SUSCEPTIBILITY OF RANDOM GRAPHS WITH GIVEN VERTEX DEGREES

SVANTE JANSON

Abstract. We study the susceptibility, i.e., the mean cluster size, in random graphs with given vertex degrees. We show, under weak assumptions, that the susceptibility converges to the expected cluster size in the corresponding branching process. In the supercritical case, a corresponding result holds for the modified susceptibility ignoring the giant component and the expected size of a finite cluster in the branching process; this is proved using a duality theorem.

The critical behaviour is studied. Examples are given where the critical exponents differ on the subcritical and supercritical sides.

1. Introduction

The *susceptibility* $\chi(G)$ of a graph G is defined as the mean size of the component containing a random vertex:

$$\chi(G) := |G|^{-1} \sum_{v \in V(G)} |\mathcal{C}(v)|, \tag{1.1}$$

where $\mathcal{C}(v)$ denotes the component of G containing the vertex v. Thus, if G has $n = |G|$ vertices and components $\mathcal{C}_i = \mathcal{C}_i(G)$, $i = 1, \dots, K$, where $K = K(G)$ is the number of components, then

$$\chi(G) := \sum_{i=1}^{K} \frac{|\mathcal{C}_i|}{n} |\mathcal{C}_i| = \frac{1}{n} \sum_{i=1}^{K} |\mathcal{C}_i|^2. \tag{1.2}$$

Although it does not matter here, we assume for later use that the components as usual are ordered with $|\mathcal{C}_1| \geq |\mathcal{C}_2| \geq \cdots$.

When the graph G is itself random, $\chi(G)$ is thus a random variable. (We do not take the expectation over G unless we explicitly write $\mathbb{E}\,\chi(G)$.)

The susceptibility (in particular its expectation) has been much studied for certain models in mathematical physics. (That is the reason for using the term susceptibility, and the notation χ, which both come from physics.) Similarly, in percolation theory, which deals with certain random infinite graphs, the corresponding quantity is the (mean) size of the open cluster containing a given vertex, and this has been extensively studied; see e.g. Bollobás and Riordan [8]. For finite random graphs, there are some papers:

Date: November 13, 2009; typo corrected 24 December 2010.

2000 *Mathematics Subject Classification.* 05C80; 60C05.

Spencer and Wormald [34] studied in a pioneering paper a class of random graph processes (including the Erdős–Rényi graph process) and used the susceptibility to study the phase transition in them. Some results for the Erdős–Rényi random graphs $G(N,p)$ and $G(n,m)$ can be regarded as folk theorems; detailed results are given by Durrett [13, Section 2.2] and Janson and Luczak [22]. Borgs, Chayes, van der Hofstad, Slade and Spencer [9] give precise results for random subgraphs of transitive graphs (including both $G(N,p)$ and, for example, random subgraphs of the hypercube); further results for random subgraphs of the hypercube are given by van der Hofstad and Slade [16, 17]. A class of inhomogeneous random graphs is studied by Janson and Riordan [26], see also Chayes and Smith [12]. Another application is given in Janson and Spencer [24]. We refer to these papers for further background. The purpose of the present paper is to study the susceptibility for the random graph $G(n,\mathbf{d})$ with given vertex degrees, where $\mathbf{d}=(d_i)_1^n$ is a given degree sequence (see Section 2 for a detailed definition). This case has earlier been studied in a heuristic way by Newman, Strogatz and Watts [33], using the branching process in Section 3 below.

The definition (1.2) is mainly interesting in the subcritical case, when all components are rather small. In the supercritical case, see Molloy and Reed [31] or Theorem 2.4 below, there is one giant component that is so large that it dominates the sum in (1.2); in fact, for some $\rho>0$, $|\mathcal{C}_1|=(\rho+o_{\mathrm{p}}(1))n$ while $|\mathcal{C}_2|=o_{\mathrm{p}}(n)$ and thus

$$\sum_{i=1}^K|\mathcal{C}_i|^2=|\mathcal{C}_1|^2+O\Bigl(|\mathcal{C}_2|\sum_{i=2}^K|\mathcal{C}_i|\Bigr)=(\rho^2+o_{\mathrm{p}}(1))n^2=(1+o_{\mathrm{p}}(1))|\mathcal{C}_1|^2.$$

It then makes sense to exclude the largest component from the definition, and we define as in [26] the *modified susceptibility* $\widehat{\chi}(G)$ of a finite graph G by

$$\widehat{\chi}(G):=\frac{1}{n}\sum_{i=2}^K|\mathcal{C}_i|^2. \tag{1.3}$$

(This is in analogy with percolation theory, where one studies the mean size of the open cluster containing a given vertex, given that this cluster is finite.)

Our main result is the following, giving the asymptotics of both χ and $\widehat{\chi}$ for $G(n,\mathbf{d})$, using notation introduced in Section 2 below,

Theorem 1.1. *Suppose that Conditions 2.1 and 2.3 hold.*

(i) *In the subcritical case $\nu_\infty<\mu_\infty$,*

$$\chi(G(n,\mathbf{d}))\xrightarrow{L^1}\chi_\infty:=1+\frac{\mu_\infty^2}{\mu_\infty-\nu_\infty},$$
$$\widehat{\chi}(G(n,\mathbf{d}))\xrightarrow{L^1}\chi_\infty.$$

(ii) *In the critical case* $\nu_\infty = \mu_\infty$,

$$\chi(G(n,\mathbf{d})), \widehat{\chi}(G(n,\mathbf{d})) \xrightarrow{\mathrm{p}} \infty.$$

(iii) *In the supercritical case* $\nu_\infty > \mu_\infty$, *with* $\varkappa \in (0,1)$ *given by* $g'(\varkappa) = \varkappa g'(1)$,

$$\chi(G(n,\mathbf{d})) \xrightarrow{\mathrm{p}} \infty,$$
$$\widehat{\chi}(G(n,\mathbf{d})) \xrightarrow{\mathrm{p}} \widehat{\chi}_\infty := g(\varkappa) + \frac{\varkappa(g'(\varkappa))^2}{g'(\varkappa) - \varkappa g''(\varkappa)} < \infty.$$

Note that the L^1-convergence in the subcritical case (i) entails both $\chi(G(n,\mathbf{d})) \xrightarrow{\mathrm{p}} \chi_\infty$ and $\mathbb{E}\,\chi(G(n,\mathbf{d})) \to \chi_\infty$. Hence, in all three cases we have

$$\chi(G(n,\mathbf{d})) \xrightarrow{\mathrm{p}} \chi_\infty := 1 + \frac{\mu_\infty^2}{(\mu_\infty - \nu_\infty)_+} \le \infty,$$
$$\widehat{\chi}(G(n,\mathbf{d})) \xrightarrow{\mathrm{p}} \widehat{\chi}_\infty := g(\varkappa)\Big(1 + \frac{\widehat{\mu}_\infty^2}{(\widehat{\mu}_\infty - \widehat{\nu}_\infty)_+}\Big) \le \infty.$$

Further, in (ii) and (iii) it follows trivially that $\mathbb{E}\,\chi(G(n,\mathbf{d})) \to \infty$. Hence also $\mathbb{E}\,\chi(G(n,\mathbf{d})) \to \chi_\infty$ holds in all three cases. However, our proof does not (at least not immediately) show convergence of $\mathbb{E}\,\widehat{\chi}(G(n,\mathbf{d}))$ in the supercritical case (iii), although we conjecture that it holds there too.

The results are based on approximation by a branching process $\mathfrak{X}$, see Section 3, as is standard when studying the component structure in both $G(n,\mathbf{d})$ and in several other random graph models (see e.g. [23] and [7]). Theorem 1.1 can be seen as saying that (under some weak conditions), the susceptibility χ and the modified susceptibility $\widehat{\chi}$ of $G(n,\mathbf{d})$ converge to the corresponding mean values for the branching process corresponding to $G(n,\mathbf{d})$; see Theorem 3.3 for details. Proofs are given in Sections 4–7.

The proof of our result for $\widehat{\chi}$ is based on a duality result, Theorem 6.1, saying that if we delete the largest component $\mathcal{C}_1$ from a supercritical $G(n,\mathbf{d})$, then the remainder is essentially another random graph of the same type, which furthermore is subcritical. (The size and vertex degrees are random, but this is not important since they are concentrated.) This was proved already by Molloy and Reed [32], but we need a slightly sharper form here. Such duality results for $G(N,p)$ go back to Bollobás [5], see also Łuczak [28], Janson, Knuth, Łuczak and Pittel [20] and the books Bollobás [6], Janson, Łuczak and Ruciński [23]; a generalization to a class of inhomogeneous random graphs is given by Bollobás, Janson and Riordan [7, Theorem 12.1] and a further generalization by Janson and Riordan [25].

Theorem 1.1 is stated as a limit result. An alternative is to formulate the result as an approximation for finite n; this version is given in Section 8. We end with some further comments. The behaviour close to criticality is studied in Section 9 for a specific situation. We show that there is symmetry between the subcritical and supercritical sides when the asymptotic degree distribution has a third moment, but not necessarily in general; the critical exponent on the subcritical side is always 1 but on the supercritical side it may be arbitrarily large. Finally, in Section 10 we give some examples showing that the main theorems may fail without our conditions.

Acknowledgement. Parts of this research has been done during visits to Centre de recherches mathématiques, Montreal (Canada), Institut Mittag-Leffler, Djursholm (Sweden) and Institut Henri Poincaré, Paris (France).

2. Preliminaries

Let $n \in \mathbb{N}$ and let $\mathbf{d} = (d_i)_{i=1}^n$ be a sequence of non-negative integers. We let $G(n, \mathbf{d})$ be a random graph with degree sequence $(d_i)_1^n$, uniformly chosen among all possibilities (tacitly assuming that there is any such graph at all; in particular, $\sum_i d_i$ has to be even).

As in many papers on these random graphs, we find it convenient to consider the corresponding random *multigraphs* generated by the *configuration model* (see Bollobás [3] and [6, Section II.4]; see also Bender and Canfield [2] and Wormald [35; 36] for related arguments): Let $n \in \mathbb{N}$ and let $(d_i)_1^n$ be a sequence of non-negative integers such that $\sum_{i=1}^n d_i$ is even. Then take a set of d_i *half-edges* for each vertex i, and combine the half-edges into pairs by a uniformly random matching of the set of all half-edges (allowing multiple edges and loops); this yields the random multigraph $G^*(n, \mathbf{d})$ with given degree sequence $(d_i)_1^n$. Conditioned on the multigraph being a (simple) graph, we obtain $G(n, \mathbf{d})$, the uniformly distributed random graph with the given degree sequence.

We assume throughout the paper that we are given a sequence $(d_i)_1^n = (d_i^{(n)})_1^n$ for each $n \in \mathbb{N}$ (or at least for some sequence $n \to \infty$); for notational simplicity we will usually not show the dependence on n explicitly for these and some other quantities. We consider asymptotics as $n \to \infty$, and all unspecified limits below are as $n \to \infty$. We say that an event holds w.h.p. (*with high probability*), if it holds with probability tending to 1 as $n \to \infty$. We use standard probabilistic notations for convergence; in particular $\xrightarrow{\mathrm{p}}$ and $\xrightarrow{\mathrm{d}}$ for convergence in probability and in distribution, and o_{p} in the standard way (see e.g. [23] or [19]): for example, if (X_n) is a sequence of random variables, then $X_n = o_{\mathrm{p}}(1)$ means that $X_n \xrightarrow{\mathrm{p}} 0$.

We write

$$n_k = n_k(n) := \#\{i : d_i = k\}, \quad k \geq 0,$$

and

$$m = m(n) := \tfrac{1}{2} \sum_{i=1}^{n} d_i = \tfrac{1}{2} \sum_{k=0}^{\infty} k n_k;$$

thus n_k is the number of vertices of degree k and m is the number of edges in the random graph $G(n, \mathbf{d})$ (or $G^*(n, \mathbf{d})$). We assume as in [21] that the given $(d_i)_1^n$ satisfy the following regularity conditions, cf. Molloy and Reed [31; 32] (where similar but not identical conditions are assumed).

