

Susceptibility in inhomogeneous random graphs

Svante Janson

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

`svante@math.uu.se`

Oliver Riordan

Mathematical Institute, University of Oxford
24–29 St Giles', Oxford OX1 3LB, UK

`riordan@maths.ox.ac.uk`

Submitted: Apr 20, 2010; Accepted: Jan 22, 2012; Published: Feb 7, 2012
Mathematics Subject Classifications: 05C80, 60C05

Abstract

We study the susceptibility, i.e., the mean size of the component containing a random vertex, in a general model of inhomogeneous random graphs. This is one of the fundamental quantities associated to (percolation) phase transitions; in practice one of its main uses is that it often gives a way of determining the critical point by solving certain linear equations. Here we relate the susceptibility of suitable random graphs to a quantity associated to the corresponding branching process, and study both quantities in various natural examples.

1 Introduction

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

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

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

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

Later we shall order the components, assuming as usual that $|\mathcal{C}_1| \geq |\mathcal{C}_2| \geq \dots$.

When the graph G is itself random, in some contexts (such as percolation) it is usual to take the expectation over G as well as over v . Here we do *not* do so: when G is random, $\chi(G)$ is a random variable.

Remark 1.1. The term susceptibility comes from physics; we therefore use the notation χ , which is standard in physics, although it usually means something else in graph theory. The connection with physics is through (e.g.) the Ising model for magnetism and the corresponding random-cluster model, which is a random graph where the susceptibility (1.2), or rather its expectation, corresponds to the magnetic susceptibility.

The susceptibility has been much studied for certain models in mathematical physics. In percolation theory, which deals with certain random infinite graphs, the corresponding quantity is the (mean) size of the open cluster containing a given vertex, and this has also been extensively studied; see e.g. Bollobás and Riordan [8]. In contrast, not much rigorous work has been done for finite random graphs. Some results for the Erdős–Rényi random graphs $G(n, p)$ and $G(n, m)$ can be regarded as folk theorems that have been known to experts for a long time. More recently, Durrett [19] proved that the susceptibility of $G(n, p)$ has expectation $\mathbb{E}\chi(G(n, p)) = (1 - \lambda)^{-1} + O(1/n)$ when $p = \lambda/n$ with $\lambda < 1$ fixed. The susceptibility of $G(n, p)$ and $G(n, m)$ has been studied in detail by Janson and Luczak [24]. For other graphs, one rigorous treatment is by Spencer and Wormald [33], who considered a class of random graph processes (including the Erdős–Rényi graph process) and used the susceptibility to study the phase transition in them.

The purpose of the present paper is to study $\chi(G^\nu(n, \kappa))$ for the inhomogeneous random graph $G^\nu(n, \kappa)$ introduced by Bollobás, Janson and Riordan [5]; this is a rather general model that includes $G(n, p)$ as a special case. In fact, much of the time we shall consider the even more general setting of [6]. We review the fundamental definitions from [5; 6] in Section 2 below.

We consider asymptotics as $n \rightarrow \infty$, and all unspecified limits are as $n \rightarrow \infty$. As usual, if G_n is a sequence of random graphs, we say that G_n has a certain property *with high probability*, or *whp*, if the probability that G_n has this property tends to 1 as $n \rightarrow \infty$.

Remark 1.2. We obtain results for $G(n, p)$ as corollaries of our general results, but note that these results are not (and cannot be, because of the generality of the model $G^\nu(n, \kappa)$) as precise as the results obtained by Janson and Luczak [24]. The proofs in the two papers are quite different: the proofs in [24] are based on studying the evolution of the susceptibility in the random graph process obtained by adding random edges one-by-one, using methods from stochastic process theory, while the present paper is based on the standard branching process approximation of the neighbourhood of a given vertex. It seems likely that the latter method can also be used to prove precise results in the special case of $G(n, p)$, but we have not attempted this. (Durrett [19] uses this method for the expectation $\mathbb{E}\chi(G(n, p))$.)

The definition (1.2) is mainly interesting in the subcritical case, when all components are rather small. In the supercritical case, there is typically one giant component that

is so large that it dominates the sum in (1.2), and thus $\chi(G) \sim |\mathcal{C}_1|^2/n$. In fact, in the supercritical case of [5, Theorem 3.1], whp $|\mathcal{C}_1| = \Theta(n)$ and $|\mathcal{C}_2| = o(n)$ and thus

$$\sum_{i=1}^K |\mathcal{C}_i|^2 = |\mathcal{C}_1|^2 + O\left(|\mathcal{C}_2| \sum_{i=2}^K |\mathcal{C}_i|\right) = |\mathcal{C}_1|^2 + O(|\mathcal{C}_2|n) = (1 + o(1))|\mathcal{C}_1|^2.$$

(See also [24, Appendix A] for $G(n, p)$.) In this case, it makes sense to exclude the largest component from the definition; this is in analogy with percolation theory, where one studies the mean size of the open cluster containing a fixed vertex, given that this cluster is finite. We thus define the *modified susceptibility* $\widehat{\chi}(G)$ of a finite graph G by

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

Note that we divide by n rather than by $n - |\mathcal{C}_1|$, which would also make sense.

In the uniform case, i.e., in the usual random graph process, one interpretation of $\widehat{\chi}(G)$ is that it gives the rate of growth of the giant component above the critical point. More generally, if we add a single new edge chosen uniformly at random to a graph G , then the probability that \mathcal{C}_i , $i \geq 2$, becomes joined to \mathcal{C}_1 is asymptotically $2|\mathcal{C}_i||\mathcal{C}_1|/n^2$, and when this happens $|\mathcal{C}_1|$ increases by $|\mathcal{C}_i|$. Thus (under suitable assumptions), the expected increase in $|\mathcal{C}_1|$ is asymptotically $2|\mathcal{C}_1| \sum_{i \geq 2} |\mathcal{C}_i|^2/n^2 = 2\widehat{\chi}(G)|\mathcal{C}_1|/n$.

The results in [5] on components of $G^\nu(n, \kappa)$ are based on approximation by a branching process \mathfrak{X}_κ , see Section 2. We define (at least when $\mu(\mathcal{S}) = 1$, see Section 2)

$$\chi(\kappa) := \mathbb{E}(|\mathfrak{X}_\kappa|) \in [0, \infty], \tag{1.4}$$

$$\widehat{\chi}(\kappa) := \mathbb{E}(|\mathfrak{X}_\kappa| 1_{\{|\mathfrak{X}_\kappa| < \infty\}}) \in [0, \infty], \tag{1.5}$$

where 1_A denotes the indicator function of an event A . Thus, $\chi(\kappa) = \widehat{\chi}(\kappa)$ when the survival probability $\rho(\kappa) := \mathbb{P}(|\mathfrak{X}_\kappa| = \infty)$ is zero (the subcritical or critical case), while $\chi(\kappa) = \infty \geq \widehat{\chi}(\kappa)$ when $\rho(\kappa) > 0$ (the supercritical case). Our main aim is to show that under suitable conditions, the [modified] susceptibility of $G^\nu(n, \kappa)$ converges to $\chi(\kappa)$ [$\widehat{\chi}(\kappa)$].

The rest of this paper is organized as follows. In Section 2 we define the models that we shall study, and recall some of their basic properties. We also describe the branching process \mathfrak{X}_κ and integral operator T_κ associated to the models, and collect together some facts from functional analysis that we shall need later. In Section 3 we relate the branching process analogues of the susceptibilities, defined in (1.4) and (1.5), to the operator T_κ . The heart of the paper is Section 4, where we show that, under certain assumptions, the susceptibilities of the random graphs are close to those of the branching process; see Theorems 4.7, 4.8 and 4.9. In Section 5 we study the behaviour of $\chi(\lambda\kappa)$ and $\widehat{\chi}(\lambda\kappa)$ as functions of the parameter $\lambda \in (0, \infty)$, and in particular the behaviour near the threshold for the existence of a giant component; this provides a way to use the susceptibility to find the threshold for the random graphs treated here. (See, e.g., Durrett [19] and Spencer and Wormald [33] for earlier uses of this method.) Finally, in Section 6 we

give some applications and examples of our results, including explicit calculations of the susceptibilities in certain much-studied special cases.

Remark 1.3. We believe that results similar to those proved here hold for the ‘higher order susceptibilities’

$$\chi_m(G) := \frac{1}{|G|} \sum_{v \in V(G)} |\mathcal{C}(v)|^m = \frac{1}{|G|} \sum_i |\mathcal{C}_i|^{m+1},$$

but we have not pursued this. (For $G(n, p)$, see [24].)

Acknowledgements. Part of this work was carried out during the programme “Combinatorics and Statistical Mechanics” at the Isaac Newton Institute, Cambridge, 2008, where SJ was supported by a Microsoft fellowship, and part during a visit of both authors to the programme “Discrete Probability” at Institut Mittag-Leffler, Djursholm, Sweden, 2009. The authors would like to thank a very thorough referee for many suggestions improving the presentation of the paper.

2 Preliminaries

We review the fundamental definitions from [5; 6], but refer to those papers for further details, as well as for references to previous work. In terms of motivation and applications, our main interest is the model $G^\mathcal{V}(n, \kappa)$ of [5], but for the proofs we sometimes need (or can handle) different generality. Throughout the paper we use standard graph theoretic notation as in [2]; for example, $|G|$ denotes the number of vertices in a graph G , and $e(G)$ the number of edges. The notation \xrightarrow{P} denotes convergence in probability. We use $x \wedge y$ as an alternative notation for $\min\{x, y\}$ when convenient.

2.1 The random graph models

All our random graphs will have vertex set $[n] = \{1, 2, \dots, n\}$.

By a *type space* (\mathcal{S}, μ) we mean a measure space with $0 < \mu(\mathcal{S}) < \infty$. The reason for the terminology is that in the graphs $G^\mathcal{V}(n, \kappa)$ defined below, each vertex i will have a *type* $x_i \in \mathcal{S}$. Often $\mathcal{S} = [0, 1]$ or $(0, 1]$, in which case μ is Lebesgue measure unless stated otherwise. Usually, $\mu(\mathcal{S}) = 1$, so (\mathcal{S}, μ) is a probability space. Then the measure μ describes the limiting distribution of the vertex types in a way that will be made precise below; the most natural case is when the types are i.i.d. with distribution μ .

The second key ingredient in all cases is a *kernel* on (\mathcal{S}, μ) , i.e., a symmetric non-negative measurable function $\kappa : \mathcal{S} \times \mathcal{S} \rightarrow [0, \infty)$. We assume throughout that all kernels κ are integrable, i.e., that $\int_{\mathcal{S}^2} \kappa(x, y) d\mu(x) d\mu(y) < \infty$.

2.1.1 The general inhomogeneous model.

To define $G^\mathcal{V}(n, \kappa)$, we assume that we are given, for each $n \geq 1$, a random or deterministic finite sequence $\mathbf{x}_n = (x_1, x_2, \dots, x_n)$ of elements of \mathcal{S} ; the elements will specify the types

of the vertices. To avoid clutter we write x_i rather than $x_i^{(n)}$, even though the vertex types for different n need not be related. We denote the triple $(\mathcal{S}, \mu, (\mathbf{x}_n)_{n \geq 1})$ by \mathcal{V} , so \mathcal{V} encodes the type space and, for each n , the (joint) distribution of the vertex types.

Given such a triple \mathcal{V} and a kernel κ on (\mathcal{S}, μ) , the random graph $G_n = G^\mathcal{V}(n, \kappa)$ is defined by first picking the type sequence $\mathbf{x}_n = (x_1, x_2, \dots, x_n)$ according to the distribution specified by \mathcal{V} and then, given \mathbf{x}_n , forming the graph on $[n]$ in which each possible edge ij , $i < j$, is present independently with probability

$$p_{ij} = \min\{\kappa(x_i, x_j)/n, 1\}. \quad (2.1)$$

(Alternatively, we may take $p_{ij} = 1 - \exp(-\kappa(x_i, x_j)/n)$; as shown in [5; 22], this is essentially equivalent.)

As in [5], to relate properties of $G^\mathcal{V}(n, \kappa)$ to properties of μ and κ some conditions must be imposed on \mathcal{V} and κ .

Let δ_x denote the measure assigning mass 1 to $x \in \mathcal{S}$ and zero elsewhere, and let ν_n be the (random) measure

$$\nu_n = n^{-1} \sum_{i=1}^n \delta_{x_i}. \quad (2.2)$$

Thus, for $A \subset \mathcal{S}$, $\nu_n(A)$ gives the (random) proportion of vertices of G_n with types in A . The most important condition is that ν_n should converge to μ in a suitable sense, meaning (roughly speaking) that μ encodes the asymptotic distribution of vertex types.

To be precise, as in [5], a *standard vertex space*, or simply a *vertex space*, is a triple $\mathcal{V} = (\mathcal{S}, \mu, (\mathbf{x}_n)_{n \geq 1})$ where \mathcal{S} is a separable metric space, μ is a Borel probability measure on \mathcal{S} , each \mathbf{x}_n is a random or deterministic sequence of n points of \mathcal{S} , and $\nu_n \xrightarrow{P} \mu$ in the sense of weak convergence of measures, where ν_n is defined by (2.2).

In the most common special cases, the convergence condition says something very simple. Firstly, if \mathcal{S} is a finite set, then $\nu_n(\{x\})$ is simply the proportion of vertices having type x , and $\nu_n \xrightarrow{P} \mu$ if and only if for each type $x \in \mathcal{S}$ this proportion converges to $\mu(\{x\})$ in probability. If $\mathcal{S} = [0, 1]$ with μ Lebesgue measure, the condition is that for each interval $[a, b]$ there are asymptotically $\mu([a, b])n = (b - a)n$ vertices with types in $[a, b]$. This last condition holds, for example, when the types are deterministic with $x_i = i/n$.

Turning to the kernel, as in [5], we say that the kernel κ is *graphical on \mathcal{V}* if κ is integrable and a.e. (almost everywhere) continuous, and $\mathbb{E} e(G^\mathcal{V}(n, \kappa))/n \rightarrow \frac{1}{2} \int_{\mathcal{S}^2} \kappa$. The last condition says roughly that the expected number of edges is as expected.

As in [5], we say that a sequence (κ_n) of kernels is *graphical on \mathcal{V} with limit κ* if (i) each κ_n is integrable and a.e. continuous, (ii) for a.e. $(y, z) \in \mathcal{S}^2$, $y_n \rightarrow y$ and $z_n \rightarrow z$ imply $\kappa_n(y_n, z_n) \rightarrow \kappa(y, z)$, and (iii) $\mathbb{E} e(G^\mathcal{V}(n, \kappa_n))/n \rightarrow \frac{1}{2} \int_{\mathcal{S}^2} \kappa$. As noted in [5] this includes the case when all $\kappa_n = \kappa$ for some graphical kernel κ .

Remark 2.1. In [5], *generalized vertex spaces* are considered, where $\mu(\mathcal{S})$ need not be equal to 1, and the number of vertices of $G^\mathcal{V}(n, \kappa)$ is not exactly n , but rather asymptotically $n\mu(\mathcal{S})$. As shown in [5, Section 8.1], the apparent extra generality is essentially

illusory: one can reduce to the vertex-space case by renormalizing appropriately. For this reason we shall not consider generalized vertex spaces further, except to note that we have chosen the normalization in our definitions of χ and $\widehat{\chi}$ in (2.10) and (2.11) below so that our main results apply verbatim to generalized vertex spaces: no extra factors of $\mu(\mathcal{S})$ appear; see Remark 2.3.

Note that we shall consider kernels κ on spaces (\mathcal{S}, μ) with $\mu(\mathcal{S}) \neq 1$; this will be useful in the supercritical case, see e.g. Section 2.3.

2.1.2 The i.i.d. case.

As noted in [7], in the special case where the types of the vertices are independent and identically distributed (i.i.d.), many of the results in [5] hold without the need for some of the technical assumptions described above. We say that $\mathcal{V} = (\mathcal{S}, \mu, (\mathbf{x}_n)_{n \geq 1})$ is an *i.i.d. vertex space* if (\mathcal{S}, μ) is an arbitrary probability space, and each sequence \mathbf{x}_n is a sequence of n i.i.d. random elements x_1, \dots, x_n of \mathcal{S} , each with distribution μ .

To unify the notation, we write

$$\kappa_n \xrightarrow{\mathcal{V}} \kappa \tag{2.3}$$

if either (i) \mathcal{V} is a standard vertex space and (κ_n) is graphical on \mathcal{V} with limit κ , or (ii) \mathcal{V} is an i.i.d. vertex space, κ is an arbitrary integrable kernel on \mathcal{V} , and $\kappa_n = \kappa$ for all n . (Many results for the i.i.d. case extend to suitable sequences of kernels, for example assuming that $\|\kappa_n - \kappa\|_1 \rightarrow 0$, as then the general setting in the next subsection applies.)

2.1.3 Cut-convergent sequences

To define the final variant we shall consider, we briefly recall some definitions. (A variant of) the Frieze–Kannan [21] *cut norm* of an integrable function $W : \mathcal{S}^2 \rightarrow \mathbb{R}$ is simply

$$\|W\|_{\square} := \sup_{\|f\|_{\infty}, \|g\|_{\infty} \leq 1} \int_{\mathcal{S}^2} f(x)W(x, y)g(y) \, d\mu(x) \, d\mu(y).$$

Given an integrable kernel κ and a measure-preserving bijection $\tau : \mathcal{S} \rightarrow \mathcal{S}$, let $\kappa^{(\tau)}$ be the corresponding *rearrangement* of κ , defined by

$$\kappa^{(\tau)}(x, y) = \kappa(\tau(x), \tau(y)).$$

We write $\kappa \sim \kappa'$ if κ' is a rearrangement of κ . Given two kernels κ, κ' on $[0, 1]$, the *cut metric* of Borgs, Chayes, Lovász, Sós and Vesztergombi [11] may be defined by

$$\delta_{\square}(\kappa, \kappa') := \inf_{\kappa'' \sim \kappa'} \|\kappa - \kappa''\|_{\square}. \tag{2.4}$$

There is also an alternative definition via couplings, which extends to kernels defined on two different probability spaces; see [11; 9].

Suppose that $A_n = (a_{ij})$ is an n -by- n symmetric matrix with non-negative entries; from now on any matrix denoted A_n is assumed to be of this form. Then there is a

random graph $G_n = G(A_n)$ naturally associated to A_n : the vertex set is $\{1, 2, \dots, n\}$, edges are present independently, and the probability that ij is an edge is $\min\{a_{ij}/n, 1\}$. Given A_n , there is a corresponding kernel κ_{A_n} on $[0, 1]$ with Lebesgue measure: divide $[0, 1]^2$ into n^2 squares of side $1/n$ in the obvious way, and take the value of κ_{A_n} on the (i, j) th square to be a_{ij} . If A_n is a matrix and κ is a kernel on $[0, 1]$, then we write $\delta_{\square}(A_n, \kappa)$ for $\delta_{\square}(\kappa_{A_n}, \kappa)$.

If A_n is itself random, then $G(A_n)$ is defined to have the conditional distribution just described, given A_n . Any results stating that if $\delta_{\square}(A_n, \kappa) \rightarrow 0$ then $G(A_n)$ has some property whp apply also if (A_n) is random with $\delta_{\square}(A_n, \kappa) \xrightarrow{\text{P}} 0$. (One way to see this is to note that there is a coupling of the distributions of the A_n in which $\delta_{\square}(A_n, \kappa) \rightarrow 0$ a.s., and we may then condition on (A_n) .)

2.1.4 Relating the models

By the *matrix of edge weights* in $G^{\mathcal{V}}(n, \kappa)$ we mean the matrix A_n whose ij th entry is $np_{ij} = \kappa(x_i, x_j) \wedge n$. Note that in general the vertex types x_i are random, so A_n is a random matrix. Recalling (2.1), we may view $G^{\mathcal{V}}(n, \kappa)$ as $G(A_n)$ for this matrix of edge weights.

The following result was proved in [6, Sections 1.2 and 1.3]. (More precisely, the result in [6] for the i.i.d. case concerns the ‘untruncated’ matrices with entries $\kappa(x_i, x_j)$; the result for A_n follows easily from this.)

Lemma 2.2. *Suppose that $G_n = G^{\mathcal{V}}(n, \kappa_n)$ where $\kappa_n \xrightarrow{\mathcal{V}} \kappa$. Then the matrix A_n of edge weights associated to G_n satisfies $\delta_{\square}(A_n, \kappa) \xrightarrow{\text{P}} 0$. \square*

This lemma shows that results applying to $G(A_n)$ when $\delta_{\square}(A_n, \kappa) \xrightarrow{\text{P}} 0$ transfer to the models $G^{\mathcal{V}}(n, \kappa_n)$ and $G^{\mathcal{V}}(n, \kappa)$ defined in Subsections 2.1.1 and 2.1.2.

