)$. More precisely, $G(n,p)$ can be obtained as $\mathcal{G}(n, -\log(1-p))$; this slight reparametrization is annoying but harmless, and it will be convenient in the proofs below.

We write $S_k(t)$ for $S_k(\mathcal{G}(n,t))$. (These and other quantities introduced below depend on n, but we choose not to show this explicitly in the notation.)

We further define, for a graph G with components $\mathcal{C}_i$ and $k, l \ge 1$,

$$S_{k,l}(G) := \sum_{i \ne j} |\mathcal{C}_i|^k |\mathcal{C}_j|^l = S_k(G)S_l(G) - S_{k+l}(G). \tag{2.1}$$

We write $S_{k,l}(t)$ for $S_{k,l}(\mathcal{G}(n,t))$.

3. The expectation

We may and will assume that the edges are added to $\mathcal{G}(n,t)$ at distinct times. If a new edge joins two different components $\mathcal{C}_i$ and $\mathcal{C}_j$ in $\mathcal{G}(n,t)$, then $S_k(t)$ increases by a jump

$$\Delta S_k(t) = \bigl(|\mathcal{C}_i| + |\mathcal{C}_j|\bigr)^k - |\mathcal{C}_i|^k - |\mathcal{C}_j|^k = \sum_{l=1}^{k-1} \binom{k}{l} |\mathcal{C}_i|^l |\mathcal{C}_j|^{k-l}. \tag{3.1}$$

For each unordered pair (i,j), the intensity of such jumps equals the number of possible edges joining the two components, i.e. $|\mathcal{C}_i||\mathcal{C}_j|$. We consider ordered pairs of components and therefore divide this by 2, and summing over all pairs we find that the drift of $S_k(t)$ is

$$V_k(t) := \sum_{i\neq j} \frac{1}{2} |\mathcal{C}_i||\mathcal{C}_j| \sum_{l=1}^{k-1} \binom{k}{l} |\mathcal{C}_i|^l |\mathcal{C}_j|^{k-l} = \sum_{l=1}^{k-1} \frac{1}{2} \binom{k}{l} S_{l+1,k+1-l}(t); \tag{3.2}$$

in other words, noting that $S_k(0) = n$,

$$M_k(t) := S_k(t) - n - \int_0^t V_k(u)\,\mathrm{d}u \tag{3.3}$$

is a martingale on $[0,\infty)$ with $M_k(0) = 0$. (Note that $M_k(t)$ is bounded for each fixed n and t in a finite interval $[0,T]$; hence, there are no problems with integrability of this martingale. The same holds for all similar martingales below.)

We define $s_k(t) := \mathbb{E}\, S_k(t)$, noting that $s_k(0) = n$, and conclude from the martingale property that $\mathbb{E}\, M_k(t) = \mathbb{E}\, M_k(0) = 0$ and thus

$$s_k(t) = \mathbb{E}\, S_k(t) = n + \int_0^t \mathbb{E}\, V_k(u)\,\mathrm{d}u. \tag{3.4}$$

In order to use this, we need information on $\mathbb{E}\, S_{k,l}(t)$.

Lemma 3.1. *For all $k,l \geq 1$:*

(i) $\mathbb{E}\, S_{k,l}(t) \leq s_k(t)s_l(t)$,
(ii) $\mathbb{E}\, S_{k,l}(t) \geq s_k(t)s_l(t) - s_{k+l}(t)$.

Proof. (i): Let $\mathcal{A}_n$ be the set of all non-empty subsets of $[n]$. If $A \in \mathcal{A}_n$, let $I_A(t) := \mathbf{1}[A \text{ is a component of } \mathcal{G}(n,t)]$. Thus,

$$S_k(t) = \sum_{A\in\mathcal{A}_n} |A|^k I_A(t)$$

and, since $I_A I_B = 0$ if $A \cap B \neq \emptyset$ but $A \neq B$,

$$S_{k,l}(t) = \sum_{A\neq B} |A|^k |B|^l I_A(t) I_B(t) = \sum_{A\in\mathcal{A}_n} |A|^k I_A(t) \sum_{B\subseteq[n]\setminus A} |B|^l I_B(t). \tag{3.5}$$

Conditioned on $I_A(t) = 1$, the conditional distribution of the restriction of $\mathcal{G}(n,t)$ to $[n]\setminus A$ is a random graph with the same distribution as $\mathcal{G}(n-|A|,t)$, apart from a relabelling of the vertices. Hence, using also Lemma 2.1,

$$\mathbb{E}\Big(\sum_{B\subseteq[n]\setminus A}|B|^l I_B(t)\;\Big|\; I_A(t)=1\Big)=\mathbb{E}\,S_l(\mathcal{G}(n-|A|,t))\le\mathbb{E}\,S_l(\mathcal{G}(n,t))=s_l(t).$$

Consequently, taking the expectation in (3.5) yields

$$\mathbb{E}\,S_{k,l}(t)\le\mathbb{E}\sum_{A\in\mathcal{A}_n}|A|^k I_A(t)s_l(t)=s_k(t)s_l(t).$$

(ii): By (2.1),

$$\mathbb{E}\,S_{k,l}(t)=\mathbb{E}\big(S_k(t)S_l(t)\big)-s_{k+l}(t),$$

and it remains to show that $\mathbb{E}\big(S_k(t)S_l(t)\big)\ge s_k(t)s_l(t)$, i.e., that $S_k(t)$ and $S_l(t)$ are positively correlated. This follows by Harris' inequality (a special case of the FKG inequality), since $S_k(t)$ and $S_l(t)$ are (by Lemma 2.1) increasing functions of the edge indicators of $\mathcal{G}(n,t)$, and these are independent. □

We use this first to find an upper bound for $s_k(t)$. Combining (3.4), (3.2) and Lemma 3.1(i), we find

$$s_k'(t)=\mathbb{E}\,V_k(t)\le\sum_{l=1}^{k-1}\frac{1}{2}\binom{k}{l}s_{l+1}(t)s_{k-l+1}(t).\tag{3.6}$$

The first cases are

$$s_2'(t)\le s_2(t)^2,\tag{3.7}$$
$$s_3'(t)\le 3s_2(t)s_3(t),\tag{3.8}$$
$$s_4'(t)\le 4s_2(t)s_4(t)+3s_3(t)^2.\tag{3.9}$$

Integrating (3.7), with the initial value $s_2(0)=n$, we find, e.g. via $(1/s_2(t))'\ge-1$ and thus $1/s_2(t)\ge 1/n-t$,

$$s_2(t)\le\frac{n}{1-nt},\qquad 0\le t<1/n.\tag{3.10}$$

Next, (3.8) and (3.10) yield $\big((1-nt)^3s_3(t)\big)'\le 0$ and thus, since $s_3(0)=n$,

$$s_3(t)\le\frac{n}{(1-nt)^3},\qquad 0\le t<1/n.\tag{3.11}$$

We can continue recursively and obtain the following bounds.

Lemma 3.2. *For every $k\ge 2$, there exists a constant C_k such that, for all n,*

$$\mathbb{E}\,S_k(t)=s_k(t)\le C_k\frac{n}{(1-nt)^{2k-3}},\qquad 0\le t<1/n.$$

Proof. We have proven this for $k = 2$ and 3. For $k \geq 4$ we use induction and assume that the lemma holds for smaller values of k; then (3.6) yields, for some constant C'_k, taking the terms $l = 1$ and $l = k - 1$ separately and using (3.10),

$$\begin{aligned} s'_k(t) &\leq k s_2(t) s_k(t) + \sum_{l=2}^{k-2} \frac{1}{2}\binom{k}{l} \frac{C_{l+1} n}{(1-nt)^{2l-1}} \frac{C_{k-l+1} n}{(1-nt)^{2k-2l-1}} \\ &\leq \frac{kn}{1-nt} s_k(t) + \frac{C'_k n^2}{(1-nt)^{2k-2}}. \end{aligned}$$

Hence, $\big((1-nt)^k s_k(t)\big)' \leq C'_k n^2 (1-nt)^{-(k-2)}$ and thus

$$\begin{aligned} (1-nt)^k s_k(t) &\leq n + \int_0^t \frac{C'_k n^2}{(1-nu)^{k-2}} \,\mathrm{d}u \leq n + \frac{C'_k n}{(k-3)(1-nt)^{k-3}} \\ &\leq \frac{C_k n}{(1-nt)^{k-3}}. \end{aligned}$$

□

We write the estimate in Lemma 3.2 as $s_k(t) = O(n(1-nt)^{3-2k})$ where, as in all similar estimates below, the implicit constant may depend on k (and later sometimes l), but not on n or t (in the given range $0 \leq t < 1/n$).

We can now use this upper bound in a more or less repetition of the same argument to obtain more precise estimates. By Lemmas 3.1 and 3.2, for $0 \leq t < 1/n$,

$$\mathbb{E}\, S_{k,l}(t) = s_k(t) s_l(t) + O(s_{k+l}(t)) = s_k(t) s_l(t) + O\Big(\frac{n}{(1-nt)^{2k+2l-3}}\Big).$$

Hence, (3.6) and (3.2) yield

$$s'_k(t) = \mathbb{E}\, V_k(t) = \sum_{l=1}^{k-1} \frac{1}{2}\binom{k}{l} s_{l+1}(t) s_{k-l+1}(t) + O\Big(\frac{n}{(1-nt)^{2k+1}}\Big). \tag{3.12}$$

The first cases are

$$s'_2(t) = s_2(t)^2 + O\big(n(1-nt)^{-5}\big), \tag{3.13}$$

$$s'_3(t) = 3 s_2(t) s_3(t) + O\big(n(1-nt)^{-7}\big), \tag{3.14}$$

$$s'_4(t) = 4 s_2(t) s_4(t) + 3 s_3(t)^2 + O\big(n(1-nt)^{-9}\big). \tag{3.15}$$

We first treat $s_2(t)$.

Theorem 3.3.

$$\mathbb{E}\, S_2(t) = s_2(t) = \frac{n}{1-nt}\Big(1 + O\Big(\frac{nt}{n(1-nt)^3}\Big)\Big), \qquad 0 \leq t < 1/n.$$

Proof. Let $T := \inf\{t : (1-nt) s_2(t) = n/2\}$. Since $f(t) := (1-nt) s_2(t)$ is continuous with $f(0) = n$ and $f(1/n) = 0$, then $0 < T < 1/n$ and for $0 \leq t \leq T$ we have $s_2(t) \geq \frac{1}{2} n/(1-nt)$ and thus, by (3.13),

$$\Big(\frac{1}{s_2(t)}\Big)' = -1 + O\Big(\frac{n}{(1-nt)^5 s_2(t)^2}\Big) = -1 + O\Big(\frac{1}{n(1-nt)^3}\Big).$$

This implies, recalling $s_2(0) = n$ and noting that $\int_0^t (1-nu)^{-3}\,\mathrm{d}u = O(t/(1-nt)^2)$ (which is, like similar integrals below, perhaps simplest seen by considering the cases $nt \le 1/2$ and $nt \ge 1/2$ separately),

$$\begin{aligned}\frac{1}{s_2(t)} &= \frac{1}{n} + \int_0^t \Big(\frac{1}{s_2(u)}\Big)'\,\mathrm{d}u = \frac{1}{n} - t + O\Big(\frac{t}{n(1-nt)^2}\Big)\\ &= \frac{1-nt}{n}\Big(1 + O\Big(\frac{nt}{n(1-nt)^3}\Big)\Big).\end{aligned} \tag{3.16}$$

Taking here $t = T$, we find $1 = O\big(1/(n(1-nT)^3)\big)$, and thus $n(1-nT)^3 = O(1)$ or $1 - nT = O(n^{-1/3})$. Choosing A large enough, we see that if $1-nt \ge An^{-1/3}$, then $t \le T$, and further the O term in (3.16) is, in absolute value, less than $1/2$. Thus (3.16) yields the result for $1 - nt \ge An^{-1/3}$. The result for $1 - nt < An^{-1/3}$ follows trivially from the bound (3.10). □

Theorem 3.3 proves (1.3) by the change of variable $t = -\log(1-p) = p + O(p^2)$ as discussed in Section 2, noting that the result is utterly trivial for $1 - np = O(n^{-1})$.

