# A POINT PROCESS DESCRIBING THE COMPONENT SIZES IN THE CRITICAL WINDOW OF THE RANDOM GRAPH EVOLUTION

SVANTE JANSON AND JOEL SPENCER

Abstract. We study a point process describing the asymptotic behavior of sizes of the largest components of the random graph $G(n,p)$ in the critical window, that is, for $p = n^{-1} + \lambda n^{-4/3}$, where $\lambda$ is a fixed real number. In particular, we show that this point process has a surprising rigidity. Fluctuations in the large values will be balanced by opposite fluctuations in the small values such that the sum of the values larger than a small $\varepsilon$ (a scaled version of the number of vertices in components of size greater than $\varepsilon n^{2/3}$) is almost constant.

## 1. Introduction

We consider the asymptotic behavior of the component sizes in the random graph $G(n,p)$, where throughout this paper $p = n^{-1} + \lambda n^{-4/3}$ for some fixed $\lambda$ with $-\infty < \lambda < \infty$. It is well-known (see Remark 1.6) that this is the critical window SJ
of $p$ where the "phase transition" occurs. It is further well-known that, for the $p$ we consider, the largest components are of order $n^{2/3}$. We therefore scale by this factor; if the components are $\mathcal{C}_1, \mathcal{C}_2, \dots, \mathcal{C}_r$, in order of decreasing size, say, and $|\mathcal{C}_i|$ is the size (order) of $\mathcal{C}_i$, we define $\xi_{ni}$ to be $n^{-2/3}|\mathcal{C}_i|$ and consider the random set $\Xi_n := \{\xi_{ni}\}_{i=1}^r$ as a point process on $(0,\infty)$ or $(0,\infty]$. See Appendix A for some technical background and note that it is convenient to define the point process formally as a random measure with point masses at the points $\xi_{ni}$; we will sometimes use this formalism, writing for example $\Xi_n[a,b]$ for the number of points in $[a,b]$, but we will also speak (and think) of point processes as random sets.

It follows immediately from Aldous [1, Corollary 2], see Lemma A.2, that as $n \to \infty$, the point processes $\Xi_n$ converge in distribution to some point process $\Xi^{(\lambda)} = \{\xi_i^{(\lambda)}\}$ on $(0,\infty]$ (in the vague topology on $(0,\infty]$, see Appendix A); this also follows from a minor extension of results in Łuczak, Pittel and Wierman [18], see Janson, Łuczak and Ruciński [12, Theorem 5.20]. Aldous [1] further gave a description of the limiting process $\Xi^{(\lambda)}$ as the set of lengths of excursions of a certain reflected Brownian motion with parabolic drift, defined as $B^\lambda(s) := W^\lambda(s) - \min_{0\le u\le s} W^\lambda(u)$, $s \ge 0$, where $W^\lambda(s) = W(s) + \lambda s - s^2/2$ for a standard Brownian motion $W$.

We will usually keep $\lambda$ fixed, and will then often omit it from the notation, thus writing $\Xi = \Xi^{(\lambda)}$ and $\xi_i = \xi_i^{(\lambda)}$. Conversely, we may write $\Xi_{n,p}$ when necessary. Note that we may regard $(\Xi^{(\lambda)})_\lambda$ as a stochastic process indexed by $\lambda \in (-\infty,\infty)$; this is the standard multiplicative coalescent as constructed by Aldous [1], except that the variables $\Xi^{(\lambda)}$ are represented as point processes while Aldous uses the equivalent representation as sequences $(\xi_i)_1^\infty$; cf. Lemma A.2, although Aldous uses a stronger topology.

*Date*: May 24, 2005; revised October 2, 2006.

2000 *Mathematics Subject Classification.* 60C05, 60K99, 05C80.

The aim of this paper is to study the limiting point process $\Xi$. The number of components in $G(n,p)$ tends to infinity (in probability) as $n \to \infty$, so we expect an infinite number of points $\xi_i$ in $\Xi$. Moreover, if we say that the *weight* of a point $x$ is $x$, the total weight of $\Xi_n$ is $\sum_i \xi_{ni} = n^{1/3}$, so we expect the total weight of $\Xi$, i.e. $\sum \xi_i = \int x\, d\Xi$, to be infinite a.s.; indeed, this is a simple consequence of Theorem 1.1. (Still, we caution that results on the limiting process $\Xi$ do not automatically follow from results on the discrete $G(n,p)$. Had we, for example, chosen the "wrong" parameterization $\xi_{ni} = n^{-0.7}|\mathcal{C}_i|$ then $\Xi$ would be almost surely empty.)

Our main result is the following. We also give in later sections various other results; several of them have been more or less well-known for a long time, but perhaps not published previously in this form.

**Theorem 1.1.** *Let $-\infty < \lambda < \infty$, and let $\Xi$ be the limiting point process defined above. Let $Z_\varepsilon := \sum_{\xi_i \ge \varepsilon} \xi_i = \int_\varepsilon^\infty x\, d\Xi(x)$ be the total weight of all points in $\Xi$ that are at least $\varepsilon$. Then, as $\varepsilon \to 0$,*

$$\mathbb{E}\, Z_\varepsilon = \Big(\frac{2}{\pi}\Big)^{1/2} \varepsilon^{-1/2} + \lambda + (2\pi)^{-1/2}\lambda^2 \varepsilon^{1/2} + O(\varepsilon) \tag{1.1}$$

*and*

$$\operatorname{Var} Z_\varepsilon = (2/\pi)^{1/2}\varepsilon^{1/2} + O(\varepsilon). \tag{1.2}$$

*In particular, $\mathbb{E}\, Z_\varepsilon \to \infty$ and $\operatorname{Var} Z_\varepsilon \to 0$ as $\varepsilon \to 0$.*

We will also give an exact, but more complicated, formulas for $\mathbb{E}\, Z_\varepsilon$ in Corollary 4.2 and $\operatorname{Var} Z_\varepsilon$ in Corollary 8.3. It seems non-trivial to obtain the asymptotics above from these formulas.

Thus, as $\varepsilon \to 0$, the variables $Z_\varepsilon$ tend to infinity, but they become more and more concentrated about their mean; hence, the random fluctuations disappear in the limit. In other words, the process $\Xi$ is very rigid, and any random fluctuation in the weights of the largest points has to be exactly balanced by opposite fluctuations in the weights of smaller points; this will be seen again in Section 8 where we consider the Palm distributions. Note that, because of the scaling, this is a non-trivial result in contrast to the corresponding fact that $\Xi_n$ has a constant total weight $n^{1/3}$. Note also that this is very far from the behaviour of a Poisson process.

We will prove Theorem 1.1 by two different methods, both classical, each giving a partial result only, in Sections 2 and 6.

In contrast to Theorem 1.1, the number of points $\ge \varepsilon$ in $\Xi$, i.e. $\Xi[\varepsilon,\infty)$, is not sharply concentrated.

**Theorem 1.2.** *Let $W_\varepsilon := \Xi[\varepsilon,\infty)$ be the number of points in $\Xi$ that are at least $\varepsilon$. Then, as $\varepsilon \to 0$,*

$$\mathbb{E}\, W_\varepsilon = \Big(\frac{2}{9\pi}\Big)^{1/2} \varepsilon^{-3/2} - (2\pi)^{-1/2}\lambda^2\varepsilon^{-1/2} + \frac{1}{4}\ln(1/\varepsilon) + O(1) \tag{1.3}$$

*and*

$$\operatorname{Var} W_\varepsilon = \Big(\frac{2}{9\pi}\Big)^{1/2} \varepsilon^{-3/2} + O(\varepsilon^{-1}) \sim \mathbb{E}\, W_\varepsilon. \tag{1.4}$$

**Remark 1.3.** It seems likely that $W_\varepsilon$ is almost Poisson distributed, in the sense that its total variation distance to a Poisson distribution with the same mean tends to 0 as $\varepsilon \to 0$, but we leave this as an open problem. If this holds, it would immediately imply asymptotic normality of $W_\varepsilon$.

The main interest in Theorem 1.1 comes from the fact that $Z_\varepsilon$ approximatively describes the large component sizes in $G(n,p)$ for large $n$. We formalize this in the following intuitively obvious result; see Section 5 for a formal verification of the technicalities.

**Proposition 1.4.** *Let $Z_{n\varepsilon} = \sum_{\xi_{ni}\ge\varepsilon} \xi_{ni} = \int_\varepsilon^\infty x\,d\Xi_n(x)$ be the total weight of all points in $\Xi_n$ that are at least $\varepsilon$; thus $Z_{n\varepsilon}$ equals $n^{-2/3}$ times the total size of all components $\ge \varepsilon n^{2/3}$ in $G(n,p)$. For every fixed $\varepsilon > 0$, as $n \to \infty$, $Z_{n\varepsilon} \xrightarrow{\mathrm{d}} Z_\varepsilon$ with convergence of all moments, i.e., for every $q \ge 0$, $\mathbb{E}\, Z_{n\varepsilon}^q \to \mathbb{E}\, Z_\varepsilon^q$.*

*The same holds for $Z'_{n\varepsilon} := \sum_{\xi_{ni}>\varepsilon} \xi_{ni}$.*

*Similarly, if $W_{n\varepsilon} := \#\{i : \xi_{ni} \ge \varepsilon\} = \Xi_n[\varepsilon,\infty)$, the number of components $\ge \varepsilon n^{2/3}$ in $G(n,p)$, then $W_{n\varepsilon} \xrightarrow{\mathrm{d}} W_\varepsilon$, with convergence of all moments.*

We introduce some more notation. Let $X_n(k)$ denote the number of components with $k$ vertices in the random graph $G(n,p)$, and let $Y_n(k) := kX_n(k)$, the total number of vertices in these components. We further define $X_n(I) := \sum_{k\in I} X_n(k)$ and $Y_n(I) := \sum_{k\in I} Y_n(k)$ for an interval $I$. (For simplicity, we omit $p$ from the notation.) Thus $Z_{n\varepsilon} = n^{-2/3} Y_n[\varepsilon n^{2/3},\infty)$, $Z'_{n\varepsilon} = n^{-2/3} Y_n(\varepsilon n^{2/3},\infty)$, and $W_{n\varepsilon} = X_n[\varepsilon n^{2/3},\infty)$. We denote falling factorials by $n^{\underline{k}} := n\cdots(n-k+1)$.

**Remark 1.5.** Although we keep $\lambda$ fixed for simplicity, it is easy to see (e.g. using the monotonicity of $G(n,p)$ in $p$) that the Proposition 1.4 holds also for a sequence $\lambda_n \to \lambda$. Moreover, as a consequence of this and monotonicity, the same holds for the random graph $G(n,m)$ with a deterministic number $m = n/2 + (\lambda + o(1))n^{2/3}/2$ edges.

**Remark 1.6.** The phase transition is, we feel, the most carefully studied, the most subtle, and the most intriguing phenomenon in the evolution of the random graph. Books ([4], [12]) cover many aspects of the phase transition in detail; see also, for SJ
example, [1], [11], [18]. Parametrize $p = n^{-1} + \lambda(n)n^{-4/3}$. When $\lambda(n) \to -\infty$ we are in the subcritical region. The largest component has size $o(n^{2/3})$, the first and second largest components have roughly the same size, and all components are trees or unicyclic. When $\lambda(n) \to +\infty$ we are in the supercritical region. The largest (sometimes called dominant) component has size $\gg n^{2/3}$ and its complexity is larger than any fixed constant. All other components have size $\ll n^{2/3}$ and are either trees or unicyclic. When (the object of our study) $\lambda = \lambda(n)$ is a constant we are in the critical window. It is during that "time" that many "small" components merge to form the dominant component. In both the subcritical and supercritical regions the size of the largest component has an asymptotic value which it achieves with high probability. In the critical window that size has a nontrivial limit distribution. SJ
Indeed the study of the behavior of many natural parameters in the critical window is most challenging.

**Acknowledgement.** We thank David Aldous for interesting discussions.

## 2. First partial proof of Theorem 1.1

If $\mu$ is a probability distribution on the non-negative integers, let $T(\mu)$ denote the (random) total progeny of a Galton–Watson process with offspring distribution $\mu$, starting with one initial particle. (Thus $T(\mu) \in \{1, 2, \dots, \infty\}$.)

**Lemma 2.1.** *Let $\lambda \in (-\infty,\infty)$ and $\varepsilon > 0$ be fixed.*

(i) *The limit* $U_\varepsilon(\lambda) := \lim_{n\to\infty} n^{1/3}\,\mathbb{P}\bigl(T(\mathrm{Po}(1+\lambda n^{-1/3})) \ge \varepsilon n^{2/3}\bigr)$ *exists, and*

$$U_\varepsilon(\lambda) = 2\max(\lambda,0) + \int_\varepsilon^\infty (2\pi)^{-1/2} x^{-3/2} e^{-\lambda^2 x/2}\,dx. \tag{2.1}$$

(ii) *More generally, if* $\lambda_n \to \lambda$, *then*

$$n^{1/3}\,\mathbb{P}\bigl(T(\mathrm{Po}(1+\lambda_n n^{-1/3})) \ge \varepsilon n^{2/3}\bigr) \to U_\varepsilon(\lambda).$$

(iii) *Moreover, for any fixed* $\delta \in \mathbb{R}$ *and any sequence* $\delta_n \to \delta$,

$$n^{1/3}\,\mathbb{P}\bigl(T(\mathrm{Bi}(\lfloor n - \delta_n n^{2/3}\rfloor, n^{-1} + \lambda n^{-4/3})) \ge \varepsilon n^{2/3}\bigr) \to U_\varepsilon(\lambda - \delta).$$

*Proof.* (i): First consider the possibility of an infinite total progeny. By elementary
SJ branching process theory, if $q := \mathbb{P}\bigl(T(\mathrm{Po}(\gamma)) = \infty\bigr)$, then $q = 0$ for $\gamma \le 1$, while $q > 0$ for $\gamma > 1$, and then $1 - q = e^{-q\gamma}$, or $\gamma = -\ln(1-q)/q = 1 + q/2 + O(q^2)$. It follows that if $\gamma \to 1$ with $\gamma > 1$, then $q \to 0$ and $q \sim 2(\gamma - 1)$. Consequently, as $n \to \infty$, for any real $\lambda$,

$$n^{1/3}\,\mathbb{P}\bigl(T(\mathrm{Po}(1+\lambda n^{-1/3})) = \infty\bigr) \to 2\max(\lambda, 0). \tag{2.2}$$

Next, consider a finite total progeny $\ge \varepsilon n^{2/3}$. By Otter [19], see also Pitman [21],

$$\mathbb{P}\bigl(T(\mu) = k\bigr) = \frac{1}{k}\,\mathbb{P}(S_k = k-1), \qquad 1 \le k < \infty, \tag{2.3}$$

where $S_k$ is the sum of $k$ independent random variables with the distribution $\mu$. In particular, for a Poisson distribution, using Stirling's formula,

$$\begin{aligned}\mathbb{P}\bigl(T(\mathrm{Po}(1+\lambda n^{-1/3})) = k\bigr) &= \frac{1}{k}\,\mathbb{P}\bigl(\mathrm{Po}(k(1+\lambda n^{-1/3})) = k-1\bigr)\\ &= \frac{k^{k-1}(1+\lambda n^{-1/3})^{k-1}}{k!} e^{-k(1+\lambda n^{-1/3})}\\ &= (2\pi)^{-1/2} k^{-3/2} (1+\lambda n^{-1/3})^k e^{-k\lambda n^{-1/3}} \bigl(1 + O(|\lambda| n^{-1/3} + k^{-1})\bigr). \end{aligned} \tag{2.4}$$

