A moment-generating formula for Erdős-Rényi component sizes

Balázs Ráth[1]

July 17, 2017

Abstract

We derive a simple formula characterizing the distribution of the size of the connected component of a fixed vertex in the Erdős-Rényi random graph which allows us to give elementary proofs of some results of [8] and [10] about the susceptibility in the subcritical graph and the CLT [4] for the size of the giant component in the supercritical graph.

KEYWORDS: Erdős-Rényi graph, generating function, susceptibility, giant component, central limit theorem
AMS MSC 2010: 60C05, 60F05, 82B26, 05A15

1 Introduction

The Erdős-Rényi graph $\mathcal{G}_{n,p}$, introduced in [7], is the random graph on n vertices where each pair of vertices is connected with probability p, independently from each other. For an introduction to this fundamental mathematical model of large networks, see [5, 11, 9].

We denote by $\mathbb{P}_{n,p}$ the law of $\mathcal{G}_{n,p}$ and $\mathbb{E}_{n,p}$ the corresponding expectation.

We assume that the vertex set of $\mathcal{G}_{n,p}$ is $[n] = \{1, \dots, n\}$ and we denote by $\mathcal{C}$ the connected component in $\mathcal{G}_{n,p}$ of the vertex indexed by 1. We denote by $|\mathcal{C}|$ the number of vertices of $\mathcal{C}$.

For any $n \in \mathbb{N}$, $p \in [0, 1]$, $j \in \mathbb{Z} \cap (-n, +\infty)$, and $k \in [n]$ we define

$$g_{n,p}(j, k) = (1 - p)^{jk} \prod_{i=0}^{k-1} \frac{n - i + j}{n - i}. \tag{1.1}$$

The central result of this short note is the following formula.

Proposition 1.1. For any $n \in \mathbb{N}$, $j \in \mathbb{Z} \cap (-n, +\infty)$ and $p \in [0, 1]$ we have

$$\mathbb{E}_{n,p} [\, g_{n,p}(j, |\mathcal{C}|) \,] = \frac{n + j}{n} (1 - \mathbb{P}_{n+j,p}[\, |\mathcal{C}| > n \,]) \,. \tag{1.2}$$

Note that if $j \le 0$ then the r.h.s. is simply $\frac{n+j}{n}$. We prove Proposition 1.1 in Section 2.

Remark 1.2. Let us define the $n \times n$ matrix M by $M_{j,k} = g_{n,p}(j, k)$ for $j \in \mathbb{Z} \cap (-n, 0]$ and $k \in [n]$. The matrix M is triangular with non-zero diagonal entries, hence it is invertible. Therefore Proposition 1.1 uniquely characterizes the distribution of $|\mathcal{C}|$ under $\mathbb{P}_{n,p}$.

Proposition 1.1 allows us to give short and self-contained proofs of some delicate results about the sizes of connected components of the Erdős-Rényi graph in the subcritical (see Theorem 1.4) as well as the supercritical (see Theorem 1.6) cases. First, we give a short non-rigorous demonstration of how our formula is used in Remark 1.3.

[1]MTA-BME Stochastics Research Group, Budapest University of Technology and Economics, Hungary. E-mail: rathb@math.bme.hu

When we study the phase transition of the Erdős-Rényi graph, it is natural to introduce a parameter $t \in \mathbb{R}_+$ and to study $\mathcal{G}_{n,p}$ for

$$p = p(t,n) = 1 - e^{-t/n}. \tag{1.3}$$

We will fix this relation between p and t throughout this paper.

For any $n \in \mathbb{N}$, $\lambda \in \mathbb{R}$, and $k \in [n]$ we define

$$f_{n,t}(\lambda, k) = \prod_{i=0}^{k-1} e^{-\lambda t} \cdot \left(1 + \frac{\lambda}{1 - \frac{i}{n}} \right), \tag{1.4}$$

so that we have $f_{n,t}(\frac{j}{n}, k) = g_{n,p}(j,k)$ if $j \in \mathbb{Z} \cap (-n, +\infty)$ and thus

$$\mathbb{E}_{n,p}\left[f_{n,t}(\lambda, |\mathcal{C}|) \right] \overset{(1.2)}{=} (1+\lambda)\left(1 - \mathbb{P}_{(1+\lambda)n,p}[\,|\mathcal{C}| > n\,]\right), \quad \lambda \in \frac{\mathbb{Z}}{n} \cap (-1, +\infty). \tag{1.5}$$

Remark 1.3. If we fix $t < 1$ and (non-rigorously) denote $G_t(z) = \lim_{n\to\infty} \mathbb{E}_{n,p(n,t)}[z^{|\mathcal{C}|}]$ for any $z \in [0,1]$, then for $\lambda = z - 1$ we (non-rigorously) obtain

$$z = 1 + \lambda \overset{(1.5)}{=} \lim_{n\to\infty} \mathbb{E}_{n,p(n,t)}\left[f_{n,t}(\lambda, |\mathcal{C}|)\right] \overset{(1.4)}{=} G_t\left(e^{-\lambda t} \cdot (1+\lambda)\right) = G_t\left(e^{(1-z)t} z\right). \tag{1.6}$$

Now it is known that if $p = 1 - e^{-t/n}$ and $n \to \infty$ then $|\mathcal{C}|$ converges in distribution to the total number of offspring in a subcritical Galton-Watson branching process with POI(t) offspring distribution (see [3, Section 10.4]), i.e., $|\mathcal{C}|$ has Borel distribution with parameter t (see [2, Section 2.2] or [10, Section 7]). The generating function G_t of the Borel distribution with parameter t is known to be characterized by the identity $G_t(z) \equiv z e^{(G_t(z)-1)t}$ (see [3, Sections 10.4 and 10.5]), which is in turn equivalent to $z \equiv G_t\left(e^{(1-z)t} z\right)$, therefore a more rigorous version of (1.6) can be used to show that the distribution $|\mathcal{C}|$ weakly converges to the Borel distribution with parameter t as $n \to \infty$.

Now we state our rigorous results. We will use the Bachmann-Landau big O notation: we write $f(n,t) = \mathcal{O}(g(n,t))$ if there exists a universal constant C such that $f(n,t) \le C g(n,t)$ for any $n \in \mathbb{N}$ and any t in an explicitly specified domain. We write $f(n) = \mathcal{O}(g(n))$ if there exists a constant C (that may depend on t) such that $f(n) \le C g(n)$ for any $n \in \mathbb{N}$.

We will give a short and self-contained proof of some results of [8] and [10]:

Theorem 1.4. For any $t \in [0, 1 - n^{-1/3}]$ we have

$$\mathbb{E}_{n,p}(|\mathcal{C}|) = \frac{1}{1-t} + \frac{\frac{t^2}{2} - t}{(1-t)^4} \frac{1}{n} + \mathcal{O}\left(\frac{1}{(1-t)^7} \frac{1}{n^2} \right), \tag{1.7}$$

$$\mathbb{E}_{n,p}(|\mathcal{C}|^2) = \frac{1}{(1-t)^3} + \mathcal{O}\left(\frac{1}{(1-t)^6} \frac{1}{n} \right). \tag{1.8}$$

We will prove Theorem 1.4 in Section 3.

Remark 1.5. $\mathbb{E}_{n,p}(|\mathcal{C}|)$ is often called the *susceptibility* of the Erdős-Rényi graph.

