A local limit theorem for the critical random graph

Remco van der Hofstad[*] Wouter Kager[†‡] Tobias Müller[‡]

June 18, 2008

Abstract

We consider the limit distribution of the orders of the k largest components in the Erdős-Rényi random graph inside the "critical window" for arbitrary k. We prove a local limit theorem for this joint distribution and derive an exact expression for the joint probability density function.

1 Introduction

The Erdős-Rényi random graph $G(n,p)$ is a random graph on the vertex-set $[n] := \{1,\dots,n\}$, constructed by including each of the $\binom{n}{2}$ possible edges with probability p, independently of all other edges. We shall be interested in the Erdős-Rényi random graph in the so-called *critical window*. That is, we fix $\lambda \in \mathbb{R}$ and for p we take

$$p = p_\lambda(n) = \frac{1}{n}\left(1 + \frac{\lambda}{n^{1/3}}\right). \tag{1.1}$$

For $v \in [n]$ we let $\mathcal{C}(v)$ denote the connected component containing the vertex v. Let $|\mathcal{C}(v)|$ denote the number of vertices in $\mathcal{C}(v)$, also called the *order* of $\mathcal{C}(v)$. For $i \geq 1$ we shall use $\mathcal{C}_i$ to denote the component of i^{th} largest order (where ties are broken in an arbitrary way), and we will sometimes also denote $\mathcal{C}_1$ by $\mathcal{C}_{\max}$.

It is well-known that, for p in the critical window (1.1),

$$\big(|\mathcal{C}_1| n^{-2/3}, \dots, |\mathcal{C}_k| n^{-2/3}\big) \xrightarrow{d} \big(C_1^\lambda, \dots, C_k^\lambda\big), \tag{1.2}$$

where $C_1^\lambda, \dots, C_k^\lambda$ are positive, absolutely continuous random variables whose (joint) distribution depends on λ. See [1, 3, 4, 5] and the references therein for the detailed history of the problem. In particular, in [5], an exact formula was found for the distribution function of the limiting variable C_1^λ, and in [1], it was shown that the limit in (1.2) can be described in terms of a certain multiplicative coalescent. The aim of this paper is to prove a *local limit theorem* for the joint probability distribution of the k largest connected components (k arbitrary) and to *investigate the joint limit distribution.* While some ideas used in this paper

[*]Department of Mathematics and Computer Science, Eindhoven University of Technology, P.O. Box 513, 5600 MB Eindhoven, The Netherlands. E-mail: rhofstad@win.tue.nl

[†]Department of Mathematics, VU University, De Boelelaan 1081a, 1081 HV Amsterdam, The Netherlands.

[‡]EURANDOM, Eindhoven University of Technology, P.O. Box 513, 5600 MB Eindhoven, The Netherlands. E-mail: {w.kager,t.muller}@tue.nl.

have also appeared in earlier work, in particular in [3, 4, 5], the results proved here have not been explicitly stated before.

Before we can state our results, we need to introduce some notation. For $n \in \mathbb{N}$ and $0 \le p \le 1$, $\mathbb{P}_{n,p}$ will denote the probability measure of the Erdős-Rényi graph of size n and with edge probability p. For $k \in \mathbb{N}$ and $x_1, \dots, x_k, \lambda \in \mathbb{R}$, we shall denote

$$F_k(x_1, \dots, x_k; \lambda) = \lim_{n\to\infty} \mathbb{P}_{n,p_\lambda(n)}\big(|\mathcal{C}_1| \le x_1 n^{2/3}, \dots, |\mathcal{C}_k| \le x_k n^{2/3}\big). \tag{1.3}$$

It has already been shown implicitly in the work of Łuczak, Pittel and Wierman [4] that this limit exists and that F_k is continuous in all of its parameters. In our proof of the local limit theorem below we will use that $F_1(x;\lambda)$ is continuous in both parameters, which can also easily be seen from the explicit formula (3.25) in [5].