2.2 The branching process associated to a kernel

Given an integrable kernel κ on a measure space (\mathcal{S}, μ) and an ‘initial type’ $x \in \mathcal{S}$, let $\mathfrak{X}_{\kappa}(x)$ be the multi-type Galton–Watson branching process defined as follows. We start with a single particle of type x in generation 0. A particle in generation t of type y gives rise to children in generation $t + 1$ whose types form a Poisson process on \mathcal{S} with intensity $\kappa(y, z) d\mu(z)$. The children of different particles are independent (given the types of their parents).

If μ is a probability measure, we also consider the branching process \mathfrak{X}_{κ} defined as above but starting with a single particle whose type has the distribution μ .

Writing $|\mathfrak{X}|$ for the total number of particles in all generations of a branching process \mathfrak{X} , let

$$\rho_k(\kappa; x) := \mathbb{P}(|\mathfrak{X}_{\kappa}(x)| = k), \quad k = 1, 2, \dots, \infty, \quad (2.5)$$

and

$$\rho_k(\kappa) := \int_{\mathcal{S}} \rho_k(\kappa; x) d\mu(x), \quad k = 1, 2, \dots, \infty. \quad (2.6)$$

Thus, when $\mu(\mathcal{S}) = 1$, $\rho_k(\kappa)$ is the probability $\mathbb{P}(|\mathfrak{X}_\kappa| = k)$.

For convenience we assume that

$$\int_{\mathcal{S}} \kappa(x, y) \, d\mu(y) < \infty \tag{2.7}$$

for all $x \in \mathcal{S}$; this implies that all sets of children are finite. This is no real restriction, since our assumption that $\int_{\mathcal{S}^2} \kappa < \infty$ implies that (2.7) holds for a.e. x , and we may impose (2.7) by changing κ on a null set, which will a.s. not affect \mathfrak{X}_κ . (Alternatively, we could work without (2.7), adding the qualifier “for a.e. x ” at some places below.)

Since all generations of $\mathfrak{X}_\kappa(x)$ are finite, it follows that $\rho_\infty(\kappa; x)$, the probability that the branching process is infinite, equals the *survival probability* of $\mathfrak{X}_\kappa(x)$, i.e., the probability that all generations are non-empty. We use the notation $\rho(\kappa; x) := \rho_\infty(\kappa; x)$; for typographical reasons we sometimes also write $\rho_\kappa(x) = \rho(\kappa; x)$. Similarly, we write $\rho(\kappa) := \rho_\infty(\kappa)$; if $\mu(\mathcal{S}) = 1$, this is the survival probability of \mathfrak{X}_κ .

We are interested in the analogue of the mean cluster size for the branching processes. For $\mathfrak{X}_\kappa(x)$, we define

$$\chi(\kappa; x) := \mathbb{E}(|\mathfrak{X}_\kappa(x)|) = \sum_{1 \leq k < \infty} k \rho_k(\kappa; x), \tag{2.8}$$

$$\widehat{\chi}(\kappa; x) := \mathbb{E}(|\mathfrak{X}_\kappa(x)| 1_{\{|\mathfrak{X}_\kappa(x)| < \infty\}}) = \sum_{1 \leq k < \infty} k \rho_k(\kappa; x); \tag{2.9}$$

thus $\chi(\kappa; x) = \widehat{\chi}(\kappa; x) \leq \infty$ if $\rho(\kappa; x) = 0$, and $\widehat{\chi}(\kappa; x) \leq \chi(\kappa; x) = \infty$ if $\rho(\kappa; x) > 0$. Let

$$\chi(\kappa) := \mu(\mathcal{S})^{-1} \int_{\mathcal{S}} \chi(\kappa; x) \, d\mu(x) = \mu(\mathcal{S})^{-1} \sum_{1 \leq k < \infty} k \rho_k(\kappa), \tag{2.10}$$

$$\widehat{\chi}(\kappa) := \mu(\mathcal{S})^{-1} \int_{\mathcal{S}} \widehat{\chi}(\kappa; x) \, d\mu(x) = \mu(\mathcal{S})^{-1} \sum_{1 \leq k < \infty} k \rho_k(\kappa). \tag{2.11}$$

If $\mu(\mathcal{S}) = 1$, we see that

$$\chi(\kappa) = \mathbb{E}(|\mathfrak{X}_\kappa|), \tag{2.12}$$

$$\widehat{\chi}(\kappa) = \mathbb{E}(|\mathfrak{X}_\kappa| 1_{\{|\mathfrak{X}_\kappa| < \infty\}}). \tag{2.13}$$

Note that when $\mu(\mathcal{S}) \neq 1$, the normalizations in (2.6) and (2.10) are different; these normalizations are chosen so that our main results, like those of [5], extend unchanged to generalized vertex spaces; see Remark 2.1.

Remark 2.3. For a generalized vertex space, where $\mu(\mathcal{S})$ may differ from 1, we may renormalize by replacing μ and κ by

$$\mu' := \mu(\mathcal{S})^{-1} \mu \quad \text{and} \quad \kappa' := \mu(\mathcal{S}) \kappa. \tag{2.14}$$

This will not affect $\mathfrak{X}_\kappa(x)$, and thus not $\chi(\kappa; x)$ and $\widehat{\chi}(\kappa; x)$; furthermore, because of our choice of normalization in (2.10) and (2.11), $\chi(\kappa)$ and $\widehat{\chi}(\kappa)$ also remain unchanged. Hence, results for generalized vertex spaces follow from the case when $\mu(\mathcal{S}) = 1$.

2.3 Integral operators

Given a kernel κ on a measure space (\mathcal{S}, μ) , let T_κ be the integral operator on (\mathcal{S}, μ) with kernel κ , defined by

$$(T_\kappa f)(x) := \int_{\mathcal{S}} \kappa(x, y) f(y) \, d\mu(y), \quad (2.15)$$

for any (measurable) function f such that this integral is defined (finite or $+\infty$) for a.e. x . (As usual, we shall assume without comment that all functions considered are measurable.) Note that $T_\kappa f$ is defined for every $f \geq 0$, with $0 \leq T_\kappa f \leq \infty$.

We define

$$\|T_\kappa\| := \sup\{\|T_\kappa f\|_2 : f \geq 0, \|f\|_2 \leq 1\} \leq \infty. \quad (2.16)$$

When finite, $\|T_\kappa\|$ is the norm of T_κ as an operator on $L^2(\mu)$ (see Section 2.7).

One of the results of [5] is that the function $\rho_\kappa : \mathcal{S} \rightarrow [0, 1]$ defined by $\rho_\kappa(x) = \rho(\kappa; x)$ is the unique maximal solution to the non-linear functional equation

$$f = 1 - e^{-T_\kappa f}, \quad f \geq 0. \quad (2.17)$$

Moreover, if $\|T_\kappa\| \leq 1$, then ρ_κ is identically 0 and thus $\rho(\kappa) = 0$, while if $\|T_\kappa\| > 1$, then $\rho_\kappa > 0$ on a set of positive measure and thus $\rho(\kappa) > 0$. The three cases $\|T_\kappa\| < 1$, $\|T_\kappa\| = 1$ and $\|T_\kappa\| > 1$, are called *subcritical*, *critical* and *supercritical*, respectively.

Given a kernel κ on a type space (\mathcal{S}, μ) , let $\hat{\mu}$ be the measure on \mathcal{S} defined by

$$d\hat{\mu}(x) := (1 - \rho(\kappa; x)) \, d\mu(x). \quad (2.18)$$

(This is interesting mainly when κ is supercritical, since otherwise $\hat{\mu} = \mu$.) The *dual kernel* $\hat{\kappa}$ is the kernel on $(\mathcal{S}, \hat{\mu})$ that is equal to κ as a function. We regard $T_{\hat{\kappa}}$ as an operator acting on the corresponding space $L^2(\hat{\mu})$. Then $\|T_{\hat{\kappa}}\| \leq 1$; typically $\|T_{\hat{\kappa}}\| < 1$ when κ is supercritical, but equality is possible, see [5, Theorem 6.7 and Example 12.4].

The definitions above imply the following explicit formula for $T_{\hat{\kappa}}f$:

$$(T_{\hat{\kappa}}f)(x) := \int_{\mathcal{S}} \hat{\kappa}(x, y) f(y) \, d\hat{\mu}(y) = \int_{\mathcal{S}} \kappa(x, y) f(y) (1 - \rho(\kappa; y)) \, d\mu(y), \quad (2.19)$$

so $T_{\hat{\kappa}}f = T_\kappa((1 - \rho_\kappa)f)$. Note also that

$$\hat{\mu}(\mathcal{S}) = \int_{\mathcal{S}} (1 - \rho(\kappa; x)) \, d\mu(x) = \mu(\mathcal{S}) - \rho(\kappa); \quad (2.20)$$

if $\mu(\mathcal{S}) = 1$, this is the extinction probability of \mathfrak{X}_κ .

2.4 Small components

Let $N_k(G)$ denote the number of *vertices* in components of order k in a graph G . Since the number of such *components* is $N_k(G)/k$, we can write the definition (1.2) as

$$\chi(G) = \frac{1}{|G|} \sum_{k=1}^{\infty} \frac{N_k(G)}{k} k^2 = \sum_{k=1}^{\infty} k \frac{N_k(G)}{|G|}. \quad (2.21)$$

Let $N_{\geq k}(G) := \sum_{j \geq k} N_j(G)$ and $\rho_{\geq k}(\kappa) := \sum_{k \leq j \leq \infty} \rho_j(\kappa)$. Collecting together basic results from [5; 7; 6], we have the following lemma.

Lemma 2.4. *Suppose either that $G_n = G^{\mathcal{V}}(n, \kappa_n)$ where $\kappa_n \xrightarrow{\mathcal{V}} \kappa$ in the sense of (2.3), or that $G_n = G(A_n)$ where A_n is a random sequence of matrices with $\delta_{\square}(A_n, \kappa) \xrightarrow{\mathbb{P}} 0$. Then, for every fixed $k \geq 1$, we have*

$$N_{\geq k}(G_n)/n \xrightarrow{\mathbb{P}} \rho_{\geq k}(\kappa) \tag{2.22}$$

and

$$N_k(G_n)/n \xrightarrow{\mathbb{P}} \rho_k(\kappa). \tag{2.23}$$

Proof. By Lemma 2.2 it suffices to consider the second case $G_n = G(A_n)$. This result follows from [6, Lemma 2.11] (the special case when the A_n are deterministic) by [6, Remark 1.5]. \square

The result for the model $G^{\mathcal{V}}(n, \kappa_n)$ was proved in [5, Theorem 9.1] and (for the i.i.d. case) [7, Lemma 21].

2.5 The giant component

As in [5], we say that a kernel κ is *reducible* if there exists $A \subset \mathcal{S}$ with $0 < \mu(A) < \mu(\mathcal{S})$ such that $\kappa(x, y) = 0$ for a.e. (x, y) with $x \in A$ and $y \in \mathcal{S} \setminus A$. Otherwise, κ is *irreducible*. Roughly speaking, κ is reducible if the set of types can be partitioned into two parts so that there will be no edges joining vertices with types in different parts.

Collecting together the results for the various models, we have the following theorem.

Theorem 2.5. *Under the assumptions of Lemma 2.4 we have*

$$|\mathcal{C}_1(G_n)|/n \xrightarrow{\mathbb{P}} \rho(\kappa) \tag{2.24}$$

and

$$|\mathcal{C}_2(G_n)|/n \xrightarrow{\mathbb{P}} 0. \tag{2.25}$$

Proof. This follows from [6, Theorem 1.1] and Lemma 2.2; see also [5, Theorems 3.1 and 3.6]. \square

2.6 Monotonicity

We note a simple monotonicity property of χ ; there is no corresponding result for $\widehat{\chi}$.

Lemma 2.6. *If H is a subgraph of G with the same vertex set, then $\chi(H) \leq \chi(G)$.*

Proof. Immediate from the definition (1.1). \square

2.7 Operators on L^2 spaces

Although our results connect two random combinatorial objects (random graphs and branching processes), in the proofs we shall need some tools from functional analysis. For the reader less familiar with this area, we collect here some of the basic facts we shall use. See e.g. [15] or [17] for proofs and further details.

Let (\mathcal{S}, μ) be a finite measure space. Two (real-valued) functions f and g on \mathcal{S} are *equal a.e.* if $\mu(\{x : f(x) \neq g(x)\}) = 0$. Formally, the elements of $L^2(\mu) = L^2(\mathcal{S}, \mu)$ are equivalence classes of functions under the relation $f \sim g$ if f and g are equal a.e. In practice one thinks of them as functions f on \mathcal{S} , bearing in mind that f is only defined up to equality a.e. Adopting this convention, $L^2(\mu)$ is simply the set of all measurable real-valued functions f on \mathcal{S} such that $\|f\|_2 = (\int_{\mathcal{S}} f(x)^2 d\mu(x))^{1/2}$ is finite.

Two key basic properties of $L^2(\mu)$ are that $\|f\|_2$ is indeed a norm on this space, i.e., $\|\lambda f\|_2 = \lambda \|f\|_2$ for λ constant and $\|f+g\|_2 \leq \|f\|_2 + \|g\|_2$, and that this norm is *complete*: if $f_j \in L^2(\mu)$ and the sequence (f_j) is Cauchy with respect to the norm then there is an $f \in L^2(\mu)$ with $\|f_j - f\|_2 \rightarrow 0$. Moreover, the norm is given by $\|f\|_2 = \langle f, f \rangle_{\mu}^{1/2}$ for the bilinear inner product $\langle f, g \rangle_{\mu} := \int_{\mathcal{S}} f(x)g(x) d\mu(x)$; thus $L^2(\mu)$ is a *Hilbert space*. [15, §I.1]

A (linear) *operator* on $L^2(\mu)$ is simply a linear function $T : L^2(\mu) \rightarrow L^2(\mu)$. (Note that if $f = g$ a.e. then we must have $Tf = Tg$ a.e.) The *operator norm* $\|T\|$ of T is then $\sup\{\|Tf\|_2 : \|f\|_2 = 1\}$. T is a *bounded operator* if $\|T\| < \infty$. The set of all bounded operators on $L^2(\mu)$ is a vector space, the operator norm is a norm on this space, and the space is complete with respect to this norm. (In other words, the set of bounded linear operators is a Banach space.) An additional property of the operator norm is that if S and T are operators on $L^2(\mu)$, then $\|ST\| \leq \|S\|\|T\|$. [15, §II.1 and Exercise III.2.1]

Note that the integral operator T_{κ} defined in Section 2.3 is an operator on $L^2(\mu)$ if and only if $\|Tf\|_2 < \infty$ for all $f \in L^2(\mu)$. In this case, $\|T_{\kappa}\|$ as defined in Section 2.3 is exactly the operator norm of T_{κ} . In particular, $\|T_{\kappa}\| < \infty$ if and only if T_{κ} is a bounded operator on $L^2(\mu)$.

An operator T on $L^2(\mu)$ is *compact* if the closure of $\{Tf : \|f\|_2 \leq 1\}$ is a compact subset of $L^2(\mu)$. An operator T is *finite rank* if its range $\{Tf : f \in L^2(\mu)\}$ has finite dimension; equivalently, if there are some $\psi_i, \varphi_i \in L^2(\mu)$ such that Tf is given by the finite sum

$$Tf = \sum_{i=1}^k \langle f, \psi_i \rangle \varphi_i. \quad (2.26)$$

A key property of compact operators is that they can be approximated by finite rank ones: if T is compact and $\varepsilon > 0$, then there is a finite rank F such that $\|T - F\| \leq \varepsilon$. [15, §II.4]

An important sufficient condition for the integral operator T_{κ} in (2.15) to be a compact operator on $L^2(\mu)$ is that the kernel is square integrable, i.e., $\int_{\mathcal{S} \times \mathcal{S}} \kappa(x, y)^2 d\mu(x) d\mu(y) < \infty$ [15, Proposition II.4.7]. Such integral operators are called *Hilbert–Schmidt*. In particular, if κ is bounded, then T_{κ} is compact (since we only consider finite measures μ).

An operator T on $L^2(\mu)$ is *self-adjoint* if $\langle Tf, g \rangle_\mu = \langle f, Tg \rangle_\mu$ for all $f, g \in L^2(\mu)$. The integral operator T_κ is always self-adjoint (if defined on $L^2(\mu)$), since κ is symmetric.

The *spectrum* $\sigma(T)$ of a bounded operator T is defined to be the complement of the set $\{\lambda : \lambda I - T \text{ is one-to-one and has a bounded inverse}\}$. (In general, one considers $\lambda \in \mathbb{C}$, and the L^2 space of complex-valued square integrable functions. For self-adjoint operators it suffices to consider real λ and real functions.)

One version of the *spectral theorem* is the following. (Theorem 2.7 holds without changes for any Hilbert space, but we state it for $L^2(\mu)$.) For simplicity, we consider only the case of a compact self-adjoint operator, which is all that we shall need in this paper; see e.g. [15, Theorem II.5.1] (with a slightly different but equivalent formulation). For the case of more general operators, see [15, §IX.2].

Theorem 2.7. *Let T be a compact self-adjoint operator on $L^2(\mu)$.*

- (i) *The spectrum $\sigma(T)$ is a non-empty compact subset of \mathbb{R} , and is either finite or consists of a sequence converging to 0.*
- (ii) *If $\lambda \neq 0$, then $\lambda \in \sigma(T)$ if and only if λ is an eigenvalue of T , and in this case the eigenspace $E_\lambda := \{f : Tf = \lambda f\}$ has finite dimension.*
- (iii) *The space $L^2(\mu)$ can be decomposed as a direct orthogonal sum $\bigoplus_{\lambda \in \sigma(T)} E_\lambda$. If $f \in L^2(\mu)$, then f thus has a decomposition $f = \sum_{\lambda \in \sigma(T)} P_\lambda f$, where P_λ is the orthogonal projection onto E_λ , and $Tf = \sum_{\lambda \in \sigma(T)} \lambda P_\lambda f$.*
- (iv) *The norm $\|T\|$ equals the spectral radius $\sup\{|\lambda| : \lambda \in \sigma(T)\}$. □*

3 Branching processes

We start by showing that the mean cluster sizes $\chi(\kappa)$ and $\widehat{\chi}(\kappa)$ can be expressed in terms of the operators T_κ and $T_{\widehat{\kappa}}$. This is doubtless well known, but we have not found the result in the literature in the generality that we need here. We write 1 for the constant function 1 on \mathcal{S} .

Lemma 3.1. *For any integrable kernel κ on a type space (\mathcal{S}, μ) we have*

$$\chi(\kappa; x) = \sum_{j=0}^{\infty} T_\kappa^j 1(x), \tag{3.1}$$

$$\chi(\kappa) = \mu(\mathcal{S})^{-1} \sum_{j=0}^{\infty} \int_{\mathcal{S}} T_\kappa^j 1(x) d\mu(x) = \mu(\mathcal{S})^{-1} \sum_{j=0}^{\infty} \langle T_\kappa^j 1, 1 \rangle_\mu, \tag{3.2}$$

$$\widehat{\chi}(\kappa; x) = (1 - \rho(\kappa; x)) \sum_{j=0}^{\infty} T_{\widehat{\kappa}}^j 1(x), \tag{3.3}$$

$$\widehat{\chi}(\kappa) = \mu(\mathcal{S})^{-1} \sum_{j=0}^{\infty} \int_{\mathcal{S}} T_{\widehat{\kappa}}^j 1(x) d\widehat{\mu}(x) = \mu(\mathcal{S})^{-1} \sum_{j=0}^{\infty} \langle T_{\widehat{\kappa}}^j 1, 1 \rangle_{\widehat{\mu}}. \tag{3.4}$$

Proof. Let $f_j(x)$ be the expected size of generation j in $\mathfrak{X}_\kappa(x)$. Then, for every $j \geq 0$, by conditioning on the first generation,

$$f_{j+1}(x) = \int_{\mathcal{S}} f_j(y) \kappa(x, y) d\mu(y) = T_\kappa f_j(x).$$

Thus, by induction, $f_j = T_\kappa^j f_0 = T_\kappa^j 1$, and (3.1) follows by summing. (Note that the sum, like $\chi(\kappa; x)$, need not be finite. However, in this and in all sums in the proof, the summands are non-negative, so the sum is certainly defined, allowing ∞ as a value.) Recalling the definition (2.10), relation (3.2) follows immediately from (3.1).

As noted in [5, page 38], it is easy to see that if we condition $\mathfrak{X}_\kappa(x)$ on extinction, then we obtain another similar branching process $\widehat{\mathfrak{X}}_\kappa(x)$ with μ replaced by $\widehat{\mu}$. Hence, T_κ is replaced by $T_{\widehat{\kappa}}$, so

$$\mathbb{E}(|\mathfrak{X}_\kappa(x)| \mid |\mathfrak{X}_\kappa(x)| < \infty) = \mathbb{E}(|\widehat{\mathfrak{X}}_\kappa(x)|).$$

Since

$$\mathbb{E}(|\mathfrak{X}_\kappa(x)| 1_{\{|\mathfrak{X}_\kappa(x)| < \infty\}}) = (1 - \rho(\kappa; x)) \mathbb{E}(|\mathfrak{X}_\kappa(x)| \mid |\mathfrak{X}_\kappa(x)| < \infty),$$

(3.3) follows from (3.1) for $\widehat{\mathfrak{X}}_\kappa(x)$. Finally, recalling the definition (2.11) and the relation (2.18) between $\widehat{\mu}$ and μ , the relation (3.4) follows from (3.3) by integrating with respect to μ . \square

Often, it is convenient to assume for simplicity that $\mu(\mathcal{S}) = 1$.