We continue with higher k.

Theorem 3.4. *The following holds for $0 \le t < 1/n$.*

$$\mathbb{E}\, S_3(t) = s_3(t) = \frac{n}{(1-nt)^3}\Big(1 + O\Big(\frac{nt}{n(1-nt)^3}\Big)\Big),$$

$$\mathbb{E}\, S_4(t) = s_4(t) = \frac{n\big(3-2(1-nt)\big)}{(1-nt)^5}\Big(1 + O\Big(\frac{nt}{n(1-nt)^3}\Big)\Big).$$

More generally, for every $k \ge 2$ there exists a polynomial p_k of degree $2k-3$ such that

$$\begin{aligned}\mathbb{E}\, S_k(t) = s_k(t) &= np_k\Big(\frac{1}{1-nt}\Big) + O\Big(\frac{nt}{(1-nt)^{2k}}\Big)\\ &= np_k\Big(\frac{1}{1-nt}\Big)\Big(1 + O\Big(\frac{nt}{n(1-nt)^3}\Big)\Big).\end{aligned} \tag{3.17}$$

We have $p_2(x) = x$, $p_3(x) = x^3$, $p_4(x) = 3x^5 - 2x^4$. In general, for $k \ge 3$, $p_k(x) = x^k q_k(x)$ for a polynomial $q_k(x)$ of degree $k - 3$ that is recursively defined by $q_k(1) = 1$ and

$$q_k'(x) = \frac{1}{2}\sum_{l=2}^{k-2}\binom{k}{l} q_{l+1}(x)q_{k-l+1}(x), \qquad k \ge 3. \tag{3.18}$$

Equivalently, $p_k(1) = 1$ and

$$p_k'(x) = \frac{1}{2x^2}\sum_{l=1}^{k-1}\binom{k}{l} p_{l+1}(x)p_{k-l+1}(x), \qquad k \ge 2. \tag{3.19}$$

A probabilistic interpretation of $p_k(x)$ and a simpler recursion formula are given in Section 7. The polynomials p_k for small k are given in Table 1. It follows from the recursion (3.18) that p_k has degree $2k - 3$ with a positive leading term; since further q_k and p_k are non-decreasing on $[1,\infty)$, for example by (3.18) again, and thus strictly positive there, it follows that $p_k(x) \asymp x^{2k-3}$ for $x \geq 1$. This shows that the two forms of (3.17), with the error term written as an absolute or a relative error, are equivalent.

$$
\begin{aligned}
p_2(x) &= x,\\
p_3(x) &= x^3,\\
p_4(x) &= 3\,x^5 - 2\,x^4,\\
p_5(x) &= 15\,x^7 - 20\,x^6 + 6\,x^5,\\
p_6(x) &= 105\,x^9 - 210\,x^8 + 130\,x^7 - 24\,x^6,\\
p_7(x) &= 945\,x^{11} - 2520\,x^{10} + 2380\,x^9 - 924\,x^8 + 120\,x^7,\\
p_8(x) &= 10395\,x^{13} - 34650\,x^{12} + 44100\,x^{11} - 26432\,x^{10} + 7308\,x^9 - 720\,x^8.
\end{aligned}
$$

TABLE 1. The polynomials $p_k(x)$ for $k \leq 8$.

Proof. We have shown the result for $k=2$, with $p_2(x) = x$ which satisfies (3.19). For larger k, we use induction and assume that (3.17) is true for smaller values of k. Then, by (3.12), taking the terms $l=1$ and $l=k-1$ separately, and (3.18),

$$
\begin{aligned}
s_k'(t) &= ks_2(t)s_k(t) + \sum_{l=2}^{k-2}\frac{1}{2}\binom{k}{l}s_{l+1}(t)s_{k-l+1}(t) + O\Bigl(\frac{n}{(1-nt)^{2k+1}}\Bigr)\\
&= \frac{kn}{1-nt}s_k(t) + n^2\sum_{l=2}^{k-2}\frac{1}{2}\binom{k}{l}p_{l+1}\Bigl(\frac{1}{1-nt}\Bigr)p_{k-l+1}\Bigl(\frac{1}{1-nt}\Bigr)\\
&\qquad + O\Bigl(\frac{n}{(1-nt)^{2k+1}}\Bigr)\\
&= \frac{kn}{1-nt}s_k(t) + \frac{n^2}{(1-nt)^{k+2}}q_k'\Bigl(\frac{1}{1-nt}\Bigr) + O\Bigl(\frac{n}{(1-nt)^{2k+1}}\Bigr).
\end{aligned}
$$

Thus,

$$
\begin{aligned}
\bigl((1-nt)^k s_k(t)\bigr)' &= \frac{n^2}{(1-nt)^2}q_k'\Bigl(\frac{1}{1-nt}\Bigr) + O\Bigl(\frac{n}{(1-nt)^{k+1}}\Bigr).\\
&= n\frac{\mathrm{d}}{\mathrm{d}t}q_k\Bigl(\frac{1}{1-nt}\Bigr) + O\Bigl(\frac{n}{(1-nt)^{k+1}}\Bigr).
\end{aligned}
$$

The result follows by integration, recalling that $s_k(0) = n$. □

4. The variance

Theorem 4.1. *For every $k \geq 2$, $\operatorname{Var}(S_k(t)) \leq s_{2k}(t)$. Hence,*

$$
\operatorname{Var}(S_k(t)) = O\bigl(n(1-nt)^{-(4k-3)}\bigr), \qquad 0 \leq t < 1/n.
$$

Proof. By (2.1) and Lemma 3.1(i),

$$\mathbb{E}(S_k(t)^2) = \mathbb{E}\, S_{k,k}(t) + \mathbb{E}\, S_{2k}(t) \le (\mathbb{E}\, S_k(t))^2 + s_{2k}(t).$$

The final estimate follows by Lemma 3.2. □

A more precise result will be given in Section 6. This will show that the bound in Theorem 4.1 is of the right order as long as $1 - nt \gg n^{-1/3}$.

Corollary 4.2. *If* $1 - nt \gg n^{-1/3}$*, then* $S_k(t) \sim_{\mathrm{p}} np_k\Bigl(\frac{1}{1-nt}\Bigr)$ *for every* $k \ge 2$*.*

Proof. By Theorem 3.4, $\mathbb{E}\, S_k(t) \sim np_k\bigl(\frac{1}{1-nt}\bigr)$. Further, Theorems 4.1 and 3.4 show that

$$\frac{\operatorname{Var}(S_k(t))}{(\mathbb{E}\, S_k(t))^2} = O\Bigl(\frac{n(1-nt)^{3-4k}}{n^2(1-nt)^{6-4k}}\Bigr) = O\Bigl(\frac{1}{n(1-nt)^3}\Bigr) = o(1),$$

and the result follows by Chebyshev's inequality. □

Proof of Theorem 1.1. As remarked above, (1.3) follows from Theorem 3.3. Similarly, the case $k = 2$ of Theorem 4.1 yields (1.4). Together, these estimates yield (1.5) for $1 - np \ge n^{-1/3}$. For $1/n \le 1 - np < n^{-1/3}$, (1.5) follows from the estimate $\mathbb{E}\,\chi(G(n,p)) = O(1/(1-np))$ one obtains from Lemma 3.2 (or (3.10)); for $0 < 1 - np < 1/n$, (1.5) follows from the trivial $\chi(G(n,p)) \le n$. □

Proof of Corollary 1.2. Let $A > 0$ be so large that the O term in (1.3) is $\le 1/2$ for $1 - np \ge An^{-1/3}$. Then (1.3) yields, for $np \le 1 - An^{-1/3}$,

$$\frac{1}{\mathbb{E}\,\chi(G(n,p))} = (1-np)\Bigl(1 + O\Bigl(\frac{1}{n(1-np)^3}\Bigr)\Bigr) = 1 - np + O\bigl(n^{-1/3}\bigr),$$

which shows (1.6) for these p. In particular, for $np = 1 - An^{-1/3}$ we find $1/\mathbb{E}\,\chi(G(n,p)) = O(n^{-1/3})$. This, and thus (1.6), then holds for all larger p too by monotonicity (Lemma 2.1).

The proof of (1.7) is similar, using (1.5). □

5. Asymptotic normality

The quadratic variation of the martingale $M_k(t)$ is

$$[M_k, M_k]_t := \sum_{0<u\le t} \Delta M_k(u)^2 = \sum_{0<u\le t} \Delta S_k(u)^2,$$

where $\Delta X(s) := X(s) - X(s-)$ denotes the jump (if any) of a process X at s. (This formula holds because M_k is a martingale with paths of finite variation and $M_k(0) = 0$; see e.g. [8] for a definition for general (semi)martingales.)

Using (3.1), we find, in analogy with (3.2), that $[M_k, M_k]_t$ has drift

$$W_k(t) := \sum_{i\neq j} \frac{1}{2}|\mathcal{C}_i||\mathcal{C}_j| \left(\sum_{l=1}^{k-1} \binom{k}{l} |\mathcal{C}_i|^l |\mathcal{C}_j|^{k-l} \right)^2$$
$$= \sum_{l=1}^{k-1} \sum_{m=1}^{k-1} \frac{1}{2} \binom{k}{l} \binom{k}{m} S_{l+m+1,\, 2k+1-l-m}(t); \tag{5.1}$$

i.e., $[M_k, M_k]_t - \int_0^t W_k(u)\,\mathrm{d}u$ is a martingale.