Condition 2.1. *For each n, $(d_i)_1^n = (d_i^{(n)})_1^n$ is a sequence of non-negative integers such that $\sum_{i=1}^n d_i$ is even. Furthermore, $(p_k)_{k=0}^\infty$ is a probability distribution independent of n such that*

(i) $n_k/n = \#\{i : d_i = k\}/n \to p_k$ *as $n \to \infty$, for every $k \geq 0$;*
(ii) $\sum_k kp_k \in (0, \infty)$*;*
(iii) $\sum_i d_i^2 = O(n)$*;*
(iv) $p_1 > 0$.

Let D_n be a random variable defined as d_I for a uniform random index $I \in \{1, \dots, n\}$: thus D_n is the degree of a random (uniformly chosen) vertex in $G(n, \mathbf{d})$ or $G^*(n, \mathbf{d})$, and

$$\mathbb{P}(D_n = k) = n_k/n. \tag{2.1}$$

Define

$$\mu_n := \mathbb{E}\, D_n = \frac{1}{n}\sum_{i=1}^n d_i = \frac{2m}{n}, \tag{2.2}$$

$$\nu_n := \mathbb{E}\, D_n(D_n - 1) = \frac{1}{n}\sum_{i=1}^n d_i(d_i - 1). \tag{2.3}$$

Further, let D_∞ be a random variable with the distribution $\mathbb{P}(D_\infty = k) = p_k$, and extend (2.2) and (2.3) to $\mu_\infty := \mathbb{E}\, D_\infty$ and $\nu_\infty := \mathbb{E}\, D_\infty(D_\infty - 1)$. Then Condition 2.1(i) can be written

$$D_n \xrightarrow{\mathrm{d}} D_\infty. \tag{2.4}$$

In other words, D_∞ describes the asymptotic distribution of the degree of a random vertex in $G(n, \mathbf{d})$. Furthermore, (ii) is $\mu_\infty = \mathbb{E}\, D_\infty \in (0, \infty)$, (iv) is $\mathbb{P}(D_\infty = 1) > 0$, and (iii) can be written

$$\mathbb{E}\, D_n^2 = O(1) \tag{2.5}$$

or, equivalently, $\nu_n = O(1)$.

Remark 2.2. In particular, (2.5) implies that the random variables D_n are uniformly integrable, and thus Condition 2.1(i), in the form (2.4), implies $\mathbb{E}\, D_n \to \mathbb{E}\, D_\infty$, i.e.

$$\mu_n = \frac{2m}{n} = \frac{1}{n}\sum_{i=1}^n d_i \to \mu_\infty, \tag{2.6}$$

see e.g. [14, Theorems 5.4.2 and 5.5.9].

We will often need an assumption that is a little stronger than Condition 2.1(iii).

Condition 2.3. *As $n \to \infty$, $\nu_n \to \nu_\infty$. (Equivalently, $\mathbb{E}\, D_n^2 \to \mathbb{E}\, D_\infty^2$.)*

This is clearly stronger than Condition 2.1(iii), see (2.5). Assuming Condition 2.1, it is by (2.4) equivalent to uniform integrability of D_n^2, cf. Remark 2.2. In particular, Condition 2.3 holds if $\sup_n \mathbb{E}\, D_n^{2+\varepsilon} < \infty$ for some $\varepsilon > 0$.

Let

$$g(x) := \mathbb{E}\, x^{D_\infty} = \sum_{k=0}^{\infty} p_k x^k,$$

the probability generating function of the probability distribution $(p_k)_{k=0}^{\infty}$. Thus $\mu_\infty = g'(1)$ and $\nu_\infty = g''(1)$.

We shall use the result by Molloy and Reed [31, 32] on existence and size of a giant component in $G(n, \mathbf{d})$; we state it in a version from [21]. For a graph G, let $v_k(G)$ be the number of vertices of degree k, $k \geq 0$.

Theorem 2.4 (Molloy and Reed)**.** *Suppose that Condition 2.1 holds. Consider the random graph $G(n, \mathbf{d})$ and let $\mathcal{C}_1$ and $\mathcal{C}_2$ be its largest and second largest components.*

(i) *If $\nu_\infty - \mu_\infty = \mathbb{E}\, D_\infty(D_\infty - 2) > 0$, then there is a unique $\varkappa \in (0, 1)$ such that $g'(\varkappa) = \mu_\infty \varkappa$. With this $\varkappa$, as $n \to \infty$,*

$$|\mathcal{C}_1|/n \xrightarrow{\text{p}} 1 - g(\varkappa) > 0,$$

and

$$v_k(\mathcal{C}_1)/n \xrightarrow{\text{p}} p_k(1 - \varkappa^k), \text{ for every } k \geq 0,$$

while $|\mathcal{C}_2|/n \xrightarrow{\text{p}} 0$.

(ii) *If $\nu_\infty - \mu_\infty = \mathbb{E}\, D_\infty(D_\infty - 2) \leq 0$, then $|\mathcal{C}_1|/n \xrightarrow{\text{p}} 0$.*

The same results hold for $G^*(n, \mathbf{d})$.

In the usual, somewhat informal, language, the theorem shows that $G(n, \mathbf{d})$ has a giant component if and only if $\nu_\infty - \mu_\infty = \mathbb{E}\, D_\infty(D_\infty - 2) > 0$. We say that $G(n, \mathbf{d})$ is *subcritical* if $\nu_\infty < \mu_\infty$ ($\mathbb{E}\, D_\infty(D_\infty - 2) < 0$), *critical* if $\nu_\infty = \mu_\infty$ ($\mathbb{E}\, D_\infty(D_\infty - 2) = 0$), *supercritical* if $\nu_\infty > \mu_\infty$ ($\mathbb{E}\, D_\infty(D_\infty - 2) > 0$).

Remark 2.5. Condition 2.1(ii),(iii) and (2.6) imply that

$$\liminf_{n\to\infty} \mathbb{P}\big(G^*(n, \mathbf{d}) \text{ is a simple graph}\big) > 0;$$

see for instance [2], [3], [6, Section II.4], [29] and [30] under some extra conditions on $\max d_i$, and [18] for the general case. Since we obtain $G(n, \mathbf{d})$ by conditioning $G^*(n, \mathbf{d})$ on being a simple graph, the results in the present paper for $G(n, \mathbf{d})$ follow from the results for $G^*(n, \mathbf{d})$ by this conditioning. (We only sometimes state the results for both $G(n, \mathbf{d})$ and $G^*(n, \mathbf{d})$ explicitly.)

Remark 2.6. Condition 2.1(iv) excludes the case $p_1 = 0$, when there are some pathologies, in particular in the critical case (which for $p_1 = 0$ occurs when $p_0+p_2=1$, i.e., $D_\infty \in \{0,2\}$ a.s.) We give some counterexamples for this case in Section 10, see also [21, Remark 2.7].

The supercritical case with $p_1=0$ (which occurs as soon as $\mathbb{P}(D_\infty \geq 3) > 0$) is better behaved. In this case, Theorem 2.4 holds with $\varkappa = 0$ and thus $|\mathcal{C}_1| = n - o(n)$, see [21, Remark 2.7]; hence $\chi(G(n,\mathbf{d})) = n - o(n)$. We conjecture that $\widehat{\chi}(G(n,\mathbf{d})) \xrightarrow{\mathrm{p}} 0$ always in this case, but we have not verified it. One important example is the random d-regular graph $G(n,d)$, when all $d_i = d$ for some fixed $d \geq 3$. In fact [4; 37], for $d \geq 3$, w.h.p. $G(n,d)$ is connected and thus trivially $\chi(G(n,d)) = n \to \infty$ and $\widehat{\chi}(G(n,d)) = 0$, in accordance with Theorem 1.1 (with $\varkappa = 0$).

3. Branching processes

For standard material on branching processes, see e.g. [1]. We review some basic facts that are important for us. The branching processes that we will use are Galton–Watson processes where the initial individual has a special offspring distribution. They are in general defined as follows.

Let ξ_0 and ξ be two given nonnegative integer-valued random variables (only their distributions matter). Start the branching process $\mathfrak{X}$ with one individual in generation 0, and give it a random number ξ_0 of children. In the sequel, give each individual a number of children that is distributed as ξ, with all these numbers independent.

We let $|\mathfrak{X}|$ denote the total population size of $\mathfrak{X}$, and define

$$\rho_k = \rho_k(\mathfrak{X}) := \mathbb{P}(|\mathfrak{X}| = k), \qquad 1 \leq k \leq \infty.$$

In particular, ρ_∞ is the survival probability of $\mathfrak{X}$, i.e., the probability that $\mathfrak{X}$ lives for ever.

Let $G_0(x) := \mathbb{E}\, x^{\xi_0}$ and $G(x) := \mathbb{E}\, x^{\xi}$ be the probability generating functions of ξ_0 and ξ. We define $\varkappa$ as the smallest non-negative solution to

$$G(\varkappa) = \varkappa. \tag{3.1}$$

For a standard Galton–Watson process ($\xi_0 = \xi$), it is well-known that this is the extinction probability. In general, by conditioning on ξ_0,

$$1 - \rho_\infty = \mathbb{P}(|\mathfrak{X}| < \infty) = \mathbb{E}\, \varkappa^{\xi_0} = G_0(\varkappa). \tag{3.2}$$

The susceptibility and modified susceptibility are defined by

$$\chi(\mathfrak{X}) := \mathbb{E}\big(|\mathfrak{X}|\big) = \sum_{1 \leq k \leq \infty} k\rho_k, \tag{3.3}$$

$$\widehat{\chi}(\mathfrak{X}) := \mathbb{E}\big(|\mathfrak{X}|; |\mathfrak{X}| < \infty\big) = \sum_{1 \leq k < \infty} k\rho_k. \tag{3.4}$$

(Note that these are expectations and not random variables.) Thus, $\chi(\mathfrak{X}) = \widehat{\chi}(\mathfrak{X})$ when the survival probability $\rho_\infty = 0$ (the subcritical or critical case), and $\chi(\mathfrak{X}) = \infty \geq \widehat{\chi}(\mathfrak{X})$ when $\rho_\infty > 0$ (the supercritical case).

For our random graph $G(n, \mathbf{d})$ with a given degree sequence satisfying Condition 2.1, we define the corresponding branching process as the Galton–Watson branching process with initial offspring distribution $\xi_0 := D_\infty$, and general offspring distribution $\xi = D_\infty^*$, where D_∞^* is the shifted size-biased version (or transform) of D_∞ defined by

$$\mathbb{P}(D_\infty^* = k) = \frac{(k+1)\,\mathbb{P}(D_\infty = k+1)}{\mathbb{E}\,D_\infty}, \qquad k \geq 0. \tag{3.5}$$

(We assume Condition 2.1, so $0 < \mathbb{E}\,D_\infty < \infty$ and then (3.5) defines a probability distribution.) The reason for this definition is the well-known fact that D_∞^* appears as the natural limit distribution when exploring components locally; the novice can see this in the proof of Lemma 4.1 below.

Note that

$$\begin{aligned}\mathbb{E}\,D_\infty^* &= \sum_{k=0}^\infty k\,\mathbb{P}(D_\infty^* = k) = \sum_{k=0}^\infty \frac{k(k+1)\,\mathbb{P}(D_\infty = k+1)}{\mathbb{E}\,D_\infty} \\ &= \frac{\mathbb{E}\,D_\infty(D_\infty - 1)}{\mathbb{E}\,D_\infty} = \frac{\nu_\infty}{\mu_\infty}.\end{aligned} \tag{3.6}$$

Hence, the standard classification of $\mathfrak{X}$ as subcritical, critical or supercritical depending on whether the expected number of children satisfies $\mathbb{E}\,\xi < 1$, $\mathbb{E}\,\xi = 1$ or $\mathbb{E}\,\xi > 1$, becomes the conditions $\nu_\infty < \mu_\infty$, $\nu_\infty = \mu_\infty$ and $\nu_\infty > \mu_\infty$ we already have seen for $G(n, \mathbf{d})$, and there is a perfect agreement between these types for $\mathfrak{X}$ and for $G(n, \mathbf{d})$. (This indicates that it really is ν_∞/μ_∞ rather than $\nu_\infty - \mu_\infty$ that is the natural parameter for criticality testing for $G(n, \mathbf{d})$; this is well-known, see e.g. [10] for generalizations.)