(i) Equation (1.15) of [8, Theorem 1.2] states that if $p = \frac{\mu}{n-1}$ and $0 < \mu < 1$ then

$$\mathbb{E}_{n,p}(|\mathcal{C}|) = \frac{1}{1-\mu} - \frac{2\mu^2 - \mu^4}{2(1-\mu)^4} \frac{1}{n} + \mathcal{O}\left(\frac{1}{n^2} \right). \tag{1.9}$$

Now (1.9) follows from (1.7) if we take into account that $\mu = (n-1)(1 - e^{-t/n})$. The proof of (1.9) in [8, Section 2] uses a coupling of the breadth-first exploration process of $\mathcal{C}$ and a process related to a branching random walk. Our proof of (1.7) is completely different as it only uses Proposition 1.1.

(ii) Equation (1.3) of [10, Theorem 1.1] follows from our (1.7). In fact it already follows from our short Lemma 3.2, see (3.7). Our (1.8) is equivalent to one of the statements about S_3 in [10, Theorem 3.4]. The proofs of these results in [10, Section 3] use differential equations (in the variable t) and are completely different from ours.

(iii) Both statements of Theorem 1.4 give something meaningful in the whole subcritical regime outside the critical window, e.g., the first term of the r.h.s. of (1.7) is much bigger than the second one, which is much bigger than the third one if $(1-t)^3 n \gg 1$.

We also give a short and self-contained proof of the central limit theorem for the size of the giant connected component of $\mathcal{G}_{n,p}$, [4, Theorem 1.1], which only uses Proposition 1.1, see Theorem 1.6 below. We begin with some notation.

Given some $t > 1$ let us define the function $\varphi : [0,1) \to \mathbb{R}$ by

$$\varphi(x) = -xt - \ln(1-x). \tag{1.10}$$

Then φ is a convex function satisfying $\varphi(0) = 0$, $\varphi'(0) < 0$ and $\varphi(1_-) = +\infty$.

Given $t > 1$ define $\theta = \theta(t) \in (0,1)$ to be unique number for which

$$\varphi(\theta) = 0, \quad \text{or, equivalently} \quad e^{t\theta}(1-\theta) = 1. \tag{1.11}$$

We note that

$$\varphi'(\theta) = -t + \frac{1}{1-\theta} > 0. \tag{1.12}$$

Recall the notion of $p = p(t,n) = 1 - e^{-t/n}$ from (1.3).

Theorem 1.6. Let us denote by $|\mathcal{C}_{max}|$ the size of the largest connected component of $\mathcal{G}_{n,p}$. For any $t > 1$ we have

$$\lim_{n\to\infty} \mathbb{P}_{n,p(t,n)} \left[\frac{|\mathcal{C}_{max}| - \theta n}{\sigma\sqrt{n}} \le x \right] = \Phi(x), \quad \text{where} \quad \sigma = \frac{\sqrt{\theta}}{\varphi'(\theta)\sqrt{1-\theta}} \tag{1.13}$$

and $\Phi(x)$ is the c.d.f. of the standard normal distribution.

We prove Theorem 1.6 in Section 4. Our proof is completely different from the proof of [4], which uses exploration processes. In contrast to [4, Theorem 1.1], our elementary proof does not provide information about the rate of convergence to the limit distribution.

Remark 1.7. We believe that (1.2) can also be used to give elementary alternative proofs of other results on the size of the connected component of a fixed vertex in the Erdős-Rényi graph, e.g. the relation between the scaling limit of $|\mathcal{C}|$ in the critical window and the excursion length measure of Brownian motion with parabolic drift (see [1, Section 5.2]) or finding the time t for which $\frac{\mathrm{d}}{\mathrm{d}t}\ln(\mathbb{E}_{n,p(n,t)}(|\mathcal{C}|))$ is maximal (see [12]).

We discuss the origins and possible extensions of (1.2) to other models in Remark 2.2.

2 Proof of Proposition 1.1

Lemma 2.1. For any $M, N \in \mathbb{N}$, $p \in [0,1]$, and $k \in \{1,\dots,N\}$ we have

$$\mathbb{P}_{M,p}[\,|\mathcal{C}| = k\,] = \mathbb{P}_{N,p}[\,|\mathcal{C}| = k\,] \cdot (1-p)^{(M-N)k} \prod_{i=1}^{k-1} \frac{M-i}{N-i}. \tag{2.1}$$

Proof. If $k > M$ then both sides of (2.1) are zero. Thus w.l.o.g. we can assume $k \le M \wedge N$. Now we observe that if we prove (2.1) for some $M \le N$, then we also obtain (2.1) for $M' = N$ and $N' = M$ by rearranging the formula (2.1), thus we may assume w.l.o.g. that $k \le M \le N$. In order to prove (2.1) it is enough to show

$$\binom{M-1}{k-1}^{-1} \mathbb{P}_{M,p}[\,|\mathcal{C}| = k\,] \cdot (1-p)^{k(N-M)} = \binom{N-1}{k-1}^{-1} \mathbb{P}_{N,p}[\,|\mathcal{C}| = k\,]. \tag{2.2}$$

Now if we denote by $V(\mathcal{C})$ the vertex set of $\mathcal{C}$ then

$$\mathbb{P}_{N,p}[\,|\mathcal{C}| = k\,] = \binom{N-1}{k-1} \mathbb{P}_{N,p}[\,V(\mathcal{C}) = [k]\,], \tag{2.3}$$

since $\mathbb{P}_{N,p}$ is invariant under the permutation of vertices and there are $\binom{N-1}{k-1}$ subsets of $[N]$ with cardinality k that contain the vertex indexed by 1. Using (2.3) for $\mathbb{P}_{N,p}$ as well as $\mathbb{P}_{M,p}$, the formula (2.2) reduces to showing

$$\mathbb{P}_{M,p}[\,V(\mathcal{C}) = [k]\,] \cdot (1-p)^{k(N-M)} = \mathbb{P}_{N,p}[\,V(\mathcal{C}) = [k]\,]. \tag{2.4}$$

Now (2.4) holds since $V(\mathcal{C}) = [k]$ in $\mathcal{G}_{N,p}$ if and only if $V(\mathcal{C}) = [k]$ in $\mathcal{G}_{M,p}$ and there are no edges in $\mathcal{G}_{N,p}$ between $[k]$ and $[N] \setminus [M]$. This completes the proof of Lemma 2.1. $\square$

Proof of Proposition 1.1. For any $n \in \mathbb{N}$, $j \in \mathbb{Z} \cap (-n, +\infty)$ and $p \in [0,1]$ we have

$$\mathbb{E}_{n,p}\left[\, g_{n,p}(j, |\mathcal{C}|)\right] \overset{(1.1)}{=} \frac{n+j}{n} \sum_{k=1}^{n} \mathbb{P}_{n,p}[\,|\mathcal{C}| = k\,] \cdot (1-p)^{jk} \prod_{i=1}^{k-1} \frac{n+j-i}{n-i} \overset{(*)}{=}$$
$$\frac{n+j}{n} \sum_{k=1}^{n} \mathbb{P}_{n+j,p}[\,|\mathcal{C}| = k\,] = \frac{n+j}{n} \left(1 - \mathbb{P}_{n+j,p}[\,|\mathcal{C}| > n\,]\right), \tag{2.5}$$

where in $(*)$ we used (2.1) with $n = N$ and $M = n + j$. The proof of (1.2) is complete. $\square$

Remark 2.2.