It turns out to be advantageous to work with a slightly different martingale. In order to cancel some terms later on, we multiply $S_k(t)$ by $(1-nt)^k$ (cf. the proof of Theorem 3.4 where we did the same with the expectation in order to simplify the differential equation); we thus define

$$\tilde{S}_k(t) := (1-nt)^k S_k(t), \tag{5.2}$$

which (by a simple instance of Ito's formula) has the drift

$$\tilde{V}_k(t) := (1-nt)^k V_k(t) - kn(1-nt)^{k-1} S_k(t). \tag{5.3}$$

Thus,

$$\tilde{M}_k(t) := \tilde{S}_k(t) - n - \int_0^t \tilde{V}_k(u)\,\mathrm{d}u \tag{5.4}$$

is a martingale with $\tilde{M}_k(0) = 0$. The quadratic variation is

$$[\tilde{M}_k, \tilde{M}_k]_t := \sum_{0<u\leq t} \Delta\tilde{M}_k(u)^2 = \sum_{0<u\leq t} \Delta\tilde{S}_k(u)^2 = \sum_{0<u\leq t} (1-nu)^{2k} \Delta S_k(u)^2.$$

This has drift

$$\tilde{W}_k(t) := (1-nt)^{2k} W_k(t), \tag{5.5}$$

and thus

$$\tilde{\tilde{M}}_k(t) := [\tilde{M}_k, \tilde{M}_k]_t - \int_0^t \tilde{W}_k(u)\,\mathrm{d}u \tag{5.6}$$

is another martingale with $\tilde{\tilde{M}}_k(0) = 0$.

We repeat the argument and find that $\tilde{\tilde{M}}_k$ has quadratic variation

$$[\tilde{\tilde{M}}_k, \tilde{\tilde{M}}_k]_t := \sum_{0<u\leq t} \Delta\tilde{\tilde{M}}_k(u)^2 = \sum_{0<u\leq t} (\Delta[\tilde{M}_k, \tilde{M}_k]_u)^2 = \sum_{0<u\leq t} \Delta\tilde{M}_k(u)^4$$
$$= \sum_{0<u\leq t} (1-nu)^{4k} \Delta S_k(u)^4,$$

which has drift, in analogy with (3.2) and (5.1),

$$\tilde{\tilde{W}}_k(t) := (1-nt)^{4k} \sum_{i\neq j} \frac{1}{2}|\mathcal{C}_i||\mathcal{C}_j| \left(\sum_{l=1}^{k-1} \binom{k}{l} |\mathcal{C}_i|^l |\mathcal{C}_j|^{k-l} \right)^4$$
$$= (1-nt)^{4k} \sum_{l_1,l_2,l_3,l_4=1}^{k-1} \frac{1}{2} \prod_{i=1}^{4} \binom{k}{l_i} \cdot S_{\sum_i l_i+1,\, 4k+1-\sum_i l_i}(t); \tag{5.7}$$

thus, $[\tilde{\tilde{M}}_k, \tilde{\tilde{M}}_k]_t - \int_0^t \tilde{\tilde{W}}_k(u)\,\mathrm{d}u$ is yet another martingale which starts at 0.

Assume in the remainder of the section that $1 - nt \geq n^{-1/3}$, i.e.

$$0 \leq t \leq n^{-1} - n^{-4/3}. \tag{5.8}$$

(Although some estimates require only $0 \leq t < 1/n$.) By Lemmas 3.1(i) and 3.2, for any $k, l \geq 2$,

$$\mathbb{E}\, S_{k,l}(t) = O\left(\frac{n^2}{(1-nt)^{2k+2l-6}}\right).$$

Hence, (5.7) yields

$$\mathbb{E}\, \tilde{\tilde{W}}_k(t) = O\Big((1-nt)^{4k} \frac{n^2}{(1-nt)^{8k-2}}\Big) = O\left(\frac{n^2}{(1-nt)^{4k-2}}\right).$$

Since $\operatorname{Var}(M(t)) = \mathbb{E}\, M(t)^2 = \mathbb{E}[M, M]_t$ for every square integrable martingale with $M(0) = 0$,

$$\begin{aligned}\mathbb{E}\big(\tilde{\tilde{M}}_k(t)\big)^2 = \mathbb{E}[\tilde{\tilde{M}}_k, \tilde{\tilde{M}}_k]_t = \mathbb{E} \int_0^t \tilde{\tilde{W}}_k(u)\,\mathrm{d}u = O\left(\int_0^t \frac{n^2}{(1-nu)^{4k-2}}\,\mathrm{d}u\right) \\ = O\left(\frac{n^2 t}{(1-nt)^{4k-3}}\right). \end{aligned} \tag{5.9}$$

We define, subtracting by (3.17) an approximation to the mean,

$$Y_k(t) := S_k(t) - np_k\Big(\frac{1}{1-nt}\Big). \tag{5.10}$$

Lemma 5.1. *For every $k \geq 2$ and $1 - nt \geq n^{-1/3}$,*

$$Y_k(t) = O_{L^2}\Big(\frac{n^{1/2}}{(1-nt)^{2k-3/2}}\Big).$$

Proof.

$$\|Y_k(t)\|_{L^2}^2 = \operatorname{Var} S_k(t) + \Big|\mathbb{E}\, S_k(t) - np_k\Big(\frac{1}{1-nt}\Big)\Big|^2,$$

and the result follows by Theorems 4.1 and 3.4, using $n(1-nt)^3 \geq 1$. □

Lemma 5.2. *For every $k, l \geq 2$ and $1 - nt \geq n^{-1/3}$,*

$$S_{k,l}(t) = n^2 p_k\Big(\frac{1}{1-nt}\Big) p_l\Big(\frac{1}{1-nt}\Big) + O_{L^1}\left(\frac{n^{3/2}}{(1-nt)^{2k+2l-9/2}}\right).$$

Proof. By (2.1) and (5.10),

$$S_{k,l}(t) = \Big(np_k\Big(\frac{1}{1-nt}\Big) + Y_k(t)\Big)\Big(np_l\Big(\frac{1}{1-nt}\Big) + Y_l(t)\Big) - S_{k+l}(t)$$

and thus, using Lemmas 5.1 and 3.2 and the Cauchy–Schwarz inequality,

$$\Big\| S_{k,l}(t) - n^2 p_k\Big(\frac{1}{1-nt}\Big) p_l\Big(\frac{1}{1-nt}\Big)\Big\|_{L^1}$$

$$= O\Big(n^{3/2}(1-nt)^{-2k-2l+9/2} + n(1-nt)^{-2k-2l+3}\Big),$$

which yields the result by our assumption $n(1-nt)^3 \geq 1$. □

Lemma 5.3. *For every $k \geq 2$, there exists a polynomial $\tilde{P}_k$ of degree $2k-2$ given by*

$$\begin{aligned}\tilde{P}_k(x) &= x^{-2k}\sum_{l=1}^{k-1}\sum_{m=1}^{k-1}\frac{1}{2}\binom{k}{l}\binom{k}{m}p_{l+m+1}(x)p_{2k+1-l-m}(x)\\ &= x^2\sum_{l=1}^{k-1}\sum_{m=1}^{k-1}\frac{1}{2}\binom{k}{l}\binom{k}{m}q_{l+m+1}(x)q_{2k+1-l-m}(x) \end{aligned}\tag{5.11}$$

such that, for $1-nt \geq n^{-1/3}$,

$$\tilde{W}_k(t) = n^2\tilde{P}_k\Big(\frac{1}{1-nt}\Big) + O_{L^1}\Big(\frac{n^{3/2}}{(1-nt)^{2k-1/2}}\Big).$$

Proof. An immediate consequence of (5.5), (5.1) and Lemma 5.2. □

Lemma 5.4. (i) *For every $k \geq 2$, there exists a polynomial $\tilde{Q}_k$ of degree $2k-3$ given by*

$$\tilde{Q}_k'(x) = x^{-2}\tilde{P}_k(x),\tag{5.12}$$

with $\tilde{Q}_k(1) = 0$, such that, for $1-nt \geq n^{-1/3}$,

$$[\tilde{M}_k,\tilde{M}_k]_t = n\tilde{Q}_k\Big(\frac{1}{1-nt}\Big) + O_{L^1}\Big(\frac{nt^{1/2}}{(1-nt)^{2k-3/2}}\Big).\tag{5.13}$$

(ii) *If $n^2t \to \infty$ and $n(1-nt)^3 \to \infty$, then*

$$[\tilde{M}_k,\tilde{M}_k]_t = n\tilde{Q}_k\Big(\frac{1}{1-nt}\Big)\big(1+o_{L^1}(1)\big) = n\tilde{Q}_k\Big(\frac{1}{1-nt}\Big)\big(1+o_{\mathrm{p}}(1)\big).$$

Proof. (i): By (5.6), Lemma 5.3 and (5.9),

$$\begin{aligned}[\tilde{M}_k,\tilde{M}_k]_t &= \int_0^t \tilde{W}_k(u)\,\mathrm{d}u + \tilde{\tilde{M}}_k(t)\\ &= \int_0^t n^2\tilde{P}_k\Big(\frac{1}{1-nu}\Big)\,\mathrm{d}u + O_{L^1}\Big(\frac{n^{3/2}t + nt^{1/2}}{(1-nt)^{2k-3/2}}\Big)\end{aligned}$$

and (5.13) follows, noting that $n^{3/2}t \leq nt^{1/2}$.

(ii): By (5.12), $\tilde{Q}_k$ is increasing for $x > 1$, and thus non-zero, and it follows that $\tilde{Q}_k\big(\frac{1}{1-nt}\big) \asymp nt(1-nt)^{3-2k}$. It remains only to verify that $nt^{1/2}(1-nt)^{3/2} = o(n^2t(1-nt)^3)$, which is obvious under our conditions if we consider $nt \leq 1/2$ and $nt \geq 1/2$ separately. □

We will use the following general result based on [8]; see [11, Proposition 9.1] for a detailed proof. (See also [9], [10] and [12] for similar versions.)

Proposition 5.5. *Assume that for each n, $M^{(n)}(x)$ is a martingale on $[0,1]$ with $M^{(n)}(0)=0$, and that $\sigma^2(x)$, $x\in[0,1]$, is a (non-random) continuous function such that for every fixed $x\in[0,1]$,*

$$[M^{(n)},M^{(n)}]_x \xrightarrow{\mathrm{p}} \sigma^2(x) \quad \text{as } n\to\infty, \tag{5.14}$$

$$\sup_n \mathbb{E}[M^{(n)},M^{(n)}]_x < \infty. \tag{5.15}$$

Then $M^{(n)} \xrightarrow{\mathrm{d}} M$ as $n\to\infty$, in $D[0,1]$, where M is a continuous Gaussian martingale with $\mathbb{E}\,M(x)=0$ and covariances

$$\mathbb{E}\big(M(x)M(y)\big)=\sigma^2(x), \quad 0\le x\le y\le 1.$$

In particular, $M^{(n)}(1) \xrightarrow{\mathrm{d}} N(0,\sigma^2(1))$.

Remark 5.6. Proposition 5.5 extends to vector-valued martingales; see the versions in [11; 12].

Remark 5.7. The versions in [11; 12] are for martingales on $[0,\infty)$; it is easily seen that the versions are equivalent by stopping the martingales at a fixed time; moreover, by a (deterministic) change of time, we may replace $[0,1]$ by any closed or half-open interval $[a,b]$ or $[a,b)\subseteq[-\infty,\infty]$.

Further, (5.15) is equivalent to $\sup_n \mathbb{E}\,|M^{(n)}(x)|^2<\infty$, the form used in e.g. [11].