Lemma 3.2. *Let κ be an integrable kernel on a type space (\mathcal{S}, μ) with $\mu(\mathcal{S}) = 1$. Then*

$$\widehat{\chi}(\kappa) = \sum_{j=0}^{\infty} \langle T_{\widehat{\kappa}}^j 1, 1 \rangle_{\widehat{\mu}} = \widehat{\mu}(\mathcal{S}) \chi(\widehat{\kappa}) = (1 - \rho(\kappa)) \chi(\widehat{\kappa}).$$

Proof. Use (3.4) for κ and μ and (3.2) for $\widehat{\kappa}$ and $\widehat{\mu}$, together with (2.20). \square

Theorem 3.3. *Let κ be an integrable kernel on a type space (\mathcal{S}, μ) with $\mu(\mathcal{S}) = 1$.*

- (i) *If κ is subcritical, i.e., $\|T_\kappa\| < 1$, then $\chi(\kappa; x) = (I - T_\kappa)^{-1} 1$ a.e., and $\chi(\kappa) = \langle (I - T_\kappa)^{-1} 1, 1 \rangle_\mu < \infty$.*
- (ii) *Suppose that κ is supercritical, i.e., $\|T_\kappa\| > 1$, and also that $\|T_{\widehat{\kappa}}\| < 1$. Then $\widehat{\chi}(\kappa; x) = (1 - \rho_\kappa)(I - T_{\widehat{\kappa}})^{-1} 1$ a.e., and $\widehat{\chi}(\kappa) = \langle (I - T_{\widehat{\kappa}})^{-1} 1, 1 \rangle_{\widehat{\mu}} < \infty$.*

The conditions of (ii) hold whenever $\|T_\kappa\| > 1$, κ is irreducible, and $\int_{\mathcal{S}^2} \kappa^2 < \infty$.

Proof. The space of operators on $L^2(\mu)$ is complete with respect to the operator norm. Since $\|T_\kappa^j\| \leq \|T_\kappa\|^j$, in case (i) the sum $\sum_{j=0}^{\infty} T_\kappa^j$ converges and (multiplying out) is the inverse of $I - T_\kappa$. Hence (i) follows from the first two parts of Lemma 3.1. Part (ii) follows similarly from the last two parts of Lemma 3.1, since now $\sum_{j=0}^{\infty} T_{\widehat{\kappa}}^j = (I - T_{\widehat{\kappa}})^{-1}$ converges as an operator on $L^2(\widehat{\mu})$. For the final statement we use [5, Theorem 6.7], which yields $\|T_{\widehat{\kappa}}\| < 1$. \square

In fact, for the last part one can replace the assumption that $\int_{\mathcal{S}^2} \kappa^2 < \infty$ by the weaker assumption that T_κ is compact; this is all that is used in the proof of [5, Theorem 6.7].

In the critical case, when $\|T_\kappa\| = 1$, we have $\chi(\kappa) = \widehat{\chi}(\kappa)$. We typically expect the common value to be infinite, but there are exceptions; see Section 6.3.

Theorem 3.4. (i) *If κ is critical and T_κ is a compact operator on $L^2(\mu)$, then $\chi(\kappa) = \infty$. In particular, this applies if $\int_{\mathcal{S}^2} \kappa(x, y)^2 d\mu(x) d\mu(y) < \infty$.*

(ii) *If κ is supercritical, then $\chi(\kappa) = \infty$.*

Proof. (i): If $\int_{\mathcal{S}^2} \kappa^2 < \infty$, then T_κ is a Hilbert–Schmidt operator and thus compact, see Section 2.7.

T_κ is always self-adjoint (when it is bounded), so if T_κ is compact and critical, then by the spectral theorem (Theorem 2.7), it has an eigenfunction ψ with eigenvalue $\pm\|T_\kappa\| = \pm 1$. Moreover, since $\kappa \geq 0$, there is at least one such eigenfunction $\psi_1 \geq 0$ with eigenvalue $+1$ (with $\|\psi_1\|_2 = 1$, say); see Lemma 5.15 in [5] and its proof, where only compactness is used. There may also be eigenfunctions with eigenvalue -1 , so we consider the positive compact operator T_κ^2 and let ψ_1, \dots, ψ_m be an orthonormal basis of the eigenspace E_1 for the eigenvalue 1 of T_κ^2 . The eigenvalues of T_κ^2 are the squares of the eigenvalues of T_κ , so $\sigma(T_\kappa^2) \subset [0, \infty)$; moreover, $\|T_\kappa^2\| = \|T_\kappa\|^2 = 1$, and thus $\sigma(T_\kappa^2) \subset [0, 1]$. By Theorem 2.7(iii), there is an orthogonal decomposition $L^2(\mu) = \bigoplus_{\lambda \in \sigma(T_\kappa^2)} E_\lambda$, so the orthogonal complement of E_1 is simply $E_1^\perp = \bigoplus_{\lambda \neq 1} E_\lambda$. Furthermore, this subspace is invariant for T_κ^2 , and if R is the norm of T_κ^2 restricted to E_1^\perp , then $R = \max\{\lambda : \lambda \in \sigma(T_\kappa^2) \setminus \{1\}\} < 1$ (using Theorem 2.7(i)). Hence, for any $j \geq 0$, T_κ^{2j} acts on E_1^\perp with norm at most R^j , so

$$\langle T_\kappa^{2j} 1, 1 \rangle = \sum_{i=1}^m \langle 1, \psi_i \rangle^2 + O(R^j) \rightarrow \sum_{i=1}^m \langle 1, \psi_i \rangle^2.$$

Since the terms in the sum are non-negative and $\langle 1, \psi_1 \rangle = \int \psi_1 d\mu > 0$, the limit is strictly positive and thus $\sum_{j=0}^\infty \langle T_\kappa^{2j} 1, 1 \rangle$ cannot converge. Since the terms in this sum are non-negative, (3.2) yields $\chi(\kappa) = \mu(\mathcal{S})^{-1} \sum_{j=0}^\infty \langle T_\kappa^{2j} 1, 1 \rangle = \infty$.

(ii): By [5, Theorem 6.1] we have $\mathbb{P}(|\mathfrak{X}_\kappa| = \infty) = \rho(\kappa) > 0$, so $\chi(\kappa) = \infty$. \square

In the subcritical case, we can find $\chi(\kappa)$ by finding $(I - T_\kappa)^{-1}1$, i.e., by solving the integral equation $f = T_\kappa f + 1$. Actually, we can do this for any κ , and can use this as a test of whether $\chi(\kappa) < \infty$.

Theorem 3.5. *Let κ be a kernel on a type space (\mathcal{S}, μ) . Then the following are equivalent:*

(i) $\chi(\kappa) < \infty$.

(ii) *There exists a function $f \geq 0$ in $L^1(\mu)$ such that (a.e.)*

$$f = T_\kappa f + 1. \tag{3.5}$$

(iii) *There exists a function $f \geq 0$ in $L^1(\mu)$ such that (a.e.)*

$$f \geq T_\kappa f + 1. \tag{3.6}$$

When the above conditions hold, there is a smallest non-negative solution f to (3.5), that is also a smallest non-negative solution to (3.6); this minimal solution f equals $\chi(\kappa; x)$, and thus $\chi(\kappa) = \mu(\mathcal{S})^{-1} \int_{\mathcal{S}} f \, d\mu$.

Proof. Recalling (3.1), let $g(x) := \chi(\kappa; x) = \sum_{j=0}^{\infty} T_{\kappa}^j 1(x)$; this is a function $g : \mathcal{S} \rightarrow [0, \infty]$ with $T_{\kappa} g = \sum_{j=1}^{\infty} T_{\kappa}^j 1 = g - 1$, so g satisfies both (3.5) and (3.6). Furthermore, $\int_{\mathcal{S}} g \, d\mu = \mu(\mathcal{S})\chi(\kappa)$ by (3.2). Hence, if (i) holds, then $g \in L^1(\mu)$; consequently, g satisfies (ii) and (iii). (Note that then g is finite a.e.)

Conversely, suppose that $f \geq 0$ solves (3.5) or (3.6). Recalling that $\kappa \geq 0$ by definition, so T_{κ} is monotone, induction on i gives

$$f \geq \sum_{j=0}^{i-1} T_{\kappa}^j 1 + T_{\kappa}^i f$$

for every $i \geq 1$. Thus $f \geq \sum_{j=0}^{i-1} T_{\kappa}^j 1$, and letting $i \rightarrow \infty$ yields $f \geq g$. Hence, if (ii) or (iii) holds, then $g \in L^1(\mu)$, and (i) holds. Furthermore, in this case, $f \geq g$, which shows that g is the smallest solution in both (ii) and (iii), completing the proof. \square

Note that in the subcritical case, (3.5) always has a solution in $L^2(\mu)$: the proof of Theorem 3.3 shows that $(I - T_{\kappa})^{-1}$ exists as an operator on $L^2(\mu)$, so $g = (I - T_{\kappa})^{-1} 1 \in L^2(\mu)$. In Section 6.3, we give an example where κ is critical and (3.5) has a solution that belongs to $L^1(\mu)$, but not to $L^2(\mu)$. (We do not know whether there can be a non-negative solution in $L^2(\mu)$ with κ critical.) Moreover, in this example, in both the subcritical and critical cases, there is more than one non-negative solution in $L^1(\mu)$. However, we can show that there is never more than one non-negative solution in $L^2(\mu)$.

Corollary 3.6. *Suppose that there exists a function $f \geq 0$ in $L^2(\mu)$ such that (3.5) holds. Then f is the unique non-negative solution to (3.5) in $L^2(\mu)$, $\chi(\kappa; x) = f(x)$ and $\chi(\kappa) = \mu(\mathcal{S})^{-1} \int_{\mathcal{S}} f \, d\mu$.*

Proof. Let g be the smallest non-negative solution, guaranteed to exist by Theorem 3.5, and let $h = f - g \geq 0$. Since $0 \leq h \leq f$, $h \in L^2(\mu)$. Then $T_{\kappa} h = T_{\kappa} f - T_{\kappa} g = (f - 1) - (g - 1) = h$, and

$$\langle f, h \rangle = \langle T_{\kappa} f + 1, h \rangle = \langle T_{\kappa} f, h \rangle + \langle 1, h \rangle = \langle f, T_{\kappa} h \rangle + \langle 1, h \rangle = \langle f, h \rangle + \langle 1, h \rangle.$$

Hence $0 = \langle 1, h \rangle = \int h \, d\mu$, so $h = 0$ a.e., and $f = g$. \square

4 Main results

In this section our overall aim is to show that the susceptibilities of suitable random graphs G_n and branching processes \mathfrak{X}_{κ} are related. Ideally, we should like to show that $\chi(G_n) \xrightarrow{P} \chi(\kappa)$ and $\widehat{\chi}(G_n) \xrightarrow{P} \widehat{\chi}(\kappa)$ for any of the random graph models G_n introduced

in Section 2. Our main results (Theorems 4.7, 4.8 and 4.9) show that this does hold in very great generality, though unfortunately not in the full generality we would like.

This section is organized as follows. Firstly we prove lower bounds on $\chi(G_n)$ and $\widehat{\chi}(G_n)$ that do hold in full generality (see Theorem 4.1). Then, in Section 4.2, we describe a general approach to proving corresponding upper bounds, based on path counting. In Sections 4.3 and 4.4 we prove our upper bound results.

4.1 The lower bound

We begin with a general asymptotic lower bound for the susceptibility. This bound depends only on convergence of the number of vertices in components of each fixed size, so it applies to any of the models considered in Section 2. More precisely, we state the bound and its consequences in the setting of Subsection 2.1.3; as noted there they then apply to $G^\mathcal{V}(n, \kappa_n)$ under the assumptions in Subsection 2.1.1 or Subsection 2.1.2.

Recall that a matrix denoted A_n is assumed to be symmetric, n -by- n and to have non-negative entries.

Theorem 4.1. *Let κ be a kernel, let (A_n) be a sequence of (random) matrices with $\delta_\square(A_n, \kappa) \xrightarrow{\text{P}} 0$, and set $G_n = G(A_n)$. Alternatively, let $G_n = G^\mathcal{V}(n, \kappa_n)$ where $\kappa_n \xrightarrow{\mathcal{V}} \kappa$ in the sense of (2.3). Then*

- (i) for every $b < \chi(\kappa)$, whp $\chi(G_n) > b$;
- (ii) for every $b < \widehat{\chi}(\kappa)$, whp $\widehat{\chi}(G_n) > b$;
- (iii) $\liminf \mathbb{E} \chi(G_n) \geq \chi(\kappa)$ and $\liminf \mathbb{E} \widehat{\chi}(G_n) \geq \widehat{\chi}(\kappa)$.

Proof. All we shall use about the graph G_n is that, for each fixed k , we have $N_k(G_n)/n \xrightarrow{\text{P}} \rho_k(\kappa)$; this holds by Lemma 2.4.

(i): Let K be a fixed positive integer. Then, by (2.21), (2.22) and (2.23),

$$\begin{aligned} \chi(G_n) &\geq \sum_{k=1}^{\infty} (k \wedge K) \frac{N_k(G_n)}{n} \\ &= \sum_{k=1}^{K-1} k \frac{N_k(G_n)}{n} + K \frac{N_{\geq K}(G_n)}{n} \\ &\xrightarrow{\text{P}} \sum_{k=1}^{K-1} k \rho_k(\kappa) + K \rho_{\geq K}(\kappa) = \sum_{1 \leq k \leq \infty} (k \wedge K) \rho_k(\kappa). \end{aligned}$$

As $K \rightarrow \infty$, the right-hand side tends to $\chi(\kappa)$ by monotone convergence and (2.10); hence we can choose a finite K such that the right-hand side is greater than b , and (i) follows.

(ii): By (1.3), if \mathcal{C}_1 is the largest component of G_n and $|\mathcal{C}_1| > K$, then

$$\widehat{\chi}(G_n) \geq \sum_{k=1}^K k \frac{N_k(G_n)}{n}.$$

On the other hand, if $|\mathcal{C}_1| \leq K$, then

$$\widehat{\chi}(G_n) = \chi(G_n) - |\mathcal{C}_1|^2/n \geq \chi(G_n) - K^2/n.$$

Hence, in both cases, using (2.23) again,

$$\widehat{\chi}(G_n) \geq \sum_{k=1}^K k \frac{N_k(G_n)}{n} - \frac{K^2}{n} \xrightarrow{\text{P}} \sum_{k=1}^K k \rho_k(\kappa). \quad (4.1)$$

As $K \rightarrow \infty$, the right-hand side tends to $\widehat{\chi}(\kappa)$, and thus we can choose K such that it exceeds b , and (ii) follows.

(iii): An immediate consequence of (i) and (ii). \square

4.2 Upper bounds: general techniques

In this section we shall show that, in the light of Theorem 4.1, to prove that $\chi(G_n) \xrightarrow{\text{P}} \chi(\kappa)$ it suffices to show that $\limsup \mathbb{E} \chi(G_n) \leq \chi(\kappa)$. Furthermore, we show that $\mathbb{E} \chi(G_n)$ can be bounded from above by counting the expected number of paths of certain types. We start with a simple general probability exercise.

Lemma 4.2. *Let (X_n) be a sequence of non-negative random variables and suppose that $a \in [0, \infty]$ is such that*

(i) *for every real $b < a$, whp $X_n \geq b$, and*

(ii) $\limsup \mathbb{E} X_n \leq a$.

Then $X_n \xrightarrow{\text{P}} a$ and $\mathbb{E} X_n \rightarrow a$. Furthermore, if $a < \infty$, then $X_n \xrightarrow{L^1} a$, i.e., $\mathbb{E} |X_n - a| \rightarrow 0$.

Proof. If $a = \infty$, (i) says that $X_n \xrightarrow{\text{P}} \infty$; this implies $\liminf \mathbb{E} X_n \geq b$ for every $b < \infty$, and thus $\mathbb{E} X_n \rightarrow \infty$.

Assume now that $a < \infty$, and let $\varepsilon \geq 0$ and $b < a$. Consider the random variable X'_n taking the value 0 when $X_n < b$, the value b when $b \leq X_n < a + \varepsilon$, and the value $a + \varepsilon$ when $X_n \geq a + \varepsilon$. Since $X_n \geq X'_n$ we have

$$\begin{aligned} \mathbb{E}(X_n - a) &\geq \mathbb{E}(X'_n - a) \\ &= \varepsilon \mathbb{P}(X_n \geq a + \varepsilon) - (a - b) \mathbb{P}(b \leq X_n < a + \varepsilon) - a \mathbb{P}(X_n < b) \\ &\geq \varepsilon \mathbb{P}(X_n \geq a + \varepsilon) - (a - b) - o(1), \end{aligned}$$

using (i) for the last step. Hence

$$\limsup \mathbb{E}(X_n - a) \geq \varepsilon \limsup \mathbb{P}(X_n \geq a + \varepsilon) - (a - b)$$

and thus, since $b < a$ is arbitrary,

$$\limsup \mathbb{E}(X_n - a) \geq \varepsilon \limsup \mathbb{P}(X_n \geq a + \varepsilon).$$

Since $\limsup \mathbb{E}(X_n - a) \leq 0$ by (ii), this yields $\limsup \mathbb{P}(X_n \geq a + \varepsilon) = 0$ for every $\varepsilon > 0$, which together with (i) yields $X_n \xrightarrow{\text{P}} a$.

Moreover, the same argument yields, for every $\varepsilon \geq 0$,

$$\liminf \mathbb{E}(X_n - a) \geq \varepsilon \liminf \mathbb{P}(X_n \geq a + \varepsilon).$$

Taking $\varepsilon = 0$ we obtain $\liminf \mathbb{E} X_n \geq a$, which together with (ii) yields $\mathbb{E} X_n \rightarrow a$. \square

We would like to apply Lemma 4.2 with $X_n = \chi(G_n)$ and $a = \chi(\kappa)$ or $X_n = \widehat{\chi}(G_n)$ and $a = \widehat{\chi}(\kappa)$. Condition (i) is satisfied by Theorem 4.1, so we only have to verify the upper bound (ii) for the expected susceptibility. For convenience, we state this explicitly.

Lemma 4.3. *Let κ and G_n be as in Theorem 4.1.*

(i) *If $\limsup \mathbb{E} \chi(G_n) \leq \chi(\kappa)$, then $\chi(G_n) \xrightarrow{\text{P}} \chi(\kappa)$ and $\mathbb{E} \chi(G_n) \rightarrow \chi(\kappa)$.*

(ii) *If $\limsup \mathbb{E} \widehat{\chi}(G_n) \leq \widehat{\chi}(\kappa)$, then $\widehat{\chi}(G_n) \xrightarrow{\text{P}} \widehat{\chi}(\kappa)$ and $\mathbb{E} \widehat{\chi}(G_n) \rightarrow \widehat{\chi}(\kappa)$.*

Proof. By Theorem 4.1 and Lemma 4.2 as discussed above. \square

Sometimes we can control the expectation only after conditioning on some (very likely) event. This still gives convergence in probability.

Lemma 4.4. *Let κ and G_n be as in Theorem 4.1, and let \mathcal{E}_n be an event in the probability space on which G_n is defined such that \mathcal{E}_n holds whp.*

(i) *If $\limsup \mathbb{E}(\chi(G_n)1_{\mathcal{E}_n}) \leq \chi(\kappa)$, then $\chi(G_n) \xrightarrow{\text{P}} \chi(\kappa)$.*

(ii) *If $\limsup \mathbb{E}(\widehat{\chi}(G_n)1_{\mathcal{E}_n}) \leq \widehat{\chi}(\kappa)$, then $\widehat{\chi}(G_n) \xrightarrow{\text{P}} \widehat{\chi}(\kappa)$.*

Proof. After conditioning on \mathcal{E}_n , we still have $N_k(G_n)/n \xrightarrow{\text{P}} \rho_k(\kappa)$ for each fixed k , which is all that was needed in the proof of Theorem 4.1. Letting $\varphi = \chi$ or $\widehat{\chi}$, since $\mathbb{E}(\varphi(G_n) | \mathcal{E}_n) \sim \mathbb{E}(\varphi(G_n)1_{\mathcal{E}_n})$, under the relevant assumption Lemma 4.2 tells us that the distribution of $\varphi(G_n)$ conditioned on \mathcal{E}_n converges in probability to $\varphi(\kappa)$. But then the unconditional distribution converges in probability. \square

For future reference, we note a consequence of Lemma 4.3.

