On the critical probability in percolation

Svante Janson* and Lutz Warnke†

November 25, 2016

Abstract

For percolation on finite transitive graphs, Nachmias and Peres suggested a characterization of the critical probability based on the logarithmic derivative of the susceptibility. As a first test-case, we study their suggestion for the Erdős–Rényi random graph $G_{n,p}$, and confirm that the logarithmic derivative has the desired properties: (i) its maximizer lies inside the critical window $p = 1/n + \Theta(n^{-4/3})$, and (ii) the inverse of its maximum value coincides with the $\Theta(n^{-4/3})$–width of the critical window. We also prove that the maximizer is not located at $p = 1/n$ or $p = 1/(n-1)$, refuting a speculation of Peres.

1 Introduction

The percolation phase transition on finite graphs is one of the most intriguing and striking phenomena at the intersection of mathematical physics, combinatorics, and probability theory. The classical Erdős–Rényi random graph $G_{n,p}$ is perhaps the most carefully studied reference model: as the edge probability p increases past the 'critical probability' $p_c = 1/n$, the global structure changes radically, from only small components to a single giant component plus small ones. More precisely, using the parametrization $p = 1/n + \lambda_n n^{-4/3}$, and for simplicity assuming $p = \Theta(1/n)$, by the inspiring work of Erdős and Rényi [11], Bollobás [4], Łuczak [27], and Aldous [3], we nowadays distinguish three qualitatively different phases of $G_{n,p}$. In the subcritical phase $\lambda_n \to -\infty$, the $r = \Theta(1)$ largest components $C_1, \ldots, C_r$ are typically all of comparable size: $|C_1| \sim |C_2| \sim \cdots \sim |C_r| = \Theta(n^{2/3}\lambda_n^{-2} \log|\lambda_n|) = o(n^{2/3})$. In the supercritical phase $\lambda_n \to \infty$, the largest component typically dominates all other components: $|C_2| \ll |C_1| = \Theta(\lambda_n n^{2/3})$. In the critical window $|\lambda_n| = O(1)$, the rescaled sizes $|C_1|/n^{2/3}, |C_2|/n^{2/3}, \cdots$ of the largest components converge in distribution to non-degenerate random variables, i.e., they are not concentrated.

In the language of mathematical physics, $G_{n,p}$ interpreted as percolation on the complete n-vertex graph is a mean-field model. Hence, we expect that the percolation phase transition of many 'high dimensional' finite graphs is similar, with the hypercube and various tori being examples of great interest (see, e.g., [2, 5, 18, 7, 8, 9, 16]). To fix notation, we assume that G is a given transitive n-vertex graph, and we write $G_p \subseteq G$ for the binomial random subgraph where each edge is included independently with probability p. As pointed out by Nachmias and Peres [29], in this general percolation setting it is a challenging problem to find a good definition of the critical probability p_c, such that for a suitable critical window around p_c, for example, the size of the largest component is not concentrated.

The folklore average degree heuristic $p_c = 1/(\deg_G(v) - 1)$ is a natural first guess (the graph G is assumed to be transitive and thus regular, so the choice of the vertex v does not matter). For the hypercube with vertex set $\{0,1\}^m$, and thus degree m, Ajtai, Komlós and Szemerédi [2] showed that there is a critical threshold $(1+o(1))/m$; this was sharpened by Bollobás, Kohayakawa, and Łuczak [5], who raised the question whether

*Department of Mathematics, Uppsala University, PO Box 480, SE-751 06 Uppsala, Sweden. E-mail: svante.janson@math.uu.se. Part of the work was done during visits to the University of Cambridge and to the Isaac Newton Institute for Mathematical Sciences during the programme Theoretical Foundations for Statistical Network Analysis (EPSCR Grant Number EP/K032208/1) and was partially supported by a grant from the Knut and Alice Wallenberg Foundation and a grant from the Simons foundation.

†Department of Pure Mathematics and Mathematical Statistics, University of Cambridge, Wilberforce Road, Cambridge CB3 0WB, UK; *and* School of Mathematics, Georgia Institute of Technology, Atlanta GA 30332, USA. E-mail: L.Warnke@dpmms.cam.ac.uk. Part of the work was done while interning with the Theory Group at Microsoft Research, Redmond; LW thanks Yuval Peres for introducing him to the topic of this paper.

the critical probability might be exactly $1/(m-1)$. However, Borgs, Chayes, van der Hofstad, Slade and Spencer [7, 8] and van der Hofstad and Nachmias [16, 17] have shown that there is critical window of width $\Theta(n^{-1/3}p_c) = \Theta(2^{-m/3}/m)$ about a critical probability p_c, which by van der Hofstad and Slade [18] satisfies $p_c = 1/(m-1) + 3.5m^{-3} + O(m^{-4})$; since the width of the window is $o(m^{-3})$, the value $1/(m-1)$ is outside the critical window.

A more sophisticated suggestion for the critical probability was pioneered by Borgs, Chayes, van der Hofstad, Slade and Spencer [7, 8] (and used for the hypercube result just described). They essentially proposed to define $p_c = p_c(G)$ as the unique solution to the polynomial equation

$$\chi_G(p) := \mathbb{E}_p|C(v)| = n^{1/3}, \tag{1.1}$$

where the susceptibility $\chi_G(p)$ denotes the expected size of the component $C(v)$ containing a fixed vertex v in G_p. (This is a widely studied key parameter in percolation theory and random graph theory, see, e.g., [1, 13, 22, 25, 32, 33]. Since G is assumed to be transitive, the choice of v does not matter.) The aforementioned technical definition is guided by Erdős–Rényi mean-field type behaviour. Indeed, in the subcritical phase we expect that $C(v)$ closely mimics a subcritical branching process, which suggests that typically $|C_1| \approx (\chi_G(p))^2$ up to logarithmic corrections (see, e.g., Section 1.2 in [15] or Proposition 5.1 in [1]). Furthermore, in the supercritical phase we expect that the largest component dominates all other components, which by transitivity suggests that $\chi_G(p) \approx \mathbb{E}_p|C_1|^2/n$. Assuming that inside the critical window we can observe subcritical and supercritical features, it thus seems plausible that the critical probability should roughly satisfy $\chi_G(p) \approx \mathbb{E}_p|C_1|^2/n \approx \chi_G(p)^4/n$, motivating the choice of equation (1.1). Borgs et al. [7, 8] showed that (a minor variant of) the discussed definition is very useful in combination with the so-called finite triangle condition: they recovered many Erdős–Rényi features under such generic mean-field assumptions (see [16, 17] for some more recent developments).

As pointed out by Peres [31], the suggestion of Borgs et al. [7, 8] builds the mean-field scaling $\Theta(n^{1/3})$ into the definition of the critical probability. It would be desirable to have a useful general definition that recovers this scaling for n-vertex mean-field graphs $G = G_n$, rather than having separate definitions for each different scaling behaviour (or, in mathematical physics jargon, for each 'universality class'). With this aim in mind, Nachmias and Peres [29] suggested to define $p_c = p_c(G)$ as the value of p which maximizes the logarithmic derivative

$$\frac{\mathrm{d}}{\mathrm{d}p} \log \chi_G(p) = \frac{\frac{\mathrm{d}}{\mathrm{d}p}\mathbb{E}_p|C(v)|}{\mathbb{E}_p|C(v)|}. \tag{1.2}$$

To motivate this definition, note that by the Margulis–Russo formula [28, 34] the derivative $\frac{\mathrm{d}}{\mathrm{d}p}\mathbb{E}_p|C(v)|$ intuitively counts the expected (weighted) number of edges of G_p which can affect the size of $|C(v)|$, see also Section 2. In other words, p_c equals the probability where the addition of a random edge has maximum relative impact on the component size $|C(v)|$. Denoting the maximum value of (1.2) by $M = M(G)$, Warnke [36] conjectured that for mean-field graphs G the width of the critical window is of order $\Theta(1/M)$. This is motivated by the fact that $\log(\chi_G(p_2)/\chi_G(p_1)) = \int_{p_1}^{p_2} \frac{\mathrm{d}}{\mathrm{d}p} \log \chi_G(p) dp \leqslant (p_2 - p_1)M$ entails that the susceptibility satisfies $\chi_G(p_2) = \Theta(\chi_G(p_1))$ for $p_2 - p_1 = O(1/M)$.

1.1 Main results

In this paper we investigate, as a first test-case, the suggested definition of Nachmias and Peres [29] for the Erdős–Rényi random graph $G_{n,p}$ i.e., the case $G = K_n$ (as proposed by Peres [31]). Here our first main result confirms that their definition of the critical probability p_c has the desired properties, i.e., that for $G_{n,p} = (K_n)_p$ the logarithmic derivative $\frac{\mathrm{d}}{\mathrm{d}p} \log \chi_{K_n}(p)$ satisfies the following:

(i) its maximizer lies inside the critical window $p = 1/n + O(n^{-4/3})$, and

(ii) the inverse of its maximum value coincides with the $\Theta(n^{-4/3})$–width of the critical window.

Theorem 1.1 (Maximizer of the logarithmic derivative for $G_{n,p}$)**.** *We have*

$$\left|\operatorname*{argmax}_{p\in(0,1)} \frac{\mathrm{d}}{\mathrm{d}p} \log \chi_{K_n}(p) - \frac{1}{n}\right| = O(n^{-4/3}), \tag{1.3}$$

$$\max_{p\in(0,1)} \frac{\mathrm{d}}{\mathrm{d}p} \log \chi_{K_n}(p) = \Theta(n^{4/3}). \tag{1.4}$$

Remark 1.2. *Theorem 1.3 shows that* (1.3) *remains valid with* $O(n^{-4/3})$ *replaced by* $\Theta(n^{-4/3})$*.*

Having established the qualitative behaviour of the logarithmic derivative for $G_{n,p}$, it is intriguing to investigate the finer scaling behaviour inside critical window. By symmetry considerations it might be tempting to believe that $p = 1/n$ or $p = 1/(n-1)$ could be the maximizer of $\frac{\mathrm{d}}{\mathrm{d}p} \log \chi_{K_n}(p)$, as speculated by Peres [31]. Our second main result refutes this tantalizing belief, instead strengthening the general feeling that $\lambda = 0$ is no special point inside the critical window of form $p = 1/n + \lambda n^{-4/3}$.

Theorem 1.3 (Scaling inside the critical window of $G_{n,p}$)**.** *Given* $\lambda \in \mathbb{R}$*, for* $p = 1/n + \big(\lambda + o(1)\big)n^{-4/3}$ *we have, as* $n \to \infty$*,*

$$\frac{\chi_{K_n}(p)}{n^{1/3}} \to f(\lambda), \tag{1.5}$$

$$\frac{\frac{\mathrm{d}}{\mathrm{d}p} \log \chi_{K_n}(p)}{n^{4/3}} \to \frac{\mathrm{d}}{\mathrm{d}\lambda} \log f(\lambda) > 0, \tag{1.6}$$

where the infinitely differentiable function $f = f_2 : \mathbb{R} \to (0, \infty)$ *is defined in* (3.1)–(3.4)*. Moreover, if* $p = 1/n + \lambda n^{-4/3}$*, then the convergence in* (1.5)–(1.6) *is uniform for* λ *in any compact interval* $[\lambda_1, \lambda_2] \subset \mathbb{R}$*. Furthermore,*

$$\frac{\mathrm{d}^2}{\mathrm{d}\lambda^2} \log f(0) \neq 0. \tag{1.7}$$

The definition of the function f appearing in Theorem 1.3 is quite involved, since it intuitively needs to capture the contribution of components with arbitrary numbers of cycles. It is easy to find asymptotics as $\lambda \to \pm\infty$; we have $f(\lambda) \sim |\lambda|^{-1}$ as $\lambda \to -\infty$ and $f(\lambda) \sim 4\lambda^2$ as $\lambda \to +\infty$; hence $\log f(\lambda) = -\log|\lambda| + o(1)$ as $\lambda \to -\infty$ and $\log f(\lambda) = 2\log\lambda + O(1)$ as $\lambda \to +\infty$; furthermore, $\frac{\mathrm{d}}{\mathrm{d}\lambda} \log f(\lambda) = O(1/|\lambda|)$ for all $\lambda \in \mathbb{R}$, see Appendix A for proofs. Theorem 1.3 also extends to convergence of higher derivatives, see Appendix C.

It would be interesting to know whether the logarithmic derivative $\frac{\mathrm{d}}{\mathrm{d}\lambda} \log f(\lambda)$ has a unique maximizer λ^*, and whether it is unimodal. Figure 1 below (which is obtained by numerical integrations) suggests that this is the case, with $\lambda^* \approx 1$ (we conjecture $\lambda^* > 1$ based on our limited precision numerical data).

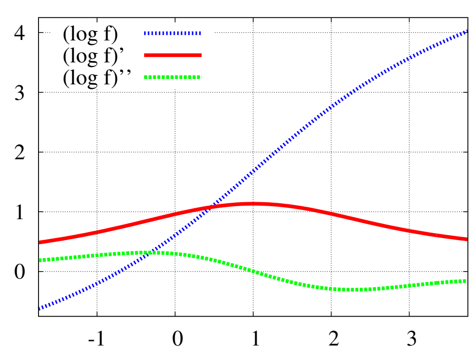

Figure 1: Plot of the functions $\log f(\lambda)$, $\frac{\mathrm{d}}{\mathrm{d}\lambda} \log f(\lambda)$ and $\frac{\mathrm{d}^2}{\mathrm{d}\lambda^2} \log f(\lambda)$ for $\lambda \in [-1.75, 3.75]$, where f is as in Theorem 1.3. It provides some evidence for our belief that $\frac{\mathrm{d}}{\mathrm{d}\lambda} \log f(\lambda)$ has a unique maximizer $\lambda^* \approx 1$.

The high-level structure of our proofs is as follows. For Theorem 1.1 our starting point is the Margulis–Russo Formula, which allows us to write $\frac{\mathrm{d}}{\mathrm{d}p}\chi_{K_n}(p)$ in terms of sums involving the squared component sizes of $G_{n,p}$. Using ideas from random graph theory we then estimate these sums, combining correlation inequalities and the ‘symmetry rule’ (also called ‘discrete duality principle’) with results for the largest component and the susceptibility of $G_{n,p}$, which eventually implies (1.3)–(1.4); see Section 2. For Theorem 1.3 with $p = 1/n + \big(\lambda + o(1)\big)n^{-4/3}$, our starting point is the well-known fact that $X_{n,2} = \sum_{j \geqslant 1} |C_j|^2/n^{4/3} \xrightarrow{\text{d}} W_{\lambda,2}$ for some random variable $W_{\lambda,2}$. Using technical arguments we then justify taking expectations and derivatives, which in view of $\chi_{K_n}(p)/n^{1/3} = \mathbb{E}_p X_{n,2}$ eventually establishes (1.5)–(1.6) with $f(\lambda) = \mathbb{E}W_{\lambda,2}$; see Section 3.1. For inequality (1.7) we show that f and its derivatives can be computed at $\lambda = 0$ by series expansions (exploiting recursive formulas for the area under a normalized Brownian excursion). Since these series converge exponentially, we can then numerically verify (1.7) by finite truncation; see Section 3.2.