Lemma 5.8. *If $n^2t\to\infty$ and $n(1-nt)^3\to\infty$, then*

$$\tilde{M}_k(t) \sim \operatorname{AsN}\Big(0, n\tilde{Q}_k\Big(\frac{1}{1-nt}\Big)\Big).$$

Proof. In order to apply Proposition 5.5, we have to change the time scale to a fixed interval so that the quadratic variation converges. By considering subsequences, we may assume that $nt\to a$ for some $a\in[0,1]$. We then define $M^{(n)}(x)$ for $x\in[0,1]$ as follows.

(i) If $0<a<1$, we let $M^{(n)}(x) := (n^2t)^{-1/2}\tilde{M}_k(xt)$, and see that Lemma 5.4(ii) implies (5.14) with $\sigma^2(x)=a^{-1}\tilde{Q}_k(1/(1-ax))$.

(ii) If $a=0$, we define $M^{(n)}(x)$ in the same way, and find now that Lemma 5.4(ii) implies (5.14) with $\sigma^2(x)=x\tilde{Q}_k'(1)$.

(iii) If $a=1$, we let $M^{(n)}(x):=n^{-1/2}(1-nt)^{k-3/2}\tilde{M}_k(t_n(x))$, where

$$t_n(x) := \begin{cases} 0, & x\le 1-nt, \\ \frac{1}{n}\big(1-\frac{1-nt}{x}\big), & x\ge 1-nt; \end{cases}$$

thus $1-nt_n(x)=\min((1-nt)/x,1)$. In this case Lemma 5.4(ii) implies (5.14) with $\sigma^2(x)=c_kx^{2k-3}$, where $c_k>0$ is the leading coefficient in $\tilde{Q}_k$.

In all cases, the same calculation yields also (5.15), because the factor $1+o_{L^1}(1)$ in Lemma 5.4 is $O_{L^1}(1)$. The result follows from the final statement in Proposition 5.5. □

Let us now consider the case $k = 2$.

Theorem 5.9. *If $n^2t \to \infty$ and $n(1-nt)^3 \to \infty$, then*

$$S_2(t) \sim \operatorname{AsN}\left(\frac{n}{1-nt}, \frac{2n^2t}{(1-nt)^5}\right).$$

Proof. By (3.2) and (2.1), $V_2(t) = S_{2,2}(t) = S_2(t)^2 - S_4(t)$, and thus (5.3) yields

$$\begin{aligned}\tilde{V}_2(t) &= (1-nt)^2V_2(t) - 2n(1-nt)S_2(t)\\ &= \bigl((1-nt)S_2(t)-n\bigr)^2 - n^2 - (1-nt)^2S_4(t).\end{aligned}$$

By Theorems 4.1 and 3.3,

$$\begin{aligned}\mathbb{E}\bigl((1-nt)S_2(t)-n\bigr)^2 &= (1-nt)^2\operatorname{Var}(S_2(t)) + \bigl((1-nt)\,\mathbb{E}\,S_2(t)-n\bigr)^2\\ &= O\Bigl(\frac{n}{(1-nt)^3}\Bigr) + O\Bigl(\frac{1}{(1-nt)^6}\Bigr) = O\Bigl(\frac{n}{(1-nt)^3}\Bigr).\end{aligned}$$

By Lemma 3.2, $\|(1-nt)^2S_4(t)\|_{L^1}$ is also estimated by $O(n(1-nt)^{-3})$. Hence,

$$\tilde{V}_2(t) = -n^2 + O_{L^1}\Bigl(\frac{n}{(1-nt)^3}\Bigr).$$

We now obtain from (5.4)

$$\tilde{S}_2(t) = \tilde{M}_2(t)+n+\int_0^t \tilde{V}_2(u)\,\mathrm{d}u = \tilde{M}_2(t)+n-n^2t+O_{L^1}\Bigl(\frac{nt}{(1-nt)^2}\Bigr). \tag{5.16}$$

For $k = 2$, (5.11) and (5.12) yield $\tilde{P}_2(x) = 2x^2q_3(x)^2 = 2x^2$ and $\tilde{Q}_2(x) = 2(x-1)$. Hence Lemma 5.8 yields

$$\tilde{M}_2(t) \sim \operatorname{AsN}\Bigl(0, \frac{2n^2t}{1-nt}\Bigr). \tag{5.17}$$

It is easily verified that $\frac{nt}{(1-nt)^2} \ll \bigl(\frac{n^2t}{1-nt}\bigr)^{1/2}$. Hence, (5.16) and (5.17) yield

$$\tilde{S}_2(t) \sim \operatorname{AsN}\Bigl(n(1-nt), \frac{2n^2t}{1-nt}\Bigr).$$

Recalling the definition $\tilde{S}_2(t) = (1-nt)^2S_2(t)$, we obtain the assertion. □

Proof of Theorem 1.3, asymptotic normality. Immediate from Theorem 5.9 by our usual relation $\chi(G(n,p)) = n^{-1}S_2(-\log(1-p))$. □

For $k > 2$, the argument is more involved, and we will be somewhat sketchy. We assume $1 - nt \gg n^{-1/3}$ and consider first $k = 3$. By (3.2) and

(2.1), $V_3(t) = 3S_{2,3}(t) = 3S_2(t)S_3(t) - 3S_5(t)$, and thus (5.3) yields, using (5.10), Lemmas 3.2 and 5.1 and the Cauchy–Schwarz inequality,

$$\begin{aligned}\tilde{V}_3(t) &= (1-nt)^3 V_3(t) - 3n(1-nt)^2 S_3(t)\\ &= 3(1-nt)^3\Big(S_2(t) - \frac{n}{1-nt}\Big)S_3(t) - 3(1-nt)^3 S_5(t)\\ &= 3(1-nt)^3 Y_2(t)S_3(t) - 3(1-nt)^3 S_5(t)\\ &= 3n(1-nt)^3 p_3\Big(\frac{1}{1-nt}\Big)Y_2(t) + O_{L^1}\Big(\frac{n}{(1-nt)^4}\Big).\end{aligned}$$

Hence, by (5.4), recalling $p_3(x) = x^3$,

$$\tilde{S}_3(t) = \tilde{M}_3(t) + n + 3n\int_0^t Y_2(u)\,\mathrm{d}u + O_{L^1}\Big(\frac{nt}{(1-nt)^3}\Big), \tag{5.18}$$

where we may ignore the O term but not the integral, unlike the corresponding expression (5.16) for $k=2$. We find from (5.16)

$$Y_2(u) = (1-nu)^{-2}\big(\tilde{S}_2(u) - n(1-nu)\big) = (1-nu)^{-2}\tilde{M}_2(u) + O_{L^1}\Big(\frac{nu}{(1-nu)^4}\Big).$$

Hence, (5.18) yields

$$\tilde{S}_3(t) - n = \tilde{M}_3(t) + 3n\int_0^t (1-nu)^{-2}\tilde{M}_2(u)\,\mathrm{d}u + O_{L^1}\Big(\frac{nt}{(1-nt)^3}\Big). \tag{5.19}$$

We applied above Proposition 5.5 to $\tilde{M}_2$, but we only used the final statement (with $x=1$) in order to obtain Lemma 5.8. Now we use the full process statement of Proposition 5.5, from which we conclude (after a change of variables as in the proof of Lemma 5.8) that $\int_0^t (1-nu)^{-2}\tilde{M}_2(u)\,\mathrm{d}u$ also has an asymptotic normal distribution. Moreover, by the vector-valued version of Proposition 5.5 mentioned in Remark 5.6, the argument in the proof of Lemma 5.8 yields joint asymptotic normality of the processes $\tilde{M}_k$ for different k; this uses a straightforward extension of Lemma 5.4 to quadratic covariations $[\tilde{M}_{k_1}, \tilde{M}_{k_2}]_t$. As a result, the first two terms on the right hand side of (5.19) are jointly normal, and the O term can be ignored. (The right normalization here is, cf. Theorem 4.1, to divide by $nt^{1/2}(1-nt)^{-3/2}$.) A careful but rather tedious (even with `Maple`) calculation of the involved covariances yields $\tilde{S}_3(t) \sim \mathrm{AsN}(n, nQ_3(1/(1-nt))$ with $Q_3(x) = 96\,x^3 - 198\,x^2 + 126\,x - 24$. Hence, with $\hat{P}_3(x) = x^6 Q_3(x) = 96\,x^9 - 198\,x^8 + 126\,x^7 - 24x^6$,

$$S_3(t) \sim \mathrm{AsN}\Big(\frac{n}{(1-nt)^3}, n\hat{P}_3\Big(\frac{1}{1-nt}\Big)\Big). \tag{5.20}$$

We can argue in the same way for $k > 3$ too. It follows from (5.3), (3.2), (2.1), (5.10) and Lemmas 3.2 and 5.1 that

$$\tilde{V}_k(t) = n(1-nt)^k \sum_{j=2}^{k-1} \binom{k}{j-1} p_{k+2-j}\Big(\frac{1}{1-nt}\Big) Y_j(t)$$

$$+\frac{1}{2}n^2(1-nt)^k\sum_{l=2}^{k-2}\binom{k}{l}p_{l+1}\Bigl(\frac{1}{1-nt}\Bigr)p_{k+1-l}\Bigl(\frac{1}{1-nt}\Bigr) \\ +O_{L^1}\Bigl(\frac{nt}{(1-nt)^{k+1}}\Bigr).$$

which, using (5.10), (5.2), (5.4) and (3.18), leads to the recursive formula (for all $k\geq 2$, cf. (5.16) and (5.18) for $k=2$ and 3)

$$Y_k(t)=(1-nt)^{-k}\tilde{M}_k(t)+n(1-nt)^{-k}\sum_{j=2}^{k-1}\binom{k}{j-1} \\ \times\int_0^t(1-nu)^k p_{k+2-j}\Bigl(\frac{1}{1-nu}\Bigr)Y_j(u)\,\mathrm{d}u+O_{L^1}\Bigl(\frac{nt}{(1-nt)^{2k}}\Bigr). \quad (5.21)$$

This yields, by induction, cf. (5.16) and (5.19) for $k=2$ and 3,

$$Y_k(t)=(1-nt)^{-k}\tilde{M}_k(t)+n(1-nt)^{-k}\sum_{j=2}^{k-1}\int_0^t\bar{P}_{k,j}\Bigl(\frac{1}{1-nu}\Bigr)\tilde{M}_j(u)\,\mathrm{d}u \\ +O_{L^1}\Bigl(\frac{nt}{(1-nt)^{2k}}\Bigr), \quad (5.22)$$

for some polynomials $\bar{P}_{k,j}(x)$ having degree at most $k+1-j$ and no terms of degree ≤ 1. The asymptotic joint normality of the processes $\tilde{M}_k$ (with a careful count of the degrees of the involved polynomials) now shows the following extension of Theorem 5.9 and (5.20).

Theorem 5.10. *There exist polynomials $\hat{P}_k(x)$ of degree (at most) $4k-3$ such that if $n^2t\to\infty$ and $1-nt\gg n^{-1/3}$, then*

$$S_k(t)\sim\operatorname{AsN}\Bigl(np_k\Bigl(\frac{1}{1-nt}\Bigr),n\hat{P}_k\Bigl(\frac{1}{1-nt}\Bigr)\Bigr),\qquad k\geq 2.$$

Furthermore, this holds jointly for all $k\geq 2$, with asymptotic covariances given by polynomials $\hat{P}_{k,l}(x)$ of degree (at most) $2k+2l-3$.