Theorem 4.5. *Let κ and G_n be as in Theorem 4.1.*

(i) *If $\chi(\kappa) = \infty$, then $\chi(G_n) \xrightarrow{\text{P}} \infty$ and $\mathbb{E} \chi(G_n) \rightarrow \infty$.*

(ii) *If $\widehat{\chi}(\kappa) = \infty$, then $\widehat{\chi}(G_n) \xrightarrow{\text{P}} \infty$ and $\mathbb{E} \widehat{\chi}(G_n) \rightarrow \infty$.*

Furthermore, the conclusion of (i) holds if κ is critical and T_κ is compact, or if κ is supercritical.

Proof. For (i) and (ii) we simply apply Lemma 4.3. For (i) the only condition to be verified is that $\limsup \mathbb{E} \chi(G_n) \leq \chi(\kappa)$, but this holds trivially since we now assume that $\chi(\kappa) = \infty$. The argument for (ii) is similar. For the final statement, Theorem 3.4 states that under the given conditions, $\chi(\kappa) = \infty$, so part (i) applies. \square

One way to obtain the upper bound on the expectation of the susceptibility required to apply Lemma 4.3 is by counting paths. Let $P_\ell = P_\ell(G)$ denote the number of sequences $v_0 v_1 \dots v_\ell$ of distinct vertices of G with $v_{i-1} v_i$ an edge of G for $i = 1, 2, \dots, \ell$. In the rest of the paper we call such a sequence a *path of length ℓ* . Note that for convenience we count directed paths, so in the usual terminology P_ℓ would be twice the number of paths of length ℓ in G when $\ell \geq 1$. P_0 is simply the number of vertices of G .

Lemma 4.6. *Let G be a graph with n vertices. Then $\chi(G) \leq \sum_{\ell=0}^{\infty} P_\ell(G)/n$.*

Proof. For each ordered pair (v, v') of vertices of G with v and v' in the same component, there is at least one path (of length ≥ 0) starting at v and ending at v' . Thus, counting all such pairs, $\sum_i |\mathcal{C}_i|^2 \leq \sum_{\ell=0}^{\infty} P_\ell$. \square

So far our arguments relied only on convergence of the number of vertices in components of a fixed size k , and so apply in very great generality. Unfortunately, bounding $\chi(G)$ from above, via Lemma 4.6 or otherwise, involves proving bounds for all k simultaneously. These bounds do not hold in general; we study two special cases where they do in the next two subsections.

4.3 Bounded kernels on general vertex spaces

In this section we consider $G_n = G^\mathcal{V}(n, \kappa_n)$ where (κ_n) is any *uniformly bounded* graphical sequence of kernels on a vertex space \mathcal{V} with limit κ . In fact, we shall consider the more general situation where $G_n = G(A_n)$ for some sequence (A_n) of uniformly bounded (random) matrices with $\delta_\square(A_n, \kappa) \xrightarrow{P} 0$. By Lemma 2.2, the graphs $G^\mathcal{V}(n, \kappa_n)$ are of this form. Note that this is the setting in which the component sizes were studied by Bollobás, Borgs, Chayes and Riordan [3].

Theorem 4.7. *Let κ be a kernel and (A_n) a uniformly bounded sequence of matrices with $\delta_\square(A_n, \kappa) \xrightarrow{P} 0$, and set $G_n = G(A_n)$. Alternatively, let $G_n = G^\mathcal{V}(n, \kappa_n)$ where $\kappa_n \xrightarrow{\mathcal{V}} \kappa$ and the κ_n are uniformly bounded.*

(i) *We have $\chi(G_n) \xrightarrow{P} \chi(\kappa)$.*

(ii) *If κ is irreducible, then $\widehat{\chi}(G_n) \xrightarrow{P} \widehat{\chi}(\kappa)$.*

The boundedness assumption is essential unless further conditions are imposed; see Example 6.9. The extra assumption in (ii) is needed to rule out the possibility that there are two or more giant components, living in different parts of the type space.

Proof. By Lemma 2.2 we may assume the setting where $G = G(A_n)$.

Let M be a constant such that all entries of all A_n are bounded by M . Coupling appropriately, we may and shall assume that $\delta_{\square}(A_n, \kappa) \rightarrow 0$. Then it is easily seen that $\kappa \leq M$ holds pointwise (ignoring a null set, if necessary).

For (i), suppose first that $\|T_{\kappa}\| \geq 1$. Then, since κ is bounded, by Theorem 3.4 we have $\chi(\kappa) = \infty$, and by Theorem 4.5 we have $\chi(G_n) \xrightarrow{P} \infty$, as required.

Suppose then that $\|T_{\kappa}\| < 1$. Let $\kappa_n = \kappa_{A_n}$ denote the piecewise constant kernel corresponding to A_n . Then, letting $\mathbf{1}$ denote the vector $(1, \dots, 1)$, and writing $A_n = (a_{ij}^{(n)})$, for $n \geq M$ we have

$$\begin{aligned} \mathbb{E} P_{\ell}(G_n) &\leq \mathbb{E} \sum_{j_0, \dots, j_{\ell}=1}^n \prod_{i=1}^{\ell} \frac{a_{j_{i-1}, j_i}^{(n)}}{n} \\ &= n \mathbb{E} \int_{\mathcal{S}^{\ell+1}} \prod_{i=1}^{\ell} \kappa_n(x_{i-1}, x_i) d\mu(x_0) \cdots d\mu(x_{\ell}) \\ &= n \langle T_{\kappa_n}^{\ell} \mathbf{1}, \mathbf{1} \rangle_{\mu}. \end{aligned} \tag{4.2}$$

Recall that κ_n and κ are uniformly bounded, and $\delta_{\square}(\kappa_n, \kappa) \rightarrow 0$. As noted in [3], or by the Riesz–Thorin interpolation theorem [17, Theorem VI.10.11] (for operators $L^{\infty} \rightarrow L^1$ and $L^1 \rightarrow L^{\infty}$), it is easy to check that this implies $\|T_{\kappa_n}\| \rightarrow \|T_{\kappa}\|$. (In fact, the normalized spectra converge; see [12].) Since $\|T_{\kappa}\| < 1$, it follows that for some $\delta > 0$ we have $\|T_{\kappa_n}\| < 1 - \delta$ for n large enough, so $\sum_{\ell} \langle T_{\kappa_n}^{\ell} \mathbf{1}, \mathbf{1} \rangle_{\mu} \leq \sum_{\ell} \|T_{\kappa_n}\|^{\ell}$ converges geometrically.

For a fixed ℓ , and kernels κ, κ' bounded by M , it is easy to check that $|\langle T_{\kappa'}^{\ell} \mathbf{1}, \mathbf{1} \rangle_{\mu} - \langle T_{\kappa}^{\ell} \mathbf{1}, \mathbf{1} \rangle_{\mu}| \leq \ell M^{\ell-1} \|\kappa' - \kappa\|_{\square}$ (see, for example, [6, Lemma 2.7]). Since $\langle T_{\kappa'}^{\ell} \mathbf{1}, \mathbf{1} \rangle_{\mu}$ is preserved by rearrangement, we may replace $\|\kappa' - \kappa\|_{\square}$ by $\delta_{\square}(\kappa', \kappa)$ in this bound. Hence, for each ℓ , we have $\langle T_{\kappa_n}^{\ell} \mathbf{1}, \mathbf{1} \rangle_{\mu} \rightarrow \langle T_{\kappa}^{\ell} \mathbf{1}, \mathbf{1} \rangle_{\mu}$. Combined with the geometric decay established above, it follows that

$$\sum_{\ell=0}^{\infty} \langle T_{\kappa_n}^{\ell} \mathbf{1}, \mathbf{1} \rangle_{\mu} \rightarrow \sum_{\ell} \langle T_{\kappa}^{\ell} \mathbf{1}, \mathbf{1} \rangle_{\mu} = \chi(\kappa).$$

By Lemma 4.6 and (4.2) we thus have

$$\limsup \mathbb{E} \chi(G_n) \leq \limsup \frac{1}{n} \sum_{\ell=0}^{\infty} \mathbb{E} P_{\ell}(G_n) \leq \limsup \sum_{\ell=0}^{\infty} \langle T_{\kappa_n}^{\ell} \mathbf{1}, \mathbf{1} \rangle_{\mu} = \chi(\kappa),$$

which with Lemma 4.3(i) gives $\chi(G_n) \xrightarrow{P} \chi(\kappa)$ as required.

We now turn to $\widehat{\chi}$, i.e., to the proof of (ii). If $\|T_{\kappa}\| \leq 1$, then $\rho(\kappa) = 0$ and $\widehat{\chi}(\kappa) = \chi(\kappa)$. On the other hand, $\widehat{\chi}(G_n) < \chi(G_n)$, so the bound above gives $\limsup \mathbb{E} \widehat{\chi}(G_n) \leq \chi(\kappa) = \widehat{\chi}(\kappa)$, and Lemma 4.3(ii) gives the result.

Now suppose that $\|T_{\kappa}\| > 1$. Let \widetilde{G}_n be the graph obtained from G_n by deleting all vertices in the largest component \mathcal{C}_1 , and let \tilde{n} be the number of vertices of \widetilde{G}_n . By the duality result of [25] (see also [5, Theorem 12.1] for the case $G_n = G^{\vee}(n, \kappa_n)$), there is a random sequence (B_n) of matrices (of random size $\tilde{n} \times \tilde{n}$) with $\delta_{\square}(B_n, \widetilde{\kappa}) \xrightarrow{P} 0$, such

that \tilde{G}_n may be coupled to agree whp with $G(B_n)$; here $\tilde{\kappa} := \hat{\kappa}'$ is $\hat{\kappa}$ renormalized as in (2.14). (Recall that $\hat{\kappa}$ is regarded as a kernel on $(\mathcal{S}, \hat{\mu})$, where $\hat{\mu}$ defined by (2.18) is not a probability measure.) By Remark 2.3, $\chi(\tilde{\kappa}) = \chi(\hat{\kappa})$.

Note that

$$\frac{|\tilde{G}_n|}{n} = \frac{n - |\mathcal{C}_1|}{n} \xrightarrow{\text{p}} 1 - \rho(\kappa) \quad (4.3)$$

by (2.24). After conditioning on the number of vertices of \tilde{G}_n we can apply part (i) to conclude that $\chi(G(B_n)) \xrightarrow{\text{p}} \chi(\tilde{\kappa}) = \chi(\hat{\kappa})$, and thus, since $\chi(\tilde{G}_n) = \chi(G(B_n))$ whp,

$$\chi(\tilde{G}_n) \xrightarrow{\text{p}} \chi(\hat{\kappa}). \quad (4.4)$$

Finally, if $\{\mathcal{C}_i\}_{i \geq 1}$ are the components of G_n , then $\{\mathcal{C}_i\}_{i \geq 2}$ are the components of \tilde{G}_n , and thus by (1.3), (1.2), (4.3), (4.4) and Lemma 3.2

$$\hat{\chi}(G_n) = \frac{\sum_{j \geq 2} |\mathcal{C}_j|^2}{n} = \frac{|\tilde{G}_n| \chi(\tilde{G}_n)}{n} \xrightarrow{\text{p}} (1 - \rho(\kappa)) \chi(\hat{\kappa}) = \hat{\chi}(\kappa). \quad \square$$

4.4 The i.i.d. case

In this section we consider the case $G_n = G^\mathcal{V}(n, \kappa)$, where \mathcal{V} is an i.i.d. vertex space and κ is an arbitrary integrable kernel on the associated probability space (\mathcal{S}, μ) . We prove two results, one for $\chi(G_n)$, and one for $\hat{\chi}(G_n)$.

Theorem 4.8. *Let κ be an integrable kernel on an i.i.d. vertex space \mathcal{V} . Then we have $\chi(G^\mathcal{V}(n, \kappa)) \xrightarrow{\text{p}} \chi(\kappa)$ and $\mathbb{E} \chi(G^\mathcal{V}(n, \kappa)) \rightarrow \chi(\kappa)$.*

Proof. Similarly to the estimate in the proof of Theorem 4.7, for any ℓ , the expected number $\mathbb{E} P_\ell$ of paths of length ℓ is

$$\begin{aligned} n \cdots (n - \ell) \int_{\mathcal{S}^{\ell+1}} \prod_{i=1}^{\ell} \min\left(\frac{\kappa(x_{i-1}, x_i)}{n}, 1\right) d\mu(x_0) \cdots d\mu(x_\ell) \\ \leq n \int_{\mathcal{S}^{\ell+1}} \prod_{i=1}^{\ell} \kappa(x_{i-1}, x_i) d\mu(x_0) \cdots d\mu(x_\ell) = n \langle T_\kappa^\ell \mathbf{1}, \mathbf{1} \rangle_\mu. \end{aligned}$$

Summing over all $\ell \geq 0$, we see by (3.2) that the expected total number of paths is at most $n\chi(\kappa)$. Hence, by Lemma 4.6,

$$\mathbb{E} \chi(G^\mathcal{V}(n, \kappa)) \leq \mathbb{E} \sum_{\ell=0}^{\infty} P_\ell / n \leq \chi(\kappa). \quad (4.5)$$

The result follows by Lemma 4.3. □

Our next aim is to prove an analogous result for $\hat{\chi}$. Unfortunately, in the proof we need an extra assumption. We shall assume that T_κ is compact, though any condition guaranteeing (4.28) below will do. We do not know whether the result holds without such an assumption.

Theorem 4.9. *Let κ be an irreducible, integrable kernel on an i.i.d. vertex space \mathcal{V} with $\|T_\kappa\| > 1$, and let $G_n = G^\mathcal{V}(n, \kappa)$. If T_κ is compact, then $\widehat{\chi}(G_n) \xrightarrow{P} \widehat{\chi}(\kappa)$.*

The proof of Theorem 4.9 will be rather involved; this is perhaps surprising given that one expects the i.i.d. case to be easy to handle for many questions. The rest of this section is devoted to the proof; we shall need various preparatory lemmas, many of which hold under more general conditions than Theorem 4.9 itself.

For the rest of the section \mathcal{V} is assumed to be an i.i.d. vertex space, and $G_n = G^\mathcal{V}(n, \kappa)$ for some kernel κ . The main idea of the proof is to count the expected number of paths P of a given length ℓ such that P is not joined to a large component of $G_n - P$. Recall that G_n is a random graph with vertex set $[n] = \{1, 2, \dots, n\}$. Since the vertex types are i.i.d., the distribution of G_n is unchanged if we permute the vertices in any fixed way. Hence it suffices to estimate the probability that the ‘first’ $\ell + 1$ vertices, i.e., the vertices $\{1, 2, \dots, \ell + 1\}$, form a path P not joined to the giant component. The obvious strategy is to ‘reveal’ the types of the vertices $1, 2, \dots, \ell + 1$, and reveal the entire graph on the vertices $\ell + 2, \dots, n$, including the types of these vertices. Then we try to argue that the latter graph will contain a giant component \mathcal{C} , and if ℓ is fairly large, it is very likely that P will be joined to this giant component. The problem is that the probability that P is joined to \mathcal{C} depends on the types of the vertices in P and in \mathcal{C} , so we need (fairly strong, as it turns out) bounds on the probability that the giant component \mathcal{C} contains ‘too few’ vertices with types in some set.

As usual, we approach results about the giant component by first considering small components. For $A \subset \mathcal{S}$ let $N_k(A)$ denote the number of vertices i of G_n such that i is in a component of order k and $x_i \in A$.

Lemma 4.10. *Let κ be an integrable kernel on an i.i.d. vertex space $\mathcal{V} = (\mathcal{S}, \mu, (\mathbf{x}_n)_{n \geq 1})$, and let A be a measurable subset of \mathcal{S} . Then*

$$N_k(A)/n \xrightarrow{P} \rho_k(A) := \int_A \rho_k(x) \, d\mu(x). \quad (4.6)$$

Moreover, the convergence is uniform in A : given any $\varepsilon > 0$ there is an n_0 such that

$$\mathbb{P}(|N_k(A)/n - \rho_k(A)| \geq \varepsilon) \leq \varepsilon$$

holds for all $n \geq n_0$ and all measurable A .

Proof. Suppose first that κ is bounded. Then the result follows easily from the local coupling argument in [7, Section 3]; for completeness, we sketch this argument.

Suppose that $\kappa(x, y) \leq M$ for all x and y . To avoid having to write $\min\{\cdot, 1\}$, we consider only $n \geq M$ in what follows.

We may construct G_n in three stages: (i) let G_n^+ be the standard Erdős–Rényi random graph $G(n, M/n)$. (ii) assign each vertex i of G_n^+ a type $x_i \in \mathcal{S}$, with the types i.i.d. with distribution μ . (iii) delete edges of G_n^+ randomly to obtain G_n , retaining each edge ij with probability $\kappa(x_i, x_j)/M$ independently of the other edges.

Similarly, we may construct \mathfrak{X}_κ , which we regard as an infinite rooted tree, by (i) starting from the single-type Galton–Watson process \mathfrak{X}^+ in which the offspring distribution is Poisson with mean M , (ii) assigning each vertex (individual) a type as in the graph case, (iii) deleting edges of the rooted tree as in the graph case, and (iv) taking for \mathfrak{X}_κ the component of the resulting forest containing the root.

It is well known that for any fixed t , the neighbourhoods of a (fixed or uniformly random) vertex v of $G(n, M/n)$ up to distance t may be coupled to agree whp with the first t generations of \mathfrak{X}^+ . The constructions above show that this coupling extends to G_n and \mathfrak{X}_κ . We shall apply this with $t = k + 1$, noting that whether the component containing v has size k or not can be determined from its $(k + 1)$ -neighbourhood. Let $\eta_{k,n}$ denote the error probability in the coupling when $t = k + 1$, so for fixed k we have $\eta_{k,n} \rightarrow 0$ as $n \rightarrow \infty$.

Let $E_{i,k,A}$ denote the event that vertex i is in a component of size k and $x_i \in A$. Let $E_{k,A}$ denote the event that the branching process \mathfrak{X}_κ has total size k and the type of its root is in A . Then the coupling shows that $\mathbb{P}(E_{i,k,A}) = \mathbb{P}(E_{k,A}) + o(1)$: more precisely, the difference is at most $\eta_{k,n}$.

Conditioning on the type of the root, we see that $\mathbb{P}(E_{k,A})$ is exactly $\rho_k(A)$. On the other hand, since all vertices are equivalent, $\mathbb{E} N_k(A) = \sum_i \mathbb{P}(E_{i,k,A}) = n \mathbb{P}(E_{1,k,A})$, so $|\mathbb{E}(N_k(A)/n) - \rho_k(A)| \leq \eta_{k,n} \rightarrow 0$.

To complete the proof in the bounded case we use a similar coupling argument starting with two vertices v and w to show that $|\mathbb{E}(N_k(A)^2/n^2) - \rho_k(A)^2| \leq \eta'_{k,n}$ for some $\eta'_{k,n}$ that tends to 0 as $n \rightarrow \infty$. Since the coupling error probabilities $\eta_{k,n}$ and $\eta'_{k,n}$ do not depend on A , the final result is uniform in A .

Using the fact that adding or deleting an edge from a graph G changes the set of vertices in components of size k in at most $2k$ places, and arguing as in [5] (see the proof of Lemma 9.9), the result for general κ follows easily. \square

Recall that $\mathcal{C}_1 = \mathcal{C}_1(G_n) \subseteq [n]$ denotes the (vertex set of) the largest component of G_n . As in [5], given G_n , let ν_n^1 denote the empirical distribution of the types of the vertices in $\mathcal{C}_1(G_n)$, so for $A \subset \mathcal{S}$ we have

$$\nu_n^1(A) = n^{-1} |\{i \in \mathcal{C}_1(G_n) : x_i \in A\}|.$$

Lemma 4.11. *Let κ be an irreducible, integrable kernel on an i.i.d. vertex space $\mathcal{V} = (\mathcal{S}, \mu, (\mathbf{x}_n)_{n \geq 1})$, and let A be a measurable subset of \mathcal{S} . Then*

$$\nu_n^1(A) \xrightarrow{\mathbb{P}} \mu_\kappa(A) := \int_A \rho(\kappa; x) d\mu(x). \tag{4.7}$$

More precisely, the convergence is uniform in A : given any $\varepsilon > 0$ there is an n_0 such that for all $n \geq n_0$ and all measurable A we have

$$\mathbb{P}(|\nu_n^1(A) - \mu_\kappa(A)| \geq \varepsilon) \leq \varepsilon.$$

Note that the first statement corresponds to Theorem 9.10 of [5], but, due to the different conditions, is not implied by it.

Proof. It suffices to prove the second statement. Recall that $\rho_{\geq k}(\kappa)$ is the probability that the branching process \mathfrak{X}_κ has total size (number of individuals in all generations together) at least k , and that $\rho(\kappa)$ is the probability that \mathfrak{X}_κ is infinite. Thus $\rho_{\geq k}(\kappa) \searrow \rho(\kappa)$ as $k \rightarrow \infty$. Fix $\varepsilon > 0$ once and for all, and choose k_0 so that $\rho_{\geq k_0}(\kappa) \leq \rho(\kappa) + \varepsilon/6$.