Furthermore, if $g_*(x) := \mathbb{E}\,x^{D_\infty^*}$ is the probability generating function of D_∞^*, then

$$G(x) = g_*(x) = \sum_{k=0}^\infty \frac{(k+1)\,\mathbb{P}(D_\infty = k+1)x^k}{\mathbb{E}\,D_\infty} = \frac{g'(x)}{\mathbb{E}\,D_\infty}. \tag{3.7}$$

(Note that $\mathbb{E}\,D_\infty = g'(1)$, so $g_*(1) = 1$ as it should.) In particular, (3.7) shows that (3.1) can be written

$$g'(\varkappa) = \mu_\infty \varkappa. \tag{3.8}$$

Thus, in the supercritical case $\varkappa$ here is the same as in Theorem 2.4. (In the subcritical and critical cases, $\varkappa = 1$, which always satisfies (3.8).) Further, by (3.2), the asymptotic relative size $1 - g(\varkappa)$ of $\mathcal{C}_1$ in Theorem 2.4(i) equals ρ_∞ for the corresponding branching process $\mathfrak{X}$. (Recall that $G = g$.)

We can easily compute the susceptibility of a Galton–Watson process by standard calculations. We consider the general version with ξ_0 and ξ, before specializing to the branching process corresponding to $G(n, \mathbf{d})$. This calculation has been done by Newman, Strogatz and Watts [33], see also Durrett [13, Section 2.3]. (Similar results for a more complicated branching process with types, but without a special initial offspring distribution, are given in [26].)

Theorem 3.1. *For a branching process $\mathfrak{X}$ defined as above by ξ_0 and ξ,*

$$\chi(\mathfrak{X}) := \mathbb{E}(|\mathfrak{X}|) = 1 + \frac{\mathbb{E}\,\xi_0}{(1-\mathbb{E}\,\xi)_+}. \tag{3.9}$$

Further, if ξ_0 and ξ have the probability generating functions G_0 and G, and $\varkappa$ is the smallest nonnegative root of $G(\varkappa) = \varkappa$, then

$$\widehat{\chi}(\mathfrak{X}) := \mathbb{E}(|\mathfrak{X}|; |\mathfrak{X}| < \infty) = G_0(\varkappa) + \frac{\varkappa G_0'(\varkappa)}{1 - G'(\varkappa)}. \tag{3.10}$$

Hence, assuming $\mathbb{E}\,\xi_0 > 0$, we have $\chi(\mathfrak{X}) = \infty$ if and only if $\mathbb{E}\,\xi \geq 1$, while $\widehat{\chi}(\mathfrak{X}) < \infty$ whenever $\mathbb{E}\,\xi \neq 1$.

Proof. As said above, this is proved by Newman, Strogatz and Watts [33] (in slightly different notation), but for completeness we give a proof. Let $\mathfrak{X}_k$ be the kth generation of $\mathfrak{X}$. Then $\mathbb{E}\,|\mathfrak{X}_k| = \mathbb{E}\,\xi_0(\mathbb{E}\,\xi)^{k-1}$ for $k \geq 1$, and thus when $\mathbb{E}\,\xi \leq 1$ (so $\varkappa = 1$ and $\rho_\infty = 0$)

$$\chi(\mathfrak{X}) = 1 + \sum_{k=1}^{\infty} \mathbb{E}\,\xi_0(\mathbb{E}\,\xi)^{k-1} = 1 + \frac{\mathbb{E}\,\xi_0}{1-\mathbb{E}\,\xi},$$

while $\chi(\mathfrak{X}) = \infty$ when $\mathbb{E}\,\xi > 1$ and thus $\rho_\infty > 0$. This shows (3.9).

For $\widehat{\chi}$ we use the standard and easily verified fact that $\hat{\mathfrak{X}} := (\mathfrak{X} \mid |\mathfrak{X}| < \infty)$, i.e. $\mathfrak{X}$ conditioned on extinction, is another branching process with initial offspring distribution $\widehat{\xi}_0$ and general offspring distribution $\widehat{\xi}$ given by

$$\begin{aligned}
\mathbb{P}(\widehat{\xi} = k) &= \frac{\varkappa^k\,\mathbb{P}(\xi = k)}{G(\varkappa)} = \varkappa^{k-1}\,\mathbb{P}(\xi = k), \\
\mathbb{P}(\widehat{\xi}_0 = k) &= \frac{\varkappa^k\,\mathbb{P}(\xi_0 = k)}{G_0(\varkappa)};
\end{aligned}$$

these have expectations

$$\mathbb{E}\,\widehat{\xi} = G'(\varkappa), \tag{3.11}$$

$$\mathbb{E}\,\widehat{\xi}_0 = \frac{\varkappa G_0'(\varkappa)}{G_0(\varkappa)}. \tag{3.12}$$

If $\mathbb{E}\,\xi \leq 1$, then $\varkappa = 1$ and $\widehat{\xi} \overset{\mathrm{d}}{=} \xi$, $\widehat{\xi}_0 \overset{\mathrm{d}}{=} \xi_0$; thus (3.10) reduces to (3.9). (Except in the trivial case $\mathbb{P}(\xi = 1) = 1$, when $|\mathfrak{X}| = 1$ or ∞ a.s.; then $\varkappa = 0$ and we interpret (3.10) as $\chi(\hat{\mathfrak{X}}) = \mathbb{P}(|\mathfrak{X}| < \infty) = G_0(0)$, which is immediate.)

Note that, by the convexity of G, if $\mathbb{E}\,\xi = G'(1) > 1$, so $\varkappa < 1$, then $\mathbb{E}\,\widehat{\xi} = G'(\varkappa) < 1$, and $\hat{\mathfrak{X}}$ is subcritical. Thus (3.2), (3.9) and (3.11)–(3.12) yield

$$\begin{aligned}
\widehat{\chi}(\mathfrak{X}) &= \mathbb{P}(|\mathfrak{X}| < \infty)\,\mathbb{E}(|\mathfrak{X}| \mid |\mathfrak{X}| < \infty) = \mathbb{P}(|\mathfrak{X}| < \infty)\,\mathbb{E}\,|\hat{\mathfrak{X}}| \\
&= \mathbb{P}(|\mathfrak{X}| < \infty)\chi(\hat{\mathfrak{X}}) = G_0(\varkappa)\Bigl(1 + \frac{\mathbb{E}\,\widehat{\xi}_0}{1-\mathbb{E}\,\widehat{\xi}}\Bigr) = G_0(\varkappa) + \frac{\varkappa G_0'(\varkappa)}{1-G'(\varkappa)}.
\end{aligned}$$

(The case $G_0(\varkappa) = 0$, which occurs when $\mathbb{P}(\xi_0 = 0) = \mathbb{P}(\xi = 0) = 0$, and entails $\varkappa = 0$ and $|\mathfrak{X}| = \infty$ a.s., is trivial and easily verified separately.) □

We now specialize to the branching process corresponding to $G(n, \mathbf{d})$.

Corollary 3.2. *Given $\mathbf{d} = (d_i)_1^n$ such that Condition 2.1 holds, let $\mathfrak{X}$ be the corresponding branching process. Then*

$$\chi(\mathfrak{X}) = 1 + \frac{\mu_\infty}{(1 - \nu_\infty/\mu_\infty)_+} = 1 + \frac{\mu_\infty^2}{(\mu_\infty - \nu_\infty)_+} \tag{3.13}$$

$$\widehat{\chi}(\mathfrak{X}) = g(\varkappa) + \frac{\varkappa g'(\varkappa)}{1 - g''(\varkappa)/\mu_\infty} = g(\varkappa) + \frac{\varkappa g'(\varkappa) g'(1)}{g'(1) - g''(\varkappa)} \tag{3.14}$$

$$= g(\varkappa) + \frac{\varkappa g'(\varkappa)^2}{g'(\varkappa) - \varkappa g''(\varkappa)}, \tag{3.15}$$

with $\widehat{\chi}(\mathfrak{X}) < \infty$ unless $\nu_\infty = \mu_\infty$.

Proof. First, $\mathbb{E}\,\xi_0 = \mathbb{E}\, D_\infty = \mu_\infty$ and, by (3.6), $\mathbb{E}\,\xi = \mathbb{E}\, D_\infty^* = \nu_\infty/\mu_\infty$; thus (3.9) yields (3.13).

Next, $G_0(x) = g(x)$ and, by (3.7), $G(x) = g'(x)/\mu_\infty$. Hence, (3.10) yields, using (3.8),

$$\widehat{\chi}(\mathfrak{X}) = g(\varkappa) + \frac{\varkappa g'(\varkappa)}{1 - g''(\varkappa)/\mu_\infty} = g(\varkappa) + \frac{\varkappa g'(\varkappa)^2}{g'(\varkappa) - \varkappa g''(\varkappa)},$$

and the result follows, recalling also $\mu_\infty = g'(1)$. □

The values $\chi(\mathfrak{X})$ and $\widehat{\chi}(\mathfrak{X})$ are thus the quantities called χ_∞ and $\widehat{\chi}_\infty$ in Theorem 1.1, and Theorem 1.1 may thus be reformulated as follows.

Theorem 3.3. *Suppose that Conditions 2.1 and 2.3 hold, and let $\mathfrak{X}$ be the corresponding branching process. Then $\chi(G(n, \mathbf{d})) \xrightarrow{\text{p}} \chi(\mathfrak{X})$ and $\widehat{\chi}(G(n, \mathbf{d})) \xrightarrow{\text{p}} \widehat{\chi}(\mathfrak{X})$. In the subcritical case $\nu_\infty < \mu_\infty$, further $\chi(G(n, \mathbf{d})) \xrightarrow{L^1} \chi(\mathfrak{X})$ and $\widehat{\chi}(G(n, \mathbf{d})) \xrightarrow{L^1} \widehat{\chi}(\mathfrak{X})$.*

4. A lower bound

We continue to assume Condition 2.1, and let $\mathfrak{X}$ be the branching process corresponding to $G(n, \mathbf{d})$ as in Section 3. Let $N_k(G)$ denote the number of vertices in components of order k in a graph G. Thus the number of such components is $N_k(G)/k$. We can write the definition (1.2) as

$$\chi(G) = \frac{1}{|G|} \sum_{k=1}^{\infty} \frac{N_k(G)}{k} k^2 = \sum_{k=1}^{\infty} k \frac{N_k(G)}{|G|}. \tag{4.1}$$

Lemma 4.1. *Suppose that Condition 2.1 holds. Then, for every fixed $k \geq 0$,*

$$N_k(G(n, \mathbf{d}))/n \xrightarrow{\text{p}} \rho_k(\mathfrak{X}) := \mathbb{P}(|\mathfrak{X}| = k). \tag{4.2}$$

The same holds for $G^*(n, \mathbf{d})$.

Proof. This is well-known (and see e.g. [10] for a more general situation), but for completeness we sketch the proof. By Remark 2.5, it suffices to consider $G^*(n, \mathbf{d})$.

The expectation $\mathbb{E}\, N_k(G^*(n, \mathbf{d}))/n$ is the probability that a random vertex v_0 belongs to a component with exactly k vertices. We let N'_k be the number of vertices in tree components of order k, and note that it is easy to see that the expected number of cycles of length $\le k$ is $O(1)$, and thus $\mathbb{E}\,|N_k(G^*(n, \mathbf{d})) - N'_k| = O(1)$; hence it suffices to consider N'_k.

