RANDOM GRAPHS WITH FORBIDDEN VERTEX DEGREES

GEOFFREY GRIMMETT AND SVANTE JANSON

Abstract. We study the random graph $G_{n,\lambda/n}$ conditioned on the event that all vertex degrees lie in some given subset $\mathcal{S}$ of the non-negative integers. Subject to a certain hypothesis on $\mathcal{S}$, the empirical distribution of the vertex degrees is asymptotically Poisson with some parameter $\widehat{\mu}$ given as the root of a certain 'characteristic equation' of $\mathcal{S}$ that maximizes a certain function $\psi_{\mathcal{S}}(\mu)$. Subject to a hypothesis on $\mathcal{S}$, we obtain a partial description of the structure of such a random graph, including a condition for the existence (or not) of a giant component. The requisite hypothesis is in many cases benign, and applications are presented to a number of choices for the set $\mathcal{S}$ including the sets of (respectively) even and odd numbers.

1. Introduction

Let $\mathcal{S}$ be a fixed nonempty set of non-negative integers. The purpose of this paper is to study the structure of random graphs having all their vertex degrees restricted to the set $\mathcal{S}$.

We call a graph an *$\mathcal{S}$-graph* if all its vertex degrees belong to $\mathcal{S}$. For example, if $\mathcal{S} = \{s\}$ is a singleton, an $\mathcal{S}$-graph is the same as a regular graph of degree s. (We are not going to say anything new about this case.) One of our main examples is the class of *Eulerian graphs*, or *even graphs*, given by the set of even numbers $\mathcal{S} = 2\mathbb{Z}_{\geq 0}$, with $\mathbb{Z}_{\geq 0}$ the set $\{0, 1, 2, \dots\}$ of non-negative integers. See Section 6 for further examples.

More precisely, we will study the random graph $G_{n,p;\mathcal{S}}$ defined as $G_{n,p}$ *conditioned on being an $\mathcal{S}$-graph*, where $G_{n,p}$ is the standard random subgraph of the (labelled) complete graph K_n where two vertices are joined by an edge with probability $p \in (0,1)$, and these $\binom{n}{2}$ events, corresponding to the edges of K_n, are independent. In other words, $G_{n,p;\mathcal{S}}$ is a random $\mathcal{S}$-subgraph of K_n such that, if G is any given subgraph of K_n that is an $\mathcal{S}$-graph, then

$$\mathbb{P}(G_{n,p;\mathcal{S}} = G) = \frac{p^{\mathrm{e}(G)}(1-p)^{\binom{n}{2}-\mathrm{e}(G)}}{\mathbb{P}(G_{n,p} \text{ is an } \mathcal{S}\text{-graph})}, \tag{1.1}$$

Date: December 1, 2007; revised June 28, 2009.
2000 *Mathematics Subject Classification.* 05C80, 05C07, 60C05.
Key words and phrases. Random graph, even graph, random-cluster model.

where $\mathrm{e}(G)$ is the number of edges of G. We are interested in asymptotics as $n \to \infty$, and we will tacitly consider only n such that there exists an $\mathcal{S}$-graph with n vertices, in other words, such that the denominator in (1.1) is non-zero. Thus, a finite number of small n may be excluded; moreover, if $\mathcal{S}$ contains only odd integers, then n has to be even. (It is easy to see that, apart from this parity restriction, all large n are allowed.)

For example, with $\mathcal{S}$ the set of even numbers, we obtain a random even graph, see Example 6.3. This is related to the random-cluster model on the complete graph K_n [6; 7], and was one of our motivations for the present study.

Remark 1.1. The choice $p = \frac{1}{2}$ gives a random $\mathcal{S}$-graph that is uniformly distributed over all $\mathcal{S}$-graphs on n labelled vertices. However, we will in this paper instead study the case when p is of order $1/n$ and the average vertex degree is bounded (for $G_{n,p}$, and as we shall see later, for $G_{n,p;\mathcal{S}}$ too).

It follows immediately from (1.1) that two $\mathcal{S}$-subgraphs of K_n with the same degree sequence are attained with the same probability. Hence, the conditional distribution of $G_{n,p;\mathcal{S}}$ given the degree sequence is uniform. We will therefore focus on studying the random degree sequence of $G_{n,p;\mathcal{S}}$; it is then possible to obtain further results on the structure of $G_{n,p;\mathcal{S}}$ by applying standard results on random graphs with given degree sequences to $G_{n,p;\mathcal{S}}$ conditioned on the degree sequence. For example, using the results by Molloy and Reed [14, 15] we obtain Theorem 3.1 below on existence of a giant component in $G_{n,p;\mathcal{S}}$.

2. Main theorem

By symmetry, the labelling of the vertices and thus the order of the degree sequence is not important, and we shall therefore study the numbers of vertices with given degrees, rather than the degree sequence itself. We introduce some notation.

Let $\mathcal{N}$ be the set of sequences $\mathbf{n} = (n_0, n_1, \dots)$ of non-negative integers $n_j \in \mathbb{Z}_{\geq 0}$ with only a finite number of non-zero terms n_j. Let

$$\mathcal{N}_{\mathcal{S}} := \{\mathbf{n} \in \mathcal{N} : n_j = 0 \text{ when } j \notin \mathcal{S}\},$$

the set of such sequences supported on $\mathcal{S}$. For a (multi)graph G, let $n_j(G)$ be the number of vertices of degree j in G, $j \in \mathbb{Z}_{\geq 0}$, and let $\mathbf{n}(G) := (n_j(G))_{j=0}^{\infty}$ be the sequence of degree counts. Thus, G is an $\mathcal{S}$-graph if and only if $\mathbf{n}(G) \in \mathcal{N}_{\mathcal{S}}$. Clearly, cf. (1.1),

$$\mathbb{P}(G_{n,p} \text{ is an } \mathcal{S}\text{-graph}) = \sum_{G:\, \mathbf{n}(G) \in \mathcal{N}_{\mathcal{S}}} p^{\mathrm{e}(G)} (1-p)^{\binom{n}{2} - \mathrm{e}(G)}. \tag{2.1}$$

Using a term that is standard in statistical physics, we shall call this summation the *partition function*, denoted as $Z_{n,p;\mathcal{S}}$.

Let $\mathcal{P}$ be the set of probability distributions on $\mathbb{Z}_{\geq 0}$. In other words, $\mathcal{P}$ is the set of sequences $\boldsymbol{\pi} = (\pi_0, \pi_1, \dots)$ of non-negative real numbers such that $\sum_j \pi_j = 1$. We regard $\mathcal{P}$ as a topological space with the usual topology of weak convergence (denoted $\overset{\mathrm{p}}{\longrightarrow}$); it is well known that this topology on $\mathcal{P}$ may be metrised by the total variation distance

$$d_{\mathrm{TV}}(\boldsymbol{\pi}, \boldsymbol{\pi}') := \tfrac{1}{2} \sum_j |\pi_j - \pi'_j|.$$

If G has $n = \sum_j n_j(G)$ vertices, let $\pi_j(G) := n_j(G)/n$, the proportion of vertices of degree j, and $\boldsymbol{\pi}(G) := \mathbf{n}(G)/n = (\pi_j(G))_{j=0}^{\infty}$. Note that $\boldsymbol{\pi}(G)$ is the probability distribution of the degree of a randomly chosen vertex in G.

Let $\phi_{\mathcal{S}}$ be the exponential generating function of $\mathcal{S}$,

$$\phi_{\mathcal{S}}(\mu) := \sum_{j \in \mathcal{S}} \frac{\mu^j}{j!}. \tag{2.2}$$

Note that this is an entire function of μ, and that $\phi_{\mathcal{S}}(\mu) > 0$ for $\mu > 0$ while $\phi_{\mathcal{S}}(0) > 0$ if and only if $0 \in \mathcal{S}$.

Let $\mathrm{Po}_{\mathcal{S}}(\mu)$ be the distribution of a $\mathrm{Po}(\mu)$ distributed variable given that it belongs to $\mathcal{S}$, i.e., $\mathcal{L}(X_\mu \mid X_\mu \in \mathcal{S})$ with $X_\mu \sim \mathrm{Po}(\mu)$. Thus, recalling (2.2),

$$\mathrm{Po}_{\mathcal{S}}(\mu)\{k\} = \frac{\mu^k/k!}{\phi_{\mathcal{S}}(\mu)}, \qquad k \in \mathcal{S}. \tag{2.3}$$

This conditional distribution is always defined for $\mu > 0$, and in the case $0 \in \mathcal{S}$ for $\mu = 0$ too (in which case it is a point mass at 0). The mean of the $\mathrm{Po}_{\mathcal{S}}(\mu)$ distribution is

$$\mathbb{E}(X_\mu \mid X_\mu \in \mathcal{S}) = \frac{1}{\phi_{\mathcal{S}}(\mu)} \sum_{k \in \mathcal{S}} \frac{k\mu^k}{k!} = \frac{\mu \phi'_{\mathcal{S}}(\mu)}{\phi_{\mathcal{S}}(\mu)}. \tag{2.4}$$

Let $\lambda > 0$. We shall refer to the equation

$$\frac{\mu \phi'_{\mathcal{S}}(\mu)}{\phi_{\mathcal{S}}(\mu)} = \frac{\mu^2}{\lambda}, \tag{2.5}$$

as the *characteristic equation* of the set $\mathcal{S}$ (for this value of λ), and we write

$$E(\lambda) := \Big\{\mu \geq 0 : \frac{\mu \phi'_{\mathcal{S}}(\mu)}{\phi_{\mathcal{S}}(\mu)} = \frac{\mu^2}{\lambda}\Big\}, \tag{2.6}$$

where we allow $\mu = 0$ only if $0 \in \mathcal{S}$. We further define the auxiliary function

$$\psi_{\mathcal{S}}(\mu) := \log \phi_{\mathcal{S}}(\mu) - \frac{\mu \phi'_{\mathcal{S}}(\mu)}{2\phi_{\mathcal{S}}(\mu)} \tag{2.7}$$

and note that when $\mu \in E(\lambda)$, $\psi_{\mathcal{S}}(\mu)$ equals the simpler function

$$\psi_{\mathcal{S},1}(\mu; \lambda) := \log \phi_{\mathcal{S}}(\mu) - \frac{\mu^2}{2\lambda}. \tag{2.8}$$

All logarithms in this paper are natural. We let $c_1, C_1, \dots$ denote positive constants, generally depending on $\mathcal{S}$ and λ (or (λ_n)) and sometimes on other parameters too (but not on n), which may be indicated by arguments. (We use C_i for large and c_i for small constants.) We sometimes assume that $n > 1$ to avoid trivialities.

Theorem 2.1. *Let $\lambda_n \to \lambda > 0$ and suppose that $E(\lambda)$ contains a unique $\widehat{\mu} = \widehat{\mu}(\lambda)$ that maximizes $\psi_{\mathcal{S}}(\mu)$ (or, equivalently, $\psi_{\mathcal{S},1}(\mu;\lambda)$) over $E(\lambda)$. Then, the following hold, as $n \to \infty$:*

(i) $\boldsymbol{\pi}(G_{n,\lambda_n/n;\mathcal{S}}) \xrightarrow{\text{p}} \mathrm{Po}_{\mathcal{S}}(\widehat{\mu})$. *In other words, for every $j \in \mathcal{S}$,*

$$\frac{n_j(G_{n,\lambda_n/n;\mathcal{S}})}{n} \xrightarrow{\text{p}} \mathrm{Po}_{\mathcal{S}}(\widehat{\mu})\{j\} = \frac{\widehat{\mu}^j/j!}{\phi_{\mathcal{S}}(\widehat{\mu})}. \tag{2.9}$$

(ii) *All moments of the random distribution $\boldsymbol{\pi}(G_{n,\lambda_n/n;\mathcal{S}})$ converge to the corresponding moments of $\mathrm{Po}_{\mathcal{S}}(\widehat{\mu})$. In other words, if $d_1, \dots, d_n$ is the degree sequence of $G_{n,\lambda_n/n;\mathcal{S}}$, and $X_\mu \sim \mathrm{Po}(\mu)$, then for every $r \in (0,\infty)$,*

$$\begin{aligned}\sum_{i=1}^{n} \frac{d_i^r}{n} &= \sum_{j=0}^{\infty} \frac{j^r n_j(G_{n,\lambda_n/n;\mathcal{S}})}{n} \\ &\xrightarrow{\text{p}} \sum_{j=0}^{\infty} j^r \,\mathrm{Po}_{\mathcal{S}}(\widehat{\mu})\{j\} = \mathbb{E}(X_{\widehat{\mu}}^r \mid X_{\widehat{\mu}} \in \mathcal{S}).\end{aligned} \tag{2.10}$$

In particular,

$$\frac{\mathrm{e}(G_{n,\lambda_n/n;\mathcal{S}})}{n} \xrightarrow{\text{p}} \tfrac{1}{2}\,\mathbb{E}(X_{\widehat{\mu}} \mid X_{\widehat{\mu}} \in \mathcal{S}) = \frac{\widehat{\mu}\phi'_{\mathcal{S}}(\widehat{\mu})}{2\phi_{\mathcal{S}}(\widehat{\mu})}. \tag{2.11}$$

(iii) *The error probabilities in* (i) *decay exponentially: for every $\varepsilon > 0$, there exists a constant $c_1 = c_1(\varepsilon,\lambda,\mathcal{S}) > 0$ such that, for all large n,*

$$\mathbb{P}\left(d_{\mathrm{TV}}\big(\boldsymbol{\pi}(G_{n,\lambda_n/n;\mathcal{S}}), \mathrm{Po}_{\mathcal{S}}(\widehat{\mu})\big) \geq \varepsilon\right) \leq e^{-c_1 n}. \tag{2.12}$$

(iv) *We have that*

$$\frac{1}{n}\log Z_{n,\lambda_n/n;\mathcal{S}} = \frac{1}{n}\log \mathbb{P}(G_{n,\lambda_n/n} \text{ is an } \mathcal{S}\text{-graph}) \to \psi_{\mathcal{S}}(\widehat{\mu}) - \tfrac{1}{2}\lambda. \tag{2.13}$$

Remark 2.2. The reason for taking n large in (2.12) is as follows. Suppose, for example, that $\mathcal{S}$ is an infinite set. Then the left hand side of (2.12) trivially equals 1 for any fixed n and sufficiently small ε (depending on n). One way around this, at least when $\lambda_n \equiv \lambda$, is to replace the right hand side of (2.12) by $2e^{-c_2 n}$ for a suitable small $c_2 > 0$. The inequality then becomes trivial for small n.

More generally, let $E_0(\lambda)$ be the subset of $E(\lambda)$ where $\psi_{\mathcal{S}}(\mu)$ (or $\psi_{\mathcal{S},1}(\mu;\lambda)$) is maximal:

$$E_0(\lambda) := \left\{\mu \in E(\lambda) : \psi_{\mathcal{S}}(\mu) = \max_{\mu' \in E(\lambda)} \psi_{\mathcal{S}}(\mu')\right\}. \tag{2.14}$$

If $E_0(\lambda)$ contains a single element, we thus take that element as $\widehat{\mu}(\lambda)$; in particular, if $|E(\lambda)| = 1$, then $E(\lambda) = E_0(\lambda) = \{\widehat{\mu}(\lambda)\}$. We shall see in

Section 4 that this is the normal case: $E(\lambda)$ is always finite and non-empty, and $|E_0(\lambda)| = 1$ except for at most a countable number of values of λ.

Further results on $\widehat{\mu}$ and the auxiliary functions are given in Section 4.

Remark 2.3. The set $E_0(\lambda)$ may contain more than one element, see Example 6.8. In this case, the theorem may not be applied, but the proof in Sections 8–9 extends to show that the random degree distribution $\pi(G_{n,\lambda_n/n;\mathcal{S}})$ approaches the finite set $F := \{\mathrm{Po}_{\mathcal{S}}(\mu) : \mu \in E_0(\lambda)\}$ in the sense that the analogue of (2.12) holds for the distance to this set, i.e.,

$$\mathbb{P}\Big(\min_{\mu \in E_0(\lambda)} d_{\mathrm{TV}}\big(\boldsymbol{\pi}(G_{n,\lambda_n/n;\mathcal{S}}), \mathrm{Po}_{\mathcal{S}}(\mu)\big) \geq \varepsilon\Big) \leq e^{-c_1 n}. \tag{2.15}$$

We can regard the distributions in F as pure phases, in analogy with the situation for many infinite systems of interest in statistical physics, but for the finite systems considered here this has to be interpreted asymptotically. Thus, for large n, the degree distribution of $G_{n,\lambda_n/n;\mathcal{S}}$ is approximately given by one of the pure phases, but we do not know which one. It follows that if we let $n \to \infty$, one of the following happens for the random degree distribution $\pi = \pi(G_{n,\lambda_n/n;\mathcal{S}})$ (regarded as an element of $\mathcal{P}$):

(i) π converges in probability to $\mathrm{Po}_{\mathcal{S}}(\mu)$ for some $\mu \in E_0(\lambda)$.
(ii) π converges in distribution to some non-degenerate distribution on F. (A mixture of two or more pure phases.)
(iii) There are oscillations and π does not converge in distribution; suitable subsequences converge as in (i) or (ii), but different subsequences may have different limits.

It is easy to show by a continuity argument that all three cases may occur in Example 6.8 for suitable sequences (λ_n) (with $\lambda_n \to \lambda_0$ defined there). We do not know whether all three cases may occur for fixed λ with $|E_0(\lambda)| > 1$.

We shall not investigate the case $|E_0(\lambda)| > 1$ further here.

We close this section with an informal explanation of the results of Theorem 2.1, several applications of which are presented in Section 6. Recall the partition function $Z_{n,\lambda/n;\mathcal{S}}$ of (2.1), considered as a summation over suitable graphs. We wish to establish which graphs are dominant in this summation. In so doing, we will treat certain discrete variables as continuous, and shall study maxima by differentiation and Lagrange multipliers. Let $\boldsymbol{\pi} = (\pi_0, \pi_1, \dots)$ be a sequence of non-negative reals satisfying $\sum_i \pi_i = 1$, and $\pi_i = 0$ for $i \notin \mathcal{S}$. We write

$$\nu = \nu(\boldsymbol{\pi}) = \sum_i i\pi_i.$$

Let $Z(\boldsymbol{\pi})$ represent the contribution to the summation of (2.1) from graphs G having, for each i, approximately $n\pi_i$ vertices with degree i. The (empirical) mean vertex-degree of such a graph is ν.

Now, $Z(\boldsymbol{\pi})$ is a summation over simple graphs subject to constraints on the vertex degrees. It may be approximated by a similar summation over certain multigraphs, and this is easier to express in closed form, as follows. The number of ways of partitioning n vertices into sets $V_0, V_1, \dots$ of respective sizes $n\pi_i$, $i \geq 0$, is

$$\frac{n!}{(n\pi_0)!(n\pi_1)!\cdots}.$$

Each vertex $v \in V_i$ will be taken to have degree i, and we therefore provide v with i 'half-edges'. Each such half-edge will be connected to some other half-edge to make a whole edge. Since half-edges are considered indistinguishable, we shall require the multiplicative factor

$$\left\{\prod_{i\in\mathcal{S}} (i!)^{-n\pi_i}\right\}$$

in order to treat them as labelled. The total number of half-edges is $2N = n\sum_i i\pi_i = n\nu$, and we assume for simplicity that N is an integer. These half-edges may be paired together in any of

$$(2N-1)!! = (2N-1)(2N-3)\cdots 3\cdot 1 = \frac{(2N)!}{2^N N!}$$

ways, and each such pairing contributes

$$\left(\frac{\lambda}{n}\right)^N \left(1-\frac{\lambda}{n}\right)^{\binom{n}{2}-N}$$

to the sum. We combine the above to obtain an approximation to $Z(\boldsymbol{\pi})$:

$$Z(\boldsymbol{\pi}) \approx \frac{n!}{(n\pi_0)!(n\pi_1)!\cdots}\left\{\prod_{i\in\mathcal{S}} (i!)^{-n\pi_i}\right\}\frac{(2N)!}{2^N (N)!}\left(\frac{\lambda}{n}\right)^N \left(1-\frac{\lambda}{n}\right)^{\binom{n}{2}-N}.$$

By Stirling's formula, as $n \to \infty$,

$$\frac{1}{n}\log Z(\boldsymbol{\pi}) \to \tfrac{1}{2}\nu\log(\lambda\nu/e) - \tfrac{1}{2}\lambda - \sum_{i\in\mathcal{S}} \pi_i \log(i!\,\pi_i). \tag{2.16}$$

We maximize the last expression subject to $\sum_i \pi_i = 1$ to find that

$$\pi_i = A\frac{\mu^i}{i!}, \qquad i \in \mathcal{S}, \tag{2.17}$$

for some constant A and some μ satisfying

$$\mu = \sqrt{\lambda\nu}. \tag{2.18}$$

Thus $\boldsymbol{\pi}$ is the mass function of the $\mathrm{Po}_{\mathcal{S}}(\mu)$ distribution and, by (2.4) and the definition of ν,

$$\nu = \frac{\mu\phi'_{\mathcal{S}}(\mu)}{\phi_{\mathcal{S}}(\mu)}. \tag{2.19}$$

We combine this with (2.18) to obtain the 'characteristic equation' (2.5).