We have

$$\ln(1+\lambda n^{-1/3}) = \lambda n^{-1/3} - \tfrac12(\lambda n^{-1/3})^2 + O(|\lambda n^{-1/3}|^3).$$

Hence, for $\varepsilon n^{2/3} \le k < a_n := (n^{1/3}|\lambda|^{-1})^{5/2}$ (with $a_n = \infty$ when $\lambda = 0$), (2.4) yields

$$\mathbb{P}\bigl(T(\mathrm{Po}(1+\lambda n^{-1/3})) = k\bigr) = (2\pi)^{-1/2} k^{-3/2} e^{-\frac12 \lambda^2 k n^{-2/3}} \bigl(1 + O(n^{-1/6})\bigr). \tag{2.5}$$

Moreover, assuming that $n$ is so large that $|\lambda n^{-1/3}| < 1/2$,

$$\ln(1+\lambda n^{-1/3}) \le \lambda n^{-1/3} - \tfrac12(\lambda n^{-1/3})^2 + \tfrac23|\lambda n^{-1/3}|^3 \le \lambda n^{-1/3} - \tfrac16\lambda^2 n^{-2/3}$$

and thus, by (2.4),

$$\mathbb{P}\bigl(T(\mathrm{Po}(1+\lambda n^{-1/3})) = k\bigr) = O\Bigl(k^{-3/2} e^{-\frac16\lambda^2 k n^{-2/3}}\Bigr). \tag{2.6}$$

Summing over $k \in [\varepsilon n^{2/3}, a_n)$ we find, by (2.5), (2.6) and dominated convergence,

$$\begin{aligned} n^{1/3}\,&\mathbb{P}\bigl(T(\mathrm{Po}(1+\lambda n^{-1/3})) \in [\varepsilon n^{2/3}, a_n)\bigr)\\ &= n^{1/3}\int_{\lceil \varepsilon n^{2/3}\rceil}^{\lceil a_n\rceil} \mathbb{P}\bigl(T(\mathrm{Po}(1+\lambda n^{-1/3})) = \lfloor x\rfloor\bigr)\,dx\\ &= \int_{n^{-2/3}\lceil \varepsilon n^{2/3}\rceil}^{n^{-2/3}\lceil a_n\rceil} n\,\mathbb{P}\bigl(T(\mathrm{Po}(1+\lambda n^{-1/3})) = \lfloor n^{2/3}x\rfloor\bigr)\,dx\\ &\to \int_\varepsilon^\infty (2\pi)^{-1/2} x^{-3/2} e^{-\lambda^2 x/2}\,dx. \end{aligned} \tag{2.7}$$

Furthermore, the sum over $k \ge a_n$ is exponentially small by (2.6). Hence, the result follows by (2.2) and (2.7).

(ii): By the same proof as (i), or by (i) and monotonicity.

(iii): One could use (2.3) and argue as above, but we will instead use a Poisson approximation. For any $N$ and $p$, we have the bound on the total variation distance

$$d_{TV}\bigl(\mathrm{Bi}(N,p), \mathrm{Po}(Np)\bigr) \le p,$$

see e.g. [2, Theorem 2.M]. Hence, using a maximal coupling of $\mathrm{Bi}(N,p)$ and $\mathrm{Po}(Np)$ in each family, we can couple the Galton–Watson processes with offspring distributions $\mathrm{Bi}(N,p)$ and $\mathrm{Po}(Np)$ such that the probability that they differ before they have reached at least $n^{1/3}$ individuals is at most $n^{1/3}p$; furthermore, conditioned on both reaching $n^{1/3}$ together, and being equal so far, the probability that they differ before they have reached at least $\varepsilon n^{2/3}$ individuals is at most $\varepsilon n^{2/3}p$. Hence, SJ

$$\begin{aligned}\bigl|\mathbb{P}\bigl(T(\mathrm{Bi}(N,p)) \ge \varepsilon n^{2/3}\bigr) - \mathbb{P}\bigl(T(\mathrm{Po}(Np)) \ge \varepsilon n^{2/3}\bigr)\bigr| \\ \le n^{1/3}p + \mathbb{P}\bigl(T(\mathrm{Po}(Np)) \ge n^{1/3}\bigr)\varepsilon n^{2/3}p. \qquad (2.8)\end{aligned}$$

Now, let $N = \lfloor n - \delta n^{2/3} \rfloor$ and $p = n^{-1} + \lambda n^{-4/3}$ and let $n \to \infty$. Then $Np \to 1$, and thus, for each fixed $M$ and $n > M^3$, SJ

$$\mathbb{P}\bigl(T(\mathrm{Po}(Np)) \ge n^{1/3}\bigr) \le \mathbb{P}\bigl(T(\mathrm{Po}(Np)) \ge M\bigr) \to \mathbb{P}\bigl(T(\mathrm{Po}(1)) \ge M\bigr).$$

Since $T(\mathrm{Po}(1))$ is finite a.s., the latter probability tends to 0 as $M \to \infty$, and it follows that $\mathbb{P}\bigl(T(\mathrm{Po}(Np) \ge n^{1/3}\bigr) \to 0$. Consequently, the right hand side of (2.8) is $O(n^{-2/3}) + o(1) \cdot O(n^{-1/3}) = o(n^{-1/3})$. The result follows from (2.8) and (ii), since

$$Np = (n - \delta n^{2/3} + O(1))(n^{-1} + \lambda n^{-4/3}) = 1 + \lambda_n n^{-1/3},$$

with $\lambda_n = \lambda - \delta + O(n^{-1/3}) \to \lambda - \delta$. □

We give alternative formulas for $U_\varepsilon(\lambda)$ defined in Lemma 2.1.

**Lemma 2.2.** *Let $-\infty < \lambda < \infty$ and $\varepsilon > 0$. Then*

$$U_\varepsilon(\lambda) = \Bigl(\frac{2}{\pi}\Bigr)^{1/2}\varepsilon^{-1/2} + \lambda + \int_0^\varepsilon (2\pi)^{-1/2}x^{-3/2}\bigl(1 - e^{-\lambda^2 x/2}\bigr)\,dx \qquad (2.9)$$

$$= \Bigl(\frac{2}{\pi}\Bigr)^{1/2}\varepsilon^{-1/2} + \lambda + (2\pi)^{-1/2}\lambda^2\varepsilon^{1/2} + O(\lambda^4\varepsilon^{3/2}). \qquad (2.10)$$

*Proof.* First note that in the case $\lambda = 0$, (2.1) yields

$$U_\varepsilon(0) = \int_\varepsilon^\infty (2\pi)^{-1/2}x^{-3/2}\,dx = (2/\pi)^{1/2}\varepsilon^{-1/2}. \qquad (2.11)$$

Since $2\max(\lambda, 0) = \lambda + |\lambda|$, (2.1) further yields

$$U_\varepsilon(\lambda) - U_\varepsilon(0) = \lambda + |\lambda| - \int_\varepsilon^\infty (2\pi)^{-1/2}x^{-3/2}\bigl(1 - e^{-\lambda^2 x/2}\bigr)\,dx. \qquad (2.12)$$

Now, for $\lambda \ne 0$, by change of variables and a standard integration by parts,

$$\begin{aligned}\int_0^\infty x^{-3/2}\bigl(1 - e^{-\lambda^2 x/2}\bigr)\,dx &= (\lambda^2/2)^{1/2}\int_0^\infty y^{-3/2}\bigl(1 - e^{-y}\bigr)\,dy \\ &= (\lambda^2/2)^{1/2}\,2\,\Gamma(1/2) = (2\pi)^{1/2}|\lambda|,\end{aligned}$$

and thus (2.12) yields

$$U_\varepsilon(\lambda) = U_\varepsilon(0) + \lambda + \int_0^\varepsilon (2\pi)^{-1/2}x^{-3/2}\bigl(1 - e^{-\lambda^2 x/2}\bigr)\,dx. \qquad (2.13)$$

This proves (2.9), and (2.10) follows by the expansion $1-e^{-\lambda^2 x/2} = \lambda^2 x/2+O(\lambda^4 x^2)$. □

**Remark 2.3.** Expression (2.1) might lead the casual reader to suppose that $\lambda = 0$ was somehow special. The equivalent expression (2.9), however, shows that $U_\varepsilon(\lambda)$ is a smooth function of $\lambda$. This corresponds to the generally held belief that there can be no further refinements of the critical window, that no value of $\lambda$ is more special than any other, and that natural functions vary smoothly with $\lambda$.

Returning to the random graphs, note that given the graph $G(n,p)$, the probability that a random vertex belongs to a component of size at least $\varepsilon n^{2/3}$ is $n^{-1}Y[\varepsilon n^{2/3},\infty) = n^{-1/3}Z_{n\varepsilon}$. Taking expectations we see that $\mathbb{E}\, Z_{n\varepsilon}$ equals $n^{1/3}$ times the probability that a given vertex $v$ belongs to a component of size at least $\varepsilon n^{2/3}$ in $G(n,p)$. We explore the component containg the given vertex by the standard breadth-first search. In this search, we explore first the neighbours of $v$, then their neighbours, and so on, see e.g. [24] or [12, Section 5.2]. When we explore the neighbours of a vertex, we find $\mathrm{Bi}(n-m,p)$ new vertices in the component, where $m$ is the number of vertices found so far. Thus, the process is dominated by a Galton–Watson process with offspring distribution $\mathrm{Bi}(n,p)$, and, if we stop when we reach $\varepsilon n^{2/3}$ vertices, dominates a Galton–Watson process with offspring distribution $\mathrm{Bi}(\lfloor n-\varepsilon n^{2/3}\rfloor, p)$; hence, the probability that we find at least $\varepsilon n^{2/3}$ vertices in the component lies between the probabilities that these Galton–Watson processes have a total progeny of at least $\varepsilon n^{2/3}$. Consequently,

$$\begin{aligned}\mathbb{P}\bigl(T(\mathrm{Bi}(\lfloor n-\varepsilon n^{2/3}\rfloor, n^{-1}+\lambda n^{-4/3})) \ge \varepsilon n^{2/3}\bigr)&\\ \le n^{-1/3}\,\mathbb{E}\, Z_{n\varepsilon} \le \mathbb{P}\bigl(T(\mathrm{Bi}(n, n^{-1}+\lambda n^{-4/3})) \ge \varepsilon n^{2/3}\bigr).&\end{aligned}$$

By Lemma 2.1 and Proposition 1.4, this yields

$$U_\varepsilon(\lambda-\varepsilon) \le \mathbb{E}\, Z_\varepsilon \le U_\varepsilon(\lambda),$$

and (1.1) follows by Lemma 2.2.

It seems more difficult to estimate $\mathrm{Var}\, Z_{n\varepsilon}$ by this method, and we will use another approach in Section 6.

## 3. Complexity

The *complexity* $c(G)$ of a graph $G$ with $v$ vertices and $e$ edges is defined by $c(G) := e - v + 1$. Thus the complexity is 0 for trees, 1 for unicyclic connected graphs, and $\ge 2$ otherwise. We say that a connected graph with complexity $\ge 2$ is *complex.*

We can refine the point processes $\Xi_n$ and $\Xi$ by considering the complexities of the components. We can think of this as giving each point in the processes a label; a point in $\Xi_n$ is labelled by the complexity of the corresponding component. Formally, we can think of the labelled versions, $\Xi_n^*$ and $\Xi^*$, say, as point processes on the space $(0,\infty]\times\mathbb{N}$, or better $(0,\infty]\times\mathbb{N}^*$, where $\mathbb{N} = \{0,1,\dots\}$ and $\mathbb{N}^*$ is the compact space $\mathbb{N}\cup\{\infty\}$. The results by Aldous [1, Corollary 2] and Łuczak, Pittel and Wierman [18] referred to above actually consider the complexity too, and show that $\Xi_n^* \xrightarrow{\mathrm{d}} \Xi^*$ as $n\to\infty$, for a suitable labelling $\Xi^*$ of $\Xi$. Aldous [1, Corollary 2] describes $\Xi^*$ by the process $B^\lambda$ defined above: introduce a process of marks on $(0,\infty)$ that, given $B^\lambda$, is a Poisson process with intensity $B^\lambda(s)\,ds$; then, as said above, the points $\xi_i$ are the lengths of the excursions of $B^\lambda$, and each excursion is labelled with the number of marks inside it. In other words, given $B^\lambda$, each point $\xi_i$ gets a label that has a Poisson distribution whose mean is the area under the corresponding excursion, and different points are labelled independently. We will give another description in Theorem 3.1 below.

We let, for $\ell \geq 0$, $\Xi_n^\ell$ be the subset $\{\xi_{ni} : c(\mathcal{C}_i) = \ell\}$ of $\Xi_n$ of points with labels $\ell$, i.e. the set of scaled sizes of components of $G(n,p)$ with complexity $\ell$. Similarly, let $\Xi^\ell$ be the subset of $\Xi$ of points with labels $\ell$. Since $\Xi_n^* \xrightarrow{\mathrm{d}} \Xi^*$, we have $\Xi_n^\ell \xrightarrow{\mathrm{d}} \Xi^\ell$ for every $\ell$.

Let $C(k,\ell)$ be the number of connected graphs with complexity $\ell$ on $k$ (labelled) vertices (they thus have $k+\ell-1$ edges). Thus $C(k,0)$ is the number of trees, and by Cayley's theorem, $C(k,0) = k^{k-2}$. More generally, Wright [25] proved that for every fixed $\ell$

$$C(k,\ell) \sim w_\ell k^{k+3\ell/2-2} \qquad \text{as } k \to \infty, \tag{3.1}$$

for some constants $w_\ell$, for which Wright [25] gave a recursion formula. (See also [11, §8] and the references there.) We have $w_0 = 1$ and $w_1 = \sqrt{\pi/8}$. It was shown in [24] that

$$w_\ell = \frac{\mathbb{E}\, L^\ell}{\ell!}, \qquad \ell \geq 0 \tag{3.2}$$

where $L$ is the area under a normalized Brownian excursion. If we introduce the moment generating function $\Psi$ of $L$, we thus have

$$\Psi(t) = \mathbb{E}\, e^{tL} = \sum_{\ell=0}^{\infty} w_\ell t^\ell. \tag{3.3}$$

The moments $\mathbb{E}\, L^\ell$ and the moment generating function $\Psi$ had earlier been studied by Louchard [16, 17]. Note that $\Psi(t)$ is finite for all $t > 0$ (and thus (3.3) holds for all complex $t$); indeed, as remarked in [5, Remark 3.1] (where $\xi = 2L$), it follows from the well-known asymptotics for $w_\ell$, see e.g. [11, §8] and [10, Theorem 3.3 and (3.8)], that $\mathbb{E}\, L^\ell \sim \sqrt{18}\, \ell\, (12e)^{-\ell/2} \ell^{\ell/2}$ as $\ell \to \infty$, and thus [5, Lemma 4.1(ii)] implies, cf. [5, Remark 4.9],

$$\Psi(t) \sim \tfrac{1}{2} t^2 e^{t^2/24} \qquad \text{as } t \to +\infty. \tag{3.4}$$

We now can state the result describing $\Xi^*$. For $x > 0$, let $P_x$ be the distribution on $\mathbb{N}$ given by

$$P_x(\ell) = \frac{w_\ell x^{3\ell/2}}{\Psi\big(x^{3/2}\big)}, \qquad \ell \geq 0. \tag{3.5}$$

**Theorem 3.1.** *The point process $\Xi^*$ on $(0,\infty] \times \mathbb{N}^*$ can be obtained from $\Xi$ by independently giving each point $\xi_i \in \Xi$ a random label with the distribution $P_{\xi_i}$.*

*Proof.* Conditioned on the vertex sets of the components of $G(n,p)$, the internal structures of the components are independent. Moreover, a component of order $k$ is distributed as $G(k,p)$ conditioned on being connected. Let $P_{k,p}(\ell)$ be the probability that such a component has complexity $\ell$. The probability that $G(k,p)$ is connected and has complexity $\ell$ is $C(k,\ell) p^{k+\ell-1} (1-p)^{\binom{k}{2}-k-\ell+1}$ and thus

$$P_{k,p}(\ell) = C(k,\ell) \Big(\frac{p}{1-p}\Big)^\ell \Big/ \sum_{\ell=0}^{\infty} C(k,\ell) \Big(\frac{p}{1-p}\Big)^\ell. \tag{3.6}$$

Consequently, the labelled process $\Xi_n^*$ can be obtained from $\Xi_n$ by giving the points $\xi_{ni}$ labels independently, such that the label of a point $x$ has the distribution $P_{xn^{2/3},p}$ given by (3.6).