Applying Lemma 4.10 for $k = 1, 2, \dots, k_0$ and summing, we see that $N_{\leq k_0}(A)/n \xrightarrow{P} \rho_{\leq k_0}(A)$, and indeed that

$$\mathbb{P}(|N_{\leq k_0}(A)/n - \rho_{\leq k_0}(A)| \geq \varepsilon/5) \leq \varepsilon/3 \quad (4.8)$$

for all large enough n and all measurable A .

By a *medium component* of G_n we mean any component of size greater than k_0 other than $\mathcal{C}_1(G_n)$. Let $M(A)$ denote the number of vertices with types in A in medium components, and $M(G_n) = M(\mathcal{S})$ the total number of vertices in medium components. Since $N_k(G_n)/n \xrightarrow{P} \rho_k(\kappa)$ for each k and $|\mathcal{C}_1(G_n)|/n \xrightarrow{P} \rho(\kappa)$, we have $M(G_n)/n \xrightarrow{P} \rho_{\geq k_0+1}(\kappa) - \rho(\kappa) \leq \varepsilon/6$. Hence, whp

$$\sup_A M(A) = M(G_n) \leq \varepsilon n/5. \quad (4.9)$$

Let $N(A)$ denote the total number of vertices of G_n with types in A . Then $N(A)$ has a binomial distribution with parameters n and $\mu(A)$, so for n large enough we have

$$\mathbb{P}(|N(A)/n - \mu(A)| \geq \varepsilon/5) \leq \varepsilon/3 \quad (4.10)$$

for all A . Finally, let $C_1(A) = n\nu_n^1(A)$ denote the number of vertices in $\mathcal{C}_1(G_n)$ with types in A . Then

$$C_1(A) = N(A) - N_{\leq k_0}(A) - M(A) + O(1), \quad (4.11)$$

with the final $O(1)$ correction term accounting for the possibility that $|\mathcal{C}_1(G_n)| \leq k_0$, so the ‘giant’ component is ‘small’.

Combining equations (4.8)–(4.11), we see that

$$\mathbb{P}(|C_1(A)/n - (\mu(A) - \rho_{\leq k_0}(A))| \geq 4\varepsilon/5) \leq \varepsilon$$

for all large enough n and all A . But

$$\mu(A) - \rho_{\leq k_0}(A) = \mu_\kappa(A) + \sum_{k=k_0+1}^{\infty} \rho_k(A).$$

The sum above is at least 0 and, by choice of k_0 , at most $\varepsilon/6$, so $\mu(A) - \rho_{\leq k_0}(A)$ is within $\varepsilon/6$ of $\mu_\kappa(A)$ and the result follows. \square

In [6, Theorem 1.4], it was shown (in a slightly different setting) that stability of the giant component under deletion of vertices implies that the distribution of the size of the giant component has an exponential tail. Parts of this argument adapt easily to the present setting.

First, as a consequence of Lemma 2.2 (or Lemma 1.7 of [6]), all results of [6] asserting that a certain conclusion holds whp when $\delta_\square(A_n, \kappa) \rightarrow 0$ apply to the random graphs $G^\mathcal{V}(n, \kappa)$. In particular, Theorem 1.3 of [6] implies the following result.

Theorem 4.12. *Let κ be an irreducible, integrable kernel on an i.i.d. vertex space \mathcal{V} , and let $G_n = G^\mathcal{V}(n, \kappa)$. For every $\varepsilon > 0$ there is a $\delta > 0$ such that whp we have*

$$\rho(\kappa) - \varepsilon \leq |\mathcal{C}_1(G'_n)|/n \leq \rho(\kappa) + \varepsilon$$

for every graph G'_n that may be obtained from G_n by deleting at most δn vertices and their incident edges, and then adding or deleting at most δn edges. \square

In the proof of our next result, we shall use the following inequality due to McDiarmid [28].

Theorem 4.13. *Let f be a real-valued function of n variables that is c -Lipschitz, i.e., changing one input changes the output by at most c . Let X_1, \dots, X_n be independent random variables. Then for any $t \geq 0$ we have*

$$\mathbb{P}(|f(X_1, \dots, X_n) - \mathbb{E} f(X_1, \dots, X_n)| \geq t) \leq e^{-2t^2/(c^2n)}. \quad \square$$

Using Theorems 4.12 and 4.13, it is easy to get an exponential lower tail bound on the number of vertices of $\mathcal{C}_1(G_n)$ with types in a given set $A \subset \mathcal{S}$. Unfortunately, there is a minor complication, due to the possible (but very unlikely) non-uniqueness of the giant component.

Given a graph G whose vertices have types in \mathcal{S} , let $\tilde{C}_1(A; G)$ denote the maximum over all components \mathcal{C} of G of the number of vertices of \mathcal{C} with types in A :

$$\tilde{C}_1(A; G) = \max\{|\{i \in \mathcal{C} : x_i \in A\}| : \mathcal{C} \text{ a component of } G\}. \quad (4.12)$$

Let $\tilde{C}_1(A) = \tilde{C}_1(A; G_n)$, so $\tilde{C}_1(A)$ is within $|\mathcal{C}_2(G_n)|$ of $C_1(A) = n\nu_n^1(A)$.

Lemma 4.14. *Let κ be an irreducible, integrable kernel on an i.i.d. vertex space $\mathcal{V} = (\mathcal{S}, \mu, (\mathbf{x}_n)_{n \geq 1})$ with $\|T_\kappa\| > 1$, and let $\varepsilon > 0$. Then there is a $c = c(\kappa, \varepsilon) > 0$ such that for all large enough n , for every subset A of \mathcal{S} we have*

$$\mathbb{P}(\tilde{C}_1(A; G_n) \leq (\mu_\kappa(A) - \varepsilon)n) \leq e^{-cn}. \quad (4.13)$$

Proof. Fix A . Given a graph G on $[n]$ where each vertex has a type in \mathcal{S} , let $D(G) = D_A(G)$ be the minimum number of vertices that must be deleted from G so that in the resulting graph G' we have

$$\tilde{C}_1(A; G') \leq (\mu_\kappa(A) - \varepsilon)n, \quad (4.14)$$

so our aim is to bound $\mathbb{P}(D(G_n) = 0)$. By Lemma 4.11, whp $\mathcal{C}_1(G_n)$ has at least $(\mu_\kappa(A) - \varepsilon/2)n$ vertices with types in A . Also, by Theorem 4.12, there is some $\delta > 0$ such that whp deleting at most δn vertices of G_n removes fewer than $\varepsilon n/2$ vertices from the (whp unique) giant component. It follows that $\mathbb{E} D(G_n) \geq \delta n/2$ for n large; moreover, this bound is uniform in A .

Since the condition (4.14) is preserved by deleting vertices, if G'' is obtained from G by adding and deleting edges all of which are incident with one vertex i , and also

perhaps changing the type of i , then $|D(G) - D(G'')| \leq 1$. We may construct G_n by taking independent variables x_1, \dots, x_n and $\{y_{ij} : 1 \leq i < j \leq n\}$ all of which are uniform on $[0, 1]$, and joining i to j if and only if $y_{ij} \leq \kappa(x_i, x_j)/n$. Modifying the variables in $S_j = \{x_j\} \cup \{y_{ij} : i < j\}$ only affects edges incident with vertex j . Considering the values of all variables in S_j as a single random variable X_j , we see that $D(G_n)$ is a 1-Lipschitz function of n independent variables, so by Theorem 4.13 we have

$$\mathbb{P}(D(G_n) = 0) \leq e^{-2(\mathbb{E} D(G_n))^2/n} \leq e^{-\delta^2 n/2},$$

completing the proof. \square

It would be nice to have an exponential bound on the upper tail of the number of vertices in ‘large’ components. Unfortunately, the argument in [6] does not seem to go through. Indeed, the corresponding result is false in this setting without an additional assumption: it is easy to find a κ for which there is a small, but only polynomially small, chance that the degree of some vertex v is of order n . In fact, one can even arrange that $\mathbb{P}(|\mathcal{C}_1(G_n)| = n)$ is only polynomially small in n .

The next lemma is the combinatorial heart of the proof of Theorem 4.9. We would like to bound the expectation of $\widehat{\chi}(G_n)$, so that we can apply Lemma 4.3. Unfortunately, we cannot do this directly; instead we bound the contribution from components with size up to some small constant times n . Formally, given a graph G with n vertices and a $\delta > 0$, let

$$\widehat{\chi}_\delta(G) := \frac{1}{n} \sum_{v \in V(G) : |\mathcal{C}(v)| \leq \delta n} |\mathcal{C}(v)| = \frac{1}{n} \sum_{i : |\mathcal{C}_i| \leq \delta n} |\mathcal{C}_i|^2. \quad (4.15)$$

Note that if $|\mathcal{C}_2| \leq \delta n < |\mathcal{C}_1|$, then $\widehat{\chi}_\delta(G) = \widehat{\chi}(G)$.

Given a kernel κ and an $M > 0$, we write κ^M for the pointwise minimum of κ and M .

Lemma 4.15. *Let κ be an irreducible, integrable kernel on an i.i.d. vertex space \mathcal{V} with $\|T_\kappa\| > 1$, and let $\varepsilon > 0$ and $M > 0$. Then there is a $\delta = \delta(\varepsilon, M, \kappa) > 0$ such that*

$$\mathbb{E} \widehat{\chi}_\delta(G^\mathcal{V}(n, \kappa)) \leq \sum_{\ell=0}^{\infty} \langle T_{\bar{\kappa}}^\ell \mathbf{1}, \mathbf{1} \rangle_{\bar{\mu}} + o(1), \quad (4.16)$$

where $\bar{\mu}$ is the measure on \mathcal{S} defined by $d\bar{\mu}(x) := f(x) d\mu(x)$ with

$$f(x) := (1 - \rho((1 - \varepsilon)\kappa^M; x) + 5\varepsilon) \wedge 1, \quad (4.17)$$

and $T_{\bar{\kappa}}$ is the integral operator on $(\mathcal{S}, \bar{\mu})$ with kernel κ .

Proof. A simple coupling argument shows that if $\kappa_1 \leq \kappa_2$ pointwise then $\rho(\kappa_1; x) \leq \rho(\kappa_2; x)$. Hence increasing M and/or decreasing ε can only increase $\rho((1 - \varepsilon)\kappa^M; x)$, and thus can only decrease $f(x)$, and hence decrease the right-hand side of (4.16). By monotone convergence $\|T_{\kappa^M}\| \rightarrow \|T_\kappa\|$ as $M \rightarrow \infty$. Hence, by first increasing M if necessary and then decreasing ε if necessary, we may assume that $(1 - \varepsilon)\kappa^M$ is supercritical, and that $\rho((1 - \varepsilon)\kappa^M) > 2\varepsilon$. We also assume that $M > 1$ and $e^{4\varepsilon} < 1 + 5\varepsilon$.

Let $0 < \delta < \varepsilon/M$ be a small constant to be chosen later, depending only on κ , ε and M , and let $Z = n\widehat{\chi}_\delta(G_n)$ denote the number of ordered pairs (v, w) of vertices of $G_n = G^\mathcal{V}(n, \kappa)$ such that v and w are in a common component of size at most δn . Also, let Z_ℓ denote the number of such pairs joined by a path of length ℓ . Since $Z \leq \sum_{\ell=0}^{\delta n-1} Z_\ell$, it suffices to show that for $0 \leq \ell < \delta n$ we have

$$\mathbb{E} Z_\ell/n \leq \langle T_{\tilde{\kappa}}^\ell \mathbf{1}, \mathbf{1} \rangle_{\tilde{\mu}} + o(1/n), \quad (4.18)$$

with the error bound uniform in ℓ .

We may bound Z_ℓ by the number of paths of length ℓ in G_n lying in components with at most δn vertices. Thus, from the symmetry of the model $G^\mathcal{V}(n, \kappa)$ under vertex permutations, $\mathbb{E} Z_\ell$ is at most $n^{\ell+1}$ times the probability that $12 \cdots (\ell+1)$ forms such a path. Let $V_0 = \{1, 2, \dots, \ell+1\}$ consist of the first $\ell+1$ vertices of G_n , and let V' consist of the last $(1 - \varepsilon/M)n$ vertices. Note that $\ell \leq \delta n < \varepsilon n/M$, so V_0 and V' are disjoint. We shall define a certain subgraph G' of $G_n[V']$, in a way that is independent of $G_n[V_0]$, and consider paths in V_0 not joined to G' .

Let $\kappa_1 = \frac{1-\varepsilon}{1-\varepsilon/M} \kappa^M$; this slightly unnatural choice simplifies some formulae below. Since $\kappa_1 \leq \kappa^M \leq \kappa$, conditional on the vertex types, the edge probabilities in $G_n^1 = G^\mathcal{V}(n, \kappa_1)$ are at most those in $G_n = G^\mathcal{V}(n, \kappa)$. Hence we may construct these graphs simultaneously in such a way that $G_n^1 \subseteq G_n$, by first constructing G_n and then deleting edges with appropriate probabilities. Coupling G_n^1 and G_n in this way, let $G' = G_n^1[V']$ be the subgraph of G_n^1 induced by V' . We shall use the following properties of G' : firstly, G' is a subgraph of G_n . Secondly, G' has the distribution

$$G' \sim G^\mathcal{V}(n', (1 - \varepsilon)\kappa^M), \quad (4.19)$$

where $n' = (1 - \varepsilon/M)n$. (We ignore the irrelevant rounding to integers.)

Let $\mathcal{A} = \mathcal{A}_\ell$ be the event that $12 \cdots (\ell+1)$ forms a path in G_n , and let $\mathcal{B} = \mathcal{B}_\ell$ be the event that some vertex in $[\ell+1]$ is joined by an edge of G_n^M to some component of G' of order at least δn . Then

$$\mathbb{E} Z_\ell \leq n^{\ell+1} \mathbb{P}(\mathcal{A} \cap \mathcal{B}^c).$$

Unfortunately, we cannot quite prove the estimate we need for the right hand side above, so we need to use a slightly less natural but stronger upper bound on $\mathbb{E} Z_\ell$.

Let Z'_ℓ be the number of ordered pairs (v, w) of vertices in $V_0 = [\ell+1]$ such that v and w are joined in G_n by a path of length ℓ lying in V_0 (and thus visiting all vertices of V_0). Since all possible sets of $\ell+1$ vertices contribute equally to the expectation $\mathbb{E} Z_\ell$, we have

$$\mathbb{E} Z_\ell \leq \binom{n}{\ell+1} \mathbb{E}(Z'_\ell \mathbf{1}_{\mathcal{B}^c}). \quad (4.20)$$

Roughly speaking, our plan is to show that with very high probability $\mathcal{C}_1(G')$ will contain almost the ‘right’ number of vertices of each type, so that, given the type y of one of the first $\ell+1$ vertices, its probability of having one or more neighbours in $\mathcal{C}_1(G')$ is almost what it ‘should be’. Unfortunately, there will certainly be sets A of types that are

‘under-represented’ in $\mathcal{C}_1(G')$, for example, the set of all types not appearing at all! So our plan is to ‘reveal’ the types $x_1, \dots, x_{\ell+1}$ of the first $\ell + 1$ vertices, and then to define certain subsets A of \mathcal{S} from these (the sets $A_{y,i}$ below). We show that these sets of types are unlikely to be under-represented, and deduce that each of the first $\ell + 1$ vertices has a not-too-small probability of being joined to $\mathcal{C}_1(G')$. There is one further complication: for technical reasons, when making this argument precise we must replace $\mathcal{C}_1(G')$ by the union of all components of G' of order at least δn .

Recall that $(1-\varepsilon)\kappa^M$ is supercritical and that $\rho((1-\varepsilon)\kappa^M) > 2\varepsilon$. Let $\mu' = \mu_{(1-\varepsilon)\kappa^M}$, so, recalling (4.7), $d\mu'(x) = \rho((1-\varepsilon)\kappa^M; x) d\mu(x)$. Recalling the definition (4.12), applying Lemma 4.14 to G' we find that there is some $c > 0$ such that for any measurable $A \subset \mathcal{S}$ we have

$$\mathbb{P}(\tilde{\mathcal{C}}_1(A; G') \leq (\mu'(A) - \varepsilon/M)n') \leq e^{-cn}.$$

Since $n' = (1 - \varepsilon/M)n$, it follows that

$$\mathbb{P}(\tilde{\mathcal{C}}_1(A; G') \leq (\mu'(A) - 2\varepsilon/M)n) \leq e^{-cn}. \quad (4.21)$$

Let

$$\delta_0 := \min\{\varepsilon/M, 1/10\} > 0,$$

and fix $0 < \delta < \delta_0$ chosen small enough that

$$(e/\delta)^\delta < e^{c/2}. \quad (4.22)$$

Let L denote the union of all components of G' of order at least δn , and let $L(A)$ be the number of vertices in L with types in A . If $\mu'(A) \geq 3\varepsilon/M$ and $\tilde{\mathcal{C}}_1(A; G') \geq (\mu'(A) - 2\varepsilon/M)n$, then since $\tilde{\mathcal{C}}_1(A; G') \geq \varepsilon n/M \geq \delta n$, we have $L(A) \geq \tilde{\mathcal{C}}_1(A; G')$. Using (4.21), it follows that

$$\mathbb{P}(L(A) \leq (\mu'(A) - 3\varepsilon/M)n) \leq e^{-cn} \quad (4.23)$$

for any A ; the condition is vacuous if $\mu'(A) < 3\varepsilon/M$.

Given $y \in \mathcal{S}$ and an integer $i \geq 0$, let $A_{y,i} = \{x \in \mathcal{S} : \kappa^M(x, y) \geq \varepsilon i\}$. Let \mathcal{E}_y be the event that $L(A_{y,i})/n \geq \mu'(A_{y,i}) - 3\varepsilon/M$ holds for all i with $1 \leq i \leq M/\varepsilon$. Applying (4.23) $M/\varepsilon = O(1)$ times, we see that, for each $y \in \mathcal{S}$,

$$\mathbb{P}(\mathcal{E}_y^c) \leq (M/\varepsilon)e^{-cn} = O(e^{-cn}). \quad (4.24)$$

From the definition of $A_{y,i}$ we see that

$$\sum_{v \in L} \kappa^M(x_v, y) \geq \sum_{i=1}^{M/\varepsilon} L(A_{y,i})\varepsilon.$$

It follows that if \mathcal{E}_y holds, then

$$\sum_{v \in L} \kappa^M(x_v, y) \geq \sum_{i=1}^{M/\varepsilon} \varepsilon(\mu'(A_{y,i}) - 3\varepsilon/M)n \geq n \sum_{i=1}^{M/\varepsilon} \varepsilon\mu'(A_{y,i}) - 3\varepsilon n.$$

Since κ^M is by definition bounded by M , the set $A_{y,i}$ is empty for $i > M/\varepsilon$, so we have

$$\begin{aligned} \sum_{i=1}^{M/\varepsilon} \varepsilon \mu'(A_{y,i}) &= \sum_{i=1}^{\infty} \varepsilon \mu' \{x : \kappa^M(x, y) \geq \varepsilon i\} = \int_{\mathcal{S}} \varepsilon \lfloor \kappa^M(x, y) / \varepsilon \rfloor d\mu'(x) \\ &\geq \int_{\mathcal{S}} \kappa^M(x, y) d\mu'(x) - \varepsilon = \int_{\mathcal{S}} \kappa^M(x, y) \rho((1 - \varepsilon)\kappa^M; x) d\mu(x) - \varepsilon. \end{aligned}$$

Putting these bounds together, writing κ' for $(1 - \varepsilon)\kappa^M$, when \mathcal{E}_y holds we have

$$\begin{aligned} \sum_{v \in L} \kappa^M(x_v, y) / n &\geq \int_{\mathcal{S}} \kappa^M(x, y) \rho(\kappa'; x) d\mu(x) - 4\varepsilon \\ &= (T_{\kappa^M \rho_{\kappa'}})(y) - 4\varepsilon \geq (T_{\kappa' \rho_{\kappa'}})(y) - 4\varepsilon. \end{aligned}$$

Recalling that κ' is supercritical, from (2.17) we have $T_{\kappa' \rho_{\kappa'}} = -\log(1 - \rho_{\kappa'})$, so when \mathcal{E}_y holds we have

$$\sum_{v \in L} \kappa^M(x_v, y) / n \geq -\log(1 - \rho(\kappa'; y)) - 4\varepsilon,$$

and hence, using $1 - z \leq e^{-z}$ and $e^{4\varepsilon} \leq 1 + 5\varepsilon$,

$$\prod_{v \in L} (1 - \kappa^M(x_v, y) / n) \leq (1 - \rho(\kappa'; y)) e^{4\varepsilon} \leq 1 - \rho(\kappa'; y) + 5\varepsilon.$$

Since κ^M is bounded by M , and the product is always at most 1, it follows that if \mathcal{E}_y holds and $n \geq M$, then, recalling the definition (4.17) of f ,

$$\prod_{v \in L} (1 - (\kappa^M(x_v, y) / n \wedge 1)) \leq f(y). \quad (4.25)$$

Let $\mathcal{E} = \mathcal{E}_{x_1} \cap \dots \cap \mathcal{E}_{x_{\ell+1}}$. Recall that by assumption the vertex types x_1, \dots, x_n are independent, and that G' involves only vertices in V' , which is disjoint from $V_0 = \{1, 2, \dots, \ell + 1\}$. Hence G' is independent of $x_1, \dots, x_{\ell+1}$. Given these types, from (4.24) we have $\mathbb{P}(\mathcal{E}) = 1 - O(\ell e^{-cn}) = 1 - O(ne^{-cn})$, with the implicit constant independent of the types. Hence, we have $\mathbb{P}(\mathcal{E}) = 1 - O(ne^{-cn})$ unconditionally. For $\ell \leq \delta n$, noting that $Z'_\ell \leq (\ell + 1)^2$ by definition, we have

$$\binom{n}{\ell + 1} \mathbb{E}(Z'_\ell 1_{\mathcal{E}^c}) \leq \binom{n}{\ell + 1} (\ell + 1)^2 \mathbb{P}(\mathcal{E}^c) \leq (e/\delta)^{\delta n} n^2 \mathbb{P}(\mathcal{E}^c) = o(1), \quad (4.26)$$

using (4.22) in the last step.