We will denote by $C(m, r)$ the number of (labeled) connected graphs with m vertices and r edges and for $l \ge -1$ we let γ_l denote Wright's constants. That is, γ_l satisfies

$$C(k, k+l) = \big(1 + o(1)\big)\gamma_l k^{k-1/2+3l/2} \qquad \text{as } k \to \infty. \tag{1.4}$$

Here l is even allowed to vary with k: as long as $l = o(k^{1/3})$, the error term $o(1)$ in (1.4) is $O(l^{3/2}k^{-1/2})$ (see [9, Theorem 2]). Moreover, the constants γ_l satisfy (see [7, 8, 9]):

$$\gamma_l = \big(1 + o(1)\big)\sqrt{\frac{3}{4\pi}}\left(\frac{e}{12l}\right)^{l/2} \qquad \text{as } l \to \infty. \tag{1.5}$$

By G we will denote the Laurent series

$$G(s) = \sum_{l=-1}^{\infty} \gamma_l s^l. \tag{1.6}$$

Note that by (1.5) the sum on the right-hand side is convergent for all $s \ne 0$. By a striking result of Spencer [6], G equals s^{-1} times the moment generating function of the scaled Brownian excursion area. For $x > 0$ and $\lambda \in \mathbb{R}$, we further define

$$\Phi(x;\lambda) = \frac{G\big(x^{3/2}\big)}{x\sqrt{2\pi}} e^{-\lambda^3/6+(\lambda-x)^3/6}. \tag{1.7}$$

The main result of this paper is the following local limit theorem for the joint distribution of the vector $(|\mathcal{C}_1|, \dots, |\mathcal{C}_k|)$ in the Erdős-Rényi random graph:

Theorem 1.1 (Local limit theorem for largest clusters)**.** *Let $\lambda \in \mathbb{R}$ and $b > a > 0$ be fixed. As $n \to \infty$, it holds that*

$$\sup_{a \le x_k \le \dots \le x_1 \le b} \Big| n^{2k/3}\, \mathbb{P}_{n,p_\lambda(n)}(|\mathcal{C}_i| = \lfloor x_i n^{2/3} \rfloor \ \forall i \le k) - \Psi_k(x_1, \dots, x_k; \lambda) \Big| \to 0, \tag{1.8}$$

where, for all $x_1 \ge \dots \ge x_k > 0$ and $\lambda \in \mathbb{R}$,

$$\Psi_k(x_1, \dots, x_k; \lambda) = \frac{F_1\big(x_k; \lambda - (x_1 + \dots + x_k)\big)}{r_1! \cdots r_m!} \prod_{i=1}^{k} \Phi\Big(x_i; \lambda - \sum_{j<i} x_j\Big), \tag{1.9}$$

and where $1 \le m \le k$ is the number of distinct values the x_i take, r_1 is the number of repetitions of the largest value, r_2 the number of repetitions of the second largest, and so on.

Theorem 1.1 gives rise to a set of explicit expressions for the probability densities f_k of the limit vectors $(C_1^\lambda, \dots, C_k^\lambda)$ with respect to k-dimensional Lebesgue measure. These densities are given in terms of the distribution function F_1 by the following corollary of Theorem 1.1:

Corollary 1.2 (Joint limiting density for largest clusters)**.** *For any $k > 0$, $\lambda \in \mathbb{R}$ and $x_1 \geq \dots \geq x_k > 0$,*

$$f_k(x_1, \dots, x_k; \lambda) = F_1\big(x_k; \lambda - (x_1 + \dots + x_k)\big) \prod_{i=1}^{k} \Phi\Big(x_i; \lambda - \sum_{j<i} x_j\Big). \tag{1.10}$$

Corollary 1.2 states a set of differential equations that the joint limiting distributions must satisfy. In particular, F_1 satisfies $\frac{\partial}{\partial x} F_1(x; \lambda) = F_1(x; \lambda - x)\Phi(x; \lambda)$. In general this differential equation has many solutions, but we will show that there is only one solution for which $x \mapsto F_1(x; \lambda)$ is a probability distribution for all λ. This leads to the following theorem:

Theorem 1.3 (Uniqueness of solution differential equation)**.** *The set of relations* (1.10) *determines the limit distributions F_k uniquely.*

2 Proof of the local limit theorem

In this section we derive the local limit theorem for the vector $(|\mathcal{C}_1|, \dots, |\mathcal{C}_k|)$ in the Erdős-Rényi random graph. We start by proving a convenient relation between the probability mass function of this vector and the one of a typical component.