By Bollobás [4, Theorem V.20], there exists a constant $c > 0$ such that, for all $k$ and $\ell$,

$$C(k,\ell) \le (c/\ell)^{\ell/2} k^{k+3\ell/2-2}. \tag{3.7}$$

Hence, if $k \le bn^{2/3}$ for some fixed $b$, and $n$ is so large that $p/(1-p) < 2/n$,

$$P_{k,p}(\ell) \le \frac{C(k,\ell)}{C(k,0)}\Big(\frac{p}{1-p}\Big)^{\ell} \le \Big(\frac{4cb^3}{\ell}\Big)^{\ell/2}. \tag{3.8}$$

Consider a sequence $k = k(n)$ such that $kn^{-2/3} \to x$ for some $x > 0$. By (3.1), for every $\ell \ge 0$, as $n \to \infty$,

$$\frac{C(k,\ell)}{C(k,0)}\Big(\frac{p}{1-p}\Big)^{\ell} = w_\ell k^{3\ell/2} n^{-\ell}\big(1+o(1)\big) \to w_\ell x^{3\ell/2}. \tag{3.9}$$

Together with (3.8), this implies by dominated convergence

$$\sum_{\ell=0}^{\infty} \frac{C(k,\ell)}{C(k,0)}\Big(\frac{p}{1-p}\Big)^{\ell} \to \sum_{\ell=0}^{\infty} w_\ell x^{3\ell/2} = \Psi\big(x^{3/2}\big). \tag{3.10}$$

Consequently, from (3.6), (3.9), (3.10) and (3.5), for $n \to \infty$ and every fixed $\ell$,

$$P_{k,p}(\ell) \to P_x(\ell).$$

Thus, the distribution $P_{k,p}$ converges to $P_x$.

Let $\Xi'$ be the labelled point process constructed in the statement of the theorem. It follows from Lemma A.2 and the Skorohod coupling theorem, see e.g. [14, Theorem 4.30], that we may assume $\Xi_n$ and $\Xi$ to be coupled such that $\xi_{ni} \to \xi_i$ a.s. for every $i$. By the description of $\Xi^*$ above and the convergence of $P_{k,p}$ to $P_x$ when $kn^{-2/3} \to x$, it follows that we may couple also the labels such that $\Xi_n^* \overset{\text{a.s.}}{\to} \Xi'$. Hence $\Xi_n^* \overset{\text{d}}{\to} \Xi'$, and thus $\Xi' \overset{\text{d}}{=} \Xi^*$. □

## 4. Intensity

Let, changing the notation slightly from [18], $X_n(k;\ell)$ denote the number of components with $k$ vertices and complexity $\ell$ in the random graph $G(n,p)$, and let $Y_n(k;\ell) := kX_n(k;\ell)$, the number of vertices in these components. We further define, for an interval $I$, $X_n(I;\ell) := \sum_{k\in I} X_n(k;\ell)$, $X_n(k;\ge\ell) := \sum_{j=\ell}^{\infty} X_n(k;j)$ and $X_n(I;\ge\ell) := \sum_{j=\ell}^{\infty} X_n(I;j)$, and similary for $Y$. Thus, for example, $X_n(k) = \sum_j X_n(k;j) = X_n(k;\ge 0)$.

Consider now a fixed $\ell \geq 0$ and $k \leq Cn^{2/3}$ for an arbitrary constant $C$. Then, by well-known calculations, uniformly for all such $k$,

$$\begin{aligned}
\mathbb{E}\, X_n(k,\ell) &= \binom{n}{k} C(k,\ell) p^{k+\ell-1} (1-p)^{(n-k)k+\binom{k}{2}-k-\ell+1} \\
&= \frac{n^k}{k!} \exp\Bigl(-\frac{k^2}{2n} - \frac{k^3}{6n^2} + O\Bigl(\frac{k}{n}\Bigr)\Bigr) C(k,\ell) n^{1-k-\ell} (1+\lambda n^{-1/3})^{k+\ell-1} \\
&\qquad\qquad\qquad\qquad \times (1-p)^{nk-k^2/2-3k/2-\ell+1} \\
&= n^{1-\ell} \frac{C(k,\ell)}{k!} \exp\Bigl(-\frac{k^2}{2n} - \frac{k^3}{6n^2} + (k+\ell-1)\lambda n^{-1/3} - \frac{1}{2} k \lambda^2 n^{-2/3} \\
&\qquad\qquad - k - k\lambda n^{-1/3} + \frac{k^2}{2n} + \frac{1}{2} k^2 \lambda n^{-4/3} + O\Bigl(\frac{k}{n}\Bigr) + O(n^{-2/3})\Bigr) \\
&= n^{1-\ell} \frac{C(k,\ell)}{k!} \exp\Bigl(-k - \frac{k^3}{6n^2} - \frac{1}{2} k \lambda^2 n^{-2/3} + \frac{1}{2} k^2 \lambda n^{-4/3} + (\ell-1)\lambda n^{-1/3} \\
&\qquad\qquad + O\Bigl(\frac{k}{n}\Bigr) + O(n^{-2/3})\Bigr) \\
&= n^{1-\ell} \frac{C(k,\ell)}{k!} e^{-k} \exp\bigl(-F(kn^{-2/3},\lambda)\bigr) \Bigl(1 + (\ell-1)\lambda n^{-1/3} + O\Bigl(\frac{k}{n}\Bigr) + O(n^{-2/3})\Bigr)
\end{aligned} \tag{4.1}$$

where

$$F(x,\lambda) := \frac{1}{6}x^3 - \frac{1}{2}x^2\lambda + \frac{1}{2}x\lambda^2 = \frac{(x-\lambda)^3 + \lambda^3}{6}. \tag{4.2}$$

Note that

$$F(x,\lambda) = \frac{x^3}{24} + x\frac{(x-2\lambda)^2}{8} \geq \frac{x^3}{24} \geq 0 \tag{4.3}$$

for all $x \geq 0$ and $-\infty < \lambda < \infty$.

In our first application of (4.1), assume $0 < a < b < \infty$ and consider only $k \in [an^{2/3}, bn^{2/3}]$. For such $k$ and fixed $\ell$, (4.1) gives, by (3.1) and Stirling's formula,

$$\begin{aligned}
\mathbb{E}\, X_n(k,\ell) &\sim n^{1-\ell} w_\ell (2\pi)^{-1/2} k^{3\ell/2-5/2} e^{-F(kn^{-2/3},\lambda)} \\
&= (2\pi)^{-1/2} w_\ell \Bigl(\frac{k^{3/2}}{n}\Bigr)^{\ell-1} e^{-F(kn^{-2/3},\lambda)} k^{-1},
\end{aligned}$$

and summing over $k$ we obtain, as $n \to \infty$,

$$\mathbb{E}\bigl(\Xi_n^\ell[a,b]\bigr) = \mathbb{E} \sum_{k=an^{2/3}}^{bn^{2/3}} X_n(k,\ell) \to (2\pi)^{-1/2} w_\ell \int_a^b \bigl(x^{3/2}\bigr)^{\ell-1} e^{-F(x,\lambda)} \,\frac{dx}{x}. \tag{4.4}$$

Since $\Xi_n^\ell \overset{\mathrm{d}}{\to} \Xi^\ell$ we have, by Lemma A.1, $\Xi_n^\ell[a,b] \overset{\mathrm{d}}{\to} \Xi^\ell[a,b]$ whenever $a$ and $b$ are continuity points of $\Xi^\ell$. In this case, by Fatou's lemma, $\mathbb{E}\,\Xi^\ell[a,b]$ is at most the right hand side of (4.4).

For any $a \in (0,\infty)$, $a \pm \varepsilon$ are continuity points of $\Xi^\ell$ for all but at most countably many $\varepsilon \in (0,a)$, and for such $\varepsilon$ we thus obtain a formula for $\mathbb{E}(\Xi^\ell[a-\varepsilon, a+\varepsilon])$ and thus an upper bound of $\mathbb{E}(\Xi^\ell\{a\})$. Letting $\varepsilon \to 0$ through such $\varepsilon$, we see that $\mathbb{E}(\Xi^\ell\{a\}) = 0$, so every point is a continuity point. Consequently, $\Xi_n^\ell[a,b] \overset{\mathrm{d}}{\to} \Xi^\ell[a,b]$ whenever $0 < a < b \leq \infty$. Summing over all $\ell$, we see that every point is a continuity point of $\Xi$ too, and thus $\Xi_n[a,b] \overset{\mathrm{d}}{\to} \Xi[a,b]$ whenever $0 < a < b \leq \infty$.

To prove convergence of the expectations, we verify uniform integrability by considering second moments. (See also the more general Lemma 5.1 below; we will give a more elementary argument here, which in any case will be needed later.)

For simplicity, fix $\ell$ and write $E_k := \mathbb{E}\, X_n(k,\ell)$. Further, let $E_{k,j}$ denote the expected number of ordered pairs of distinct components of complexity $\ell$, of orders $k$ and $j$, respectively, in $G(n,p)$. Thus, if $k \neq j$ then $E_{k,j} = \mathbb{E}\big(X_n(k;\ell)X_n(j;\ell)\big)$, while

$$E_{k,k} = \mathbb{E}\big(X_n(k;\ell)(X_n(k;\ell)-1)\big) = \mathbb{E}\big(X_n(k;\ell)\big)^2 - E_k.$$

Consequently,

$$\mathbb{E}\big(X_n\big([an^{2/3}, bn^{2/3}];\ell\big)\big)^2 = \mathbb{E}\Big(\sum_{k=an^{2/3}}^{bn^{2/3}} X_n(k;\ell)\Big)^2 = \sum_{k=an^{2/3}}^{bn^{2/3}} \sum_{j=an^{2/3}}^{bn^{2/3}} E_{k,j} + \sum_{k=an^{2/3}}^{bn^{2/3}} E_k. \tag{4.5}$$

We have, cf. (4.1), by simple calculations, assuming, say, $k+j \le n/2$,

$$\begin{aligned}
E_{k,j} &= \binom{n}{k+j}\binom{k+j}{k} C(k,\ell)C(j,\ell) p^{k+\ell-1+j+\ell-1}(1-p)^{n(k+j)-(k+j)^2/2-3(k+j)/2-2\ell+2}\\
&= \frac{n^{\underline{k+j}}}{n^{\underline{k}}n^{\underline{j}}}(1-p)^{-kj}E_kE_j\\
&= E_kE_j \exp\Big(\sum_{i=0}^{j-1} \ln\Big(1-\frac{k}{n-i}\Big) - kj\ln(1-p)\Big)\\
&= E_kE_j \exp\Big(\sum_{i=0}^{j-1} -\Big(\frac{k}{n}+\frac{ki}{n^2}+\frac{k^2}{2n^2}+O\Big(\frac{ki^2+k^2i+k^3}{n^3}\Big)\Big) + kjp + O\Big(\frac{kj}{n^2}\Big)\Big)\\
&= E_kE_j \exp\Big(\lambda kjn^{-4/3} - \frac{k^2j+kj^2}{2n^2} + O\Big(\frac{kj}{n^2}+\frac{kj(k^2+j^2)}{n^3}\Big)\Big).
\end{aligned} \tag{4.6}$$

In particular, for $k,j \le Cn^{2/3}$, $E_{k,j} = O(E_kE_j)$, and (4.5) implies, for fixed $\ell$, $a$ and $b$, with $0<a<b<\infty$, recalling that $\mathbb{E}\,\Xi_n^\ell[a,b] = O(1)$ by (4.4),

$$\begin{aligned}
\mathbb{E}\big(\Xi_n^\ell[a,b]\big)^2 &= \mathbb{E}\big(X_n\big([an^{2/3}, bn^{2/3}];\ell\big)\big)^2\\
&= O\Big(\big(\mathbb{E}\, X_n\big([an^{2/3}, bn^{2/3}];\ell\big)\big)^2 + \mathbb{E}\, X_n\big([an^{2/3}, bn^{2/3}];\ell\big)\Big)\\
&= O\Big(\big(\mathbb{E}\,\Xi_n^\ell[a,b]\big)^2 + \mathbb{E}\,\Xi_n^\ell[a,b]\Big) = O(1).
\end{aligned}$$

Thus, the random variables $\Xi_n^\ell[a,b]$ are uniformly integrable, and $\Xi_n^\ell[a,b] \xrightarrow{\mathrm{d}} \Xi^\ell[a,b]$ implies $\mathbb{E}\,\Xi_n^\ell[a,b] \to \mathbb{E}\,\Xi^\ell[a,b]$, see e.g. [6, Theorems 5.4.2 and 5.5.9]. Consequently, $\mathbb{E}\,\Xi^\ell[a,b]$ equals the right hand side of (4.4). This leads to the following result. Recall that $\Psi$ denotes the moment generating function (3.3) of the Brownian excursion area.