Estimating Z'_ℓ by the number of paths of length ℓ lying in V_0 ,

$$\binom{n}{\ell + 1} \mathbb{E}(Z'_\ell 1_{\mathcal{B}^c \cap \mathcal{E}}) \leq \binom{n}{\ell + 1} (\ell + 1)! \mathbb{P}(\mathcal{A} \cap \mathcal{B}^c \cap \mathcal{E}) \leq n^{\ell+1} \mathbb{P}(\mathcal{A} \cap \mathcal{B}^c \cap \mathcal{E}). \quad (4.27)$$

To estimate the final probability let us condition on G' and also on the vertex types $x_1, \dots, x_{\ell+1}$, assuming as we may that \mathcal{E} holds. Note that we have not yet ‘looked at’ edges within V_0 , or edges from V_0 to V' . The conditional probability of \mathcal{A} is then exactly

$$\prod_{i=1}^{\ell} (\kappa(x_i, x_{i+1})/n \wedge 1) \leq n^{-\ell} \prod_{i=1}^{\ell} \kappa(x_i, x_{i+1}).$$

For $i \leq \ell + 1$ and $v \in L$, the edge iv is present with probability $\min\{\kappa(x_i, x_v)/n, 1\} \geq \min\{\kappa^M(x_i, x_v)/n, 1\}$, since $\kappa^M \leq \kappa$. Since \mathcal{E}_{x_i} holds we have from (4.25) that the probability that i has no neighbours in L is thus at most $f(x_i)$. These events are (conditionally) independent for different i , so

$$\mathbb{P}(\mathcal{A} \cap \mathcal{B}^c \cap \mathcal{E} \mid x_1, \dots, x_{\ell+1}, G') \leq n^{-\ell} \prod_{i=1}^{\ell} \kappa(x_i, x_{i+1}) \prod_{i=1}^{\ell+1} f(x_i).$$

Since the right-hand side is independent of G' , the same bound holds conditional only on $x_1, \dots, x_{\ell+1}$. Integrating out it follows that

$$\begin{aligned} n^{\ell+1} \mathbb{P}(\mathcal{A} \cap \mathcal{B}^c \cap \mathcal{E}) &\leq n \int_{\mathcal{S}^{\ell+1}} \prod_{i=1}^{\ell} \kappa(x_i, x_{i+1}) \prod_{i=1}^{\ell+1} f(x_i) \, d\mu(x_1) \cdots d\mu(x_{\ell+1}) \\ &= n \langle T_{\tilde{\kappa}}^{\ell} 1, 1 \rangle_{\tilde{\mu}}. \end{aligned}$$

From (4.27) it follows that $\binom{n}{\ell+1} \mathbb{E}(Z_{\ell}^{\mathcal{A} \cap \mathcal{B}^c \cap \mathcal{E}}) \leq n \langle T_{\tilde{\kappa}}^{\ell} 1, 1 \rangle_{\tilde{\mu}}$. Combined with (4.26) and (4.20) this establishes (4.18); as noted earlier, the result follows. \square

Taking, say, $M = 1/\varepsilon$ and defining $f_{\varepsilon}(x)$ by (4.17), as $\varepsilon \rightarrow 0$ we have $(1 - \varepsilon)\kappa^M \nearrow \kappa$ pointwise. Hence, by [5, Theorem 6.4], $\rho((1 - \varepsilon)\kappa^M; x) \nearrow \rho(\kappa; x)$ pointwise. Thus $f_{\varepsilon}(x) \searrow 1 - \rho(\kappa; x)$ pointwise. If we know that $\langle T_{\tilde{\kappa}}^{\ell} 1, 1 \rangle_{\tilde{\mu}} < \infty$ for some $\varepsilon > 0$, then by dominated convergence it follows that $\langle T_{\tilde{\kappa}}^{\ell} 1, 1 \rangle_{\tilde{\mu}} \searrow \langle T_{\tilde{\kappa}}^{\ell} 1, 1 \rangle_{\tilde{\mu}}$ as $\varepsilon \rightarrow 0$. Furthermore, if we have

$$\sum_{\ell=0}^{\infty} \langle T_{\tilde{\kappa}}^{\ell} 1, 1 \rangle_{\tilde{\mu}} < \infty \tag{4.28}$$

for some $\varepsilon > 0$, then by dominated convergence, as $\varepsilon \rightarrow 0$ we have

$$\sum_{\ell=0}^{\infty} \langle T_{\tilde{\kappa}}^{\ell} 1, 1 \rangle_{\tilde{\mu}} \searrow \sum_{\ell=0}^{\infty} \langle T_{\tilde{\kappa}}^{\ell} 1, 1 \rangle_{\tilde{\mu}} = \widehat{\chi}(\kappa). \tag{4.29}$$

Unfortunately we need some assumption on κ to establish (4.28).

Proof of Theorem 4.9. Suppose for the moment that (4.28) holds for some $\varepsilon > 0$, where $\tilde{\mu}$ is defined using $f_{\varepsilon}(x)$, which is in turn given by (4.17) with $M = 1/\varepsilon$, say.

From (4.29) it follows that, given any $\eta > 0$, choosing ε small enough and $M = 1/\varepsilon$ we have $\sum_{\ell=0}^{\infty} \langle T_{\tilde{\kappa}}^{\ell} 1, 1 \rangle_{\tilde{\mu}} \leq \widehat{\chi}(\kappa) + \eta$. Lemma 4.15 then gives $\mathbb{E} \widehat{\chi}_{\alpha(\eta)}(G_n) \leq \widehat{\chi}(\kappa) + 2\eta$ if n is large enough, for some $\alpha(\eta) > 0$.

Suppose that $\delta(n)$ tends to zero. Since $\widehat{\chi}_\delta$ is an increasing function of δ , for any $\eta > 0$ we see that if n is large enough, then $\mathbb{E} \widehat{\chi}_{\delta(n)}(G_n) \leq \mathbb{E} \widehat{\chi}_{\alpha(\eta)}(G_n) \leq \widehat{\chi}(\kappa) + 2\eta$. Hence,

$$\limsup \mathbb{E} \widehat{\chi}_{\delta(n)}(G_n) \leq \widehat{\chi}(\kappa). \quad (4.30)$$

Since κ is supercritical we have $\rho(\kappa) > 0$, and by (2.24) we have $|\mathcal{C}_1(G_n)| \geq \rho(\kappa)n/2$ whp. For any fixed $\delta > 0$, by (2.25) we have $|\mathcal{C}_2(G_n)| < \delta n$ whp; this also holds if $\delta = \delta(n)$ tends to zero sufficiently slowly. Given a function $\delta(n)$, let \mathcal{E}_n be the event that $|\mathcal{C}_2(G_n)| \leq n\delta(n) < |\mathcal{C}_1(G_n)|$. Then, provided $\delta(n)$ tends to zero slowly enough, \mathcal{E}_n holds whp. When \mathcal{E}_n holds we have $\widehat{\chi}_{\delta(n)}(G_n) = \widehat{\chi}(G_n)$, so $\mathbb{E}(\widehat{\chi}(G_n)1_{\mathcal{E}_n}) \leq \mathbb{E} \widehat{\chi}_{\delta(n)}(G_n)$, and (4.30) gives $\limsup \mathbb{E}(\widehat{\chi}(G_n)1_{\mathcal{E}_n}) \leq \widehat{\chi}(\kappa)$. By Lemma 4.4 this implies that $\widehat{\chi}(G_n) \xrightarrow{P} \widehat{\chi}(\kappa)$, which is our goal. It thus suffices to establish that (4.28) holds for some $\varepsilon > 0$.

Recall that $f_\varepsilon(x) \leq 1$ and $f_\varepsilon \searrow f_0 = 1 - \rho_\kappa$ as $\varepsilon \rightarrow 0$. Recall also that $T_{\widehat{\kappa}}$ is defined as the integral operator on $L^2(\widehat{\mu})$ with

$$(T_{\widehat{\kappa}}g)(x) = \int \kappa(x, y)g(y) d\widehat{\mu}(y) = \int \kappa(x, y)f_\varepsilon(y)g(y) d\mu(y).$$

The map $g(x) \mapsto g(x)f_\varepsilon(x)^{1/2}$ is an isometry of $L^2(\widehat{\mu})$ onto $L^2(\mu)$, and thus $T_{\widehat{\kappa}}$ is unitarily equivalent to the integral operator T_ε on $L^2(\mu)$ with kernel $f_\varepsilon(x)^{1/2}\kappa(x, y)f_\varepsilon(y)^{1/2}$. In particular, $\|T_{\widehat{\kappa}}\| = \|T_\varepsilon\|$, and for the special case $\varepsilon = 0$, when $T_{\widehat{\kappa}} = T_{\widehat{\kappa}}$, $\|T_{\widehat{\kappa}}\| = \|T_0\|$.

Fix $\eta > 0$. Since $T_{\widehat{\kappa}}$ is compact, there is a finite rank operator F such that $\Delta = T_{\widehat{\kappa}} - F$ satisfies $\|\Delta\| < \eta$. Let F_ε and Δ_ε denote the operators on $L^2(\mu)$ obtained by multiplying the kernels of F and Δ by $f_\varepsilon(x)^{1/2}f_\varepsilon(y)^{1/2}$; thus $T_\varepsilon = F_\varepsilon + \Delta_\varepsilon$. Since $f_\varepsilon \leq 1$ holds pointwise, we have

$$\|\Delta_\varepsilon\| \leq \|\Delta\| < \eta.$$

For any $g \in L^2(\mu)$ we have $f_\varepsilon^{1/2}g - f_0^{1/2}g \rightarrow 0$ pointwise and hence (by dominated convergence) in $L^2(\mu)$. Applying this with g each of the functions ψ_i, φ_i in the expression (2.26) for F , it follows that $\|F_\varepsilon - F_0\| \rightarrow 0$, and hence that

$$\limsup_{\varepsilon \rightarrow 0} \|T_\varepsilon - T_0\| \leq \limsup_{\varepsilon \rightarrow 0} \|F_\varepsilon - F_0\| + 2\eta = 2\eta.$$

Since $\eta > 0$ was arbitrary, this gives $\|T_\varepsilon - T_0\| \rightarrow 0$, and in particular $\|T_{\widehat{\kappa}}\| = \|T_\varepsilon\| \rightarrow \|T_0\| = \|T_{\widehat{\kappa}}\|$. Furthermore, $\|T_{\widehat{\kappa}}\| < 1$ by [5, Theorem 6.7] (and its proof; [5, Theorem 6.7] is stated only for the case when $T_{\widehat{\kappa}}$ is Hilbert–Schmidt, but the proof assumes only that $T_{\widehat{\kappa}}$ is compact). Hence, there exists $\varepsilon > 0$ such that $\|T_{\widehat{\kappa}}\| < 1$. But then (4.28) holds, because $\langle T_{\widehat{\kappa}}^\ell 1, 1 \rangle_{\widehat{\mu}} \leq \|T_{\widehat{\kappa}}\|^\ell$. \square

Remark 4.16. Chayes and Smith [14] have recently proved a result related to Theorem 4.7(i) and Theorem 4.8, for the special case where the type space \mathcal{S} is finite. Their model has a fixed number of vertices of each type, which makes essentially no difference in this finite-type case. Chayes and Smith consider (in effect) the number of ordered pairs (v, w) of vertices with v of type i , w of type j , and v and w in the same component, normalized by dividing by n , showing convergence to the relevant branching process quantity.

These numbers sum to give the susceptibility, so such a result is more refined than the corresponding result for the susceptibility itself.

In our setting, the analogue is to fix arbitrary measurable subsets S and T of the type space, and consider $\chi_{S,T}(G_n)$, defined as $1/n$ times the number of pairs (v, w) in the same component with the type of v lying in S and that of w in T . The corresponding branching process quantity is just $\chi_{S,T}(\kappa)$, the integral over $x \in S$ of the expected number of particles in $\mathfrak{X}_\kappa(x)$ with types in T . In analogy with Theorem 3.3, in the subcritical case this quantity may be expressed as $\chi_{S,T}(\kappa) = \langle (I - T_\kappa)^{-1} 1_S, 1_T \rangle_\mu < \infty$. It is not hard to see that the proof of Theorem 4.8 in fact shows that

$$\chi_{S,T}(G_n) \xrightarrow{P} \chi_{S,T}(\kappa), \quad (4.31)$$

where $G_n = G^\mathcal{V}(n, \kappa)$ is defined on an i.i.d. vertex space. The key point is that, in the light of Theorem 4.1 and its proof, it suffices to prove a convergence result for the contribution to $\chi_{S,T}(G_n)$ from components of a fixed size k . For all the models we consider here, this may be proved by adapting the methods used to prove convergence of $N_k(G_n)/n$; we omit the details. Once we have such convergence, we also obtain the analogue of (4.31) for $\widehat{\chi}$, so all our results in this section may be extended in this way, with the proviso that when considering $G^\mathcal{V}(n, \kappa)$ with a general vertex space \mathcal{V} as in [5], we must assume that S and T are μ -continuity sets.

Remark 4.17. We believe that all the results in this section extend, with suitable modifications, to the random graphs with clustering introduced by Bollobás, Janson and Riordan [7], and generalized (to a form analogous to $G(A_n)$) in [6]; these may be seen as the simple graphs obtained from an appropriate random hypergraph by replacing each hyperedge by a complete graph on its vertex set. Note that in this case the appropriate limiting object is a hyperkernel (for the definitions see [7]), and the corresponding branching process is now a (multi-type, of course) compound Poisson one.

A key observation is that in such a graph, which is the union of certain complete graphs, two vertices are in the same component if and only if they are joined by a path which uses at most one edge from each of these complete graphs. Roughly speaking, this means that we need consider only the individual edge probabilities, and not their correlations, and then arguments such as the proof of Theorem 4.8 and (at least the first part of) Theorem 4.7 go through with little change. It also tells us that the susceptibility of a hyperkernel is simply that of the corresponding edge kernel; this is no surprise, since for the expected total size of the branching process all that matters is (informally) the expected number of type- y children of each type- x individual, not the details of the distribution. This does not extend to the modified susceptibility $\widehat{\chi}$, since this depends on the (type-dependent) survival probability $\rho(\kappa; x)$, which certainly is sensitive to the details of the offspring distribution.

Adapting the proof of Theorem 4.9 needs more work, but we believe it should be possible. Most of the time, one can work with bounded hyperkernels, where not only are the individual (hyper)matrix entries uniformly bounded, but there is a maximum edge cardinality. Taking the r -uniform case for simplicity, one needs to show that the number

of $(r-1)$ -tuples of vertices in the giant component in some subset of \mathcal{S}^{r-1} is typically close to what it should be, since, in the proof of Lemma 4.15, the sets $A_{y,i}$ should (presumably) be replaced by corresponding subsets of \mathcal{S}^{r-1} . For strong concentration, one argues as here but using the appropriate stability result from [6] in place of Theorem 4.12. Needless to say, since we have not checked the details, there is always the possibility of unforeseen complications!

5 Behaviour near the threshold

In this section we consider the behaviour of χ and $\widehat{\chi}$ for a family $\lambda\kappa$ of kernels, with κ fixed and the ‘scaling factor’ λ ranging from 0 to ∞ . Since $\|T_{\lambda\kappa}\| = \lambda\|T_\kappa\|$, then, as discussed in [5], $\lambda\kappa$ is subcritical, critical and supercritical for $\lambda < \lambda_{\text{cr}}$, $\lambda = \lambda_{\text{cr}}$ and $\lambda > \lambda_{\text{cr}}$, respectively, where $\lambda_{\text{cr}} = \|T_\kappa\|^{-1}$. Note that if $\|T_\kappa\| < \infty$, then $\lambda_{\text{cr}} > 0$, while if $\|T_\kappa\| = \infty$, then $\lambda_{\text{cr}} = 0$, so $\lambda\kappa$ is supercritical for any $\lambda > 0$.

Note also that Theorem 3.5 provides an alternative way to find λ_{cr} (and thus $\|T_\kappa\|$): we can try to solve the integral equation $f = 1 + T_{\lambda\kappa}f = 1 + \lambda T_\kappa f$ and see whether there are any integrable positive solutions. By Theorem 3.5 this tells us whether $\chi(\lambda\kappa)$ is finite; since, by Theorems 3.3 and 3.4, the susceptibility is finite in the subcritical case and infinite in the supercritical case, this information determines λ_{cr} . The advantage of this approach over attempting to solve (2.17) itself is that the equation is linear; this is one of the main motivations for studying χ . (Another is that it tends to evolve very simply in time in suitably parameterized models.)

In the subcritical case, $\lambda < \lambda_{\text{cr}}$, we have the following simple result. As usual, when we say that a function f defined on the reals is *analytic* at a point x , we mean that there is a neighbourhood of x in which f is given by the sum of a convergent power series; equivalently, f extends to a complex analytic function in a complex neighbourhood of x .

Theorem 5.1. *Let κ be a kernel. Then the function $\lambda \mapsto \chi(\lambda\kappa) = \widehat{\chi}(\lambda\kappa)$ is increasing and analytic on $(0, \lambda_{\text{cr}})$, with a singularity at λ_{cr} . Furthermore, $\chi(\lambda\kappa) \nearrow \chi(\lambda_{\text{cr}}\kappa) = \widehat{\chi}(\lambda_{\text{cr}}\kappa) \leq \infty$ as $\lambda \nearrow \lambda_{\text{cr}}$, and $\chi(\lambda\kappa; x) \nearrow \chi(\lambda_{\text{cr}}\kappa; x)$ pointwise.*

Proof. By (3.2),

$$\chi(\lambda\kappa) = \mu(\mathcal{S})^{-1} \sum_{j=0}^{\infty} \langle T_\kappa^j 1, 1 \rangle \lambda^j, \quad (5.1)$$

which converges for $0 < \lambda < \lambda_{\text{cr}}$ by Theorem 3.3. Hence, $\chi(\lambda\kappa)$ is increasing and analytic on $(0, \lambda_{\text{cr}})$. Moreover, by Theorem 3.4(ii), the sum in (5.1) diverges for $\lambda > \lambda_{\text{cr}}$; hence the radius of convergence of this power series is λ_{cr} . Since the coefficients are non-negative, this implies that $\chi(\lambda\kappa)$ is not analytic at λ_{cr} .

Finally, $\chi(\lambda\kappa) \nearrow \chi(\lambda_{\text{cr}}\kappa)$ as $\lambda \nearrow \lambda_{\text{cr}}$ by (5.1) and monotone convergence. Similarly, $\chi(\lambda\kappa; x) \nearrow \chi(\lambda_{\text{cr}}\kappa; x)$ by (3.1) and monotone convergence. \square

We shall see in Section 6.3 that it is possible to have $\chi(\lambda_{\text{cr}}\kappa) < \infty$. As we shall now show, if T_κ is compact, then $\chi(\lambda_{\text{cr}}\kappa) = \infty$, and the critical exponent of χ is -1 as $\lambda \nearrow \lambda_{\text{cr}}$.

Theorem 5.2. *Suppose that T_κ is compact (for example, that $\int \kappa^2 < \infty$). Then for some constant a , $0 < a \leq 1$, we have*

$$\chi(\lambda\kappa) = \widehat{\chi}(\lambda\kappa) = \frac{a\lambda_{\text{cr}}}{\lambda_{\text{cr}} - \lambda} + O(1), \quad 0 < \lambda < \lambda_{\text{cr}},$$

and $\chi(\lambda_{\text{cr}}\kappa) = \widehat{\chi}(\lambda_{\text{cr}}\kappa) = \infty$.