1.2 Remarks on some other graphs

In the present paper we discuss only the Erdős–Rényi random graph $G_{n,p} = (K_n)_p$, i.e., percolation on the complete n-vertex graph. In particular, Theorem 1.1 shows that the definition of the critical probability p_c suggested by Nachmias and Peres [29, 31] ‘works’ in this case. It is an interesting open problem to establish analogous results for other finite graphs.

For example, consider again the hypercube with vertex set $\{0,1\}^m$ discussed above, see [7, 8, 16, 17]. In the subcritical phase $p = (1-\varepsilon)p_c$ with $\varepsilon^3 n \to \infty$, [7, Proposition A.1] combined with [8, Theorem 1.3 and Theorem 1.5] show that $\frac{\text{d}}{\text{d}p}\chi_G(p) \sim m(\chi_G(p))^2$ and $\chi_G(p) \sim \varepsilon^{-1}$, and thus

$$\frac{\text{d}}{\text{d}p}\log\chi_G(p) \sim m\chi_G(p) = \Theta(\varepsilon^{-1}m) = o\big(n^{1/3}m\big).$$

In the supercritical phase $p = (1+\varepsilon)p_c$ with $\varepsilon^3 n \to \infty$ and $\varepsilon = \varepsilon(n) \to 0$, we have $\chi_G(p) \sim 4\varepsilon^2 n$ according to [16, Theorem 1.1]; hence it is natural to conjecture that the logarithmic derivative satisfies

$$\frac{\text{d}}{\text{d}p}\log\chi_G(p) = p_c^{-1}\frac{\text{d}}{\text{d}\varepsilon}\log\chi_G(p) \approx \frac{2}{\varepsilon p_c} = \Theta(\varepsilon^{-1}m) = o\big(n^{1/3}m\big),$$

in the supercritical phase too, and, moreover, that the logarithmic derivative has a maximum of order $\Theta(n^{1/3}m)$ which is attained inside the critical window. Proving this, however, remains a challenging problem.

Another important example would be random d-regular graphs with $d = d(n) \to \infty$.

Moreover, it would be conceptually very interesting to start with the maximizer of (1.2) and then derive properties of the phase transition of G_p (rather than, as in this paper, using known results for G_p to verify properties of the maximizer).

It also seems highly desirable to better understand the critical probability p_c for finite transitive graphs G which do not exhibit the mean-field behavior of the complete graph K_n or the hypercube $\{0,1\}^m$. Here the perhaps simplest example is percolation on the n-vertex cycle, $n \geqslant 3$, for which it is not difficult to check that there are three different phases: (i) for $1-p = o(n^{-1})$ we typically have $|C_1| = n$, (ii) for $1-p = \omega(n^{-1})$ we typically have $|C_1| = o(n)$, and (iii) for $1-p = \Theta(n^{-1})$ the rescaled sizes $|C_1|/n, \cdots, |C_r|/n$ of the $r = \Theta(1)$ largest components are not concentrated. Hence the critical window is parametrized by $p = 1 - \lambda_n n^{-1}$ with $\lambda_n = \Theta(1)$. For $p \in (0,1)$ it is routine to see that

$$\mathbb{E}_p|C(v)| = 1 + \sum_{1\leqslant j<n}(p^j + p^{n-j} - p^n) = 1 + \sum_{1\leqslant j<n}(2p^j - p^n),$$

$$\frac{\text{d}}{\text{d}p}\mathbb{E}_p|C(v)| = \sum_{1\leqslant j<n}(2jp^{j-1} - np^{n-1}) = \sum_{1\leqslant j<n}2jp^{j-1}(1-p^{n-j}).$$

A short calculation shows that $\mathbb{E}_p|C(v)| = \Theta(n^{1/3})$ implies $p = 1-\Theta(n^{-1/3})$. Furthermore, $\frac{\text{d}}{\text{d}p}\log\mathbb{E}_p|C(v)| = \Theta(n)$ for $1-p = \Theta(n^{-1})$, and $\frac{\text{d}}{\text{d}p}\log\mathbb{E}_p|C(v)| = \Theta(\min\{(1-p)n^2, (1-p)^{-1}\}) = o(n)$ otherwise. For the critical probability p_c of n-vertex cycles, it follows that the mean-field definition of Borgs et al. fails (as expected, since cycles are not ‘high dimensional’). By contrast, the definition based on the maximizer of the logarithmic derivative of the susceptibility does correctly predict $p_c = 1-\Theta(n^{-1})$ and the $\Theta(n^{-1})$–width of the critical window, supporting the hope that this definition might work beyond the mean-field case.

1.3 Some notation

For emphasis, we will often use $\mathbb{P}_{n,p}$ and $\mathbb{E}_{n,p}$ for probability and expectation with respect to $G_{n,p}$. We let C_i denote the components of $G_{n,p}$ in order of decreasing sizes, $|C_1| \geqslant |C_2| \geqslant \cdots$ (resolving ties by taking the component with the smallest vertex label first, for definiteness). Finally, convergence in distribution is denoted $\xrightarrow{\mathrm{d}}$, and unspecified limits are as $n \to \infty$.

2 Maximizer of the logarithmic derivative

In this section we prove Theorem 1.1. Our arguments combine the Margulis–Russo formula with results and ideas from random graph theory. For mathematical convenience we shall work with the 'rescaled' susceptibility parameters

$$S(G_{n,p}) := \sum_{v\in[n]} |C(v)| = \sum_i |C_i|^2, \tag{2.1}$$

where the component $C(v)$ is with respect to $G_{n,p}$, as usual, and

$$S_n(p) := \mathbb{E}S(G_{n,p}) = \mathbb{E}_{n,p}\Big(\sum_i |C_i|^2\Big). \tag{2.2}$$

Recall that $\chi_{K_n}(p) := \mathbb{E}_{n,p}|C(v)|$, which is the same for every $v \in [n]$ by symmetry, and thus by (2.1)–(2.2)

$$S_n(p) = n\chi_{K_n}(p), \tag{2.3}$$

which implies

$$\frac{\mathrm{d}}{\mathrm{d}p}\log\chi_{K_n}(p) = \frac{\mathrm{d}}{\mathrm{d}p}\log S_n(p). \tag{2.4}$$

Theorem 1.1 follows from equation (2.4) and inequalities (2.5)–(2.6) of Theorem 2.1 below. (In fact, in the lower bound (2.6), it suffices to consider, for example, $\lambda = 0$.)

Theorem 2.1 (Bounds for the logarithmic derivative)**.** *There is a constant $C > 0$ such that, for all $n \geqslant 1$ and $p \in (0,1)$,*

$$\frac{\mathrm{d}}{\mathrm{d}p}\log S_n(p) \leqslant C \cdot \min\big\{|p - 1/n|^{-1},\ n^{4/3}\big\}. \tag{2.5}$$

Furthermore, for every $\lambda \in \mathbb{R}$ there is a constant $D_\lambda > 0$ such that, for all $n \geqslant 2$ and $p = 1/n + \lambda n^{-4/3} \in (0,1)$,

$$\frac{\mathrm{d}}{\mathrm{d}p}\log S_n(p) \geqslant D_\lambda n^{4/3}. \tag{2.6}$$

The remainder of this section is devoted to the proof of Theorem 2.1, and we start by studying a combinatorial form of $\frac{\mathrm{d}}{\mathrm{d}p}S_n(p)$. Writing $v \leftrightarrow w$ for the event that v and w are connected (which trivially holds if $v = w$), note that $S(G) = \sum_{v,w\in V(G)} \mathbb{1}_{\{v\leftrightarrow w\}}$ and thus, by taking the expectation, see (2.2),

$$S_n(p) = \sum_{v,w\in[n]} \mathbb{P}_{n,p}(v \leftrightarrow w). \tag{2.7}$$

We now record the following simple monotonicity property, which is obvious from (2.7).

Lemma 2.2. *If $p \leqslant p'$ and $n \leqslant n'$, then $S_n(p) \leqslant S_{n'}(p')$.* □

We say that an edge $e \in E(K_n)$ is *pivotal* for $v \leftrightarrow w$, if $v \leftrightarrow w$ in $G_{n,p} + e$ and $v \not\leftrightarrow w$ in $G_{n,p} - e$ (i.e., in the possibly modified graphs where e is added and removed, respectively). Recalling the form of (2.7), for $p \in (0,1)$ the Margulis–Russo Formula [28, 34] gives

$$\frac{\mathrm{d}}{\mathrm{d}p}S_n(p) = \sum_{v,w\in[n]} \frac{\mathrm{d}}{\mathrm{d}p}\mathbb{P}_{n,p}(v \leftrightarrow w) = \sum_{v,w\in[n]}\sum_{e\in E(K_n)} \mathbb{P}_{n,p}(e \text{ is pivotal for } v \leftrightarrow w). \tag{2.8}$$

Let $\mathcal{P}_{e,v,w}$ denote the event that (i) $e \notin G_{n,p}$ and (ii) e is pivotal for $v \leftrightarrow w$. Since being pivotal does not depend on the status of e, it follows that

$$\frac{\mathrm{d}}{\mathrm{d}p} S_n(p) = \frac{\mathbb{E}_{n,p}\left(\sum_{v,w\in[n]} \sum_{e\in E(K_n)} \mathbb{1}_{\{\mathcal{P}_{e,v,w}\}}\right)}{1-p}. \tag{2.9}$$

An edge not present in $G_{n,p}$ is pivotal for $v \leftrightarrow w$ if and only if one of its endpoints is in $C(v)$ and the other is in $C(w) \neq C(v)$. Hence $\sum_{e\in E(K_n)} \mathbb{1}_{\{\mathcal{P}_{e,v,w}\}} = \mathbb{1}_{\{C(v)\neq C(w)\}}|C(v)||C(w)|$. Consequently,

$$\sum_{v,w\in[n]} \sum_{e\in E(K_n)} \mathbb{1}_{\{\mathcal{P}_{e,v,w}\}} = \sum_{v\in[n]} |C(v)| \sum_{w\notin C(v)} |C(w)| = \sum_{i\neq j} |C_i|^2|C_j|^2, \tag{2.10}$$

and thus, by (2.9),

$$\frac{\mathrm{d}}{\mathrm{d}p} S_n(p) = \frac{\mathbb{E}_{n,p}\left(\sum_{v\in[n]} |C(v)| \sum_{w\notin C(v)} |C(w)|\right)}{1-p} = \frac{\mathbb{E}_{n,p}\left(\sum_{i\neq j} |C_i|^2|C_j|^2\right)}{1-p}, \tag{2.11}$$

which eventually allows us to bring random graph theory into play.

2.1 Upper bounds

In this subsection we prove the upper bound (2.5) from Theorem 2.1.

We shall use some more or less well-known results for the susceptibility and the size of the largest component of $G_{n,p}$ in near-critical cases, which we state as the following theorem. (See, e.g., [7, 22, 25, 6] for similar or related results.)

Theorem 2.3. (i) *There is a constant $D > 0$ such that, for all $n \geqslant 1$, $p \in [0,1]$, and $\varepsilon > 0$,*

$$S_n(p) \leqslant \varepsilon^{-1} n \qquad \text{if } np \leqslant 1-\varepsilon, \tag{2.12}$$

$$S_n(p) \leqslant D\varepsilon^2 n^2 \qquad \text{if } np \leqslant 1+\varepsilon \text{ and } \varepsilon^3 n \geqslant 1. \tag{2.13}$$

(ii) *For any $A > 0$ there are constants $a, B, n_0 > 0$ such that, for all $n \geqslant n_0$, $p \in [0,1]$, $\varepsilon \in (0,A]$, and $\delta \in (0,1/2]$ satisfying $np = 1+\varepsilon$ and $\delta^2\varepsilon^3 n \geqslant B$,*

$$\mathbb{P}_{n,p}\big(\big||C_1| - \rho(\varepsilon)n\big| \geqslant \delta\rho(\varepsilon)n\big) \leqslant e^{-a\delta^2\varepsilon^3 n}, \tag{2.14}$$

where $\rho(\varepsilon) > 0$ is the positive solution to $1-\rho(\varepsilon) = e^{-(1+\varepsilon)\rho(\varepsilon)}$.

(iii) *Furthermore, for any $A > 0$ there are constants $\delta \in (0,1/2)$ and $c > 0$ such that, for all $\varepsilon \in (0,A]$,*

$$0 < \big(1-(1-\delta)\rho(\varepsilon)\big)\cdot(1+\varepsilon) \leqslant 1 - c\varepsilon. \tag{2.15}$$

Proof. The subcritical upper bound (2.12) for the susceptibility is simple and well-known. The supercritical upper bound (2.13) is intuitively clear, since in the supercritical range, the susceptibility ought to be dominated by $\mathbb{E}_{n,p}|C_1|^2$ and $|C_1|$ is with high probability $\Theta(n\varepsilon)$ when $np = 1+\varepsilon$. However, we are unaware of a reference which contains a short proof of (2.13), and thus for completeness we give in Appendix B proofs of both upper bounds (2.12)–(2.13) for the susceptibility.

The tail bound (2.14) follows from [6, Theorem 4, (10) and Remark 3].

The estimate (2.15) follows for small ε, say $\varepsilon \leqslant \varepsilon_0$, from the fact that $\rho(\varepsilon) = 2\varepsilon + o(\varepsilon)$ as $\varepsilon \to 0$, and for $\varepsilon \in [\varepsilon_0, A]$ from the fact that (with $\rho = \rho(\varepsilon)$) $(1-\rho)(1+\varepsilon) = -(1-\rho)\log(1-\rho)/\rho < 1$ for $\varepsilon > 0$ together with the continuity of $\rho(\varepsilon)$. (Cf. e.g. [22, Lemma A.2].) ☐

Corollary 2.4. *There are constants $n_0, \pi_0, b > 0$ such that, for all $n \geqslant n_0$ and $p \in [0,1]$ satisfying $np \geqslant \pi_0$, we have $\mathbb{P}(|C_1| \leqslant n/2) \leqslant e^{-bn}$ and $S_n(p) \geqslant n^2/8$.*

Proof. Choose ε such that $\rho(\varepsilon) \geqslant 3/4$ and let $\pi_0 := 1+\varepsilon$. Then the tail estimate (2.14) and monotonicity yield $\mathbb{P}_{n,p}(|C_1| \leqslant n/2) \leqslant \mathbb{P}_{n,\pi_0/n}(|C_1| \leqslant n/2) \leqslant e^{-bn}$. The second conclusion follows from $S_n(p) \geqslant \mathbb{E}|C_1|^2 \geqslant (n/2)^2\mathbb{P}(|C_1| \geqslant n/2)$. ☐

We next prove two convenient auxiliary estimates.