We have, for example, $\hat{P}_2(x)=2x^5-2x^4$, $\hat{P}_3(x)=96\,x^9-198\,x^8+126\,x^7-24x^6$ (as said above), and $\hat{P}_{2,3}(x)=12\,x^7-18\,x^6+6\,x^5$. To find $\hat{P}_k=\hat{P}_{k,k}$ and $\hat{P}_{k,l}$ in general by this method seems quite difficult, although it is in principle possible using computer algebra. In the next section we will, by a different method, find the asymptotics of the covariances of the variables $S_k(t)$. It is natural to conjecture that these coincide with the asymptotic covariances in Theorem 5.10, which by general probability theory, e.g. [7, Theorem 5.5.9], is equivalent to uniform square integrability of each of the standardized variables $(S_k(t)-\mathbb{E}\,S_k(t))/\operatorname{Var}(S_k(t))^{1/2}$ as $n\to\infty$. This is very plausible (and thus verified for $k=2$ and 3 by our calculations of $\hat{P}_2$ and $\hat{P}_3$), but we have so far been unable to verify it in general, and we leave

$$
\begin{aligned}
P_{2,2}(x) &= 2\,x^5 - 2\,x^4\\
P_{3,3}(x) &= 96\,x^9 - 198\,x^8 + 126\,x^7 - 24\,x^6\\
P_{4,4}(x) &= 10170\,x^{13} - 34050\,x^{12} + 43520\,x^{11} - 26192\,x^{10} + 7272\,x^9 - 720\,x^8\\
P_{3,2}(x) &= 12\,x^7 - 18\,x^6 + 6\,x^5\\
P_{4,2}(x) &= 90\,x^9 - 190\,x^8 + 124\,x^7 - 24\,x^6\\
P_{4,3}(x) &= 900\,x^{11} - 2430\,x^{10} + 2322\,x^9 - 912\,x^8 + 120\,x^7
\end{aligned}
$$

TABLE 2. The polynomials $P_{k,l}(x)$ for $k, l \le 4$.

this as an open problem and conjecture. (It would suffice to consider the case $nt \le 1/2$, say, and show for example that then $\mathbb{E}\,|S_k(t) - \mathbb{E}\,S_k(t)|^4 = O(n^2)$.)

Conjecture 5.11. *$\hat{P}_{k,l}$ equals the polynomial $P_{k,l}$ defined in* (6.1)*.*

Remark 5.12. The purpose of introducing $\tilde{S}_k$ in (5.2) is that if we argued directly with S_k and M_k, we would obtain an equation similar to (5.21), but with $Y_k(u)$ in one of the integrals on the right hand side. Thus, to derive the asymptotic normality of $Y_(t)$ from the asymptotic normality of the processes M_k, we would have to invert a Volterra equation (also for $k = 2$). This is effectively what we do by introducing $\tilde{S}_k$.

6. The variance again

In Theorem 4.1 we gave a simple upper bound for the variance $S_k(t)$. We shall now, using a more involved argument, find the precise asymptotics.

Theorem 6.1. *For every $k, l \ge 2$ and $0 \le t < 1/n$,*

$$\operatorname{Cov}\bigl(S_k(t), S_l(t)\bigr) = nP_{k,l}\Bigl(\frac{1}{1-nt}\Bigr) + O\Bigl(\frac{nt}{(1-nt)^{2k+2l}}\Bigr),$$

where $P_{k,l}$ is a polynomial of degree $2k + 2l - 3$ given by

$$P_{k,l}(x) = p_{k+l}(x) - \frac{p_{k+1}(x)p_{l+1}(x)}{x}. \tag{6.1}$$

Some polynomials $P_{k,l}$ are given in Table 2. In particular, $P_{2,2}(1/y) = 2(1-y)/y^5$ and thus

$$\operatorname{Var}(S_2(t)) = \frac{2n^2t}{(1-nt)^5}\biggl(1 + O\Bigl(\frac{1}{n(1-nt)^3}\Bigr)\biggr). \tag{6.2}$$

For $1 - nt < n^{-1/3}$, Theorem 6.1 is a trivial (and uninteresting) consequence of Theorem 4.1 and the Cauchy–Schwarz inequality, so we assume in the sequel that $1 - nt \ge n^{-1/3}$. We precede the proof by several lemmas;

we begin by defining, extending (2.1),

$$S_{k_1,\dots,k_m}(G) := \sum_{i_1,\dots,i_m}^{*} |\mathcal{C}_{i_1}|^{k_1} \cdots |\mathcal{C}_{i_m}|^{k_m},$$

where $\sum^*$ denotes the sum over distinct indices only. For $G = \mathcal{G}(n,t)$ we write $S_{k_1,\dots,k_m}(t)$ and have the following estimate, cf. Lemma 3.1.

Lemma 6.2. *For each $m \geq 1$, $k_1, \dots, k_m \geq 2$ and $1 - nt \geq n^{-1/3}$,*

$$\mathbb{E}\, S_{k_1,\dots,k_m}(t) = n^m p_{k_1}\Big(\frac{1}{1-nt}\Big) \cdots p_{k_m}\Big(\frac{1}{1-nt}\Big)\Big(1 + O\Big(\frac{1}{n(1-nt)^3}\Big)\Big).$$

Proof. First, the argument in the proof of Lemma 3.1(i) extends to show that $\mathbb{E}\, S_{k_1,\dots,k_m}(t) \leq \mathbb{E}\, S_{k_1,\dots,k_{m-1}}(t)\, \mathbb{E}\, S_k(t)$. Thus, by induction, $\mathbb{E}\, S_{k_1,\dots,k_m}(t) \leq \prod_{i=1}^m \mathbb{E}\, S_{k_i}(t)$, which by Theorem 3.4 yields an upper bound of the required type.

Secondly, we have, for any graph G,

$$\prod_{i=1}^m S_{k_i}(G) = \sum_{i_1,\dots,i_m} |\mathcal{C}_{i_1}|^{k_1} \cdots |\mathcal{C}_{i_m}|^{k_m}. \tag{6.3}$$

Each vector $(i_1, \dots, i_m)$ defines a partition $\pi(i_1, \dots, i_m)$ of $\{1, \dots, m\}$ into the sets $\{j : i_j = i\}$, $i \geq 1$ (ignoring empty such sets). For any given partition π of $\{1, \dots, m\}$, let $S_{k_1,\dots,k_m;\pi}(G)$ be the sum of all terms in (6.3) with $\pi(i_1, \dots, i_m) = \pi$. Thus,

$$\prod_{i=1}^m S_{k_i}(G) = \sum_{\pi} S_{k_1,\dots,k_m;\pi}(G). \tag{6.4}$$

The term with $\pi = \pi_0 := \{1\} \cdots \{m\}$ equals $S_{k_1,\dots,k_m}(G)$, and any other term equals $S_{k'_1,\dots,k'_l}(G)$ for some $l < m$ ($l = |\pi|$, the number of parts of π) and some k'_j with $\sum_j k'_j = \sum_i k_i$ (each k'_j is the sum $\sum_{i \in A_j} k_i$ over a set $A_j \in \pi$).

We now take $G = \mathcal{G}(n,t)$ and use induction on m. Recall that $p_k(x) \asymp x^{2k-3}$ for $x \geq 1$. Thus

$$n^m p_{k_1}\Big(\frac{1}{1-nt}\Big) \cdots p_{k_m}\Big(\frac{1}{1-nt}\Big) \asymp \big(n(1-nt)^3\big)^m (1-nt)^{-\sum_i k_i},$$

and we find, by induction on m, for every $\pi \neq \pi_0$ and with $l = |\pi| < m$,

$$\mathbb{E}\, S_{k_1,\dots,k_m;\pi}(t) = O\Big(n^m p_{k_1}\Big(\frac{1}{1-nt}\Big) \cdots p_{k_m}\Big(\frac{1}{1-nt}\Big)\big(n(1-nt)^3\big)^{l-m}\Big).$$

Hence, (6.4) yields

$$\begin{aligned}\mathbb{E}\, S_{k_1,\dots,k_m}(t) &= \mathbb{E} \prod_{i=1}^m S_{k_i}(t) \\ &\quad + O\Big(n^m p_{k_1}\Big(\frac{1}{1-nt}\Big) \cdots p_{k_m}\Big(\frac{1}{1-nt}\Big)\big(n(1-nt)^3\big)^{-1}\Big).\end{aligned}$$

By Harris' inequality, $\mathbb{E}\prod_{i=1}^m S_{k_i}(t) \ge \prod_{i=1}^m \mathbb{E}\, S_{k_i}(t)$, and we obtain by Theorem 3.4 a lower bound of the required type. $\square$

We write $S_k(t;n)$ when needed to show the number of vertices explicitly.

Lemma 6.3. *For each $k \ge 2$ and $1-nt \ge n^{-1/3}$,*

$$\mathbb{E}\, S_k(t;n+1) - \mathbb{E}\, S_k(t;n) = p_k^*\Bigl(\frac{1}{1-nt}\Bigr) + O\Bigl(\frac{t}{(1-nt)^{2k+1}}\Bigr), \tag{6.5}$$

where p_k^* is a polynomial of degree $2k-2$ given by

$$p_k^*(x) := p_k(x) + (x^2-x)p_k'(x) = x^{-1}p_{k+1}(x). \tag{6.6}$$

The formula (6.5) is, not surprisingly, essentially what a formal differentiation of (3.17) with respect to n would give.

Proof. Let $\mathcal{G}(n,t)$ have the components $\mathcal{C}_1,\dots,\mathcal{C}_K$. Add a new vertex and add edges to it with the correct probabilities, and let $\Delta S_k := S_k(t;n+1) - S_k(t;n)$ be the resulting increase of $S_k(t)$. Let J_i be the indicator of the event that there is an edge between the new vertex and $\mathcal{C}_i$. Then

$$\Delta S_2 = 1 + \sum_i 2|\mathcal{C}_i|J_i + \frac12 \sum_{i,j}{}^{*}\, 2|\mathcal{C}_i||\mathcal{C}_j|J_iJ_j,$$

$$\begin{aligned}\Delta S_3 = 1 &+ \sum_i (3|\mathcal{C}_i| + 3|\mathcal{C}_i|^2)J_i + \frac12 \sum_{i,j}{}^{*}\,(3|\mathcal{C}_i|^2|\mathcal{C}_j| + 3|\mathcal{C}_i||\mathcal{C}_j|^2 + 6|\mathcal{C}_i||\mathcal{C}_j|)J_iJ_j \\ &+ \frac16 \sum_{i,j,k}{}^{*}\, 6|\mathcal{C}_i||\mathcal{C}_j||\mathcal{C}_k|J_iJ_jJ_k,\end{aligned}$$

and so on. Given the components $\mathcal{C}_1,\mathcal{C}_2,\dots$, the indicators J_i are independent with $\mathbb{E}\, J_i = 1 - e^{-|\mathcal{C}_i|t} = |\mathcal{C}_i|t + O(|\mathcal{C}_i|^2t^2)$. Hence, for $k=2$, using $|\mathcal{C}_i|t \le nt < 1$ to simplify terms like $|\mathcal{C}_i|^2t^2|\mathcal{C}_j|^2t^2$,

$$\mathbb{E}\bigl(\Delta S_2 \mid \mathcal{G}(n,t)\bigr) = 1 + 2tS_2(t) + O(t^2S_3(t)) + t^2S_{2,2}(t) + O(t^3S_{3,2}(t)).$$

Taking the expectation we find, using Lemma 6.2,

$$\mathbb{E}\,\Delta S_2 = 1 + 2ntp_2\Bigl(\frac{1}{1-nt}\Bigr) + (nt)^2p_2\Bigl(\frac{1}{1-nt}\Bigr)^2 + O\Bigl(\frac{t}{(1-nt)^5}\Bigr). \tag{6.7}$$

The same argument applies to every k, and yields an expression for $\mathbb{E}\,\Delta S_k$ where the main terms are of the type $c(nt)^m p_{k_1+1}\bigl(\frac{1}{1-nt}\bigr)\cdots p_{k_m+1}\bigl(\frac{1}{1-nt}\bigr)$, where c is a positive combinatorial constant, $0 \le m < k$, $1 \le k_i \le k-1$ and $\sum_i k_i \le k$; the error terms are all $O(1/(n(1-nt)^3)$ of some such terms. The main terms are polynomials in $1/(1-nt)$ of degree $\sum_i(2k_i-1) \le 2k-2$, so the result can be written as (6.5) for some polynomial p_k^*.