Let $\mathcal{C}$ be the component containing the random vertex v_0. We explore $\mathcal{C}$ by breadth-first search, using a predetermined order of the half-edges at each vertex. In this way, $\mathcal{C}$ is exhibited as an ordered (or plane) tree, possibly with some extra edges, and with root v_0. Let T be a given tree with k vertices, and let us compute the probability that $\mathcal{C}$ equals T (as an ordered, rooted, unlabelled tree). If T has a root of degree d_0 and $k-1$ other vertices of outdegrees $d_1, \dots, d_k$ (in breadth-first order), then there are n_{d_0} choices of v_0 and, for $i = 1, \dots, k$, $n_{d_i} - O(1)$ choices of the ith vertex. Moreover, for $i \ge 1$ we also have $d_i + 1$ choices of half-edge to connect to, out of $2m - O(1)$ remaining half-edges. The probability is thus, using (2.6) and (3.5),

$$\frac{n_{d_0}}{n} \prod_{i=1}^{k-1} \frac{(d_i+1)n_{d_i+1}}{2m} + o(1) = \mathbb{P}(D_\infty = d_0) \prod_{i=1}^{k-1} \mathbb{P}(D^*_\infty = d_i) + o(1),$$

which equals, except for $o(1)$, the probability that the family tree of $\mathfrak{X}$ (considered as an ordered tree) equals T. Summing over all trees T of order k, we find

$$\begin{aligned}\mathbb{E}\, N'_k/n = \mathbb{P}(\mathcal{C} \text{ is a tree of order } k) = \mathbb{P}(|\mathfrak{X}| = k) + o(1) = \rho_k(\mathfrak{X}) + o(1)\\ \to \rho_k(\mathfrak{X}).\end{aligned} \tag{4.3}$$

The same argument, but starting with two independent random vertices, shows that $\mathbb{E}\, N'_k(N'_k - k)/n^2 \to \rho_k(\mathfrak{X})^2$. Hence, $\operatorname{Var}(N'_k/n) \to 0$, and thus, by (4.3), $N'_k/n \xrightarrow{\text{p}} \rho_k(\mathfrak{X})$, which as said above completes the proof. □

Lemma 4.2. *Suppose that Condition 2.1 holds, and let a be a real number.*

(i) *If $a < \chi(\mathfrak{X})$, then $\chi(G^*(n, \mathbf{d})) > a$ w.h.p.*
(ii) *If further $\nu_\infty \le \mu_\infty$, then also $\widehat{\chi}(G^*(n, \mathbf{d})) > a$ w.h.p.*

Proof. In the supercritical case $\nu_\infty > \mu_\infty$, $\chi(G^*(n, \mathbf{d})) \ge \frac{1}{n}|\mathcal{C}_1|^2 = n(\rho_\infty^2 + o_{\text{p}}(1)) > a$ w.h.p. by Theorem 2.4.

Assume thus $\nu_\infty \le \mu_\infty$, and thus $\rho_\infty = 0$. By the assumption, (3.3) and $\rho_\infty = 0$, then $a < \chi(\mathfrak{X}) = \sum_{1 \le k < \infty} k\rho_k$. Hence there exists k_0 such that $\sum_{k=1}^{k_0} k\rho_k > a$. By (4.1) and Lemma 4.1,

$$\chi(G^*(n, \mathbf{d})) := \frac{1}{n} \sum_{k=1}^{\infty} kN_k \ge \sum_{k=1}^{k_0} \frac{kN_k}{n} \xrightarrow{\text{p}} \sum_{k=1}^{k_0} k\rho_k > a,$$

and thus $\chi(G^*(n, \mathbf{d})) > a$ w.h.p.

Finally, considering the two cases $|\mathcal{C}_1| > k_0$ and $|\mathcal{C}_1| \le k_0$ separately, we see that

$$\widehat{\chi}(G^*(n,\mathbf{d})) \ge \frac{1}{n}\sum_{k=1}^{k_0} kN_k - \frac{k_0^2}{n} \overset{\mathrm{p}}{\longrightarrow} \sum_{k=1}^{k_0} k\rho_k > a,$$

and thus also $\widehat{\chi}(G^*(n,\mathbf{d})) > a$ w.h.p. $\square$

5. An upper bound

A path of length $\ell \ge 0$ in a multigraph is a sequence $i_0 e_1 i_1 \cdots e_\ell i_\ell$ of alternating vertices and edges that are distinct and such that each e_j has endpoints i_{j-1} and i_j.

Lemma 5.1. *Let P_ℓ be the number of paths of length ℓ in $G^*(n,\mathbf{d})$. Then, for every $\ell \ge 1$,*

$$\mathbb{E}\,P_\ell \le \frac{(n\nu_n)^{\ell-1}}{(n\mu_n)^{\ell-2}} = n\frac{\nu_n^{\ell-1}}{\mu_n^{\ell-2}}. \tag{5.1}$$

Proof. If $\ell = 1$, then (5.1) says that $\mathbb{E}\,P_1 \le n\mu_n = \sum_i d_i = 2m$; this is trivially true: $P_1 \le 2m$ because a path of length 1 is a single edge and each edge that is not a loop yields two paths in the opposite direction.

Let $\ell \ge 2$ and $\ell \le m$. A path $i_0 \cdots i_\ell$ contains one half-edge at i_0 and at i_ℓ, and two half-edges at each of $i_1, \dots, i_{\ell-1}$; these may be chosen arbitrarily, and for each choice, the probability that they are connected to each other in the right way is $(2m-1)^{-1} \cdots (2m-2\ell+1)^{-1}$. Hence,

$$\mathbb{E}\,P_\ell = \frac{\sum^*_{i_0,\dots,i_\ell} d_{i_0} \cdot \prod_{j=1}^{\ell-1} d_{i_j}(d_{i_j}-1) \cdot d_{i_\ell}}{\prod_{j=1}^{\ell}(2m-2j+1)},$$

where $\sum^*$ denotes the sum over *distinct* indices.

For each choice of distinct $i_0, \dots, i_{\ell-1}$ with $d_{i_0} \ge 1$ and $d_{i_j} \ge 2$, $1 \le j \le \ell-1$, the sum $\sum d_{i_\ell}$ over $i_\ell \notin \{i_0, \dots, i_{\ell-1}\}$ equals

$$\sum_{i=1}^n d_i - d_{i_0} - \sum_{j=1}^{\ell-1} d_{i_j} \le 2m - 1 - 2(\ell-1) = 2m - 2\ell + 1.$$

Hence,

$$\mathbb{E}\,P_\ell \le \frac{\sum^*_{i_0,\dots,i_{\ell-1}} d_{i_0} \prod_{j=1}^{\ell-1} d_{i_j}(d_{i_j}-1)}{\prod_{j=1}^{\ell-1}(2m-2j+1)}.$$

Similarly, for each choice of distinct $i_1, \dots, i_{\ell-1}$ with $d_{i_j} \ge 2$, $1 \le j \le \ell-1$, the sum $\sum d_{i_0}$ over $i_0 \notin \{i_1, \dots, i_{\ell-1}\}$ equals

$$\sum_{i=1}^n d_i - \sum_{j=1}^{\ell-1} d_{i_j} \le 2m - 2(\ell-1) < 2m - 2\ell + 3.$$

Hence,

$$\mathbb{E}\, P_\ell \le \frac{\sum^*_{i_1,\dots,i_{\ell-1}} \prod_{j=1}^{\ell-1} d_{i_j}(d_{i_j}-1)}{\prod_{j=1}^{\ell-2}(2m-2j+1)} \le 2^{-(\ell-2)} \frac{\sum^*_{i_1,\dots,i_{\ell-1}} \prod_{j=1}^{\ell-1} d_{i_j}(d_{i_j}-1)}{\prod_{j=1}^{\ell-2}(m-j)} \tag{5.2}$$

Let $a_i := d_i(d_i-1)$ and let R be the set of indices i such that $a_i > 0$, i.e., $d_i \ge 2$. Let $r := |R|$. By an inequality of Maclaurin [15, Theorem 52], for $2 \le \ell \le r+1$,

$$\left(\frac{(r-\ell+1)!}{r!} \sum^*_{i_1,\dots,i_{\ell-1}\in R} \prod_{j=1}^{\ell-1} a_{i_j}\right)^{1/(\ell-1)} \le \frac{1}{r}\sum_{i\in R} a_i \le \frac{1}{r}\sum_{i=1}^n a_i = \frac{n\nu_n}{r}.$$

Hence,

$$\begin{aligned}\sum^*_{i_1,\dots,i_{\ell-1}} \prod_{j=1}^{\ell-1} d_{i_j}(d_{i_j}-1) &= \sum^*_{i_1,\dots,i_{\ell-1}\in R} \prod_{j=1}^{\ell-1} a_{i_j} \le \Bigl(\frac{n\nu_n}{r}\Bigr)^{\ell-1} \frac{r!}{(r-\ell+1)!} \\ &= (n\nu_n)^{\ell-1} \prod_{j=0}^{\ell-2}\Bigl(1-\frac{j}{r}\Bigr).\end{aligned}$$

Further, $2m = \sum_{i=1}^n d_i \ge 2r$, so $r \le m$. Consequently, (5.2) yields

$$\mathbb{E}\, P_\ell \le (2m)^{-(\ell-2)} \frac{\prod_{j=1}^{\ell-2}(1-j/r)}{\prod_{j=1}^{\ell-2}(1-j/m)} \cdot (n\nu_n)^{\ell-1} \le \frac{(n\nu_n)^{\ell-1}}{(2m)^{\ell-2}} = \frac{(n\nu_n)^{\ell-1}}{(n\mu_n)^{\ell-2}},$$

which proves the result when $2 \le \ell \le m$ and $\ell \le r+1$. Since trivially $P_\ell = 0$ if $\ell > m$ or $\ell - 1 > r$, this completes the proof. $\square$

Lemma 5.2. *For any* $\mathbf{d}$,

$$\mathbb{E}\bigl(\chi(G^*(n,\mathbf{d}))\bigr) \le 1 + \frac{\mu_n^2}{(\mu_n-\nu_n)_+}. \tag{5.3}$$

Proof. By [26, Lemma 4.6] (which trivially extends to multigraphs),

$$\chi(G^*(n,\mathbf{d})) \le \frac{1}{n}\sum_{\ell=0}^\infty P_\ell(G^*(n,\mathbf{d})) = 1 + \frac{1}{n}\sum_{\ell=1}^\infty P_\ell(G^*(n,\mathbf{d})).$$

Hence, using Lemma 5.1 for $\ell \ge 1$,

$$\mathbb{E}\bigl(\chi(G^*(n,\mathbf{d}))\bigr) \le 1 + \sum_{\ell=1}^\infty \frac{\mathbb{E}\, P_\ell(G^*(n,\mathbf{d}))}{n} \le 1 + \sum_{\ell=1}^\infty \frac{\nu_n^{\ell-1}}{\mu_n^{\ell-2}}.$$

If $\mu_n > \nu_n$, the geometric series sums to $\mu_n/(1-\nu_n/\mu_n) = \mu_n^2/(\mu_n-\nu_n)$, and the result follows. The case $\nu_n \ge \mu_n > 0$ is trivial, since the right hand side of (5.3) is ∞, and the case $\mu_n = 0$ is trivial too, since then there are no edges at all and thus $\chi(G^*(n,\mathbf{d})) = 1$. $\square$

6. Duality

Let, as before, $\mathcal{C}_1$ be the largest component of $G(n, \mathbf{d})$; if there is a tie we for definiteness choose the component with maximal size that contains the vertex with the largest label. Consider the complement of $\mathcal{C}_1$; we denote this random graph by $\widehat{G}(n, \mathbf{d}) := G(n, \mathbf{d}) \setminus \mathcal{C}_1$. This graph has thus the random vertex set $[n] \setminus \mathcal{C}_1^\circ$, where $[n] := \{1, \dots, n\}$ and we in this section denote the vertex set of a graph $\mathcal{C}$ by $\mathcal{C}^\circ$.

We have defined $G(n, \mathbf{d})$ with the vertex set $[n]$. Of course, the definition generalizes to an arbitrary finite vertex set A and a degree sequence $\mathbf{d} = (d_i)_{i \in A}$; we denote this random graph by $G(A, \mathbf{d})$. If A is a subset of $[n]$, and $\mathbf{d} = (d_i)_1^n$, let $\mathbf{d}|_A$ be the sequence $(d_i)_{i \in A}$.