Lemma 2.5. *For all $n \geqslant 1$ and $p \in [0,1]$,*

$$\mathbb{E}_{n,p}\Big(\sum_{i\neq j} |C_i|^2|C_j|^2\Big) \leqslant S_n(p)^2 = \Big(\mathbb{E}_{n,p}\big(\sum_i |C_i|^2\big)\Big)^2, \tag{2.16}$$

$$\mathbb{E}_{n,p}\Big(\sum_{i,j} |C_i|^2|C_j|^2\Big) \leqslant \Big(S_n(p) + 3\big[n^{-1}S_n(p)\big]^4\Big) \cdot S_n(p). \tag{2.17}$$

Proof. We start with (2.16) and fix any vertex $v \in [n]$. Conditioning on the vertex set of $C(v)$ in $G_{n,p}$, the remaining graph with vertex set $[n] \setminus C(v)$ has the same distribution as $G_{n-|C(v)|,p}$ (up to relabeling of the vertices). Since $S_{n-|C(v)|}(p) \leqslant S_n(p)$ by Lemma 2.2, using (2.1) it follows that

$$\mathbb{E}_{n,p}\Big(|C(v)| \sum_{w\notin C(v)} |C(w)| \;\Big|\; C(v)\Big) = |C(v)| \cdot S_{n-|C(v)|}(p) \leqslant |C(v)| \cdot S_n(p). \tag{2.18}$$

Taking the expectation and summing over all vertices $v \in [n]$, we obtain, recalling (2.10) and (2.1)–(2.2),

$$\mathbb{E}_{n,p}\Big(\sum_{i\neq j} |C_i|^2|C_j|^2\Big) = \mathbb{E}_{n,p}\Big(\sum_{v\in[n]} |C(v)| \sum_{w\notin C(v)} |C(w)|\Big) \leqslant \mathbb{E}_{n,p}\Big(\sum_{v\in[n]} |C(v)|\Big) \cdot S_n(p) = S_n(p)^2, \tag{2.19}$$

which is (2.16).

For (2.17) we rely on the classical tree–graph inequalities [1, (5.3)–(5.4)] of Aizenman and Newman from 1984 (see also [13, (6.85)–(6.96)] for a modern exposition). As noted in [1, p. 123], their proofs apply directly to percolation on any finite transitive graph. For any integer $k \geqslant 1$ and vertex $v \in [n]$, these inequalities state (in our notation) that

$$\mathbb{E}_{n,p}\big(|C(v)|^k\big) \leqslant (2k-3)!! \cdot \big(\mathbb{E}_{n,p}|C(v)|\big)^{2k-1}. \tag{2.20}$$

Recalling $S_n(p) = n\mathbb{E}_{n,p}|C(v)|$, see (2.3), by summing (2.20) with $k = 3$ over all vertices $v \in [n]$ we infer

$$\mathbb{E}_{n,p}\Big(\sum_i |C_i|^4\Big) = \sum_{v\in[n]} \mathbb{E}_{n,p}\big(|C(v)|^3\big) \leqslant 3 \cdot \big[n^{-1}S_n(p)\big]^5 \cdot n,$$

which together with (2.16) establishes (2.17). ☐

Proof of (2.5) *of Theorem 2.1.* We shall distinguish five (somewhat overlapping) ranges of np that will be treated separately. We begin by noting that (2.11) and (2.16) together imply

$$\frac{\mathrm{d}}{\mathrm{d}p} \log S_n(p) \leqslant \frac{S_n(p)}{1-p}, \tag{2.21}$$

which will be useful in the subcritical and critical cases.

Let π_0 and b be as in Corollary 2.4 and pick $A \geqslant \max\{\pi_0, 2\}$ such that $exe^{-x/2} \leqslant 1/2$ for $x \geqslant A$. Let $a, B, c > 0$ and $\delta \in (0, 1/2]$ be the constants given in Theorem 2.3(ii)–(iii). We set $\Lambda := \max\{(B/\delta^2)^{1/3}, 1\}$, and henceforth assume that n is large enough whenever necessary. (This is no loss of generality since (2.1)–(2.2) and (2.8) imply $S_n(p) \geqslant n$ and $\frac{\mathrm{d}}{\mathrm{d}p}S_n(p) \leqslant n^4$ while $\min\{|p - 1/n|^{-1}, n^{4/3}\} \geqslant 1$ for every $n \geqslant 1$, and thus (2.5) trivially holds for any fixed n if C is large enough.)

Case 1: $np = 1 - \varepsilon$ with $\varepsilon^3 n \geqslant 1$. By (2.12),

$$S_n(p) \leqslant \varepsilon^{-1}n = |p - 1/n|^{-1}. \tag{2.22}$$

Since $p \leqslant 1/2$ (for $n \geqslant 2$) and $|p - 1/n|^{-1} \leqslant n^{4/3}$, now (2.21) and (2.22) imply (2.5).

Case 2: $|np - 1| \leqslant \Lambda n^{-1/3}$. Noting that $np \leqslant 1 + \varepsilon$ with $\varepsilon = \Lambda n^{-1/3}$ and using the supercritical upper bound (2.13) for $S_n(p)$ it follows that

$$S_n(p) \leqslant D\varepsilon^2 n^2 = D\Lambda^2 \cdot n^{4/3} \leqslant D\Lambda^3 \cdot |p - 1/n|^{-1}.$$

Since $p \leqslant 1/2$ (for $n \geqslant 4\Lambda$, say), now (2.21) implies (2.5).

Case 3: $np = 1+\varepsilon$ with $\Lambda n^{-1/3} \leqslant \varepsilon \leqslant A$. This is a more difficult range. We shall be guided by the so-called 'symmetry rule', which intuitively states the following: after removing the largest component from the supercritical random graph $G_{n,p}$ with $np = 1+\varepsilon$, the remaining graph resembles a subcritical random graph $G_{n',p}$ with suitable n' and $n'p = 1-\varepsilon'$, see [21, Section 5.6].

Let

$$\alpha(\varepsilon) = (1-\delta)\rho(\varepsilon), \tag{2.23}$$

so that $(n-\alpha(\varepsilon)n)\cdot p \leqslant 1 - c\varepsilon$ by (2.15). Using the subcritical estimate (2.12) of Theorem 2.3, it follows that for $D_1 := c^{-1}$ we have

$$S_{\lfloor n-\alpha(\varepsilon)n\rfloor}(p) \leqslant (c\varepsilon)^{-1}n = D_1\varepsilon^{-1}n. \tag{2.24}$$

Note that $|C_1| \leqslant \alpha(\varepsilon)n$ is a decreasing event, and that $S(G_{n,p}) = \sum_i |C_i|^2$ and thus $S(G_{n,p})^2 = \sum_{i,j}|C_i|^2|C_j|^2$ are increasing functions of the edge indicators. By Harris's inequality (a special case of the FKG-inequality), it follows that

$$\mathbb{E}\Big(\mathbb{1}_{\{|C_1|\leqslant\alpha(\varepsilon)n\}}\sum_{i,j}|C_i|^2|C_j|^2\Big) \leqslant \mathbb{P}\big(|C_1| \leqslant \alpha(\varepsilon)n\big)\cdot\mathbb{E}\Big(\sum_{i,j}|C_i|^2|C_j|^2\Big). \tag{2.25}$$

Combining (2.25) with (2.23), the tail estimate (2.14) and the inequality (2.17), using the upper bound (2.13) for $S_n(p)$, it follows that

$$\mathbb{E}\Big(\mathbb{1}_{\{|C_1|\leqslant\alpha(\varepsilon)n\}}\sum_{i,j}|C_i|^2|C_j|^2\Big) \leqslant e^{-a\delta^2\varepsilon^3 n}\cdot\Big(D\varepsilon^2n^2 + 3\big[D\varepsilon^2 n\big]^4\Big)\cdot S_n(p) = O(\varepsilon^{-1}n)\cdot S_n(p), \tag{2.26}$$

where we used $e^{-x}(x+x^3) \leqslant 2$ for the last inequality (and that a, δ, D are constants).

Conditioning on (the vertex set of) the largest component C_1 of $G_{n,p}$, the remaining graph with vertex set $[n]\setminus C_1$ has the same distribution as $G_{n-|C_1|,p}$ conditioned on the event $\mathcal{D}_{C_1}$ that all components have size at most $|C_1|$ and that there is no component of size exactly $|C_1|$ with a smaller vertex label than C_1. Similarly to (2.18), it follows that

$$\mathbb{E}\Big(|C_1|^2\sum_{i\geqslant 2}|C_i|^2 \;\Big|\; C_1\Big) = |C_1|^2\cdot\mathbb{E}\big(S(G_{n-|C_1|,p}) \;\big|\; C_1,\, \mathcal{D}_{C_1}\big). \tag{2.27}$$

For any given C_1, $\mathcal{D}_{C_1}$ is a decreasing event for the random graph $G_{n-|C_1|,p}$, while $S(G_{n-|C_1|,p})$ is an increasing function. Hence, as in (2.25), by Harris's inequality, it follows that

$$\mathbb{E}\big(S(G_{n-|C_1|,p}) \mid C_1,\, \mathcal{D}_{C_1}\big) \leqslant \mathbb{E}\big(S(G_{n-|C_1|,p}) \mid C_1\big) = S_{n-|C_1|}(p). \tag{2.28}$$

By (2.27)–(2.28) and the monotonicity of Lemma 2.2 together with (2.24), we infer

$$\begin{aligned}\mathbb{E}_{n,p}\Big(\mathbb{1}_{\{|C_1|\geqslant\alpha(\varepsilon)n\}}|C_1|^2\sum_{i\geqslant 2}|C_i|^2 \;\Big|\; C_1\Big) &\leqslant \mathbb{1}_{\{|C_1|\geqslant\alpha(\varepsilon)n\}}|C_1|^2 S_{n-|C_1|}(p)\\ &\leqslant S_{\lfloor n-\alpha(\varepsilon)n\rfloor}(p)\cdot|C_1|^2 \leqslant D_1\varepsilon^{-1}n|C_1|^2\end{aligned} \tag{2.29}$$

and thus, by taking the expectation and using (2.2),

$$\mathbb{E}_{n,p}\Big(\mathbb{1}_{\{|C_1|\geqslant\alpha(\varepsilon)n\}}|C_1|^2\sum_{i\geqslant 2}|C_i|^2\Big) \leqslant D_1\varepsilon^{-1}n\mathbb{E}_{n,p}|C_1|^2 \leqslant D_1\varepsilon^{-1}nS_n(p). \tag{2.30}$$

Similarly to (2.27)–(2.30), by combining (2.17) with the upper bound (2.24) for $S_{\lfloor n-\alpha(\varepsilon)n\rfloor}(p)$, we deduce

$$\begin{aligned}\mathbb{E}\Big(\mathbb{1}_{\{|C_1|\geqslant\alpha(\varepsilon)n\}}\sum_{i,j\geqslant 2}|C_i|^2|C_j|^2\Big) &\leqslant \mathbb{E}_{\lfloor n-\alpha(\varepsilon)n\rfloor,p}\Big(\sum_{i,j}|C_i|^2|C_j|^2\Big)\\ &\leqslant \Big(D_1\varepsilon^{-1}n + 3\big[D_1\varepsilon^{-1}\big]^4\Big)\cdot S_{\lfloor n-\alpha(\varepsilon)n\rfloor}(p) = O(\varepsilon^{-1}n)\cdot S_n(p),\end{aligned} \tag{2.31}$$

where we used $\varepsilon^3 n \geqslant \Lambda^3 \geqslant 1$ and $S_{\lfloor n-\alpha(\varepsilon)n\rfloor}(p) \leqslant S_n(p)$ (see Lemma 2.2) for the final inequality.

In view of (2.11), using $p \leqslant 1/2$ (for $n \geqslant 2(1+A)$, say) our estimates (2.26), (2.30) and (2.31) imply

$$\frac{\mathrm{d}}{\mathrm{d}p} S_n(p) = O(\varepsilon^{-1} n) \cdot S_n(p),$$

which due to $\varepsilon^{-1} n = |p - 1/n|^{-1}$ and $|p-1/n|^{-1} \leqslant n^{4/3}/\Lambda$ yields (2.5) in this case too.

Case 4: $A \leqslant np \leqslant n/2$. In this range many technicalities from the previous case simplify. By distinguishing the events $|C_1| \leqslant n/2$ and $n/2 < |C_1| \leqslant n$ (in which case $|C_2| \leqslant n - |C_1| < n/2$), using $\sum_i |C_i| = n$ we infer

$$\sum_{i \neq j} |C_i|^2 |C_j|^2 \leqslant \mathbb{1}_{\{|C_1| \leqslant n/2\}} n^4 + 2n^2 \sum_i \mathbb{1}_{\{|C_i| \leqslant n/2\}} |C_i|^2. \tag{2.32}$$

As $enpe^{-np/2} \leqslant 1/2$ by the choice of A, standard component counting arguments from random graph theory and Stirling's formula ($k! \geqslant \sqrt{2\pi k}(k/e)^k$) yield

$$\begin{aligned} \mathbb{E}\Big(\sum_i \mathbb{1}_{\{|C_i| \leqslant n/2\}} |C_i|^2\Big) &\leqslant \sum_{1 \leqslant k \leqslant n/2} k^2 \cdot \binom{n}{k} k^{k-2} p^{k-1} (1-p)^{k(n-k)} \\ &\leqslant \sum_{k \geqslant 1} \frac{(knp)^k e^{-knp/2}}{k!\, p} \leqslant \frac{1}{p} \sum_{k \geqslant 1} \frac{\big(enpe^{-np/2}\big)^k}{\sqrt{2\pi k}} \leqslant \frac{1}{p}. \end{aligned} \tag{2.33}$$

Since $np \geqslant \pi_0$, by Corollary 2.4 we see that for large n we also have

$$\mathbb{E}(\mathbb{1}_{\{|C_1| \leqslant n/2\}} n^4) = \mathbb{P}(|C_1| \leqslant n/2) \cdot n^4 \leqslant n^4 e^{-bn} = O(1). \tag{2.34}$$

Inserting (2.32)–(2.34) into (2.11) and using $(1-p)^{-1} \leqslant 2$ we obtain

$$\frac{\mathrm{d}}{\mathrm{d}p} S_n(p) = O\big(n^2 p^{-1}\big). \tag{2.35}$$

Since $S_n(p) \geqslant n^2/8$ by Corollary 2.4, this yields $\frac{\mathrm{d}}{\mathrm{d}p} \log S_n(p) = O(p^{-1})$, which establishes (2.5) because now $p^{-1} \leqslant |p - 1/n|^{-1}$ and $p^{-1} = O(n) = O(n^{4/3})$.