To identify p_k^*, fix $y \in (0,1/2)$ and a rational $\varepsilon \in (0,1)$, consider only n such that εn is an integer and let $t = y/n$ and repeat (6.5) εn times. This yields

$$\mathbb{E}\, S_k(t;(1+\varepsilon)n) - \mathbb{E}\, S_k(t;n) = \varepsilon n\Bigl(p_k^*\Bigl(\frac{1}{1-y}\Bigr) + O(\varepsilon)\Bigr) + O(\varepsilon),$$

and thus, by Theorem 3.4,

$$(1+\varepsilon)np_k\Big(\frac{1}{1-(1+\varepsilon)y}\Big)-np_k\Big(\frac{1}{1-y}\Big)=\varepsilon np_k^*\Big(\frac{1}{1-y}\Big)+O(\varepsilon^2 n)+O(1).$$

Divide by n and let $n\to\infty$; this gives

$$\varepsilon p_k^*\Big(\frac{1}{1-y}\Big)=(1+\varepsilon)p_k\Big(\frac{1}{1-(1+\varepsilon)y}\Big)-p_k\Big(\frac{1}{1-y}\Big)+O(\varepsilon^2).$$

Divide by ε and let $\varepsilon\to 0$; this gives, with $x=1/(1-y)$,

$$p_k^*(x)=p_k(x)+\frac{y}{(1-y)^2}p_k'(x)=p_k(x)+(x^2-x)p_k'(x).$$

The final identification of this as $x^{-1}p_{k+1}(x)$ follows by (7.8) proved in Section 7 below. Alternatively, the proof of Theorem 6.1 below and the symmetry of $\operatorname{Cov}(S_k,S_l)$ shows that $p_{k+l}-p_{k+1}p_l^*=p_{k+l}-p_{l+1}p_k^*$, and thus, choosing $l=2$, $p_k^*=p_{k+1}\cdot p_2^*/p_3$, which yields the formula, since it follows from (6.7) that $p_2^*(x)=x^2$. (This thus gives an alternative proof of (7.8).) □

Proof of Theorem 6.1. Let $\mathcal{A}_n$ and $I_A(t)$ be as in the proof of Lemma 3.1. Conditioned on $I_A(t)=1$, the complement of A is a random graph equivalent to $\mathcal{G}(n-|A|,t)$. Thus,

$$\begin{aligned}\operatorname{Cov}\big(S_k(t),S_l(t)\big)&=\mathbb{E}\Big(\sum_{A\in\mathcal{A}_n}|A|^kI_A(t)\sum_{B\in\mathcal{A}_n}|B|^lI_B(t)\Big)-\mathbb{E}\,S_k(t)\,\mathbb{E}\,S_l(t)\\&=\mathbb{E}\sum_{A\in\mathcal{A}_n}|A|^{k+l}I_A(t)+\mathbb{E}\sum_{A\in\mathcal{A}_n}|A|^kI_A(t)\Big(\sum_{B\cap A=\emptyset}|B|^lI_B(t)-\mathbb{E}\,S_l(t)\Big)\\&=\mathbb{E}\,S_{k+l}(t)+\mathbb{E}\sum_{A\in\mathcal{A}_n}|A|^kI_A(t)\Big(\mathbb{E}\,S_l(t;n-|A|)-\mathbb{E}\,S_l(t;n)\Big).\end{aligned}$$

By Lemma 6.3, for some $\theta\in[0,1]$,

$$\begin{aligned}\mathbb{E}\,S_l(t;n)-\mathbb{E}\,S_l(t;n-|A|)&=|A|p_l^*\Big(\frac{1}{1-nt+\theta|A|t}\Big)+O\Big(\frac{|A|t}{(1-nt)^{2l+1}}\Big)\\&=|A|p_l^*\Big(\frac{1}{1-nt}\Big)+O\Big(\frac{t|A|^2}{(1-nt)^{2l-1}}\Big)+O\Big(\frac{t|A|}{(1-nt)^{2l+1}}\Big).\end{aligned}$$

Consequently,

$$\begin{aligned}\operatorname{Cov}(S_k(t),S_l(t))&=\mathbb{E}\,S_{k+l}(t)-\mathbb{E}\,S_{k+1}(t)p_l^*\Big(\frac{1}{1-nt}\Big)\\&\quad+O\Big(\frac{t}{(1-nt)^{2l-1}}\,\mathbb{E}\,S_{k+2}(t)\Big)+O\Big(\frac{t}{(1-nt)^{2l+1}}\,\mathbb{E}\,S_{k+1}(t)\Big),\end{aligned}$$

and the result follows by Theorem 3.4. □

In the case $nt\to 1$, only the leading term of $P_{k,l}$ is significant in Theorem 6.1. Since the leading term of p_k is $(2k-5)!!\,x^{2k-3}$, as follows by (7.8) in Section 7, we have the following corollary.

Corollary 6.4. *For every $k, l \geq 2$, if $nt \to 1$ with $1 - nt \gg n^{-1/3}$, then*

$$\operatorname{Cov}\bigl(S_k(t), S_l(t)\bigr) \sim c_{k,l} n(1-nt)^{3-2k-2l},$$

with $c_{k,l} := (2k+2l-5)!! - (2k-3)!!\,(2l-3)!!$.

In particular, under these conditions,

$$\begin{aligned}
\operatorname{Var}(S_2(t)) &\sim 2n(1-nt)^{-5},\\
\operatorname{Var}(S_3(t)) &\sim 96\,n(1-nt)^{-9},\\
\operatorname{Var}(S_4(t)) &\sim 10170\,n(1-nt)^{-11},
\end{aligned}$$

cf. Table 2 and (6.2).

Proof of Theorem 1.3, asymptotic variance. Immediate from Theorem 6.1, see (6.2). □

7. The Borel distribution

Let $T(z)$ be the *tree function*

$$T(z) := \sum_{j=1}^{\infty} \frac{j^{j-1} z^j}{j!}, \qquad |z| \leq e^{-1},$$

and recall the well-known formulas $T(z)e^{-T(z)} = z$ ($|z| \leq e^{-1}$), $T(\alpha e^{-\alpha}) = \alpha$ ($0 \leq \alpha \leq 1$), and

$$T'(z) = \frac{T(z)}{z\bigl(1 - T(z)\bigr)}. \tag{7.1}$$

A random variable B_λ has the *Borel distribution* $\operatorname{Bo}(\lambda)$ with parameter $\lambda \in [0,1]$ if

$$\mathbb{P}(B_\lambda = j) = \frac{j^{j-1}}{j!}\lambda^{j-1}e^{-j\lambda} = \frac{1}{T(\lambda e^{-\lambda})}\frac{j^{j-1}}{j!}(\lambda e^{-\lambda})^j, \qquad j = 1, 2, \dots \tag{7.2}$$

The probability generating function of the Borel distribution is thus

$$\mathbb{E}\, z^{B_\lambda} = \sum_{l=1}^{\infty} \mathbb{P}(B_\lambda = l) z^l = \frac{T(\lambda e^{-\lambda} z)}{T(\lambda e^{-\lambda})} = \frac{T(\lambda e^{-\lambda} z)}{\lambda}. \tag{7.3}$$

It is well-known that $\operatorname{Bo}(\lambda)$ is the distribution of the total progeny of a Galton–Watson branching process where each individual has $\operatorname{Po}(\lambda)$ children; for this and related results, see e.g. [4; 17; 16; 22; 20; 6; 21; 18; 15].

Now consider $G(n,p)$ with $p = \lambda/n$ for a fixed $\lambda < 1$, and let $\mathcal{C}_v$ be the component containing a fixed vertex v. It is easily seen that as $n \to \infty$, for every fixed $j \geq 1$, $\mathbb{P}(|\mathcal{C}_v| = j) \to \mathbb{P}(B_\lambda = j)$ given by (7.2), either by the usual branching process approximation and the result just quoted, or by a direct estimation of the probability, using Cayley's formula for the number of trees of order j and the fact that w.h.p. the component $\mathcal{C}_v$ is a tree. In other words,

$|\mathcal{C}_v| \xrightarrow{\mathrm{d}} B_\lambda$. For any integer m, the moment $\mathbb{E}\,|\mathcal{C}_v|^m = \mathbb{E}\, S_{m+1}(G(n,p))/n$, and Theorem 3.4 shows, with $t = -\log(1-p)$ and thus $nt \to \lambda$, that

$$\mathbb{E}\,|\mathcal{C}_v|^m = \frac{\mathbb{E}\, S_{m+1}(G(n,\lambda/n))}{n} \to p_{m+1}\Big(\frac{1}{1-\lambda}\Big).$$

Since thus $|\mathcal{C}_v|$ converges in distribution and all moments converge (to finite limits), the moments have to converge to the moments of the limit distribution. We have thus shown the following.

Theorem 7.1. *The polynomials p_k describe the moments of the Borel distribution* $\mathrm{Bo}(\lambda)$ *by the formula*

$$\mathbb{E}\, B_\lambda^m = p_{m+1}\Big(\frac{1}{1-\lambda}\Big), \qquad m \geq 1.$$

For example, as is well-known, $\mathbb{E}\, B_\lambda = (1-\lambda)^{-1}$ and $\mathbb{E}\, B_\lambda^2 = (1-\lambda)^{-3}$.

Remark 7.2. By Theorem 7.1, Corollary 4.2 can be written

$$S_k(t) \sim_{\mathrm{p}} n\,\mathbb{E}\, B_{nt}^{k-1}, \qquad 1 - nt \gg n^{-1/3}.$$

This is not surprising since we have $S_k(G) = \sum_v |\mathcal{C}_v|^{k-1}$, and we expect only a weak dependence between the components $\mathcal{C}_v$ in this range, so this is a kind of law of large numbers.