If there exists a unique μ satisfying the characteristic equation, then we are done. If there is more than one, we pick the value that maximizes the right hand side of (2.16). That is to say, the exponential asymptotics of

$Z_{n,\lambda/n;\mathcal{S}}$ are dominated by the contributions from graphs with $\boldsymbol{\pi}$ satisfying (2.17) with μ chosen to satisfy the characteristic equation and to maximize $\psi_{\mathcal{S}}(\mu)$.

Note from (2.16) that

$$\frac{1}{n}\log Z(\boldsymbol{\pi}) \to \psi_{\mathcal{S}}(\mu) - \tfrac{1}{2}\lambda, \tag{2.20}$$

and part (iv) of Theorem (2.1) is explained.

We make the above argument rigorous in the forthcoming proof of Sections 8–9, a substantial part of which is devoted to proving that the conditional distribution of vertex degrees is concentrated near its mode. The 'half-edge' method has been used by many authors dating back to [4] or earlier, and references will be included at appropriate places later.

3. The giant and the core

We show next how to apply Theorem 2.1, in conjunction with results of Molloy and Reed [14, 15] and Janson and Luczak [10, 11], to identify the sizes of the giant cluster and the k-core of $G_{n,\lambda_n/n;\mathcal{S}}$. The proofs are deferred to Section 10.

We consider first the existence or not of a giant component in the random $\mathcal{S}$-graph $G_{n,\lambda_n/n;\mathcal{S}}$ as $\lambda_n \to \lambda > 0$ (i.e., a component of size comparable to n). Let $\boldsymbol{\pi} = (\pi_0, \pi_1, \dots)$ be a vector of non-negative reals with sum 1, and write $\nu = \sum_j j\pi_j$. As explained in [14; 15], if we consider the random graph with a given degree sequence $\mathbf{d} = (d_i)_1^n$, in which there are $n(1+\mathrm{o}(1))\pi_j$ vertices with degree j, the quantity that is key to the existence of a giant component is

$$Q(\boldsymbol{\pi}) := \sum_j j(j-2)\pi_j.$$

Subject to certain conditions, if $Q(\boldsymbol{\pi}) > 0$, there exists a giant component, while there is no giant component when $Q(\boldsymbol{\pi}) \le 0$.

We shall apply this with $\pi_j = \mathrm{Po}_{\mathcal{S}}(\widehat{\mu})\{j\}$, and to that end we introduce some further notation. Let, see (2.4),

$$\nu(\mu) := \sum_j j\,\mathrm{Po}_{\mathcal{S}}(\mu)\{j\} = \frac{\mu\phi'_{\mathcal{S}}(\mu)}{\phi_{\mathcal{S}}(\mu)}, \tag{3.1}$$

$$Q(\mu) := \sum_j j(j-2)\,\mathrm{Po}_{\mathcal{S}}(\mu)\{j\} = \frac{\mu^2\phi''_{\mathcal{S}}(\mu) - \mu\phi'_{\mathcal{S}}(\mu)}{\phi_{\mathcal{S}}(\mu)}, \tag{3.2}$$

$$\chi_\mu(\xi) := \sum_j j\,\mathrm{Po}_{\mathcal{S}}(\mu)\{j\}(\xi^2 - \xi^j) = \nu(\mu)\xi^2 - \mu\xi\frac{\phi'_{\mathcal{S}}(\mu\xi)}{\phi_{\mathcal{S}}(\mu)}. \tag{3.3}$$

Note that $\chi_\mu(0) = \chi_\mu(1) = 0$. Furthermore, the only possibly negative term in the sums in (3.2) and (3.3) for $\xi \in (0,1)$ are those with $j = 1$, while the terms with $j = 0$ and $j = 2$ always vanish and the others are positive unless $\mathrm{Po}_{\mathcal{S}}(\mu) = 0$.

Let $\Gamma_{n,p;\mathcal{S}}$ be the component of $G_{n,p;\mathcal{S}}$ with the largest number of vertices, and let $\Gamma^{(2)}_{n,p;\mathcal{S}}$ be the second largest. (Break ties by any rule.) We write $\mathrm{v}(H)$ for the number of vertices in a graph H.

Theorem 3.1. *Suppose that $\mathcal{S} \not\subseteq \{0,2\}$. Let $\lambda_n \to \lambda > 0$ and suppose that $E_0(\lambda)$ contains a unique element $\widehat{\mu} \geq 0$. Then, $G_{n,\lambda_n/n;\mathcal{S}}$ has a giant component if and only if $Q(\widehat{\mu}) > 0$, i.e., if and only if $\widehat{\mu}\phi''_{\mathcal{S}}(\widehat{\mu}) > \phi'_{\mathcal{S}}(\widehat{\mu})$. More precisely, as $n \to \infty$,*

$$n^{-1}\mathrm{v}(\Gamma_{n,\lambda_n/n;\mathcal{S}}) \xrightarrow{\mathrm{p}} \widehat{\gamma} \geq 0, \qquad n^{-1}\mathrm{e}(\Gamma_{n,\lambda_n/n;\mathcal{S}}) \xrightarrow{\mathrm{p}} \widehat{\zeta} \geq 0,$$
$$n^{-1}\mathrm{v}(\Gamma^{(2)}_{n,\lambda_n/n;\mathcal{S}}) \xrightarrow{\mathrm{p}} 0, \qquad n^{-1}\mathrm{e}(\Gamma^{(2)}_{n,\lambda_n/n;\mathcal{S}}) \xrightarrow{\mathrm{p}} 0,$$

where

$$\widehat{\gamma} = \widehat{\gamma}(\widehat{\mu}) := 1 - \sum_j \widehat{\xi}^j \,\mathrm{Po}_{\mathcal{S}}(\widehat{\mu})\{j\} = 1 - \frac{\phi_{\mathcal{S}}(\widehat{\xi}\widehat{\mu})}{\phi_{\mathcal{S}}(\widehat{\mu})}, \tag{3.4}$$

$$\widehat{\zeta} = \widehat{\zeta}(\widehat{\mu}) := \tfrac{1}{2}\nu(\widehat{\mu})\bigl(1 - \widehat{\xi}^2\bigr) \tag{3.5}$$

with $\widehat{\xi} = \widehat{\xi}(\widehat{\mu}) \in [0,1]$ given as follows:

(i) *if $Q(\widehat{\mu}) > 0$ and $1 \in \mathcal{S}$, then $\widehat{\xi} \in (0,1)$ is the unique solution to $\chi_{\widehat{\mu}}(\widehat{\xi}) = 0$ with $0 < \widehat{\xi} < 1$, and $\widehat{\gamma}, \widehat{\zeta} > 0$;*

(ii) *if $Q(\widehat{\mu}) > 0$ and $1 \notin \mathcal{S}$, then $\widehat{\xi} = 0$ and $\widehat{\gamma} = 1 - 1/\phi_{\mathcal{S}}(\widehat{\mu}) > 0$, $\widehat{\zeta} = \frac{1}{2}\nu(\widehat{\mu}) > 0$;*

(iii) *if $Q(\widehat{\mu}) \leq 0$, then $\widehat{\xi} = 1$ and $\widehat{\gamma} = \widehat{\zeta} = 0$.*

Remark 3.2. It is easily seen, using (3.3), that $\widehat{\xi}$ equals the extinction probability of a Galton–Watson process with offspring distribution

$$\begin{aligned}\mathbb{P}(X = j-1) &= \frac{j\,\mathrm{Po}_{\mathcal{S}}(\widehat{\mu})\{j\}}{\nu} \\ &= \frac{\widehat{\mu}^j}{(j-1)!\,\phi_{\mathcal{S}}(\widehat{\mu})}\Bigl(\frac{\widehat{\mu}\phi'_{\mathcal{S}}(\widehat{\mu})}{\phi_{\mathcal{S}}(\widehat{\mu})}\Bigr)^{-1}, \qquad j \in \mathcal{S}\setminus\{0\},\end{aligned}$$

that is, the distribution $\mathrm{Po}_{\mathcal{S}-1}(\widehat{\mu})$ where $\mathcal{S}-1 := \{k \geq 0 : k+1 \in \mathcal{S}\}$. (Note that $\phi_{\mathcal{S}-1}(\mu) = \phi'_{\mathcal{S}}(\mu)$.) Hence $\widehat{\gamma}$, the asymptotic relative size of $\Gamma_{n,\lambda_n/n;\mathcal{S}}$, equals by (3.4) the survival probability of a Galton–Watson process with offspring distribution $\mathrm{Po}_{\mathcal{S}-1}(\widehat{\mu})$ and initial distribution $\mathrm{Po}_{\mathcal{S}}(\widehat{\mu})$.

Remark 3.3. If $\widehat{\mu} = 0$, then $Q(\widehat{\mu}) = 0$ and we are in Case (iii) with no giant component; in fact, by Theorem 2.1, $n_0/n \xrightarrow{\mathrm{p}} 1$ so almost all vertices are isolated. In this case $\chi_{\widehat{\mu}}(\xi) = 0$ for all ξ.

If $\widehat{\mu} > 0$, then $\chi_{\widehat{\mu}} < 0$ on $(0, \widehat{\xi})$ and $\chi_{\widehat{\mu}} > 0$ on $(\widehat{\xi}, 1)$ in all three cases, as follows from [11, Lemma 5.5], which yields another characterization of $\widehat{\xi}$.

Remark 3.4. If $1 \notin \mathcal{S}$, then $Q(\widehat{\mu}) > 0$ as soon as $\widehat{\mu} > 0$. Hence we are in Case (ii) if $\widehat{\mu} > 0$ and in Case (iii) in $\widehat{\mu} = 0$.

Remark 3.5. We have excluded the cases $\mathcal{S} \subseteq \{0,2\}$, i.e., the trivial case $\mathcal{S} = \{0\}$ and the cases $\{2\}$ and $\{0,2\}$ that are exceptional; in the latter cases $n^{-1}\mathrm{v}(\Gamma_{n,\lambda_n/n;\mathcal{S}})$ has a continuous limiting distribution and thus not a constant limit when $\widehat{\mu} > 0$, see Examples 6.2 and 6.7. We note also that $Q(\mu) = 0$ for all μ in the excluded cases.

The k-core of a graph G is the largest induced subgraph having mimimum vertex degree at least k. The k-core of an Erdős–Reńyi random graph has attracted much attention; see [10] and the references therein. Theorem 2.1 may be applied in conjunction with Theorem 2.4 of Janson and Luczak [10] to obtain the asymptotics of the k-core of $G_{n,\lambda_n/n;\mathcal{S}}$. Let $K^{(k)}_{n,p;\mathcal{S}}$ denote the k-core of $G_{n,p;\mathcal{S}}$. We shall require some further notation in order to state our results for $K^{(k)}_{n,\lambda_n/n;\mathcal{S}}$.

Let $k \in \{2,3,\dots\}$. Let $\mu \geq 0$, and let W_μ be a random variable with the $\mathrm{Po}_{\mathcal{S}}(\mu)$ distribution. For $r \in [0,1]$, let $W_{\mu,r}$ be obtained by 'thinning' W_μ at rate $1-r$ so that, conditional on W_μ, $W_{\mu,r}$ has the binomial distribution $\mathrm{Bin}(W_\mu, r)$. For $k \in \{2,3,\dots\}$, let

$$\begin{aligned} h_{\mu,k}(r) &= \mathbb{E}(W_{\mu,r} I_{\{W_{\mu,r} \geq k\}}) = \sum_{l=k}^{\infty} l\,\mathbb{P}(W_{\mu,r} = l), \\ \overline{h}_{\mu,k}(r) &= \mathbb{P}(W_{\mu,r} \geq k). \end{aligned}$$

Theorem 3.6. *Let $\lambda_n \to \lambda > 0$ and suppose that $E_0(\lambda)$ contains a unique element $\widehat{\mu}$. Let $k \geq 2$, and let, with $\nu = \nu(\widehat{\mu})$ as above,*

$$\widehat{r} = \sup\{r \leq 1 : \nu r^2 = h_{\widehat{\mu},k}(r)\}.$$

As $n \to \infty$:

(i) *if $\widehat{r} = 0$,*

$$\frac{1}{n}\mathrm{v}(K^{(k)}_{n,\lambda_n/n;\mathcal{S}}) \xrightarrow{\mathrm{p}} 0, \qquad \frac{1}{n}\mathrm{e}(K^{(k)}_{n,\lambda_n/n;\mathcal{S}}) \xrightarrow{\mathrm{p}} 0;$$

if, further, $k \geq 3$, then

$$\mathbb{P}(K^{(k)}_{n,\lambda_n/n;\mathcal{S}} \text{ is empty}) \to 1;$$

(ii) *if $\widehat{r} > 0$, and in addition $\nu r^2 < h_{\widehat{\mu},k}(r)$ on some non-empty interval $(\widehat{r} - \varepsilon, \widehat{r})$, then*

$$\begin{aligned} &\frac{1}{n}\mathrm{v}(K^{(k)}_{n,\lambda_n/n;\mathcal{S}}) \xrightarrow{\mathrm{p}} \overline{h}_{\widehat{\mu},k}(\widehat{r}), \\ &\frac{1}{n}\mathrm{e}(K^{(k)}_{n,\lambda_n/n;\mathcal{S}}) \xrightarrow{\mathrm{p}} \frac{1}{2}h_{\widehat{\mu},k}(\widehat{r}) = \frac{1}{2}\nu\widehat{r}^2. \end{aligned}$$

Remark 3.7. Let $\mathcal{X}_\mu$ be the Galton–Watson process with offspring distribution $\mathrm{Po}_{\mathcal{S}-1}(\mu)$, started with a single individual o, and let $\overline{\mathcal{X}}_\mu$ be the modified process where the first generation has distribution $\mathrm{Po}_{\mathcal{S}}(\mu)$, cf. Remark 3.2. It may be seen that $\widehat{r}$ is the probability that the family tree of $\mathcal{X}_\mu$ contains an infinite subtree with root o and every node having $k-1$ children. Similarly,

$\overline{h}_{\widehat{\mu},k}(\widehat{r})$ equals the probability that $\overline{\mathcal{X}}_\mu$ contains an infinite k-regular subtree with root o (the root has k children and all other vertices have $k-1$). It is easy to see heuristically that this yields the asymptotic probability that a random vertex belongs to the k-core, see Pittel, Spencer and Wormald [16] (for $\mathcal{S} = \mathbb{Z}_{\geq 0}$), but it is difficult to make a proof based on branching process theory; see Riordan [17] where this is done rigorously for another random graph model.

4. Roots of the characteristic equation

To avoid some trivial complications, we assume throughout this section that $\mathcal{S} \neq \{0\}$, thus excluding the trivial case $\mathcal{S} = \{0\}$ for which $G_{n,\lambda_n/n;\mathcal{S}}$ comprises isolated vertices only.

Lemma 4.1. *$\phi'_{\mathcal{S}}(\mu) \leq C_1 \phi_{\mathcal{S}}(\mu)$ for all $\mu \geq 1$.*

Proof. By (2.2),

$$\phi'_{\mathcal{S}}(\mu) = \sum_{j \in \mathcal{S}} j \frac{\mu^{j-1}}{j!}.$$

We split the sum into two parts. For $j \leq 4\mu$,

$$\sum_{j \in \mathcal{S},\, j \leq 4\mu} \frac{j\mu^{j-1}}{j!} \leq \sum_{j \in \mathcal{S}} \frac{4\mu^j}{j!} = 4\phi_{\mathcal{S}}(\mu).$$

For $j > 4\mu$, Stirling's formula implies

$$\frac{j\mu^{j-1}}{j!} \leq j\mu^{j-1}\Big(\frac{e}{j}\Big)^j \leq 4^{1-j}e^j$$

and thus, for $\mu \geq 1$,

$$\sum_{j \in \mathcal{S},\, j > 4\mu} \frac{j\mu^{j-1}}{j!} \leq \sum_{j=1}^{\infty} 4\Big(\frac{e}{4}\Big)^j = C_2 = C_3\phi_{\mathcal{S}}(1) \leq C_3\phi_{\mathcal{S}}(\mu). \qquad \square$$

Theorem 4.2. *For each $\lambda > 0$, the set $E(\lambda)$ is finite and non-empty.*

Proof. The characteristic equation (2.5) may be written as $h(\mu) = 0$ where $h(\mu) = \lambda\mu\phi'_{\mathcal{S}}(\mu) - \mu^2\phi_{\mathcal{S}}(\mu)$. By Lemma 4.1, for $\mu \geq 1$,

$$h(\mu) \leq (C_1\lambda\mu - \mu^2)\phi_{\mathcal{S}}(\mu), \tag{4.1}$$

and thus $h(\mu) < 0$ for $\mu > C_4 = \max\{C_1\lambda, 1\}$.

Since h is an entire function, and does not vanish identically by what we just have shown, it has only finitely many zeros in each bounded subset of the complex plane, and in particular in the interval $[0, C_4]$. Hence $E(\lambda)$ is finite.

To see that $E(\lambda)$ is non-empty, let s be the smallest element of $\mathcal{S}$. If $s = 0$, then $0 \in E(\lambda)$. If $s > 0$, then $h(\mu) \sim \lambda s\mu^s/s!$ as $\mu \to 0$, so $h(\mu) > 0$ for small positive μ. Since further $h(\mu)$ is negative for large μ, h possesses a zero on the positive real axis. $\square$

We have defined $\widehat{\mu}$ as the maximum point of $\psi_{\mathcal{S}}$ or $\psi_{\mathcal{S},1}$ on $E(\lambda)$. The next theorem shows that, alternatively, it can be defined as the maximum point of $\psi_{\mathcal{S},1}$ on $[0,\infty)$ (but not of $\psi_{\mathcal{S}}$). Furthermore, instead of $\psi_{\mathcal{S},1}$, we can use the function

$$\psi_{\mathcal{S},2}(\mu;\lambda) := \log\phi_{\mathcal{S}}(\mu) + \frac{\mu\phi'_{\mathcal{S}}(\mu)}{2\phi_{\mathcal{S}}(\mu)}\Bigl(\log\frac{\lambda\phi'_{\mathcal{S}}(\mu)}{\mu\phi_{\mathcal{S}}(\mu)} - 1\Bigr), \tag{4.2}$$

that arises as follows. In Section 7, we will indicate the use of multigraphs in proving Theorem 2.1, of which we shall derive a multigraph equivalent at Theorem 7.3. With $Z^*_{n,\lambda/n;\mathcal{S}}$ denoting the multigraph partition function, we shall see in the proof of Theorem 7.3 that $\psi_{\mathcal{S},2}(\mu;\lambda)$ represents the contribution to $n^{-1}\log Z^*_{n,\lambda/n;\mathcal{S}}$ from multigraphs with degree distribution close to $\mathrm{Po}_{\mathcal{S}}(\mu)$, see Remark 8.3. For this reason, $\psi_{\mathcal{S},2}$ is a more natural function than $\psi_{\mathcal{S},1}$, although it has a more complicated formula. We shall have to exclude the trivial case when $\mathcal{S}=\{s\}$ is a singleton; in this case $\psi_{\mathcal{S},2}(\mu;\lambda) = \frac12 s\bigl(\log(\lambda s)-1\bigr) - \log(s!)$ is constant.

It is easily seen that $\psi_{\mathcal{S},2}(\mu;\lambda) \ge \psi_{\mathcal{S},1}(\mu,\lambda)$ for all μ and $\lambda>0$, with equality if and only if $\mu\in E(\lambda)$.

We regard $\psi_{\mathcal{S},1}$ and $\psi_{\mathcal{S},2}$ as functions of μ, with λ considered a fixed parameter. These functions are evidently analytic on $(0,\infty)$. Note that if $0\in\mathcal{S}$, then $\phi_{\mathcal{S}}(0)=1$ and $\psi_{\mathcal{S}}(\mu)$, $\psi_{\mathcal{S},1}(\mu;\lambda)$ and $\psi_{\mathcal{S},2}(\mu;\lambda)$ are continuous at $\mu=0$ with $\psi_{\mathcal{S}}(0)=\psi_{\mathcal{S},1}(0;\lambda)=\psi_{\mathcal{S},2}(0;\lambda)=0$. On the other hand, if $0\notin\mathcal{S}$, then $\phi_{\mathcal{S}}(0)=0$ and $\psi_{\mathcal{S},1}(\mu;\lambda)\to-\infty$, while a simple calculation yields $\psi_{\mathcal{S},2}(\mu;\lambda)\to\frac12 s\bigl(\log(\lambda s)-1\bigr)-\log(s!)$ as $\mu\to0$, where $s=\min\mathcal{S}$.