(i) Our original proof of Proposition 1.1 used the so-called *rigid representation* of the time evolution of the component size structure of the Erdős-Rényi graph, see [13, Section 6.1.1, Case 1]. In a nutshell, if $Y_k = t - X_k$, $k \in [n]$, where $X_1, X_2, \dots, X_n$ denote independent exponentially distributed random variables $X_k \sim \mathrm{EXP}\left(1 - \frac{k}{n}\right)$, then $\tau = \min\{\, k \,:\, Y_1 + \dots + Y_k < 0 \,\}$ has the same distribution as $|\mathcal{C}|$ under $\mathbb{P}_{n,p}$, $p = 1 - e^{-t/n}$. We chose to include an elementary proof instead in order to keep the paper self-contained.

(ii) It is formally possible to extend Proposition 1.1 (in particular (2.4)) to the inhomogeneous random graph model of [6], however it is hard to assess at this point whether the resulting generalization of (1.2) is simple enough to be useful.

3 Proof of Theorem 1.4

The basic idea is to treat $\mathbb{E}_{n,p}\left[g_{n,p}(j, |\mathcal{C}|)\right]$ as the generating function of $|\mathcal{C}|$, c.f. Remark 1.3. Thus if we want to obtain information about the first and second moments of $|\mathcal{C}|$, we have to "differentiate" with respect to the variable j twice. Since j can only take integer values, we have to consider the first order discrete differences $g_{n,p}(j, |\mathcal{C}|) - g_{n,p}(0, |\mathcal{C}|)$ for $j = -1$ and $j = -2$ in the proof of Lemmas 3.2 and 3.4, and the second order discrete difference (i.e., the difference of the first order differences) in the proof of Lemma 3.5.

The statement of Lemma 3.1 is equivalent to [10, Lemma 3.2] (which is proved using differential equations), moreover it also classically follows from the fact that $|\mathcal{C}|$ is stochastically dominated by a subcritical branching process if $t < 1$. Despite of this, we chose to include a proof of Lemma 3.1 which only uses Proposition 1.1 in order to keep the paper self-contained.

Recall our convention $p = 1 - e^{-t/n}$ from (1.3).

Lemma 3.1. If $t \in (0,1)$ then

$$\mathbb{E}_{n,p}\left(|\mathcal{C}|^i\right) = \mathcal{O}\left(\frac{1}{(1-t)^{2i-1}}\right), \qquad i \in \mathbb{N}. \tag{3.1}$$

Proof. W.l.o.g. we assume $t \in [\frac{1}{2}, 1)$ and $\frac{100}{1-t} \leq n$. For any $j \geq 0$ we have

$$\mathbb{E}_{n,p}\left[\left(e^{-tj/n}\left(1+\frac{j}{n}\right)\right)^{|\mathcal{C}|}\right] \overset{(1.4)}{\leq} \mathbb{E}_{n,p}\left[f_{n,t}\left(\frac{j}{n}, |\mathcal{C}|\right)\right] \overset{(1.5)}{\leq} 1 + \frac{j}{n}. \tag{3.2}$$

Note that if we let $\widetilde{\lambda} = \frac{1}{t} - 1$ then we have $\max_\lambda e^{-\lambda t}(1+\lambda) = e^{-\widetilde{\lambda} t}(1+\widetilde{\lambda}) = \frac{1}{t}e^{t-1} > e^{\frac{1}{2}(1-t)^2}$. Thus choosing $j^* = \lfloor n \cdot \left(\frac{1}{t} - 1\right)\rfloor$ we can use $\frac{100}{1-t} \leq n$ to infer

$$e^{-tj^*/n}\left(1 + \frac{j^*}{n}\right) \geq e^{\frac{1}{4}(1-t)^2}, \tag{3.3}$$

therefore for any $i \in \mathbb{N}$ we have

$$1 + \frac{1}{i!}\frac{1}{4^i}(1-t)^{2i}\mathbb{E}_{n,p}\left(|\mathcal{C}|^i\right) \leq \mathbb{E}_{n,p}\left[\sum_{\ell=0}^{\infty} \frac{\left(\frac{1}{4}(1-t)^2|\mathcal{C}|\right)^\ell}{\ell!}\right] =$$
$$\mathbb{E}_{n,p}\left[e^{\frac{1}{4}(1-t)^2|\mathcal{C}|}\right] \overset{(3.3)}{\leq} \mathbb{E}_{n,p}\left[\left(e^{-tj^*/n}\left(1+\frac{j^*}{n}\right)\right)^{|\mathcal{C}|}\right] \overset{(3.2)}{\leq} 1 + \frac{j^*}{n} \leq \frac{1}{t}, \tag{3.4}$$

from which (3.1) follows if $t \in [\frac{1}{2}, 1)$. □

Lemma 3.2. For any $t \in [0,1)$ we have

$$1 = (1-t)\mathbb{E}_{n,p}(|\mathcal{C}|) + \left(t - \frac{t^2}{2}\right)\frac{\mathbb{E}_{n,p}(|\mathcal{C}|^2)}{n} + \left(\frac{t^2}{2} - \frac{t^3}{6}\right)\frac{\mathbb{E}_{n,p}(|\mathcal{C}|^3)}{n^2} + \mathcal{O}\left(\frac{1}{(1-t)^7 n^3}\right). \tag{3.5}$$

Before we prove Lemma 3.2, let us state the immediate

Corollary 3.3. Applying (3.1) to $\mathbb{E}_{n,p}(|\mathcal{C}|^3)$ in (3.5) we obtain

$$\mathbb{E}_{n,p}(|\mathcal{C}|) = \frac{1}{1-t} + \frac{\frac{t^2}{2} - t}{1-t}\frac{\mathbb{E}_{n,p}(|\mathcal{C}|^2)}{n} + \mathcal{O}\left(\frac{1}{(1-t)^6 n^2}\right), \quad t \in [0, 1 - n^{-1/3}]. \tag{3.6}$$

Applying (3.1) to $\mathbb{E}_{n,p}(|\mathcal{C}|^2)$ in (3.6) we obtain

$$\mathbb{E}_{n,p}(|\mathcal{C}|) = \frac{1}{1-t} + \mathcal{O}\left(\frac{1}{(1-t)^4 n}\right), \quad t \in [0, 1 - n^{-1/3}]. \tag{3.7}$$

Proof of Lemma 3.2. Let $k \in [n]$. We begin with with observing that (1.1) is a telescopic product if $j = -1$ and then we apply Taylor expansion:

$$g_{n,p}(-1,k) \overset{(1.1),(1.3)}{=} e^{tk/n}\left(1 - \frac{k}{n}\right) = \left(\sum_{i=0}^{3}\frac{1}{i!}\frac{t^i k^i}{n^i} + \mathcal{O}\left(\frac{k^4}{n^4}\right)\right)\left(1 - \frac{k}{n}\right)$$
$$= 1 + (t-1)\frac{k}{n} + \left(\frac{t^2}{2} - t\right)\frac{k^2}{n^2} + \left(\frac{t^3}{6} - \frac{t^2}{2}\right)\frac{k^3}{n^3} + \mathcal{O}\left(\frac{k^4}{n^4}\right). \tag{3.8}$$

Combining (3.8) with Proposition 1.1 we obtain

$$1-\frac{1}{n}=1+(t-1)\frac{\mathbb{E}_{n,p}(|\mathcal{C}|)}{n}+ \left(\frac{t^2}{2}-t\right)\frac{\mathbb{E}_{n,p}(|\mathcal{C}|^2)}{n^2}+\left(\frac{t^3}{6}-\frac{t^2}{2}\right)\frac{\mathbb{E}_{n,p}(|\mathcal{C}|^3)}{n^3}+\mathcal{O}\left(\frac{\mathbb{E}_{n,p}(|\mathcal{C}|^4)}{n^4}\right). \quad (3.9)$$