We construct, given n and $\mathbf{d} = (d_i)_1^n$, a random graph $\widetilde{G}(n, \mathbf{d})$ by first constructing $G(n, \mathbf{d})$ and finding its largest component $\mathcal{C}_1$; we then, given $\mathcal{C}_1$, let $A := [n] \setminus \mathcal{C}_1$ and construct a new random graph $G(A, \mathbf{d}|_A)$ and take that as our random graph $\widetilde{G}(n, \mathbf{d})$. Hence $\widetilde{G}(n, \mathbf{d})$ is a random graph where both the vertex set and the edge set are random, but conditioned on the vertex set, it is a uniform random graph with given vertex degrees. Our version of the duality theorem is that $\widehat{G}(n, \mathbf{d})$ and $\widetilde{G}(n, \mathbf{d})$ are equal w.h.p., with a suitable coupling. This is a precise version of saying that $\widehat{G}(n, \mathbf{d})$ conditioned on its vertex set almost is a uniform random graph with given vertex degrees.

Theorem 6.1. *Suppose that Condition 2.1 holds, and that $\nu_\infty > \mu_\infty$. With the notations above, it is possible to couple $\widehat{G}(n, \mathbf{d})$ and $\widetilde{G}(n, \mathbf{d})$ such that they coincide w.h.p. Furthermore, we may assume, by another coupling, that $\widetilde{G}(n, \mathbf{d})$ conditioned on its order and degree sequence satisfies Condition 2.1, with p_k replaced by $\widehat{p}_k := p_k \varkappa^k / g(\varkappa)$. Let $\widehat{D}_\infty$ be a random variable with this distribution:*

$$\mathbb{P}(\widehat{D}_\infty = k) = \widehat{p}_k := \frac{p_k \varkappa^k}{g(\varkappa)}, \qquad k \geq 0. \tag{6.1}$$

Then $\widehat{D}_\infty$ has probability generating function

$$\widehat{g}(x) := \mathbb{E}\, x^{\widehat{D}_\infty} = \frac{g(\varkappa x)}{g(\varkappa)}, \tag{6.2}$$

and

$$\widehat{\mu}_\infty := \mathbb{E}\, \widehat{D}_\infty = \widehat{g}'(1) = \frac{\varkappa g'(\varkappa)}{g(\varkappa)} = \frac{\varkappa^2 \mu_\infty}{g(\varkappa)}, \tag{6.3}$$

$$\widehat{\nu}_\infty := \mathbb{E}\, \widehat{D}_\infty(\widehat{D}_\infty - 1) = \widehat{g}''(1) = \frac{\varkappa^2 g''(\varkappa)}{g(\varkappa)}. \tag{6.4}$$

Moreover, $\widehat{\nu}_\infty < \widehat{\mu}_\infty$, so $\widetilde{G}(n, \mathbf{d})$ is subcritical.

Before giving the proof, we give a simple and well-known result on conditioning. Recall that the *total variation distance* between two random variables X and Y (taking values in any common space) is

$$d_{\mathrm{TV}}(X,Y) := \sup_A \bigl(\mathbb{P}(X \in A) - \mathbb{P}(Y \in A)\bigr),$$

taking the supremum over all measurable sets A. Recall further that the existence of a coupling with $\widehat{G}(n,\mathbf{d}) = \widetilde{G}(n,\mathbf{d})$ w.h.p. is equivalent to $d_{\mathrm{TV}}(\widehat{G}(n,\mathbf{d}), \widetilde{G}(n,\mathbf{d})) \to 0$.

Lemma 6.2. *If X is any random variable (with values in any space) and $\mathcal{E}$ is any event with $\mathbb{P}(\mathcal{E}) > 0$, then $d_{\mathrm{TV}}\bigl((X \mid \mathcal{E}), X\bigr) \le 1 - \mathbb{P}(E)$.*

Proof. For any event $\mathcal{A}$ of the type $\{X \in A\}$,

$$\mathbb{P}(\mathcal{A} \mid \mathcal{E}) - \mathbb{P}(\mathcal{A}) = \frac{\mathbb{P}(\mathcal{A} \cap \mathcal{E})}{\mathbb{P}(\mathcal{E})} - \mathbb{P}(\mathcal{A}) \le \frac{\mathbb{P}(\mathcal{A}) \wedge \mathbb{P}(\mathcal{E})}{\mathbb{P}(\mathcal{E})} - \mathbb{P}(\mathcal{A}).$$

The right-hand side is a function of $\mathbb{P}(\mathcal{A})$ that is maximal for $\mathbb{P}(\mathcal{A}) = \mathbb{P}(\mathcal{E})$, when it equals $1 - \mathbb{P}(\mathcal{E})$. $\square$

Proof of Theorem 6.1. Define a total order $\prec$ on the subsets of $[n]$ by defining $A \prec B$ if $|A| < |B|$ or $|A| = |B|$ and $\max A < \max B$. Thus $\mathcal{C}_1$ is by definition the component of $G(n,\mathbf{d})$ whose vertex set is maximal in this order.

Let $A \subseteq [n]$. Conditioned on $\mathcal{C}_1^\circ = A$, the complement $G(n,\mathbf{d}) \setminus \mathcal{C}_1 = G(n,\mathbf{d})|_{[n]\setminus A}$ is a random graph on the vertex set $[n] \setminus A$ with a given degree sequence $(d_i)_{i \in [n]\setminus A}$. Moreover, it may be any such graph except that it must not contain a component $\mathcal{C}$ with $\mathcal{C}^\circ \succ A$; furthermore, all permitted graphs have the same probability. Thus, conditioned on $\mathcal{C}_1^\circ = A$,

$$\widehat{G}(n,\mathbf{d}) = G(n,\mathbf{d})|_{[n]\setminus A} \overset{\mathrm{d}}{=} \bigl(G([n] \setminus A, \mathbf{d}|_{[n]\setminus A}) \bigm| \widehat{\mathcal{E}}_A^{\mathrm{c}}\bigr),$$

where $\widehat{\mathcal{E}}_A^{\mathrm{c}}$ is the complement of the event $\widehat{\mathcal{E}}_A$ that $G([n] \setminus A, \mathbf{d}|_{[n]\setminus A})$ contains a component $\mathcal{C}$ with $\mathcal{C}^\circ \succ A$.

On the other hand, by definition, conditioned on $\mathcal{C}_1^\circ = A$ we have $\widetilde{G}(n,\mathbf{d}) = G([n] \setminus A, \mathbf{d}|_{[n]\setminus A})$. Hence, by Lemma 6.2, the total variation distance between $\widehat{G}(n,\mathbf{d})$ and $\widetilde{G}(n,\mathbf{d})$, both conditioned on $\mathcal{C}_1^\circ = A$, is

$$\begin{aligned} &d_{\mathrm{TV}}\Bigl(\bigl(\widehat{G}(n,\mathbf{d}) \mid \mathcal{C}_1^\circ = A\bigr), \bigl(\widetilde{G}(n,\mathbf{d}) \mid \mathcal{C}_1^\circ = A\bigr)\Bigr) \\ &= d_{\mathrm{TV}}\Bigl(\bigl(G([n] \setminus A, \mathbf{d}|_{[n]\setminus A}) \bigm| \widehat{\mathcal{E}}_A^{\mathrm{c}}\bigr), G([n] \setminus A, \mathbf{d}|_{[n]\setminus A})\Bigr) \le 1 - \mathbb{P}(\widehat{\mathcal{E}}_A^{\mathrm{c}}) = \mathbb{P}(\widehat{\mathcal{E}}_A). \end{aligned}$$

Taking the expectation over $\mathcal{C}_1^\circ$ we find

$$\begin{aligned} d_{\mathrm{TV}}\bigl(\widehat{G}(n,\mathbf{d}), \widetilde{G}(n,\mathbf{d})\bigr) &\le \mathbb{E}\, d_{\mathrm{TV}}\Bigl(\bigl(\widehat{G}(n,\mathbf{d}) \mid \mathcal{C}_1^\circ\bigr), \bigl(\widetilde{G}(n,\mathbf{d}) \mid \mathcal{C}_1^\circ\bigr)\Bigr) \\ &\le \mathbb{E}\, \mathbb{P}(\widehat{\mathcal{E}}_{\mathcal{C}_1^\circ}) = \sum_{A \subseteq [n]} \mathbb{P}(\widehat{\mathcal{E}}_A)\, \mathbb{P}(\mathcal{C}_1^\circ = A). \end{aligned} \tag{6.5}$$

We split this sum into two parts. Let $r := \rho_\infty/2$; thus $r > 0$ and $\mathbb{P}(|\mathcal{C}_1| \ge rn) \to 1$. Further, let $\mathcal{E}_A$ be the event that $A = \mathcal{C}^\circ$ for some component $\mathcal{C}$

of $G(n, \mathbf{d})$, and note that $\mathcal{C}_1^\circ = A$ implies $\mathcal{E}_A$. Thus

$$\sum_{A \subseteq [n]} \mathbb{P}(\widehat{\mathcal{E}}_A)\,\mathbb{P}(\mathcal{C}_1^\circ = A) \le \mathbb{P}(|\mathcal{C}_1| < rn) + \sum_{|A| \ge rn} \mathbb{P}(\widehat{\mathcal{E}}_A)\,\mathbb{P}(\mathcal{E}_A). \tag{6.6}$$

Conditioned on $\mathcal{E}_A$, the complement $G(n, \mathbf{d})|_{[n]\setminus A}$ of A has the same distribution as $G([n] \setminus A, \mathbf{d}_{[n]\setminus A})$, and thus

$$\begin{aligned}\mathbb{P}(\widehat{\mathcal{E}}_A) &= \mathbb{P}\bigl(G(n, \mathbf{d}) \text{ contains a component } \mathcal{C} \text{ with } \mathcal{C}^\circ \succ A \bigm| \mathcal{E}_A\bigr) \\ &= \mathbb{P}\Bigl(\bigcup_{B \succ A} \mathcal{E}_B \Bigm| \mathcal{E}_A\Bigr).\end{aligned}$$

Hence,

$$\sum_{|A| \ge rn} \mathbb{P}(\widehat{\mathcal{E}}_A)\,\mathbb{P}(\mathcal{E}_A) = \sum_{|A| \ge rn} \mathbb{P}\bigl(\bigcup_{B \succ A} \mathcal{E}_B \cap \mathcal{E}_A\bigr) = \mathbb{E} \sum_{|A| \ge rn} \mathbf{1}\bigl[\mathcal{E}_A \cap \bigcup_{B \succ A} \mathcal{E}_B\bigr]. \tag{6.7}$$

Let N be the number of components of size $\ge rn$ in $G(n, \mathbf{d})$. If we order these components as $A_1 \prec \cdots \prec A_N$, then the indicator in (6.7) is 1 exactly when A is one of $A_1, \dots, A_{N-1}$, so the sum is $(N-1)_+$. Further, since components are disjoint, $N \le n/(rn) = r^{-1}$, and thus $(N-1)_+ \le r^{-1}\mathbf{1}[N \ge 2]$. However, $N \ge 2$ if and only if the *second* largest component $\mathcal{C}_2$ is larger than rn. Consequently, by (6.5), (6.6), (6.7),

$$\begin{aligned}d_{\mathrm{TV}}\bigl(\widehat{G}(n, \mathbf{d}), \widetilde{G}(n, \mathbf{d})\bigr) &\le \mathbb{P}(|\mathcal{C}_1| < rn) + r^{-1}\,\mathbb{P}(N \ge 2) \\ &= \mathbb{P}(|\mathcal{C}_1| < rn) + r^{-1}\,\mathbb{P}(|\mathcal{C}_2| \ge rn),\end{aligned}$$

where, by Theorem 2.4, both terms on the right-hand side tend to 0. This shows the existence of a coupling with $\widehat{G}(n, \mathbf{d}) = \widetilde{G}(n, \mathbf{d})$ w.h.p.