Let, cf. (7.3), for $|t|$ small enough,

$$\psi(t;\lambda) = \mathbb{E}\, e^{tB_\lambda} = \sum_{m=0}^{\infty} \frac{t^m}{m!}\,\mathbb{E}\, B_\lambda^m = \frac{T(\lambda e^{-\lambda} e^t)}{\lambda} \tag{7.4}$$

be the moment generating function of $B_\lambda \sim \mathrm{Bo}(\lambda)$. The moments of B_λ can be obtained by differentiation of $\psi(t;\lambda)$ at $t = 0$.

Lemma 7.3. *For each $m \geq 0$ there exists a polynomial r_m such that*

$$\frac{\mathrm{d}^m}{\mathrm{d}t^m}\psi(t;\lambda) = \frac{T(\lambda e^{-\lambda} e^t)}{\lambda} r_m\Big(\frac{1}{1 - T(\lambda e^{-\lambda} e^t)}\Big). \tag{7.5}$$

We have $r_0(x) = 1$ and

$$r_{m+1}(x) = x r_m(x) + (x^3 - x^2) r_m'(x), \qquad m \geq 0. \tag{7.6}$$

By (7.6), $r_1(x) = x$, and it follows by induction that r_m has degree $2m-1$ for $m \geq 1$.

Proof. For $m = 0$, (7.5) is just (7.4).

Suppose that (7.5) holds for some $m \geq 0$. Then, by the chain rule and (7.1), with $T = T(\lambda e^{-\lambda} e^t)$,

$$\begin{aligned}
\frac{\mathrm{d}^{m+1}}{\mathrm{d}t^{m+1}}\psi(t;\lambda) &= \frac{\mathrm{d}}{\mathrm{d}T}\Big(\frac{T}{\lambda}r_m\Big(\frac{1}{1-T}\Big)\Big)\cdot\frac{T}{1-T} \\
&= \frac{1}{\lambda}\frac{T}{1-T}r_m\Big(\frac{1}{1-T}\Big) + \frac{T^2}{\lambda(1-T)^3}r_m'\Big(\frac{1}{1-T}\Big) \\
&= \frac{T}{\lambda}\left(\frac{1}{1-T}r_m\Big(\frac{1}{1-T}\Big) + \Big(\frac{1}{(1-T)^3}-\frac{1}{(1-T)^2}\Big)r_m'\Big(\frac{1}{1-T}\Big)\right),
\end{aligned}$$

which verifies (7.5) for $m+1$ with r_{m+1} given by (7.6). □

Setting $t = 0$ in (7.5) yields

$$\mathbb{E}\, B_\lambda^m = \frac{\mathrm{d}^m}{\mathrm{d}t^m}\psi(t;\lambda)\big|_{t=0} = r_m\Big(\frac{1}{1-\lambda}\Big), \qquad m \geq 0.$$

Consequently, Theorem 7.1 shows that

$$r_m(x) = p_{m+1}(x), \qquad m \geq 1. \tag{7.7}$$

In particular, (7.6) yields the simple linear recursion

$$p_{k+1}(x) = xp_k(x) + (x^3 - x^2)p_k'(x), \qquad k \geq 2. \tag{7.8}$$

It is evident from (7.8) and induction that, for $k \geq 2$, the leading term of p_k is $(2k-5)!!\, x^{2k-3}$ (with the standard interpretation $(-1)!! = 1$) and that for $k \geq 3$, the lowest order non-zero term is $(-1)^{k-1}(k-2)!\, x^k$, see Table 1.

Remark 7.4. The quadratic recursion (3.19) can be seen to be equivalent to the quadratic partial differential equation

$$\frac{\partial}{\partial\lambda}\psi(t;\lambda) = \big(\psi(t;\lambda) - 1\big)\frac{\partial}{\partial t}\psi(t;\lambda),$$

while the linear recursion (7.8) is equivalent to the linear partial differential equation

$$\frac{\partial\psi}{\partial t}(t;\lambda) = \frac{1}{1-\lambda}\psi(t;\lambda) + \frac{\lambda}{1-\lambda}\frac{\partial\psi}{\partial\lambda}(t;\lambda).$$

Remark 7.5. By Theorem 7.1, the recursion (3.19) can be written

$$\frac{\mathrm{d}}{\mathrm{d}\lambda}\mathbb{E}\, B_\lambda^{k-1} = (1-\lambda)^{-2}p_k'\Big(\frac{1}{1-\lambda}\Big) = \frac{1}{2}\sum_{l=1}^{k-1}\binom{k}{l}\mathbb{E}\, B_\lambda^l\, \mathbb{E}\, B_\lambda^{k-l},$$

or, if B'_λ and B''_λ are independent copies of B_λ, using $\frac{\mathrm{d}}{\mathrm{d}\lambda}\mathbb{P}(B_\lambda = j) = \bigl(\frac{j-1}{\lambda} - j\bigr)\mathbb{P}(B_\lambda = j)$ from (7.2),

$$\begin{aligned}
\mathbb{E}(B'_\lambda + B''_\lambda)^k &= 2\,\mathbb{E}\,B_\lambda^k + 2\frac{\mathrm{d}}{\mathrm{d}\lambda}\,\mathbb{E}\,B_\lambda^{k-1} \\
&= \sum_{j=1}^\infty \mathbb{P}(B_\lambda = j)\cdot\Bigl(2j^k + 2j^{k-1}\bigl(\frac{j-1}{\lambda} - j\bigr)\Bigr) \\
&= \sum_{j=1}^\infty \mathbb{P}(B_\lambda = j)\cdot\frac{2(j-1)}{j\lambda}j^k,
\end{aligned}$$

which is equivalent to the well-known formula

$$\mathbb{P}(B'_\lambda + B''_\lambda = j) = \frac{2(j-1)}{j\lambda}\,\mathbb{P}(B_\lambda = j) = 2\frac{j^{j-3}}{(j-2)!}\lambda^{j-2}e^{-j\lambda}, \qquad j \ge 2;$$

see e.g. [22; 21; 18; 15] and note that $B'_\lambda + B''_\lambda$ can be seen as the total progeny of a Galton–Watson process with $\mathrm{Po}(\lambda)$ offspring started with 2 individuals, or as the limit distribution of $|\mathcal{C}_v \cup \mathcal{C}_w|$ if $\mathcal{C}_v$ and $\mathcal{C}_w$ are the components containg two given vertices in $G(n, \lambda/n)$.

Remark 7.6. The *cumulants* $\varkappa_m$ of the Borel distribution $\mathrm{Bo}(\lambda)$ are the Taylor coefficients of $\log\psi(t;\lambda)$ at $t = 0$ (times $m!$). Since $T(z) = ze^{T(z)}$, (7.4) yields

$$\log\psi(t;\lambda) = T(\lambda e^{-\lambda}e^t) - \lambda + t = \lambda\psi(t;\lambda) - \lambda + t,$$

and thus

$$\varkappa_m(B_\lambda) = \frac{\mathrm{d}^m}{\mathrm{d}t^m}\log\psi(t;\lambda)\big|_{t=0} = \lambda\,\mathbb{E}\,B_\lambda^m = \lambda p_{m+1}\Bigl(\frac{1}{1-\lambda}\Bigr), \qquad m \ge 2,$$

while, of course, $\varkappa_1(B_\lambda) = \mathbb{E}\,B_\lambda = (1-\lambda)^{-1}$.

We can interpret the asymptotic covariances and the polynomials $P_{k,l}$ in Section 6 by introducing the *size-biased Borel distribution* $\widehat{B}_\lambda$ defined by

$$\mathbb{P}(\widehat{B}_\lambda = j) = \frac{j\,\mathbb{P}(B_\lambda = j)}{\mathbb{E}\,B_\lambda} = (1-\lambda)\frac{j^j}{j!}\lambda^{j-1}e^{-j\lambda}. \tag{7.9}$$

Then

$$\mathbb{E}\,\widehat{B}_\lambda^m = \mathbb{E}\,B_\lambda^{m+1}/\,\mathbb{E}\,B_\lambda = (1-\lambda)p_{m+2}\Bigl(\frac{1}{1-\lambda}\Bigr), \qquad m \ge 0, \tag{7.10}$$

and thus, by (6.1),

$$P_{k,l}\Bigl(\frac{1}{1-nt}\Bigr) = \frac{1}{1-nt}\,\mathrm{Cov}\bigl(\widehat{B}_{nt}^{k-1}, \widehat{B}_{nt}^{l-1}\bigr). \tag{7.11}$$

Hence, by Theorem 6.1, the random variables $n^{-1/2}(1-nt)^{1/2}S_k(t)$, $k \ge 2$, have asymptotically the same covariance structure as $\widehat{B}_{nt}^{k-1}$.

Appendix A. The supercritical case

Consider $G(n,p)$ with $np-1 \gg n^{-1/3}$. It is well-known, see e.g. [13, Chapter 5], that w.h.p. $G(n,p)$ has a unique giant component. More precisely, there is a deterministic function $\rho > 0$ on $(1,\infty)$ such that, if the components $\mathcal{C}_1, \mathcal{C}_2, \dots$ of $G(n,p)$ are ordered with $|\mathcal{C}_1| \ge |\mathcal{C}_2| \ge \dots$, then $|\mathcal{C}_1| \sim_{\mathrm{p}} n\rho(np) \gg n^{2/3}$, while $|\mathcal{C}_2| = o_{\mathrm{p}}(n^{2/3})$. The function $\rho(\lambda)$ is the survival probability of a Galton–Watson branching process with $\mathrm{Po}(\lambda)$ offspring, and is given by the equation

$$\rho(\lambda) = 1 - e^{-\lambda\rho(\lambda)}. \tag{A.1}$$

The largest component is thus much larger than the others, and it turns out that it dominates all other terms in the sums S_k. We write in this appendix $S_k(n,p)$ for $S_k(G(n,p))$, and continue to let $\mathcal{C}_1$ denote the largest component of $G(n,p)$.

Theorem A.1. *If $np-1 \gg n^{-1/3}$, then for every $k \ge 2$,*

$$S_k(n,p) = |\mathcal{C}_1|^k + O_{\mathrm{p}}\left(\frac{n}{(np-1)^{2k-3}}\right) \sim_{\mathrm{p}} |\mathcal{C}_1|^k \sim_{\mathrm{p}} \big(n\rho(np)\big)^k.$$

In particular, then $\chi(G(n,p)) \sim_{\mathrm{p}} n\rho(np)^2$. We first prove a technical lemma.

Lemma A.2. *There exists a function $\alpha : (1,\infty) \to (0,1)$ such that the following holds, for some $c > 0$:*

(i) *For any $p = p(n)$ with $np - 1 \gg n^{-1/3}$, w.h.p. $|\mathcal{C}_1| > \alpha(np)n$.*
(ii) *$\alpha(\lambda) \ge c\min(\lambda - 1, 1)$.*
(iii) *If $1 < \lambda \le 2$, then $\lambda(1-\alpha(\lambda)) \le 1 - c(\lambda - 1)$.*
(iv) *If $\lambda \ge 2$, then $\lambda(1-\alpha(\lambda)) \le 1 - c$.*
(v) *For each $m \ge 0$, $1 - \alpha(\lambda) = O(\lambda^{-m})$.*

Proof. Note first that (ii) follows from (iii) and (iv) because $1 - \alpha(\lambda) \le \lambda(1-\alpha(\lambda))$.