Subtracting one from both sides of (3.9), multiplying the result by $-n$ and applying (3.1) to $\mathbb{E}_{n,p}(|\mathcal{C}|^4)$, we obtain (3.5). □

Lemma 3.4. For any $t\in[0,1-n^{-1/3}]$ we have

$$-2=(2t-2)\mathbb{E}_{n,p}[|\mathcal{C}|]+(1-4t+2t^2)\frac{\mathbb{E}_{n,p}[|\mathcal{C}|^2]}{n}-\frac{\mathbb{E}_{n,p}[|\mathcal{C}|]}{n}+ (2t-4t^2+\frac{4}{3}t^3)\frac{\mathbb{E}_{n,p}[|\mathcal{C}|^3]}{n^2}+\mathcal{O}\left(\frac{1}{(1-t)^7n^3}\right). \quad (3.10)$$

Proof. Let $k\in[n]$. We begin with a calculation similar to (3.8):

$$g_{n,p}(-2,k)\overset{(1.1),(1.3)}{=}e^{2tk/n}\left(1-\frac{k}{n}\right)\left(1-\frac{k}{n-1}\right)= \left(\sum_{i=0}^{3}\frac{1}{i!}\frac{2^it^ik^i}{n^i}+\mathcal{O}\left(\frac{k^4}{n^4}\right)\right)\left(1-\frac{k}{n}\right)\left(1-\frac{k}{n}\left(\sum_{j=0}^{2}\frac{1}{n^j}+\mathcal{O}\left(\frac{1}{n^3}\right)\right)\right)=$$

$$1+(2t-2)\frac{k}{n}+(1-4t+2t^2)\frac{k^2}{n^2}-\frac{k}{n^2}+(2t-4t^2+\frac{4}{3}t^3)\frac{k^3}{n^3}+(1-2t)\frac{k^2}{n^3}-\frac{k}{n^3}+\mathcal{O}\left(\frac{k^4}{n^4}\right). \quad (3.11)$$

From (3.11) and Proposition 1.1 we obtain

$$1-\frac{2}{n}=1+(2t-2)\frac{\mathbb{E}_{n,p}[|\mathcal{C}|]}{n}+(1-4t+2t^2)\frac{\mathbb{E}_{n,p}[|\mathcal{C}|^2]}{n^2}-\frac{\mathbb{E}_{n,p}[|\mathcal{C}|]}{n^2}+ (2t-4t^2+\frac{4}{3}t^3)\frac{\mathbb{E}_{n,p}[|\mathcal{C}|^3]}{n^3}+(1-2t)\frac{\mathbb{E}_{n,p}[|\mathcal{C}|^2]}{n^3}-\frac{\mathbb{E}_{n,p}[|\mathcal{C}|]}{n^3}+\mathcal{O}\left(\frac{\mathbb{E}_{n,p}[|\mathcal{C}|^4]}{n^4}\right). \quad (3.12)$$

Subtracting one from both sides of (3.12), multiplying the result by n and applying (3.1) to the last three terms of (3.12), we obtain (3.10).

□

Lemma 3.5. For any $t\in[0,1-n^{-1/3}]$ we have

$$\mathbb{E}_{n,p}(|\mathcal{C}|^2)=\frac{\mathbb{E}_{n,p}(|\mathcal{C}|)}{(1-t)^2}+\mathcal{O}\left(\frac{1}{(1-t)^6n}\right). \quad (3.13)$$

Proof. Adding (3.10) to twice (3.5) we obtain

$$0=(1-2t+t^2)\frac{\mathbb{E}_{n,p}(|\mathcal{C}|^2)}{n}-\frac{\mathbb{E}_{n,p}(|\mathcal{C}|)}{n}+(t^3-3t^2+2t)\frac{\mathbb{E}_{n,p}(|\mathcal{C}|^3)}{n^2}+\mathcal{O}\left(\frac{1}{(1-t)^7n^3}\right). \quad (3.14)$$

Rearranging (3.14) and multiplying by n we obtain

$$\mathbb{E}_{n,p}(|\mathcal{C}|)=(1-t)^2\mathbb{E}_{n,p}(|\mathcal{C}|^2)+t(t-1)(t-2)\frac{\mathbb{E}_{n,p}(|\mathcal{C}|^3)}{n}+\mathcal{O}\left(\frac{1}{(1-t)^7n^2}\right). \quad (3.15)$$

Dividing both sides of (3.15) by $(1-t)^2$ we use (3.1) to obtain (3.13). □

Proof of Theorem 1.4. From (3.7) and (3.13) we obtain (1.8).
Plugging (1.8) into (3.6) we obtain (1.7).

□

4 Proof of Theorem 1.6

We will deduce Theorem 1.6 (i.e., the CLT for $|\mathcal{C}_{max}|$) from Lemma 4.1 (i.e., the CLT for $|\mathcal{C}|$). We deduce Lemma 4.1 from Lemmas 4.2 and 4.3 using that convergence of moment generating functions implies weak convergence of probability distributions. We prove Lemmas 4.2 and 4.3 by viewing (1.5) as moment generating function identity. The crux of the proof of Lemma 4.2 is (4.24) and the crux of the proof of Lemma 4.3 is (4.38).

Throughout this section we fix $t > 1$. Recall the notion of $\varphi : [0, 1) \to \mathbb{R}$ from (1.10) and $\theta = \theta(t) \in (0, 1)$ from (1.11). Recall the notion of $p = p(t, n) = 1 - e^{-t/n}$ from (1.3).

We will often use the shorthand $\mathbb{P}$ for $\mathbb{P}_{n,p(n,t)}$ and $\mathbb{E}$ for $\mathbb{E}_{n,p(n,t)}$.

If X is a random variable and A is an event, we will denote $\mathbb{E}(X; A) := \mathbb{E}(X\mathbb{1}_A)$.

Lemma 4.1. Let us define σ as in (1.13). For any $x \in \mathbb{R}$ we have

$$\lim_{n\to\infty} \mathbb{P}_{n,p(n,t)}\left[\frac{|\mathcal{C}| - \theta n}{\sigma\sqrt{n}} \le x\right] = (1-\theta) + \theta\Phi(x). \tag{4.1}$$

Before we prove Lemma 4.1, we use it to prove Theorem 1.6.

Proof of Theorem 1.6. Given some $\gamma \in \mathbb{R}_+$, let us define the events

$$A_{n,\gamma} = \{\, \mathcal{G}_{n,p} \text{ has a component bigger than } \theta n + \gamma\sqrt{n}\,\}, \tag{4.2}$$
$$B_{n,\gamma} = \{\, \mathcal{G}_{n,p} \text{ has at least two components bigger than } \theta n - \gamma\sqrt{n}\,\}, \tag{4.3}$$
$$C_{n,\gamma} = \{\, \mathcal{G}_{n,p} \text{ has a component bigger than } \theta n - \gamma\sqrt{n}\,\}. \tag{4.4}$$

Before we deduce Theorem 1.6 from Lemma 4.1, we will first show

$$\lim_{\gamma\to\infty} \limsup_{n\to\infty} \mathbb{P}_{n,p(n,t)}[A_{n,\gamma}] = 0, \tag{4.5}$$
$$\lim_{\gamma\to\infty} \limsup_{n\to\infty} \mathbb{P}_{n,p(n,t)}[B_{n,\gamma}] = 0, \tag{4.6}$$
$$\lim_{\gamma\to\infty} \liminf_{n\to\infty} \mathbb{P}_{n,p(n,t)}[C_{n,\gamma}] = 1. \tag{4.7}$$

Proof of (4.5)**:** Since $|\mathcal{C}|$ is a size-biased sample from the collection of component sizes of the graph $\mathcal{G}_{n,p}$, we have $\mathbb{P}\left[|\mathcal{C}| > \theta n + \gamma\sqrt{n}\right] = \mathbb{P}\left[|\mathcal{C}| > \theta n + \gamma\sqrt{n} \,\middle|\, A_{n,\gamma}\right] \mathbb{P}[A_{n,\gamma}] \ge \theta\mathbb{P}[A_{n,\gamma}]$, thus from (4.1) we can infer (4.5).