Case 5: $(\log n)^2 \leqslant np < n$. This is a less interesting range since with very high probability, $G_{n,p}$ is connected and thus $\sum_i |C_i|^2 = |C_1|^2 = n^2$. To obtain rigorous estimates, let $\mathcal{E}$ denote the monotone increasing event that $G_{n,p}$ is 2-edge connected (after deleting any edge the resulting graph remains connected). It is well-known that $\mathbb{P}_{n,2(\log n)/n}(\neg\mathcal{E}) = o(1)$ holds (see, e.g., [12]), so a multi-round exposure argument yields $\mathbb{P}_{n,p}(\neg\mathcal{E}) \leqslant \mathbb{P}_{n,2\log n/n}(\neg\mathcal{E})^{\lfloor np/2\log n\rfloor} \leqslant n^{-\omega(1)}$. Observe that if $\mathcal{E}$ holds, then no edge can be pivotal for the event $v \leftrightarrow w$. Using (2.8) we infer

$$\frac{\mathrm{d}}{\mathrm{d}p} S_n(p) \leqslant n^4 \cdot \mathbb{P}_{n,p}(\neg\mathcal{E}) \leqslant n^{-\omega(1)},$$

which together with $S_n(p) \geqslant 1$ and $|p - 1/n| \geqslant 1/n$ completes the proof of (2.5). □

2.2 Lower bound

In this subsection we focus on the lower bound (2.6) in Theorem 2.1. Our proof strategy is to consider the event that $G_{n,p}$ contains two distinct components of size $\Theta(n^{2/3})$.

Lemma 2.6. *Let $\mathcal{L}$ be the event that $|C_2| \geqslant n^{2/3}$, i.e., that $G_{n,p}$ contains two distinct components with at least $n^{2/3}$ vertices each. For every $\lambda \in \mathbb{R}$ there exist constants $\delta_\lambda, n_0 > 0$ such that, for all $n \geqslant n_0$, if $p = 1/n + \lambda n^{-4/3}$, then*

$$\mathbb{P}_{n,p}(\mathcal{L}) \geqslant \delta_\lambda. \tag{2.36}$$

Proof. This follows immediately from [3, Corollary 2]. (See also [21, Theorem 5.20], there stated for $G(n,m)$.) □

Proof of (2.6) *of Theorem 2.1.* As for the upper bound, we may assume that n is large enough, since (2.6) trivially holds (if D_λ is chosen small enough) for every fixed $n \geqslant 2$ because $S_n(p)$ and $\frac{\mathrm{d}}{\mathrm{d}p}S_n(p)$ are positive functions on $(0,1)$.

With $\mathcal{L}$ as in Lemma 2.6, we have

$$\sum_{i\neq j} |C_i|^2|C_j|^2 \geqslant \mathbb{1}_{\{\mathcal{L}\}}\big(n^{2/3}\big)^4,$$

and thus by (2.9)–(2.10) and (2.36)

$$\frac{\mathrm{d}}{\mathrm{d}p}S_n(p) \geqslant \mathbb{E}_{n,p}\Big(\sum_{i\neq j}|C_i|^2|C_j|^2\Big) \geqslant \mathbb{P}_{n,p}(\mathcal{L})n^{8/3} \geqslant \delta_\lambda \cdot n^{8/3}. \tag{2.37}$$

By (2.13) (with $\varepsilon := \max\{1,\lambda\}n^{-1/3}$) we also know that $p = 1/n + \lambda n^{-4/3}$ implies $S_n(p) \leqslant C_\lambda n^{4/3}$, establishing (2.6) with $D_\lambda = \delta_\lambda/C_\lambda$ for n sufficiently large. □

Remark 2.7. *Although we have stated* (2.6) *and Lemma 2.6 for a fixed λ, the results hold uniformly for λ in any compact interval, i.e., we can take D_λ and δ_λ independent of $\lambda \in [-\Lambda,\Lambda]$ for any $\Lambda > 0$, provided we assume for example $n \geqslant 2\max\{1,\Lambda^3\}$ (to guarantee that $p \in (0,1)$). This follows from the more refined Theorem 1.3, but it can also be seen from the simple proof above by noting that the result in [3, Corollary 2], although stated for $p = 1/n + \lambda n^{-4/3}$ for a fixed λ, also holds (by the same proof) more generally for $p = 1/n + (\lambda + o(1))n^{-4/3}$; it then follows from Lemma 3.3 below that for $\lambda \in [-\Lambda,\Lambda]$, $\mathbb{P}_{n,n^{-1}+\lambda n^{-4/3}}(\mathcal{L})$ converges uniformly to a continuous positive function, which yields a uniform lower bound in* (2.36)*, and thus in* (2.6).

3 Scaling inside the critical window

In this section we prove Theorem 1.3. Our arguments exploit that inside the critical window, the rescaled sizes of the largest components converge to some random variables (as mentioned in the introduction).

Following [23], we define

$$\Lambda^{(\lambda)}(x) := (2\pi)^{-1/2}x^{-5/2}e^{-F(x,\lambda)}\sum_{\ell\geqslant 0} w_\ell x^{3\ell/2}, \tag{3.1}$$

where

$$F(x,\lambda) := \big((x-\lambda)^3 + \lambda^3\big)/6, \tag{3.2}$$

and w_ℓ, $\ell \geqslant 0$, are Wright's constants [37], which as shown by Spencer [35] can be expressed as

$$w_\ell = \mathbb{E}(\mathcal{B}_{\mathrm{ex}}^\ell)/\ell!, \tag{3.3}$$

where the random variable $\mathcal{B}_{\mathrm{ex}}$ is the area under a normalized Brownian excursion, see also the survey [19]. As shown in [23, Theorem 4.1], $\Lambda^{(\lambda)}(x)$ is the intensity of the point process that by [3] describes asymptotically the sequence $(|C_i|/n^{2/3})_{i\geqslant 1}$, and we define the corresponding moments

$$f_k(\lambda) = \int_0^\infty x^k\Lambda^{(\lambda)}(x)\mathrm{d}x, \qquad k \geqslant 2. \tag{3.4}$$

As remarked in [23, after Corollary 4.2], $\Lambda^{(\lambda)}(x)$ decreases exponentially as $x \to \infty$, and is $\Theta\big(x^{-5/2}\big)$ as $x \to 0$; hence the integral (3.4) converges so $0 < f_k(\lambda) < \infty$ for every $k \geqslant 2$ and $\lambda \in \mathbb{R}$.

By (3.2) we have $\frac{\partial}{\partial\lambda}F(x,\lambda) = -x^2/2 + \lambda x$ and thus by (3.1) $\frac{\partial}{\partial\lambda}\Lambda^{(\lambda)}(x) = \big(\frac{x^2}{2} - \lambda x\big)\Lambda^{(\lambda)}(x)$. Hence, by differentiating inside the integral in (3.4) (which is easily justified, e.g. using dominated convergence), $f_k(\lambda)$ is differentiable and

$$\frac{\mathrm{d}}{\mathrm{d}\lambda}f_k(\lambda) = \int_0^\infty x^k\frac{\partial\Lambda^{(\lambda)}(x)}{\partial\lambda}\mathrm{d}x = \int_0^\infty x^k\Big(\frac{x^2}{2} - \lambda x\Big)\Lambda^{(\lambda)}(x)\mathrm{d}x = \frac{1}{2}f_{k+2}(\lambda) - \lambda f_{k+1}(\lambda). \tag{3.5}$$

By induction, $f_k(\lambda)$ is infinitely differentiable for every $k \geqslant 2$.

Recall now (2.2), and note that (2.11) can be written as

$$\frac{\mathrm{d}}{\mathrm{d}p} S_n(p) = \frac{\mathbb{E}_{n,p}\big(\big(\sum_i |C_i|^2\big)^2\big)}{1-p} - \frac{\mathbb{E}_{n,p}\big(\sum_i |C_i|^4\big)}{1-p}. \tag{3.6}$$

To treat such sums, we first note the following fact, which is stated in [22, Theorem B1 and Remark B2] as an immediate consequence of results of Aldous [3] and Janson and Spencer [23].

Lemma 3.1 ([3, 23, 22]). *Let $\lambda \in \mathbb{R}$ and $k \in \mathbb{N}$ with $k \geqslant 2$. Then there exists a random variable $W_{\lambda,k}$ with*

$$\mathbb{E} W_{\lambda,k} = f_k(\lambda), \tag{3.7}$$

such that for $p = 1/n + (\lambda + o(1))n^{-4/3}$ we have

$$\frac{\sum_i |C_i|^k}{n^{2k/3}} \xrightarrow{\mathrm{d}} W_{\lambda,k}. \tag{3.8}$$

□

In Section 3.1 we justify taking expectations, higher moments and derivatives in (3.8), and use this to establish the convergence results (1.5)–(1.6) of Theorem 1.3 with $f = f_2$. (In Appendix C we extend this argument to higher derivatives.) Finally, in Section 3.2 we complete the proof of Theorem 1.3 by showing $\frac{\mathrm{d} \log f_2}{\mathrm{d}\lambda}(0) \neq 0$ numerically via a series expansion (that converges exponentially).

3.1 Convergence

In this subsection we prove the convergence results (1.5)–(1.6) of Theorem 1.3 using the distributional convergence (3.8) from Lemma 3.1 and the following auxiliary result.

Theorem 3.2. *Let $D, \lambda \in \mathbb{R}$ and $k, q \in \mathbb{N}$ with $k \geqslant 2$ and $q \geqslant 1$.*

(i) *There exists $C = C(D,k,q)$ such that, for all $n \geqslant 1$ and $p \in (0,1)$ satisfying $p \leqslant 1/n + Dn^{-4/3}$,*

$$0 \leqslant \mathbb{E}_{n,p}\left(\left(\frac{\sum_i |C_i|^k}{n^{2k/3}}\right)^q\right) \leqslant C. \tag{3.9}$$

(ii) *For $p = 1/n + (\lambda + o(1))n^{-4/3}$ we have*

$$\mathbb{E}_{n,p}\left(\left(\frac{\sum_i |C_i|^k}{n^{2k/3}}\right)^q\right) \xrightarrow{n\to\infty} \mathbb{E}\big((W_{\lambda,k})^q\big) < \infty, \tag{3.10}$$

where the random variable $W_{\lambda,k}$ is defined as in Lemma 3.1. Moreover, the limit in (3.10) is a continuous function of λ, and if $p = 1/n + \lambda n^{-4/3}$, then the convergence in (3.10) is uniform for λ in any compact interval $[\lambda_1, \lambda_2] \subset \mathbb{R}$.

Proof. We start with the uniform moment bound (3.9). Since $\sum_i |C_i|^k$ does not decrease if any edge is added, the expectation is a monotone function of p; thus it suffices to consider $p = 1/n + Dn^{-4/3}$. As a warm-up, we first consider the special case $q = 2$. Similarly to (2.10) we have

$$\Big(\sum_i |C_i|^k\Big)^2 = \sum_{v\in[n]} |C(v)|^{k-1} \sum_{w \notin C(v)} |C(w)|^{k-1} + \sum_{v \in [n]} |C(v)|^{2k-1}.$$

Mimicking the conditioning and monotonicity arguments leading to (2.16), see (2.18), we infer that

$$\mathbb{E}_{n,p}\left(\left(\frac{\sum_i |C_i|^k}{n^{2k/3}}\right)^2\right) \leqslant \frac{\big(\mathbb{E}_{n,p} \sum_{v\in[n]} |C(v)|^{k-1}\big)^2}{n^{4k/3}} + \frac{\mathbb{E}_{n,p} \sum_{v\in[n]} |C(v)|^{2k-1}}{n^{4k/3}}.$$

Generalizing the above argument, for every integer $q \geqslant 1$ there is a constant $A_{k,q}$ such that

$$\mathbb{E}_{n,p}\Bigl(\Bigl(\frac{\sum_i |C_i|^k}{n^{2k/3}}\Bigr)^q\Bigr) \leqslant A_{k,q} \sum_{1\leqslant r\leqslant q} \sum_{\substack{j_1+\cdots j_r=qk:\\ k|j_i}} \frac{\prod_{1\leqslant i\leqslant r} \mathbb{E}_{n,p} \sum_{v\in[n]} |C(v)|^{j_i-1}}{n^{2qk/3}}. \tag{3.11}$$

By [23, Corollary 5.3] (or by inserting $\mathbb{E}_{n,p}|C(v)| = n^{-1}S_n(p) = O(n^{1/3})$, see (2.3) and (2.13), into (2.20)) there are constants $(B_{j,D})_{j\geqslant 2}$ such that

$$\mathbb{E}_{n,p} \sum_{v\in[n]} |C(v)|^{j-1} = \mathbb{E}_{n,p} \sum_i |C_i|^j \leqslant B_{j,D} n^{2j/3}. \tag{3.12}$$

Since $j_i \geqslant k \geqslant 2$, (3.12) applies to each factor in each product in (3.11), so (3.9) follows for suitable $C = C(D, k, q)$.

We next turn to (ii), and thus assume $p = 1/n + (\lambda + o(1))n^{-4/3}$. For brevity we write

$$X_{n,k} := \frac{\sum_i |C_i|^k}{n^{2k/3}}. \tag{3.13}$$

The upper bound (3.9), with $2q$, say, shows that the random variables $\bigl((X_{n,k})^q\bigr)_{n\geqslant 1}$ are uniformly integrable for fixed $k \geqslant 2$ and $q \geqslant 1$, see e.g. [14, Theorem 5.4.2]. Since $X_{n,k} \overset{\mathrm{d}}{\to} W_{\lambda,k}$ by (3.8), and thus $(X_{n,k})^q \overset{\mathrm{d}}{\to} (W_{\lambda,k})^q$ by the continuous mapping theorem [14, Theorem 5.10.4]), it thus follows that $\mathbb{E}_{n,p}((X_{n,k})^q) \to \mathbb{E}((W_{\lambda,k})^q) < \infty$ as $n \to \infty$, see [14, Theorem 5.5.9], which completes the proof of (3.10).