Theorem 4.3. *The following hold for every fixed $\lambda>0$ and $j=1$ or 2, except for $j=2$ in the trivial case $|\mathcal{S}|=1$.*

(i) *$E(\lambda)$ is the set of stationary points of $\psi_{\mathcal{S},j}$, possibly with 0 added:*

$$E(\lambda)\cap(0,\infty) = \Bigl\{\mu : \frac{d}{d\mu}\psi_{\mathcal{S},j}(\mu;\lambda)=0\Bigr\}.$$

(ii) *$E_0(\lambda)$ is the set of global maximum points of $\psi_{\mathcal{S},j}$:*

$$E_0(\lambda) = \Bigl\{\mu : \psi_{\mathcal{S},j}(\mu;\lambda) = \max_{\mu'\ge0}\psi_{\mathcal{S},j}(\mu';\lambda)\Bigr\}.$$

Proof. (i): Differentiation yields

$$\frac{d}{d\mu}\psi_{\mathcal{S},1}(\mu;\lambda) = \frac{\phi'_{\mathcal{S}}(\mu)}{\phi_{\mathcal{S}}(\mu)} - \frac{\mu}{\lambda} \tag{4.3}$$

and, after some simplifications,

$$\frac{d}{d\mu}\psi_{\mathcal{S},2}(\mu;\lambda) = \frac{d}{d\mu}\Bigl(\frac{\mu\phi'_{\mathcal{S}}(\mu)}{2\phi_{\mathcal{S}}(\mu)}\Bigr)\log\Bigl(\frac{\lambda\phi'_{\mathcal{S}}(\mu)}{\mu\phi_{\mathcal{S}}(\mu)}\Bigr). \tag{4.4}$$

By the Cauchy–Schwarz inequality, provided $\mu>0$ and $|\mathcal{S}|\ge2$,

$$\mu\frac{d}{d\mu}\Bigl(\frac{\mu\phi'_{\mathcal{S}}(\mu)}{\phi_{\mathcal{S}}(\mu)}\Bigr) = \mu\frac{d}{d\mu}\Bigl(\frac{\sum_{j\in\mathcal{S}} j\mu^j/j!}{\sum_{j\in\mathcal{S}}\mu^j/j!}\Bigr)$$

$$= \frac{\left(\sum_{j\in\mathcal{S}} j^2\mu^j/j!\right)\left(\sum_{j\in\mathcal{S}} \mu^j/j!\right) - \left(\sum_{j\in\mathcal{S}} j\mu^j/j!\right)^2}{\left(\sum_{j\in\mathcal{S}} \mu^j/j!\right)^2} > 0. \quad (4.5)$$

(See Theorem 5.2 below for a more general result.) By (4.3)–(4.5), for $\mu > 0$, $\psi'_{\mathcal{S},j}(\mu;\lambda) = 0$ if and only if $\phi'_{\mathcal{S}}(\mu)/\phi_{\mathcal{S}}(\mu) = \mu/\lambda$, i.e., the characteristic equation (2.5) holds.

(ii): By (4.3)–(4.5) and Lemma 4.1, $\psi_{\mathcal{S},j}$ is decreasing for large μ. Furthermore, by the remarks prior to the theorem, $\psi_{\mathcal{S},j}$ is either continuous at 0 or tends to $-\infty$ there. This implies that $\psi_{\mathcal{S},j}$ has a finite maximum, attained at one or several points in $[0,\infty)$. It remains to show that the maximum points belong to $E(\lambda)$; it then follows that $E_0(\lambda)$ equals the set of maximum points.

If $\mu > 0$ is a maximum point of $\psi_{\mathcal{S},j}$, then $\psi'_{\mathcal{S},j}(\mu) = 0$ and $\mu \in E(\lambda)$ by (i).

If 0 is a maximum point and $0 \in \mathcal{S}$, then $0 \in E(\lambda)$ by definition. Finally, if $0 \notin \mathcal{S}$, and $s := \min\mathcal{S} > 0$, then $\phi'_{\mathcal{S}}(\mu)/\phi_{\mathcal{S}}(\mu) \sim s/\mu \to \infty$ as $\mu \to 0$, and it follows from (4.3)–(4.5) that $\psi'_{\mathcal{S},j} > 0$ for small μ; hence 0 is not a maximum point in this case. (In fact, for $j = 1$, $\psi_{\mathcal{S},1}(0) = -\infty$ when $0 \notin \mathcal{S}$, as remarked above.) $\square$

We define $\widehat{\mu}_*(\lambda) := \min E_0(\lambda)$ and $\widehat{\mu}^*(\lambda) := \max E_0(\lambda)$; thus $|E_0(\lambda)| = 1$ (and Theorem 2.1 applies) if and only if $\widehat{\mu}_* = \widehat{\mu}^*$, and in that case $\widehat{\mu} = \widehat{\mu}_* = \widehat{\mu}^*$. We have defined $\widehat{\mu}$ only when $|E_0(\lambda)| = 1$; for convenience we extend the definition to all $\lambda > 0$ by letting $\widehat{\mu}(\lambda)$ be any element of $E_0(\lambda)$ (for example $\widehat{\mu}_*(\lambda)$ or $\widehat{\mu}^*(\lambda)$).

Corollary 4.4. *For every $\lambda > 0$ and $j = 1, 2$,*

$$\psi_{\mathcal{S}}(\widehat{\mu}(\lambda)) = \psi_{\mathcal{S},j}(\widehat{\mu}(\lambda);\lambda) = \max_{\mu\ge 0} \psi_{\mathcal{S},j}(\mu;\lambda).$$

Theorem 4.5. *If $0 < \lambda_1 < \lambda_2$, then $\widehat{\mu}(\lambda_1) \le \widehat{\mu}(\lambda_2)$, with equality only if $\widehat{\mu}(\lambda_1) = \widehat{\mu}(\lambda_2) = 0$.*

Proof. If $\mu < \widehat{\mu}^*(\lambda_1) \in E_0(\lambda_1)$, then $\psi_{\mathcal{S},1}(\mu;\lambda_1) \le \psi_{\mathcal{S},1}(\widehat{\mu}^*(\lambda_1);\lambda_1)$ by Theorem 4.3(ii), and thus

$$\begin{aligned}\psi_{\mathcal{S},1}(\mu;\lambda_2) &= \psi_{\mathcal{S},1}(\mu;\lambda_1) + \frac{\mu^2}{2}\Big(\frac{1}{\lambda_1} - \frac{1}{\lambda_2}\Big)\\ &< \psi_{\mathcal{S},1}(\widehat{\mu}^*(\lambda_1);\lambda_1) + \frac{\widehat{\mu}^*(\lambda_1)^2}{2}\Big(\frac{1}{\lambda_1} - \frac{1}{\lambda_2}\Big) = \psi_{\mathcal{S},1}(\widehat{\mu}^*(\lambda_1);\lambda_2),\end{aligned}$$

so μ is not a global maximum point of $\psi_{\mathcal{S},1}(\mu;\lambda_2)$ and, by Theorem 4.3(ii) again, $\mu \notin E_0(\lambda_2)$. Hence, $\widehat{\mu}(\lambda_2) \ge \widehat{\mu}_*(\lambda_2) \ge \widehat{\mu}^*(\lambda_1) \ge \widehat{\mu}(\lambda_1)$. Equality is possible only if $\mu := \widehat{\mu}(\lambda_1) = \widehat{\mu}(\lambda_2) \in E(\lambda_1) \cap E(\lambda_2)$, and then the characteristic equation (2.5) is satisfied with μ and both λ_1 and λ_2; hence $\mu^2/\lambda_1 = \mu^2/\lambda_2$ and $\mu = 0$. $\square$

Theorem 4.6. (i) *For every $\lambda > 0$, $\widehat{\mu}(\lambda') \nearrow \widehat{\mu}_*(\lambda)$ as $\lambda' \nearrow \lambda$ and $\widehat{\mu}(\lambda') \searrow \widehat{\mu}^*(\lambda)$ as $\lambda' \searrow \lambda$.*

(ii) *$\widehat{\mu}(\lambda) \to 0$ as $\lambda \to 0$.*

(iii) *$\widehat{\mu}(\lambda) \to \infty$ as $\lambda \to \infty$.*

Proof. (i): Let $\mu_0 := \lim_{\lambda' \nearrow \lambda} \widehat{\mu}(\lambda')$; the limit exist by the monotonicity in Theorem 4.5. For any fixed μ, $\psi_{\mathcal{S},1}(\mu; \lambda') \le \psi_{\mathcal{S},1}(\widehat{\mu}(\lambda'); \lambda')$ by Theorem 4.3 and it follows by continuity that $\psi_{\mathcal{S},1}(\mu; \lambda) \le \psi_{\mathcal{S},1}(\mu_0; \lambda)$. Hence, by Theorem 4.3 again, $\mu_0 \in E_0(\lambda)$, and Theorem 4.5 implies that $\mu_0 = \widehat{\mu}_*(\lambda)$.

The second statement is proved similarly.

(ii): This is similar. Assume $\mu_0 := \lim_{\lambda \to 0} \widehat{\mu}(\lambda) > 0$, and let $\mu_1 := \mu_0/2$. Then $\psi_{\mathcal{S},1}(\mu_1; \lambda) \le \psi_{\mathcal{S},1}(\widehat{\mu}(\lambda); \lambda)$ for all λ by Theorem 4.3 which contradicts the fact that, by (2.8),

$$\begin{aligned}\psi_{\mathcal{S},1}(\mu_1; \lambda) - \psi_{\mathcal{S},1}(\widehat{\mu}(\lambda); \lambda) &= \frac{\widehat{\mu}(\lambda)^2 - \mu_1^2}{2\lambda} + \mathrm{O}(1) \\ &\ge \frac{\mu_0^2 - \mu_1^2}{2\lambda} + \mathrm{O}(1) \to \infty\end{aligned}$$

as $\lambda \to 0$.

(iii): Assume $\mu_0 := \lim_{\lambda \to \infty} \widehat{\mu}(\lambda) < \infty$, and let $\mu_1 := \mu_0 + 1$. Then $\psi_{\mathcal{S},1}(\mu_1; \lambda) \le \psi_{\mathcal{S},1}(\widehat{\mu}(\lambda); \lambda)$ for all λ by Theorem 4.3 and it follows from (2.8) by continuity, letting $\lambda \to \infty$, that $\log \phi_{\mathcal{S}}(\mu_1) \le \log \phi_{\mathcal{S}}(\mu_0)$, a contradiction since $\phi_{\mathcal{S}}$ is strictly increasing. □

Remark 4.7. For large λ, we have the estimates $c_3 \lambda^{1/2} \le \widehat{\mu}(\lambda) \le C_5 \lambda$, where the lower bound follows from (2.5) and the upper from (4.1). Examples 6.1 and 6.6 show that both these orders of growth can be attained.

We see from Theorem 4.6 that $|E_0(\lambda)| > 1$ exactly when $\widehat{\mu}(\lambda)$ is discontinuous, and that all discontinuities are jump discontinuities: $\widehat{\mu}$ jumps from $\widehat{\mu}_*(\lambda)$ to $\widehat{\mu}^*(\lambda)$. In accordance with Remark 2.3, we interpret these discontinuities as *phase transitions* of $G_{n,\lambda_n/n;\mathcal{S}}$. More generally, we say that we have a phase transition at each λ where $\widehat{\mu}$ is not analytic. (See, further, Theorem 4.15.) We show that there is only a countable number of phase transitions with jump discontinuities, and we note from Example 6.9 that the number may be infinite. (This notion of phase transition is different from the question whether there is a giant component or not.)

Theorem 4.8. *The set $\widetilde{\Lambda} := \{\lambda > 0 : |E_0(\lambda)| > 1\} = \{\lambda : \widehat{\mu}_*(\lambda) < \widehat{\mu}^*(\lambda)\}$ of λ such that Theorem 2.1 does not apply is at most countable.*

Proof. By Theorem 4.5, the open intervals $\big(\widehat{\mu}_*(\lambda), \widehat{\mu}^*(\lambda)\big)$, $\lambda > 0$, are disjoint, and thus at most a countable number of them are non-empty. □

It follows from Theorem 4.5 and its corollaries that $\widehat{\mu}$ is the *inverse function* of a continuous non-decreasing function with graph $\{(\mu, \lambda) : \mu \in [\widehat{\mu}_*(\lambda), \widehat{\mu}^*(\lambda)]\}$; the exceptional set $\widetilde{\Lambda}$ consists of the values taken by this function on the intervals where it is constant.

For $\mu > 0$ we can rewrite the characteristic equation (2.5) as

$$\lambda = \hat{\lambda}(\mu) := \frac{\mu \phi_{\mathcal{S}}(\mu)}{\phi'_{\mathcal{S}}(\mu)}. \tag{4.6}$$

Thus, $E(\lambda) = \{\mu > 0 : \hat{\lambda}(\mu) = \lambda\}$ or $\{\mu > 0 : \hat{\lambda}(\mu) = \lambda\} \cup \{0\}$. Note that our assumption $\mathcal{S} \neq \{0\}$ implies that $\phi'_{\mathcal{S}}(\mu) > 0$ for all $\mu > 0$, so $\hat{\lambda}$ is well-defined.

Lemma 4.9. *The function $\hat{\lambda}$ is positive and analytic on $(0, \infty)$, with $\lim_{\mu\to\infty} \hat{\lambda}(\mu) = \infty$ and*

$$\lim_{\mu \to 0} \hat{\lambda}(\mu) = \begin{cases} 0, & 0 \notin \mathcal{S}, \\ 0, & 0 \in \mathcal{S},\ 1 \in \mathcal{S}, \\ 1, & 0 \in \mathcal{S},\ 1 \notin \mathcal{S},\ 2 \in \mathcal{S}, \\ \infty, & 0 \in \mathcal{S},\ 1 \notin \mathcal{S},\ 2 \notin \mathcal{S}. \end{cases}$$

Proof. That $\hat{\lambda}$ is analytic and positive is evident. By Lemma 4.1, $\hat{\lambda}(\mu) \geq c_4 \mu$ for large μ, and the behaviour as $\mu \to 0$ follows by looking at the first non-zero terms in the Taylor expansions of $\phi_{\mathcal{S}}$ and $\phi'_{\mathcal{S}}$. □

Lemma 4.10. *$\psi'_{\mathcal{S}}(\mu)$ and $\hat{\lambda}'(\mu)$ have the same sign for every $\mu > 0$. Hence $\psi_{\mathcal{S}}$ and $\hat{\lambda}$ have the same stationary points and are increasing or decreasing on the same intervals.*

Proof. By (2.7) and (2.8), $\psi_{\mathcal{S}}(\mu) = \psi_{\mathcal{S},1}(\mu; \hat{\lambda}(\mu))$. Differentiating and using (4.3) we obtain, for all $\mu > 0$,

$$\psi'_{\mathcal{S}}(\mu) = \frac{\partial}{\partial \mu} \psi_{\mathcal{S},1}(\mu, \hat{\lambda}(\mu)) + \frac{\partial}{\partial \lambda} \psi_{\mathcal{S},1}(\mu, \hat{\lambda}(\mu)) \hat{\lambda}'(\mu) = 0 + \frac{\mu^2}{2\hat{\lambda}(\mu)^2} \hat{\lambda}'(\mu). \quad \square$$

Theorem 4.11. (i) *If $\hat{\lambda}$ is decreasing on an interval (μ_1, μ_2) with $0 \leq \mu_1 < \mu_2$, then there exists $\lambda \in \widetilde{\Lambda}$ with $\widehat{\mu}_*(\lambda) \leq \mu_1 < \mu_2 \leq \widehat{\mu}^*(\lambda)$.*

(ii) *$\widetilde{\Lambda} = \emptyset$ if and only if $\hat{\lambda}$ is increasing on $(0, \infty)$. In this case, for every $\lambda > 0$ either $E(\lambda) = \{\widehat{\mu}(\lambda)\}$ with $\widehat{\mu}(\lambda) \geq 0$ or $E(\lambda) = \{0, \widehat{\mu}(\lambda)\}$ with $\widehat{\mu}(\lambda) > 0$.*

Proof. (i): Suppose that $\widehat{\mu}_*(\lambda) \in (\mu_1, \mu_2)$ for some $\lambda > 0$. Taking a sequence $\lambda_n \nearrow \lambda$, we have $\widehat{\mu}(\lambda_n) \nearrow \widehat{\mu}_*(\lambda)$ by Theorem 4.6, so, for large n, $\mu_1 < \widehat{\mu}(\lambda_n) < \widehat{\mu}_*(\lambda) < \mu_2$ and hence $\lambda_n = \hat{\lambda}(\widehat{\mu}(\lambda_n)) > \hat{\lambda}(\widehat{\mu}_*(\lambda)) = \lambda$, a contradiction. Hence, $\widehat{\mu}_*(\lambda) \notin (\mu_1, \mu_2)$ for all $\lambda > 0$.

Let $\lambda_0 := \sup\{\lambda : \widehat{\mu}_*(\lambda) \leq \mu_1\}$. Thus, $\widehat{\mu}_*(\lambda) \geq \mu_2$ for $\lambda > \lambda_0$. By Theorem 4.6(ii)(iii), $0 < \lambda_0 < \infty$, and by Theorem 4.6(i) $\widehat{\mu}_*(\lambda_0) \leq \mu_1$ and $\widehat{\mu}^*(\lambda_0) \geq \mu_2$. In particular, $\widehat{\mu}_*(\lambda_0) < \widehat{\mu}^*(\lambda_0)$ so $\lambda_0 \in \widetilde{\Lambda}$.

(ii): If $\hat{\lambda}$ is not increasing, then $\hat{\lambda}'(\mu) < 0$ for some $\mu > 0$ and (i) applies to some interval $(\mu - \varepsilon, \mu + \varepsilon)$ and shows that $\widetilde{\Lambda} \neq \emptyset$.

If $\hat{\lambda}$ is increasing, then $\psi_{\mathcal{S}}$ is (strictly) increasing on $(0, \infty)$ by Lemma 4.10; if $0 \in \mathcal{S}$, then $\psi_{\mathcal{S}}$ is continuous at 0 and thus increasing on $[0, \infty)$ also.

Consequently, $E(\lambda)$ contains a unique $\widehat{\mu} = \max E(\lambda)$ that maximizes $\psi_{\mathcal{S}}$. Further, when $\hat{\lambda}$ is increasing, there is at most one positive solution to $\lambda = \hat{\lambda}(\mu)$, and thus to (2.5), and the result on $E(\lambda)$ follows. $\square$

We next study whether $\widehat{\mu} = 0$ is possible. Note that this is a rather degenerate case, when Theorem 2.1 shows that $G_{n,\lambda_n/n;\mathcal{S}}$ is very sparse with $o_p(n)$ edges and $n_0/n \xrightarrow{\text{p}} 1$, which is to say that $n(1 - o_p(1))$ vertices are isolated.

Theorem 4.12. (i) *If $0 \notin \mathcal{S}$, then $\widehat{\mu} > 0$ for every $\lambda > 0$.*

(ii) *If $0 \in \mathcal{S}$ and $1 \in \mathcal{S}$, then $\widehat{\mu} > 0$ for every $\lambda > 0$.*

(iii) *If $0 \in \mathcal{S}$ and $1 \notin \mathcal{S}$, then there exists $\lambda_0 > 0$ such that $\widehat{\mu} = 0$ for every $\lambda < \lambda_0$, but $\widehat{\mu} > 0$ for every $\lambda > \lambda_0$.*

Proof. (i): Trivial, since $0 \notin E(\lambda)$ in this case.

(ii): In this case, $\phi_{\mathcal{S}}(0) = \phi'_{\mathcal{S}}(0) = 1$ and thus (4.3) shows that

$$\frac{d}{d\mu}\psi_{\mathcal{S},1}(\mu;\lambda) \to 1 > 0$$

as $\mu \to 0$; hence $\psi_{\mathcal{S}}(\mu;\lambda)$ is increasing for small μ and 0 is not a maximum point. By Theorem 4.3(ii), $0 \notin E_0(\lambda)$.

(iii): By Lemma 4.9, $\hat{\lambda}(\mu)$ tends to ∞ as $\mu \to \infty$, and to either 1 or ∞ as $\mu \to 0$. Hence $\lambda_1 := \inf_{\mu>0} \hat{\lambda}(\mu) > 0$. If $\lambda < \lambda_1$, there is thus no positive solution to (2.5), so $E(\lambda) = \{0\}$ and $\widehat{\mu} = 0$. The existence of $\lambda_0 \geq \lambda_1$ as asserted now follows from Theorem 4.5 and Theorem 4.6. $\square$

In case (iii), $\widehat{\mu}_*(\lambda_0) = 0$ by Theorem 4.6; it is possible both that $E_0(\lambda_0) = \{0\}$ so that $\widehat{\mu}^*(\lambda_0) = 0$, and that $|E_0(\lambda_0)| > 1$ so that $\lambda_0 \in \widetilde{\Lambda}$ and $\widehat{\mu}^*(\lambda_0) > 0$. We can classify these subcases too.