Proof of (4.6)**:** Let $v = \min\{[n] \setminus \mathcal{C}\}$ and let $\mathcal{C}^*$ denote the connected component of the vertex v in $\mathcal{G}_{n,p}$. If $\mathcal{C} = [n]$, let $\mathcal{C}^* = \emptyset$. Note that $\mathcal{C} \cap \mathcal{C}^* = \emptyset$ and that we have

$$\mathbb{P}_{n,p}\left[\, |\mathcal{C}^*| = \ell \,\middle|\, |\mathcal{C}| = k \,\right] = \mathbb{P}_{n-k,p}\left[\, |\mathcal{C}| = \ell \,\right]. \tag{4.8}$$

Let us denote $\widetilde{n} = \lfloor n - \theta n + \gamma\sqrt{n} \rfloor$. We have

$$\left(\theta - \frac{\gamma}{\sqrt{n}}\right)^2 \mathbb{P}[B_{n,\gamma}] \le \mathbb{P}\left[\left\{\frac{|\mathcal{C}|}{n} > \theta - \frac{\gamma}{\sqrt{n}}\right\} \cap \left\{\frac{|\mathcal{C}^*|}{n} > \theta - \frac{\gamma}{\sqrt{n}}\right\} \cap B_{n,\gamma}\right] =$$
$$\mathbb{P}\left[\, |\mathcal{C}^*| > \theta n - \gamma\sqrt{n} \,\middle|\, |\mathcal{C}| > \theta n - \gamma\sqrt{n}\,\right] \mathbb{P}\left[\, |\mathcal{C}| > \theta n - \gamma\sqrt{n}\,\right] \le$$
$$\mathbb{P}\left[\, |\mathcal{C}^*| > \theta n - \gamma\sqrt{n} \,\middle|\, |\mathcal{C}| > \theta n - \gamma\sqrt{n}\,\right] \overset{(4.8)}{\le} \mathbb{P}_{\widetilde{n},p}\left[\, |\mathcal{C}| > \theta n - \gamma\sqrt{n}\,\right] \le \frac{\mathbb{E}_{\widetilde{n},p}\left[|\mathcal{C}|\right]}{\theta n - \gamma\sqrt{n}}. \tag{4.9}$$

Now we observe that $\mathcal{G}_{\widetilde{n},p}$ is a *subcritical* Erdős-Rényi graph, since

$$\lim_{n\to\infty} \widetilde{n}p \overset{(1.3)}{=} \lim_{n\to\infty} \lfloor n - \theta n + \gamma\sqrt{n} \rfloor \cdot (1 - e^{-t/n}) = (1-\theta)t \overset{(1.12)}{<} 1.$$

Now $\mathbb{E}_{\widetilde{n},p}\left[|\mathcal{C}|\right]$ remains bounded as $n \to \infty$ by (3.1), hence (4.6) follows from (4.9).

Proof of (4.7)**:** Let us define $D_{n,\gamma} = A^c_{n,\gamma} \cap B^c_{n,\gamma} \cap C_{n,\gamma}$.

Now we bound the probability of $C_{n,\gamma}$ from below.

$$\mathbb{P}\left[|\mathcal{C}| > \theta n - \gamma\sqrt{n}\right] = \mathbb{P}\left[\left\{|\mathcal{C}| > \theta n - \gamma\sqrt{n}\right\} \cap C_{n,\gamma}\right] \le \\ \mathbb{P}\left[\, |\mathcal{C}| > \theta n - \gamma\sqrt{n} \;\middle|\; D_{n,\gamma}\,\right]\mathbb{P}[\,D_{n,\gamma}\,] + \mathbb{P}[A_{n,\gamma} \cap B_{n,\gamma}] \le \\ \left(\theta + \frac{\gamma}{\sqrt{n}}\right)\mathbb{P}[D_{n,\gamma}] + \mathbb{P}[A_{n,\gamma} \cap B_{n,\gamma}] \le \left(\theta + \frac{\gamma}{\sqrt{n}}\right)\mathbb{P}[C_{n,\gamma}] + \mathbb{P}[A_{n,\gamma} \cap B_{n,\gamma}]. \quad (4.10)$$

From (4.1), (4.5), (4.6) and (4.10) we obtain (4.7) and

$$\lim_{\gamma\to\infty}\lim_{n\to\infty}\mathbb{P}_{n,p(n,t)}[D_{n,\gamma}] = 1. \quad (4.11)$$

Proof of (1.13)**:** We first note that if the event $D_{n,\gamma}$ occurs then $\mathcal{G}_{n,p}$ has exactly one component $\mathcal{C}_{max}$ which satisfies $|\mathcal{C}_{max}| \in [\,n\theta - \gamma\sqrt{n}, n\theta + \gamma\sqrt{n}\,]$ and $\mathcal{C}_{max}$ is the unique largest component in $\mathcal{G}_{n,p}$. Hence, for any $a < b \in \mathbb{R}$ we have

$$\theta(\Phi(b) - \Phi(a)) \overset{(4.1),(4.11)}{=} \lim_{\gamma\to\infty}\lim_{n\to\infty}\mathbb{P}_{n,p(n,t)}\left[\frac{|\mathcal{C}| - \theta n}{\sigma\sqrt{n}} \in [a,b],\, D_{n,\gamma}\right] = \\ \lim_{\gamma\to\infty}\lim_{n\to\infty}\mathbb{P}_{n,p(n,t)}\left[\mathcal{C} = \mathcal{C}_{max},\, \frac{|\mathcal{C}_{max}| - \theta n}{\sigma\sqrt{n}} \in [a,b],\, D_{n,\gamma}\right] \overset{(4.11)}{=} \\ \theta\lim_{n\to\infty}\mathbb{P}_{n,p(n,t)}\left[\frac{|\mathcal{C}_{max}| - \theta n}{\sigma\sqrt{n}} \in [a,b]\right]. \quad (4.12)$$

This completes the proof of Theorem 1.6 given Lemma 4.1. □

We will deduce Lemma 4.1 from Lemmas 4.2 and 4.3 below.

Let us subdivide the interval $[1, n)$ into five disjoint sub-intervals:

$$I_n = [1, n^{1/4}), \quad J_n = [n^{1/4}, n^{3/4}), \quad K_n = [n^{3/4}, \theta n - n^{5/8}), \quad (4.13)$$

$$\widetilde{I}_n = [\theta n - n^{5/8}, \theta n + n^{5/8}), \quad \widetilde{K}_n = [\theta n + n^{5/8}, n). \quad (4.14)$$

The choice of the exponents $\frac{1}{4}$, $\frac{3}{4}$ and $\frac{5}{8}$ above is somewhat arbitrary.