Lemma 2.1 (Probability mass function of largest clusters)**.** *Fix $l_1 \geq l_2 \geq \dots \geq l_k > 0$, $n > l_1 + \dots + l_k$ and $p \in [0, 1]$. Let $1 \leq m \leq k$ be the number of distinct values the l_i take, and let r_1 be the number of repetitions of the largest value, r_2 the number of repetitions of the second largest, and so on up to r_m. Then*

$$\mathbb{P}_{n,p}(|\mathcal{C}_i| = l_i \ \forall i \leq k, |\mathcal{C}_{k+1}| < l_k) = \frac{\mathbb{P}_{m_k,p}(|\mathcal{C}_{\max}| < l_k)}{r_1! \cdots r_m!} \prod_{i=0}^{k-1} \frac{m_i}{l_{i+1}} \mathbb{P}_{m_i,p}(|\mathcal{C}(1)| = l_{i+1}), \tag{2.1}$$

where $m_i = n - \sum_{j \leq i} l_j$ for $i = 1, \dots, k$ and $m_0 = n$. Moreover,

$$\mathbb{P}_{n,p}(|\mathcal{C}_i| = l_i \ \forall i \leq k) \leq \frac{1}{r_1! \cdots r_m!} \prod_{i=0}^{k-1} \frac{m_i}{l_{i+1}} \mathbb{P}_{m_i,p}(|\mathcal{C}(1)| = l_{i+1}). \tag{2.2}$$

Proof. For A an event, we denote by $I(A)$ the indicator function of A. For the graph $G(n, p)$, let E_k be the event that $|\mathcal{C}_i| = l_i$ for all $i \leq k$, and notice that

$$I(E_k, |\mathcal{C}_{k+1}| < l_k) = \frac{1}{r_1 l_1} \sum_{v=1}^{n} I(|\mathcal{C}(v)| = l_1, E_k, |\mathcal{C}_{k+1}| < l_k). \tag{2.3}$$

Since $\mathbb{P}_{n,p}(|\mathcal{C}(v)| = l_1, E_k, |\mathcal{C}_{k+1}| < l_k)$ is the same for every vertex v, it follows by taking expectations on both sides of the previous equation that

$$\mathbb{P}_{n,p}(E_k, |\mathcal{C}_{k+1}| < l_k) = \frac{n}{r_1 l_1} \mathbb{P}_{n,p}(|\mathcal{C}(1)| = l_1, E_k, |\mathcal{C}_{k+1}| < l_k). \tag{2.4}$$

Next we observe, by conditioning on $\mathcal{C}(1)$, that

$$\mathbb{P}_{n,p}\big(E_k, |\mathcal{C}_{k+1}| < l_k \bigm| |\mathcal{C}(1)| = l_1\big) = \mathbb{P}_{n-l_1,p}(|\mathcal{C}_1| = l_2, \dots, |\mathcal{C}_{k-1}| = l_k, |\mathcal{C}_k| < l_k). \tag{2.5}$$

Combining (2.4) and (2.5), we thus get

$$\mathbb{P}_{n,p}(E_k, |\mathcal{C}_{k+1}| < l_k) = \frac{n\,\mathbb{P}_{n,p}(|\mathcal{C}(1)| = l_1)}{r_1 l_1}\,\mathbb{P}_{n-l_1,p}(|\mathcal{C}_i| = l_{i+1}\ \forall i < k, |\mathcal{C}_k| < l_k). \tag{2.6}$$

The relation (2.1) now follows by a straightforward induction argument. To see that (2.2) holds, notice that

$$I(|\mathcal{C}_i| = l_i\ \forall i \le k) \le \frac{1}{r_1 l_1}\sum_{v=1}^{n} I(|\mathcal{C}(v)| = l_1, |\mathcal{C}_i| = l_i\ \forall i \le k). \tag{2.7}$$

Proceeding analogously as before leads to (2.2). □

Lemma 2.2 (Scaling function cluster distribution)**.** *Let $\beta > \alpha$ and $b > a > 0$ be arbitrary. As $n \to \infty$,*

$$\sup_{\substack{a\le x\le b\\ \alpha\le\lambda\le\beta}} \Big| n\,\mathbb{P}_{n,p_\lambda(n)}(|\mathcal{C}(1)| = \lfloor xn^{2/3}\rfloor) - x\,\Phi(x;\lambda)\Big| \to 0. \tag{2.8}$$

Proof. For convenience let us write $k := \lfloor xn^{2/3}\rfloor$ and $p = p_\lambda(n)$, with $a \le x \le b$ and $\alpha \le \lambda \le \beta$ arbitrary. Throughout this proof, $o(1)$ denotes error terms tending to 0 with n *uniformly* over all x, λ considered. First notice that

$$\mathbb{P}_{n,p}(|\mathcal{C}(1)| = k) = \binom{n-1}{k-1}\sum_{l=-1}^{\binom{k}{2}-k} C(k,k+l)p^{k+l}(1-p)^{\binom{k}{2}-(k+l)+k(n-k)}. \tag{2.9}$$