The final claims now follow by the following elementary calculus lemma. □

Lemma 3.3. *Suppose that $h_n(\lambda)$ and $h(\lambda)$ are real-valued functions on $\mathbb{R}$ such that if $\lambda \in \mathbb{R}$ and $\lambda_n = \lambda+o(1)$, then $h_n(\lambda_n) \to h(\lambda)$ as $n \to \infty$. Then $h(\lambda)$ is continuous and $h_n(\lambda) \to h(\lambda)$ uniformly for λ in any compact set.*

Proof. First, suppose that h is discontinuous at some λ. Then there exist $\varepsilon > 0$ and a sequence $\lambda_k \to \lambda$ such that $|h(\lambda_k) - h(\lambda)| > \varepsilon$ for all k. Since $h_n(\lambda_k) \to h(\lambda_k)$, we may find an increasing sequence n_k such that $|h_{n_k}(\lambda_k) - h(\lambda_k)| < \varepsilon/2$. Then $|h_{n_k}(\lambda_k) - h(\lambda)| > \varepsilon/2$. On the other hand, the assumption implies $h_{n_k}(\lambda_k) \to h(\lambda)$, a contradiction.

Similarly, assume that $h_n(\lambda)$ does not converge uniformly to $h(\lambda)$ on the compact set K. Then there exist $\varepsilon > 0$ and sequences $n_k \to \infty$ and $\lambda_k \in K$ such that $|h_{n_k}(\lambda_k) - h(\lambda_k)| > \varepsilon$. Since K is compact, we may select a subsequence such that, along this subsequence, $\lambda_k \to \lambda$ for some λ. Then the assumption and the continuity of h just shown imply that, along the subsequence, $h_{n_k}(\lambda_k) \to h(\lambda)$ and $h(\lambda_k) \to h(\lambda)$, and thus $h_{n_k}(\lambda_k) - h(\lambda_k) \to 0$, a contradiction. □

Remark 3.4. *Lemma 3.3 is valid for functions on any metric space. (Also the ranges of the functions may be in an arbitrary metric space.) Furthermore, the converse of the lemma also holds (and is easy): if $h_n(\lambda) \to h(\lambda)$ uniformly on compact sets and h is continuous, then $h_n(\lambda_n) \to h(\lambda)$ whenever $\lambda_n \to \lambda$.*

Proof of (1.5)–(1.6) *of Theorem 1.3.* Recall that $\chi_{K_n}(p) = S_n(p)/n$ by (2.3). Define $X_{n,k}$ as in (3.13). Using also (2.2), (3.10) and (3.7), we have

$$\frac{\chi_{K_n}(p)}{n^{1/3}} = \frac{S_n(p)}{n^{4/3}} = \mathbb{E}_{n,p} X_{n,2} \xrightarrow{n\to\infty} \mathbb{E} W_{\lambda,2} = f_2(\lambda) > 0, \tag{3.14}$$

where $f_2(\lambda) > 0$ follows from the definition (3.4). This proves (1.5).

Similarly, by (3.6), (3.10) and (3.7), we have

$$\frac{\frac{\mathrm{d}}{\mathrm{d}p}\chi_{K_n}(p)}{n^{5/3}} = \frac{\frac{\mathrm{d}}{\mathrm{d}p}S_n(p)}{n^{8/3}} = \frac{\mathbb{E}_{n,p}(X_{n,2}^2)}{1-p} - \frac{\mathbb{E}_{n,p}X_{n,4}}{1-p} \xrightarrow{n\to\infty} \mathbb{E}(W_{\lambda,2}^2) - \mathbb{E}W_{\lambda,4} =: g(\lambda). \tag{3.15}$$

By Lemma 3.3, $f_2(\lambda)$ and $g(\lambda)$ are continuous, and for $p = 1/n + \lambda n^{-4/3}$, the limits hold uniformly on compact sets. Combining (3.14)–(3.15) we infer

$$\frac{\frac{\mathrm{d}}{\mathrm{d}p}\log\chi_{K_n}(p)}{n^{4/3}} \xrightarrow{n\to\infty} \frac{g(\lambda)}{f_2(\lambda)},$$

and thus $g(\lambda)/f_2(\lambda) > 0$ follows from (2.4) and (2.6).

It remains to prove that $g(\lambda) = \frac{\mathrm{d}}{\mathrm{d}\lambda} f_2(\lambda)$ holds. To this end we fix $\lambda_1, \lambda_2 \in \mathbb{R}$ with $\lambda_1 < \lambda_2$, and set $p_i = 1/n + \lambda_i n^{-4/3}$. By (3.14) we have

$$\int_{p_1}^{p_2} \frac{\frac{\mathrm{d}}{\mathrm{d}p}\chi_{K_n}(p)}{n^{1/3}}\,\mathrm{d}p = \frac{\chi_{K_n}(p_2)}{n^{1/3}} - \frac{\chi_{K_n}(p_1)}{n^{1/3}} \xrightarrow{n\to\infty} f_2(\lambda_2) - f_2(\lambda_1). \tag{3.16}$$

On the other hand, by substituting $p = 1/n + \lambda n^{-4/3}$ we have by the uniform convergence in (3.15) just shown (or by dominated convergence and (3.9)) that

$$\int_{p_1}^{p_2} \frac{\frac{\mathrm{d}}{\mathrm{d}p}\chi_{K_n}(p)}{n^{1/3}}\,\mathrm{d}p = \int_{\lambda_1}^{\lambda_2} \frac{\frac{\mathrm{d}}{\mathrm{d}p}\chi_{K_n}(p)\big|_{p=1/n+\lambda n^{-4/3}}}{n^{5/3}}\,\mathrm{d}\lambda \xrightarrow{n\to\infty} \int_{\lambda_1}^{\lambda_2} g(\lambda)\,\mathrm{d}\lambda. \tag{3.17}$$

It follows from (3.16)–(3.17) that $f_2(\lambda_2) - f_2(\lambda_1) = \int_{\lambda_1}^{\lambda_2} g(\lambda)\,\mathrm{d}\lambda$. Since $\lambda_1 < \lambda_2$ were arbitrary, and $g(\lambda)$ is continuous, it follows that $g(\lambda) = \frac{\mathrm{d}}{\mathrm{d}\lambda} f_2(\lambda)$ for all $\lambda \in \mathbb{R}$, completing the proof. $\square$

3.2 Explicit bounds for $\lambda = 0$

In this subsection we complete the proof of Theorem 1.3, and by the arguments of Section 3.1 it remains to prove the following technical lemma.

Lemma 3.5. *Define the function f_2 as in (3.4). Then $\frac{\mathrm{d}^2}{\mathrm{d}\lambda^2} \log f_2(0) \neq 0$.*

Remark 3.6. *The proof of Lemma 3.5 shows that $\frac{\mathrm{d}^2}{\mathrm{d}\lambda^2} \log f_2(0) \approx 0.296833365232$.*

The idea is to give (in the special case $\lambda = 0$) rigorous numerical estimates for the right hand side of

$$\frac{\mathrm{d}^2}{\mathrm{d}\lambda^2} \log f_2(\lambda) = \frac{f_2(\lambda)\frac{\mathrm{d}^2}{\mathrm{d}\lambda^2} f_2(\lambda) - \big(\frac{\mathrm{d}}{\mathrm{d}\lambda} f_2(\lambda)\big)^2}{\big(f_2(\lambda)\big)^2}. \tag{3.18}$$

The derivatives can be computed by (3.5). In the special case $\lambda = 0$ we obtain

$$\frac{\mathrm{d}}{\mathrm{d}\lambda} f_2(0) = \tfrac{1}{2} f_4(0), \tag{3.19}$$

$$\frac{\mathrm{d}^2}{\mathrm{d}\lambda^2} f_2(0) = \tfrac{1}{4} f_6(0) - f_3(0). \tag{3.20}$$

Furthermore, by [23, Remark 6] we also have (in our notation) the identity $f_3(\lambda) = 2 + 2\lambda f_2(\lambda)$, so that $f_3(0) = 2$. Hence (3.18) yields

$$\frac{\mathrm{d}^2}{\mathrm{d}\lambda^2} \log f_2(0) = \frac{f_2(0)\big(f_6(0) - 8\big) - f_4(0)^2}{4 f_2(0)^2}, \tag{3.21}$$

and due to $0 < f_2(\lambda) < \infty$ our task is reduced to showing that

$$f_2(0) f_6(0) - 8 f_2(0) - f_4(0)^2 \neq 0. \tag{3.22}$$

To evaluate the terms in (3.22), note that by Tonelli's theorem, the function f_k defined in (3.4) can be written as

$$f_k(\lambda) = \sum_{\ell \geqslant 0} (2\pi)^{-1/2} w_\ell \int_0^\infty x^{k+3\ell/2-5/2} e^{-F(x,\lambda)}\mathrm{d}x. \tag{3.23}$$

Since $F(x,0) = x^3/6$, see (3.2), we can in the case $\lambda = 0$ evaluate the integral in (3.23) using the gamma function $\Gamma(z) := \int_0^\infty x^{z-1} e^{-x}\mathrm{d}x$. We define, using the substitution $y = x^3/6$,

$$I_{k,\ell} := \int_0^\infty x^{k+3\ell/2-5/2} e^{-F(x,0)}\mathrm{d}x = \frac{1}{3} \cdot 6^{k/3+(\ell-1)/2} \cdot \Gamma(k/3 + (\ell-1)/2), \tag{3.24}$$

and by (3.23) thus have

$$f_k(0) = \sum_{\ell \geqslant 0} (2\pi)^{-1/2} w_\ell I_{k,\ell}. \tag{3.25}$$

The plan is to truncate the infinite sum in (3.25) with the help of the following uniform estimates.

Lemma 3.7. *For all $k \geqslant 2$ and $\ell \geqslant 1$ we have $I_{k,\ell}, w_\ell \geqslant 0$ and*

$$I_{k,\ell} \leqslant 2\pi^{1/2}(3\ell)^{k/3-1}\Big(\frac{3\ell}{e}\Big)^{\ell/2} \cdot e^{2k^2/9\ell}, \tag{3.26}$$

$$w_\ell \leqslant 8\pi^{-1/2}\sqrt{\ell}\Big(\frac{e}{12\ell}\Big)^{\ell/2}. \tag{3.27}$$

The upper bound (3.27) is off from the asymptotic value in [19, (52)] only by a factor 8/3.

Proof. The lower bounds $I_{k,\ell} \geqslant 0$ and $w_\ell \geqslant 0$ are trivial. Turning to upper bounds, we start with $I_{k,\ell}$. We use the well-known Stirling-type estimate (see, e.g., [30, (5.6.1)])

$$1 \leqslant \frac{\Gamma(m)}{\sqrt{2\pi/m}(m/e)^m} \leqslant e^{1/12m}. \tag{3.28}$$

Inserting (3.28) into (3.24), it follows by a simple calculation that, using $k \geqslant 2$ and and $1 + x \leqslant e^x$,

$$\begin{aligned}
I_{k,\ell} &\leqslant \frac{1}{3} \cdot \frac{\sqrt{2\pi}}{\sqrt{k/3 + (\ell-1)/2}} \cdot \Big(\frac{2k + 3(\ell-1)}{e}\Big)^{k/3+(\ell-1)/2} \cdot e^{1/4k} \\
&\leqslant \frac{\sqrt{4\pi}}{3\sqrt{\ell}} \cdot \Big(\frac{3\ell}{e}\Big)^{k/3+(\ell-1)/2} \cdot \Big(1 + \frac{2k-3}{3\ell}\Big)^{k/3+\ell/2} \cdot e^{1/4k}. \\
&\leqslant \frac{\sqrt{4\pi}}{3\sqrt{\ell}} \cdot \Big(\frac{3\ell}{e}\Big)^{k/3+(\ell-1)/2} \cdot e^{1/4k+2k^2/9\ell+k/3-1/2} \\
&\leqslant \frac{\sqrt{4\pi}}{\sqrt{3}} \cdot (3\ell)^{k/3-1} \cdot \Big(\frac{3\ell}{e}\Big)^{\ell/2} \cdot e^{1/8+2k^2/9\ell},
\end{aligned}$$

which due to $\sqrt{4/3}e^{1/8} < 2$ completes the proof of (3.26).

For w_ℓ, we combine (3.3) with a recurrence formula by Louchard [26] for the Brownian excursion area, see [19, (4) and (5)]: in the recursion [19, (5)] all γ_i are positive by [19, (4)], so the first term on the right hand side of [19, (5)] gives an upper bound, which implies

$$w_\ell = \frac{1}{\ell!} \cdot \mathbb{E}(\mathcal{B}_{\text{ex}}^\ell) \leqslant \frac{1}{\ell!} \cdot \frac{2\sqrt{\pi}}{(36\sqrt{2})^\ell \Gamma((3\ell-1)/2)} \cdot \frac{12\ell}{6\ell-1}\frac{\Gamma(3\ell+1/2)}{\Gamma(\ell+1/2)}. \tag{3.29}$$

Using $\ell! = \ell\Gamma(\ell)$ and the estimate (3.28) four times, it follows by a simple (but slightly tedious) calculation that

$$\begin{aligned}
w_\ell &\leqslant \frac{1}{\sqrt{\pi\ell}}\Big(\frac{e}{36\sqrt{2}\ell}\Big)^\ell \cdot \frac{3\ell-1}{2\sqrt{e}}\Big(\frac{2e}{3\ell-1}\Big)^{3\ell/2} \cdot \frac{12\ell}{6\ell-1}\Big(\frac{3\ell+1/2}{e}\Big)^{3\ell}\Big(\frac{e}{\ell+1/2}\Big)^\ell e^{1/36\ell} \\
&\leqslant \frac{3e^{1/36\ell}}{\sqrt{\pi e}} \cdot \sqrt{\ell} \cdot \Big(\frac{e^2}{36\sqrt{2}\ell^2}\Big)^\ell \cdot 2^{3\ell/2}\Big(1 + \frac{3/2}{3\ell-1}\Big)^{3\ell/2} \cdot \Big(\frac{3\ell}{e}\Big)^{3\ell/2}\Big(1 + \frac{1}{6\ell}\Big)^{3\ell/2} \\
&\leqslant \frac{3e^{(1/36+9/8+3/12-1/2)}}{\sqrt{\pi}} \cdot \sqrt{\ell} \cdot \Big(\frac{e}{12\ell}\Big)^{\ell/2} \leqslant 8\pi^{-1/2}\sqrt{\ell}\Big(\frac{e}{12\ell}\Big)^{\ell/2},
\end{aligned}$$

completing the proof of (3.27). □

Lemma 3.8. *For every real $s \geqslant 0$ and integer $\ell_0 \geqslant 2s$,*

$$\sum_{\ell > \ell_0} \ell^s 2^{-\ell} \leqslant 5\ell_0^s 2^{-\ell_0}.$$

Proof. Let $a_\ell := l^s 2^{-\ell}$. For $\ell \geqslant \ell_0$, we have

$$\frac{a_{\ell+1}}{a_\ell} = \Big(1+\frac{1}{\ell}\Big)^s 2^{-1} \leqslant e^{s/\ell} 2^{-1} \leqslant e^{1/2} 2^{-1} < \frac{5}{6}.$$