Lemma 4.2. We have

$$\lim_{n\to\infty}\mathbb{P}_{n,p(n,t)}\left(|\mathcal{C}| \in I_n\right) = 1 - \theta. \quad (4.15)$$

Lemma 4.3. For any $\alpha \in \mathbb{R}$ we have

$$\lim_{n\to\infty}\left(\mathbb{P}_{n,p(n,t)}\left(|\mathcal{C}| \in I_n\right) + \mathbb{E}_{n,p(n,t)}\left(\alpha\varphi'(\theta)\frac{|\mathcal{C}| - \theta n}{\sqrt{n}} - \frac{\alpha^2}{2}\frac{\theta}{1-\theta}\,;\, |\mathcal{C}| \in \widetilde{I}_n\right)\right) = 1. \quad (4.16)$$

Before we prove Lemmas 4.2 and 4.3, let us deduce Lemma 4.1 from them.

Proof of Lemma 4.1. First note that it follows from (4.15) and the $\alpha = 0$ case of (4.16) that

$$\lim_{n\to\infty}\mathbb{P}_{n,p(n,t)}\left(|\mathcal{C}| \in \widetilde{I}_n\right) = \theta. \quad (4.17)$$

Denote by μ_n the conditional distribution of $\frac{|\mathcal{C}|-\theta n}{\sqrt{n}}$ given $|\mathcal{C}| \in \widetilde{I}_n$. We have

$$\lim_{n\to\infty}\int\exp\left(\alpha\varphi'(\theta)x\right)\mathrm{d}\mu_n(x) \overset{(4.15),(4.16),(4.17)}{=} \exp\left(\frac{\alpha^2}{2}\frac{\theta}{1-\theta}\right), \qquad \alpha \in \mathbb{R}. \quad (4.18)$$

The r.h.s. of (4.18) is the moment generating function of $\mathcal{N}\left(0, \frac{\theta}{1-\theta}\right)$, thus it classically follows from (4.18) that μ_n weakly converges to $\mathcal{N}\left(0, \sigma^2\right)$ as $n \to \infty$, where σ appears in (1.13). Together with (4.15) and (4.17) this implies Lemma 4.1, given Lemmas 4.2 and 4.3. □

We will prove Lemma 4.2 in Section 4.1 and Lemma 4.3 in Section 4.2. The proofs will make excessive use of (1.5). Let us now introduce some notation that will be used throughout.

For any $\lambda \in (-1, +\infty)$ and any $n \in \mathbb{N}$ let us define

$$\lambda_n^* = \frac{1}{n} \lfloor n\lambda \rfloor. \tag{4.19}$$

Now $\lambda_n^* \in \frac{\mathbb{Z}}{n} \cap (-1, +\infty)$, which is required if we want to use (1.5).

Having fixed $t > 1$, we note that λ_n^* approximates λ well, i.e., we have

$$f_{n,t}(\lambda_n^*, k) \overset{(1.4)}{=} f_{n,t}(\lambda, k) \exp\left(\mathcal{O}\left(\frac{k}{n}\right)\right), \quad e^{-t} - 1 \le \lambda \le 1, \quad 1 \le k \le \frac{e^{-t}}{2} n. \tag{4.20}$$

We will often implicitly use that for any $\lambda > -1$ we have

$$f_{n,t}(\lambda_n^*, k) = 0 \quad \text{if} \quad n + \lfloor n\lambda \rfloor < k \le n \quad \text{and} \quad f_{n,t}(\lambda_n^*, k) \ge 0 \quad \text{if} \quad k \in \{1, \dots, n\}. \tag{4.21}$$

We note that if we let

$$\widetilde{\lambda} := \frac{1}{t} - 1 \quad \text{then we have} \quad \widetilde{x} := \max_{\lambda} e^{-\lambda t}(1 + \lambda) = e^{-\widetilde{\lambda} t}(1 + \widetilde{\lambda}) = \frac{1}{t} e^{t-1} > 1. \tag{4.22}$$

In Sections 4.1 and 4.2 we will dominate $f_{n,t}(\lambda, k)$ by $f_{n,t}(\widetilde{\lambda}_n^*, k)$ for $k \in J_n$ (defined in (4.13)) in order to show that "nothing interesting happens" in the interval J_n.

We will write $f(n) = \Omega(g(n))$ if there exists a constant $c > 0$ (that may depend on t) such that $f(n) \ge c g(n)$ for any $n \in \mathbb{N}$.

4.1 Proof of Lemma 4.2

Before we outline the strategy of the proof of Lemma 4.2 in the paragraph below (4.23), we need to introduce some notation. Let us abbreviate

$$X = f_{n,t}(-\theta, |\mathcal{C}|) \qquad \text{and} \qquad X^* = f_{n,t}((-\theta)_n^*, |\mathcal{C}|).$$

Recalling the definition of the intervals I_n and J_n from (4.13), we have

$$1 + (-\theta)_n^* \overset{(1.5)}{=} \mathbb{E}\left[X^*;\, |\mathcal{C}| \in I_n \right] + \mathbb{E}\left[X^*;\, |\mathcal{C}| \in J_n \right] + \mathbb{E}\left[X^*;\, n^{3/4} \le |\mathcal{C}| \right]. \tag{4.23}$$

We will estimate the three terms on the r.h.s. of (4.23). We will show that the first term approximates $\mathbb{P}(|\mathcal{C}| \in I_n)$ as $n \to \infty$, while the second and third terms vanish as $n \to \infty$.

Before we start estimating the three terms of (4.23), we observe

$$f_{n,t}(-\theta, k) \overset{(1.4),(1.11)}{=} \prod_{i=0}^{k-1} \left(1 - \frac{\theta}{1-\theta} \frac{\frac{i}{n}}{1 - \frac{i}{n}}\right), \qquad k \in [n]. \tag{4.24}$$

Note that $(-\theta)_n^* > e^{-t} - 1$ for large enough n, since $\theta < 1 - e^{-t}$ by (1.10) and (1.11), so we can apply (4.20) in (4.25) and (4.29) below. Now we bound the three terms of (4.23).

First term:

$$\mathbb{E}\left[X^*;\, |\mathcal{C}| \in I_n \right] \overset{(4.13),(4.20)}{=} \mathbb{E}\left[X e^{\mathcal{O}\left(n^{-3/4}\right)};\, |\mathcal{C}| \in I_n \right] \overset{(4.13),(4.24)}{=} \mathbb{P}(|\mathcal{C}| \in I_n) + \mathcal{O}\left(\frac{1}{\sqrt{n}}\right). \tag{4.25}$$

Second term ($\mathbb{E}[\,X^*;\ |\mathcal{C}|\in J_n]$):

$$e^{-(-\theta)^*_n t}\left(1+\frac{(-\theta)^*_n}{1-\frac{i}{n}}\right) \overset{(1.11),(4.19),(4.22)}{\leq} \left(\frac{1+\widetilde{x}}{2}\right)^{-1} e^{-\widetilde{\lambda}^*_n t}\left(1+\frac{\widetilde{\lambda}^*_n}{1-\frac{i}{n}}\right), \qquad 1\leq i\leq n^{3/4}, \tag{4.26}$$

$$\mathbb{E}[\,X^*;\ |\mathcal{C}|\in J_n] \overset{(1.4),(4.13),(4.26)}{\leq} \mathbb{E}\left[\left(\frac{1+\widetilde{x}}{2}\right)^{-|\mathcal{C}|} f_{n,t}(\widetilde{\lambda}^*_n,|\mathcal{C}|);\ |\mathcal{C}|\in J_n\right] \overset{(4.13)}{\leq}$$
$$\left(\frac{1+\widetilde{x}}{2}\right)^{-n^{1/4}} \mathbb{E}\left[f_{n,t}(\widetilde{\lambda}^*_n,|\mathcal{C}|)\right] \overset{(1.5)}{=} \left(\frac{1+\widetilde{x}}{2}\right)^{-n^{1/4}} \left(1+\widetilde{\lambda}^*_n\right) \overset{(4.19),(4.22)}{\leq} e^{-\Omega(n^{1/4})}. \tag{4.27}$$