Stirling's approximation $m! = \big(1 + O(m^{-1})\big)\sqrt{2\pi m}\,(m/e)^m$ gives us that

$$\binom{n-1}{k-1} = \frac{k}{n}\binom{n}{k} = \big(1+o(1)\big)\frac{n^k k^{1/2-k}}{n\sqrt{2\pi}}\left(1-\frac{k}{n}\right)^{k-n}. \tag{2.10}$$

Next we use the expansion $1 + x = \exp\big(x - x^2/2 + x^3/3 + O(x^4)\big)$ for each factor on the left of the following equation, to obtain

$$\left(1-\frac{k}{n}\right)^{k-n} p^k\,(1-p)^{\binom{k}{2}-k+k(n-k)} = \big(1+o(1)\big)n^{-k}\exp\left(\frac{\lambda k^2}{2n^{4/3}} - \frac{\lambda^2 k}{2n^{2/3}} - \frac{k^3}{6n^2}\right). \tag{2.11}$$

Using that $k = \lfloor xn^{2/3}\rfloor$, combining (2.9)–(2.11) and substituting (1.4) leads to

$$\begin{aligned} n\,\mathbb{P}_{n,p}(|\mathcal{C}(1)| = k) &= \big(1+o(1)\big)e^{(\lambda x^2-\lambda^2 x)/2 - x^3/6}\sum_{l=-1}^{\binom{k}{2}-k}\frac{C(k,k+l)k^{1/2-k}}{n^l\sqrt{2\pi}}\left(\frac{np}{1-p}\right)^l \\ &= \big(1+o(1)\big)e^{(\lambda-x)^3/6-\lambda^3/6}\left[\sum_{l=-1}^{\lfloor\log n\rfloor}\frac{\gamma_l x^{3l/2}}{\sqrt{2\pi}} + \frac{R(n,k)}{\sqrt{2\pi}}\right], \end{aligned} \tag{2.12}$$

where

$$R(n,k) = \sum_{l=\lfloor \log n\rfloor+1}^{\binom{k}{2}-k} C(k,k+l)k^{1/2-k}\left(\frac{p}{1-p}\right)^l. \tag{2.13}$$

Clearly, Lemma 2.2 follows from (2.12) if we can show that in the limit $n\to\infty$, $R(n,k)$ tends to 0 uniformly over all x,λ considered.

To show this, we recall that by [2, Corollary 5.21], there exists an absolute constant $c>0$ such that

$$C(k,k+l) \le cl^{-l/2}k^{k+(3l-1)/2} \tag{2.14}$$

for all $1\le l\le \binom{k}{2}-k$. Substituting this bound into (2.13) gives

$$R(n,k) \le c\sum_{l>\lfloor \log n\rfloor}\left(\frac{k^{3/2}}{l^{1/2}}\frac{p}{1-p}\right)^l \le c\sum_{l>\lfloor \log n\rfloor}\left(\frac{\text{const}}{\sqrt{\log n}}\right)^l \le c\left(\frac{\text{const}}{\sqrt{\log n}}\right)^{\log n}\sum_{l>1}\frac{1}{2^l}, \tag{2.15}$$

where we have used that $k^{3/2}p/(1-p)$ is bounded uniformly by a constant, and the last inequality holds for n sufficiently large. Hence $R(n,k)=o(1)$, which completes the proof. □

Lemma 2.3 (Uniform weak convergence largest cluster)**.** *Let $\beta>\alpha$ and $b>a>0$ be arbitrary. As $n\to\infty$,*

$$\sup_{\substack{a\le x\le b\\ \alpha\le\lambda\le\beta}}\left|\mathbb{P}_{n,p_\lambda(n)}(|\mathcal{C}_{\max}|<xn^{2/3})-F_1(x;\lambda)\right|\to 0. \tag{2.16}$$

Proof. Fix $\varepsilon>0$. Recall that F_1 is continuous in both arguments, as follows for instance from [5, (3.25)]. Therefore, F_1 is *uniformly* continuous on $[a,b]\times[\alpha,\beta]$, and hence we can choose $a=x_1<\cdots<x_m=b$ and $\alpha=\lambda_1<\cdots<\lambda_m=\beta$ such that for all $1\le i,j\le m-1$,

$$\sup\left\{\left|F_1(x;\lambda)-F_1(x_i;\lambda_j)\right| : (x,\lambda)\in[x_i,x_{i+1}]\times[\lambda_j,\lambda_{j+1}]\right\}<\varepsilon. \tag{2.17}$$