For any fixed $M > 1$, we can take $\alpha(\lambda) = (1-\varepsilon)\rho(\lambda)$ for $1 < \lambda \le M$, if ε is sufficiently small. This choice satisfies (i), in this range (v) is trivial, and it is easily seen that (iii) and (iv) follow (provided ε is small enough) from the facts that $\rho(\lambda) \sim 2(\lambda - 1)$ and $\lambda(1-\rho(\lambda)) = 1 - (\lambda - 1) + O(\lambda - 1)^2$ as $\lambda \searrow 1$, and $\lambda(1-\rho(\lambda)) < 1$ for $\lambda > 1$. (All three are easily verified by writing (A.1) as $\lambda = -\log(1-\rho)/\rho$.)

For large λ, we argue as follows. Take $\gamma < \rho(2)$. Thus, w.h.p. $G(n, 2/n)$ has a giant component of order at least γn. For $\lambda = np > 2$, construct $G(n,p)$ by the usual two-round method: first take $G(n,2/n)$ and then add further edges independently in a second round with probabilities $p - 2/n$ (or, to be precise, $(np-2)/(n-2) > p - 2/n$). If we obtain a component of order at least γn in the first round, then the probability that a given vertex will *not* be joined to this component in the second round is less than $\exp(-\gamma n(p - 2/n)) = \exp(2\gamma - \gamma\lambda)$. Hence, w.h.p. the number of such vertices is less than $n\exp(2\gamma-\gamma\lambda/2)$; for $\lambda=O(1)$ by concentration of the binomial distribution and for $\lambda\to\infty$ by Markov's inequality. Consequently, there is w.h.p. a component with more than $n-n\exp(2\gamma-\gamma\lambda/2)$ vertices; hence (i) holds with $\alpha(\lambda)=1-\exp(2\gamma-\gamma\lambda/2)$. This α satisfies (v) too, and (iv) for large enough λ. We thus can use this α for $\lambda>M$ for some large M, and the first construction for smaller λ. $\square$

Proof of Theorem A.1. Let $\alpha=\alpha(np)$ be as in Lemma A.2, and use the notation of the proof of Lemma 3.1. By Lemma A.2(ii) and the assumption $np-1\gg n^{-1/3}$, we have $\alpha\gg n^{-1/3}$.

Let N be the number of components of size $>\alpha n$ in $G(n,p)$ (thus w.h.p. $N\geq 1$ by Lemma A.2(i)), and let

$$Z_k:=\sum_{|A|>\alpha n}\sum_{B\cap A=\emptyset}|B|^k I_A I_B.$$

Then

$$\mathbb{E}\,Z_k:=\mathbb{E}\sum_{|A|>\alpha n}I_A\,\mathbb{E}\,S_k(n-|A|,p)\leq \mathbb{E}\,N\,\mathbb{E}\,S_k(n-\lceil\alpha n\rceil,p). \tag{A.2}$$

If $1<np\leq 2$, then by Lemma A.2(iii), $(n-\lceil\alpha n\rceil)p\leq np(1-\alpha)\leq 1-c(np-1)$, and thus by Lemma 3.2,

$$\mathbb{E}\,S_k(n-\lceil\alpha n\rceil,p)=O\biggl(\frac{n-\lceil\alpha n\rceil}{(np-1)^{2k-3}}\biggr)=O\biggl(\frac{n}{(np-1)^{2k-3}}\biggr).$$

If instead $np>2$, then by Lemma A.2(iv), $(n-\lceil\alpha n\rceil)p\leq np(1-\alpha)\leq 1-c$, and thus by Lemmas 3.2 and A.2(v), with $m=2k-3$,

$$\mathbb{E}\,S_k(n-\lceil\alpha n\rceil,p)=O(n-\lceil\alpha n\rceil)=O\biggl(\frac{n}{(np)^{2k-3}}\biggr).$$

Hence, for all np,

$$\mathbb{E}\,S_k(n-\lceil\alpha n\rceil,p)=O\biggl(\frac{n}{(np-1)^{2k-3}}\biggr)=o\bigl(n^{2k/3}\bigr). \tag{A.3}$$

Note first that $Z_k\geq N(N-1)(\alpha n)^k$. Hence, by (A.2) and (A.3),

$$\mathbb{E}\,N(N-1)\leq(\alpha n)^{-k}\,\mathbb{E}\,Z_k=o(\mathbb{E}\,Nn^{2k/3}(\alpha n)^{-k})=o(\mathbb{E}\,N).$$

Since $N\leq 1+N(N-1)$, it follows that $\mathbb{E}\,N(N-1)=o(1)$ and $\mathbb{E}\,N=O(1)$; hence (A.2) and (A.3) yield $\mathbb{E}\,Z_k=O(n/(np-1)^{2k-3})$. By Lemma A.2(i), w.h.p. $|\mathcal{C}_1|>\alpha n$; in this case, $|\mathcal{C}_1|^k\leq S_k(n,p)\leq|\mathcal{C}_1|^k+Z_k$, and the result follows. $\square$

Appendix B. The critical case

The critical case is $np=1+O(n^{-1/3})$. By considering subsequences, it suffices to consider the case $n^{1/3}(np-1)\to\tau$ for some $\tau\in(-\infty,\infty)$, i.e., $np=1+(\tau+o(1))n^{-1/3}$.

We continue to use the notations of Appendix A. It is well-known that in the critical case, $|\mathcal{C}_1|$ is of the order $n^{2/3}$, in the sense that $|\mathcal{C}_1|/n^{2/3}$ converges in distribution to some non-degenerate random variable, and the same holds for $|\mathcal{C}_2|$, $|\mathcal{C}_3|$, ... Moreover, Aldous [1] has shown that, with notations as in Appendix A, the sequence $(n^{-2/3}|\mathcal{C}_1|, n^{-2/3}|\mathcal{C}_2|, \dots)$ (extended by an infinite number of 0's) converges in distribution to a certain random sequence $(C^\tau(1), C^\tau(2), \dots)$ that can be described as the sequence of excursion lengths of a certain reflecting Brownian motion with inhomogeneous drift (depending on τ) that is defined in [1]. The convergence is in the ℓ^2-topology, and thus immediately implies convergence of the sums of squares. Moreover, convergence in ℓ^2 implies convergence in ℓ^k for every $k \geq 2$, and thus we also have convergence of the sums of kth powers. Consequently,

Theorem B.1. *If $np = 1 + (\tau + o(1))n^{-1/3}$ with $-\infty < \tau < \infty$, then for every $k \geq 2$,*

$$n^{-2k/3} S_k(n,p) \xrightarrow{\mathrm{d}} W_k := \sum_i C^\tau(i)^k.$$

Note that we here have limits that are non-degenerate random variables and not constants, unlike the subcritical and supercritical cases where $S_k(n,p) \sim_{\mathrm{p}} a_n$ for a suitable sequence a_n.

Remark B.2. Janson and Spencer [14] give a related description of the limit of the component sizes as a point process $\Xi^{(\tau)}$ on $(0,\infty)$. It follows that we also have $W_k = \int_0^\infty x^k \, \mathrm{d}\Xi^{(\tau)}(x)$, and thus $\mathbb{E}\, W_k = \int_0^\infty x^k \, \mathrm{d}\Lambda^{(\tau)}(x)$, where $\Lambda^{(\tau)}$ is the intensity of $\Xi^{(\tau)}$ given in [14, Theorem 4.1].

References

[1] D. Aldous, Brownian excursions, critical random graphs and the multiplicative coalescent. *Ann. Appl. Probab.* **25** (1997), 812–854.
[2] B. Bollobás, *Random Graphs*, 2nd ed., Cambridge Univ. Press, Cambridge, 2001.
[3] B. Bollobás, S. Janson and O. Riordan, Susceptibility in inhomogeneous random graphs. In preparation.
[4] É. Borel, Sur l'emploi du théorème de Bernoulli pour faciliter le calcul d'une infinité de coefficients. Application au problème de l'attente à un guichet. *C. R. Acad. Sci. Paris* **214** (1942), 452–456.
[5] R. Durrett, *Random Graph Dynamics.* Cambridge University Press, Cambridge, 2007.
[6] M. Dwass, The total progeny in a branching process and a related random walk. *J. Appl. Probab.* **6** (1969), 682–686.
[7] A. Gut, *Probability: A Graduate Course.* Springer, New York, 2005.
[8] J. Jacod & A. N. Shiryaev, *Limit Theorems for Stochastic Processes.* Springer, Berlin, 1987.
[9] S. Janson, A functional limit theorem for random graphs with applications to subgraph count statistics. *Random Struct. Alg.* **1** (1990), 15–37.

[10] S. Janson, *Orthogonal Decompositions and Functional Limit Theorems for Random Graph Statistics*. Mem. Amer. Math. Soc., vol. 111, no. 534, Amer. Math. Soc., Providence, RI, 1994.

[11] S. Janson, Functional limit theorems for multitype branching processes and generalized Pólya urns. *Stochastic Processes Appl.* **110** (2004), no. 2, 177–245.

[12] S. Janson & M. Luczak, Asymptotic normality of the k-core in random graphs. *Ann. Appl. Probab.* **18** (2008), no. 3, 1085–1137.

[13] S. Janson, T. Łuczak & A. Ruciński, *Random Graphs*, Wiley, New York, 2000.

[14] S. Janson & J. Spencer, A point process describing the component sizes in the critical window of the random graph evolution. *Combin. Probab. Comput.* **16** (2007), 631–658.

[15] N. L. Johnson, A. W. Kemp & S. Kotz, *Univariate discrete distributions*. Third edition. Wiley-Interscience, Hoboken, NJ, 2005.

[16] D.G. Kendall, Some problems in the theory of queues. *J. Roy. Statist. Soc. Ser. B.* **13** (1951), 151–185.

[17] R. Otter, The multiplicative process. *Ann. Math. Statist.* **20** (1949), 206–224.

[18] J. Pitman, Enumerations of trees and forests related to branching processes and random walks. *Microsurveys in discrete probability (Princeton, NJ, 1997)*, 163–180, DIMACS Ser. Discrete Math. Theoret. Comput. Sci., 41, Amer. Math. Soc., Providence, RI, 1998.

[19] J. Spencer & N. Wormald, Birth control for giants. *Combinatorica* **27** (2007), no. 5, 587–628.

[20] L. Takács, *Combinatorial Methods in the Theory of Stochastic Processes*. Wiley, New York, 1967.

[21] L. Takács, Ballots, queues and random graphs. *J. Appl. Probab.* **26** (1989), no. 1, 103–112.

[22] J.C. Tanner, A derivation of the Borel distribution. *Biometrika* **48** (1961), 222–224.

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

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

URL: `http://www.math.uu.se/~svante/`

Department of Mathematics, London School of Economics, Houghton Street, London WC2A 2AE, United Kingdom

E-mail address: `m.j.luczak@lse.ac.uk`

URL: `http://www.lse.ac.uk/people/m.j.luczak@lse.ac.uk/`