Third term ($\mathbb{E}\left[\,X^*;\ n^{3/4}\leq|\mathcal{C}|\right]$):

$$e^{-(-\theta)^*_n t}\left(1+\frac{(-\theta)^*_n}{1-\frac{i}{n}}\right) \leq e^{\theta t}(1-\theta) \overset{(1.11)}{=} 1 \quad \text{for any} \quad i\geq \lceil n^{3/4}\rceil, \tag{4.28}$$

$$\mathbb{E}\left[X^*;\ n^{3/4}\leq|\mathcal{C}|\right] \overset{(1.4),(4.28)}{\leq} f_{n,t}((-\theta)^*_n,\lceil n^{3/4}\rceil) \overset{(4.20)}{=} f_{n,t}(-\theta,\lceil n^{3/4}\rceil)e^{\mathcal{O}\left(n^{-1/4}\right)} \overset{(4.24)}{\leq}$$
$$\exp\left(-\frac{\theta}{1-\theta}\sum_{i=0}^{\lceil n^{3/4}\rceil-1}\frac{i}{n}\right)e^{\mathcal{O}\left(n^{-1/4}\right)} \leq e^{-\Omega(\sqrt{n})}. \tag{4.29}$$

The statement of Lemma 4.2 follows from (4.23), (4.25), (4.27) and (4.29).

4.2 Proof of Lemma 4.3

Before we outline the strategy of the proof of Lemma 4.3 in the paragraph below (4.31), we need to introduce some notation. If we define

$$\widetilde{\alpha}_n := \frac{\lfloor\sqrt{n}\alpha\rfloor}{\sqrt{n}} \quad \text{then} \quad \left(\frac{\alpha}{\sqrt{n}}\right)^*_n \overset{(4.19)}{=} \frac{\widetilde{\alpha}_n}{\sqrt{n}} \quad \text{and} \quad |\widetilde{\alpha}_n-\alpha|\leq\frac{1}{\sqrt{n}}. \tag{4.30}$$

Let us abbreviate

$$Y^* = f_{n,t}\left(\frac{\widetilde{\alpha}_n}{\sqrt{n}},|\mathcal{C}|\right).$$

Recall the definitions of the five intervals from (4.13) and (4.14). We have

$$\left(1+\frac{\widetilde{\alpha}_n}{\sqrt{n}}\right)\left(1-\mathbb{P}_{n+\lfloor\sqrt{n}\alpha\rfloor,p}[\,|\mathcal{C}|>n\,]\right) \overset{(1.5)}{=} \mathbb{E}[Y^*;|\mathcal{C}|\in I_n]+\mathbb{E}[Y^*;|\mathcal{C}|\in J_n]+$$
$$\mathbb{E}[Y^*;|\mathcal{C}|\in K_n]+\mathbb{E}[Y^*;|\mathcal{C}|\in\widetilde{I}_n]+\mathbb{E}[Y^*;|\mathcal{C}|\in\widetilde{K}_n]. \tag{4.31}$$

We will estimate the five terms on the r.h.s. of (4.31). We will show that the terms corresponding to I_n and $\widetilde{I}_n$ in (4.31) approximate the terms corresponding to I_n and $\widetilde{I}_n$ in (4.16) as $n\to\infty$, while the terms corresponding to J_n, K_n and $\widetilde{K}_n$ in (4.31) vanish as $n\to\infty$.

Before we start estimating the five terms of (4.31), we note that if $k \in I_n \cup J_n \cup K_n \cup \widetilde{I}_n$ then we can use Taylor expansion of $\ln(1+x)$ to obtain for any $\alpha \in \mathbb{R}$ the formula

$$f_{n,t}\left(\frac{\alpha}{\sqrt{n}}, k\right) \overset{(1.4)}{=} \exp\left(-\frac{\alpha}{\sqrt{n}}kt + \sum_{i=0}^{k-1} \ln\left(1 + \frac{\frac{\alpha}{\sqrt{n}}}{1-\frac{i}{n}}\right)\right) =$$

$$\exp\left(-\frac{\alpha}{\sqrt{n}}kt + \sum_{i=0}^{k-1} \frac{\alpha}{\sqrt{n}}\frac{1}{1-\frac{i}{n}} - \frac{1}{2}\sum_{i=0}^{k-1}\frac{\alpha^2}{n}\frac{1}{(1-\frac{i}{n})^2} + \mathcal{O}\left(\frac{1}{\sqrt{n}}\right)\right) =$$

$$\exp\left(\frac{\alpha}{\sqrt{n}}\left(-kt + n\int_0^{\frac{k}{n}} \frac{1}{1-x}\,\mathrm{d}x\right) - \frac{\alpha^2}{2}\int_0^{\frac{k}{n}}\frac{1}{(1-x)^2}\,\mathrm{d}x + \mathcal{O}\left(\frac{1}{\sqrt{n}}\right)\right)$$

$$\overset{(1.10)}{=} \exp\left(\alpha\sqrt{n}\varphi\left(\frac{k}{n}\right) - \frac{\alpha^2}{2}\frac{\frac{k}{n}}{1-\frac{k}{n}} + \mathcal{O}\left(\frac{1}{\sqrt{n}}\right)\right). \quad (4.32)$$

Now we can estimate the five terms on the r.h.s. of (4.31).

First term:

$$\mathbb{E}[Y^*; |\mathcal{C}| \in I_n] \overset{(1.10),(4.13),(4.32)}{=} \mathbb{P}\left(|\mathcal{C}| \in I_n\right) + \mathcal{O}\left(n^{-1/4}\right). \quad (4.33)$$

Second term: The bound

$$\mathbb{E}[Y^*; |\mathcal{C}| \in J_n] \le e^{-\Omega(n^{1/4})} \quad (4.34)$$

can be deduced analogously to (4.27) using that for large enough n we have

$$e^{-\frac{\widetilde{\alpha}_n}{\sqrt{n}}t}\left(1 + \frac{\frac{\widetilde{\alpha}_n}{\sqrt{n}}}{1-\frac{i}{n}}\right) \overset{(4.19),(4.22)}{\le} \left(\frac{1+\widetilde{x}}{2}\right)^{-1} e^{-\widetilde{\lambda}_n^* t}\left(1 + \frac{\widetilde{\lambda}_n^*}{1-\frac{i}{n}}\right), \qquad 1 \le i \le n^{3/4}. \quad (4.35)$$