By the Skorohod coupling theorem [27, Theorem 4.30], we may assume that the random graphs for different n are coupled such that the limits in Theorem 2.4(i) hold a.s. Let $\tilde{\mathbf{d}} := \mathbf{d}|_{[n]\setminus \mathcal{C}_1^\circ}$ be the degree sequence used to define $\widetilde{G}^*(n, \mathbf{d})$, let $\widetilde{n} := |\widetilde{G}^*(n, \mathbf{d})|$ be its length and let $\widetilde{n}_k$ be the number of elements $d_i = k$ in it. Then $\widetilde{n} = n - |\mathcal{C}_1|$ and $\widetilde{n}_k = n_k - v_k(\mathcal{C}_1)$, and thus, by Theorem 2.4 with the assumed coupling, $\widetilde{n}/n \xrightarrow{\text{a.s.}} g(\varkappa)$ and $\widetilde{n}_k/n \xrightarrow{\text{a.s.}} p_k \varkappa^k$, $k \ge 0$, and thus $\widetilde{n}_k/\widetilde{n} \xrightarrow{\text{a.s.}} \widehat{p}_k := p_k \varkappa^k / g(\varkappa)$. Consequently, conditioned on the order and degree sequence of $\widetilde{G}^*(n, \mathbf{d})$, Condition 2.1 then holds a.s., with p_k replaced by $\widehat{p}_k$.

The formulas (6.2)–(6.4) are straightforward.

Finally, since $G(n, \mathbf{d})$ is supercritical, $p_k > 0$ for some $k \ge 3$, and thus $g''(x)$ is strictly increasing. Hence, recalling $g'(1) = \mu_\infty$ and $g'(\varkappa) = \varkappa \mu_\infty$,

$$(1 - \varkappa)\mu_\infty = g'(1) - g'(\varkappa) = \int_\varkappa^1 g''(x)\,\mathrm{d}x > (1 - \varkappa) g''(\varkappa).$$

Consequently, $\mu_\infty > g''(\varkappa)$ and (6.3)–(6.4) show that $\widehat{\mu}_\infty > \widehat{\nu}_\infty$. □

7. Proof of main theorems

Proof of Theorems 1.1 and 3.3. By Remark 2.5, it suffices to prove Theorems 1.1 and 3.3 for $G^*(n,\mathbf{d})$.

Consider first the subcritical case $\nu_\infty < \mu_\infty$. By Corollary 3.2, $\chi_\infty = \chi(\mathfrak{X})$ and $\widehat{\chi}_\infty = \widehat{\chi}(\mathfrak{X})$. By Lemma 4.2(i), if $a < \chi_\infty := \chi(\mathfrak{X})$, then $\chi(G^*(n,\mathbf{d})) > a$ w.h.p., while Lemma 5.2, (2.6) and Condition 2.3 show that

$$\mathbb{E}\,\chi(G^*(n,\mathbf{d})) \le 1 + \frac{\mu_n^2}{(\mu_n-\nu_n)_+} \to 1 + \frac{\mu_\infty^2}{\mu_\infty - \nu_\infty} = \chi_\infty.$$

These upper and lower bounds imply, see Janson and Riordan [26, Lemma 4.2], that $\chi(G^*(n,\mathbf{d})) \xrightarrow{L^1} \chi_\infty$. The same argument holds for $\widehat{\chi}(G^*(n,\mathbf{d}))$, by Lemmas 4.2(ii) and 5.2 together with $\widehat{\chi} \le \chi$.

In the critical case $\nu_\infty = \mu_\infty$, Lemma 4.2 yields $\chi(G^*(n,\mathbf{d})) \ge \widehat{\chi}(G^*(n,\mathbf{d})) > a$ w.h.p. for any finite a, and thus $\chi(G^*(n,\mathbf{d}))$, $\widehat{\chi}(G^*(n,\mathbf{d})) \xrightarrow{\text{p}} \infty$.

In the supercritical case, Lemma 4.2 shows $\chi(G^*(n,\mathbf{d})) \xrightarrow{\text{p}} \infty$. For $\widehat{\chi}$ we consider $G(n,\mathbf{d})$ and note that

$$\widehat{\chi}(G(n,\mathbf{d})) := \frac{1}{n}\sum_{\mathcal{C}_i \ne \mathcal{C}_1} |\mathcal{C}_i|^2 = \frac{n-|\mathcal{C}_1|}{n}\chi\bigl(\widehat{G}(n,\mathbf{d})\bigr). \tag{7.1}$$

Here $(n-|\mathcal{C}_1|)/n \xrightarrow{\text{p}} g(\varkappa)$ by Theorem 2.4, and by Theorem 6.1, we may couple $\widehat{G}(n,\mathbf{d})$ and $\widetilde{G}(n,\mathbf{d})$ such that w.h.p. they coincide and thus

$$\chi\bigl(\widehat{G}(n,\mathbf{d})\bigr) = \chi\bigl(\widetilde{G}(n,\mathbf{d})\bigr) \qquad \text{w.h.p.} \tag{7.2}$$

We may by Theorem 6.1 assume that $(n-|\mathcal{C}_1|)/n \to g(\varkappa)$ a.s. and that Condition 2.1 holds for $\widetilde{G}(n,\mathbf{d})$, conditioned on its order and degree sequence, with p_k replaced by $\widehat{p}_k$. Further, for any constant A,

$$\frac{1}{|\widetilde{G}(n,\mathbf{d})|}\sum_{i\in V(\widetilde{G}(n,\mathbf{d}))} d_i^2 \mathbf{1}[d_i \ge A] \le \frac{1}{n(g(\varkappa)+o(1))}\sum_i d_i^2\mathbf{1}[d_i > A],$$

so the uniform integrability of D_n implies that also Condition 2.3 holds for the random graph $\widetilde{G}(n,\mathbf{d})$ conditioned on its order and degree sequence.

We can thus apply the already proven result for χ to $\widetilde{G}(n,\mathbf{d})$, conditioned on its order and degree sequence, and conclude that

$$\chi\bigl(\widetilde{G}(n,\mathbf{d})\bigr) \xrightarrow{\text{p}} 1 + \frac{\widehat{\mu}_\infty}{1-\widehat{\nu}_\infty/\widehat{\mu}_\infty}. \tag{7.3}$$

By (7.1), (7.2), (7.3), (6.3), (6.4) and (3.14), we thus obtain

$$\widehat{\chi}\bigl(G(n,\mathbf{d})\bigr) \xrightarrow{\text{p}} g(\varkappa)\Bigl(1+\frac{\widehat{\mu}_\infty}{1-\widehat{\nu}_\infty/\widehat{\mu}_\infty}\Bigr) = g(\varkappa)+\frac{\varkappa g'(\varkappa)}{1-g''(\varkappa)/\mu_\infty} = \widehat{\chi}(\mathfrak{X}). \qquad \square$$

8. Approximation

We have assumed Condition 2.1, including convergence of the degree distribution D_n. This is convenient, but it is also interesting to regard the result as an approximation for finite n, without assuming convergence of D_n. For simplicity we consider only χ, leaving the similar but notationally more complicated result for $\widehat{\chi}$ to the reader.

In order to treat convergence to ∞, we let δ be a metric on the compact space $[1, \infty]$, for example $\delta(x, y) := |x^{-1} - y^{-1}|$.

Theorem 8.1. *Suppose that* $\mathbf{d} = (d_i)_1^n$ *are given for* $n \geq 1$ *such that the random variables* D_n *are uniformly square integrable and that* $\liminf_{n\to\infty} \mathbb{P}(D_n = 1) > 0$. *Then*

$$\delta\Big(\chi(G(n, \mathbf{d})),\ 1 + \frac{\mu_n^2}{(\mu_n - \nu_n)_+}\Big) \xrightarrow{\mathrm{p}} 0. \tag{8.1}$$

If further $\mu_n - \nu_n \geq c > 0$, *for some fixed* c, *then also*

$$\mathbb{E}\Big|\chi(G(n, \mathbf{d})) - \Big(1 + \frac{\mu_n^2}{\mu_n - \nu_n}\Big)\Big| \to 0, \tag{8.2}$$

and thus $\chi(G(n, \mathbf{d})) = 1 + \mu_n^2/(\mu_n - \nu_n) + o_{\mathrm{p}}(1)$.

Proof. The uniform square integrability implies that $\sup_n \mathbb{E}\, D_n^2 < \infty$; hence the variables D_n are tight, and we may by considering a subsequence assume that $D_n \xrightarrow{\mathrm{d}} D_\infty$ for some random variable D_∞ on $\mathbb{Z}_{\geq 0}$. However, this is exactly Condition 2.1(i); furthermore $\mathbb{P}(D_\infty = 1) > 0$, and the uniform square integrability of D_n implies that Conditions 2.1 and 2.3 hold (along the subsequence). In particular, $\mu_n \to \mu_\infty$ and $\nu_n \to \nu_\infty$.

If $\nu_\infty < \mu_\infty$, then, by Theorem 1.1 applied to the subsequence, $\chi(G(n, \mathbf{d})) \xrightarrow{L^1} 1 + \mu_\infty^2/(\mu_\infty - \nu_\infty)$. Further, $\mu_n^2/(\mu_n - \nu_n)_+ \to \mu_\infty^2/(\mu_\infty - \nu_\infty)$, and both claims follow.

If $\nu_\infty \geq \mu_\infty$, we have $\chi(G(n, \mathbf{d})) \xrightarrow{\mathrm{p}} \infty$ by Theorem 1.1; further, $\mu_n \to \mu_\infty > 0$ and $(\mu_n - \nu_n)_+ \to 0$, and thus $\mu_n^2/(\mu_n - \nu_n)_+ \to \infty$, and (8.1) follows in this case too.

Hence, there is always a subsequence along which the results hold. Since we may start by taking an arbitrary subsequence, the results hold generally by the standard subsubsequence principle, see e.g. [23, p. 12]. $\square$

9. Approaching criticality

In order to study the critial behaviour more closely, we consider a family of random graphs parametrized by a parameter λ besides n. We consider asymptotics as $n \to \infty$, and investigate how the limits depend on λ. More precisely, we consider for simplicity the following case.

Let $h(x)$ be the probability generating function of a non-negative integer-valued random variable D such that $\infty > h''(1) = \mathbb{E}\, D(D-1) > h'(1) = \mathbb{E}\, D$; this thus corresponds to a supercritical $G(n, \mathbf{d})$. Now, for a fixed $\lambda \in (0, 1]$, add an even number $n(1-\lambda)/\lambda + O(1)$ of vertices of degree 1;

this gives a random graph $G(n_\lambda, \mathbf{d}_\lambda)$ with a corresponding asymptotic degree distribution D_∞ that has the probability generating function

$$g_\lambda(x) = (1-\lambda)x + \lambda h(x). \tag{9.1}$$

In particular (we will often omit the argument λ from the notation),

$$\mu_\infty = \mu_\infty(\lambda) = g'_\lambda(1) = (1-\lambda) + \lambda h'(1), \tag{9.2}$$

$$\nu_\infty = \nu_\infty(\lambda) = g''_\lambda(1) = \lambda h''(1). \tag{9.3}$$

Consequently, the random graph $G(n_\lambda, \mathbf{d}_\lambda)$ is critical when $1-\lambda+\lambda h'(1) = \lambda h''(1)$, i.e., when $\lambda = \lambda_{\mathsf{c}}$ given by

$$\lambda_{\mathsf{c}} := \frac{1}{1-h'(1)+h''(1)} \in (0,1), \tag{9.4}$$

while it is subcritical for $\lambda < \lambda_{\mathsf{c}}$ and supercritical for $\lambda > \lambda_{\mathsf{c}}$.

We consider the limits $\chi_\infty(\lambda)$ and $\widehat{\chi}_\infty(\lambda)$ of $\chi(G(n_\lambda, \mathbf{d}_\lambda))$ and $\widehat{\chi}(G(n_\lambda, \mathbf{d}_\lambda))$ as $n \to \infty$, and investigate how they depend on λ as $\lambda \to \lambda_{\mathsf{c}}$. (We thus let first $n \to \infty$ and then $\lambda \to \lambda_{\mathsf{c}}$. A related problem, not considered here, is to let $\lambda \to \lambda_{\mathsf{c}}$ and $n \to \infty$ simultaneously.)