**Theorem 4.1.** *The point process $\Xi^\ell$ has intensity $\Lambda_\ell := (2\pi)^{-1/2}w_\ell x^{3\ell/2-5/2}e^{-F(x,\lambda)}$ on $(0,\infty)$. Their sum $\Xi$ has the intensity, for $0<x<\infty$,*

$$\Lambda(x) = \Lambda^{(\lambda)}(x) := \sum_{\ell=0}^{\infty} \Lambda_\ell(x) = (2\pi)^{-1/2}x^{-5/2}\Psi(x^{3/2})e^{-F(x,\lambda)}. \tag{4.7}$$

*Proof.* We have shown that $\mathbb{E}\,\Xi^\ell[a,b] = \int_a^b \Lambda_\ell(x)\,dx$ when $0<a<b<\infty$, which by definition shows that $\Lambda_\ell(x)$ is the intensity of $\Xi^\ell$. The second part follows by summing over $\ell$. ◻

**Corollary 4.2.**

$$\mathbb{E}\, Z_\varepsilon = \int_\varepsilon^\infty (2\pi)^{-1/2} x^{-3/2} \Psi(x^{3/2}) e^{-F(x,\lambda)}\, dx. \tag{4.8}$$

*Proof.* $\mathbb{E}\, Z_\varepsilon = \int_\varepsilon^\infty x\Lambda(x)\, dx$. □

We already know that the expectation in (4.8) is finite; that the integral converges follows also by (3.4) and (4.3), which imply that $\Lambda(x)$ decreases exponentially as $x \to \infty$. Note further that the intensity $\Lambda^{(\lambda)}(x) \sim (2\pi)^{-1/2} x^{-5/2}$ as $x \to 0$, for every $\lambda$.

**Remark 4.3.** The intensity $\Lambda_\ell$ has a finite integral over $(0,\infty)$ precisely when the exponent $3\ell/2 - 5/2 > -1$. Thus, for $\ell = 0$ and $\ell = 1$, $\Xi^\ell$ has an infinite expected number of points; indeed, it is easily shown from (7.1) and (7.2) that $\Xi^\ell$ a.s. has an infinite number of points. On the other hand, for any $\ell \geq 2$, $\Xi^\ell$ has a finite number of points. Further, $\sum_{\ell \geq 2} \Xi^\ell$, the point process for the complex components, has a finite number of points. One may view this in an evolutionary way. Roughly speaking, when $\lambda$ is large negative complex components have not yet formed. When $\lambda$ is large positive a "dominant component" will have formed which is complex. But there will not usually be other complex components as components get "sucked into" the dominant component before becoming complex. In [11] it is shown that with probability converging to $5\pi/18 \approx 0.87$ there is never more than one complex component in the entire evolution of the random graph.

**Remark 4.4.** Considering the difference $\mathbb{E}\, Z_\varepsilon^{(\lambda)} - \mathbb{E}\, Z_\varepsilon^{(0)}$ and letting $\varepsilon \to 0$, we find from Theorem 1.1 and Corollary 4.2 the identity

$$\int_0^\infty (2\pi)^{-1/2} x^{-3/2} \Psi(x^{3/2}) \Big( e^{-F(x,\lambda)} - e^{-F(x,0)} \Big)\, dx = \int_0^\infty x\big(\Lambda^{(\lambda)}(x) - \Lambda^{(0)}(x)\big)\, dx = \lambda. \tag{4.9}$$

Differentiating with respect to $\lambda$ we further find $-\int_0^\infty x \frac{\partial F}{\partial \lambda}(x,\lambda) \Lambda^{(\lambda)}(x)\, dx = 1$ or

$$\int_0^\infty x^2 (x - 2\lambda) \Lambda^{(\lambda)}(x)\, dx = 2,$$

and thus $\mathbb{E} \sum_i \xi_i^3 = 2 + 2\lambda\, \mathbb{E} \sum_i \xi_i^2$.

**Remark 4.5.** We similarly find expressions for the expectations of sums of all points in $\Xi^*$ with a given label. For example, for label 1, corresponding to unicyclic components, we obtain for the expectation of the total weight of $\Xi^1$

$$\int_0^\infty x\Lambda_1(x)\, dx = \tfrac{1}{4} \int_0^\infty e^{-F(x,\lambda)}\, dx. \tag{4.10}$$

Thus, cf. Remark 4.3, the $\Xi^1$ process has an infinite number of points with finite sum. See also [18, Lemma 2.2], which implies both (4.10) and

$$\int_0^\infty x\big(\Lambda^{(\lambda)}(x) - \Lambda_0^{(\lambda)}(x)\big)\, dx = (2\pi)^{-1/2} \int_0^\infty x^{-3/2} \big(1 - e^{-F(x,\lambda)}\big)\, dx + \lambda,$$

which indeed also easily follows from (1.1) and (4.8). The expectation of the sum of all points with label at least 2 (corresponding to the total size of the complex components in $G(n,p)$) is $\int_0^\infty x\big(\Lambda^{(\lambda)}(x) - \Lambda_0^{(\lambda)}(x) - \Lambda_1^{(\lambda)}(x)\big)\, dx$; an evaluation in terms of hypergeometric functions is given in [11, (15.12)].

## 5. An estimate for $G(n,p)$

We prove in this section an estimate for the components of $G(n,p)$ that we will need. This estimate is known, at least in principle, but we do not know any precise reference. Recall that we keep $\lambda$ fixed.

**Lemma 5.1.** *For any fixed $\varepsilon > 0$ and integer $q \ge 0$,*

$$\mathbb{E}\big((X_n[\varepsilon n^{2/3},\infty))^q\big) = O(1), \qquad \mathbb{E}\big((Y_n[\varepsilon n^{2/3},\infty))^q\big) = O\big(n^{2q/3}\big).$$

*In other words, $W_{n\varepsilon}$, $Z_{n\varepsilon}$ and $Z'_{n\varepsilon}$ have moments that are bounded, uniformly in $n$.*

*Proof.* It suffices to prove the result for $Y_n$, since $Y_n[\varepsilon n^{2/3},\infty) \ge \varepsilon n^{2/3} X_n[\varepsilon n^{2/3},\infty)$.

Let us begin with the complex components; in this case we do not need a lower bound on the size of the components. (This is not surprising, since typically there are no small complex components.) Let $n_c(G)$ denote the number of vertices in complex components of a graph $G$. Thus $n_c\big(G(n,p)\big) = Y_n\big([1,\infty),\ge 2\big)$.

The *excess* of a graph, as defined in [11, §13], equals the complexity minus the number of complex components. (Thus, a component of complexity $\ell$ contributes $\max(\ell-1,0)$.) We first claim the following estimate, where we momentarily consider the random graph $G(n,m)$ with a fixed number $m$ of edges.

**Lemma 5.2.** *There exists $\eta > 0$ such that if $q \ge 0$ and $\mu \in \mathbb{R}$ are fixed, and $m = \lfloor \frac{n}{2}(1+\mu n^{-1/3})\rfloor$, then*

$$\mathbb{E}\Big(n_c\big(G(n,m)\big)^q\,\mathbf{1}\big[\operatorname{excess}\big(G(n,m)\big) = r\big]\Big) = O\big(n^{2q/3} r^q e^{-\eta r}\big), \tag{5.1}$$

*uniformly in $n$ and $r \ge 1$. Consequently, still for fixed $q$ and $\mu$,*

$$\mathbb{E}\big(n_c\big(G(n,m)\big)^q\big) = O\big(n^{2q/3}\big). \tag{5.2}$$

*Proof.* The case $q = 0$ of (5.1) is a special case of [11, Lemma 5] (with $d = 0$). Similarly, the case $q = 1$ is given on [11, pages 299–300], in connection with a detailed study of $\mathbb{E}\big(n_c(G(n,m))\big)$; the proof is not given in detail in [11], but as remarked there, the result follows by a simple modification of the (not so simple) proof of Lemma 5 in [11], and the same is true for general $q \ge 1$. More precisely, we may as in [11] assume $\mu \ge 1$, and we take $d = 0$. To incorporate $n_c\big(G(n,m)\big)^q$, we replace $E_r$ in the first line on page 296 of [11] by $\vartheta^q E_r$, where $\vartheta = z\frac{d}{dz}$. We use the upper bound in [11, (15.2)], and observe that, for $k \ge 1$, $\vartheta(T/(1-T)^k) \le kT/(1-T)^{k+2}$. Hence this replacement gives us a factor $O(r^q)$ and replaces $(1-T)^{-3r}$ by $(1-T)^{-3r-2q}$; we note that the latter gives the same result in the second line on page 296 as if we replace $d$ by $-2q$, except in $e_{rd} = e_{r0} = e_r$, and the rest of the proof is exactly as in [11]. (Since we take $\mu$ fixed, the case $r \le \mu^3$ is only a finite number of cases, and does not have to be stated separately as in [11].)

We then obtain (5.2) by summing (5.1) over $r \ge 1$. $\square$

Returning to the proof of Lemma 5.1, choose $\mu > \lambda$ and observe that by a standard Chernoff estimate, see for example [12, Theorem 2.1], the probability that $G(n,p)$ has more than $m$ edges is $O(e^{-\delta n^{1/3}})$ for some $\delta > 0$. Since $n_c$ is monotone if we add edges, any coupling of $G(n,p)$ and $G(n,m)$ thus gives, for fixed $q \ge 0$,

$$\begin{aligned}\mathbb{E}\Big(Y_n\big([1,\infty),\ge 2\big)^q\Big) &= \mathbb{E}\big(n_c\big(G(n,p)\big)^q\big) \le \mathbb{E}\big(n_c\big(G(n,m)\big)^q\big) + O\big(n^q e^{-\delta n^{1/3}}\big)\\ &= O\big(n^{2q/3}\big).\end{aligned} \tag{5.3}$$

For components of complexity 0 or 1, i.e. trees and unicyclic components, it is possible to argue as for the second moment in Section 4, but we will instead use a trick together with the result just proved.

Let $P(n;k_1,\dots,k_j;\ell)$ be the probability that $j$ given disjoint subsets of the vertex set of $G(n,p)$, with sizes $k_1,\dots,k_j$ respectively, all are the vertex sets of components with complexities $\ell$. Thus, with $k=k_1+\dots+k_j$,

$$P(n;k_1,\dots,k_j;\ell)=(1-p)^{nk-k^2/2-3k/2-j\ell+j}\prod_{i=1}^{j}C(k_i,\ell)p^{k_i+\ell-1}. \tag{5.4}$$

It is easily seen that for any integer $q\ge 1$,

$$\mathbb{E}\big(Y_n\big([A,\infty);\ell\big)^q\big)=\sum_{j=1}^{q}\sum_{k_1,\dots,k_j\ge A}c(n;q;j;k_1,\dots,k_j)P(n;k_1,\dots,k_j;\ell), \tag{5.5}$$

for some combinatorial coefficients $c(n;q;j;k_1,\dots,k_j)$ not depending on $\ell$. For fixed $\ell$ and $\ell'$, we have by (5.4) and (3.1),

$$\frac{P(n;k_1,\dots,k_j;\ell')}{P(n;k_1,\dots,k_j;\ell)}=\prod_{i=1}^{j}\frac{C(k_i,\ell')}{C(k_i,\ell)}\Big(\frac{p}{1-p}\Big)^{\ell'-\ell}=\Theta\Big(\prod_{i=1}^{j}\frac{k_i^{3(\ell'-\ell)/2}}{n^{\ell'-\ell}}\Big).$$

Hence, if $\ell'\le\ell$ and $k_i\ge\varepsilon n^{2/3}$, we have $P(n;k_1,\dots,k_j;\ell')=O\big(P(n;k_1,\dots,k_j;\ell)\big)$ (recall that $\varepsilon$ is fixed), and (5.5) yields

$$\mathbb{E}\big(Y_n\big([\varepsilon n^{2/3},\infty);\ell'\big)^q\big)=O\Big(\mathbb{E}\big(Y_n\big([\varepsilon n^{2/3},\infty);\ell\big)^q\big)\Big)$$

We apply this with $\ell'=0$ and 1 and $\ell=2$, and obtain from (5.3) the required estimates for $\mathbb{E}\big(Y_n\big([\varepsilon n^{2/3},\infty);0\big)^q\big)$ and $\mathbb{E}\big(Y_n\big([\varepsilon n^{2/3},\infty);1\big)^q\big)$, which together with (5.3) complete the proof. $\square$

Before we proceed, we point out a simple consequence. Peres [20] recently gave a simple proof (with an explicit bound) of the case $q=2$; this case is equivalent to $\mathbb{E}\,|\mathcal{C}(v)|=O\big(n^{1/3}\big)$ where $\mathcal{C}(v)$ is the component containing a given (or random) vertex $v$.

**Corollary 5.3.** *Let $q>3/2$. Then $\mathbb{E}\sum_i\xi_{ni}^q=O(1)$; equivalently, for $G(n,p)$, $\mathbb{E}\sum_i|\mathcal{C}_i|^q=O\big(n^{2q/3}\big)$ for any $q>3/2$.*

*Proof.* First, $\sum_{i:|\mathcal{C}_i|>n^{2/3}}|\mathcal{C}_i|^q\le Y_n[n^{2/3},\infty)^q$, whose mean is $O\big(n^{2q/3}\big)$ by Lemma 5.1.

Similarly, the sum over the complex components has expectation $O\big(n^{2q/3}\big)$ by (5.3). It thus remains only to consider components of size at most $n^{2/3}$ with complexity 0 or 1. Let $t_k,u_k$ denote respectively the expected number of trees and unicyclic components in $G(n,p)$. The corresponding sum has expectation $\sum_{k=1}^{n^{2/3}}k^q(t_k+u_k)$, SJ
which is $O\big(n^{2q/3}\big)$ by (6.2) and (6.7). $\square$

*Proof of Proposition* 1.4. It is an easy consequence of Lemma A.2 that the mappings $\{\xi_i\}_i\mapsto\sum_{\xi_i\ge\varepsilon}\xi_i$, $\{\xi_i\}_i\mapsto\sum_{\xi_i>\varepsilon}\xi_i$ and $\{\xi_i\}_i\mapsto\#\{i:\xi_i\ge\varepsilon\}$ are measurable on $\mathfrak{N}(0,\infty]$ and continuous at every $\{\xi_i\}$ such that $\varepsilon\notin\{\xi_i\}$.

Since $\Xi_n\xrightarrow{\mathrm{d}}\Xi$ and $\mathbb{P}(\varepsilon\in\Xi)=0$ by Theorem 4.1, the results $Z_{n\varepsilon}\xrightarrow{\mathrm{d}}Z_\varepsilon$, $Z'_{n\varepsilon}\xrightarrow{\mathrm{d}}Z_\varepsilon$ and $W_{n\varepsilon}\xrightarrow{\mathrm{d}}W_\varepsilon$ follow by the continuous mapping theorem, see e.g. [3, Theorem 5.1].

Since the moments of $Z_{n\varepsilon}$, $Z'_{n\varepsilon}$ and $W_{n\varepsilon}$ are bounded uniformly in $n$ by Lemma 5.1, this further implies convergence of all moments, see e.g. [6, Theorems 5.4.2 and 5.5.9]. ◻

## 6. Second partial proof of Theorem 1.1

In this proof we do the calculations with the small components, and consider complexities 0 and 1 separately. Let throughout $0 < \varepsilon < 1$.