Hence, $a_\ell \leqslant (5/6)^{\ell-\ell_0} a_{\ell_0}$ and the result follows by summing a geometric series. □

Corollary 3.9. *For all integers $k \geqslant 2$ and $\ell_0 \geqslant 2k/3 - 1$,*

$$0 \leqslant f_k(0) - \sum_{0\leqslant \ell \leqslant \ell_0} (2\pi)^{-1/2} w_\ell I_{k,\ell} \leqslant 11\, e^{2k^2/9\ell_0} 3^{k/3} \ell_0^{k/3-1/2} 2^{-\ell_0}. \tag{3.30}$$

Proof. The lower bound in (3.30) is trivial by (3.23), (3.24) and $I_{k,\ell}, w_\ell \geqslant 0$. Turning to the upper bound, (3.26)–(3.27) yield

$$(2\pi)^{-1/2} w_\ell I_{k,\ell} \leqslant \frac{16}{\sqrt{2\pi}} e^{2k^2/9\ell} 3^{k/3-1} \ell^{k/3-1/2} 2^{-\ell}.$$

Lemma 3.8 thus gives

$$\sum_{\ell>\ell_0} (2\pi)^{-1/2} w_\ell I_{k,\ell} \leqslant \frac{80}{3\sqrt{2\pi}} e^{2k^2/9\ell_0} 3^{k/3} \ell_0^{k/3-1/2} 2^{-\ell_0},$$

and the result follows. □

The constants w_ℓ are easily computed by recursion, see, e.g., [19, (4)–(5) or (6)–(7)], so the finite sum $\sum_{0\leqslant\ell\leqslant\ell_0} (2\pi)^{-1/2} w_\ell I_{k,\ell}$ can be computed numerically (with arbitrary precision) for any ℓ_0 that is not too large. Together with the estimate in Corollary 3.9 of the remainder, which can be made arbitrarily small by choosing a suitable ℓ_0, this enables us to compute $f_k(0)$ with arbitrary precision for any $k \geqslant 2$.

Proof of Lemma 3.5. Choosing $\ell_0 = 75$, the right hand side of (3.30) is less than 10^{-17} for all $2 \leqslant k \leqslant 6$, with room to spare. Proceeding as discussed above, we then obtain (using Maple)

$$f_2(0) \doteq 1.830470321422761, \tag{3.31}$$
$$f_4(0) \doteq 3.514851319980978, \tag{3.32}$$
$$f_6(0) \doteq 16.922562003970612, \tag{3.33}$$

where $\doteq$ means equality for all but the last digit (which might be off by one). Hence

$$f_2(0) f_6(0) - 8 f_2(0) - f_4(0)^2 \doteq 3.9783051377505, \tag{3.34}$$

which shows (3.22) and thus completes the proof of Lemma 3.5. Remark 3.6 follows by inserting (3.34) and (3.31) into (3.21). □

A Asymptotics of $f(\lambda)$ as $\lambda \to \pm\infty$

In this appendix we prove the asymptotics of the function $f(\lambda) = f_2(\lambda)$ stated after Theorem 1.3, and extend the results to $f_k(\lambda)$ for arbitrary $k \geqslant 2$.

Theorem A.1. *Define $f_k : \mathbb{R} \to (0,\infty)$ as in (3.4). For any fixed $k \geqslant 2$, $f_k(\lambda)$ has the asymptotics*

$$f_k(\lambda) = \frac{(2k-5)!!}{|\lambda|^{2k-3}} \big(1 + O(|\lambda|^{-3})\big) \qquad \text{as } \lambda \to -\infty, \tag{A.1}$$
$$f_k(\lambda) = (2\lambda)^k \big(1 + o(1)\big), \qquad \text{as } \lambda \to +\infty. \tag{A.2}$$

Here, $(2k-5)!!$ is the usual semifactorial, i.e., $(2k-5)!! = \prod_{j=1}^{k-2}(2j-1)$, with $(-1)!! = 1$. In particular, for $f = f_2$, we have $f(\lambda) \sim |\lambda|^{-1}$ as $\lambda \to -\infty$ and $f(\lambda) \sim 4\lambda^2$ as $\lambda \to +\infty$, as said in the introduction.

Proof. For $\lambda < 0$, we use results from [22]. The parametrization there is slightly different, so we define, given $\lambda < 0$, first $p := 1/n + \lambda n^{-4/3}$ and then $t := -\log(1-p) = p + O(p^2) = n^{-1} + (\lambda + o_n(1))n^{-4/3}$, where $o_n(1)$ denotes a quantity that tends to 0 as $n \to \infty$ for fixed λ. Note that (for large n) $t < 1/n$ and $1 - nt = (|\lambda| + o_n(1))n^{-1/3}$. By [22, Theorem 3.4], there exists a polynomial p_k of degree $2k-3$ such that

$$\begin{aligned}\mathbb{E}_{n,p}\sum_i |C_i|^k &= np_k\Big(\frac{1}{1-nt}\Big)\Big(1+O\Big(\frac{1}{n(1-nt)^3}\Big)\Big)\\ &= np_k\big((|\lambda|^{-1}+o_n(1))n^{1/3}\big)\big(1+O\big((|\lambda|+o_n(1))^{-3}\big)\big).\end{aligned} \tag{A.3}$$

Letting a_k be the leading coefficient of p_k, so $p_k(x) \sim a_k x^{2k-3}$ as $x \to \infty$, we obtain by letting $n \to \infty$ in (A.3) and using (3.7) and (3.10),

$$f_k(\lambda) = \lim_{n\to\infty}\Big(n^{-2k/3}\mathbb{E}_{n,p}\sum_i |C_i|^k\Big) = a_k|\lambda|^{-(2k-3)}\big(1+O\big(|\lambda|^{-3}\big)\big). \tag{A.4}$$

Note that this estimate holds uniformly in all $\lambda < 0$. Finally, we note that $a_k = (2k-5)!!$, as remarked in [22, after (7.8)], and (A.1) follows.

For $\lambda > 0$, we use results from [23]. The idea is that as $\lambda \to +\infty$, we approach the supercritical regime, where there is a single giant component C_1 that dominates the sum $\sum_i |C_i|^k$, and that $|C_1| \approx 2\lambda$.

It is shown in [23, Lemma 9.5] that as $\lambda \to +\infty$, the intensity $\Lambda^{(\lambda)}(x)$ is well approximated by the density function of the normal distribution $N(2\lambda, 2\lambda^{-1})$ (except for small x), and (A.2) follows easily by (3.4) and estimates as in the proof of [23, Lemma 9.5]; we omit the details. □

For $\lambda \to -\infty$, we can combine Theorem A.1 with (3.5) and obtain $\frac{\mathrm{d}}{\mathrm{d}\lambda}f(\lambda) = |\lambda|^{-2}\big(1+O(|\lambda|^{-3})\big)$, and similarly for larger k. This extends by induction to higher derivatives; the result shows that we can formally take any number of derivatives in (A.1) (keeping the multiplicative error term $O(|\lambda|^{-3})$).

For $\lambda \to +\infty$, the estimates for $\Lambda^{(\lambda)}(x)$ used in [23] are not precise enough to yield as precise results for derivatives (note that in (3.5) we expect the leading terms of $f_{k+2}(\lambda)/2 \sim (2\lambda)^{k+2}/2$ and $\lambda f_{k+1} \sim \lambda(2\lambda)^{k+1}$ to cancel); we conjecture that here too we can take derivatives formally in (A.2), but we have not tried to prove it. (This would require more precise estimates of $\Lambda^{(\lambda)}(x)$ and thus by (3.1) and (3.3) of $\sum_{\ell\geqslant 0} w_\ell x^{3\ell/2} = \mathbb{E}\exp\big(x^{3/2}\mathcal{B}_{\mathrm{ex}}\big)$. Such estimates can possibly be derived from the asymptotic expansions for the distribution of $\mathcal{B}_{\mathrm{ex}}$ in [20], but we leave this as an open problem.) We note only that (1.6) and (2.5) imply $\frac{\mathrm{d}}{\mathrm{d}\lambda}\log f(\lambda) = O(\min\{|\lambda|^{-1}, 1\})$ for all $\lambda \in \mathbb{R}$.

B Simple bounds for the susceptibility

In this appendix we give complete proofs of (2.12)–(2.13), which we restate as the theorem below. (The bounds are sharp up to constant factors when $np = 1 \mp \varepsilon$, $\varepsilon = O(1)$ and $\varepsilon^3 n \geqslant 1$.)

Theorem B.1. (i) *For all $n \geqslant 1$, $p \in [0,1]$, and $\varepsilon > 0$ satisfying $np \leqslant 1 - \varepsilon$,*

$$\mathbb{E}_{n,p}|C(v)| \leqslant \varepsilon^{-1}. \tag{B.1}$$

(ii) *There is a constant $D > 0$ such that, for all $n \geqslant 1$, $p \in [0,1]$, and $\varepsilon \geqslant 0$ satisfying $np \leqslant 1 + \varepsilon$,*

$$\mathbb{E}_{n,p}|C(v)| \leqslant D\max\{\varepsilon^2 n, n^{1/3}\}. \tag{B.2}$$

Part (i) is easy and well-known, and included for completeness. For part (ii), we do not know any reference with a short proof; the bound is proved in [7] as a special case of a more general and involved result. We give here a more direct argument which adapts recent ideas from percolation theory [24, 16] to the simpler $G_{n,p}$ case.

We start by recalling some well-known branching processes results (we include proofs for completeness). Let $\mathfrak{X}_{n,p}$ denote a Galton–Watson branching process with $\mathrm{Bin}(n,p)$ offspring distribution, starting with a single individual, and let $|\mathfrak{X}_{n,p}|$ be its total size. We define $\mathfrak{X}_\lambda$ and $|\mathfrak{X}_\lambda|$ analogously, with $\mathrm{Bin}(n,p)$ replaced by $\mathrm{Po}(\lambda)$.

Lemma B.2. (i) *For all $n \geqslant 1$, $p \in [0,1]$ and $\lambda \geqslant -n\log(1-p)$, $|\mathfrak{X}_\lambda|$ stochastically dominates $|\mathfrak{X}_{n,p}|$.*

(ii) *There exists a constant $C > 0$ such that, for all $\lambda \geqslant 0$ and $k \geqslant 1$, we have*

$$\mathbb{P}(k \leqslant |\mathfrak{X}_\lambda| \leqslant \infty) \leqslant C\big(\max\{\lambda - 1, 0\} + k^{-1/2}\big). \tag{B.3}$$

Proof. (i): To prove that $|\mathfrak{X}_\lambda|$ stochastically dominates $|\mathfrak{X}_{n,p}|$, it suffices to show that $\mathrm{Po}(-n\log(1-p))$ stochastically dominates $\mathrm{Bin}(n,p)$. Taking n independent couplings, it thus is enough to prove that $X \sim \mathrm{Po}(-\log(1-p))$ stochastically dominates $Y \sim \mathrm{Bin}(1,p)$. This is immediate since $\mathbb{P}(X=0) = 1-p = \mathbb{P}(Y=0)$, $\mathbb{P}(X \geqslant 1) = p = \mathbb{P}(Y=1)$ and $\mathbb{P}(Y \geqslant 2) = 0$.

(ii): Let $(\xi_i)_{i\geqslant 1}$ be a sequence of independent random variables with $\mathrm{Po}(\lambda)$ distribution. For all $k \geqslant 1$, using the classical Otter–Dwass formula [10] and Stirling's formula ($k! \geqslant \sqrt{2\pi k}(k/e)^k$) we infer

$$\mathbb{P}(|\mathfrak{X}_\lambda| = k) = \frac{\mathbb{P}(\xi_1 + \cdots + \xi_k = k-1)}{k} = \frac{\mathbb{P}(\mathrm{Po}(k\lambda) = k-1)}{k} = \frac{e^{-\lambda k}}{k\lambda}\frac{(\lambda k)^k}{k!} \leqslant \frac{(\lambda e^{1-\lambda})^k}{\sqrt{2\pi}k^{3/2}\lambda} \leqslant \frac{e}{\sqrt{2\pi}k^{3/2}},$$

where the last inequality follows by noting $\lambda e^{1-\lambda} \leqslant 1$ and thus $(\lambda e^{1-\lambda})^k \leqslant \lambda e^{1-\lambda} \leqslant e\lambda$. Summing this inequality, we see that there is a constant C such that

$$\mathbb{P}(k \leqslant |\mathfrak{X}_\lambda| < \infty) \leqslant Ck^{-1/2}. \tag{B.4}$$

It is easy to see that $\rho := \mathbb{P}(|\mathfrak{X}_\lambda| = \infty)$ satisfies $1 - \rho = \mathbb{E}(1-\rho)^{\mathrm{Po}(\lambda)} = e^{-\lambda\rho}$. Using Taylor series we infer $\lambda\rho = -\log(1-\rho) \geqslant \rho + \rho^2/2$, so that either $\rho = 0$ or $0 < \rho \leqslant 2(\lambda - 1)$, which together with (B.4) completes the proof of (B.3). ☐

Given a graph H, we write $C_H(v)$ for the component containing the vertex v in H, and $\mathrm{d}_H(v,w)$ for the length of the shortest path between v and w in H (setting $\mathrm{d}_H(v,w) = \infty$ if there is no such path). Define $B_H(v,r) = \{w \in V(H) : \mathrm{d}_H(v,w) \leqslant r\}$ and $\partial B_H(v,r) = \{w \in V(H) : \mathrm{d}_H(v,w) = r\}$.