For the subcritical case $\lambda < \lambda_{\mathsf{c}}$ we have, by Theorem 1.1 and (9.2)–(9.4),

$$\chi_\infty = 1 + \frac{\mu_\infty^2}{\mu_\infty - \nu_\infty} = 1 + \frac{\mu_\infty(\lambda)^2}{1-\lambda+\lambda h'(1) - \lambda h''(1)} = 1 + \frac{\mu_\infty(\lambda)^2}{1-\lambda/\lambda_{\mathsf{c}}}.$$

For $\lambda \nearrow \lambda_{\mathsf{c}}$, it follows that, with $\mu_{\mathsf{c}} := \mu_\infty(\lambda_{\mathsf{c}}) = \nu_\infty(\lambda_{\mathsf{c}}) > 0$,

$$\chi_\infty \sim \frac{\lambda_{\mathsf{c}}\mu_\infty(\lambda_{\mathsf{c}})^2}{\lambda_{\mathsf{c}} - \lambda} = \frac{\lambda_{\mathsf{c}}\mu_{\mathsf{c}}^2}{\lambda_{\mathsf{c}} - \lambda}. \tag{9.5}$$

In the supercritical case $\lambda > \lambda_{\mathsf{c}}$, the parameter $\varkappa = \varkappa(\lambda)$ is given by $g'_\lambda(\varkappa) = \varkappa g'_\lambda(1)$, or, by (9.1),

$$1 - \lambda + \lambda h'(\varkappa) = \varkappa(1-\lambda+\lambda h'(1)); \tag{9.6}$$

equivalently,

$$1 - \varkappa = \lambda(1 - h'(\varkappa) - \varkappa + \varkappa h'(1)). \tag{9.7}$$

We use $\varepsilon := 1 - \varkappa$ as parameter, and have thus $\varkappa = 1-\varepsilon$ and, by (9.7) and (9.4),

$$\frac{1}{\lambda} = \frac{\varepsilon + h'(1) - \varepsilon h'(1) - h'(1-\varepsilon)}{\varepsilon} = 1 - h'(1) + \frac{h'(1)-h'(1-\varepsilon)}{\varepsilon},$$

$$\frac{1}{\lambda_{\mathsf{c}}} - \frac{1}{\lambda} = h''(1) - \frac{h'(1)-h'(1-\varepsilon)}{\varepsilon}. \tag{9.8}$$

As $\lambda \searrow \lambda_{\mathsf{c}}$, $\varkappa \nearrow 1$ and thus $\varepsilon = 1 - \varkappa \searrow 0$. Moreover, $g'_\lambda(\varkappa) - \varkappa g''_\lambda(\varkappa) \to g'_{\lambda_{\mathsf{c}}}(1) - g''_{\lambda_{\mathsf{c}}}(1) = 0$, and thus by Theorem 1.1(iii) and $g'_\lambda(\varkappa) = \varkappa g'_\lambda(1)$,

$$\widehat{\chi}_\infty \sim \frac{\varkappa(g'_\lambda(\varkappa))^2}{g'_\lambda(\varkappa) - \varkappa g''_\lambda(\varkappa)} = \frac{(g'_\lambda(\varkappa))^2}{g'_\lambda(1) - g''_\lambda(\varkappa)} \sim \frac{\mu_{\mathsf{c}}^2}{g'_\lambda(1) - g''_\lambda(\varkappa)}, \tag{9.9}$$

where by (9.1)

$$g'_\lambda(1) - g''_\lambda(\varkappa) = 1 - \lambda + \lambda h'(1) - \lambda h''(\varkappa) = \lambda\Big(\frac{1}{\lambda} - \frac{1}{\lambda_{\mathsf{c}}} + h''(1) - h''(\varkappa)\Big). \tag{9.10}$$

Let $r(\varepsilon)$ be the remainder term in the Taylor expansion

$$h(1-\varepsilon) = 1 - \varepsilon h'(1) + \frac{\varepsilon^2}{2} h''(1) - r(\varepsilon),$$

and note that $r(0) = r'(0) = r''(0) = 0$, while $r'''(\varepsilon) = h'''(1-\varepsilon)$. Then (9.8) and (9.10) can be written

$$\frac{1}{\lambda_{\mathsf{c}}} - \frac{1}{\lambda} = \frac{r'(\varepsilon)}{\varepsilon}, \tag{9.11}$$

$$g'_\lambda(1) - g''_\lambda(\varkappa) \sim \lambda\Big(-\frac{r'(\varepsilon)}{\varepsilon} + r''(\varepsilon)\Big). \tag{9.12}$$

These yield, using (9.9),

$$\lambda - \lambda_{\mathsf{c}} \sim \lambda_{\mathsf{c}}^2 \frac{r'(\varepsilon)}{\varepsilon},$$

$$\widehat{\chi}_\infty \sim \frac{\mu_{\mathsf{c}}^2}{g'_\lambda(1) - g''_\lambda(\varkappa)} \sim \lambda_{\mathsf{c}}^{-1} \mu_{\mathsf{c}}^2 \Big(-\frac{r'(\varepsilon)}{\varepsilon} + r''(\varepsilon)\Big)^{-1}$$

and thus, as $\lambda \searrow \lambda_{\mathsf{c}}$,

$$(\lambda - \lambda_{\mathsf{c}})\widehat{\chi}_\infty \sim \lambda_{\mathsf{c}} \mu_{\mathsf{c}}^2 \frac{r'(\varepsilon)}{\varepsilon r''(\varepsilon) - r'(\varepsilon)}. \tag{9.13}$$

Since $r''(\varepsilon) = \int_0^\varepsilon r'''(t)\,\mathrm{d}t$ and $r'(\varepsilon) = \int_0^\varepsilon (\varepsilon - t) r'''(t)\,\mathrm{d}t$, (9.13) can be written

$$\widehat{\chi}_\infty \sim \frac{\lambda_{\mathsf{c}} \mu_{\mathsf{c}}^2}{\lambda - \lambda_{\mathsf{c}}} \cdot \frac{\int_0^\varepsilon (\varepsilon - t) r'''(t)\,\mathrm{d}t}{\int_0^\varepsilon t r'''(t)\,\mathrm{d}t}. \tag{9.14}$$

If $\mathbb{E}\, D^3 < \infty$, then with $b := h'''(1) = \mathbb{E}\, D(D-1)(D-2) > 0$ (because h is supercritical), as $\varepsilon \to 0$, $r'''(\varepsilon) \sim b$, $r''(\varepsilon) \sim b\varepsilon$ and $r'(\varepsilon) \sim b\varepsilon^2/2$; thus (9.13) or (9.14) yields, as $\lambda \searrow \lambda_{\mathsf{c}}$

$$\widehat{\chi}_\infty \sim \frac{\lambda_{\mathsf{c}} \mu_{\mathsf{c}}^2}{\lambda - \lambda_{\mathsf{c}}}. \tag{9.15}$$

Combining (9.15) and (9.5), we obtain as $\lambda \to \lambda_{\mathsf{c}}$ from any side the symmetric asymptotic, assuming $\mathbb{E}\, D^3 < \infty$,

$$\widehat{\chi}_\infty \sim \frac{\lambda_{\mathsf{c}} \mu_{\mathsf{c}}^2}{|\lambda - \lambda_{\mathsf{c}}|}. \tag{9.16}$$

(Hence, in terminology from percolation theory and mathematical physics, the critical exponent γ equals 1.) Such symmetry between the subcritical and supercritical sides has been observed in many different models, but there are also exceptions: for example, for the CHKNS model, $\widehat{\chi}$ has finite limits as the parameter λ increases or decreases to the critical value, but the limits are different and the derivative is finite on the supercritical side but not on the subcritical side where there is a square-root singularity, see Callaway, Hopcroft, Kleinberg, Newman and Strogatz [11] and [26, Section 6.3]. We shall see that also for $G(n,\mathbf{d})$ it is possible to have asymmetric asymptotics when $\mathbb{E}\,D^3=\infty$. We consider some examples, with less and less integrability of D beyond $\mathbb{E}\,D^2<\infty$, which we always assume. We let for convenience $c_1,c_2,\dots$ denote some positive constants (depending on D) whose values we do not want to write explicitly.

Example 9.1. Let $\mathbb{P}(D=k)\sim k^{-3-\alpha}$ as $k\to\infty$, for some $\alpha\in(0,1)$. Then $r'''(\varepsilon)\sim\Gamma(1-\alpha)\varepsilon^{\alpha-1}$ as $\varepsilon\to0$, and thus (9.14) yields, as $\lambda\searrow\lambda_{\mathsf{c}}$,

$$\widehat{\chi}_\infty\sim\frac{\lambda_{\mathsf{c}}\mu_{\mathsf{c}}^2}{\lambda-\lambda_{\mathsf{c}}}\cdot\frac{\int_0^\varepsilon(\varepsilon-t)t^{\alpha-1}\,\mathrm{d}t}{\int_0^\varepsilon t^\alpha\,\mathrm{d}t}=\frac{\lambda_{\mathsf{c}}\mu_{\mathsf{c}}^2}{\lambda-\lambda_{\mathsf{c}}}\cdot\frac{\int_0^1(1-t)t^{\alpha-1}\,\mathrm{d}t}{\int_0^1 t^\alpha\,\mathrm{d}t}=\frac{\alpha^{-1}\lambda_{\mathsf{c}}\mu_{\mathsf{c}}^2}{\lambda-\lambda_{\mathsf{c}}},$$

with the same exponent as on the subcritical side but a different constant.

Example 9.2. Let $\mathbb{P}(D=k)\sim k^{-3}(\log k)^{-\alpha}$ as $k\to\infty$, for some $\alpha>1$. Then $r'''(\varepsilon)\sim1/(\varepsilon|\log\varepsilon|^\alpha)$ as $\varepsilon\to0$, and thus

$$\varepsilon r''(\varepsilon)-r'(\varepsilon)=\int_0^\varepsilon tr'''(t)\,\mathrm{d}t\sim\frac{\varepsilon}{|\log\varepsilon|^\alpha},$$

while $r'(\varepsilon)\sim(\alpha-1)^{-1}\varepsilon/|\log\varepsilon|^{\alpha-1}$. Hence, by (9.13),

$$\widehat{\chi}_\infty\sim\frac{c_1}{\lambda-\lambda_{\mathsf{c}}}|\log\varepsilon|.$$

Further, (9.11) yields

$$\lambda-\lambda_{\mathsf{c}}\sim\lambda_{\mathsf{c}}^{-2}\frac{r'(\varepsilon)}{\varepsilon}\sim\frac{c_2}{|\log\varepsilon|^{\alpha-1}}$$

and thus $|\log\varepsilon|\sim c_3(\lambda-\lambda_{\mathsf{c}})^{-1/(\alpha-1)}$. Consequently, as $\lambda\searrow\lambda_{\mathsf{c}}$,

$$\widehat{\chi}_\infty\sim c_4(\lambda-\lambda_{\mathsf{c}})^{-\alpha/(\alpha-1)},$$

with a critical exponent $\alpha/(\alpha-1)>1$, in contrast to the subcritical case (9.5).

Example 9.3. Let $\mathbb{P}(D=k)\sim k^{-3}(\log k)^{-1}(\log\log k)^{-2}$ as $k\to\infty$. Then, $r'''(\varepsilon)\sim\varepsilon^{-1}\log(1/\varepsilon)^{-1}(\log\log(1/\varepsilon))^{-2}$, and thus $r''(\varepsilon)\sim1/(\log\log(1/\varepsilon))$ and $r'(\varepsilon)\sim\varepsilon/(\log\log(1/\varepsilon))$, while

$$\varepsilon r''(\varepsilon)-r'(\varepsilon)=\int_0^\varepsilon tr'''(t)\,\mathrm{d}t\sim\frac{\varepsilon}{\log(1/\varepsilon)\cdot(\log\log(1/\varepsilon))^2}.$$

Hence, (9.13) yields, as $\lambda\searrow\lambda_{\mathsf{c}}$,

$$\widehat{\chi}_\infty\sim\frac{c_5}{\lambda-\lambda_{\mathsf{c}}}\log(1/\varepsilon)\cdot\log\log(1/\varepsilon).\tag{9.17}$$

Furthermore, (9.11) yields

$$\lambda-\lambda_{\mathsf{c}}\sim\lambda_{\mathsf{c}}^2\frac{r'(\varepsilon)}{\varepsilon}\sim\frac{\lambda_{\mathsf{c}}^2}{\log\log(1/\varepsilon)}.$$

and thus $\log(1/\varepsilon) = \exp\bigl((c_6 + o(1))/(\lambda - \lambda_{\mathsf{c}})\bigr)$. Consequently (9.17) yields

$$\widehat{\chi}_\infty = e^{(c_6+o(1))/(\lambda-\lambda_{\mathsf{c}})},$$

with a more rapid growth than any power of $(\lambda - \lambda_{\mathsf{c}})^{-1}$. (The critical exponent is ∞.)