Consider first the tree components. Let $t_k = \mathbb{E}\, X_n(k;0)$ be the expected number of tree components of order $k$. By (4.1), for $k \le n^{2/3}$,

$$t_k = n\frac{k^{k-2}}{k!}e^{-k}\exp\bigl(-F(kn^{-2/3},\lambda)\bigr)\Bigl(1-\lambda n^{-1/3}+O\Bigl(\frac{k}{n}\Bigr)+O(n^{-2/3})\Bigr). \tag{6.1}$$

In particular, with $t_k^* := nk^{k-2}e^{-k}/k!$,

$$t_k = O(t_k^*) = O\Bigl(n\frac{k^{k-2}}{k!}e^{-k}\Bigr) = O\Bigl(\frac{n}{k^{5/2}}\Bigr). \tag{6.2}$$

Note further that, for any fixed real $\alpha$ and all $\varepsilon > 0$,

$$\sum_{k=1}^{\varepsilon n^{2/3}} k^\alpha t_k^* = \sum_{k=1}^{\varepsilon n^{2/3}} O\bigl(nk^{\alpha-5/2}\bigr) = \begin{cases} O\bigl(n(\varepsilon n^{2/3})^{\alpha-3/2}\bigr) = O\bigl(n^{2\alpha/3}\varepsilon^{\alpha-3/2}\bigr), & \alpha > 3/2,\\ O(n), & \alpha < 3/2. \end{cases} \tag{6.3}$$

By (6.1) and (6.3) we obtain,

$$\begin{aligned}\mathbb{E}\, Y([1,\varepsilon n^{2/3}],0) &= \sum_{k=1}^{\varepsilon n^{2/3}} kt_k = n\sum_{k=1}^{\varepsilon n^{2/3}}\frac{k^{k-1}}{k!}e^{-k}\bigl(1-\lambda n^{-1/3}\bigr)\\ &\qquad + O\Bigl(\sum_{k=1}^{\varepsilon n^{2/3}} kt_k^*\bigl(F(kn^{-2/3},\lambda)+kn^{-1}+n^{-2/3}\bigr)\Bigr)\\ &= \bigl(n-\lambda n^{2/3}\bigr)\sum_{k=1}^{\varepsilon n^{2/3}}\frac{k^{k-1}}{k!}e^{-k} + O\Bigl(\sum_{k=1}^{\varepsilon n^{2/3}} t_k^*(k^4n^{-2}+k^2n^{-2/3})\Bigr)\\ &= \bigl(n-\lambda n^{2/3}\bigr)\sum_{k=1}^{\varepsilon n^{2/3}}\frac{k^{k-1}}{k!}e^{-k} + O\Bigl(\varepsilon^{1/2}n^{2/3}\Bigr).\end{aligned} \tag{6.4}$$

Moreover, using the fact that $\sum_1^\infty \frac{k^{k-1}}{k!}e^{-k} = 1$, and Stirling's formula,

$$\begin{aligned}\sum_{k=1}^{\varepsilon n^{2/3}}\frac{k^{k-1}}{k!}e^{-k} &= 1 - \sum_{k>\varepsilon n^{2/3}}\frac{k^{k-1}}{k!}e^{-k}\\ &= 1-\sum_{k>\varepsilon n^{2/3}}(2\pi k^3)^{-1/2} + O\Bigl(\sum_{k>\varepsilon n^{2/3}} k^{-5/2}\Bigr)\\ &= 1-\int_{\varepsilon n^{2/3}}^\infty (2\pi x^3)^{-1/2}\,dx + O(\varepsilon^{-3/2}n^{-1})\\ &= 1-\sqrt{\tfrac{2}{\pi}}\varepsilon^{-1/2}n^{-1/3}+O(\varepsilon^{-3/2}n^{-1}).\end{aligned} \tag{6.5}$$

Consequently, combining (6.4) and (6.5),

$$\begin{aligned}&\mathbb{E}\,Y_n\big([1,\varepsilon n^{2/3}];0\big)\\&\qquad= n-\sqrt{\tfrac{2}{\pi}}\varepsilon^{-1/2}n^{2/3}-\lambda n^{2/3}+O\Big(\varepsilon^{1/2}n^{2/3}+\varepsilon^{-3/2}+\varepsilon^{-1/2}n^{1/3}\Big).\end{aligned}\tag{6.6}$$

Next, let $u_k=\mathbb{E}\,X_n(k;1)$ be the expected number of unicyclic components of order $k$. We have, cf. (4.1) and (3.1),

$$\begin{aligned}u_k&=\binom{n}{k}C(k,1)p^k(1-p)^{(n-k)k+\binom{k}{2}-k}\\&=\frac{C(k,1)}{k^{k-2}}p(1-p)^{-1}t_k\\&=O\Big(n^{-1}k^{3/2}t_k\Big),\end{aligned}\tag{6.7}$$

and thus, by (6.3),

$$\mathbb{E}\,Y_n\big([1,\varepsilon n^{2/3}];1\big)=\sum_{k=1}^{\varepsilon n^{2/3}}ku_k=O\Big(n^{-1}\sum_{k=1}^{\varepsilon n^{2/3}}k^{5/2}t_k\Big)=O(n^{2/3}\varepsilon).\tag{6.8}$$

For complex components we use the well-known fact that $\mathbb{E}\,X_n\big([1,\infty);\geq 2\big)$ is bounded; see the stronger result in [7], [12, Theorem 5.8(i)]. (As a bound we can take 1.2, say, at least for large $n$, and possibly 1, as conjectured in [18].) Hence,

$$\mathbb{E}\,Y_n\big([1,\varepsilon n^{2/3}];\geq 2\big)\leq\varepsilon n^{2/3}\,\mathbb{E}\,X_n\big([1,\varepsilon n^{2/3}];\geq 2\big)=O(n^{2/3}\varepsilon).\tag{6.9}$$

Adding (6.6), (6.8) and (6.9), we find, since the sum of all component sizes $Y_n([1,\infty))=n$,

$$\begin{aligned}\mathbb{E}\,Z'_{n\varepsilon}&=n^{-2/3}\,\mathbb{E}\,Y_n\big(\varepsilon n^{2/3},\infty\big)=n^{-2/3}\big(n-\mathbb{E}\,Y_n\big([1,\varepsilon n^{2/3}]\big)\big)\\&=\sqrt{\tfrac{2}{\pi}}\varepsilon^{-1/2}+\lambda+O\Big(\varepsilon^{1/2}+\varepsilon^{-3/2}n^{-2/3}+\varepsilon^{-1/2}n^{-1/3}\Big).\end{aligned}$$

Thus, letting $n\to\infty$, by Proposition 1.4,

$$\mathbb{E}\,Z_\varepsilon=\sqrt{\tfrac{2}{\pi}}\varepsilon^{-1/2}+\lambda+O\Big(\varepsilon^{1/2}\Big),$$

which is (1.1) with the weaker error term $O(\varepsilon^{1/2})$.

Next, consider the variance of $Y_n\big([1,\varepsilon n^{2/3}];0\big)$. Similarly to (4.5) we have, with $\ell=0$, and $E_k=t_k$,

$$\mathbb{E}\big(Y_n\big([1,\varepsilon n^{2/3}];\ell\big)\big)^2=\mathbb{E}\Big(\sum_{k=1}^{\varepsilon n^{2/3}}kX_n(k;\ell)\Big)^2=\sum_{k=1}^{\varepsilon n^{2/3}}\sum_{j=1}^{\varepsilon n^{2/3}}kjE_{k,j}+\sum_{k=1}^{\varepsilon n^{2/3}}k^2E_k.\tag{6.10}$$

Hence, using (4.6) and letting $A_2 := \sum_{k=1}^{\varepsilon n^{2/3}} k^2 t_k = O\big(n^{4/3}\varepsilon^{1/2}\big)$, by (6.3),

$$\begin{aligned}
\operatorname{Var}\big(Y_n\big([1,\varepsilon n^{2/3}];0\big)\big) &= \sum_{k=1}^{\varepsilon n^{2/3}} k^2 t_k + \sum_{k=1}^{\varepsilon n^{2/3}}\sum_{j=1}^{\varepsilon n^{2/3}} kj(E_{k,j}-t_kt_j)\\
&= \sum_{k=1}^{\varepsilon n^{2/3}} k^2 t_k + \sum_{k=1}^{\varepsilon n^{2/3}}\sum_{j=1}^{\varepsilon n^{2/3}} kjt_kt_jO\Big(kjn^{-4/3}\Big)\\
&= \sum_{k=1}^{\varepsilon n^{2/3}} k^2 t_k + O\Big(\sum_{k=1}^{\varepsilon n^{2/3}}\sum_{j=1}^{\varepsilon n^{2/3}} k^2j^2t_kt_jn^{-4/3}\Big)\\
&= A_2 + O\big(A_2^2 n^{-4/3}\big) = A_2 + O\big(n^{4/3}\varepsilon\big). \qquad (6.11)
\end{aligned}$$

In particular, by (6.3), this variance is $O\big(n^{4/3}\varepsilon^{1/2}\big)$.

The variance of $Y_n\big([1,\varepsilon n^{2/3}];1\big)$ can be computed in the same way, with $t_k$ replaced by $u_k$. Since $u_k = O(\varepsilon^{3/2}t_k)$ for $k \le \varepsilon n^{2/3}$ by (6.7), we obtain the estimate

$$\operatorname{Var}\big(Y_n\big([1,\varepsilon n^{2/3}];1\big)\big) = O\big(\varepsilon^{3/2}A_2 + \varepsilon^3 A_2^2 n^{-4/3}\big) = O\big(n^{4/3}\varepsilon^2\big). \qquad (6.12)$$

For the complex components we now use the fact that also $\mathbb{E}\big(X_n\big([1,\infty);\ge 2\big)^2\big)$ is bounded [7], and thus

$$\begin{aligned}
\operatorname{Var}\big(Y_n\big([1,\varepsilon n^{2/3}];\ge 2\big)\big) &\le \mathbb{E}\big(Y_n\big([1,\varepsilon n^{2/3}];\ge 2\big)^2\big) \le \varepsilon^2 n^{4/3}\,\mathbb{E}\big(X_n\big([1,\varepsilon n^{2/3}];\ge 2\big)^2\big)\\
&= O(n^{4/3}\varepsilon^2). \qquad (6.13)
\end{aligned}$$

By the Cauchy–Schwarz inequality, the three covariances between the three variables in (6.11), (6.12) and (6.13) are all $O(n^{4/3}\varepsilon^{5/4})$, so summing the variables we find from these formulas that

$$\operatorname{Var}\big(Y_n(\varepsilon n^{2/3},\infty)\big) = \operatorname{Var}\big(Y_n[1,\varepsilon n^{2/3}]\big) = A_2 + O\big(n^{4/3}\varepsilon\big). \qquad (6.14)$$

Moreover, by (6.1), (6.3) and Stirling's formula,

$$\begin{aligned}
A_2 &= \sum_{k=1}^{\varepsilon n^{2/3}} k^2 t_k = \sum_{k=1}^{\varepsilon n^{2/3}} k^2 t_k^*\big(1+O(kn^{-2/3}+n^{-1/3})\big)\\
&= \sum_{k=1}^{\varepsilon n^{2/3}} n\big((2\pi k)^{-1/2} + O(k^{-3/2})\big) + O\big(n^{4/3}\varepsilon^{3/2} + n\varepsilon^{1/2}\big)\\
&= (2/\pi)^{1/2}\varepsilon^{1/2}n^{1+1/3} + O\big(n + n^{4/3}\varepsilon^{3/2}\big).
\end{aligned}$$

Thus, (6.14) yields

$$\operatorname{Var} Z'_{n\varepsilon} = n^{-4/3}\operatorname{Var}\big(Y_n(\varepsilon n^{2/3},\infty)\big) = (2/\pi)^{1/2}\varepsilon^{1/2} + O\big(\varepsilon + n^{-1/3}\big).$$

and (1.2) follows by Proposition 1.4.

## 7. Proof of Theorem 1.2

By Theorem 4.1 and (4.2), for $0 < \varepsilon \le 1$ and fixed $\lambda$,

$$\begin{aligned}\mathbb{E}\, W_\varepsilon &= \int_\varepsilon^\infty \Lambda(x)\, dx = \mathbb{E}\, W_1 + \int_\varepsilon^1 (2\pi)^{-1/2} x^{-5/2} \Psi(x^{3/2}) e^{-F(x,\lambda)}\, dx \\ &= O(1) + \int_\varepsilon^1 (2\pi)^{-1/2} x^{-5/2} \bigl(1 + w_1 x^{3/2} + O(x^3)\bigr)\bigl(1 - F(x,\lambda) + O(x^2)\bigr)\, dx \\ &= \int_\varepsilon^1 (2\pi)^{-1/2} x^{-5/2} \bigl(1 + w_1 x^{3/2} - x\lambda^2/2 + O(x^2)\bigr)\, dx + O(1),\end{aligned}$$

and (1.3) follows. (Recall that $w_1 = (\pi/8)^{1/2}$.)

For the variance, we use (4.5) with $a = \varepsilon$ and $b = 1$ and argue as in Section 6. Calculations similar to (6.11) yield

$$\operatorname{Var}\bigl(X_n([\varepsilon n^{2/3}, n^{2/3}]; 0)\bigr) = \mathbb{E}\bigl(X_n([\varepsilon n^{2/3}, n^{2/3}]; 0)\bigr) + O(\varepsilon^{-1}) = O(\varepsilon^{-3/2}) \tag{7.1}$$

and

$$\operatorname{Var}\bigl(X_n([\varepsilon n^{2/3}, n^{2/3}]; 1)\bigr) = \mathbb{E}\bigl(X_n([\varepsilon n^{2/3}, n^{2/3}]; 1)\bigr) + O(1) = O\bigl(\ln(1/\varepsilon)\bigr); \tag{7.2}$$

we omit the details. Using again the fact that $\mathbb{E}\bigl(X_n([1,\infty]; \ge 2)^2\bigr) = O(1)$ together with $\mathbb{E}\bigl(\bigl(X_n(n^{2/3},\infty]\bigr)^2\bigr) = O(1)$ (Lemma 5.1) and the Cauchy–Schwarz inequality, we obtain

$$\operatorname{Var}\bigl(X_n[\varepsilon n^{2/3}, \infty]\bigr) = \mathbb{E}\bigl(X_n[\varepsilon n^{2/3}, \infty]\bigr) + O(\varepsilon^{-1}).$$

Letting $n \to \infty$, we find using Proposition 1.4, $\operatorname{Var} W_\varepsilon = \mathbb{E}\, W_\varepsilon + O(\varepsilon^{-1})$, and (1.4) follows.

## 8. The Palm distribution

The Palm distributions of a point process $\Xi$ in a suitable space $\mathfrak{S}$ are the conditional distributions $\mathcal{L}(\Xi \mid s \in \Xi)$ given the presence of a given point $s$, $s \in \mathfrak{S}$. (Usually, $s \in \mathfrak{S}$ is an event of probability 0, so this must be interpreted with some care, see [13, Chapter 10]. In particular, note that the Palm distribution is uniquely determined only for a.e. $s$.)

In our case, the Palm distribution is obtained by a simple shift of the parameter $\lambda$; we thus write $\Xi^{(\lambda)} = \{\xi_i^{(\lambda)}\}_i$ in this section. Recall that we regard $\Xi^{(\lambda)}$ as a random measure on $(0,\infty)$ that is the sum of the pointmasses $\delta_{\xi_i^{(\lambda)}}$, see Appendix A.