Theorem 4.13. *Suppose that $0 \in \mathcal{S}$ and $1 \notin \mathcal{S}$.*

(i) *If $2 \in \mathcal{S}$ and $3 \notin \mathcal{S}$ ($\mathcal{S} = \{0,2,s,\dots\}$ with $s \geq 4$, or $\{0,2\}$), then $\widehat{\mu}(\lambda) = 0$ for $\lambda \leq \lambda_0 = 1$ and $\widehat{\mu}(\lambda) > 0$ for $\lambda > 1$, with $\widehat{\mu}^*(1) = 0$ and thus $1 \notin \widetilde{\Lambda}$ and $\widehat{\mu}(\lambda) \searrow 0$ as $\lambda \searrow 1$.*

(ii) *If $2 \in \mathcal{S}$ and $3 \in \mathcal{S}$, or if $2 \notin \mathcal{S}$ ($\mathcal{S} = \{0,2,3,\dots\}$ or $\{0,s,\dots\}$ with $s \geq 3$), then there exists $\lambda_0 > 0$ such that $\widehat{\mu}(\lambda) = 0$ for $\lambda < \lambda_0$ and $\widehat{\mu}(\lambda) > 0$ for $\lambda > \lambda_0$, with $\widehat{\mu}^*(\lambda_0) > 0 = \widehat{\mu}_*(\lambda_0)$ and thus $\lambda_0 \in \widetilde{\Lambda}$ and $\lim_{\lambda\searrow\lambda_0} \widehat{\mu}(\lambda) > 0$.*

Proof. (ii): If $2, 3 \in \mathcal{S}$, then the Taylor series $\phi_{\mathcal{S}}(\mu) = 1 + \frac{1}{2}\mu^2 + \dots$ and $\phi'_{\mathcal{S}}(\mu) = \mu + \frac{1}{2}\mu^2 + \dots$ yield $\hat{\lambda}(\mu) = 1 - \frac{1}{2}\mu + \dots$ for small μ. If $2 \notin \mathcal{S}$, then $1/\hat{\lambda}(\mu) \to 0$ as $\mu \to 0$ by Lemma 4.9. In both cases, $1/\hat{\lambda}$ is analytic in a neighbourhood of 0 and increases on an interval $(0,\mu_0)$, so $\hat{\lambda}$ decreases there and the result follows by Theorems 4.11(i) and 4.12(iii).

(i): Taylor expansions as in the proof of (ii) show that $\hat{\lambda}(\mu) = 1 + \frac{1}{3}\mu^2 + \dots$ (when $4 \in \mathcal{S}$) or $\hat{\lambda}(\mu) = 1 + \frac{1}{2}\mu^2 + \dots$ (when $4 \notin \mathcal{S}$) so $\hat{\lambda}$ increases for small

μ, say in an interval $(0, \mu_0)$. By Lemma 4.10, $\psi_{\mathcal{S}}$ increases in $(0, \mu_0)$, and since $\psi_{\mathcal{S}}$ is continuous at 0 we have $\psi_{\mathcal{S}}(\mu) > \psi_{\mathcal{S}}(0)$ for $0 < \mu < \mu_0$.

However, we also have to consider larger μ, and we use Lemma 4.14 below which implies that $\lambda_1 := \inf_{\mu \geq \mu_0} \hat{\lambda}(\mu) > 1$. It follows that if $\lambda \leq 1$, then $E(\lambda) = \{0\}$, and in particular $\widehat{\mu}^*(1) = 0$. Similarly, if $1 < \lambda < \lambda_1$, then $E(\lambda) = \{0, \mu\}$ for the unique $\mu \in (0, \mu_0)$ with $\hat{\lambda}(\mu) = \lambda$; in this case, $\psi_{\mathcal{S}}(\mu) > \psi_{\mathcal{S}}(0)$ and we have $\widehat{\mu} = \mu > 0$. Finally, by Theorem 4.6, $\widehat{\mu}(\lambda) \searrow \widehat{\mu}^*(1) = 0$ as $\lambda \searrow 1$. $\square$

Lemma 4.14. *Under the assumptions $0, 2 \in \mathcal{S}$ and $1, 3 \notin \mathcal{S}$ of Theorem 4.13(i), $\hat{\lambda}(\mu) > 1$ for every $\mu > 0$.*

Proof. By (4.6), the claimed inequality is equivalent to $\phi'_{\mathcal{S}}(\mu) < \mu\phi_{\mathcal{S}}(\mu)$, where $\phi'_{\mathcal{S}}(\mu) = \sum_{k \in \mathcal{S}} \mu^{k-1}/(k-1)!$. First, we use the trivial estimates

$$\phi'_{\mathcal{S}}(\mu) \leq \phi'_{\mathbb{Z}_{\geq 0} \setminus \{1,3\}}(\mu) = e^{\mu} - 1 - \tfrac{1}{2}\mu^2$$

and

$$\phi_{\mathcal{S}}(\mu) \geq 1 + \tfrac{1}{2}\mu^2.$$

We may verify numerically (by `Maple`, or otherwise) that $e^{\mu} - 1 - \frac{1}{2}\mu^2 < \mu(1 + \frac{1}{2}\mu^2)$ for $0 < \mu < 3.38$, and thus the claim holds in this range.

For larger μ, we split the sum for $\phi'_{\mathcal{S}}(\mu)$ into two parts. With $K := \lfloor \mu^2 \rfloor$, we have that

$$\begin{aligned}\sum_{k \in \mathcal{S},\, k \leq K} \frac{\mu^{k-1}}{(k-1)!} &\leq \mu + \sum_{k \in \mathcal{S},\, 4 \leq k \leq K} \frac{k}{\mu} \cdot \frac{\mu^k}{k!} \leq \mu + \frac{K}{\mu} \sum_{k \in \mathcal{S},\, k \geq 4} \frac{\mu^k}{k!} \\ &= \mu + \frac{K}{\mu}\big(\phi_{\mathcal{S}}(\mu) - 1 - \tfrac{1}{2}\mu^2\big) \\ &\leq \mu\phi_{\mathcal{S}}(\mu) - \tfrac{1}{2}\mu^3. \end{aligned} \tag{4.7}$$

For $k > K$ we use the Chernoff bound for the Poisson distribution, see e.g. [12, Theorem 2.1 and Remark 2.6]:

$$\begin{aligned}\sum_{k \in \mathcal{S},\, k \geq K+1} \frac{\mu^{k-1}}{(k-1)!} &\leq \sum_{k=K}^{\infty} \frac{\mu^k}{k!} = e^{\mu}\, \mathbb{P}\big(\mathrm{Po}(\mu) \geq K\big) \\ &\leq \exp\big(\mu - K\log(K/\mu) + K - \mu\big) = \Big(\frac{e\mu}{K}\Big)^K. \end{aligned} \tag{4.8}$$

For $\mu \geq 3$, $K/\mu = \lfloor \mu^2 \rfloor/\mu > 9/\sqrt{10} > e$, so the sum in (4.8) is less than 1. We combine this with (4.7) to obtain $\phi'_{\mathcal{S}}(\mu) < \mu\phi_{\mathcal{S}}(\mu)$ for $\mu \geq 3$. $\square$

Theorem 4.15. *The set $\{\lambda : \widehat{\mu}$ is not analytic at $\lambda\}$ of phase transitions is at most countable. Each phase transition is of one of the following types.*

(i) *A jump discontinuity: $\lambda \in \widetilde{\Lambda}$ and $\widehat{\mu}_*(\lambda) < \widehat{\mu}^*(\lambda)$.*

(ii) *A continuous phase transition with $\widehat{\mu}_*(\lambda) = \widehat{\mu}^*(\lambda) = 0$ but $\widehat{\mu}(\lambda') > 0$ for $\lambda' > \lambda$. This can happen only at $\lambda = 1$, where it happens if and only if $0, 2 \in \mathcal{S}$ but $1, 3 \notin \mathcal{S}$.*

(iii) *A continuous phase transition with* $\widehat{\mu} > 0$*; in this case,* $\lambda = \hat{\lambda}(\mu)$ *for some* $\mu > 0$ *with* $\hat{\lambda}'(\mu) = 0$ *but* $\hat{\lambda}$ *increasing in a neighbourhood of* μ*; thus* μ *is an inflection point of* $\hat{\lambda}$*.*

Examples of type (i) and (ii) are given in Section 6. We do not know whether (iii) actually occurs.

Proof. Since $\hat{\lambda}(\widehat{\mu}(\lambda)) = \lambda$ when $\widehat{\mu}(\lambda) > 0$ and $\hat{\lambda}$ is analytic, the implicit function theorem shows that $\widehat{\mu}$ is analytic at every point where it is continuous and positive and $\hat{\lambda}'(\widehat{\mu}) \neq 0$. Hence we have only the three given possibilities; in (iii), $\hat{\lambda}$ has to be increasing in a neighbourhood of μ since otherwise we would have a jump discontinuity by Theorem 4.11(i).

The further characterization in (ii) follows by Theorems 4.12 and 4.13.

The number of jump discontinuities is countable by Theorem 4.8, and so is the number of inflection points of $\hat{\lambda}$, while there is at most one phase transition of type (ii). □

In Case (ii), by the proof of Theorem 4.13, $\hat{\lambda}(\mu) = 1+c\mu^2+o(\mu^2)$ as $\mu \to 0$, so $\widehat{\mu}(\lambda) \sim c'\sqrt{\lambda - 1}$ as $\lambda \searrow 1$ and we have a square-root type singularity. In Case (iii), provided it happens at all, if the inflection point is μ_0, then $\hat{\lambda}(\mu) - \hat{\lambda}(\mu_0) \sim c(\mu - \mu_0)^m$ as $\mu \to \mu_0$ for some odd $m \geq 3$ (presumably $m = 3$), and thus $\widehat{\mu}(\lambda) - \mu_0 \sim c'(\lambda - \lambda_0)^{1/m}$ as $\lambda \to \lambda_0 = \hat{\lambda}(\mu_0)$.

Theorems 4.11 and 4.15 show that phase transitions, except the possible one of type (ii), occur when $\hat{\lambda}$ ceases to be increasing at some points, or at least almost ceases to be, in the form of an inflection point. (Recall that $\hat{\lambda} \to \infty$, so it increases in the long run.) We have no criterion for when this happens, but it seems likely that it occurs whenever there are large gaps in $\mathcal{S}$.

Problems 4.16. Several open problems remain. For example:

(i) Is there ever any phase transition of type (iii) in Theorem 4.15 (with an inflection point of $\hat{\lambda}$)?

(ii) Does the set of phase transitions lack accumulation points? In other words, if there is an infinite number of phase transitions (as in Example 6.9), can we always order them in an increasing sequence $\lambda_1 < \lambda_2 < \dots$ with $\lambda_n \to \infty$?

(iii) Is $|E_0(\lambda)|$ always 1 or 2? In the latter case, is always $|E(\lambda)| = 3$, as in Example 6.8, with one intermediate point that is a minimum rather than a maximum of $\psi_{\mathcal{S},1}$ and $\psi_{\mathcal{S},2}$?

5. Monotonicity

We begin with a general result that is a simple consequence of standard results, and may be known already. Recall that if X and Y are two random variables, we say that X is *stochastically smaller* than Y, and write $X \leq_{\mathrm{st}} Y$, if $\mathbb{P}(X > x) \leq \mathbb{P}(Y > x)$ for every real x; it is well-known that this is equivalent to the existence of a coupling (X', Y') of (X, Y) with $X' \le Y'$ a.s. See, for example, [13, Section IV.1].

Lemma 5.1. *Let Y be a random variable taking values in $[0, \infty)$ with distribution function F and moment generating function $M(\theta) = \int_0^\infty e^{\theta y}\, dF$ that is finite for all $\theta < \overline{\theta} \le \infty$, and let, for $\theta < \overline{\theta}$, Y_θ have the conjugate (or tilted) distribution function F_θ satisfying $dF_\theta/dF = e^{\theta y}/M(\theta)$.*

(i) *If $f : [0, \infty) \to \mathbb{R}$ is a non-decreasing function such that $\mathbb{E}\,|f(Y_\theta)| < \infty$ for every $\theta < \overline{\theta}$, then $\frac{d}{d\theta}\,\mathbb{E}\, f(Y_\theta) \ge 0$, with strict inequality except in the trivial case when $f(Y_\theta)$ is almost surely constant. When the last holds for some θ, it holds for all $\theta < \overline{\theta}$.*

(ii) *If $\theta_1 \le \theta_2 < \overline{\theta}$ then $Y_{\theta_1} \le_{\mathrm{st}} Y_{\theta_2}$.*

Proof. (i): For $\theta < \overline{\theta}$,

$$\begin{aligned}\frac{d}{d\theta}\,\mathbb{E}\, f(Y_\theta) &= \frac{d}{d\theta}\frac{\int f(y)e^{\theta y}\, dF}{M(\theta)}\\ &= \frac{\int y f(y)e^{\theta y}\, dF}{M(\theta)} - \frac{\big(\int f(y)e^{\theta y}\, dF\big)\big(\int y e^{\theta y}\, dF\big)}{M(\theta)^2}\\ &= \mathbb{E}\big(Y_\theta f(Y_\theta)\big) - \mathbb{E}(Y_\theta)\,\mathbb{E}\big(f(Y_\theta)\big) = \operatorname{Cov}\big(f(Y_\theta), Y_\theta\big) \ge 0,\end{aligned}$$

since, as is well-known, the two non-decreasing functions $f(Y_\theta)$ and Y_θ of Y_θ are positively correlated, for example by a calculation of $\mathbb{E}\,\big[(f(Y_\theta) - f(Y_\theta'))(Y_\theta - Y_\theta')\big] \ge 0$ with Y_θ' an independent copy of Y_θ. The same proof yields strict inequality if $f(Y_\theta)$ is not a.s. constant. The final remark of (i) holds since the Lebesgue–Stieltjes measures of F_{θ_1} and F_{θ_2} are absolutely continuous with respect to one another.

(ii): By (i), $\mathbb{P}(Y_\theta > x)$ is a non-decreasing function of θ for every x. $\square$

Let, for $\mu > 0$, $X_\mu \sim \mathrm{Po}(\mu)$ and $X_{\mu,\mathcal{S}} \sim \mathrm{Po}_{\mathcal{S}}(\mu)$. Applying Lemma 5.1 to $Y = X_{1,\mathcal{S}}$ with $\mu = e^\theta$, we obtain the following.

Theorem 5.2. (i) *If $f : \mathbb{Z}_{\ge 0} \to \mathbb{R}$ is a non-decreasing function such that $\mathbb{E}\,|f(X_\mu)| < \infty$ for every $\mu > 0$, then*

$$\frac{d}{d\mu}\,\mathbb{E}\, f(X_{\mu,\mathcal{S}}) = \frac{d}{d\mu}\,\mathbb{E}\big(f(X_\mu) \mid X_\mu \in \mathcal{S}\big) \ge 0,$$

with strict inequality except in the trivial case when f is constant on $\mathcal{S}$.

(ii) *If $\mu_1 \le \mu_2$ then $X_{\mu_1,\mathcal{S}} \le_{\mathrm{st}} X_{\mu_2,\mathcal{S}}$.*

This shows, in conjunction with Theorem 2.1, that the asymptotic degree distribution of $G_{n,\lambda_n/n;\mathcal{S}}$ is stochastically increasing in $\widehat{\mu}(\lambda)$ and thus, by Theorem 4.5, in λ. In particular, the asymptotic edge density, which is given by $\mathbb{E}\, X_{\widehat{\mu}(\lambda),\mathcal{S}} = \nu(\widehat{\mu}(\lambda))$, is an increasing function of λ (except that it is constant in the trivial case $|\mathcal{S}| = 1$). This holds for finite n too.

Theorem 5.3. *If $0 < p_1 \le p_2 < 1$, then* $\mathrm{e}(G_{n,p_1;\mathcal{S}}) \le_{\mathrm{st}} \mathrm{e}(G_{n,p_2;\mathcal{S}})$.

Proof. This is another application of Lemma 5.1, since it follows from (1.1) that $\mathrm{e}(G_{n,p;\mathcal{S}})$ has the conjugate distribution $Y_{p/(1-p)}$ with $Y = \mathrm{e}(G_{n,\frac{1}{2};\mathcal{S}})$. □

Unfortunately, if we consider the entire random graph $G_{n,p;\mathcal{S}}$ (and not just the number of its edges), it is in general *not* stochastically increasing in p.

Example 5.4. Let $n = 4$ and let $\mathcal{S} = \{0, 2\}$ (or the set of all even numbers). The $\mathcal{S}$-graphs are, ignoring the labelling: (i) E_4, the empty graph with no edges, (ii) $C_3 + E_1$, a 3-cycle plus an isolated vertex, (iii) C_4, a 4-cycle. We have $\mathbb{P}(G_{n,p;\mathcal{S}} = E_4) \to 1$ as $p \to 0$ and $\mathbb{P}(G_{n,p;\mathcal{S}} = C_4) \to 1$ as $p \to 1$. Hence, if $f(G)$ is the number of 3-cycles in G, then $\mathbb{E}\, f(G_{n,p;\mathcal{S}}) = \mathbb{P}(G_{n,p;\mathcal{S}} = C_3 + E_1)$ tends to 0 both as $p \to 0$ and $p \to 1$, so this expectation is not monotone in p.

Problem 5.5. Is the random multigraph $G^*_{n,\nu;\mathcal{S}}$ defined in Section 7 stochastically increasing in ν? (Its number of edges is, by the same argument as for Theorem 5.3.)

For the existence of a giant component, we note that the crucial quantity $Q(\mu)$ in (3.2) is *not* always monotone in μ, not even in the classical case $\mathcal{S} = \mathbb{Z}_{\geq 0}$ (when $Q(\mu) = \mu^2 - \mu$). Nevertheless, the condition $Q(\mu) > 0$ is monotone.

Theorem 5.6. *If $\mu_1 \leq \mu_2$ and $Q(\mu_1) > 0$, then $Q(\mu_2) > 0$.*

Moreover, assuming $\mathcal{S} \not\subseteq \{0, 2\}$, if $\lambda_1 \leq \lambda_2$ and thus $\widehat{\mu}(\lambda_1) \leq \widehat{\mu}(\lambda_2)$, then $\widehat{\xi}(\widehat{\mu}(\lambda_1)) \geq \widehat{\xi}(\widehat{\mu}(\lambda_2))$ and $\widehat{\gamma}(\widehat{\mu}(\lambda_1)) \leq \widehat{\gamma}(\widehat{\mu}(\lambda_2))$. Hence, if $G_{n,\lambda_1/n;\mathcal{S}}$ has a giant component, then so has $G_{n,\lambda_2/n;\mathcal{S}}$ for all $\lambda_2 \geq \lambda_1$, and it is (asymptotically) at least as large.

Proof. The condition $Q(\mu) > 0$ is equivalent to $\mu\phi''_{\mathcal{S}}(\mu)/\phi'_{\mathcal{S}}(\mu) > 1$. Since $\phi'_{\mathcal{S}}(\mu) = \phi_{\mathcal{S}-1}(\mu)$,

$$\frac{\mu\phi''_{\mathcal{S}}(\mu)}{\phi'_{\mathcal{S}}(\mu)} = \frac{\mu\phi'_{\mathcal{S}-1}(\mu)}{\phi_{\mathcal{S}-1}(\mu)} = \mathbb{E}\, X_{\mu,\mathcal{S}-1}, \tag{5.1}$$

which is non-decreasing by Theorem 5.2(i).

The monotonicity of $\widehat{\xi}$ and $\widehat{\gamma}$ follows from the branching process interpretations in Remark 3.2 together with the stochastic monotonicity Theorem 5.2(ii) (for both $\mathcal{S}$ and $\mathcal{S} - 1$) and Theorem 4.5. □

Remark 5.7. Theorem 3.1 and (5.1) yield the curious relation that, provided $\mathcal{S} \not\subseteq \{0, 2\}$, $G_{n,\lambda/n;\mathcal{S}}$ has a giant component if and only if $\nu_{\mathcal{S}-1}(\widehat{\mu}) := \mathbb{E}\, X_{\widehat{\mu},\mathcal{S}-1} > 1$, with $\widehat{\mu} = \widehat{\mu}(\lambda)$ calculated for $\mathcal{S}$. Cf. Remarks 3.2 and 3.7.

It follows similarly from Remark 3.7 that the existence and size of a k-core, for any fixed $k \geq 3$, is monotone in λ.

6. Examples

Example 6.1. $\mathcal{S} = \mathbb{Z}_{\geq 0}$, the Erdős–Rényi random graph. We have in this much studied case that $\phi_{\mathcal{S}}(\mu) = e^{\mu}$. The characteristic equation (2.5) becomes $\mu = \mu^2/\lambda$, with solutions $\mu = 0$, $\mu = \lambda$. By (2.7), $\psi_{\mathcal{S}}(\mu) = \frac{1}{2}\mu$, so $\psi_{\mathcal{S}}(0) < \psi_{\mathcal{S}}(\lambda)$. Therefore, and in accordance with Theorem 4.12(ii), $\widehat{\mu} = \lambda$, so that the number n_i of vertices of degree i satisfies $n_i/n \to \lambda^i e^{-\lambda}/i!$, $i \geq 0$. This is a simple instance of Theorem 4.11(ii). There is no phase transition of $G_{n,\lambda_n/n;\mathcal{S}}$. We have $\psi_{\mathcal{S},1}(\mu;\lambda) = \mu - \mu^2/(2\lambda)$ and $\psi_{\mathcal{S},2}(\mu;\lambda) = \frac{1}{2}\mu\bigl(\log(\lambda/\mu) + 1\bigr)$. Details of the application of Theorem 3.1 to this well understood case may be found in [15]. Similarly, the application of Theorem 3.6 is described in [10].