Third term ($\mathbb{E}[Y^*; |\mathcal{C}| \in K_n]$): We note

$$\frac{f_{n,t}\left(\frac{\widetilde{\alpha}_n}{\sqrt{n}}, k\right)}{f_{n,t}\left(\frac{\widetilde{(\alpha-1)}_n}{\sqrt{n}}, k\right)} \overset{(4.32)}{=} \exp\left(\sqrt{n}\varphi\left(\frac{k}{n}\right) + \mathcal{O}(1)\right) \overset{(1.10),(1.11),(4.13)}{\le} e^{-\Omega(n^{1/8})}, \quad k \in K_n, \quad (4.36)$$

$$\mathbb{E}[Y^*; |\mathcal{C}| \in K_n] \overset{(4.36)}{\le} e^{-\Omega(n^{1/8})}\mathbb{E}\left[f_{n,t}\left(\frac{\widetilde{(\alpha-1)}_n}{\sqrt{n}}, |\mathcal{C}|\right); |\mathcal{C}| \in K_n\right] \overset{(1.5)}{\le} 2e^{-\Omega(n^{1/8})}. \quad (4.37)$$

Fourth term ($\mathbb{E}[Y^*; |\mathcal{C}| \in \widetilde{I}_n]$): If $x \in [-n^{1/8}, n^{1/8}]$, i.e., if $k = \lfloor \theta n + x\sqrt{n}\rfloor \in \widetilde{I}_n$ then

$$f_{n,t}\left(\frac{\alpha}{\sqrt{n}}, k\right) \overset{(1.11),(4.32)}{=} \exp\left(\alpha\varphi'(\theta)x - \frac{\alpha^2}{2}\frac{\theta}{1-\theta} + \mathcal{O}\left(n^{-1/4}\right)\right), \quad (4.38)$$

$$f_{n,t}\left(\frac{\widetilde{\alpha}_n}{\sqrt{n}}, k\right) \overset{(4.30),(4.38)}{=} \exp\left(\alpha\varphi'(\theta)x - \frac{\alpha^2}{2}\frac{\theta}{1-\theta}\right) + \mathcal{O}\left(n^{-1/4} f_{n,t}\left(\frac{\widetilde{\alpha}_n}{\sqrt{n}}, k\right)\right), \quad (4.39)$$

$$\mathbb{E}[Y^*; |\mathcal{C}| \in \widetilde{I}_n] \overset{(1.5),(4.14),(4.39)}{=} \mathbb{E}\left[\exp\left(\alpha\varphi'(\theta)\frac{|\mathcal{C}|-\theta n}{\sqrt{n}} - \frac{\alpha^2}{2}\frac{\theta}{1-\theta}\right);\ |\mathcal{C}| \in \widetilde{I}_n\right] + \mathcal{O}\left(n^{-1/4}\right). \quad (4.40)$$

Fifth term ($\mathbb{E}[Y^*; |\mathcal{C}| \in \widetilde{K}_n]$): We observe that

$$\frac{f_{n,t}\left(\frac{\widetilde{\alpha}_n}{\sqrt{n}}, \lfloor \theta n + n^{1/8}\sqrt{n}\rfloor\right)}{f_{n,t}\left(\frac{\widetilde{(\alpha+1)}_n}{\sqrt{n}}, \lfloor \theta n + n^{1/8}\sqrt{n}\rfloor\right)} \overset{(4.38)}{=} \exp\left(-\varphi'(\theta)n^{1/8} + \mathcal{O}(1)\right) \overset{(1.12)}{\le} e^{-\Omega(n^{1/8})}, \quad (4.41)$$

$$\exp\left(-\frac{\widetilde{\alpha}_n}{\sqrt{n}}t\right)\left(1+\frac{\frac{\widetilde{\alpha}_n}{\sqrt{n}}}{1-\frac{i}{n}}\right) \overset{(1.12)}{\leq} \exp\left(-\frac{\widetilde{(\alpha+1)}_n}{\sqrt{n}}t\right)\left(1+\frac{\frac{\widetilde{(\alpha+1)}_n}{\sqrt{n}}}{1-\frac{i}{n}}\right), \quad \lfloor \theta n + n^{1/8}\sqrt{n}\rfloor \leq i, \tag{4.42}$$

$$f_{n,t}\left(\frac{\widetilde{\alpha}_n}{\sqrt{n}}, k\right) \overset{(1.4),(4.41),(4.42)}{\leq} e^{-\Omega(n^{1/8})} f_{n,t}\left(\frac{\widetilde{(\alpha+1)}_n}{\sqrt{n}}, k\right), \quad \lfloor \theta n + n^{1/8}\sqrt{n}\rfloor \leq k, \tag{4.43}$$

$$\mathbb{E}[Y^*; |\mathcal{C}| \in \widetilde{K}_n] \overset{(4.14),(4.43)}{\leq} e^{-\Omega(n^{1/8})}\mathbb{E}\left[f_{n,t}\left(\frac{\widetilde{(\alpha+1)}_n}{\sqrt{n}}, |\mathcal{C}|\right); |\mathcal{C}| \in \widetilde{K}_n\right] \overset{(1.5)}{\leq} 2e^{-\Omega(n^{1/8})}. \tag{4.44}$$

Finally, the proof of the fact that the error term $\mathbb{P}_{n+\lfloor\sqrt{n}\alpha\rfloor,p}[\,|\mathcal{C}| > n\,]$ that appears on the l.h.s. of (4.31) goes to zero as $n \to \infty$ is analogous to the $\alpha = 0$ case of (4.44). The statement of Lemma 4.3 follows from (4.31), (4.33), (4.34), (4.37), (4.40) and (4.44).

Acknowledgements: The author thanks James Martin, Dominic Yeo and Bálint Tóth for valuable discussions. This work is partially supported by OTKA (Hungarian National Research Fund) grant K109684, the Postdoctoral Fellowship of NKFI (National Research, Development and Innovation Office) and the Bolyai Research Scholarship of the Hungarian Academy of Sciences.

References

[1] D. J. Aldous. Brownian excursions, critical random graphs and the multiplicative coalescent. *The Annals of Probability* 812-854, 1997.

[2] D. J. Aldous. Deterministic and stochastic models for coalescence (aggregation and coagulation): a review of the mean-field theory for probabilists. *Bernoulli* 5.1: 3-48, 1999.

[3] N. Alon, J. H. Spencer. The probabilistic method *(Second edition)*. Wiley, 2000.

[4] D. Barraez, S. Boucheron, W.F. De La Vega. On the fluctuations of the giant component. *Combinatorics, Probability and Computing* 9.4, 28-304, 2000.

[5] B. Bollobás. Random graphs. No. 73. *Cambridge university press*, 2001.

[6] B. Bollobás, S. Janson, O. Riordan. The phase transition in inhomogeneous random graphs. *Random Structures and Algorithms* 31.1: 3-122, 2007.

[7] P. Erdős, and A. Rényi. On the evolution of random graphs. *Publ. Math. Inst. Hung. Acad. Sci* 5.1: 17-60, 1960.

[8] L. Federico, R. van der Hofstad, F. den Hollander, T. Hulshof. Expansion of percolation critical points for Hamming graphs. *arXiv:1701.02099*, 2017.

[9] R. van der Hofstad. Random graphs and complex networks. *Cambridge Series in Statistical and probabilistic Mathematics* 43, 2016.

[10] S. Janson, M. J. Luczak. Susceptibility in subcritical random graphs. *Journal of Mathematical Physics* 49.12, 125207, 2008.

[11] S. Janson, T. Luczak, A. Rucinski. Random graphs. Vol. 45. John Wiley and Sons, 2011.

[12] S. Janson, L. Warnke. On the critical probability in percolation. *arXiv:1611.08549*, 2016.

[13] J. B. Martin, B. Ráth. Rigid representations of the multiplicative coalescent with linear deletion. *(submitted), arXiv:1610.00891*, 2016.