Lemma B.3. *There is a constant $C > 0$ such that, for all $n \geqslant 1$, $p \in [0,1]$, and $\varepsilon > 0$ satisfying $np \leqslant 1 + \varepsilon$ and $\varepsilon n \geqslant 1$, the following holds for all $1 \leqslant r \leqslant \lceil \varepsilon^{-1} \rceil$:*

$$\Gamma(r) := \max_{G \subseteq K_n} \max_{v \in V(G)} \mathbb{P}(\partial B_{G_p}(v,r) \neq \emptyset) \leqslant Cr^{-1}. \tag{B.5}$$

Proof. Assuming $C \geqslant 9$, note that for $\varepsilon \geqslant 1/9$ inequality (B.5) holds trivially for all $1 \leqslant r \leqslant \lceil \varepsilon^{-1} \rceil$. It thus suffices to consider the case $\varepsilon \leqslant 1/9$. Let $\lambda := -n\log(1-p)$. Note that $p \leqslant (1+\varepsilon)/n \leqslant 2\varepsilon < 1/2$ and thus $\lambda \leqslant np(1+p) \leqslant 1 + 4\varepsilon$. Let $k_0 \geqslant 1$ satisfy $3^{k_0-1} \leqslant \lceil \varepsilon^{-1} \rceil < 3^{k_0}$.

It is well-known and easy to see that for any subgraph $G \subseteq K_n$, $|C_{G_p}(v)| \leqslant |C_{G_{n,p}}(v)|$ is stochastically dominated by $|\mathfrak{X}_{n,p}|$. By Lemma B.2 it follows that there exists a constant $B > 0$ such that, for all $1 \leqslant K \leqslant 9^{k_0}$,

$$\max_{G \subseteq K_n} \max_{v \in V(G)} \mathbb{P}(|C_{G_p}(v)| \geqslant K) \leqslant \mathbb{P}(|\mathfrak{X}_{n,p}| \geqslant K) \leqslant \mathbb{P}(|\mathfrak{X}_\lambda| \geqslant K) \leqslant \mathbb{P}(|\mathfrak{X}_{1+4\varepsilon}| \geqslant K) \leqslant BK^{-1/2}. \tag{B.6}$$

Mimicking [24, Section 3.2], we now show by induction on $k \in \mathbb{N}$ that $D := (3^3 + B)^3$ satisfies

$$\Gamma(3^k) \leqslant D3^{-k} \qquad \text{for all } 0 \leqslant k \leqslant k_0, \tag{B.7}$$

which readily implies (B.5) with $C = 3D$ (for any $1 \leqslant r \leqslant \lceil \varepsilon^{-1} \rceil$ there is $1 \leqslant k \leqslant k_0$ with $3^{k-1} \leqslant r < 3^k$, so $\Gamma(r) \leqslant \Gamma(3^{k-1}) \leqslant 3Dr^{-1}$ follows). The base case $k = 0$ holds trivially since since $\Gamma(1) \leqslant 1 \leqslant D$.

For the induction step, let $1 \leqslant k \leqslant k_0$ and assume that (B.7) holds for $k-1$. Fix $G \subseteq K_n$ and $v \in V(G)$. Set $\delta := D^{-4/3} \leqslant 1$. Then, by (B.6) we see that

$$\mathbb{P}\big(\partial B_{G_p}(v,3^k) \neq \emptyset\big) \leqslant \mathbb{P}\big(\partial B_{G_p}(v,3^k) \neq \emptyset \text{ and } |C_{G_p}(v)| < \delta 9^k\big) + B\delta^{-1/2}3^{-k}. \tag{B.8}$$

By the pigeonhole principle, if $\partial B_{G_p}(v,3^k) \neq \emptyset$ and $|C_{G_p}(v)| < \delta 9^k$, then at least one level j with $3^{k-1} \leqslant j \leqslant 2{\cdot}3^{k-1}$ satisfies $0 < |\partial B_{G_p}(v,j)| \leqslant \delta 3^{k+1}$; let J denote the smallest such level. (If no such j exists, let $J := \infty$.) Note that, for any given non-empty sets of vertices $W \subseteq [n]$ and $\partial W \subseteq W$, using a breadth-first-search neighbourhood exploration algorithm, we can determine whether $B_{G_p}(v,J) = W$ and $\partial B_{G_p}(v,J) = \partial W$

by testing the status (in G_p) only of edges with at least one endpoint in $W \setminus \partial W$. Furthermore, if this event holds, then this determines J and $J < \infty$. Consequently, if H is the induced subgraph of G with vertex set $[n] \setminus (W \setminus \partial W)$, then after conditioning on $B_{G_p}(v,J) = W$ and $\partial B_{G_p}(v,J) = \partial W$, the remaining random graph $G_p \cap H$ has the same distribution as the *unconditional* random graph H_p. Furthermore, by construction, the shortest path in G_p from ∂W to $\partial B_{G_p}(v,3^k)$ contains only edges in H, so $\partial B_{G_p}(v,3^k) \neq \emptyset$ implies the existence of a vertex $w \in \partial W$ with $\partial B_{G_p \cap H}(w, 3^k - J) \neq \emptyset$. Combining $|\partial B_{G_p}(v,J)| \leqslant \delta 3^{k+1}$ and $3^k - J \geqslant 3^{k-1}$ with $H \subseteq G \subseteq K_n$ and $\partial W \subseteq V(H)$, it follows that

$$\begin{aligned}\mathbb{P}\big(\partial B_{G_p}(v,3^k) \neq \emptyset \mid B_{G_p}(v,J) = W,\, \partial B_{G_p}(v,J) = \partial W\big) &\leqslant \sum_{w \in \partial W} \mathbb{P}\big(\partial B_{H_p}(w, 3^{k-1}) \neq \emptyset\big) \\ &\leqslant \delta 3^{k+1} \cdot \Gamma(3^{k-1}).\end{aligned} \tag{B.9}$$

Since the bound in (B.9) does not depend on W and ∂W, it follows that

$$\mathbb{P}\big(\partial B_{G_p}(v,3^k) \neq \emptyset \mid J < \infty\big) \leqslant \delta 3^{k+1} \Gamma(3^{k-1}).$$

Consequently, using the induction hypothesis (and recalling that $J \geqslant 3^{k-1}$),

$$\begin{aligned}\mathbb{P}\big(\partial B_{G_p}(v,3^k) \neq \emptyset \text{ and } |C_{G_p}(v)| < \delta 9^k\big) &\leqslant \mathbb{P}\big(\partial B_{G_p}(v,3^k) \neq \emptyset \text{ and } J < \infty\big) \\ &\leqslant \delta 3^{k+1} \Gamma(3^{k-1}) \cdot \mathbb{P}(J < \infty) \\ &\leqslant \delta 3^{k+1} \Gamma(3^{k-1}) \cdot \mathbb{P}\big(\partial B_{G_p}(v, 3^{k-1}) \neq \emptyset\big) \\ &\leqslant \delta 3^{k+1} \Gamma(3^{k-1})^2 \leqslant \delta 3^{k+1} (D 3^{1-k})^2.\end{aligned} \tag{B.10}$$

After inserting (B.10) into (B.8), by recalling $\delta = D^{-4/3}$ and $D = (3^3 + B)^3$ we infer

$$\mathbb{P}\big(\partial B_{G_p}(v,3^k) \neq \emptyset\big) \leqslant \big(3^3 \delta D^2 + B \delta^{-1/2}\big) 3^{-k} = (3^3 + B) D^{2/3} 3^{-k} = D 3^{-k},$$

completing the proof of the induction step (since $G \subseteq K_n$ and $v \in V(G)$ were arbitrary). □

Proof of Theorem B.1. (i): Since $|\mathfrak{X}_{n,p}|$ stochastically dominates $|C(v)|$, using $np \leqslant 1 - \varepsilon$ we infer

$$\mathbb{E}|C(v)| \leqslant \mathbb{E}|\mathfrak{X}_{n,p}| = \sum_{j \geqslant 0} (np)^j = (1 - np)^{-1} \leqslant \varepsilon^{-1}.$$

(ii): Suppose first that $\varepsilon \leqslant 1$ and $\varepsilon^3 n \geqslant 1$ (the upper bound $\varepsilon \leqslant 1$ conveniently ensures $\varepsilon \lceil \varepsilon^{-1} \rceil \leqslant 2$). We set $r := \lceil \varepsilon^{-1} \rceil$, and proceed by a case distinction similar to [16, Lemma 2.3]. Observe that

$$\mathbb{E}_{n,p}|C(v)| = \sum_{w \in [n]} \mathbb{P}_{n,p}(w \in C(v)) = \sum_{w \in [n]} \big[\mathbb{P}(\mathrm{d}_{G_{n,p}}(v,w) \leqslant 2r) + \mathbb{P}(2r < \mathrm{d}_{G_{n,p}}(v,w) < \infty)\big]. \tag{B.11}$$

Since $|B_{\mathfrak{X}_{n,p}}(v,r)|$ stochastically dominates $|B_{G_{n,p}}(v,r)|$, using $np \leqslant 1 + \varepsilon$ and $\varepsilon \leqslant 1$ we deduce

$$\begin{aligned}\sum_{w \in [n]} \mathbb{P}\big(\mathrm{d}_{G_{n,p}}(v,w) \leqslant 2r\big) &= \mathbb{E}|B_{G_{n,p}}(v,2r)| \leqslant \mathbb{E}|B_{\mathfrak{X}_{n,p}}(v,2r)| \\ &\leqslant \sum_{0 \leqslant j \leqslant 2r} (1+\varepsilon)^j \leqslant \varepsilon^{-1}(1+\varepsilon)^{2r+1} \leqslant \varepsilon^{-1} e^5.\end{aligned} \tag{B.12}$$

Note that $2r < \mathrm{d}_{G_{n,p}}(v,w) < \infty$ implies $B_{G_{n,p}}(v,r) \cap B_{G_{n,p}}(w,r) = \emptyset$ and $\partial B_{G_{n,p}}(v,r), \partial B_{G_{n,p}}(w,r) \neq \emptyset$. By conditioning on $B_{G_{n,p}}(v,r)$, and letting $H := G_{n,p} \setminus B_{G_{n,p}}(v,r)$, it follows that, similarly to (B.9),

$$\mathbb{P}\big(2r < \mathrm{d}_{G_{n,p}}(v,w) < \infty \mid B_{G_{n,p}}(v,r)\big) \leqslant \mathbb{P}(\partial B_H(w,r) \neq \emptyset) \mathbb{1}_{\{\partial B_{G_{n,p}}(v,r) \neq \emptyset\}} \leqslant \Gamma(r) \mathbb{1}_{\{\partial B_{G_{n,p}}(v,r) \neq \emptyset\}}$$

and consequently by taking the expectation and using Lemma B.3,

$$\mathbb{P}\big(2r < \mathrm{d}_{G_{n,p}}(v,w) < \infty\big) \leqslant \Gamma(r) \cdot \mathbb{P}\big(\partial B_{G_{n,p}}(v,r) \neq \emptyset\big) \leqslant \Gamma(r)^2 \leqslant (C r^{-1})^2 \leqslant C^2 \varepsilon^2. \tag{B.13}$$

By (B.11)–(B.13), and $\varepsilon^3 n \geqslant 1$, there thus is a constant $D = D(C)$ such that

$$\mathbb{E}_{n,p}|C(v)| \leqslant e^5 \varepsilon^{-1} + C^2 \varepsilon^2 n \leqslant D \varepsilon^2 n. \tag{B.14}$$

This proves (B.2) when $\varepsilon \leqslant 1$ and $\varepsilon^3 n \geqslant 1$. When $\varepsilon \geqslant 1$, the bound in (B.2) holds trivially (since $|C(v)| \leqslant n$), assuming as we may $D \geqslant 1$.

In the remaining case $\varepsilon^3 n < 1$, we observe that $np \leqslant 1 + \varepsilon \leqslant 1 + n^{-1/3}$. Hence (B.14) with $\varepsilon := n^{-1/3}$ implies $\mathbb{E}_{n,p}|C(v)| \leqslant D n^{1/3}$, completing the proof of (B.2). □

C Higher derivatives of the susceptibility

In this appendix we extend the method of proof from Section 3.1 to higher derivatives, using arguments from [22]. The key fact is that, extending (3.15), any mixed moment of $X_{n,k}$ defined in (3.13), $k \geqslant 2$, has a derivative that can be expressed as a linear combination of such moments, and thus by induction the same holds for higher derivatives as well. We illustrate the general method by some examples, leaving the details in the general case to the reader. For notational convenience, we write

$$\mathrm{D_{n,p}} := n^{-4/3}(1-p)\frac{\mathrm{d}}{\mathrm{d}p}.$$

Note that $\mathrm{D_{n,p}} = n^{-4/3}\frac{\mathrm{d}}{\mathrm{d}t}$ for the parametrization $p = 1 - e^{-t}$ used in [22]. Note also that the factor $1-p$, which is needed in the exact formulas below, disappear asymptotically, since $p = o(1)$, and that apart from this factor, $\mathrm{D_{n,p}} = \frac{\mathrm{d}}{\mathrm{d}\lambda}$ for our usual parametrization $p = n^{-1} + \lambda n^{-4/3}$.

First, consider $\mathrm{D_{n,p}}\bigl(\mathbb{E}_{n,p}(X_{n,k})\bigr)$ for an arbitrary $k \geqslant 2$. As noted in [22, (3.1)], if $v \not\leftrightarrow w$, then adding the edge vw to the graph increases $\sum_i |C_i|^k$ by

$$\Delta_{vw}\Bigl(\sum_i |C_i|^k\Bigr) := (|C(v)| + |C(w)|)^k - |C(v)|^k - |C(w)|^k = \sum_{\ell=1}^{k-1}\binom{k}{\ell}|C(v)|^\ell |C(w)|^{k-\ell}. \tag{C.1}$$

(And, trivially, the change $\Delta_{vw}(\sum_i |C_i|^k) = 0$ if $v \leftrightarrow w$.) Recalling $X_{n,k} = n^{-2k/3}\sum_i |C_i|^k$, see (3.13), it follows by a modification of the argument leading to (2.11) that (similar to [13, Theorem 2.32]), with a factor $\frac12$ because each edge is counted twice,

$$\begin{aligned}
\mathrm{D_{n,p}}\bigl(\mathbb{E}_{n,p}(X_{n,k})\bigr) &= \tfrac12 n^{-2(k+2)/3}\sum_{v,w\in[n]}\mathbb{E}_{n,p}\left(\mathbb{1}_{\{vw\notin G_{n,p}\}}\Delta_{vw}\Bigl(\sum_i |C_i|^k\Bigr)\right)\\
&= \tfrac12 n^{-2(k+2)/3}\mathbb{E}_{n,p}\left(\sum_{\ell=1}^{k-1}\binom{k}{\ell}\sum_{i\neq j}|C_i|^{\ell+1}|C_j|^{k-\ell+1}\right)\\
&= \tfrac12\sum_{\ell=1}^{k-1}\binom{k}{\ell}\mathbb{E}_{n,p}\bigl(X_{n,\ell+1}X_{n,k-\ell+1}\bigr) - (2^{k-1}-1)\mathbb{E}_{n,p}(X_{n,k+2}).
\end{aligned}\tag{C.2}$$

Theorem 3.2(ii) extends to mixed moments, because the convergence (3.8) holds jointly for different $k \geqslant 2$ (by the same proof), and the uniform moment bound (3.9) extends to mixed moments by Hölder's inequality. Thus we obtain from (C.2), if $p = 1/n + \bigl(\lambda + o(1)\bigr)n^{-4/3}$,

$$\mathrm{D_{n,p}}\bigl(\mathbb{E}_{n,p}(X_{n,k})\bigr) \to \tfrac12\sum_{\ell=1}^{k-1}\binom{k}{\ell}\mathbb{E}\bigl(W_{\lambda,\ell+1}W_{\lambda,k-\ell+1}\bigr) - (2^{k-1}-1)\mathbb{E}(W_{\lambda,k+2}). \tag{C.3}$$

The special case $k = 2$ is given above in (3.15).