Example 6.2. $\mathcal{S} = \{s\}$, where $s \geq 1$. We have $\phi_{\mathcal{S}}(\mu) = \mu^s/s!$ and the characteristic equation (2.5) becomes $s = \mu^2/\lambda$, with solution $\widehat{\mu} = \sqrt{s\lambda}$. However, in this case, the value of $\widehat{\mu}$ is in fact immaterial, since $\mathrm{Po}_{\mathcal{S}}(\mu)$ is a point mass at s for every $\mu > 0$. Moreover, the graph $G_{n,\lambda_n/n;\mathcal{S}}$ is a random regular graph with all vertices of degree s. It is immediate that $Q(\mu) = s(s-2)$, and so there exists a giant component if $s > 2$, and not if $s = 1$. In fact, if $s = 1$, the graph consists of isolated edges only, while if $s \geq 3$, it is well-known that the graph is connected with probability tending to 1, see Bollobás [5], Section VII.6. In the remaining case $s = 2$, the graph consists of cycles, of which the largest has a length that divided by n converges to some non-degenerate distribution on $[0,1]$, see, e.g., Arratia, Barbour and Tavaré [1]; this is thus an exceptional case where we do not have convergence in probability of the proportion of vertices in the giant cluster, as in Theorem 3.1.

Example 6.3. $\mathcal{S} = 2\mathbb{Z}_{\geq 0}$, the even numbers. In this case,

$$\phi_{\mathcal{S}}(\mu) = \sum_{k=0}^{\infty} \frac{\mu^{2k}}{(2k)!} = \cosh\mu.$$

The characteristic equation (2.5) is

$$\frac{\mu\sinh\mu}{\cosh\mu} = \frac{\mu^2}{\lambda},$$

so either $\mu = 0$ or

$$\lambda = \hat{\lambda}(\mu) = \frac{\mu}{\tanh\mu}. \tag{6.1}$$

Since $\hat{\lambda}(\mu) = \mu/\tanh\mu$ increases (strictly) from 1 to ∞ for $\mu \in [0,\infty)$, it follows that: if $\lambda \leq 1$, $\mu = 0$ is the only solution, while if $\lambda > 1$, there is also a positive solution. We have

$$\psi_{\mathcal{S}}(\mu) = \log(\cosh\mu) - \tfrac{1}{2}\mu\tanh\mu.$$

Therefore,

$$\psi'_{\mathcal{S}}(\mu) = \frac{\sinh(2\mu) - 2\mu}{4\cosh^2\mu} > 0$$

for $\mu > 0$. Hence, $\psi_{\mathcal{S}}(\mu) > \psi_{\mathcal{S}}(0)$ for $\mu > 0$, whence $\widehat{\mu}$ is the unique positive solution of (6.1) when $\lambda > 1$, cf. Lemma 4.10, Theorem 4.11(ii) and Theorem 4.13(i).

We thus have a continuous phase transition at $\lambda = 1$ with $\widehat{\mu}(1) = 0$; there is a unique $\widehat{\mu}$ (and thus Theorem 2.1 applies) for every λ, and $\widehat{\mu}(\lambda)$ is a continuous function, but it is not differentiable at $\lambda = 1$. This is the only phase transition, and $\widetilde{\Lambda} = \emptyset$.

The asymptotic edge density (i.e., the number of edges per vertex, see (2.11)) is

$$\nu(\widehat{\mu}) = \widehat{\mu} \tanh \widehat{\mu}. \tag{6.2}$$

Since $1 \notin \mathcal{S}$, Theorem 3.1 shows that there is a giant component as soon as $\widehat{\mu} > 0$, i.e., if $\lambda > 1$. In fact, it is easily seen that

$$Q(\mu) = \mu \tanh \mu \left(\frac{\mu}{\tanh \mu} - 1 \right).$$

One may study a random even subgraph of a general graph G. It turns out that the random even subgraph with parameter $p \in [0, \frac{1}{2}]$ is related to the random-cluster model on G with edge-parameter $2p$ and cluster-weighting factor $q = 2$. When G is a planar graph, the random even subgraph may be identified as the dual graph of the $+/-$ boundary of the Ising model on the (Whitney) dual graph of G with an appropriate parameter-value. This relationship is especially fruitful when G is part of a planar lattice such as the square lattice $\mathbb{Z}^2$. See [6] for a general account of the random-cluster model, and [7] for its relationship with the random even subgraph and the Ising model.

Example 6.4. $\mathcal{S} = \{1, 3, 5, \dots\}$, the odd numbers. This time, $\phi_{\mathcal{S}}(\mu) = \sinh \mu$, and the characteristic equation is

$$\frac{\mu \cosh \mu}{\sinh \mu} = \frac{\mu^2}{\lambda}$$

with $\hat{\lambda}(\mu) = \mu \tanh \mu$. Since $0 \notin \mathcal{S}$ and $\hat{\lambda}$ is increasing, the unique solution $\widehat{\mu}$ is given as the unique positive solution of $\mu \tanh \mu = \lambda$. Cf. Theorems 4.12(i) and 4.11(ii). There is no phase transition. This time,

$$Q(\mu) = \frac{\mu}{\tanh \mu}(\mu \tanh \mu - 1).$$

Thus $Q(\widehat{\mu}) > 0$ if and only if $\widehat{\mu} \tanh \widehat{\mu} > 1$; since $\widehat{\mu} \tanh \widehat{\mu} = \lambda$, it follows that there is a giant component for $\lambda > 1$, and not for $\lambda \le 1$. (This also follows from Remark 5.7 and (6.2).) In the critical case $\lambda = 1$ we have $\widehat{\mu} \tanh \widehat{\mu} = 1$ and numerically $\widehat{\mu} \approx 1.19968$ and (asymptotic) edge density $\nu(\widehat{\mu}) = \widehat{\mu}^2/\lambda = \widehat{\mu}^2 \approx 1.43923$.

Example 6.5. $\mathcal{S} = \{1, 2, 3, \dots\} = \mathbb{Z}_{\ge 1}$, graphs without isolated vertices. We have that $\phi_{\mathcal{S}}(\mu) = e^{\mu} - 1$, and the characteristic equation is

$$\frac{\mu e^{\mu}}{e^{\mu} - 1} = \frac{\mu^2}{\lambda}.$$

Since $0 \notin \mathcal{S}$, we seek strictly positive solutions, which is to say that $\lambda = \hat{\lambda}(\mu) = \mu(1-e^{-\mu})$. Since $\mu(1-e^{-\mu})$ is increasing on $(0,\infty)$, there is a unique such solution $\widehat{\mu}$ for every $\lambda > 0$. Cf. Theorem 4.11(ii).

We have that

$$Q(\mu) = \frac{\mu(\mu-1)e^{\mu}}{e^{\mu}-1}.$$

Thus $Q(\widehat{\mu}) > 0$ if and only if $\widehat{\mu} > 1$, which is to say that $\lambda > \hat{\lambda}(1) = 1 - e^{-1}$. There is a giant component when $\lambda > 1 - e^{-1}$, and not when $\lambda \le 1 - e^{-1}$. In the critical case $\lambda = 1 - e^{-1}$, $\widehat{\mu} = 1$ and the critical (asymptotic) edge density is $\nu(\widehat{\mu}) = \widehat{\mu}^2/\lambda = e/(e-1) \approx 1.58198$.

Example 6.6. $\mathcal{S} = \{0,1\}$, matchings. We have that $\phi_{\mathcal{S}}(\mu) = 1 + \mu$, and the characteristic equation is

$$\frac{\mu}{1+\mu} = \frac{\mu^2}{\lambda}.$$

Either $\mu = 0$ or $\lambda = \hat{\lambda}(\mu) = \mu(1+\mu)$, so the solutions for given λ are $\mu = 0$ and $\mu = -\frac{1}{2} + \sqrt{\lambda + \frac{1}{4}}$. Since $\hat{\lambda}$ is increasing, Theorem 4.11(ii) applies and shows that $\widehat{\mu} = -\frac{1}{2} + \sqrt{\lambda + \frac{1}{4}}$ for all $\lambda > 0$. This can also easily be verified directly, using

$$\psi_{\mathcal{S}}(\mu) = \log(1+\mu) - \frac{\mu}{2(1+\mu)}$$

which yields $\psi'_{\mathcal{S}}(\mu) > 0$ for $\mu \ge 0$ (cf. Lemma 4.10).

By Theorem 2.1, as $n \to \infty$,

$$\frac{n_0}{n} \xrightarrow{\mathrm{p}} \mathrm{Po}_{\mathcal{S}}(\widehat{\mu})\{0\} = \frac{1}{\phi_{\mathcal{S}}(\widehat{\mu})} = \frac{1}{1+\widehat{\mu}} = \frac{1}{\lambda}\left\{\sqrt{\lambda + \frac{1}{4}} - \frac{1}{2}\right\}.$$

Obviously there is no giant component. Indeed, $Q(\mu) = -\mathrm{Po}_{\mathcal{S}}(\mu)\{1\} < 0$ for $\mu > 0$.

Example 6.7. $\mathcal{S} = \{0,2\}$, isolated cycles. We have that $\phi_{\mathcal{S}}(\mu) = 1 + \frac{1}{2}\mu^2$, and the characteristic equation is

$$\frac{\mu^2}{1+\frac{1}{2}\mu^2} = \frac{\mu^2}{\lambda}.$$

Therefore, either $\mu = 0$ or $\lambda = \hat{\lambda}(\mu) = 1 + \frac{1}{2}\mu^2$, so that the solutions for a given λ are $\mu = 0$ and, when $\lambda > 1$, $\mu = \sqrt{2(\lambda-1)}$. Again, $\hat{\lambda}$ is an increasing function, and so is

$$\psi_{\mathcal{S}}(\mu) = \log(1 + \tfrac{1}{2}\mu^2) - \frac{\frac{1}{2}\mu^2}{1+\frac{1}{2}\mu^2},$$

by Lemma 4.10 or direct calculations. Thus, see Theorem 4.11(ii),

$$\widehat{\mu} = \begin{cases} 0 & \text{when } \lambda \le 1, \\ \sqrt{2(\lambda-1)} & \text{when } \lambda > 1. \end{cases}$$

We thus have a continuous phase transition at $\lambda = 1$, of the same type as in Example 6.3, see Theorem 4.12(iii) and Theorem 4.13(i). There is no other phase transition.

It is easily seen that $Q(\mu) = 0$ for all μ, which may be interpreted as saying that the random graph is, in a certain sense, critical whenever $\lambda > 1$. If we remove the isolated vertices, and condition on the number of remaining vertices, we obtain a random regular graph with degree 2. Hence, for $\widehat{\mu} > 0$, i.e., for $\lambda > 1$, we see that the largest component behaves as for $\mathcal{S} = \{2\}$, see Example 6.2, with convergence of $\mathrm{v}(\Gamma_{n,\lambda_n/n;\mathcal{S}})/n$ to a distribution but not to a constant.

Example 6.8. $\mathcal{S} = \{0, 3\}$. This time, $\phi_{\mathcal{S}}(\mu) = 1 + \frac{1}{6}\mu^3$, and the characteristic equation is

$$\frac{\frac{1}{2}\mu^3}{1 + \frac{1}{6}\mu^3} = \frac{\mu^2}{\lambda}.$$

Either $\mu = 0$, or

$$\lambda = \hat{\lambda}(\mu) = \frac{1 + \frac{1}{6}\mu^3}{\frac{1}{2}\mu}. \tag{6.3}$$

This is a convex function of μ with a minimum of $3^{2/3}$ at the point $\mu = 3^{1/3}$. Hence, the characteristic equation has no positive root when $\lambda < 3^{2/3}$, one such root if $\lambda = 3^{2/3}$, and two such roots if $\lambda > 3^{2/3}$. We have

$$\psi_{\mathcal{S}}(\mu) = \log(1 + \tfrac{1}{6}\mu^3) - \frac{\frac{1}{4}\mu^3}{1 + \frac{1}{6}\mu^3}.$$

Unlike the previous examples, $\psi_{\mathcal{S}}$ is *not* monotone, cf. Lemma 4.10. In fact,

$$\psi_{\mathcal{S}}(3^{1/3}) = \log(1 + \tfrac{1}{2}) - \tfrac{1}{2} < 0 = \psi_{\mathcal{S}}(0),$$

so the correct root is $\widehat{\mu} = 0$ (rather than $3^{1/3}$) when $\lambda = 3^{2/3}$. The function ψ has a minimum at $\mu = 3^{1/3}$; ψ decreases on $[0, 3^{1/3}]$ and increases on $[3^{1/3}, \infty)$. There exists thus a unique $\mu_0 > 3^{1/3}$ such that $\psi_{\mathcal{S}}(\mu_0) = 0$, and we set $\lambda_0 = 2(1 + \frac{1}{6}\mu_0^3)/\mu_0$. (Numerically, $\mu_0 \approx 2.03134$ and $\lambda_0 = \hat{\lambda}(\mu_0) \approx 2.36002$.) We deduce that $\widehat{\mu} = 0$ for $\lambda < \lambda_0$ while, for $\lambda > \lambda_0$, $\widehat{\mu}$ is the largest root of (6.3). For $\lambda = \lambda_0$, there are two roots μ of (6.3) with the same value of $\psi_{\mathcal{S}}(\mu)$, so we have a jump phase transition and Theorem 2.1 does not apply; see Theorems 4.12(iii) and 4.13(ii). There is no other phase transition, and $\widetilde{\Lambda} = \{\lambda_0\}$.

Since $1 \notin \mathcal{S}$, by Theorem 3.1(ii), there exists a giant component whenever $\lambda > \lambda_0$. Indeed,

$$Q(\mu) = \frac{\frac{1}{2}\mu^2}{1 + \frac{1}{6}\mu^3} > 0$$

for every $\mu > 0$.

The case $\mathcal{S} = \{0, k\}$, $k \geq 4$ is similar; we omit the details.

Example 6.9. $\mathcal{S} = \{1, 2, 4, 8, \dots\} = \{2^j : j \geq 0\}$. We claim that as $j \to \infty$,

$$\hat{\lambda}(2^j x) = 2^j x^2 \big(1 + \mathrm{o}(1)\big), \tag{6.4}$$

for every $x \in (2/e, 4/e)$. (In fact, this holds uniformly on every closed subinterval of $(2/e, 4/e)$.) It follows that if $a = 4/e$ and $\varepsilon > 0$ is small and fixed, then for large j, $\hat{\lambda}((a-\varepsilon)2^j) \approx (a-\varepsilon)^2 2^j$ and $\hat{\lambda}((a+\varepsilon)2^j) = \hat{\lambda}((a/2+\varepsilon/2)2^{j+1}) \approx (a+\varepsilon)^2 2^{j-1}$, so $\hat{\lambda}$ drops by a factor of about 2 in the vicinity of $a2^j$. Consequently, for all large j, there is an interval $I_j \subset \big((a-\varepsilon)2^j, (a+\varepsilon)2^j\big)$ where $\hat{\lambda}$ is decreasing, and thus by Theorem 4.11 there exists $\lambda \in \widetilde{\Lambda}$ such that $\widehat{\mu}^*(\lambda) \geq \max I_j \geq 2^j$. Hence the set $\{\widehat{\mu}^*(\lambda) : \lambda \in \widetilde{\Lambda}\}$ is unbounded and thus infinite, so $\widetilde{\Lambda}$ is infinite and there is an infinite number of phase transitions.

To verify (6.4) we show that if $\mu = 2^j x \in (a2^{j-1}, a2^j)$ then $\phi_{\mathcal{S}}(\mu) = \sum_{k \in \mathcal{S}} \mu^k / k!$ and $\phi'_{\mathcal{S}}(\mu) = \sum_{k \in \mathcal{S}} k\mu^k/(k-1)!$ are dominated by the terms with $k = 2^j$:

$$\phi_{\mathcal{S}}(\mu) = \big(1 + \mathrm{o}(1)\big)\frac{\mu^{2^j}}{2^j!}, \qquad \phi'_{\mathcal{S}}(\mu) = \big(1 + \mathrm{o}(1)\big)\frac{\mu^{2^j-1}}{(2^j-1)!} \tag{6.5}$$

as $j \to \infty$; this yields that $\phi'_{\mathcal{S}}(\mu) \sim (2^j/\mu)\phi_{\mathcal{S}}(\mu)$ and

$$\hat{\lambda}(\mu) = \frac{\mu}{\phi'_{\mathcal{S}}(\mu)/\phi_{\mathcal{S}}(\mu)} \sim \frac{\mu}{2^j/\mu} = \frac{\mu^2}{2^j} = 2^j x^2.$$

Finally, to show (6.5), we observe by Stirling's inequality that, as $k \to \infty$,

$$\frac{\mu^k}{k!} = \big(1 + \mathrm{o}(1)\big)(2\pi k)^{-1/2}\Big(\frac{e\mu}{k}\Big)^k$$

and thus, with $k_i = 2^i$,

$$\begin{aligned}\frac{\mu^{k_{i+1}}}{k_{i+1}!} \Big/ \frac{\mu^{k_i}}{k_i!} &= \big(1 + \mathrm{o}(1)\big)2^{-1/2}\Big(\frac{e\mu}{2k_i}\Big)^{2k_i}\Big(\frac{e\mu}{k_i}\Big)^{-k_i} \\ &= \big(2^{-1/2} + \mathrm{o}(1)\big)\Big(\frac{e\mu}{4k_i}\Big)^{k_i},\end{aligned}$$

which for $\mu = 2^j x \in (a2^{j-1}, a2^j)$ is exponentially small if $i \geq j$ and exponentially large if $i < j$. The estimate (6.5) for $\phi_{\mathcal{S}}(\mu)$ follows, and a similar calculation with $k_i = 2^i - 1$ yields the result for $\phi'_{\mathcal{S}}(\mu)$.

7. Multigraphs

As explained at the end of Section 2, we shall count multigraphs with certain properties, and shall later relate our conclusions to simple graphs. Let $\mathcal{G}_n^*$ be the (infinite) set of all multigraphs on the vertex set $\{1, 2, \dots, n\}$, and let $\mathcal{G}_{n;\mathcal{S}}^*$ be the subset of $\mathcal{S}$-multigraphs on $\{1, 2, \dots, n\}$ (we extend the definitions above to multigraphs in the obvious way, noting that a loop counts two towards the degree of the vertex in question).

Let $\nu \geq 0$. We define a random multigraph $G^*_{n,\nu}$ by taking $\mathrm{Po}(\nu)$ edges between each pair of vertices and $\mathrm{Po}(\frac{1}{2}\nu)$ loops at each vertex, these random numbers being independent of one another. It is easily seen that this is equivalent to assigning to each multigraph $G \in \mathcal{G}^*_n$ the probability

$$\mathbb{P}(G^*_{n,\nu} = G) = w(G)\nu^{\mathrm{e}(G)}e^{-n^2\nu/2}, \tag{7.1}$$

where

$$w(G) := 2^{-\ell}\prod_{j\geq 2} j!^{-m_j},$$

with ℓ the number of loops of G, and m_j the number of j-fold multiple edges (including multiple loops). That is, $m_j = a_j + b_j$ where a_j is the number of distinct pairs of vertices joined by exactly j parallel edges, and b_j is the number of vertices having exactly j loops. See, e.g., Janson, Knuth, Łuczak and Pittel [9].

Note that the total number of edges $\mathrm{e}(G^*_{n,\nu})$ is Poisson-distributed with parameter $\binom{n}{2}\nu + \frac{1}{2}n\nu = \frac{1}{2}n^2\nu$. We further define the random $\mathcal{S}$-multigraph $G^*_{n,\nu;\mathcal{S}}$ by conditioning $G^*_{n,\nu}$ on being an $\mathcal{S}$-multigraph. Thus, for any multigraph $G \in \mathcal{G}^*_{n;\mathcal{S}}$, by (7.1),

$$\mathbb{P}\big(G^*_{n,\nu;\mathcal{S}} = G\big) = \frac{1}{Z^*_{n,\nu;\mathcal{S}}} w(G)\nu^{\mathrm{e}(G)}, \tag{7.2}$$

where

$$Z^*_{n,\nu;\mathcal{S}} := \sum_{G\in\mathcal{G}^*_{n;\mathcal{S}}} w(G)\nu^{\mathrm{e}(G)} = e^{n^2\nu/2}\,\mathbb{P}(G^*_{n,\nu} \text{ is an } \mathcal{S}\text{-multigraph}). \tag{7.3}$$

We shall assume, of course, that $\mathcal{G}^*_{n;\mathcal{S}} \neq \emptyset$. It is easy to see that this holds for all n if $\mathcal{S}$ contains some even number, but if all elements of $\mathcal{S}$ are odd, then n has to be even. We tacitly assume this in the sequel.

If the multigraph $G \in \mathcal{G}^*_n$ is simple, i.e. has no loops and no multiple edges, then $w(G) = 1$ and (7.1) yields $\mathbb{P}(G^*_{n,\nu} = G) \asymp \nu^{\mathrm{e}(G)} \asymp \mathbb{P}(G_{n,p} = G)$ when $\nu = p/(1-p)$, i.e., $p = \nu/(1+\nu)$. Hence, assuming this relation between ν and p, $G^*_{n,\nu}$ conditioned on being simple has the same distribution as $G_{n,p}$. Conditioning further on being $\mathcal{S}$-graphs, we obtain the following.

Lemma 7.1. *If $\nu = p/(1-p)$, then*

$$G_{n,p;\mathcal{S}} \overset{\mathrm{d}}{=} \big(G^*_{n,\nu;\mathcal{S}} \mid G^*_{n,\nu;\mathcal{S}} \text{ is simple}\big).$$

We are interested in the case $np \to \lambda < \infty$, and note that $np \to \lambda$ and $n\nu \to \lambda$ are equivalent.