**Theorem 8.1.** *The Palm distribution $\mathcal{L}(\Xi^{(\lambda)} \mid s \in \Xi^{(\lambda)})$ equals for every $s > 0$ the distribution of $\Xi^{(\lambda - s)} + \delta_s$.*

*Proof.* Given that $G(n,p)$ has a component of size $m$ on a certain set of vertices, the remainder of the graph is distributed as $G(n-m,p)$. Hence, if $\mathfrak{N} = \mathfrak{N}(0,\infty]$ is the space of locally finite integer-valued measures on $\mathfrak{S} = (0,\infty]$ defined in Appendix A, and $f : (0,\infty] \to \mathbb{R}$ and $g : \mathfrak{N} \to \mathbb{R}$ are bounded continuous functions and $f$ has compact support, then

$$\mathbb{E}\Bigl(g(\Xi_n) \int_0^\infty f\, d\Xi_n\Bigr) = \mathbb{E}\Bigl(g(\Xi_n) \sum_i f(\xi_{ni})\Bigr) = \mathbb{E}\Bigl(\sum_i h_n(\xi_{ni})\Bigr), \tag{8.1}$$

where $h_n(s) = f(s)\, \mathbb{E}\, g(\Xi_{n - sn^{2/3}, p} + \delta_s)$. If $n \to \infty$ and $s_n \to s$, then $(n - s_n n^{2/3})p = 1 + (\lambda - s + o(1)) n^{-1/3}$ and $\Xi_{n - s_n n^{2/3}, p} \xrightarrow{\mathrm{d}} \Xi^{(\lambda - s)}$, and thus $h_n(s_n) \to h(s) :=$

$f(s)\,\mathbb{E}\,g(\Xi^{(\lambda-s)}+\delta_s)$. It follows, using Lemma A.2 and [3, Theorem 5.5], that $\sum_i h_n(\xi_{ni}) \xrightarrow{\mathrm{d}} \sum_i h(\xi_i)$, and thus by (8.1) and dominated convergence,

$$\begin{aligned}\mathbb{E}\Bigl(g(\Xi^{(\lambda)})\int_0^\infty f\,d\Xi^{(\lambda)}\Bigr) &= \lim_{n\to\infty}\mathbb{E}\Bigl(g(\Xi_n)\int_0^\infty f\,d\Xi_n\Bigr) = \lim_{n\to\infty}\mathbb{E}\Bigl(\sum_i h_n(\xi_{ni})\Bigr)\\ &= \mathbb{E}\Bigl(\sum_i h(\xi_i)\Bigr) = \int_0^\infty h(s)\,d\,\mathbb{E}\,\Xi^{(\lambda)}(s) = \int_0^\infty f(s)\,\mathbb{E}\,g(\Xi^{(\lambda-s)}+\delta_s)\,d\,\mathbb{E}\,\Xi^{(\lambda)}(s),\end{aligned} \tag{8.2}$$

where $d\,\mathbb{E}\,\Xi^{(\lambda)}(s) = \Lambda^{(\lambda)}(s)\,ds$ by Theorem 4.1. It follows by a monotone class argument (e.g. [8, Theorem A.1]) that the first and last terms are equal for any bounded measurable $g$, and the result follows, see [13, (10.2)]. $\square$

Note that Theorems 8.1 and 1.1 imply that for any fixed $\lambda$ and $s>0$, for small $\varepsilon$ (so that $\varepsilon<s$), $\mathbb{E}(Z_\varepsilon^{(\lambda)} \mid s\in\Xi^{(\lambda)}) = \mathbb{E}\,Z_\varepsilon^{(\lambda-s)}+s = \mathbb{E}\,Z_\varepsilon^{(\lambda)}+O(\varepsilon^{1/2})$. Hence the existence of a certain point in $\Xi$ asymptotically does not influence $\mathbb{E}\,Z_\varepsilon$ for small $\varepsilon$, showing the rigidity of $\Xi$.

Theorem 8.1 can be put in a computational form as follows. Let, as above, $\mathfrak{N}=\mathfrak{N}(0,\infty]$ be the space of integer-valued measures defined in Appendix A.

**Theorem 8.2.** *For any bounded or non-negative measurable function $F:(0,\infty)\times\mathfrak{N}\to[0,\infty]$,*

$$\mathbb{E}\sum_i F\bigl(\xi_i^{(\lambda)},\Xi^{(\lambda)}\bigr) = \int_0^\infty \mathbb{E}\,F\bigl(x,\Xi^{(\lambda-x)}+\delta_x\bigr)\Lambda^{(\lambda)}(x)\,dx, \tag{8.3}$$

*where $\Lambda^{(\lambda)}(x)$ is given by* (4.7)*.*

*Proof.* First consider $F$ of the special form $F(x,\Xi)=f(x)g(\Xi)$, where, as in the proof of Theorem 8.1, $f:(0,\infty]\to\mathbb{R}$ and $g:\mathfrak{N}\to\mathbb{R}$ are bounded continuous functions and $f$ has compact support. Then (8.2) holds, which can be written

$$\mathbb{E}\int_0^\infty F(x,\Xi^{(\lambda)})\,d\Xi^{(\lambda)}(x) = \int_0^\infty \mathbb{E}\,F(x,\Xi^{(\lambda-x)}+\delta_x)\Lambda^{(\lambda)}(x)\,dx. \tag{8.4}$$

By another monotone class argument (e.g. [8, Theorem A.1]), (8.4) holds for every bounded measurable $F$, and thus by monotone convergence for every non-negative measurable $F$ too.

The integral on the left hand side of (8.4) equals $\sum_i F(\xi_i^{(\lambda)},\Xi^{(\lambda)})$, which yields (8.3). $\square$

We give some applications.

**Corollary 8.3.** *Let $\Lambda^{(\lambda)}(x)$ be given by* (4.7)*. Then, for every $\varepsilon>0$,*

$$\mathbb{E}\,Z_\varepsilon^2 = \int_\varepsilon^\infty x^2\Lambda^{(\lambda)}(x)\,dx + \int_\varepsilon^\infty\int_\varepsilon^\infty xy\Lambda^{(\lambda)}(x)\Lambda^{(\lambda-x)}(y)\,dy\,dx$$

*and thus*

$$\begin{aligned}\operatorname{Var} Z_\varepsilon &= \int_\varepsilon^\infty x^2\Lambda^{(\lambda)}(x)\,dx - \int_\varepsilon^\infty\int_\varepsilon^\infty xy\Lambda^{(\lambda)}(x)\bigl(\Lambda^{(\lambda)}(y)-\Lambda^{(\lambda-x)}(y)\bigr)\,dy\,dx\\ &= \int_\varepsilon^\infty\int_0^\varepsilon xy\Lambda^{(\lambda)}(x)\bigl(\Lambda^{(\lambda)}(y)-\Lambda^{(\lambda-x)}(y)\bigr)\,dy\,dx.\end{aligned}$$

*Proof.* Take $F(x,\Xi) = x\int_\varepsilon^\infty y\,d\Xi(y)\mathbf{1}[x\geq\varepsilon]$ in (8.3), or $f(x) = x\mathbf{1}[x\geq\varepsilon]$ and $g(\Xi) = \int_\varepsilon^\infty y\,d\Xi(y)$ in (8.2), to find

$$\mathbb{E}\,Z_\varepsilon^2 = \int_\varepsilon^\infty x\,\mathbb{E}\Bigl(\int_\varepsilon^\infty y\,d\Xi^{(\lambda-x)}(y) + x\Bigr)\Lambda^{(\lambda)}(x)\,dx,$$

which yields the formula for $\mathbb{E}\,Z_\varepsilon^2$ by Theorem 4.1 (or Corollary 4.2) applied with $\lambda - x$.

The first formula for $\operatorname{Var} Z_\varepsilon$ follows immediately, and the second follows because (4.9) implies

$$x = \int_0^\infty y\bigl(\Lambda^{(\lambda)}(y) - \Lambda^{(\lambda-x)}(y)\bigr)\,dy$$

and thus

$$\int_\varepsilon^\infty x^2\Lambda^{(\lambda)}(x)\,dx = \int_\varepsilon^\infty\int_0^\infty xy\bigl(\Lambda^{(\lambda)}(y) - \Lambda^{(\lambda-x)}(y)\bigr)\Lambda^{(\lambda)}(x)\,dy\,dx. \qquad\square$$

**Corollary 8.4.** *$\Xi^{(\lambda)}$ is a.s. simple, i.e. lacks multiple points.*

*Proof.* Take $F(x,\Xi) := \mathbf{1}[\Xi\{x\}\geq 2]$ in (8.3). The left hand side becomes the expected number of multiple points (with multiplicities), while the right hand side is 0 because, for each $x$, $\mathbb{E}\,F(x,\Xi^{(\lambda-x)}+\delta_x) = \mathbb{P}\bigl(\Xi^{(\lambda-x)}\{x\}\geq 1\bigr) = 0$, using Theorem 4.1 which shows that the intensity of $\Xi^{(\lambda-x)}$ is absolutely continuous. $\square$

**Corollary 8.5.** *The largest point $\xi_1^{(\lambda)}$ in $\Xi^{(\lambda)}$ has a distribution with the density function $h_1^{(\lambda)}(x) := \mathbb{P}\bigl(\Xi^{(\lambda-x)}(x,\infty) = 0\bigr)\Lambda^{(\lambda)}(x)$.*

*Proof.* Let $f : (0,\infty) \to [0,\infty]$ be a measurable function and take $F(x,\Xi) := f(x)\mathbf{1}[\Xi(x,\infty) = 0]$ in Theorem 8.2. Since $\Xi^{(\lambda)}$ is simple by Corollary 8.4, the left hand side of (8.3) becomes $\mathbb{E}\,f(\xi_1^{(\lambda)})$, and the right hand side is $\int_0^\infty f(x)h_1^{(\lambda)}(x)\,dx$. Since $f$ is arbitrary, the result follows. $\square$

The proof immediately extends to the following, more general, result.

**Corollary 8.6.** *For any $k\geq 1$, the $k$:th largest point $\xi_k^{(\lambda)}$ in $\Xi^{(\lambda)}$ has a distribution with the density function $h_k^{(\lambda)}(x) := \mathbb{P}\bigl(\Xi^{(\lambda-x)}(x,\infty) = k-1\bigr)\Lambda^{(\lambda)}(x)$.* $\square$

**Corollary 8.7.** *For any Borel set $B\subseteq(0,\infty)$ and $k\geq 1$,*

$$\mathbb{E}\bigl(\Xi^{(\lambda)}(B)^{\underline{k}}\bigr) = \int_B\cdots\int_B \Lambda^{(\lambda)}(x_1)\Lambda^{(\lambda-x_1)}(x_2)\cdots\Lambda^{(\lambda-x_1-\cdots-x_{k-1})}(x_k)\,dx_k\cdots dx_1.$$

*Proof.* For $k=1$, this is just the definition of intensity, see Theorem 4.1. For $k\geq 2$, we use Theorem 8.2 with $F(x,\Xi) := \mathbf{1}[x\in B]\bigl(\Xi(B)-1\bigr)^{\underline{k-1}}$, which yields

$$\mathbb{E}\bigl(\Xi^{(\lambda)}(B)^{\underline{k}}\bigr) = \int_0^\infty \mathbf{1}[x\in B]\,\mathbb{E}\bigl(\Xi^{(\lambda-x)}(B)\bigr)^{\underline{k-1}}\Lambda^{(\lambda)}(x)\,dx,$$

and the result follows by induction. $\square$

**Remark 8.8.** It follows immediately that if $B\subseteq[a,b]$ with $0<a<b<\infty$, then $\mathbb{E}\bigl(\Xi^{(\lambda)}(B)^{\underline{k}}\bigr) = O(C^k)$ as $k\to\infty$, for some $C<\infty$ depending on $B$ and $\lambda$; with only a little more effort, the same can be shown also for $B\subseteq[a,\infty]$. This implies $\mathbb{E}\,e^{t\Xi^{(\lambda)}(B)} < \infty$ for every such $B$ and $t<\infty$. In particular, the distribution of $\Xi^{(\lambda)}(B)$ is determined by its (factorial) moments. Hence the formula in Corollary 8.7 in principle determines the distribution of $\Xi^{(\lambda)}(B)$ for any relatively compact $B\subset$

$(0,\infty]$. Moreover, the formula in Corollary 8.7 easily extends to mixed factorial moments of $\Xi^{(\lambda)}(B_1),\dots,\Xi^{(\lambda)}(B_m)$ when $B_1,\dots,B_m$ are disjoint relatively compact Borel sets. This extension, which we leave to the reader, characterizes the joint distribution of $\Xi^{(\lambda)}(B_1),\dots,\Xi^{(\lambda)}(B_m)$, and thus [13, Theorem 3.1] the distribution of $\Xi^{(\lambda)}$.

**Remark 8.9.** If $B\subseteq[a,\infty]$ for some $a>0$, we have, using the estimate in Remark 8.8, the standard formula

$$\mathbb{P}\bigl(\Xi^{(\lambda)}(B)=0\bigr)=\sum_{k=0}^{\infty}\frac{(-1)^k}{k!}\,\mathbb{E}\bigl(\Xi^{(\lambda)}(B)^{\underline{k}}\bigr), \tag{8.5}$$

which together with Corollary 8.7 (and perhaps the Bonferroni inequalities) can be used for numerical evaluation of $\mathbb{P}\bigl(\Xi^{(\lambda)}(B)=0\bigr)$, and thus, in particular, of the density function in Corollary 8.5.

**Remark 8.10.** It follows easily from Theorem 3.1 that a result analogous to Theorem 8.1 holds for $\Xi^*$ too.

## 9. Limits as $\lambda\to\pm\infty$

In this section we consider limit results for $\Xi^{(\lambda)}$, and in particular for the largest point $\xi_1^{(\lambda)}$, as $\lambda\to\pm\infty$. These results are equivalent to limit results for $G(n,p)$ with $p=n^{-1}+\lambda(n)n^{-4/3}$ with $\lambda(n)\to\pm\infty$ slowly, but we get in this way no information on the allowed range of $\lambda(n)$.

Consider first $\lambda\to-\infty$. By (4.2), $F(x,\lambda)\to\infty$ for every $x>0$ and $F(x,\lambda)$ is monotone in $\lambda$ for $\lambda\le 0$. Recalling the notation $W_\varepsilon:=\Xi^{(\lambda)}[\varepsilon,\infty)$, it follows by dominated convergence that, for every fixed $\varepsilon>0$, $\mathbb{E}\,W_\varepsilon=\int_\varepsilon^\infty\Lambda^{(\lambda)}(x)\,dx\to 0$. Hence, $\mathbb{P}(W_\varepsilon>0)\to 0$ and $\mathbb{P}(\Xi[\varepsilon,\infty)=\emptyset)\to 1$. Consequently, $\Xi\xrightarrow{\mathrm{p}}\emptyset$ (in the vague topology, see Appendix A) and $\xi_1^{(\lambda)}\xrightarrow{\mathrm{p}}0$.