For all $(x,\lambda)\in[a,b]\times[\alpha,\beta]$ set $g_n(x,\lambda)=\mathbb{P}_{n,p_\lambda(n)}(|\mathcal{C}_{\max}|<xn^{2/3})$. Note that $g_n(x,\lambda)$ is non-decreasing in x and non-increasing in λ. By definition (1.3) of F_1, there exists an $n_0=n_0(\varepsilon)$ such that for all $n\ge n_0$, $|F_1(x_i;\lambda_j)-g_n(x_i,\lambda_j)|<\varepsilon$ for every $1\le i,j\le m$. Therefore, if $(x,\lambda)\in[x_i,x_{i+1}]\times[\lambda_j,\lambda_{j+1}]$, then for all $n\ge n_0$,

$$g_n(x,\lambda)-F_1(x;\lambda)<g_n(x_{i+1},\lambda_j)-F_1(x_{i+1};\lambda_j)+\varepsilon<2\varepsilon, \tag{2.18}$$

and likewise $F_1(x;\lambda)-g_n(x,\lambda)<2\varepsilon$. Hence $g_n\to F_1$ uniformly on $[a,b]\times[\alpha,\beta]$. □

Proof of Theorem 1.1. We start by introducing some notation. Fix $a\le x_k\le\cdots\le x_1\le b$, and for $i=1,\dots,k$ set $l_i=l_i(n)=\lfloor x_i n^{2/3}\rfloor$. Now for $i=0,\dots,k$, let $m_i=m_i(n)=n-\sum_{j\le i}l_j$ and define $\lambda_i=\lambda_i(n)$ so that $p_{\lambda_i}(m_i)=p_\lambda(n)$, that is,

$$p_\lambda(n)=\frac{1}{n}\left(1+\lambda n^{-1/3}\right)=\frac{1}{m_i}\left(1+\lambda_i m_i^{-1/3}\right)=p_{\lambda_i}(m_i). \tag{2.19}$$

Finally, for $i=1,\dots,k$ let $y_i=y_i(n)$ be chosen such that $\lfloor y_i m_{i-1}^{2/3}\rfloor=\lfloor x_i n^{2/3}\rfloor=l_i$. It is straightforward to verify that $\lambda_i=\lambda-(x_1+\cdots+x_i)+o(1)$ and $y_i=x_i+o(1)$, where the error terms $o(1)$ are *uniform* over all choices of the x_i in $[a,b]$. Throughout this proof, the notation $o(1)$ will be used in this meaning.

Note that for all sufficiently large n, the y_i are all contained in a compact interval of the form $[a-\varepsilon, b+\varepsilon]$ for some $0<\varepsilon<a$, and the λ_i are also contained in a compact interval. Hence, since $l_{i+1} = \lfloor y_{i+1} m_i^{2/3} \rfloor$, it follows from Lemma 2.2 that for $i = 0, \dots, k-1$,

$$m_i \, \mathbb{P}_{m_i, p_{\lambda_i}(m_i)}(|\mathcal{C}(1)| = l_{i+1}) = y_{i+1} \, \Phi(y_{i+1}; \lambda_i) + o(1). \tag{2.20}$$

But because $\Phi(x;\lambda)$ is uniformly continuous on a compact set, the function on the right tends uniformly to $x_{i+1} \, \Phi\big(x_{i+1}; \lambda - \sum_{j<i} x_j\big)$. We conclude that

$$m_i \frac{n^{2/3}}{l_{i+1}} \, \mathbb{P}_{m_i, p_{\lambda_i}(m_i)}(|\mathcal{C}(1)| = l_{i+1}) = \Phi\Big(x_i; \lambda - \sum_{j<i} x_j\Big) + o(1). \tag{2.21}$$

Similarly, using that F_1 is uniformly continuous on a compact set, from Lemma 2.3 we obtain

$$\mathbb{P}_{m_k, p_{\lambda_k}(m_k)}(|\mathcal{C}_{\max}| < l_k) = F_1\big(x_k; \lambda - (x_1 + \cdots + x_k)\big) + o(1). \tag{2.22}$$

By Lemma 2.1, we see that we are interested in the product of the left-hand sides of (2.21) and (2.22). Since the right-hand sides of these equations are bounded uniformly over the x_i considered, it follows immediately that

$$n^{2k/3} \, \mathbb{P}_{n, p_\lambda(n)}(|\mathcal{C}_i| = l_i \ \forall i \le k, |\mathcal{C}_{k+1}| < l_k) = \Psi_k(x_1, \dots, x_k; \lambda) + o(1). \tag{2.23}$$