We shall also use the configuration model for random multigraphs with given vertex degrees introduced by Bollobás [4], see Bollobás [5], Section II.4. (See Bender and Canfield [2] and Wormald [18, 19] for related arguments.) To be precise, let us fix the vertex degrees to be some non-negative integers $d_1, d_2, \dots, d_n$ (assuming tacitly that $\sum_i d_i$ is even); equivalently, we fix a degree sequence $\mathbf{d} = (d_i)_1^n$. We attach d_i *half-edges* (or *stubs*) to vertex i.

The total number of half-edges is thus $2N := \sum_{i=1}^n d_i$, and a *configuration* is one of the $(2N-1)!! = (2N)!/(2^N N!)$ partitions of the set of half-edges into N pairs. Each configuration defines a multigraph in $\mathcal{G}_n^*$ by combining each pair of half-edges to an edge; this multigraph has vertex degrees $d_1, \dots, d_n$ and $N = \frac{1}{2}\sum_i d_i$ edges. By taking a uniformly random configuration we thus obtain a random multigraph $G_{n,\mathbf{d}}^*$ with the given degree sequence $\mathbf{d}$.

It is easily seen that every multigraph $G \in \mathcal{G}_n^*$ with the given vertex degrees $d_1, \dots, d_n$ arises from exactly $w(G)\prod_{i=1}^n d_i!$ configurations. We obtain therefore that the contribution to $Z_{n,\nu;\mathcal{S}}^*$ in (7.3) from a set of multigraphs with given vertex degrees $d_1, \dots, d_n \in \mathcal{S}$ is given by summing $\nu^N / \prod_{i=1}^n d_i!$ over all corresponding configurations. In particular, since the number of configurations is $(2N-1)!!$, the contribution to $Z_{n,\nu;\mathcal{S}}^*$ from all multigraphs with vertex degrees $d_1, \dots, d_n \in \mathcal{S}$ equals

$$\frac{(2N-1)!!\,\nu^N}{\prod_{i=1}^n d_i!}. \tag{7.4}$$

Moreover, for given $\mathbf{d} = (d_i)_1^n$, the factor $\nu^N / \prod_{i=1}^n d_i!$ is a constant, so by (7.1), the probability that $G_{n,\nu}^*$ belongs to any given set of multigraphs with this degree sequence $\mathbf{d}$ is proportional to the number of corresponding configurations. Consequently, if $d_i(G)$ denotes the degree of vertex i in a (multi)graph G, and $\mathbf{d}(G) := (d_i(G))_{i=1}^n$ for $G \in \mathcal{G}_n^*$, we obtain the following well-known fact. (This is another reason for the weights $w(G)$ in (7.1).)

Lemma 7.2. *For any given degree sequence $\mathbf{d} = (d_i)_1^n$ and any $\nu > 0$, the random multigraph $G_{n,\nu}^*$ conditioned on having degree sequence $\mathbf{d}$ has the distribution given by the configuration model; in other words,*

$$\big(G_{n,\nu}^* \mid \mathbf{d}(G_{n,\nu}^*) = \mathbf{d}\big) \overset{\mathrm{d}}{=} G_{n,\mathbf{d}}^*.$$

As a consequence, if every $d_i \in \mathcal{S}$, the same holds for $G_{n,\nu;\mathcal{S}}^*$.

We are interested in the case $\nu = \lambda_n/n$, with $\lambda_n \to \lambda > 0$.

Theorem 7.3. *The results of Theorem 2.1(i)–(iii) hold with $G_{n,\lambda_n/n;\mathcal{S}}$ replaced by $G_{n,\lambda_n/n;\mathcal{S}}^*$. Furthermore,*

$$\frac{1}{n}\log \mathbb{P}(G_{n,\lambda_n/n}^* \text{ is an } \mathcal{S}\text{-graph}) \to \psi_{\mathcal{S}}(\widehat{\mu}) - \tfrac{1}{2}\lambda \tag{7.5}$$

and $n^{-1}\log Z_{n,\lambda_n/n;\mathcal{S}}^* \to \psi_{\mathcal{S}}(\widehat{\mu})$.

We will prove Theorem 7.3 in the following section, and then obtain Theorem 2.1 as a consequence using Lemma 7.1 and the following technical result, which is proved in Section 9.

Lemma 7.4. *If $\lambda_n \to \lambda > 0$, then $\liminf_{n\to\infty} \mathbb{P}\big(G_{n,\lambda_n/n;\mathcal{S}}^* \text{ is simple}\big) > 0$.*

8. Proof of Theorem 7.3

For notational convenience, we shall consider only the case $\lambda_n = \lambda$ for all n, while noting that our estimates may be extended to the general case $\lambda_n \to \lambda$. (The "constants" below generally depend on λ, but they may be chosen uniformly for λ lying in any compact subset of $(0, \infty)$. Uniformity as $\lambda \to 0$ is less obvious, and perhaps not always true, but it is remarked a few times when it is important for later proofs.)

Let $\mathcal{N}_{\mathcal{S}}^n$ denote the set of all $\mathbf{n} = (n_0, n_1, \dots) \in \mathcal{N}_{\mathcal{S}}$ such that: $\sum_j n_j = n$ and $\sum_j jn_j$ is even. We write $z(\mathbf{n})$ for the contribution to $Z^*_{n,\lambda/n;\mathcal{S}}$ from all multigraphs with n_j vertices of degree j, which is to say that

$$Z^*_{n,\lambda/n;\mathcal{S}} = \sum_{\mathbf{n} \in \mathcal{N}_{\mathcal{S}}^n} z(\mathbf{n}) \tag{8.1}$$

and

$$\mathbb{P}\bigl(\mathbf{n}(G^*_{n,\lambda/n;\mathcal{S}}) = \mathbf{n}\bigr) = \frac{z(\mathbf{n})}{Z^*_{n,\lambda/n;\mathcal{S}}}, \qquad \mathbf{n} \in \mathcal{N}_{\mathcal{S}}^n. \tag{8.2}$$

By (7.4) with $N = \frac{1}{2}\sum_j jn_j$ and $\nu = \lambda/n$,

$$z(\mathbf{n}) = \frac{n!\,\bigl(\sum_j jn_j - 1\bigr)!!}{\prod_j n_j!\,\prod_j j!^{n_j}} \Bigl(\frac{\lambda}{n}\Bigr)^N. \tag{8.3}$$

We note by Stirling's formula that

$$(2N-1)!! = \frac{(2N)!}{2^N N!} = \Bigl(\frac{2N}{e}\Bigr)^N \bigl(\sqrt{2} + \mathrm{O}(N^{-1})\bigr), \tag{8.4}$$

and it is easily verified that

$$(2N-1)!! \ge \Bigl(\frac{2N}{e}\Bigr)^N, \qquad N \ge 0. \tag{8.5}$$

Let $\widehat{\mathbf{n}} = \widehat{\mathbf{n}}(n)$ be a mode of the random sequence $\mathbf{n}(G^*_{n,\lambda/n;\mathcal{S}})$, i.e., by (8.2), a sequence in $\mathcal{N}_{\mathcal{S}}^n$ that maximizes $z(\mathbf{n})$. (In the case of a tied maximum we make an arbitrary choice.) We write $\widehat{N} = \widehat{N}(n) := \frac{1}{2}\sum_j j\widehat{n}_j$.

We begin with a coarse but useful quantitative estimate, obtained by considering only regular multigraphs. Let $\mathbf{e}_s := (\delta_{is})_{i=0}^\infty$, where δ_{is} is the Kronecker delta. In the following lemma we take an even number $s \in \mathcal{S}$, if $\mathcal{S}$ contains such a number. If not, we pick an odd $s \in \mathcal{S}$ and must then, as noted in the introduction, restrict ourselves to even values of n.

Lemma 8.1. *Let $s \in \mathcal{S}$, and assume that s is even if possible. Then*

$$c_5(s)^n \le z(n\mathbf{e}_s) \le z(\widehat{\mathbf{n}}) \le Z^*_{n,\lambda/n;\mathcal{S}} \le C_6^n. \tag{8.6}$$

As a consequence,

$$\mathbb{P}(G^*_{n,\lambda/n} \text{ is an } \mathcal{S}\text{-multigraph}) \ge c_6^n,$$

and, for any set $\mathcal{H}$ of multigraphs,

$$\mathbb{P}(G^*_{n,\lambda/n;\mathcal{S}} \in \mathcal{H}) \le C_7^n\,\mathbb{P}(G^*_{n,\lambda/n} \in \mathcal{H}). \tag{8.7}$$

Proof. By (8.3) and (8.5),

$$z(n\mathbf{e}_s) = (s!)^{-n}(ns-1)!!\Big(\frac{\lambda}{n}\Big)^{ns/2} \ge (s!)^{-n}\Big(\frac{ns}{e}\cdot\frac{\lambda}{n}\Big)^{ns/2},$$

which yields the first inequality in (8.6) with $c_5 = (s!)^{-1}(s\lambda/e)^{s/2}$. The second and third inequalities are trivial, and the fourth follows from (7.3) (with $\nu = \lambda/n$), which yields $Z^*_{n,\lambda/n;\mathcal{S}} \le e^{n^2\nu/2} = e^{n\lambda/2}$.

By (7.3) and (8.6),

$$\mathbb{P}(G^*_{n,\lambda/n} \text{ is an } \mathcal{S}\text{-multigraph}) = e^{-\lambda n/2} Z^*_{n,\lambda/n;\mathcal{S}} \ge c_6^n,$$

with $c_6 = c_5 e^{-\lambda/2}$. Consequently, (8.7) follows with $C_7 = c_6^{-1}$ by the definition of conditional probabilities. □

Lemma 8.2. *There exists a constant $B = B(\mathcal{S},\lambda)$ such that*

$$\sum_{\mathbf{n}\in\mathcal{N}^n_{\mathcal{S}}:\, N>Bn} z(\mathbf{n}) < e^{-n} z(\widehat{\mathbf{n}}) \le e^{-n} Z^*_{n,\lambda/n;\mathcal{S}},$$

where $N = \frac12\sum_i in_i$. Hence, $\mathbb{P}\big(\mathrm{e}(G^*_{n,\lambda/n;\mathcal{S}}) > Bn\big) < e^{-n}$.

More generally, for any $x \ge Bn$,

$$\mathbb{P}\big(\mathrm{e}(G^*_{n,\lambda/n;\mathcal{S}}) > x\big) < e^{-x/B}.$$

Moreover, for any $\lambda_0 > 0$, the constant B can be chosen uniformly for all $\lambda \le \lambda_0$.

Proof. By (7.1),

$$\sum_{\mathbf{n}\in\mathcal{N}^n_{\mathcal{S}}:\, N>x} z(\mathbf{n}) \le \sum_{\mathbf{n}\in\mathcal{N}^n:\, N>x} z(\mathbf{n}) = e^{n\lambda/2}\,\mathbb{P}\big(\mathrm{e}(G^*_{n,\lambda/n}) > x\big). \tag{8.8}$$

Since the number of edges $\mathrm{e}(G^*_{n,\lambda/n}) \sim \mathrm{Po}(\frac12\lambda n)$, it follows by standard Chernoff estimates for the Poisson distribution, see e.g. [12, Corollary 2.4 and Remark 2.6], that if $B \ge 4\lambda$ and $x \ge Bn \ge 4\lambda n$, then

$$\mathbb{P}\big(\mathrm{e}(G^*_{n,\lambda/n}) > x\big) = \mathbb{P}\big(\mathrm{Po}(\tfrac12\lambda n) > x\big) < e^{-x}. \tag{8.9}$$

We choose

$$B \ge \max\{4\lambda,\, \tfrac12\lambda + 1 - \log c_5(s)\},$$

and find by (8.8), (8.9) and (8.6), since $(B-1)n \le (B-1)x/B = x - x/B$,

$$\sum_{\mathbf{n}\in\mathcal{N}^n_{\mathcal{S}}:\, N>x} z(\mathbf{n}) < e^{n\lambda/2 - x - n\log c_5(s)} z(\widehat{\mathbf{n}}) \le e^{(B-1)n-x} z(\widehat{\mathbf{n}}) \le e^{-x/B} z(\widehat{\mathbf{n}}).$$

The results follow by this and (8.2). □

Let $\mathbf{n} \in \mathcal{N}_{\mathcal{S}}$, and let j and k be two different indices in $\mathcal{S}$ such that $n_k \ge 2$, and define $\mathbf{n}' \in \mathcal{N}_{\mathcal{S}}$ by $n'_j = n_j + 2$, $n'_k = n_k - 2$, and $n'_i = n_i$ for $i \ne j,k$; in other words, we replace two vertices of degree k by vertices of degree j. By (8.3), with $N = \frac{1}{2}\sum_i in_i$ and $N' = \frac{1}{2}\sum_i in'_i = N + j - k$,

$$\begin{aligned}\frac{z(\mathbf{n}')}{z(\mathbf{n})} &= \frac{(2N'-1)!!}{(2N-1)!!} \cdot \frac{n_j!\,n_k!}{n'_j!\,n'_k!} \cdot \frac{k!^2}{j!^2} \cdot \Bigl(\frac{\lambda}{n}\Bigr)^{N'-N} \\ &= \frac{n_k(n_k-1)k!^2}{(n_j+1)(n_j+2)j!^2}(2N)^{j-k}\Bigl(\frac{\lambda}{n}\Bigr)^{j-k}\left(1 + \mathrm{O}(|j-k|^2/N)\right).\end{aligned}$$

For $\mathbf{n} = \widehat{\mathbf{n}}$ and any $j, k \in \mathcal{S}$ with $\widehat{n}_k \geq 2$, this quotient is ≤ 1. Hence, for all $j, k \in \mathcal{S}$ (also, trivially, if $\widehat{n}_k < 2$ or $j = k$),

$$\widehat{n}_k(\widehat{n}_k - 1)k!^2 \leq (\widehat{n}_j + 1)(\widehat{n}_j + 2)j!^2 \Bigl(\frac{2\lambda\widehat{N}}{n}\Bigr)^{k-j} \Bigl(1 + \mathrm{O}(|j-k|^2/\widehat{N})\Bigr). \tag{8.10}$$

Furthermore, in the case $k > j$, we have $N' < N$ and

$$(2N-1)!! < (2N'-1)!!\,(2N)^{k-j},$$

and we obtain in the same way the sharper inequality

$$\widehat{n}_k(\widehat{n}_k - 1)k!^2 \leq (\widehat{n}_j + 1)(\widehat{n}_j + 2)j!^2 \Bigl(\frac{2\lambda\widehat{N}}{n}\Bigr)^{k-j}. \tag{8.11}$$

By Lemma 8.2,

$$\widehat{N} \leq Bn. \tag{8.12}$$

Now let $n \to \infty$. Since $\widehat{N}/n$ is bounded by (8.12), each subsequence has a subsequence such that $\widehat{N}/n$ converges. Consider such a subsequence, and assume that $2\widehat{N}/n \to \nu \geq 0$. Furthermore, let $\widehat{p}_j := \widehat{n}_j/n$. Then $\widehat{\mathbf{p}} := (\widehat{p}_j)_0^\infty$ is a probability distribution (the distribution of the degree of a random vertex in a graph G with $\mathbf{n}(G) = \widehat{\mathbf{n}}$). Since the mean of this distribution is $\sum_j j\widehat{p}_j = 2\widehat{N}/n$, which is bounded by (8.12) as $n \to \infty$, this sequence of distributions is tight, and by taking a further subsequence we may assume that the distributions converge, i.e. that $\widehat{n}_j/n \to \overline{p}_j$ for some probability distribution $(\overline{p}_j)_0^\infty$ and every $j \geq 0$. Clearly, this probability distribution is supported on $\mathcal{S}$ in that $\overline{p}_j = 0$ when $j \notin \mathcal{S}$.

We treat the cases $\nu > 0$ and $\nu = 0$ separately. Assume first that $\nu > 0$. Divide (8.10) by n^2 and let $n \to \infty$ to find that

$$\overline{p}_k^2 k!^2 \leq \overline{p}_j^2 j!^2\,(\lambda\nu)^{k-j}, \qquad j, k \in \mathcal{S},$$

and thus

$$\overline{p}_k^2 k!^2 (\lambda\nu)^{-k} \leq \overline{p}_j^2 j!^2\,(\lambda\nu)^{-j}, \qquad j, k \in \mathcal{S}. \tag{8.13}$$

Interchanging j and k we obtain equality in (8.13). Thus $\overline{p}_j^2 j!^2 (\lambda\nu)^{-j}$ does not depend on $j \in \mathcal{S}$. Writing C_8^2 for its common value, and $\mu := \sqrt{\lambda\nu}$, we deduce that

$$\overline{p}_j = C_8 \frac{\mu^j}{j!}, \qquad j \in \mathcal{S}. \tag{8.14}$$

If, instead, $\nu = 0$, then $\sum_{i\geq 1} \widehat{n}_i \leq 2\widehat{N} = o(n)$, so $\widehat{n}_0/n \to 1$ and $\widehat{n}_i/n \to 0$, $i \geq 1$; hence $\overline{p}_0 = 1$, and $\overline{p}_j = 0$ for $j > 0$. (Thus, $\nu = 0$ implies that $0 \in \mathcal{S}$.) Equation (8.14) holds in this case also, this time with $\mu = 0$.

Hence, (8.14) holds in all cases. Summing over j and recalling (2.2), we find that $1 = C_8\phi_{\mathcal{S}}(\mu)$ and thus $C_8 = \phi_{\mathcal{S}}(\mu)^{-1}$. In particular, $\phi_{\mathcal{S}}(\mu) > 0$, another demonstration that $\mu = 0$ is possible only when $0 \in \mathcal{S}$.

In summary, along the selected subsequence,

$$\widehat{p}_j := \frac{\widehat{n}_j}{n} \to \overline{p}_j = \frac{\mu^j/j!}{\phi_{\mathcal{S}}(\mu)} = \mathrm{Po}_{\mathcal{S}}(\mu)\{j\}, \qquad j \in \mathcal{S}, \tag{8.15}$$

which is to say that every subsequence possesses a subsequence along which

$$\widehat{\mathbf{p}} = (\widehat{p}_j)_0^\infty \to \mathrm{Po}_{\mathcal{S}}(\mu) \tag{8.16}$$

for some μ. We next identify μ, and show that it is the same for all subsequences.

We constructed $\mathbf{n}'$ above by changing by 2 the degrees of two vertices; the reason was that this ensures that $\sum_i in'_i$ remains even. If j and k have the same parity, i.e. $j - k \equiv 0 \pmod 2$, then we may also argue as above changing just one vertex degree from k to j. If further $k \geq j$, this leads as in (8.11) to the inequality

$$\widehat{n}_k k! \leq (\widehat{n}_j + 1)j! \left(\frac{2\lambda\widehat{N}}{n}\right)^{(k-j)/2}. \tag{8.17}$$

We consider again a subsequence along which (8.16) holds for some μ. For every k we apply (8.17) with j the smallest number in $\mathcal{S}$ of the same parity as k. Using (8.15) for these (at most two) j, we obtain, with $\mu_n := (2\lambda\widehat{N}/n)^{1/2} \to (\lambda\nu)^{1/2} = \mu$, uniformly for all $k \in \mathcal{S}$,

$$\begin{aligned}\widehat{n}_k &\leq \frac{1}{k!} n\bigl(\overline{p}_j + o(1)\bigr) j!\, \mu_n^{k-j} = \frac{n}{\phi_{\mathcal{S}}(\mu)k!}\mu_n^{k-j}\bigl(\mu^j + o(1)\bigr) \\ &\leq \frac{n}{\phi_{\mathcal{S}}(\mu)k!}(\mu+1)^k\bigl(1+o(1)\bigr),\end{aligned}$$

since $\mu_n < \mu + 1$ for large n. Consequently, for every exponent $r > 0$,

$$\sum_{k\in\mathcal{S}} k^r \widehat{p}_k = \sum_{k\in\mathcal{S}} k^r \frac{\widehat{n}_k}{n} \leq \bigl(1+o(1)\bigr)\phi_{\mathcal{S}}(\mu)^{-1} \sum_{k\in\mathcal{S}} \frac{k^r}{k!}(\mu+1)^k = O(1).$$

In other words, for every $r \in (0,\infty)$, the distributions $\widehat{\mathbf{p}}$ have rth moments that are uniformly bounded in n. It follows that all moments converge in (8.16), i.e., for every $r > 0$,

$$\sum_{k\in\mathcal{S}} k^r \widehat{p}_k = \sum_{k\in\mathcal{S}} k^r \frac{\widehat{n}_k}{n} \to \sum_{k\in\mathcal{S}} k^r \overline{p}_k = \sum_{k\in\mathcal{S}} k^r \,\mathrm{Po}_{\mathcal{S}}(\mu)\{k\}. \tag{8.18}$$

In particular, $r = 1$ yields, using (2.4)

$$\frac{2\widehat{N}}{n} = \sum_{k\in\mathcal{S}} k\frac{\widehat{n}_k}{n} \to \sum_{k\in\mathcal{S}} k\overline{p}_k = \frac{\mu\phi'_{\mathcal{S}}(\mu)}{\phi_{\mathcal{S}}(\mu)}. \tag{8.19}$$

On the other hand, we have assumed $2\widehat{N}/n \to \nu$ and $\mu = \sqrt{\lambda\nu}$, whence $\nu = \mu^2/\lambda$, so we have the consistency relation

$$\frac{\mu\phi'_{\mathcal{S}}(\mu)}{\phi_{\mathcal{S}}(\mu)} = \frac{\mu^2}{\lambda}. \tag{8.20}$$

In other words, with $E(\lambda)$ defined by (2.6), we have $\mu \in E(\lambda)$.