We can by much more precise. For $|\lambda|>1$, let

$$a_\lambda:=3\ln|\lambda|-\tfrac52\ln\ln|\lambda|-\tfrac12\ln(2^4 3^5\pi), \tag{9.1}$$

so that, as $\lambda\to\pm\infty$, $a_\lambda\sim 3\ln|\lambda|$ and

$$e^{-a_\lambda}=|\lambda|^{-3}(\ln|\lambda|)^{5/2}(2^4 3^5\pi)^{1/2}\sim 4\pi^{1/2}|\lambda|^{-3}a_\lambda^{5/2}. \tag{9.2}$$

**Theorem 9.1.** *As $\lambda\to-\infty$,*

$$\frac{|\lambda|^2}{2}\xi_1^{(\lambda)}-a_\lambda\xrightarrow{\mathrm{d}}V,$$

*where $V$ has the Gumbel (extreme value) distribution $\mathbb{P}(V\le s)=e^{-e^{-s}}$.*

*Proof.* Fix a real $s$, and let $N^{(\lambda)}(x):=\Xi^{(\lambda)}(x,\infty)$, the number of points in $\Xi^{(\lambda)}$ larger than $x$. Thus $\mathbb{E}\,N^{(\lambda)}(x)=\mathbb{E}\,\Xi^{(\lambda)}(x,\infty)=\int_x^\infty\Lambda^{(\lambda)}(y)\,dy$. With the change of variables $y=2\lambda^{-2}(a_\lambda+t)$, we obtain

$$\mathbb{E}\,N^{(\lambda)}\bigl(2\lambda^{-2}(a_\lambda+s)\bigr)=\int_s^\infty 2\lambda^{-2}\Lambda^{(\lambda)}\bigl(2\lambda^{-2}(a_\lambda+t)\bigr)\,dt. \tag{9.3}$$

For $\lambda\le 0$ and any real $t$ we have, by (4.2),

$$F\bigl(2\lambda^{-2}(a_\lambda+t),\lambda\bigr)=\tfrac86\lambda^{-6}(a_\lambda+t)^3+2|\lambda|^{-3}(a_\lambda+t)^2+(a_\lambda+t)=a_\lambda+t+o(1),$$

as $\lambda \to -\infty$ with $t$ fixed. Since $\Psi(x) \to 1$ as $x \to 0$, it follows from this, (4.7) and (9.2) that

$$\begin{aligned}
&2\lambda^{-2}\Lambda^{(\lambda)}\bigl(2\lambda^{-2}(a_\lambda + t)\bigr)\\
&\qquad = (2\pi)^{-1/2}2^{-3/2}|\lambda|^3(a_\lambda + t)^{-5/2}\Psi\bigl(2^{3/2}|\lambda|^{-3}(a_\lambda + t)^{3/2}\bigr)e^{-F(2\lambda^{-2}(a_\lambda+t),\lambda)}\\
&\qquad = \tfrac{1}{4}\pi^{-1/2}|\lambda|^3 a_\lambda^{-5/2}e^{-a_\lambda - t + o(1)} \to e^{-t}.
\end{aligned}$$

Moreover, for $t \ge s$ and $\lambda < 0$ with $|\lambda|$ so large that $a_\lambda > 2|s|$ we also obtain, using $\Psi(x) = O\bigl(e^{x^2/6}\bigr)$ from (3.4) and $F(x,\lambda) \ge x^3/6 + x\lambda^2/2$ from (4.2),

$$2\lambda^{-2}\Lambda^{(\lambda)}\bigl(2\lambda^{-2}(a_\lambda + t)\bigr) = O\Bigl(|\lambda|^3 a_\lambda^{-5/2}e^{-(a_\lambda+t)}\Bigr) = O\bigl(e^{-t}\bigr).$$

Consequently, we can use dominated convergence in (9.3) and thus

$$\mathbb{E}\, N^{(\lambda)}\bigl(2\lambda^{-2}(a_\lambda + s)\bigr) \to \int_s^\infty e^{-t}\,dt = e^{-s}.$$

Higher factorial moments can be computed similarly using Corollary 8.7, with $B = (2\lambda^{-2}(a_\lambda + s), \infty)$ and $x_j = 2\lambda^{-2}(a_\lambda + t_j)$. Note that, for fixed $t_j$, $x_j \to 0$, and thus $F(x_j, \lambda - x_1 - \cdots - x_{j-1}) = F(x_j, \lambda) + o(1)$. Note further that, for $\lambda < 0$ and every $u, x \ge 0$, $\Lambda^{(\lambda-u)}(x) \le \Lambda^{(\lambda)}(x)$; hence the bound used to verify dominated convergence above applies to each factor in this multivariate setting too. Consequently, for every $k \ge 1$,

$$\mathbb{E}\bigl(N^{(\lambda)}(2\lambda^{-2}(a_\lambda + s))\bigr)^{\underline{k}} \to \int_s^\infty \cdots \int_s^\infty e^{-t_1}\cdots e^{-t_k}\,dt_k\cdots dt_1 = \bigl(e^{-s}\bigr)^k.$$

By the method of moments, this implies $N^{(\lambda)}\bigl(2\lambda^{-2}(a_\lambda + s)\bigr) \xrightarrow{\mathrm{d}} \mathrm{Po}\bigl(e^{-s}\bigr)$, and thus

$$\mathbb{P}\bigl(\xi_1^{(\lambda)} \le 2\lambda^{-2}(a_\lambda + s)\bigr) = \mathbb{P}\bigl(N^{(\lambda)}\bigl(2\lambda^{-2}(a_\lambda + s)\bigr) = 0\bigr) \to e^{-e^{-s}}. \qquad \square$$

**Remark 9.2.** The proof yields also the asymptotic distribution of $\xi_2^{(\lambda)}$, $\xi_3^{(\lambda)}$, .... In fact, for every fixed $i$, as $\lambda \to -\infty$,

$$\mathbb{P}\bigl(\xi_i^{(\lambda)} \le 2\lambda^{-2}(a_\lambda + s)\bigr) = \mathbb{P}\bigl(N^{(\lambda)}\bigl(2\lambda^{-2}(a_\lambda + s)\bigr) < i\bigr) \to \sum_{j=0}^{i-1}\frac{e^{-js}}{j!}e^{-e^{-s}};$$

if we write the right hand side as $\mathbb{P}(V_i \le s)$, this can be written

$$\frac{|\lambda|^2}{2}\xi_i^{(\lambda)} - a_\lambda \xrightarrow{\mathrm{d}} V_i.$$

Note that these asymptotic distributions are the same as for the $i$:th records of suitable i.i.d. sequences, see [15, Section 2.2].

More generally, the proof above is easily extended to show that $\Xi^{(\lambda)}$ with the points rescaled as above, converges in distribution to a Poisson process on $(-\infty, \infty)$ with intensity $e^{-s}$. (This holds in $\mathfrak{N}[-a, \infty]$ for every $a$, say; we cannot use $\mathfrak{N}(-\infty, \infty]$ directly, since the rescaled processes are not elements of this space.) Thus for $\lambda \to -\infty$, the point process $\Xi^{(\lambda)}$ becomes Poisson-like.

In particular, $\xi_i^{(\lambda)}$ is roughly $2\lambda^{-2}a_\lambda \sim 6\ln|\lambda|/\lambda^2$ for every fixed $i \ge 1$. This can be made precise in the following form, where we use the notation that $X_\lambda \sim_p x_\lambda$ if $X_\lambda/x_\lambda \xrightarrow{\mathrm{p}} 1$.

**Corollary 9.3.** *As $\lambda \to -\infty$, $\xi_i^{(\lambda)} \sim_p 6\ln|\lambda|/\lambda^2$ for every fixed $i \ge 1$.* $\square$

**Remark 9.4.** The asymptotic results for large negative $\lambda$ in Theorem 9.1 and Remark 9.2 have a natural interpretation. The results on the asymptotic distribution of $\xi_i^{(\lambda)}$ are what they would be if $\Xi^{(\lambda)}$ were replaced by a Poisson point process with intensity $\Lambda^{(\lambda)}$. This corresponds to the view that as one moves in the critical window toward the subcritical phase the largest components become "local phenomenon" and their interaction becomes negligible.

Let us now turn to $\lambda \to +\infty$. It is well-known that in this case, with probability tending to 1, $\Xi^{(\lambda)}$ contains exactly one large point. In fact, $\xi_1^{(\lambda)} \overset{\mathrm{p}}{\to} \infty$ and $\xi_2^{(\lambda)} \overset{\mathrm{p}}{\to} 0$ as $\lambda \to +\infty$. Again, we can be much more precise.

Let $X$ and $Y$ by two random variables. The *total variation distance* between the distributions of $X$ and $Y$ is defined as

$$d_{\mathrm{TV}}(X,Y) := \sup_B |\mathbb{P}(X \in B) - \mathbb{P}(Y \in B)|,$$

taking the supremum over all Borel sets $B$. Note that this only depends on the distributions $\mathcal{L}(X)$ and $\mathcal{L}(Y)$, although we for simplicity use the notation $d_{\mathrm{TV}}(X,Y)$ instead of $d_{\mathrm{TV}}(\mathcal{L}(X),\mathcal{L}(Y))$; we will also write $d_{\mathrm{TV}}(X,\mu)$ when $Y$ has distribution $\mu$. Note also that $d_{\mathrm{TV}}$ is a very strong measure of distance between distributions; for example, for a sequence $X_n$, $d_{\mathrm{TV}}(X_n,Y) \to 0$ is much stronger than $X_n \overset{\mathrm{d}}{\to} Y$, and thus (i) below is stronger than asymptotic normality in the standard form $(\xi_1^{(\lambda)} - 2\lambda)/\sqrt{2/\lambda} \overset{\mathrm{d}}{\to} N(0,1)$.

**Theorem 9.5.** *If $\lambda \to +\infty$, then*

(i) $d_{\mathrm{TV}}\big(\xi_1^{(\lambda)}, N(2\lambda, 2\lambda^{-1})\big) \to 0$;

(ii) $\frac{\lambda^2}{2}\xi_2^{(\lambda)} - a_\lambda \overset{\mathrm{d}}{\to} V$, *with $a_\lambda$ as in* (9.1) *and $V$ as in Theorem* 9.1.

The proof below also shows that $\frac{1}{2}|\lambda|^2\xi_i^{(\lambda)} - a_\lambda \overset{\mathrm{d}}{\to} V_{i-1}$ for every $i \ge 2$, with $V_i$ as in Remark 9.2.

**Corollary 9.6.** *As $\lambda \to +\infty$, $\xi_1^{(\lambda)} \sim_p 2\lambda$ and $\xi_i^{(\lambda)} \sim_p 6\ln\lambda/\lambda^2$ for every fixed $i \ge 2$.* □

To prove Theorem 9.5, we begin with two lemmas. Let $\varphi_\lambda$ denote the density function of $N(2\lambda, 2\lambda^{-1})$; thus, $\varphi_\lambda(x) = (\lambda/4\pi)^{1/2}e^{-\lambda(x-2\lambda)^2/4}$.

**Lemma 9.7.** *As $\lambda \to +\infty$,*

$$\int_\lambda^\infty \big|\Lambda^{(\lambda)}(x) - \varphi_\lambda(x)\big|\,dx \to 0.$$

The lower limit $\lambda$ is for convenience only; it can easily be replaced by, e.g., 1.

*Proof.* For $x \ge \lambda$, (3.4) yields $\Psi(x^{3/2}) = \frac{1}{2}x^3e^{x^3/24}\big(1+o(1)\big)$, with $o(1) \to 0$ as $\lambda \to +\infty$, uniformly in $x \ge \lambda$. Using $|a^{1/2} - b^{1/2}| = |a-b|/(a^{1/2}+b^{1/2})$ and $|e^a - e^b| \le |a-b|e^{\max\{a,b\}}$, we thus find from (4.7) and (4.3), for $x \ge \lambda$,

$$\begin{aligned}\Lambda^{(\lambda)}(x) &= (2\pi)^{-1/2}\tfrac{1}{2}x^{1/2}e^{-x(x-2\lambda)^2/8}\big(1+o(1)\big)\\ &= (8\pi)^{-1/2}(2\lambda)^{1/2}e^{-x(x-2\lambda)^2/8}\big(1+o(1)\big) + O\Big(|x-2\lambda|\lambda^{-1/2}e^{-\lambda(x-2\lambda)^2/8}\Big)\\ &= \varphi_\lambda(x)\big(1+o(1)\big) + O\Big(\big(\lambda^{1/2}|x-2\lambda|^3 + \lambda^{-1/2}|x-2\lambda|\big)e^{-\lambda(x-2\lambda)^2/8}\Big).\end{aligned}$$

The result follows by integrating; the $O$ term yields, if we let $Z \sim N(2\lambda, 4\lambda^{-1})$, $O\big(\mathbb{E}(\lambda^{1/2}|Z-2\lambda|^3 + \lambda^{-1/2}|Z-2\lambda|)\big) = O(\lambda^{-1})$. □

**Lemma 9.8.** *For any random variables $X$ and $Y$ with density functions $f_X$ and $f_Y$, and any Borel set $B \subseteq \mathbb{R}$,*

$$d_{\mathrm{TV}}(X,Y) \le \int_B |f_X(x) - f_Y(x)|\,dx + \mathbb{P}(Y \notin B).$$

*Proof.* It is well-known, and easy to verify, that

$$d_{\mathrm{TV}}(X,Y) = \tfrac12 \int_{-\infty}^{\infty} |f_X(x) - f_Y(x)|\,dx.$$

Since $\int f_X = 1 = \int f_Y$, we have

$$\begin{aligned}\int_{B^c} |f_X(x) - f_Y(x)|\,dx &\le \int_{B^c} \bigl(f_X(x) + f_Y(x)\bigr)\,dx \\ &= 2\,\mathbb{P}(Y \notin B) + \int_{B^c} \bigl(f_X(x) - f_Y(x)\bigr)\,dx = 2\,\mathbb{P}(Y \notin B) - \int_B \bigl(f_X(x) - f_Y(x)\bigr)\,dx\end{aligned}$$

and thus $\int_{\mathbb{R}} |f_X(x) - f_Y(x)|\,dx \le 2\,\mathbb{P}(Y \notin B) + 2\int_B |f_X(x) - f_Y(x)|\,dx$. $\square$

*Proof of Theorem* 9.5. If $x \ge \lambda \ge 0$, then $\Lambda^{(\lambda-x)}(y) \le \Lambda^{(0)}(y)$ for $y \ge 0$ and thus

$$\begin{aligned}\mathbb{P}\bigl(\Xi^{(\lambda-x)}(\lambda,\infty) \ge 1\bigr) \le \mathbb{E}\,\Xi^{(\lambda-x)}(\lambda,\infty) &= \int_\lambda^\infty \Lambda^{(\lambda-x)}(y)\,dy \\ &\le \int_\lambda^\infty \Lambda^{(0)}(y)\,dy = \mathbb{E}\,\Xi^{(0)}(\lambda,\infty) \to 0\end{aligned}$$

as $\lambda \to \infty$. Hence, by Corollary 8.5, $\xi_1^{(\lambda)}$ has a density function $h_1^{(\lambda)}$ with $h_1^{(\lambda)}(x) = \bigl(1-o(1)\bigr)\Lambda^{(\lambda)}(x)$ as $\lambda \to \infty$, uniformly in $x \ge \lambda$. Since Lemma 9.7 implies $\int_\lambda^\infty \Lambda^{(\lambda)} = O(1)$, this yields

$$\int_\lambda^\infty \bigl|h_1^{(\lambda)}(x) - \Lambda^{(\lambda)}(x)\bigr|\,dx \le \mathbb{E}\,\Xi^{(0)}(\lambda,\infty) \int_\lambda^\infty \Lambda^{(\lambda)}(x)\,dx \to 0.$$

Hence Lemma 9.7 yields $\int_\lambda^\infty \bigl|h_1^{(\lambda)}(x) - \varphi_\lambda(x)\bigr|\,dx \to 0$ as $\lambda \to \infty$, and (i) follows by Lemma 9.8, with $B = (\lambda,\infty)$.