If, in addition, κ is irreducible, then $a = (\int_{\mathcal{S}} \psi)^2 / \int_{\mathcal{S}} \psi^2$, where ψ is any non-negative eigenfunction of T_κ .

Proof. Since a compact operator is bounded, $\lambda_{\text{cr}} > 0$. We may assume that $\mu(\mathcal{S}) = 1$ by Remark 2.3. Furthermore, we may replace κ by $\lambda_{\text{cr}}\kappa$ and may thus assume, for convenience, that $\|T_\kappa\| = 1$ and $\lambda_{\text{cr}} = 1$.

We use Theorem 2.7. Since $\|T_\kappa\| = 1$, we have $\sigma(T_\kappa) \subset [-1, 1]$. Let E_1 be the eigenspace $\{f \in L^2(\mu) : T_\kappa f = f\}$ of T_κ , and let P_1 be the orthogonal projection onto E_1 . Since T_κ is compact and self-adjoint, E_1 and its orthogonal complement are T_κ -invariant. Furthermore, if $r := \max\{s \in \sigma(T_\kappa) \setminus \{1\}\}$, then $r < 1$ and for $0 \leq \lambda \leq 1$,

$$\|(I - \lambda T_\kappa)^{-1}(I - P_1)\| = \sup_s \{(1 - \lambda s)^{-1} : s \in \sigma(T_\kappa) \setminus \{1\}\} = \frac{1}{1 - \lambda r} \leq \frac{1}{1 - r}.$$

On the other hand, since T_κ is the identity on E_1 , we have $(I - \lambda T_\kappa)^{-1}P_1 = (1 - \lambda)^{-1}P_1$ for $\lambda < 1$. Consequently, by Theorem 3.3,

$$\chi(\lambda\kappa) = \langle (I - \lambda T_\kappa)^{-1}1, 1 \rangle = (1 - \lambda)^{-1} \langle P_1 1, 1 \rangle + O(1).$$

Let $a := \langle P_1 1, 1 \rangle = \|P_1 1\|_2^2 \geq 0$; then $a \leq \|1\|_2^2 = 1$, so $0 \leq a \leq 1$. If $a = 0$, then $P_1 1 = 0$, so the constant function 1 is orthogonal to E_1 . But this contradicts the fact that E_1 always contains a non-zero eigenfunction $\psi \geq 0$, see the proof of Theorem 3.4 and [5, Lemma 5.15]. Hence, $a > 0$.

The fact that $\chi(\lambda_{\text{cr}}\kappa) = \infty$ now follows from Theorem 5.1.

Furthermore, if κ is irreducible, then E_1 is one-dimensional, see again [5, Lemma 5.15 and its proof], so $P_1 f = \|\psi\|_2^{-2} \langle f, \psi \rangle \psi$, and the formula for a follows, noting that every non-negative eigenfunction is a multiple of this ψ . \square

In a moment, we shall discuss the supercritical case. First we state a lemma from perturbation theory that we shall need in the proof. The exact form of this lemma is adapted to our purposes; see [17, Section VII.6] or [26] for similar arguments and many related results.

Lemma 5.3. *Let T be a compact self-adjoint operator on $L^2(\mu)$, such that T has a largest eigenvalue 1 that is simple, with a corresponding normalized eigenfunction ψ . Then there exists $\eta > 0$ such that if T' is any self-adjoint operator with $\|T' - T\| < \eta$ such that $I - T'$ is invertible, then*

$$\langle (I - T')^{-1}f, g \rangle = \frac{\langle f, \psi \rangle \langle \psi, g \rangle + O(\|T' - T\|)}{1 - \langle T'\psi, \psi \rangle + O(\|T' - T\|^2)} + O(1) \quad (5.2)$$

uniformly for all $f, g \in L^2(\mu)$ with $\|f\|, \|g\| \leq 1$.

Proof. In this proof (only!) we use the complex version of $L^2(\mu)$, since the spectral theory is more complete in the complex case. The result then holds in the real case too. We assume for simplicity that T' too is compact; this holds in our application below. (The general case is similar but uses more advanced spectral theory, see [15, §§VII.2 and IX.2] or [17, VII.3].)

By Theorem 2.7(i), the spectrum $\sigma(T)$ is contained in $(-\infty, 1 - \delta] \cup \{1\}$ for some $\delta > 0$. Let γ be the circle $\{z : |z - 1| = \delta/2\}$. Then the minimum distance between γ and $\sigma(T)$ is $\delta/2 > 0$. For $z \in \gamma$, on each eigenspace E_λ of T the map $(zI - T)^{-1}$ is simply multiplication by $(z - \lambda)^{-1}$. By Theorem 2.7(iii), $L^2(\mu)$ is the orthogonal direct sum of the E_λ . Since the factors $(z - \lambda)^{-1}$ are uniformly bounded (by $2/\delta$), it follows that for any $f \in L^2(\mu)$ we have

$$(zI - T)^{-1}f = \sum_{\lambda \in \sigma(T)} (z - \lambda)^{-1}P_\lambda f, \quad (5.3)$$

where P_λ is the orthogonal projection onto E_λ . Furthermore, the sum converges uniformly in $z \in \gamma$. Integrating, and using Cauchy's integral formula on each term, we see that the 'spectral projection'

$$Q_0 := \frac{1}{2\pi i} \oint_\gamma (zI - T)^{-1} dz \quad (5.4)$$

is the orthogonal projection onto the sum of the eigenspaces E_λ of T for λ in the interior of γ . (See [17, §VII.3] for more general results.) In our case the only eigenvalue inside γ is 1, so $Q_0 = P_1$, the projection onto E_1 . Furthermore, by assumption E_1 is the one-dimensional space spanned by ψ , so, for any $f \in L^2(\mu)$,

$$Q_0 f = P_1 f = \langle f, \psi \rangle \psi. \quad (5.5)$$

Let $A = T' - T$, and suppose that $\|A\| \leq \eta := \delta/4$. For $z \in \gamma$, let $X = zI - T$, so $zI - T - A = X - A = X(I - X^{-1}A)$. From the formula (5.3) we have $\|X^{-1}\| \leq 2/\delta$, so $\|X^{-1}A\| \leq \|X^{-1}\|\|A\| \leq 1/2$. Hence the sum $\sum_{r \geq 0} (X^{-1}A)^r X^{-1}$ converges in the space of operators on $L^2(\mu)$ and (multiplying out) is the inverse of $zI - T - A$. Note that $(zI - T - A)^{-1} - (zI - T)^{-1} = \sum_{r \geq 1} (X^{-1}A)^r X^{-1}$ has norm at most $\sum_{r \geq 1} (2/\delta)^{r+1} \|A\|^r = O(\|A\|)$, uniformly in A with $\|A\| \leq \eta$.

Let

$$Q_A := \frac{1}{2\pi i} \oint_\gamma (zI - T - A)^{-1} dz. \quad (5.6)$$

Thus Q_A is the spectral projection for $T + A$ associated to the interior of γ . It follows from (5.4), (5.6) and the estimate on $\|(zI - T - A)^{-1} - (zI - T)^{-1}\|$ above that $\|Q_A - Q_0\| = O(\|A\|)$, so, reducing η if necessary, we have $\|Q_A - Q_0\| < 1$ whenever $\|A\| \leq \eta$. Then by [17, Lemma VII.6.7] Q_A too has rank 1; thus Q_A must be the orthogonal projection onto a one-dimensional space spanned by an eigenfunction ψ_A of $T + A$ with eigenvalue λ_A , with $|\lambda_A - 1| < \delta/2$. Moreover, if $I - T' = I - T - A$ is invertible, then $\lambda_A \neq 1$ so, since all other eigenvalues of $T + A$ lie outside γ ,

$$(I - (T + A))^{-1} = (1 - \lambda_A)^{-1}Q_A + R_A, \quad (5.7)$$

with $\|R_A\| \leq 2/\delta = O(1)$.

Since $\|Q_A\psi - \psi\| = \|(Q_A - Q_0)\psi\| = O(\|A\|)$, $Q_A\psi \neq 0$ (provided η is small enough), and thus we can take $\psi_A = Q_A\psi$. Hence $\|\psi_A - \psi\| = \|Q_A\psi - \psi\| = O(\|A\|)$ and, assuming $\|A\| \leq \eta$ (with η small),

$$\begin{aligned}\langle \psi_A, \psi \rangle &= \langle \psi, \psi \rangle + O(\|A\|) = 1 + O(\|A\|), \\ \langle T\psi_A, \psi \rangle &= \langle \psi_A, T\psi \rangle = \langle \psi_A, \psi \rangle, \\ \langle A\psi_A, \psi \rangle &= \langle A\psi, \psi \rangle + O(\|A\|^2),\end{aligned}$$

and thus, recalling that $A = T' - T$ and $T\psi = \psi$,

$$\begin{aligned}\lambda_A &= \frac{\langle (T + A)\psi_A, \psi \rangle}{\langle \psi_A, \psi \rangle} = 1 + \frac{\langle A\psi_A, \psi \rangle}{\langle \psi_A, \psi \rangle} = 1 + \langle A\psi, \psi \rangle + O(\|A\|^2) \\ &= \langle T\psi, \psi \rangle + \langle A\psi, \psi \rangle + O(\|A\|^2) = \langle T'\psi, \psi \rangle + O(\|A\|^2).\end{aligned}\tag{5.8}$$

Furthermore, for any f and g with $\|f\|, \|g\| \leq 1$, using (5.5),

$$\langle Q_A f, g \rangle = \langle Q_0 f, g \rangle + O(\|A\|) = \langle f, \psi \rangle \langle \psi, g \rangle + O(\|A\|).\tag{5.9}$$

The result follows from (5.7), (5.8) and (5.9). \square

In the supercritical case, only $\widehat{\chi}$ is of interest. If we allow reducible κ , we can have several singularities, coming from different parts of the type space; see Example 6.8. We therefore assume that κ is irreducible. Even in that case, it is possible that the dual kernel $\widehat{\kappa}$ is critical, see [5, Example 12.4]; in this example it is not hard to check that $\widehat{\chi}(\kappa)$ is infinite.

We conjecture that when κ is irreducible, $\widehat{\chi}(\lambda\kappa)$ is analytic for all $\lambda \neq \lambda_{\text{cr}}$ under very weak conditions, but we have only been able to show this under the rather stringent condition (5.10) below. (See also the examples in Section 6.) Under this condition, we can also show that the behaviour of $\widehat{\chi}$ is symmetric at λ_{cr} to first order: the asymptotic behaviour is the same on the subcritical and supercritical sides. As seen in Examples 6.2 and 6.3, this does not hold for all κ , even if we assume the Hilbert–Schmidt condition $\int \kappa^2 < \infty$. (Furthermore, we shall see in Sections 6.1 and 6.2 that the second order terms generally differ between the two sides.)

Theorem 5.4. *Suppose that κ is irreducible, and that*

$$\sup_x \int_{\mathcal{S}} \kappa(x, y)^2 d\mu(y) < \infty.\tag{5.10}$$

(i) *The function $\lambda \mapsto \widehat{\chi}(\lambda\kappa)$ is analytic except at $\lambda_{\text{cr}} := \|T_\kappa\|^{-1}$.*

(ii) *As $\lambda \rightarrow \lambda_{\text{cr}}$,*

$$\widehat{\chi}(\lambda\kappa) = \frac{b\lambda_{\text{cr}}}{|\lambda - \lambda_{\text{cr}}|} + O(1),$$

with $b = (\int_{\mathcal{S}} \psi)^2 / \int_{\mathcal{S}} \psi^2 > 0$, where ψ is any non-negative eigenfunction of T_κ .

Proof. Note that (5.10) implies that T_κ is Hilbert–Schmidt and thus compact. The subcritical case $\lambda < \lambda_{\text{cr}}$ thus follows from Theorem 5.2, so we assume $\lambda > \lambda_{\text{cr}}$. We may also assume that $\mu(\mathcal{S}) = 1$.

(i): Let $\lambda_0 > \lambda_{\text{cr}}$. By [5, Section 15], there exists an analytic function $z \mapsto \rho_z^+$ defined in a complex neighbourhood U of λ_0 with values in the Banach space $L^2(\mu)$ such that $\rho_z^+ = \rho_{z\kappa}$ when z is real, and (2.17) extends to

$$\rho_z^+ = 1 - e^{-zT_\kappa\rho_z^+}. \quad (5.11)$$

Shrinking U if necessary, we may assume that $\|\rho_z^+\|_2$ is bounded in U . Then, by (5.10) and Cauchy–Schwartz, $\|T_\kappa(\rho_z^+)\|_\infty = O(1)$ in U , and thus, by (5.11), $|1 - \rho_z^+|$ is bounded above and below, uniformly for $z \in U$. In particular, for every $\lambda\kappa$ with real $\lambda \in U$, $L^2(\widehat{\mu}) = L^2(\mu)$, with uniformly equivalent norms. We can therefore regard $T_{\widehat{\lambda\kappa}}$ as an operator on $L^2(\mu)$.

We define, for $z \in U$, $\widehat{T}_z f := zT_\kappa((1 - \rho_z^+)f)$; thus $\widehat{T}_\lambda = T_{\widehat{\lambda\kappa}}$ for real $\lambda \in U$ by (2.19). Note that $z \mapsto \widehat{T}_z$ is an analytic map of U into the Banach space of bounded operators on $L^2(\mu)$.

By Theorem 3.3, $I - T_{\widehat{\lambda_0\kappa}}$ is invertible. By continuity, we may assume that $I - \widehat{T}_z$ is invertible in U . Then $f(z) := \langle (I - \widehat{T}_z)^{-1}1, 1 - \rho_z^+ \rangle_\mu$ is an analytic function on U , and $f(\lambda) = \widehat{\chi}(\lambda\kappa)$ for real $\lambda \in U$ by Theorem 3.3(ii). Hence $\widehat{\chi}(\lambda\kappa)$ is analytic at λ_0 .

(ii): We may rescale and assume that $\lambda_{\text{cr}} = \|T_\kappa\| = 1$, i.e., κ is critical.

It will be convenient to use the fixed Hilbert space $L^2(\mu)$ rather than $L^2(\widehat{\mu})$; recall that $\widehat{\mu}$ depends on λ . Define a self-adjoint operator \widetilde{T}_λ on $L^2(\mu)$ by

$$\widetilde{T}_\lambda f := (1 - \rho_{\lambda\kappa})^{1/2} \lambda T_\kappa (f(1 - \rho_{\lambda\kappa})^{1/2}), \quad (5.12)$$

and note that if U_λ is the unitary mapping $f \mapsto (1 - \rho_{\lambda\kappa})^{1/2} f$ of $L^2(\widehat{\mu})$ onto $L^2(\mu)$, then $\widetilde{T}_\lambda = U_\lambda T_{\widehat{\lambda\kappa}} U_\lambda^{-1}$ by (2.19). Hence, \widetilde{T}_λ as an operator on $L^2(\mu)$ is unitarily equivalent to $T_{\widehat{\lambda\kappa}}$ on $L^2(\widehat{\mu})$. Further, by Theorem 3.3(ii),

$$\widehat{\chi}(\lambda\kappa) = \langle (I - T_{\widehat{\lambda\kappa}})^{-1}1, 1 \rangle_{\widehat{\mu}} = \langle (I - \widetilde{T}_\lambda)^{-1}U_\lambda 1, U_\lambda 1 \rangle_\mu. \quad (5.13)$$

Note that $\rho_\kappa = 0$, and thus $\widetilde{T}_1 = T_\kappa$, which has a simple eigenvalue 1, with a positive eigenfunction ψ [5, Lemma 5.15], and all other eigenvalues strictly smaller. We may assume that $\|\psi\|_2 = 1$.

We apply Lemma 5.3 with $T = \widetilde{T}_1$ and $T' = \widetilde{T}_\lambda$, with $\lambda = 1 + \varepsilon$ for small $\varepsilon > 0$. By [5, Section 15], $\|\rho_{\lambda\kappa}\|_\infty = O(\varepsilon)$, and more precisely, $\rho_{\lambda\kappa} = a_\varepsilon\psi + \rho_\varepsilon^*$ with $\|\rho_\varepsilon^*\|_2 = O(\varepsilon^2)$ and

$$a_\varepsilon = \frac{2}{\int_{\mathcal{S}} \psi^3 d\mu} \varepsilon + O(\varepsilon^2). \quad (5.14)$$

It follows (recalling that ψ is bounded because $\psi = T_\kappa\psi$ and (5.10) holds) that $(1 - \rho_{\lambda\kappa})^{1/2}\psi = \psi - \frac{1}{2}a_\varepsilon\psi^2 + r_\varepsilon$, with $\|r_\varepsilon\|_2 = O(\varepsilon^2)$. Consequently, (5.12) implies that $\|\widetilde{T}_\lambda -$

$\|\tilde{T}_1\| = O(\varepsilon)$ and, using $\langle T_\kappa \psi^2, \psi \rangle = \langle \psi^2, T_\kappa \psi \rangle = \langle \psi^2, \psi \rangle = \int_S \psi^3 d\mu$ and (5.14),

$$\begin{aligned} \langle \tilde{T}_\lambda \psi, \psi \rangle &= \lambda \langle T_\kappa((1 - \rho_{\lambda\kappa})^{1/2} \psi), (1 - \rho_{\lambda\kappa})^{1/2} \psi \rangle \\ &= \lambda (\langle T_\kappa \psi, \psi \rangle - \frac{1}{2} a_\varepsilon \langle T_\kappa \psi, \psi^2 \rangle - \frac{1}{2} a_\varepsilon \langle T_\kappa \psi^2, \psi \rangle + O(\varepsilon^2)) \\ &= (1 + \varepsilon)(1 - 2\varepsilon + O(\varepsilon^2)) \\ &= 1 - \varepsilon + O(\varepsilon^2). \end{aligned}$$

Further, $U_\lambda 1 = (1 - \rho_{\lambda\kappa})^{1/2} = 1 + O(\varepsilon)$. Hence, (5.13) and (5.2) yield

$$\hat{\chi}((1 + \varepsilon)\kappa) = \frac{\langle 1, \psi \rangle^2 + O(\varepsilon)}{\varepsilon + O(\varepsilon^2)} + O(1) = \frac{\langle 1, \psi \rangle^2}{\varepsilon} + O(1),$$

which is the desired result. □

6 Examples

In this section we give several examples illustrating the results in the rest of the paper and their limits. We sometimes drop κ from the notation: we let ρ_k denote the function $x \mapsto \rho_k(x) = \rho_k(\kappa; x)$. (But we continue to denote the number $\int_S \rho_k d\mu$ by $\rho_k(\kappa)$, in order to distinguish it from the function ρ_k .)

Note first that the probabilities $\rho_k(x)$ can in principle be calculated by recursion and integration. The number of children of an individual of type x in the branching process \mathfrak{X}_κ is Poisson with mean $\int \kappa(x, y) d\mu(y) = T_\kappa 1(x)$, and thus (in somewhat informal language)

$$\rho_1(x) = \mathbb{P}(x \text{ has no child}) = e^{-T_\kappa 1(x)}. \quad (6.1)$$

Next, $|\mathfrak{X}_\kappa(x)| = 2$ if and only if x has a single child, which is childless. Hence, by conditioning on the offspring of x ,

$$\begin{aligned} \rho_2(x) &= e^{-T_\kappa 1(x)} \int_S \kappa(x, y) \mathbb{P}(|\mathfrak{X}_\kappa(y)| = 1) d\mu(y) = e^{-T_\kappa 1(x)} T_\kappa(\rho_1)(x) \\ &= \rho_1(x) T_\kappa(\rho_1)(x). \end{aligned} \quad (6.2)$$

Similarly, considering the two ways to get $|\mathfrak{X}_\kappa(x)| = 3$,

$$\begin{aligned} \rho_3(x) &= e^{-T_\kappa 1(x)} \int_S \kappa(x, y) \rho_2(y) d\mu(y) \\ &\quad + e^{-T_\kappa 1(x)} \frac{1}{2} \int_S \kappa(x, y) \rho_1(y) d\mu(y) \int_S \kappa(x, z) \rho_1(z) d\mu(z) \\ &= \rho_1(x) T_\kappa(\rho_2)(x) + \frac{1}{2} \rho_1(x) (T_\kappa(\rho_1)(x))^2, \end{aligned} \quad (6.3)$$

and the three ways to get $|\mathfrak{X}_\kappa(x)| = 4$,

$$\rho_4 = \rho_1 T(\rho_3) + \rho_1 T(\rho_1) T(\rho_2) + \frac{1}{6} \rho_1 (T(\rho_1))^3, \quad (6.4)$$

and so on. In general, in the expression for ρ_k , $k \geq 2$, there is one term $\rho_1 \prod_j T(\rho_j)^{m_j}/m_j!$ for each partition $1^{m_1}2^{m_2} \dots$ of $k - 1$.

The numbers $\rho_k(\kappa)$ are then obtained by integration.