We summarize the result so far. *Each subsequence of n possesses a subsequence such that* (8.16) *and* (8.18) *hold for some $\mu \in E(\lambda)$, i.e., $\mu \geq 0$ satisfies* (8.20) *and, further, $\mu = 0$ only if $0 \in \mathcal{S}$.*

The next step is to find the right solution of (8.20) in the case when $E(\lambda)$ contains two or more points.

We continue to consider a subsequence for which (8.16) holds. By applying again (8.17) with j the smallest odd or even number in $\mathcal{S}$ as appropriate, and using $\widehat{n}_j \leq n$ and (8.12), we see that for some constants C_9, C_{10},

$$\widehat{n}_k k! \leq C_9 n C_{10}^k. \tag{8.21}$$

If $k \geq \log n$, then by Stirling's formula, for large n,

$$\log\bigl(k!\,C_{10}^{-k}\bigr) \geq k\log k - k(1+\log C_{10}) \geq 2k > \log(C_9 n),$$

and thus (8.21) yields $\widehat{n}_k < 1$. Consequently, for large n,

$$\widehat{n}_k = 0 \text{ for all } k \geq \log n. \tag{8.22}$$

Let us now estimate $z(\widehat{\mathbf{n}}) = \max_{\mathbf{n}} z(\mathbf{n})$. By (8.3) and Stirling's formula, recalling (8.4),

$$\begin{aligned}\log z(\widehat{\mathbf{n}}) = n\log n - n + \mathrm{O}(\log n) - \sum_i \bigl(\widehat{n}_i \log \widehat{n}_i - \widehat{n}_i + \mathrm{O}(\log(\widehat{n}_i+1))\bigr)\\ - \sum_i \widehat{n}_i \log(i!) + \widehat{N}\bigl(\log(2\widehat{N}) - 1\bigr) + \mathrm{O}(1) + \widehat{N}\log(\lambda/n).\end{aligned}$$

By (8.22), we only have to sum over $i \leq \log n$, and thus the sum of all O terms is $\mathrm{O}(\log^2 n)$. Thus, with $y := \sum_i i\widehat{p}_i = 2\widehat{N}/n$,

$$\begin{aligned}\frac{1}{n}\log z(\widehat{\mathbf{n}}) &= \log n - 1 - \sum_i \widehat{p}_i\bigl(\log n + \log \widehat{p}_i - 1 + \log i!\bigr)\\ &\qquad + \frac{y}{2}\bigl(\log y + \log\lambda - 1\bigr) + \mathrm{o}(1)\\ &= -\sum_i \widehat{p}_i \log(\widehat{p}_i i!) + \frac{y}{2}\bigl(\log y + \log\lambda - 1\bigr) + \mathrm{o}(1).\end{aligned} \tag{8.23}$$

For each i, $\widehat{p}_i \to \overline{p}_i$ by (8.16), and in addition, by (8.19) and (8.20),

$$y = \sum_i i\widehat{p}_i \to \sum_i i\overline{p}_i = \mu^2/\lambda. \tag{8.24}$$

Now, $x\log x \geq -e^{-1}$ on $[0,\infty)$, whence $\widehat{p}_i \log(\widehat{p}_i i!) \geq -e^{-1}/i!$, and by (8.21),

$$\widehat{p}_i \log(\widehat{p}_i i!) \leq \widehat{p}_i \log i! \leq \widehat{p}_i i \log i \leq \widehat{p}_i i^2 = \mathrm{O}\bigl(i^2 C_{10}^i/i!\bigr).$$

Consequently, by dominated convergence,

$$\frac{1}{n}\log z(\widehat{\mathbf{n}}) \to -\sum_i \overline{p}_i \log(\overline{p}_i i!) + \frac{\mu^2}{2\lambda}\Bigl(\log\frac{\mu^2}{\lambda} + \log\lambda - 1\Bigr). \tag{8.25}$$

Furthermore, by (8.15) and (8.24), if $\mu > 0$,

$$\sum_i \overline{p}_i \log(\overline{p}_i i!) = \sum_i \overline{p}_i\bigl(i\log\mu - \log\phi_{\mathcal{S}}(\mu)\bigr) = \frac{\mu^2}{\lambda}\log\mu - \log\phi_{\mathcal{S}}(\mu); \tag{8.26}$$

if $\mu = 0$ this holds trivially with all terms zero. Hence, (8.25) yields, using (8.20),

$$\begin{aligned}\frac{1}{n}\log z(\widehat{\mathbf{n}}) &\to -\frac{\mu^2}{\lambda}\log\mu + \log\phi_{\mathcal{S}}(\mu) + \frac{\mu^2}{2\lambda}\bigl(2\log\mu - 1\bigr)\\ &= \log\phi_{\mathcal{S}}(\mu) - \frac{\mu^2}{2\lambda} = \log\phi_{\mathcal{S}}(\mu) - \frac{\mu\phi'_{\mathcal{S}}(\mu)}{2\phi_{\mathcal{S}}(\mu)}. \end{aligned}\tag{8.27}$$

Conversely, take any finite sequence $(p_i)_{i=0}^M$ with $p_i \ge 0$, $\sum_i p_i = 1$ and $p_i = 0$ when $i \notin \mathcal{S}$. Define $n_i = nx_i$, rounded up or down to integers, preserving $\sum_i n_i = n$ and possibly adjusting two of them by ± 1 so that $\sum_i in_i$ is even. As $n \to \infty$, we then obtain as in (8.23)–(8.25) (but simpler, since the sums are finite), with $\nu := \sum_i ip_i$,

$$\frac{1}{n}\log z(\mathbf{n}) \to -\sum_i p_i \log(p_i i!) + \frac{\nu}{2}\Bigl(\log\nu + \log\lambda - 1\Bigr).$$

Since, by definition, $z(\widehat{\mathbf{n}})$ is maximal, $z(\widehat{\mathbf{n}}) \ge z(\mathbf{n})$, and thus

$$\liminf_{n\to\infty}\frac{\log z(\widehat{\mathbf{n}})}{n} \ge -\sum_i p_i\log(p_i i!) + \frac{\nu}{2}\Bigl(\log\nu + \log\lambda - 1\Bigr). \tag{8.28}$$

We have shown (8.28) for any probability distribution (p_i) on $\mathcal{S}$ with finite support. More generally, let (p_i) be a probability distribution supported on $\mathcal{S}$ with $\nu := \sum_i ip_i < \infty$ and $\sum_i p_i\log(p_i i!) < \infty$. For $M \ge \min\mathcal{S}$, let $p_i^{(M)} := p_i/\sum_{j\le M} p_j$ for $i \le M$ and apply (8.28) to $(p_i^{(M)})_{i=0}^M$. It is easily seen that the right hand side of (8.28) converges as $M \to \infty$ to the corresponding value for (p_i), showing that (8.28) holds for (p_i) also.

In particular, for any $\mu \in E(\lambda)$, we can use (8.28) with $p_i = \mathrm{Po}_{\mathcal{S}}(\mu)\{i\}$ given by (2.3) and, by (2.4),

$$\nu := \sum_i i\,\mathrm{Po}_{\mathcal{S}}(\mu)\{i\} = \frac{\mu\phi'_{\mathcal{S}}(\mu)}{\phi_{\mathcal{S}}(\mu)} = \frac{\mu^2}{\lambda}.$$

Hence, (8.28) yields, by the calculations in (8.26) and (8.27),

$$\liminf_{n\to\infty}\frac{\log z(\widehat{\mathbf{n}})}{n} \ge \log\phi_{\mathcal{S}}(\mu) - \frac{\mu\phi'_{\mathcal{S}}(\mu)}{2\phi_{\mathcal{S}}(\mu)} = \psi_{\mathcal{S}}(\mu), \tag{8.29}$$

for every $\mu \in E(\lambda)$. Comparing this to (8.27), we see that if (8.16) holds for some subsequence and some $\mu \in E(\lambda)$, then this μ must maximize $\psi_{\mathcal{S}}(\mu) = \psi_{\mathcal{S},1}(\mu;\lambda)$ over $E(\lambda)$, in other words, $\mu = \widehat{\mu}$ as defined in Theorem 2.1. In particular, this shows that every subsequence possesses a subsequence such that (8.16) holds with a fixed $\mu = \widehat{\mu}$; hence (8.16) holds for the full sequence of $n \to \infty$, and

$$\widehat{\mathbf{p}} = (\widehat{p}_j)_0^\infty \to \mathrm{Po}_{\mathcal{S}}(\widehat{\mu}). \tag{8.30}$$

Remark 8.3. We have for simplicity considered $\mu \in E(\lambda)$ only in (8.29); for general $\mu \geq 0$ the lower bound obtained from this argument takes the form, with $\nu = \mu\phi'_{\mathcal{S}}(\mu)/\phi_{\mathcal{S}}(\mu)$,

$$\log \phi_{\mathcal{S}}(\mu) + \frac{\nu}{2}\Bigl(\log \frac{\nu\lambda}{\mu^2} - 1\Bigr) = \psi_{\mathcal{S},2}(\mu;\lambda)$$

defined and studied in Section 4.

We have so far studied the mode $\widehat{\mathbf{p}}$ of the degree distribution. We now show that the distribution is concentrated close to the mode.

Lemma 8.4. *For every $\varepsilon > 0$, there exists $c_7 = c_7(\varepsilon) > 0$ such that, if n is large enough then for every $\mathbf{n} \in \mathcal{N}_{\mathcal{S}}^n$ with $d_{\mathrm{TV}}\bigl(\mathbf{n}/n, \mathrm{Po}_{\mathcal{S}}(\widehat{\mu})\bigr) \geq \varepsilon$, $z(\mathbf{n}) \leq e^{-c_7 n} z(\widehat{\mathbf{n}})$.*

We will first show a weaker statement.

Lemma 8.5. *For every $\varepsilon > 0$, there exists $c_8 = c_8(\varepsilon) > 0$ such that, if n is large enough then for every $\mathbf{n} \in \mathcal{N}_{\mathcal{S}}^n$ with $d_{\mathrm{TV}}\bigl(\mathbf{n}/n, \mathrm{Po}_{\mathcal{S}}(\widehat{\mu})\bigr) \geq \varepsilon$, either $z(\mathbf{n}) \leq e^{-c_8 n} z(\widehat{\mathbf{n}})$, or there exists $\mathbf{n}' \in \mathcal{N}_{\mathcal{S}}^n$ with $d_{\mathrm{TV}}(\mathbf{n}/n, \mathbf{n}'/n) \leq 2/n$ and $z(\mathbf{n}) \leq (1 - c_8) z(\mathbf{n}')$.*

Proof. Suppose this fails. Then there exists $\varepsilon > 0$ and a sequence $\mathbf{n} = \mathbf{n}^{(n)} \in \mathcal{N}_{\mathcal{S}}^n$ with $n \to \infty$, such that $d_{\mathrm{TV}}\bigl(\mathbf{n}/n, \mathrm{Po}_{\mathcal{S}}(\widehat{\mu})\bigr) \geq \varepsilon$, $z(\mathbf{n}) = e^{-\mathrm{o}(n)} z(\widehat{\mathbf{n}})$ and $z(\mathbf{n}') \leq (1 + \mathrm{o}(1)) z(\mathbf{n})$ for all $\mathbf{n}' \in \mathcal{N}_{\mathcal{S}}^n$ with $d_{\mathrm{TV}}(\mathbf{n}/n, \mathbf{n}'/n) \leq 2/n$, i.e., for all $\mathbf{n}' \in \mathcal{N}_{\mathcal{S}}^n$ with $\sum_i |n_i - n'_i| \leq 4$.

We now repeat much of the arguments presented above for the mode $\widehat{\mathbf{n}}$. First, we obtain that (8.10), (8.11) and (8.17) hold for these $\mathbf{n}$, with an extra factor $1 + \mathrm{o}(1)$ on the right hand sides, uniformly in all $j, k \in \mathcal{S}$ (with $k \geq j$ for (8.11) and $k \geq j$ and $k \equiv j \pmod 2$ for (8.17)). Furthermore, by Lemma 8.2 and the assumption $z(\mathbf{n}) = e^{-\mathrm{o}(n)} z(\widehat{\mathbf{n}})$, $N := \frac{1}{2}\sum_i i n_i \leq Bn$ (for large n). It follows, as above, that by considering a subsequence we may assume that $2N/n \to \nu \in [0, \infty)$ and $\mathbf{n}/n \to \overline{\mathbf{p}}$ for some probability distribution $\overline{\mathbf{p}}$, where, again as above, necessarily $\overline{\mathbf{p}} = \mathrm{Po}_{\mathcal{S}}(\mu)$ for some $\mu \in E(\lambda)$ and (8.27) holds for $\mathbf{n}$.

Since $\log z(\mathbf{n}) = \mathrm{o}(n) + \log z(\widehat{\mathbf{n}})$, this shows that $\psi_{\mathcal{S}}(\mu) = \psi_{\mathcal{S}}(\widehat{\mu})$, and thus $\mu = \widehat{\mu}$, since $\widehat{\mu}$ is assumed to be a unique maximum point. Consequently, $\mathbf{n}/n \xrightarrow{\mathrm{p}} \mathrm{Po}_{\mathcal{S}}(\widehat{\mu})$, which contradicts $d_{\mathrm{TV}}\bigl(\mathbf{n}/n, \mathrm{Po}_{\mathcal{S}}(\widehat{\mu})\bigr) \geq \varepsilon$. ∎

Proof of Lemma 8.4. By Lemma 8.5, if n is large enough, $d_{\mathrm{TV}}\bigl(\mathbf{n}/n, \mathrm{Po}_{\mathcal{S}}(\widehat{\mu})\bigr) \geq \varepsilon$, and $z(\mathbf{n}) > e^{-c_8 n} z(\widehat{\mathbf{n}})$, there exists $\mathbf{n}^{(1)} = \mathbf{n}'$ such that $d_{\mathrm{TV}}(\mathbf{n}/n, \mathbf{n}^{(1)}/n) \leq 2/n$ and $z(\mathbf{n}) \leq (1 - c_8) z(\mathbf{n}^{(1)})$; in particular, $z(\mathbf{n}^{(1)}) > z(\mathbf{n})$.

If also $d_{\mathrm{TV}}\big(\mathbf{n}^{(1)}/n, \mathrm{Po}_{\mathcal{S}}(\widehat{\mu})\big) \geq \varepsilon$, we iterate and find $\mathbf{n}^{(2)}$, and so on. This gives a sequence $\mathbf{n}^{(0)} = \mathbf{n}, \mathbf{n}^{(1)}, \dots, \mathbf{n}^{(L)}$, where for $l < L$ we have $d_{\mathrm{TV}}(\mathbf{n}^{(l)}/n, \mathbf{n}^{(l+1)}/n) \leq 2/n$ and $z(\mathbf{n}^{(l)}) \leq (1-c_8)z(\mathbf{n}^{(l+1)})$, while

$$d_{\mathrm{TV}}\big(\mathbf{n}^{(L)}/n, \mathrm{Po}_{\mathcal{S}}(\widehat{\mu})\big) < \varepsilon.$$

If further $d_{\mathrm{TV}}\big(\mathbf{n}/n, \mathrm{Po}_{\mathcal{S}}(\widehat{\mu})\big) \geq 2\varepsilon$, it follows that $d_{\mathrm{TV}}\big(\mathbf{n}/n, \mathbf{n}^{(L)}/n\big) > \varepsilon$, and thus the number of steps $L > \varepsilon n/2$. Consequently,

$$z(\mathbf{n}) \leq (1-c_8)^L z\big(\mathbf{n}^{(L)}\big) \leq (1-c_8)^L z(\widehat{\mathbf{n}}) \leq \exp\big(-\tfrac{1}{2}c_8\varepsilon n\big) z(\widehat{\mathbf{n}}).$$

This proves Lemma 8.4 for 2ε, with $c_7(2\varepsilon) = \min\{1, \frac{1}{2}\varepsilon\}c_8(\varepsilon)$. □

We now complete the proof of Theorem 7.3. Let $\varepsilon \geq 0$ and let, with B as in Lemma 8.2 and $N = \frac{1}{2}\sum_i in_i$,

$$A_1 := \{\mathbf{n} \in \mathcal{N}_{\mathcal{S}}^n : N > Bn\},$$
$$A_2^{\varepsilon} := \{\mathbf{n} \in \mathcal{N}_{\mathcal{S}}^n : N \leq Bn \text{ and } d_{\mathrm{TV}}\big(\mathbf{n}/n, \mathrm{Po}_{\mathcal{S}}(\widehat{\mu})\big) \geq \varepsilon\}.$$

Lemma 8.6. *For every $\varepsilon \geq 0$, $|A_2^{\varepsilon}| \leq |A_2^0| = e^{\mathrm{o}(n)}$.*

Proof. Suppose $\mathbf{n} = (n_i)_i \in A_2^0$. For each i, $n_i \leq n$, and thus the number of choices of $(n_i)_{i\leq\sqrt{n}}$ is at most $(n+1)^{\sqrt{n}+1} = \exp\big(\mathrm{O}(\sqrt{n}\log n)\big)$. Furthermore, $\sum_i in_i = 2N \leq 2Bn$, and thus $n_i = 0$ for $i > 2Bn$ and

$$\sum_{i>\sqrt{n}} n_i \leq \frac{2Bn}{\sqrt{n}} = 2B\sqrt{n},$$

so $(n_i)_{\sqrt{n}<i\leq 2BN}$ may be described by a sequence of at most $2B\sqrt{n}$ numbers in the range $[\sqrt{n}, 2Bn]$ (the degrees of the corresponding vertices). Hence, the number of choices of $(n_i)_{i>\sqrt{n}}$ is at most $(2Bn)^{2B\sqrt{n}} = \exp\big(\mathrm{O}(\sqrt{n}\log n)\big)$.

Combining the two parts, $|A_2^0| = \exp\big(\mathrm{O}(\sqrt{n}\log n)\big)$. □

Now, fix $\varepsilon > 0$. By Lemmas 8.2, 8.4 and 8.6,

$$\mathbb{P}\left(d_{\mathrm{TV}}\big(\boldsymbol{\pi}(G_{n,\lambda_n/n;\mathcal{S}}), \mathrm{Po}_{\mathcal{S}}(\widehat{\mu})\big) \geq \varepsilon\right) \leq e^{-n} + \frac{|A_2^{\varepsilon}|e^{-c_7 n}z(\widehat{\mathbf{n}})}{Z^*_{n,\lambda/n;\mathcal{S}}} \leq e^{-c_1 n},$$

for some $c_1 > 0$ and all large n. This proves (2.12) and hence (2.9) (for $G^*_{n,\lambda/n;\mathcal{S}}$).

A similar calculation with $\varepsilon = 0$ yields

$$z(\widehat{\mathbf{n}}) \leq Z^*_{n,\lambda/n;\mathcal{S}} = \sum_{\mathbf{n}\in A_1} z(\mathbf{n}) + \sum_{\mathbf{n}\in A_2^0} z(\mathbf{n}) \leq e^{\mathrm{o}(n)} z(\widehat{\mathbf{n}}),$$

and thus $\log Z^*_{n,\lambda/n;\mathcal{S}} = \log z(\widehat{\mathbf{n}}) + \mathrm{o}(n)$, which together with (8.27) implies $n^{-1}\log Z^*_{n,\lambda/n;\mathcal{S}} \to \psi_{\mathcal{S}}(\widehat{\mu})$. By (7.3), this further yields (7.5).

Lemma 8.7. *Uniformly in all $k \geq 0$,*

$$\mathbb{E}\, n_k(G^*_{n,\lambda/n;\mathcal{S}}) \leq C_{11} n e^{-c_9 k}.$$

Moreover, for any λ_0, this holds uniformly in $\lambda \leq \lambda_0$.