To complete the proof, set $l_{k+1} = l_k$, and note that, by Lemma 2.1 and (2.21),

$$\begin{aligned} &n^{2k/3} \, \mathbb{P}_{n, p_\lambda(n)}(|\mathcal{C}_i| = l_i \ \forall i \le k, |\mathcal{C}_{k+1}| = l_k) \\ &\qquad \le n^{-2/3} \prod_{i=0}^{k} \left(m_i \frac{n^{2/3}}{l_{i+1}} \mathbb{P}_{m_i, p_{\lambda_j}(m_i)}(|\mathcal{C}(1)| = l_{i+1}) \right) = o(1). \end{aligned} \tag{2.24}$$

Because $n^{2k/3} \, \mathbb{P}_{n, p_\lambda(n)}(|\mathcal{C}_i| = l_i \ \forall i \le k)$ is the sum of the left-hand sides of (2.23) and (2.24), this completes the proof of Theorem 1.1. □

Proof of Corollary 1.2. For any $x = (x_1, \dots, x_k) \in \mathbb{R}^k$, set

$$g_n(x) = n^{2k/3} \, \mathbb{P}_{n, p_\lambda(n)}(|\mathcal{C}_i| = \lfloor x_i n^{2/3} \rfloor \ \forall i \le k), \tag{2.25}$$

and notice that g_n is then a probability density with respect to k-dimensional Lebesgue measure. Let $X_n = (X_n^1, \dots, X_n^k)$ be a random vector having this density, and define the vector Y_n on the same space by setting $Y_n = \big(\lfloor X_n^1 n^{2/3} \rfloor n^{-2/3}, \dots, \lfloor X_n^k n^{2/3} \rfloor n^{-2/3}\big)$. Then Y_n has the same distribution as the vector $(|\mathcal{C}_1| n^{-2/3}, \dots, |\mathcal{C}_k| n^{-2/3})$ in $G\big(n, p_\lambda(n)\big)$. Now recall that by [1, Corollary 2], this vector converges in distribution to a limit which lies a.s. in $(0,\infty)^k$. Let P_λ be the law of the limit vector. Since $|X_n - Y_n| \to 0$ almost surely, P_λ is also the weak limit law of the X_n.

By Theorem 1.1, g_n converges pointwise to $\Psi_k(\,\cdot\,;\lambda)$ on $(0,\infty)^k$, and hence $\Psi_k(\,\cdot\,;\lambda)$ is integrable on $(0,\infty)^k$ by Fatou's lemma. Now let A be any compact set in $(0,\infty)^k$. Then g_n converges *uniformly* to $\Psi_k(\,\cdot\,;\lambda)$ on A, so we can apply dominated convergence to see that

$$\int_A g_n(x) \, dx \to \int_A \Psi_k(x;\lambda) \, dx = P_\lambda(A). \tag{2.26}$$

Since this holds for any compact A in $(0,\infty)^k$, it follows that $\Psi_k(\,\cdot\,;\lambda)$ is the density of P_λ with respect to Lebesgue measure. □

3 Unique identification of the limit distributions

In this section we will show that the system of differential equations (1.10) identifies the joint limiting distributions uniquely. Let us first observe that it suffices to show that there is only one solution to the differential equation

$$\frac{\partial}{\partial x} F_1(x;\lambda) = \Phi(x;\lambda) F_1(x;\lambda - x), \tag{3.1}$$

such that $x \mapsto F_1(x;\lambda)$ is the distribution function of a probability distribution for all $\lambda \in \mathbb{R}$. In the remainder of this section we will show that if F_1 satisfies (3.1) and $x \mapsto F_1(x;\lambda)$ is the distribution function of a probability distribution for all $\lambda \in \mathbb{R}$ then F_1 can be written as

$$F_1(x;\lambda) = 1 + e^{-\lambda^3/6} \sum_{k=1}^{\infty} \frac{(-1)^k}{k!} \int_x^\infty \cdots \int_x^\infty \prod_{i=1}^k \varphi(x_i) e^{(\lambda - x_1 - \cdots - x_k)^3/6} dx_1 \cdots dx_k, \tag{3.2}$$

where $\varphi(x) = G(x^{3/2})/x\sqrt{2\pi}$. This will prove Theorem 1.3 by our previous observation. To this end, we first note that it can be seen from Stirling's approximation and (1.5) that $G(s) = \exp\big(s^2/24 + o(s^2)\big)$ as $s \to \infty$, so that

$$\int_x^\infty \Phi(y;\lambda)\, dy = \int_x^\infty \exp[-\Omega(y^3)]\, dy < \infty \tag{3.3}$$

for all $\lambda \in \mathbb{R}$. To prove (3.2), we will make use of the following bound:

Lemma 3.1. *Let $a > \delta > 0$, $\lambda \in \mathbb{R}$ and $k > \lambda/\delta$, and write $\varphi(x) = G(x^{3/2})/x\sqrt{2\pi}$. Denote by $d_k x$ integration with respect to $x_1, \dots, x_k$. Then*

$$\begin{aligned} \int \cdots \int_{a < x_1 < \cdots < x_k} \prod_{i=1}^k \Phi\Big(x_i; \lambda - \sum_{j<i} x_j\Big)\, d_k x &= \frac{e^{-\lambda^3/6}}{k!} \int \cdots \int_{a < x_1, \dots, x_k} \prod_{i=1}^k \varphi(x_i) e^{\left(\lambda - \sum_{j \le k} x_j\right)^3/6}\, d_k x \\ &\le \frac{e^{-\lambda^3/6}}{k!} \Big(e^{\delta^3/6} \int_a^\infty \Phi(y;\delta)\, dy\Big)^k. \end{aligned} \tag{3.4}$$

Proof. Notice that $\Phi(x;\lambda) = \varphi(x) \exp\big(-\lambda^3/6 + (\lambda - x)^3/6\big)$, and that therefore we get

$$\prod_{i=1}^k \Phi\Big(x_i; \lambda - \sum_{j<i} x_j\Big) = \varphi(x_1) \cdots \varphi(x_k) \exp\Big(-\lambda^3/6 + \Big(\lambda - \sum_{j \le k} x_k\Big)^3/6\Big). \tag{3.5}$$

The equality in (3.4) now follows from the fact that the integrand is invariant under permutations of the variables. Next notice that if $\lambda < k\delta$ and $x_1, \dots, x_k > a > \delta$, then

$$(\lambda - x_1 - \cdots - x_k)^3 \le \big((\delta - x_1) + \cdots + (\delta - x_k)\big)^3 \le \sum_{i \le k} (\delta - x_i)^3, \tag{3.6}$$

since $\delta - x_i < 0$ for all $i = 1, \dots, k$ and $(u+v)^3 \ge u^3 + v^3$ for all $u, v \ge 0$. So it follows that

$$\int \cdots \int_{a < x_1, \dots, x_k} \prod_{i=1}^k \varphi(x_i) e^{\left(\lambda - \sum_{j \le k} x_j\right)^3/6}\, d_k x \le e^{k\delta^3/6} \int \cdots \int_{a < x_1, \dots, x_k} \prod_{i=1}^k \varphi(x_i) e^{-\delta^3/6 + (\delta - x_i)^3/6}\, d_k x, \tag{3.7}$$

which gives us the inequality in (3.4). □

Proof of Theorem 1.3. Applying (3.1) twice, we see that

$$\begin{aligned} F_1(x;\lambda) &= 1 - \int_x^\infty \Phi(x_1;\lambda) F_1(x_1;\lambda - x_1)\, dx_1 \\ &= 1 - \int_x^\infty \Phi(x_1;\lambda) \Big(1 - \int_{x_1}^\infty \Phi(x_2, \lambda - x_1) F_1(x_2;\lambda - x_1 - x_2)\, dx_2\Big)\, dx_1, \end{aligned} \tag{3.8}$$

and repeating this $m-2$ more times leads to

$$\begin{aligned} F_1(x;\lambda) = 1 + \sum_{k=1}^{m-1} (-1)^k \int\cdots\int_{x<x_1<\cdots<x_k} \prod_{i=1}^{k} \Phi\Big(x_i;\lambda - \sum_{j<i} x_j\Big)\, dx_1 \cdots dx_k \\ + (-1)^m \int\cdots\int_{x<x_1<\cdots<x_m} \prod_{i=1}^{m} \Phi\Big(x_i;\lambda - \sum_{j<i} x_j\Big) F_1\Big(x_m;\lambda - \sum_{j=1}^{m} x_j\Big)\, dx_1 \cdots dx_m. \end{aligned} \tag{3.9}$$