For higher moments, we give for notational convenience just one example of the method. Adding an edge vw with $v \not\leftrightarrow w$ increases $X_{n,2} = n^{-4/3}\sum_i |C_i|^2$ by $\Delta(X_{n,2}) = 2n^{-4/3}|C(v)||C(w)|$, see (C.1), and thus increases $X_{n,2}^2$ by

$$\begin{aligned}
\Delta\bigl(X_{n,2}^2\bigr) &= \bigl(X_{n,2} + \Delta X_{n,2}\bigr)^2 - X_{n,2}^2 = 2X_{n,2}\Delta X_{n,2} + (\Delta X_{n,2})^2\\
&= 4n^{-4/3}X_{n,2}|C(v)||C(w)| + 4n^{-8/3}|C(v)|^2|C(w)|^2,
\end{aligned}\tag{C.4}$$

leading to

$$\begin{aligned}
\mathrm{D_{n,p}}\bigl(\mathbb{E}_{n,p}(X_{n,2}^2)\bigr) &= \tfrac12\mathbb{E}_{n,p}\Bigl(4n^{-8/3}X_{n,2}\sum_{i\neq j}|C_i|^2|C_j|^2 + 4n^{-12/3}\sum_{i\neq j}|C_i|^3|C_j|^3\Bigr)\\
&= 2\mathbb{E}_{n,p}\bigl(X_{n,2}\bigl(X_{n,2}^2 - X_{n,4}\bigr)\bigr) + 2\mathbb{E}_{n,p}\bigl(X_{n,3}^2 - X_{n,6}\bigr)\\
&= 2\mathbb{E}_{n,p}\bigl(X_{n,2}^3\bigr) - 2\mathbb{E}_{n,p}\bigl(X_{n,4}X_{n,2}\bigr) + 2\mathbb{E}_{n,p}\bigl(X_{n,3}^2\bigr) - 2\mathbb{E}_{n,p}\bigl(X_{n,6}\bigr).
\end{aligned}\tag{C.5}$$

This together with the cases $k = 2$ and $k = 4$ of (C.2) yield, after simplifications,

$$\begin{aligned}(\mathrm{D_{n,p}})^2\Big(\mathbb{E}_{n,p}\big(X_{n,2}\big)\Big) &= \mathrm{D_{n,p}}\Big(\mathbb{E}_{n,p}\big(X_{n,2}^2\big) - \mathbb{E}_{n,p}\big(X_{n,4}\big)\Big) \\ &= 2\mathbb{E}_{n,p}\big(X_{n,2}^3\big) - 6\mathbb{E}_{n,p}\big(X_{n,4}X_{n,2}\big) - \mathbb{E}_{n,p}\big(X_{n,3}^2\big) + 5\mathbb{E}_{n,p}\big(X_{n,6}\big)\end{aligned} \tag{C.6}$$

and thus we obtain, if $p = 1/n + \big(\lambda + o(1)\big)n^{-4/3}$,

$$(\mathrm{D_{n,p}})^2\Big(\mathbb{E}_{n,p}\big(X_{n,2}\big)\Big) \to 2\mathbb{E}\big(W_{\lambda,2}^3\big) - 6\mathbb{E}\big(W_{\lambda,4}W_{\lambda,2}\big) - \mathbb{E}\big(W_{\lambda,3}^2\big) + 5\mathbb{E}\big(W_{\lambda,6}\big). \tag{C.7}$$

The general case is similar. In particular, this leads to the following extension of Theorem 1.3.

Theorem C.1. *Define the infinitely differentiable function $f = f_2 : \mathbb{R} \to (0,\infty)$ as in (3.4). Given $\lambda \in \mathbb{R}$, for $p = 1/n + \big(\lambda + o(1)\big)n^{-4/3}$ we have, as $n \to \infty$, for every fixed m,*

$$n^{-(4m+1)/3}\frac{\mathrm{d}^m}{\mathrm{d}p^m}\chi_{K_n}(p) \to \frac{\mathrm{d}^m}{\mathrm{d}\lambda^m}f(\lambda), \qquad (m \geqslant 0), \tag{C.8}$$

$$n^{-4m/3}\frac{\mathrm{d}^m}{\mathrm{d}p^m}\log\chi_{K_n}(p) \to \frac{\mathrm{d}^m}{\mathrm{d}\lambda^m}\log f(\lambda) \qquad (m \geqslant 1). \tag{C.9}$$

Moreover, if $p = 1/n + \lambda n^{-4/3}$, then the convergence is uniform for λ in any compact set $[\lambda_1, \lambda_2] \subset \mathbb{R}$.

Proof. For $p = 1/n + \big(\lambda + o(1)\big)n^{-4/3}$, the argument above shows that for every $m \geqslant 0$,

$$(\mathrm{D_{n,p}})^m\Big(\mathbb{E}_{n,p}\big(X_{n,2}\big)\Big) \to g_m(\lambda) \tag{C.10}$$

for some function $g_m(\lambda)$, with

$$g_0(\lambda) = f(\lambda). \tag{C.11}$$

Recalling the definition of $\mathrm{D_{n,p}}$, using (C.10), the product rule, and induction, it is easy to see that for every $m \geqslant 0$ there are constants $c_{j,m} \in \mathbb{R}$ with $c_{m,m} = 1$ such that

$$(\mathrm{D_{n,p}})^m\Big(\mathbb{E}_{n,p}\big(X_{n,2}\big)\Big) = \sum_{0\leqslant j\leqslant m} c_{j,m}(1-p)^j n^{-4m/3}\frac{\mathrm{d}^j}{\mathrm{d}p^j}\Big(\mathbb{E}_{n,p}\big(X_{n,2}\big)\Big). \tag{C.12}$$

Combining (C.10) and (C.12) with $p = o(1)$, by another induction on $m \geqslant 0$ we now infer

$$n^{-4m/3}\frac{\mathrm{d}^m}{\mathrm{d}p^m}\Big(\mathbb{E}_{n,p}\big(X_{n,2}\big)\Big) \to g_m(\lambda), \tag{C.13}$$

since in (C.12) any summand with $j < m$ is $O(n^{-4m/3} \cdot n^{4j/3}) = o(1)$ by the induction hypothesis.

We now change parametrization and define $h_n(\lambda) := \mathbb{E}_{n,p}\big(X_{n,2}\big)$ for $p = 1/n + \lambda n^{-4/3}$, so that (C.13) translates into

$$\frac{\mathrm{d}^m}{\mathrm{d}\lambda^m}h_n(\lambda) = n^{-4m/3}\frac{\mathrm{d}^m}{\mathrm{d}p^m}h_n(\lambda) \to g_m(\lambda). \tag{C.14}$$

Lemma 3.3 shows that $g_m(\lambda)$ is continuous, and that (C.14) holds uniformly for λ in any compact set. Hence we can integrate, as in (3.16)–(3.17), and obtain $g_m(\lambda_2) - g_m(\lambda_1) = \int_{\lambda_1}^{\lambda_2} g_{m+1}(\lambda)\,\mathrm{d}\lambda$ whenever $\lambda_1 < \lambda_2$, and thus $g_{m+1}(\lambda) = \frac{\mathrm{d}}{\mathrm{d}\lambda}g_m(\lambda)$. By induction and (C.11), we infer, for every $m \geqslant 0$,

$$g_m(\lambda) = \frac{\mathrm{d}^m}{\mathrm{d}\lambda^m}g_0(\lambda) = \frac{\mathrm{d}^m}{\mathrm{d}\lambda^m}f(\lambda). \tag{C.15}$$

By (C.14) and (C.15),

$$\frac{\mathrm{d}^m}{\mathrm{d}\lambda^m}h_n(\lambda) \to g_m(\lambda) = \frac{\mathrm{d}^m}{\mathrm{d}\lambda^m}f(\lambda), \qquad m \geqslant 0, \tag{C.16}$$

and thus also, by expanding the derivatives of the logarithms on both sides and applying (C.16) to each term,

$$\frac{\mathrm{d}^m}{\mathrm{d}\lambda^m}\log h_n(\lambda) \to \frac{\mathrm{d}^m}{\mathrm{d}\lambda^m}\log f(\lambda), \qquad m \geqslant 1. \tag{C.17}$$

Moreover, the convergence in (C.16) and (C.17) is uniform on any compact set. Consequently, (C.16)–(C.17) hold also with $h_n(\lambda_n)$ on the left-hand side, for any sequence $\lambda_n \to \lambda$, see Remark 3.4. The result (C.8)–(C.9) now follows for any $p = 1/n + \big(\lambda + o(1)\big)n^{-4/3}$ by taking $\lambda_n := n^{4/3}(p - 1/n) = \lambda + o(1)$. ∎

References

[1] M. Aizenman and C. M. Newman. Tree graph inequalities and critical behavior in percolation models. *J. Statist. Phys.* **36** (1984), 107–143.

[2] M. Ajtai, J. Komlós and E. Szemerédi. Largest random component of a k-cube. *Combinatorica* **2** (1982), 1–7.

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

[4] B. Bollobás. The evolution of random graphs. *Trans. Amer. Math. Soc.* **286** (1984), 257–274.

[5] B. Bollobás, Y. Kohayakawa and T. Łuczak. The evolution of random subgraphs of the cube. *Rand. Struct. & Algor.* **3** (1992), 55–90.

[6] B. Bollobás and O. Riordan. Exploring hypergraphs with martingales. *Rand. Struct. & Algor.*, to appear. `arXiv:1403.6558`

[7] C. Borgs, J.T. Chayes, R. van der Hofstad, G. Slade and J. Spencer. Random subgraphs of finite graphs. I. The scaling window under the triangle condition. *Rand. Struct. & Algor.* **27** (2005), 137–184.

[8] C. Borgs, J. T. Chayes, R. van der Hofstad, G. Slade and J. Spencer. Random subgraphs of finite graphs. II. The lace expansion and the triangle condition. *Ann. Probab.* **33** (2005), 1886–1944.

[9] C. Borgs, J. T. Chayes, R. van der Hofstad, G. Slade and J. Spencer. Random subgraphs of finite graphs. III. The phase transition for the n-cube. *Combinatorica* **26** (2006), 395–410.

[10] M. Dwass. The total progeny in a branching process and a related random walk. *J. Appl. Probability* **6** (1969), 682–686.

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

[12] P. Erdős and A. Rényi. On the strength of connectedness of a random graph. *Acta Math. Acad. Sci. Hungar.* **12** (1961), 261–267.

[13] G. Grimmett. *Percolation.* Second edition, Springer, Berlin (1999).

[14] A. Gut. *Probability: a Graduate Course.* Springer, New York (2005).

[15] R. van der Hofstad and M. J. Luczak. Random subgraphs of the 2D Hamming graph: the supercritical phase. *Probab. Theory & Related Fields* **147** (2010), 1–41.

[16] R. van der Hofstad and A. Nachmias. Hypercube percolation. *J. Eur. Math. Soc.*, to appear. `arXiv:1201.3953`

[17] R. van der Hofstad and A. Nachmias. Unlacing hypercube percolation: a survey. *Metrika* **77** (2014), 23–50.

[18] R. van der Hofstad and G. Slade. Expansion in n^{-1} for percolation critical values on the n-cube and $\mathbb{Z}^n$: the first three terms. *Combin. Probab. Comput.* **15** (2006), 695–713.

[19] S. Janson. Brownian excursion area, Wright's constants in graph enumeration, and other Brownian areas. *Probab. Surv.* (2007), 80–145.

[20] S. Janson and G. Louchard. Tail estimates for the Brownian excursion area and other Brownian areas. *Electronic J. Probab.* **12** (2007), no. 58, 1600–1632.

[21] S. Janson, T. Łuczak and A. Ruciński. *Random Graphs.* Wiley-Interscience, New York (2000).

[22] S. Janson and M. J. Luczak. Susceptibility in subcritical random graphs. *J. Math. Phys.* **49** (2008), 125207.

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

[24] G. Kozma and A. Nachmias. The Alexander–Orbach conjecture holds in high dimensions. *Invent. Math.* **178** (2009), 635–654.

[25] Y. Long, A. Nachmias, W. Ning and Y. Peres. A power law of order 1/4 for critical mean field Swendsen–Wang dynamics. *Mem. Amer. Math. Soc.* **232** (2014).

[26] G. Louchard. The Brownian excursion area: a numerical analysis. *Comput. Math. Appl.* **10** (1984), 413–417.

[27] T. Łuczak. Component behavior near the critical point of the random graph process. *Rand. Struct. & Algor.* **1** (1990), 287–310.

[28] G. A. Margulis. Probabilistic characteristics of graphs with large connectivity. (Russian.) *Problemy Peredači Informacii* **10** (1974), 101–108.

[29] A. Nachmias and Y. Peres. Critical random graphs: diameter and mixing time. *Ann. Probab.* **36** (2008), 1267–1286.

[30] *NIST Handbook of Mathematical Functions.* Eds. F. W. J. Olver, D. W. Lozier, R. F. Boisvert and C. W. Clark. Cambridge University Press (2010). Also available as *NIST Digital Library of Mathematical Functions*, http://dlmf.nist.gov/.

[31] Y. Peres. Personal communication (2012).

[32] O. Riordan and L. Warnke. The evolution of subcritical Achlioptas processes. *Rand. Struct. & Algor.* **47** (2015), 174–203.

[33] O. Riordan and L. Warnke. The phase transition in bounded-size Achlioptas processes. In preparation.

[34] L. Russo. On the critical percolation probabilities. *Z. Wahrsch. Verw. Gebiete* **56** (1981), 229–237.

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

[36] L. Warnke. Percolation thoughts. MSR-Internship Report (2012).

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