It seems that in this way we can find examples where $\widehat{\chi}_\infty$ grows arbitrarily fast as $\lambda \searrow \lambda_{\mathsf{c}}$.

10. Some counterexamples

We give some examples where Conditions 2.1 and 2.3 are not satisfied, in order to show that the results in general do not hold without these conditions.

Example 10.1. Let $d_i = 1$ for $i \geq 2$, and $d_1 \sim a\sqrt{n}$ for some $a > 0$. Thus $G(n, \mathbf{d})$ has $n-1$ vertices of degree 1 and a single vertex of higher degree d_1. Consequently, the components are a single star with $d_1 + 1$ vertices and $(n_1 - 1 - d_1)/2$ isolated edges. Hence (deterministically),

$$\chi\bigl(G(n, \mathbf{d})\bigr) = \frac{1}{n}\bigl((d_1+1)^2 + 2(n-1-d_1)\bigr) \to a^2 + 2. \tag{10.1}$$

Condition 2.1 holds with $p_1 = 1$ and $p_k = 0$, $k \neq 1$ (i.e., $D_\infty = 1$ a.s.), so $\mu_\infty = 1$ and $\nu_\infty = 0$. Further, $\mu_n = 1 + O(\sqrt{n}) \to \mu_\infty$ but $\nu_n = \frac{1}{n}d_1(d_1 - 1) \to a^2 \neq \nu_\infty$ so Condition 2.3 fails. We have $1 + \mu_\infty^2/(\mu_\infty - \nu_\infty) = 2$, and thus the conclusion of Theorem 1.1(i) does not hold. This shows that Theorem 1.1 can fail if Condition 2.3 does not hold.

If further $a < 1$, then

$$1 + \frac{\mu_n^2}{\mu_n - \nu_n} = 1 + \frac{1}{1-a^2} + o(1) = \frac{2-a^2}{1-a^2} + o(1).$$

Consequently, Theorem 8.1 too fails for this example, which shows that the theorem does not hold in general without the assumption of uniform square integrability. Similarly, in the case $a > 1$, $\nu_n > \mu_n$ (at least for large n), and Theorem 8.1 would predict that $\chi(G(n, \mathbf{d})) \xrightarrow{\mathrm{p}} \infty$, while (10.1) shows that in fact it has a finite limit.

Example 10.2. Modify Example 10.1 by taking two vertices with high degree, for example $d_1 = d_2 \sim a\sqrt{n}$ and $d_i = 1$ for $i \geq 3$, for some $a > 0$. Thus the components of $G(n, \mathbf{d})$ are either (when there is no edge 12) two stars of order $d_1 + 1 \sim a\sqrt{n}$ plus $n/2 + o(n)$ isolated edges, or (when there is an edge 12) one component of order $2d_1 \sim 2a\sqrt{n}$ plus $n/2 + o(n)$ isolated edges. Both events occur with positive limiting probabilities. (In fact, a simple calculation of the number of labelled graphs of the two types shows that the probability of an edge 12 converges to $a^2/(a^2 + 2)$.) Consequently, $\chi(G(n, \mathbf{d}))$ is either $4a^2 + 2 + o(1)$ or $2a^2 + 2 + o(1)$, and $\chi(G(n, \mathbf{d}))$ converges (in distribution) to a two-point distribution and not to a constant. Similarly, $\widehat{\chi}(G(n, \mathbf{d}))$ is either $2 + o(1)$ or $a^2 + 2 + o(1)$, and again there is a limiting two-point distribution. Consequently, the conclusions of Theorem 1.1(i) fail for both χ and $\widehat{\chi}$.

As in Example 10.1, Condition 2.1 holds with $p_1 = 1$ but Condition 2.3 fails.

In the remaining examples $p_1 = 0$, and thus Condition 2.1(iv) does not hold.

Example 10.3. The random 2-regular graph $G(n, 2)$ is critical, with $p_2 = 1$ and $p_k = 0$, $k \neq 2$ (i.e., $D_\infty = 2$ a.s.). It is well-known that $G(n, 2)$ has w.h.p. several large components with sizes $\Theta_p(n)$, and it follows that $\chi(G(n, 2)), \widehat{\chi}(G(n, 2)) \xrightarrow{\text{p}} \infty$, in accordance with Theorem 1.1(ii).

Now perturb this example by adding a suitable number of vertices of degree 4, say $n_4 = \lfloor n^{0.9} \rfloor$ and $n_2 = n - n_4$. By ignoring all vertices of degree 2 in $G^*(n, \mathbf{d})$(contracting their adjacent edges to a single edge), we obtain a random 4-regular multigraph, which w.h.p. is connected, cf. Remark 2.6. Hence, w.h.p. all vertices of degree 4 belong to a single component. Moreover, by splitting each vertex of degree 4 into two vertices of degree 2, considering the (2-regular) configuration model for these vertices, and then recombining the vertices of degree 4, it is easily seen that w.h.p. this is the giant component $\mathcal{C}_1$ and that it contains all vertices except a small number of cycles with a total size $\leq n^{0.2}$, say. Thus w.h.p. $\widehat{\chi}(G(n, 2)) \leq n^{0.4-1}$ so $\widehat{\chi}(G(n, 2)) \xrightarrow{\text{p}} 0$, although this example is critical, and so Theorem 1.1(ii) fails for it.

Example 10.4. If $p_0 = 1$ (i.e., $D_\infty = 0$ a.s.), then Condition 2.1(iv) and (ii) do not hold. We have $\mu_\infty = 0$ and $|\mathfrak{X}| = 1$ a.s., so $\chi_\infty = 1$. However, $\chi(G(n, \mathbf{d})) \xrightarrow{\text{p}} 1$ may fail.

For example, let $d_i = 3$ for $i \leq n_3$ and $d_i = 0$ for $n_3 < i \leq n$, where we choose $n_3 := 2\lfloor a\sqrt{n} \rfloor$ for some $a > 0$. Then Conditions 2.1 and 2.3 hold except for Condition 2.1(ii),(iv), with $p_0 = 1$. $G(n, \mathbf{d})$ consists of $n_0 = n - n_3$ isolated vertices together with a random cubic graph on n_3 vertices. The latter is w.h.p. connected, see Remark 2.6, and thus w.h.p.

$$\chi\big(G(n, \mathbf{d})\big) = \frac{1}{n}(n - n_3 + n_3^2) \to 1 + 4a^2 > 1.$$

Hence, Theorem 1.1 fails for this example.

References

[1] K. B. Athreya & P. E. Ney, *Branching Processes*, Springer-Verlag, Berlin, 1972.

[2] E. A. Bender & E. R. Canfield, The asymptotic number of labeled graphs with given degree sequences. *J. Combin. Theory Ser. A* **24** (1978), no. 3, 296–307.

[3] B. Bollobás, A probabilistic proof of an asymptotic formula for the number of labelled regular graphs, *European J. Comb.* **1** (1980), 311–316.

[4] B. Bollobás, Random graphs. *Combinatorics (Swansea, 1981)*, pp. 80–102, London Math. Soc. Lecture Note Ser., 52, Cambridge Univ. Press, Cambridge, 1981.

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

[6] B. Bollobás, *Random Graphs*, 2nd ed., Cambridge University Press, Cambridge, 2001.

[7] B. Bollobás, S. Janson & O. Riordan, The phase transition in inhomogeneous random graphs. *Random Struct. Alg.* **31** (2007), 3–122.

[8] B. Bollobás & O. Riordan, *Percolation.* Cambridge University Press, Cambridge, 2006.

[9] C. Borgs, J. T. Chayes, R. van der Hofstad, G. Slade & J. Spencer, Random subgraphs of finite graphs. I. The scaling window under the triangle condition. *Random Structures Algorithms* **27** (2005), no. 2, 137–184.

[10] T. Britton, S. Janson & A. Martin-Löf, Graphs with specified degree distributions, simple epidemics and local vaccination strategies. *Advances Appl. Probab.* **39** (2007), no. 4, 922–948.

[11] D. S. Callaway, J. E. Hopcroft, J. M. Kleinberg, M. E. J. Newman & S. H. Strogatz, Are randomly grown graphs really random? *Phys. Rev. E* **64** (2001), 041902.

[12] L. Chayes & S. A. Smith, Layered percolation on the complete graph. Preprint (2009). `http://www.math.ucla.edu/~lchayes/`

[13] R. Durrett, *Random Graph Dynamics.* Cambridge University Press, Cambridge, 2007.

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

[15] G. H. Hardy, J. E. Littlewood & G. Pólya, *Inequalities.* 2nd ed., Cambridge, at the University Press, 1952.

[16] R. van der Hofstad & G. Slade, Asymptotic expansions in n^{-1} for percolation critical values on the n-cube and $\mathbb{Z}^n$. *Random Structures Algorithms* **27** (2005), no. 3, 331–357.

[17] R. van der Hofstad & 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), no. 5, 695–713.

[18] S. Janson, The probability that a random multigraph is simple. *Combin. Probab. Comput.* **1**8 (2009), 205–225.

[19] S. Janson, Probability asymptotics: notes on notation. Institut Mittag-Leffler preprint 31, 2009 spring.

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

[21] S. Janson & M. Luczak, A new approach to the giant component problem. *Random Struct. Alg.* **3**4 (2008), 197–216.

[22] S. Janson & M. Luczak, Susceptibility in subcritical random graphs. *J. Math. Phys.* **49**:12 (2008), 125207.

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

[24] S. Janson & J. Spencer, Phase transitions for modified Erdős-Rényi processes. `arXiv:1005.4494.`

[25] S. Janson & O. Riordan, Duality in inhomogeneous random graphs, and the cut metric. *Random Struct. Alg.*, to appear.

[26] S. Janson & O. Riordan, Susceptibility in inhomogeneous random graphs. `arXiv:0905.0437.`

[27] O. Kallenberg, *Foundations of Modern Probability*, 2nd ed., Springer-Verlag, New York, 2002.

[28] T. Łuczak, Component behavior near the critical point of the random graph process, *Random Struct. Alg.* **1** (1990), 287–310.

[29] B. D. McKay, Asymptotics for symmetric 0–1 matrices with prescribed row sums. *Ars Combin.* **19A** (1985), 15–25.

[30] B. D. McKay & N. C. Wormald, Asymptotic enumeration by degree sequence of graphs with degrees $o(n^{1/2})$. *Combinatorica* **11** (1991), no. 4, 369–382.

[31] M. Molloy & B. Reed, A critical point for random graphs with a given degree sequence, *Random Struct. Alg.* **6** (1995), no. 2–3, 161–179.

[32] M. Molloy & B. Reed, The size of the giant component of a random graph with a given degree sequence. *Combin. Probab. Comput.* **7** (1998), 295–305.

[33] M. E. J. Newman, S. H. Strogatz, and D. J. Watts, Random graphs with arbitrary degree distributions and their applications. *Phys. Rev. E* **64** (2001), 026118.

[34] J. Spencer & N. Wormald, Birth control for giants. *Combinatorica* **27** (2007), 587–628.

[35] N. C. Wormald, Some problems in the enumeration of labelled graphs. Ph. D. thesis, University of Newcastle, 1978.

[36] N. C. Wormald, The asymptotic distribution of short cycles in random regular graphs. *J. Combin. Theory Ser. B* **31** (1981), no. 2, 168–182.

[37] N. C. Wormald, The asymptotic connectivity of labelled regular graphs. *J. Combin. Theory Ser. B* **31** (1981), no. 2, 156–167.

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

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

URL: `http://www.math.uu.se/∼svante/`