Proof. Let j be the smallest element of $\mathcal{S}$ with the same parity as k. Given any $\mathbf{n} \in \mathcal{N}_{\mathcal{S}}^n$, let $\mathbf{n}' \in \mathcal{N}_{\mathcal{S}}^n$ be given by $n'_j := n_j + 1$, $n'_k := n_k - 1$ and $n'_i := n_i$, $i \neq j, k$ (assuming $j < k$ and $n_k \geq 1$; otherwise $\mathbf{n}' = \mathbf{n}$). By (8.3), cf. the argument yielding (8.17),

$$z(\mathbf{n}) \leq z(\mathbf{n}') \frac{(n_j+1)j!}{n_k k!} \Big(\frac{2\lambda N}{n}\Big)^{(k-j)/2},$$

and thus

$$n_k z(\mathbf{n}) \leq \frac{C_{12} n}{k!} \Big(\frac{2\lambda N}{n}\Big)^{(k-j)/2} z(\mathbf{n}'). \tag{8.31}$$

Lemma 8.2 and (8.31) imply

$$\begin{aligned}
\sum_{\mathbf{n} \in \mathcal{N}_{\mathcal{S}}^n} n_k z(\mathbf{n}) &= \sum_{\mathbf{n} \in A_1} n_k z(\mathbf{n}) + \sum_{\mathbf{n} \in A_2^0,\, n_k > 0} n_k z(\mathbf{n}) \\
&\leq n e^{-n} z(\widehat{\mathbf{n}}) + \frac{C_{12} n}{k!} (2B\lambda)^{(k-j)/2} \sum_{\mathbf{n}' \in \mathcal{N}_{\mathcal{S}}^n} z(\mathbf{n}') \\
&\leq n e^{-n} Z^*_{n,\lambda/n;\mathcal{S}} + C_{13} \frac{C_{14}^k n}{k!} Z^*_{n,\lambda/n;\mathcal{S}}
\end{aligned}$$

and the result for $0 \leq k \leq 2Bn$ follows by dividing by $Z^*_{n,\lambda/n;\mathcal{S}}$, with $c_9 = 1/(2B)$.

Finally, if $k > 2Bn$, then for every $i \geq 1$ we have $n_k \geq i \implies N \geq kn_k/2 \geq ki/2 > Bn$, and thus by Lemma 8.2

$$\mathbb{P}\big(n_k(G^*_{n,\lambda/n;\mathcal{S}}) \geq i\big) \leq \mathbb{P}\big(\mathrm{e}(G^*_{n,\lambda/n;\mathcal{S}}) \geq ki/2\big) \leq e^{-ki/(2B)}.$$

Hence, $\mathbb{E}\, n_k(G^*_{n,\lambda/n;\mathcal{S}}) = \sum_{i=1}^{\infty} \mathbb{P}\big(n_k(G^*_{n,\lambda/n;\mathcal{S}}) \geq i\big) \leq 2e^{-k/(2B)}$. □

Let $X_{n,K} := n^{-1} \sum_{k=0}^{K} k^r n_k(G^*_{n,\lambda/n;\mathcal{S}})$ be a partial sum of the sum in (2.10) for $G^*_{n,\lambda/n;\mathcal{S}}$. Then, for every fixed K, by Lemma 8.7,

$$\mathbb{E}(X_{n,\infty} - X_{n,K}) = \mathbb{E}\left(\sum_{k=K+1}^{\infty} \frac{k^r n_k(G^*_{n,\lambda/n;\mathcal{S}})}{n}\right) \leq \sum_{k=K+1}^{\infty} C_{11} k^r e^{-c_9 k},$$

which can be made arbitrarily small by choosing K large. Since $X_{n,K} \xrightarrow{\mathrm{p}} \sum_{k=0}^{K} k^r \operatorname{Po}_{\mathcal{S}}(\widehat{\mu})\{k\}$ as $n \to \infty$ for every fixed K, (2.10) follows by standard arguments. (See, for example, the much more general [3, Theorem 4.2].) This completes the proof of Theorem 7.3.

9. Proof of Lemma 7.4 and Theorem 2.1

Proof of Lemma 7.4. We use Lemma 7.2 together with the result of [8] (with previous partial results by many authors) that states that, for a sequence of degree sequences $\mathbf{d} = \mathbf{d}^{(n)}$ satisfying $\sum_i d_i \to \infty$, if $\sum_i d_i^2 = O\big(\sum_i d_i\big)$, then

$\liminf \mathbb{P}\bigl(G^*_{n,\mathbf{d}} \text{ is simple}\bigr) > 0$. (The converse holds also, see [8].) In other words, for every K there exist constants a_K and $b_K > 0$ such that, if

$$\text{(i)} \quad \sum_i d_i \geq a_K \qquad \text{and} \qquad \text{(ii)} \quad \sum_i d_i^2 \leq K \sum_i d_i, \tag{9.1}$$

then

$$\mathbb{P}\bigl(G^*_{n,\mathbf{d}} \text{ is simple}\bigr) \geq b_K. \tag{9.2}$$

Let $p(\mathbf{d}) := \mathbb{P}\bigl(G^*_{n,\mathbf{d}} \text{ is simple}\bigr)$. By Lemma 7.2, for every K,

$$\begin{aligned} \mathbb{P}\bigl(G^*_{n,\lambda_n/n;\mathcal{S}} \text{ is simple}\bigr) &= \mathbb{E}\, p\bigl(\mathbf{d}(G^*_{n,\lambda_n/n;\mathcal{S}})\bigr) \\ &\geq b_K\, \mathbb{P}\bigl(\mathbf{d}(G^*_{n,\lambda_n/n;\mathcal{S}}) \text{ satisfies (9.1)}\bigr). \end{aligned} \tag{9.3}$$

Thus, it suffices to show that $\liminf \mathbb{P}\bigl(\mathbf{d}(G^*_{n,\lambda_n/n;\mathcal{S}}) \text{ satisfies (9.1)}\bigr) > 0$.

First, consider the case $\widehat{\mu} > 0$. By Theorem 7.3,

$$\frac{1}{n} \sum_i d_i(G^*_{n,\lambda_n/n;\mathcal{S}})^r \xrightarrow{\text{p}} A_r, \qquad r = 1, 2,$$

for some constants $A_r > 0$. Hence, taking $K := A_2/A_1 + 1$,

$$\mathbb{P}\bigl(\mathbf{d}(G^*_{n,\lambda_n/n;\mathcal{S}}) \text{ satisfies (9.1)}\bigr) \to 1$$

and the result follows in this case.

Now suppose that $\widehat{\mu} = 0$, which can occur only if $0 \in \mathcal{S}$. Although the graphs are sparser in this case, and intuitively it seems more probable that they are simple, we have not found a really simple proof and have to work harder in this case. (The proof above is not valid since now $A_1 = A_2 = 0$.) Let $\mathcal{S}' := \mathcal{S} \setminus \{0\}$, and let $M := n - n_0$ be the number of non-isolated vertices in $G^*_{n,\lambda_n/n;\mathcal{S}}$.

Let $0 \leq m \leq n$ and let V be any subset of $\{1, 2, \dots, n\}$ with $|V| = m$. If we condition $G^*_{n,\lambda_n/n;\mathcal{S}}$ on having the set of non-isolated vertices equal to V, we evidently get a random multigraph $G^*_{m,\lambda_n/n;\mathcal{S}'}$ on V (up to relabelling the vertices) with $n - m$ isolated vertices added. It follows that, for $r > 0$,

$$\Bigl(\sum_i d_i(G^*_{n,\lambda_n/n;\mathcal{S}})^r \Bigm| M = m\Bigr) \overset{\text{d}}{=} \sum_i d_i(G^*_{m,\lambda_n/n;\mathcal{S}'})^r. \tag{9.4}$$

Note that the relevant parameter of $G^*_{m,\lambda_n/n;\mathcal{S}'}$ is $m\lambda_n/n$, not λ_n. Since we consider $0 \leq m \leq n$, and the case $m = 0$ is trivial and thus can be ignored, we have $0 < m\lambda_n/n \leq \lambda_n \leq C_{15}$, and thus Lemma 8.7 implies that

$$\mathbb{E} \sum_i d_i(G^*_{m,\lambda_n/n;\mathcal{S}'})^2 = \mathbb{E} \sum_{k=1}^{\infty} k^2 n_k\bigl(G^*_{m,\lambda_n/n;\mathcal{S}'}\bigr) \leq C_{16} m, \qquad m \leq n,$$

for some constant C_{16} not depending on m or n. Furthermore, since $0 \notin \mathcal{S}'$, each vertex degree is at least 1 and $\sum_i d_i(G^*_{m,\lambda_n/n;\mathcal{S}'}) \geq m$. Consequently, choosing $K = 4C_{16}$, it follows by Markov's inequality that, for every $m \le n$, with probability at least $\frac{3}{4}$,

$$\sum_i d_i(G^*_{m,\lambda_n/n;\mathcal{S}'})^2 \le 4C_{16}m \le K \sum_i d_i(G^*_{m,\lambda_n/n;\mathcal{S}'}).$$

Consequently, by conditioning on M and using (9.4),

$$\mathbb{P}\big((9.1)\text{(ii) holds for } G^*_{n,\lambda_n/n;\mathcal{S}}\big) \ge \tfrac{3}{4}.$$

Hence, whenever

$$\mathbb{P}\bigg(\sum_i d_i(G^*_{n,\lambda_n/n;\mathcal{S}}) \ge a_K\bigg) \ge \tfrac{1}{2}, \tag{9.5}$$

then (9.1) fails for $G^*_{n,\lambda_n/n;\mathcal{S}}$ with probability at most $\frac{1}{2} + \frac{1}{4} = \frac{3}{4}$, and thus by (9.3)

$$\mathbb{P}\big(G^*_{n,\lambda_n/n;\mathcal{S}} \text{ is simple}\big) \ge \tfrac{1}{4} b_K.$$

The only remaining case is when $\widehat{\mu} = 0$ and (9.5) is false. We recall $\sum_i d_i(G) = \sum_j j n_j(G)$ and define $\mathcal{N}^* := \{\mathbf{n} \in \mathcal{N}^n_{\mathcal{S}} : \sum_j j n_j < a_K\}$. Thus we now have, using (8.2),

$$\tfrac{1}{2} < \mathbb{P}\big(\mathbf{n}(G^*_{n,\lambda_n/n;\mathcal{S}}) \in \mathcal{N}^*\big) = \frac{1}{Z^*_{n,\lambda_n/n;\mathcal{S}}} \sum_{\mathbf{n} \in \mathcal{N}^*} z(\mathbf{n}). \tag{9.6}$$

Further, $\mathcal{N}^*$ is a finite set (with $n_j = 0$ for $j \ge a_K$), and (8.3) yields

$$z(\mathbf{n}) \le \frac{n!}{n_0!}(\lceil a_K \rceil)!! \Big(\frac{\lambda_n}{n}\Big)^{\frac{1}{2}\sum j n_j} \le C_{17} n^{n-n_0} \Big(\frac{\lambda_n}{n}\Big)^{\frac{1}{2}\sum j n_j}, \quad \mathbf{n} \in \mathcal{N}^*. \tag{9.7}$$

We now use Theorem 4.12, which shows that $\widehat{\mu} = 0$ is possible only when $1 \notin \mathcal{S}$. Thus, if $\mathbf{n} \in \mathcal{N}^* \subset \mathcal{N}_{\mathcal{S}}$, then $n_1 = 0$ and $\frac{1}{2}\sum_j j n_j \ge \sum_2^\infty n_j = n - n_0$; hence, since $\lambda_n = O(1) = O(n)$, by (9.7),

$$z(\mathbf{n}) \le C_{18} n^{n-n_0} \Big(\frac{\lambda_n}{n}\Big)^{n-n_0} \le C_{19}, \qquad \mathbf{n} \in \mathcal{N}^*.$$

Therefore, by (9.6),

$$Z^*_{n,\lambda_n/n;\mathcal{S}} \le 2 \sum_{\mathbf{n} \in \mathcal{N}^*} z(\mathbf{n}) \le C_{20}.$$

However, if E_n is the empty graph with n vertices and no edges, then by (7.2), $\mathbb{P}\big(G^*_{n,\lambda_n/n;\mathcal{S}} = E_n\big) = 1/Z^*_{n,\lambda_n/n;\mathcal{S}} \ge C_{20}^{-1}$. Now E_n is simple, and so the result follows in this case too. $\square$

Proof of Theorem 2.1. By Lemmas 7.1 and 7.4, for any event $\mathcal{E}$ and n large enough,

$$\mathbb{P}\big(G_{n,\lambda_n/n;\mathcal{S}} \in \mathcal{E}\big) \le \frac{\mathbb{P}(G^*_{n,\lambda_n/n;\mathcal{S}} \in \mathcal{E})}{\mathbb{P}(G^*_{n,\lambda_n/n;\mathcal{S}} \text{ is simple})} \le C_{21}\, \mathbb{P}(G^*_{n,\lambda_n/n;\mathcal{S}} \in \mathcal{E}).$$

Hence parts (i)–(iii) follow directly from Theorem 7.3. Similarly,

$$\begin{aligned}
&\mathbb{P}\big(G_{n,\lambda_n/n} \text{ is an } \mathcal{S}\text{-graph}\big) = \mathbb{P}\big(G^*_{n,\lambda_n/n} \text{ is an } \mathcal{S}\text{-graph} \mid G^*_{n,\lambda_n/n} \text{ is simple}\big)\\
&\quad = \mathbb{P}\big(G^*_{n,\lambda_n/n} \text{ is simple} \mid G^*_{n,\lambda_n/n} \text{ is an } \mathcal{S}\text{-graph}\big)\frac{\mathbb{P}\big(G^*_{n,\lambda_n/n} \text{ is an } \mathcal{S}\text{-graph}\big)}{\mathbb{P}\big(G^*_{n,\lambda_n/n} \text{ is simple}\big)}\\
&\quad = \frac{\mathbb{P}\big(G^*_{n,\lambda_n/n;\mathcal{S}} \text{ is simple}\big)}{\mathbb{P}\big(G^*_{n,\lambda_n/n} \text{ is simple}\big)}\,\mathbb{P}\big(G^*_{n,\lambda_n/n} \text{ is an } \mathcal{S}\text{-graph}\big)
\end{aligned}$$

and (iv) follows by Theorem 7.3 and Lemma 7.4 (applied both to $\mathcal{S}$ and with $\mathcal{S}$ replaced by $\mathbb{Z}_{\geq 0}$). ☐

10. Proofs of Theorems 3.1 and 3.6

Proof of Theorem 3.1. The case $\widehat{\mu} = 0$ is trivial by Theorem 2.1, as remarked in Remark 3.3, so we will assume $\widehat{\mu} > 0$. We use the results of Molloy and Reed [14, 15] in the following version, see Janson and Luczak [11], Theorem 2.3 and Remark 2.7; we only consider the limiting degree distribution $(p_k)_{k=0}^{\infty}$ given by $p_k = \mathrm{Po}_{\mathcal{S}}(\widehat{\mu})\{k\}$.

Let $\nu = \nu(\widehat{\mu})$, $Q = Q(\widehat{\mu})$, $\widehat{\xi}$, $\widehat{\gamma}$ and $\widehat{\zeta}$ be as in Section 3; the existence of a unique solution $\widehat{\xi} \in (0, 1)$ in (i) follows by [11, Lemma 5.5]. By assumption, $p_0 + p_2 < 1$. Further, let $G_{n,\mathbf{d}}$ be the random graph with given degree sequence $\mathbf{d}$, chosen uniformly among all such graphs (assuming that there is at least one), and let $\Gamma_{n,\mathbf{d}}$ and $\Gamma^{(2)}_{n,\mathbf{d}}$ be the largest and second largest components of $G_{n,\mathbf{d}}$.

Theorem 10.1. *Suppose that, for each n, $\mathbf{d} = (d_i)_1^n$ is a sequence of non-negative integers such that $\sum_{i=1}^n d_i$ is even, and that*

(i) *$|\{i : d_i = k\}|/n \to p_k$ as $n \to \infty$, for every $k \geq 0$;*
(ii) *$\sum_{i=1}^n d_i^2 = \mathrm{O}(n)$.*

Then, the following hold for the random graph $G_{n,\mathbf{d}}$, as $n \to \infty$:

$$\mathrm{v}(\Gamma_{n,\mathbf{d}})/n \xrightarrow{\mathrm{p}} \widehat{\gamma}, \qquad \mathrm{e}(\Gamma_{n,\mathbf{d}})/n \xrightarrow{\mathrm{p}} \widehat{\zeta},$$
$$\mathrm{v}(\Gamma^{(2)}_{n,\mathbf{d}})/n \xrightarrow{\mathrm{p}} 0, \qquad \mathrm{e}(\Gamma^{(2)}_{n,\mathbf{d}})/n \xrightarrow{\mathrm{p}} 0.$$

This theorem is stated as a limit result, but it can be reformulated as follows.

Theorem 10.2. *For every $\varepsilon > 0$ and $C < \infty$, there exists $\delta > 0$ such that if $n > \delta^{-1}$ and $\mathbf{d} = (d_i)_1^n$ is a degree sequence such that $\sum_{i=1}^n d_i$ is even and*

(i) *$\sum_{k=0}^{\infty} \big||\{i : d_i = k\}|/n - p_k\big| < \delta$,*
(ii) *$\sum_{i=1}^n d_i^2 \leq Cn$,*

then

$$\mathbb{P}\big(|\mathrm{v}(\Gamma_{n,\mathbf{d}})/n - \widehat{\gamma}| > \varepsilon\big) < \varepsilon, \qquad \mathbb{P}\big(|\mathrm{e}(\Gamma_{n,\mathbf{d}})/n - \widehat{\zeta}| > \varepsilon\big) < \varepsilon,$$
$$\mathbb{P}\big(\mathrm{v}(\Gamma^{(2)}_{n,\mathbf{d}})/n > \varepsilon\big) < \varepsilon, \qquad \mathbb{P}\big(\mathrm{e}(\Gamma^{(2)}_{n,\mathbf{d}})/n > \varepsilon\big) < \varepsilon.$$

By Theorem 2.1, for every $\varepsilon > 0$, a suitable C and sufficiently large n, the random degree sequence $\mathbf{d}(G_{n,\lambda_n/n;\mathcal{S}})$ satisfies the conditions (i) and (ii) of Theorem 10.2 with probability at least $1-\varepsilon$.

Since $(G_{n,\lambda_n/n;\mathcal{S}} \mid \mathbf{d}(G_{n,\lambda_n/n;\mathcal{S}} = \mathbf{d}) \overset{\mathrm{d}}{=} G_{n,\mathbf{d}}$ by Lemmas 7.2 and 7.1, it follows that $\mathbb{P}\bigl(|\mathrm{v}(\Gamma_{n,\lambda_n/n;\mathcal{S}})/n - \widehat{\gamma}| > \varepsilon\bigr) < 2\varepsilon$ if n is large enough, and similarly for $\mathrm{e}(\Gamma_{n,\lambda_n/n;\mathcal{S}})$ and for $\Gamma^{(2)}_{n,\lambda_n/n;\mathcal{S}}$, which proves Theorem 3.1. □

Proof of Theorem 3.6. This proof is similar, using [10, Theorem 2.4]; we omit the details while noting that we now need condition (i) of Theorem 10.2, and in addition the condition (ii′) (stronger than (ii) above) that $\sum_{i=1}^n e^{cd_i} \le Cn$ for some $c > 0$. This holds for $\mathbf{d}(G_{n,\lambda_n/n;\mathcal{S}})$ with probability $> 1-\varepsilon$ for suitable c and C (that may depend on ε), as a consequence of the following corollary of Lemma 8.7.

Lemma 10.3. *Assume that $\lambda_n \to \lambda > 0$. If $c < c_9$, then*

$$\mathbb{E}\sum_{i=1}^n e^{cd_i(G^*_{n,\lambda_n/n;\mathcal{S}})} = \mathbb{E}\sum_{k=0}^\infty n_k(G^*_{n,\lambda_n/n;\mathcal{S}})e^{ck} \le C_{22}n.$$

By Lemmas 7.1 and 7.4, the conclusion of the lemma is valid for $G_{n,\lambda_n/n;\mathcal{S}}$ also. □

Acknowledgement

This research was mainly done during a visit by SJ to the University of Cambridge, partly funded by Trinity College.

References

[1] R. Arratia, A. D. Barbour and S. Tavaré, *Logarithmic Combinatorial Structures: a Probabilistic Approach*, EMS, Zürich, 2003.

[2] E. A. Bender & E. R. Canfield, The asymptotic number of labeled graphs with given degree sequences. *J. Combin. Theory Ser. A*, **24** (1978), 296–307.

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

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

[5] B. Bollobás, *Random Graphs*, 2nd edn, Cambridge Univ. Press, Cambridge, 2001.

[6] G. R. Grimmett, *The Random-Cluster Model*, Springer Verlag, 2006.

[7] G. R. Grimmett & S. Janson, Random even graphs and the Ising model. *Electronic J. Combin.* **16** (2009), Paper R46.

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

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

[10] S. Janson & M. Luczak, A simple solution to the k-core problem. *Random Struct. Alg.* **30** (2007), 50–62.

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

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

[13] T. Lindvall, *Lectures on the Coupling Method*, Dover Publications, New York, 2002.

[14] M. Molloy & B. Reed, A critical point for random graphs with a given degree sequence, *Random Struct. Alg.* **6** (1995), 161–179.

[15] 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.

[16] B. Pittel, J. Spencer & N. Wormald, Sudden emergence of a giant k-core in a random graph, *J. Combin. Theor. Ser. B* **67** (1996), 111–151.

[17] O. Riordan, The k-core and branching processes. *Combin. Probab. Comput.* **1**7 (2008), 111–136.

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

[19] N. C. Wormald, The asymptotic distribution of short cycles in random regular graphs. *J. Combin. Theory Ser. B* **31** (1981), 168–182.

Statistical Laboratory, Centre for Mathematical Sciences, Cambridge University, Wilberforce Road, Cambridge CB3 0WB, UK

E-mail address: `g.r.grimmett@statslab.cam.ac.uk`

URL: `http://www.statslab.cam.ac.uk/~grg/`

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/`