From Lemma 3.1 we see that for any $\varepsilon > 0$ we can choose $m = m(\varepsilon)$ such that

$$\int\cdots\int_{x<x_1<\cdots<x_m} \prod_{i=1}^{m} \Phi\Big(x_i;\lambda - \sum_{j<i} x_j\Big) F_1\Big(x_m;\lambda - \sum_{j=1}^{m} x_j\Big)\, dx_1 \cdots dx_m < \varepsilon, \tag{3.10}$$

where we have used that $F_1 \leq 1$. Hence (3.2) follows from (3.9) and Lemma 3.1. □

4 Discussion

We end the paper by mentioning a possibly useful extension of our results. Recall that the surplus of a connected component $\mathcal{C}$ is equal to the number of edges in $\mathcal{C}$ minus the number of vertices plus one, so that the surplus of a tree equals zero. There has been considerable interest in the surplus of the connected components of the Erdős-Rényi random graph (see e.g. [1, 3, 4] and the references therein). For example, in [1] and with $\sigma_n(k)$ denoting the surplus of $\mathcal{C}_k$, it is shown that

$$\Big(\big(|\mathcal{C}_1| n^{-2/3}, \sigma_n(1)\big), \ldots, \big(|\mathcal{C}_k| n^{-\frac{2}{3}}, \sigma_n(k)\big) \Big) \xrightarrow{d} \Big(\big(C_1^\lambda, \sigma(1)\big), \ldots, \big(C_k^\lambda, \sigma(k)\big) \Big), \tag{4.1}$$

for some bounded random variables $\sigma(k)$. A straightforward adaption of our proof of Theorem 1.1 will give that

$$\begin{aligned} n^{2k/3} \mathbb{P}_{n,p_\lambda(n)}\big(|\mathcal{C}_1| = \lfloor x_1 n^{2/3} \rfloor, \ldots, |\mathcal{C}_k| = \lfloor x_k n^{2/3} \rfloor, \sigma_n(1) = \sigma_1, \ldots, \sigma_n(k) = \sigma_k\big) \\ = \Psi_k(x_1, \ldots, x_k, \sigma_1, \ldots, \sigma_k; \lambda) + o(1), \end{aligned} \tag{4.2}$$

where $o(1)$ now is uniform in $\sigma_1, \ldots, \sigma_k$ and in $x_1, \ldots, x_k$ satisfying $a \leq x_1 \leq \cdots \leq x_k \leq b$ for some $0 < a < b$, and where we define

$$\Psi_k(x_1, \ldots, x_k, \sigma_1, \ldots, \sigma_k; \lambda) = \frac{F_1\big(x_k;\lambda - (x_1 + \cdots + x_k)\big)}{r_1! \cdots r_m!} \prod_{i=1}^{k} \Phi_{\sigma_i}\Big(x_i;\lambda - \sum_{j<i} x_j\Big), \tag{4.3}$$

with

$$\Phi_\sigma(x;\lambda) = \frac{\gamma_{\sigma-1} x^{3(\sigma-1)/2}}{x\sqrt{2\pi}} e^{-\lambda^3/6 + (\lambda - x)^3/6}. \tag{4.4}$$

Acknowledgements.

We thank Philippe Flajolet, Tomasz Łuczak and Boris Pittel for helpful discussions. The work of RvdH was supported in part by the Netherlands Organization for Scientific Research (NWO).

References

[1] D. Aldous. Brownian excursions, critical random graphs and the multiplicative coalescent. *Ann. Probab.*, 25(2):812–854, 1997.

[2] B. Bollobás. *Random graphs*, volume 73 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, second edition, 2001.

[3] S. Janson, D.E. Knuth, T. Łuczak, and B. Pittel. The birth of the giant component. *Random Structures Algorithms*, 4(3):231–358, 1993. With an introduction by the editors.

[4] T. Łuczak, B. Pittel, and J. C. Wierman. The structure of a random graph at the point of the phase transition. *Trans. Amer. Math. Soc.*, 341(2):721–748, 1994.

[5] B. Pittel. On the largest component of the random graph at a nearcritical stage. *J. Combin. Theory Ser. B*, 82(2):237–269, 2001.

[6] J. Spencer. Enumerating graphs and Brownian motion. *Comm. Pure Appl. Math.*, 50(3):291–294, 1997.

[7] E. M. Wright. The number of connected sparsely edged graphs. *J. Graph Theory*, 1(4):317–330, 1977.

[8] E. M. Wright. The number of connected sparsely edged graphs. II. Smooth graphs and blocks. *J. Graph Theory*, 2(4):299–305, 1978.

[9] E. M. Wright. The number of connected sparsely edged graphs. III. Asymptotic results. *J. Graph Theory*, 4(4):393–407, 1980.