For (ii), we observe that, by a simple extension of the proof of Corollary 8.5, the conditional distribution $\mathcal{L}(\Xi^{(\lambda)} \mid \xi_1^{(\lambda)} = x)$ equals the conditional distribution $\mathcal{L}\bigl(\delta_x + \Xi^{(\lambda-x)} \mid \Xi^{(\lambda-x)}(x,\infty) = 0\bigr)$. Since the second largest point in $\delta_x + \Xi^{(\lambda-x)}$, when $\Xi^{(\lambda-x)}(x,\infty) = 0$, is the largest point $\xi_1^{(\lambda-x)}$ in $\Xi^{(\lambda-x)}$, we have, in particular,

$$\mathcal{L}\bigl(\xi_2^{(\lambda)} \mid \xi_1^{(\lambda)} = x\bigr) = \mathcal{L}\bigl(\xi_1^{(\lambda-x)} \mid \xi_1^{(\lambda-x)} \le x\bigr). \tag{9.4}$$

Let $\lambda \to +\infty$, and assume $\lambda > 2$. By (i), $\mathbb{P}(|\xi_1^{(\lambda)} - 2\lambda| < 1) \to 1$. If $|x - 2\lambda| < 1$ and $\lambda' := \lambda - x$, then $|\lambda' - (-\lambda)| < 1$ and thus, by (9.1), $|a_\lambda - a_{\lambda'}| = O(1/\lambda)$. Hence, it follows from Theorem 9.1 that

$$\tfrac12\lambda^2\xi_1^{(\lambda-x)} - a_\lambda = |\lambda/\lambda'|^2\bigl(\tfrac12|\lambda'|^2\xi_1^{(\lambda')} - a_{\lambda'}\bigr) + \bigl(|\lambda/\lambda'|^2 - 1\bigr)a_{\lambda'} + a_{\lambda'} - a_\lambda \xrightarrow{\mathrm{d}} V. \tag{9.5}$$

Furthermore, $\mathbb{P}\bigl(\xi_1^{(\lambda-x)} \le x\bigr) \to 1$, again by Theorem 9.1, and thus (9.5) holds also for the conditional distribution given $\xi_1^{(\lambda-x)} \le x$. By (9.4) and $\mathbb{P}(|\xi_1^{(\lambda)} - 2\lambda| < 1) \to 1$, this yields, for every $y$,

$$\mathbb{P}\bigl(\tfrac12\lambda^2\xi_2^{(\lambda)} - a_\lambda \le y\bigr) = \mathbb{E}\,\mathbb{P}\bigl(\tfrac12\lambda^2\xi_2^{(\lambda)} - a_\lambda \le y \mid \xi_1^{(\lambda)}\bigr) \to \mathbb{P}(V \le y),$$

which proves (ii). $\square$

**Remark 9.9.** Note that it is a fallacy to believe that Theorem 8.1 implies that $\Xi^{(\lambda)}$ conditioned on $\xi_1^{(\lambda)} = x$ has the distribution of $\delta_x + \Xi^{(\lambda-x)}$; as is seen in the proof above, the correct conclusion requires conditioning on $\Xi^{(\lambda-x)}(x,\infty) = 0$. Nevertheless, the proof also shows that the erroneous statement is asymptotically correct as $\lambda \to +\infty$: If $\xi$ has the distribution of $\xi_1^{(\lambda)}$ given in Corollary 8.5, or simply $\xi \sim N(2\lambda, 2\lambda^{-1})$, and given $\xi$ we take a random $\Xi^{(\lambda-\xi)}$, then the distribution of $\delta_\xi + \Xi^{(\lambda-\xi)}$ approximates that of $\Xi^{(\lambda)}$, and $d_{\mathrm{TV}}\bigl(\delta_\xi + \Xi^{(\lambda-\xi)}, \Xi^{(\lambda)}\bigr) \to 0$ as $\lambda \to +\infty$.

**Remark 9.10.** As remarked above, Theorem 9.5 implies asymptotic normality of the size of the largest component in $G(n,p)$ in the case $p = n^{-1} + \lambda(n)n^{-4/3}$ with $\lambda(n) \to \infty$ slowly (without specifying the allowed rate). Indeed, asymptotic normality has been shown for all $\lambda(n)$ in the range $\lambda(n) \to \infty$ but $\lambda(n) = O(n^{1/3})$, i.e. $p = O(n^{-1})$, by Pittel [22] ($p = c/n$) and Pittel and Wormald [23] (the general case).

## Appendix A. Appendix: Point processes

We give here some technical remarks on point processes; see e.g. [13] and [9, Section 4] for further details and proofs.

Let $\mathfrak{S}$ be a 'nice' topological space (more precisely, a locally compact Polish space); in this paper we only consider the intervals $(0,\infty)$ and $(0,\infty]$ and their products with $\mathbb{N}$ or $\mathbb{N}^*$. Although we regard a point process as a random (multi)set $\{\xi_i\}_i \subset \mathfrak{S}$, it is technically convenient to formally define it as a random measure $\sum_i \delta_{\xi_i}$. Hence, if $\Xi$ denotes the point process $\{\xi_i\}$, we write $\Xi(A)$ for the number of points $\xi_i$ that belong to a subset $A \subseteq \mathfrak{S}$; similarly, for suitable functions $f$ on $\mathfrak{S}$, $\int f\,d\Xi = \sum_i f(\xi_i)$.

Thus, let $\mathfrak{N} = \mathfrak{N}(\mathfrak{S})$ be the class of all Borel measures $\mu$ on $\mathfrak{S}$ such that $\mu(A)$ is a (finite) integer $0, 1, \dots$ for every relatively compact Borel set $A$; this coincides with the class of all finite or countably infinite sums of the type $\sum_i \delta_{x_i}$, where $x_i \in \mathfrak{S}$ and each compact subset of $\mathfrak{S}$ contains only a finite number of $x_i$, and we identify such a sum with the (multi)set $\{x_i\}$.

The standard topology on $\mathfrak{N}$ (known as the *vague topology*) is defined such that, for $\mu, \mu_1, \mu_2, \dots \in \mathfrak{N}$, $\mu_n \to \mu$ if and only if $\int f\,d\mu_n \to \int f\,d\mu$ for every $f \in C_c(\mathfrak{S})$, the space of (real-valued) continuous functions on $\mathfrak{S}$ with compact support. (This is a metrizable topology and $\mathfrak{N}$ is a Polish space, see [13, Section 15.7].)

A point process on $\mathfrak{S}$ is a random element of $\mathfrak{N}$. If $\Xi$ is a point process on $\mathfrak{S}$, there exists a unique Borel measure $\nu$ on $\mathfrak{S}$ such that $\mathbb{E}\,\Xi(A) = \nu(A)$ for every Borel set $A$, and more generally $\mathbb{E}\int h\,d\Xi = \int h\,d\nu$ for every positive measurable function $h$. This measure $\nu$ is called the *intensity* of $\Xi$. In the cases we consider, $\mathfrak{S}$ is an interval or a union of intervals, and $\nu$ is absolutely continuous; then also the function $d\nu/dx$ is called the intensity.

If $\Xi_n$ and $\Xi$ are point processes on $\mathfrak{S}$, then $\Xi_n \xrightarrow{\mathrm{d}} \Xi$ (w.r.t. the vague topology just defined) if and only if $\int f\,d\Xi_n \xrightarrow{\mathrm{d}} \int f\,d\Xi$ (as real-valued random variables) for every $f \in C_c(\mathfrak{S})$. It is also true that $\Xi_n \xrightarrow{\mathrm{d}} \Xi$ if and only if $\Xi_n(A) \xrightarrow{\mathrm{d}} \Xi(A)$ for every relatively compact Borel set $A \subseteq \mathfrak{S}$ such that $\Xi(\partial A) = 0$ a.s., and moreover joint convergence holds for every finite collection of such sets $A$.

We state a particular case that we need. Say that a point $x$ is a *continuity point* of a point process $\Xi$ if $x$ is a continuity point of $\mathbb{E}\,\Xi$, i.e. if $\mathbb{E}\,\Xi\{x\} = 0$, or equivalently, $x \notin \Xi$ a.s.

**Lemma A.1.** *If $\Xi_n \xrightarrow{\mathrm{d}} \Xi$ as point processes on an interval $J$, then $\Xi_n[a,b] \xrightarrow{\mathrm{d}} \Xi[a,b]$ for every interval $[a,b] \subset J$ such that $a$ and $b$ are continuity points of $\Xi$.* □

Note that the definitions of both point processes and convergence of them are sensitive to the choice of $\mathfrak{S}$, since a point process is not allowed to have any cluster point in $\mathfrak{S}$. Hence, it matters whether boundary points are included in $\mathfrak{S}$, even if they are not attained by any point. For example, if $\mathfrak{S}$ is the closed interval $[0,\infty]$ (or any compact set), then every point process is finite. If, instead, $\mathfrak{S}$ is the half-open interval $(0,\infty]$, then an element $\mu \in \mathfrak{N}$ is finite on every interval $[a,\infty]$, and thus every point process may be written as a (finite or infinite) set $\{\xi_i\}$ with $\infty \geq \xi_1 \geq \xi_2 \geq \dots$ and, if the set is infinite, $\xi_i \to 0$ as $i \to \infty$. Similarly, a point process on the open interval $(0,\infty)$ may have both 0 and $\infty$ as cluster points. By including one or both endpoints, we thus get stronger conditions, and, similarly, we get a stronger mode of convergence. It may thus be advantageous to consider (when possible) a set of points in $(0,\infty)$ as a point process on $[0,\infty)$, $(0,\infty]$ or $[0,\infty]$.

For point processes on a closed or half-open interval, with the points ordered as above, convergence is equivalent to joint convergence of the individual points. We state this for the case we are interested in.

**Lemma A.2.** *There is a bijection between $\mathfrak{N}(0,\infty]$ and the space of sequences $(\xi_i)_1^\infty$ with $\xi_1 \geq \xi_2 \geq \dots \geq 0$ and $\lim_{i\to\infty} \xi_i = 0$, such that $\Xi = \{\xi_i\}_{i=1}^N \in \mathfrak{N}$ (or, more formally, $\Xi = \sum_{i=1}^N \delta_{\xi_i}$), with $\xi_1 \geq \xi_2 \geq \dots$ and $0 \leq N \leq \infty$, corresponds to the sequence $(\xi_i)_1^\infty$ where we define $\xi_i := 0$ for $i > N$. This bijection is a homeomorphism between $\mathfrak{N}$ with the vague topology and the space of sequences with component-wise convergence (i.e., the restriction of the product topology on $[0,\infty]^\infty$).*

*Consequently, if $\Xi_n$, $1 \leq n \leq \infty$, are point processes on the interval $(0,\infty]$, and we write $\Xi_n = \{\xi_{ni}\}_{i=1}^{N_n}$ with $\xi_{n1} \geq \xi_{n2} \geq \dots$ and $0 \leq N_n \leq \infty$, and if some $N_n < \infty$, we further define $\xi_{ni} = 0$ for $i > N_n$, then $\Xi_n \xrightarrow{\mathrm{d}} \Xi_\infty$ if and only if $(\xi_{n1}, \xi_{n2}, \dots) \xrightarrow{\mathrm{d}} (\xi_{\infty 1}, \xi_{\infty 2}, \dots)$, in the standard sense that all finite dimensional distributions converge.* □

## References


[1] D. Aldous, Brownian excursions, critical random graphs and the multiplicative coalescent. *Ann. Probab.* **25** (1997), 812–854.

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

[3] P. Billingsley, *Convergence of Probability Measures.* Wiley, New York, 1968.

[4] B. Bollobás, *Random Graphs.* Academic Press, London, 1985.

[5] P. Chassaing & S. Janson, The center of mass of the ISE and the Wiener index of trees. *Electronic Comm. Probab.*, **9** (2004), paper 20, 178–187.

[6] A. Gut, *Probability: A Graduate Course.* Springer, New York, 2005.

[7] S. Janson, Multicyclic components in a random graph process. *Random Struct. Alg.* **4** (1993), 71–84.

[8] S. Janson, *Gaussian Hilbert Spaces.* Cambridge Univ. Press, Cambridge, 1997.

[9] S. Janson, Cycles and unicyclic components in random graphs. *Combin. Probab. Comput.* **12** (2003), 27–52.

[10] S. Janson, The Wiener index of simply generated random trees. *Random Struct. Alg.* **22** (2003), no. 4, 337–358.

[11] S. Janson, D.E. Knuth, T. Łuczak & B. Pittel, The birth of the giant component. *Random Struct. Alg.* **3** (1993), 233–358.

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

[13] O. Kallenberg, *Point Processes.* Akademie-Verlag, Berlin, 1983.

[14] O. Kallenberg, *Foundations of Modern Probability.* 2nd ed., Springer-Verlag, New York, 2002.
[15] M.R. Leadbetter, G. Lindgren & H. Rootzén, *Extremes and Related Properties of Random Sequences and Processes.* Springer-Verlag, New York, 1983.
[16] G. Louchard, Kac's formula, Lévy's local time and Brownian excursion. *J. Appl. Probab.* **21** (1984), no. 3, 479–499.
[17] G. Louchard, The Brownian excursion area: a numerical analysis. *Comput. Math. Appl.* **10** (1984), no. 6, 413–417. Erratum: *Comput. Math. Appl. Part A* **12** (1986), no. 3, 375.
[18] T. Łuczak, B. Pittel & J.C. Wierman, The structure of a random graph near the point of the phase transition. *Trans. Amer. Math. Soc.* **341** (1994), 721–748.
[19] R. Otter, The multiplicative process. *Ann. Math. Statistics* **20** (1949), 206–224.
[20] Y. Peres, Personal communication, 2005.
[21] J. Pitman, Enumerations of trees and forests related to branching processes and random walks. *Microsurveys in discrete probability (Princeton, NJ, 1997)*, DIMACS Ser. Discrete Math. Theoret. Comput. Sci., 41, Amer. Math. Soc., Providence, RI, 1998, pp. 163–180.
[22] B. Pittel, On tree census and the giant component in sparse random graphs. *Random Struct. Alg.* **1** (1990), 311–342.
[23] B. Pittel & N.C. Wormald, Counting connected graphs inside-out. *J. Comb. Th. B* **93** (2005), 127–172.
[24] J. Spencer, Enumerating graphs and Brownian motion. *Comm. Pure Appl. Math.* **50** (1997), no. 3, 291–294.
[25] E.M. Wright, The number of connected sparsely edged graphs. *J. Graph Th.* **1** (1977), 317–330.

Department of Mathematics, Uppsala University, PO Box 480, S-751 06 Uppsala, Sweden
*E-mail address*: `svante.janson@math.uu.se`
*URL*: `http://www.math.uu.se/~svante/`

Joel Spencer, Courant Institute, 251 Mercer St., New York, NY 10012, USA
*E-mail address*: `spencer@cs.nyu.edu`
*URL*: `http://www.cs.nyu.edu/cs/faculty/spencer/`