Alternatively, a similar recursion can be given for the probability that $\mathfrak{X}_\kappa(x)$ has the shape of a given tree; this can then be summed over all trees of a given size.

6.1 The Erdős–Rényi case

Let \mathcal{S} consist of a single point, with $\mu(\mathcal{S}) = 1$. Thus, κ is a positive number. (More generally, a constant κ on any probability space (\mathcal{S}, μ) yields the same results. See [5, Example 4.1].) We keep to more traditional notation by letting $\kappa = \lambda > 0$; then $G^\nu(n, \kappa) = G(n, p)$ with $p = \lambda/n$. We write in this case e.g. T_λ , $\chi(\lambda)$ and $\rho(\lambda)$.

Since T_λ is just multiplication by λ , $\|T_\lambda\| = \lambda$, and, as is well-known, λ is subcritical if $\lambda < 1$, critical if $\lambda = 1$, and supercritical if $\lambda > 1$.

In the subcritical case, by (3.2) or Theorem 3.3(i),

$$\chi(\lambda) = \frac{1}{1 - \lambda}, \quad \lambda < 1. \quad (6.5)$$

Theorem 4.7 or Theorem 4.8 shows that $\chi(G(n, \lambda/n)) \xrightarrow{p} (1 - \lambda)^{-1}$ for every constant $\lambda < 1$. (This and more detailed results are shown by Janson and Luczak [24] by another method. See also Durrett [19, Section 2.2] for the expectation $\mathbb{E} \chi(G(n, \lambda/n))$.)

Similarly, if $\lambda \geq 1$ then $\chi(G(n, \lambda/n)) \xrightarrow{p} \chi(\lambda) = \infty$ by Theorem 3.4 and any of Theorems 4.5, 4.7 or 4.8.

For $\widehat{\chi}$, we have the same results for $\lambda \leq 1$. In the supercritical case $\lambda > 1$, $T_{\widehat{\kappa}}$ is multiplication by $\lambda(1 - \rho(\lambda)) < 1$, where $1 - \rho(\lambda) = \exp(-\lambda\rho(\lambda))$ by (2.17). Hence, by Theorems 4.7 and 3.3, or (3.4), for $\lambda > 1$,

$$\widehat{\chi}(G(n, \lambda/n)) \xrightarrow{p} \widehat{\chi}(\lambda) = \frac{\widehat{\mu}(\mathcal{S})}{1 - \lambda(1 - \rho(\lambda))} = \frac{1 - \rho(\lambda)}{1 - \lambda(1 - \rho(\lambda))}. \quad (6.6)$$

More generally, Theorem 4.7 shows that $\widehat{\chi}(G(n, \lambda_n/n)) \xrightarrow{p} \widehat{\chi}(\lambda)$ for every sequence $\lambda_n \rightarrow \lambda > 0$.

For $\lambda = 1 + \varepsilon$, $\varepsilon > 0$, we have the Taylor expansion

$$\rho(1 + \varepsilon) = 2\varepsilon - \frac{8}{3}\varepsilon^2 + \frac{28}{9}\varepsilon^3 - \frac{464}{135}\varepsilon^4 + \dots \quad (6.7)$$

and thus

$$\widehat{\chi}(1 + \varepsilon) = \varepsilon^{-1} - \frac{4}{3} + \frac{4}{3}\varepsilon - \frac{176}{135}\varepsilon^2 + \dots \quad (6.8)$$

Combining (6.5) and (6.8), we see that, as shown by Theorem 5.4, $\widehat{\chi}(\lambda) \sim 1/|\lambda - 1|$ for λ on both sides of 1, but the second order terms are different for $\lambda \nearrow 1$ and $\lambda \searrow 1$.

We can also obtain $\chi(\lambda)$ and $\widehat{\chi}(\lambda)$ from ρ_k and the formulae (2.10) and (2.11). In this case, \mathfrak{X}_κ is an ordinary, single-type, Galton–Watson process with Poisson distributed offspring, and it is well-known, see e.g. [10; 29; 35; 20; 34; 30], that $|\mathfrak{X}_\kappa|$ has a Borel distribution (degenerate if $\lambda > 1$), i.e.,

$$\rho_k(\lambda) = \rho_k(x) = \frac{k^{k-1}}{k!} \lambda^{k-1} e^{-k\lambda}, \quad k \geq 1. \quad (6.9)$$

Consequently, if $\mathcal{T}(z) := \sum_{k=1}^{\infty} \frac{k^{k-1}}{k!} z^k$ is the tree function, then

$$\rho(\lambda) = 1 - \sum_{1 \leq k < \infty} \rho_k(\lambda) = 1 - \frac{\mathcal{T}(\lambda e^{-\lambda})}{\lambda} \quad (6.10)$$

and, using the well-known identity $z\mathcal{T}'(z) = \mathcal{T}(z)/(1 - \mathcal{T}(z))$, see e.g. [23],

$$\widehat{\chi}(\lambda) = \sum_{1 \leq k < \infty} k \rho_k(\lambda) = \sum_{k=1}^{\infty} \frac{k^k}{k!} \lambda^{k-1} e^{-k\lambda} = \lambda^{-1} \frac{\mathcal{T}(\lambda e^{-\lambda})}{1 - \mathcal{T}(\lambda e^{-\lambda})}. \quad (6.11)$$

In the subcritical case, when $\lambda < 1$, we have $\mathcal{T}(\lambda e^{-\lambda}) = \lambda$, and we recover (6.5). In general, (6.10) and (6.11) yield (6.6).

Remark 6.1. Consider the random graph $G(n, m)$ with a given number m of edges. In the subcritical case $m \sim \lambda n/2$ with $0 < \lambda < 1$, we obtain $\chi(G(n, m)) \xrightarrow{P} \chi(\lambda) = 1/(1 - \lambda)$ by comparison with $G(n, p)$ with $p = \lambda_n/n$ for $\lambda_n = 2m/n \pm n^{-1/3}$, say, using Lemma 2.6. In the supercritical case $\lambda > 1$, one can use standard results on the numbers of vertices and edges in the giant component: conditioning on the giant component assuming typical values, the rest of the graph is essentially a subcritical instance of $G(n, m)$ with different parameters; this may be compared with $G(n, p)$ as above. Consequently, for $m \sim \lambda n/2$ with $\lambda > 1$, $\widehat{\chi}(G(n, m)) \xrightarrow{P} \widehat{\chi}(\lambda)$, where $\widehat{\chi}(\lambda)$ is given by (6.6) and (6.11), just as for $G(n, p)$ with $p = \lambda/n$.

6.2 The rank 1 case

Suppose that $\kappa(x, y) = \psi(x)\psi(y)$ for some positive integrable function ψ on \mathcal{S} . This is the *rank 1 case* studied in [5, Section 16.4]; note that T_κ is the rank 1 operator $f \mapsto \langle f, \psi \rangle \psi$, with ψ as eigenfunction, provided $\psi \in L^2(\mu)$.

We assume, for simplicity, that $\mu(\mathcal{S}) = 1$. As in Section 5 we consider the family of kernels $\lambda\kappa$, $\lambda > 0$. In this case, $\|T_\kappa\| = \|\psi\|_2^2 = \int_{\mathcal{S}} \psi^2$, and thus $\lambda_{\text{cr}} = \|\psi\|_2^{-2}$.

In the subcritical case, $\lambda < \lambda_{\text{cr}} = (\int_{\mathcal{S}} \psi^2)^{-1}$, which entails $\int_{\mathcal{S}} \psi^2 < \infty$, we have by induction

$$T_{\lambda\kappa}^j 1(x) = \lambda^j \left(\int_{\mathcal{S}} \psi^2 d\mu \right)^{j-1} \int_{\mathcal{S}} \psi d\mu \cdot \psi(x), \quad j \geq 1,$$

and thus by (3.2) (or by solving (3.5))

$$\begin{aligned}\chi(\lambda\kappa) = \widehat{\chi}(\lambda\kappa) &= 1 + \frac{\lambda (\int \psi)^2}{1 - \lambda \int \psi^2} = 1 + \frac{\lambda (\int \psi)^2}{1 - \lambda/\lambda_{\text{cr}}} \\ &= \frac{(\int \psi)^2 / \int \psi^2}{1 - \lambda/\lambda_{\text{cr}}} + 1 - \frac{(\int \psi)^2}{\int \psi^2}.\end{aligned}\tag{6.12}$$

In particular, this verifies the formula in Theorem 5.2.

In the supercritical case, we first note that the equation (2.17) for $\rho = \rho_{\lambda\kappa}$ becomes

$$\rho = 1 - e^{-\lambda T_{\kappa}\rho} = 1 - e^{-\lambda\langle\rho,\psi\rangle\psi}.\tag{6.13}$$

We define $\xi \in (0, \infty)$ by $\xi := \lambda\langle\rho, \psi\rangle$, and thus have

$$\rho = 1 - e^{-\xi\psi},\tag{6.14}$$

with ξ given by the implicit equation

$$\xi = \lambda \int_S \rho(x)\psi(x) \, d\mu(x) = \lambda \int_S \psi(x) \left(1 - e^{-\xi\psi(x)}\right) \, d\mu(x).\tag{6.15}$$

(See [5, Section 16.4], where the notation is somewhat different.) We know, by [5, Theorem 6.1], that (6.13) has a unique positive solution ρ for every $\lambda > \lambda_{\text{cr}}$; thus (6.15) has a unique solution $\xi = \xi(\lambda) > 0$ for every $\lambda > \lambda_{\text{cr}}$.

It is easier to use ξ as a parameter; by (6.15) we have

$$\lambda = \frac{\xi}{\int (1 - e^{-\xi\psi}) \psi}.\tag{6.16}$$

The denominator is finite for every $\xi > 0$ since $\psi \in L^1$; moreover, $\int (1 - e^{-\xi\psi})\psi < \int \xi\psi^2$, and thus (6.16) yields $\lambda > 1/\int \psi^2 = \lambda_{\text{cr}}$. Consequently, (6.15) and (6.16) give a bijection between $\lambda \in (\lambda_{\text{cr}}, \infty)$ and $\xi \in (0, \infty)$. Furthermore, differentiation of (6.16) shows that $\lambda = \lambda(\xi)$ is differentiable, and it follows easily from $\int (1 - e^{-\xi\psi})\psi > \int \xi\psi^2 e^{-\xi\psi}$ that $d\lambda/d\xi > 0$. Hence, the function $\lambda(\xi)$ and its inverse $\xi(\lambda)$ are both strictly increasing and continuous. In particular, $\lambda \searrow \lambda_{\text{cr}} \iff \xi \searrow 0$. Moreover, the denominator in (6.16) is an analytic function of *complex* ξ with $\text{Re } \xi > 0$; hence $\lambda(\xi)$ and its inverse $\xi(\lambda)$ are analytic, for $\xi > 0$ and $\lambda > \lambda_{\text{cr}}$, respectively.

We note also the following equivalent formula, provided $\int_S \psi^2 < \infty$:

$$\frac{1}{\lambda_{\text{cr}}} - \frac{1}{\lambda} = \xi^{-1} \int_S (e^{-\xi\psi} - 1 + \xi\psi) \psi.\tag{6.17}$$

By (2.19) and (6.14),

$$T_{\widehat{\lambda\kappa}} f = T_{\lambda\kappa}((1 - \rho)f) = \lambda \langle (1 - \rho)f, \psi \rangle \psi = \lambda \int_S e^{-\xi\psi(x)} \psi(x) f(x) \, d\mu(x) \psi.\tag{6.18}$$

Hence $T_{\widehat{\lambda\kappa}}$ too is a rank 1 operator, with eigenfunction ψ and eigenvalue (take $f = \psi$ in (6.18))

$$\gamma = \lambda \int_{\mathcal{S}} e^{-\xi\psi(x)} \psi(x)^2 d\mu(x) = \frac{\xi \int e^{-\xi\psi} \psi^2}{\int (1 - e^{-\xi\psi}) \psi}. \quad (6.19)$$

Since $y^2 e^{-y} < y(1 - e^{-y})$ for $y > 0$, it follows that $0 < \gamma < 1$. (When $\int \psi^2 < \infty$, this follows also from the general result [5, Theorem 6.7], cf. Theorem 3.3.) Hence $I - T_{\widehat{\lambda\kappa}}$ is invertible (in, for example, $L^2(\widehat{\mu})$), and by Theorem 3.3(ii),

$$\widehat{\chi}(\lambda\kappa; x) = (1 - \rho(x))(I - T_{\widehat{\lambda\kappa}})^{-1}1(x) = e^{-\xi\psi(x)}(I - T_{\widehat{\lambda\kappa}})^{-1}1(x). \quad (6.20)$$

Let us write $g := (I - T_{\widehat{\lambda\kappa}})^{-1}1$. Then, by (6.18), $1 = (I - T_{\widehat{\lambda\kappa}})g = g - \zeta\psi$, with $\zeta = \lambda \int_{\mathcal{S}} e^{-\xi\psi} \psi g$. Hence, $g = 1 + \zeta\psi$ and, using (6.19),

$$\zeta = \lambda \int_{\mathcal{S}} e^{-\xi\psi} \psi g = \lambda \int_{\mathcal{S}} e^{-\xi\psi} \psi + \lambda\zeta \int_{\mathcal{S}} e^{-\xi\psi} \psi^2 = \lambda \int_{\mathcal{S}} e^{-\xi\psi} \psi + \zeta\gamma.$$

Hence, using (6.16) and (6.19),

$$\zeta = \frac{\lambda \int e^{-\xi\psi} \psi}{1 - \gamma} = \frac{\xi \int e^{-\xi\psi} \psi}{\int (1 - e^{-\xi\psi}) \psi - \xi \int e^{-\xi\psi} \psi^2}.$$

Finally, by (6.20),

$$\begin{aligned} \widehat{\chi} &= \int_{\mathcal{S}} \widehat{\chi}(\lambda\kappa; x) d\mu(x) = \int_{\mathcal{S}} e^{-\xi\psi} g = \int_{\mathcal{S}} e^{-\xi\psi} + \zeta \int_{\mathcal{S}} e^{-\xi\psi} \psi \\ &= \int_{\mathcal{S}} e^{-\xi\psi} + \frac{\xi (\int e^{-\xi\psi} \psi)^2}{\int (1 - e^{-\xi\psi} (1 + \xi\psi)) \psi}. \end{aligned} \quad (6.21)$$

We observe that (6.21) shows that $\widehat{\chi}$ is an analytic function of $\xi \in (0, \infty)$, and thus of $\lambda \in (\lambda_{\text{cr}}, \infty)$. (So in the rank 1 case, at least, the condition (5.10) is not required for Theorem 5.4(i).)

Next, suppose that $\int_{\mathcal{S}} \psi^3 < \infty$. In this case, we can differentiate twice under the integral signs in (6.16) and (6.21) using dominated convergence (comparing with $\int_{\mathcal{S}} \psi^3$), and taking Taylor expansions we see that as $\xi \rightarrow 0$ we have

$$\lambda = \frac{\xi}{\xi \int \psi^2 - \frac{1}{2}\xi^2 \int \psi^3 + o(\xi^2)} = \lambda_{\text{cr}} + \frac{1}{2}\xi \frac{\int \psi^3}{(\int \psi^2)^2} + o(\xi) \quad (6.22)$$

and

$$\widehat{\chi} = O(1) + \frac{\xi (\int \psi + O(\xi))^2}{\frac{1}{2}\xi^2 \int \psi^3 + o(\xi^2)} \sim \frac{2 (\int \psi)^2}{\int \psi^3} \xi^{-1} \sim \frac{(\int \psi)^2 / (\int \psi^2)^2}{\lambda - \lambda_{\text{cr}}}, \quad (6.23)$$

where we used (6.22) in the last step.

Note that (6.12) and (6.23) show that the behaviour of $\widehat{\chi}$ at the critical point λ_{cr} is symmetrical to the first order:

$$\widehat{\chi}(\lambda_{\kappa}) \sim \frac{(\int \psi)^2 / (\int \psi^2)^2}{|\lambda - \lambda_{\text{cr}}|} = \frac{(\int \psi)^2 / \int \psi^2}{|\lambda / \lambda_{\text{cr}} - 1|}, \quad \lambda \rightarrow \lambda_{\text{cr}}, \quad (6.24)$$

at least when $\int \psi^3 < \infty$. (This is the same first order asymptotics as given by Theorem 5.4(ii), but note that the latter applies only when ψ is bounded, since (5.10) fails otherwise.) The second order terms are different on the two sides of λ_{cr} , though: if $\int \psi^4 < \infty$, then carrying the Taylor expansions above one step further leads to

$$\begin{aligned} \widehat{\chi}(\lambda_{\kappa}) &= \frac{(\int \psi)^2 / \int \psi^2}{\lambda / \lambda_{\text{cr}} - 1} + 1 + \frac{(\int \psi)^2}{\int \psi^2} - \frac{4 \int \psi \int \psi^2}{\int \psi^3} + \frac{2 (\int \psi)^2 \int \psi^4}{3 (\int \psi^3)^2} \\ &\quad + o(1), \quad \lambda \searrow \lambda_{\text{cr}}, \end{aligned} \quad (6.25)$$

in contrast to (6.12) for $\lambda < \lambda_{\text{cr}}$.

To see what may happen if $\int_{\mathcal{S}} \psi^3 = \infty$, we look at a few specific examples.

Example 6.2. Let $2 < q < 3$ and take $\mathcal{S} = [1, \infty)$ with $d\mu(x) = qx^{-q-1} dx$, and take $\psi(x) = x$; note that $\int_{\mathcal{S}} \psi^p < \infty$ if and only if $p < q$; in particular $\int_{\mathcal{S}} \psi^2 < \infty$ but $\int_{\mathcal{S}} \psi^3 = \infty$. By (6.17), and standard integration by parts of Gamma integrals, as $\xi \rightarrow 0$ we have

$$\begin{aligned} \frac{1}{\lambda_{\text{cr}}} - \frac{1}{\lambda} &= \xi^{-1} \int_1^{\infty} (e^{-\xi x} - 1 + \xi x) qx^{-q} dx = q\xi^{q-2} \int_{\xi}^{\infty} (e^{-y} - 1 + y) y^{-q} dy \\ &\sim q\xi^{q-2} \int_0^{\infty} (e^{-y} - 1 + y) y^{-q} dy = q\xi^{q-2} \Gamma(1 - q), \end{aligned}$$

or $\lambda - \lambda_{\text{cr}} \sim q\Gamma(1 - q)\lambda_{\text{cr}}^2 \xi^{q-2}$. Similarly, by another integration by parts,

$$\begin{aligned} \int_{\mathcal{S}} (1 - e^{-\xi\psi}(1 + \xi\psi))\psi d\mu &= \int_1^{\infty} (1 - e^{-\xi x}(1 + \xi x))qx^{-q} dx \\ &= q\xi^{q-1} \int_{\xi}^{\infty} (1 - e^{-y}(1 + y))y^{-q} dy \sim q\xi^{q-1} \int_0^{\infty} (1 - e^{-y}(1 + y))y^{-q} dy \\ &= \frac{q\xi^{q-1}}{q-1} \Gamma(3 - q) = q(q-2)\xi^{q-1} \Gamma(1 - q), \end{aligned}$$

and thus by (6.21),

$$\widehat{\chi} \sim \frac{\xi (\int \psi)^2}{q(q-2)\xi^{q-1}\Gamma(1 - q)} \sim \frac{(\int \psi)^2 \lambda_{\text{cr}}^2}{(q-2)(\lambda - \lambda_{\text{cr}})}, \quad \lambda \searrow \lambda_{\text{cr}},$$

which still has power -1 , but differs by a factor $(q-2)^{-1}$ from the subcritical asymptotics in (6.12) and Theorem 5.2. Hence, (6.24) does *not* hold in general without assuming $\int_{\mathcal{S}} \psi^3 < \infty$. (Although this integral does not appear in the formula.)

Example 6.3. We see in Example 6.2 that $\widehat{\chi}$ is relatively large in the barely supercritical phase when ψ is only a little more than square integrable. We can pursue this further by taking the same \mathcal{S} and ψ , and $d\mu(x) = c(\log x + 1)^{-q}x^{-3} dx$ with $q > 1$ and a normalization constant c . Similar calculations using (6.17) and (6.23) (we omit the details) show that, letting c denote different positive constants (depending on q), as $\xi \rightarrow 0$ we have $\lambda - \lambda_{\text{cr}} \sim c(\log(1/\xi))^{-(q-1)}$ and $\widehat{\chi} \sim c(\log(1/\xi))^q$, and thus

$$\widehat{\chi}(\lambda\kappa) \sim c(\lambda - \lambda_{\text{cr}})^{-q/(q-1)}, \quad \lambda \searrow \lambda_{\text{cr}},$$

with an exponent $-q/(q-1)$, which can be any real number in $(-\infty, -1)$.

Taking instead $d\mu(x) = c(\log \log x)^{-2}(\log x)^{-1}x^{-3} dx$, $x > 3$, we similarly find $\lambda - \lambda_{\text{cr}} \sim c(\log \log(1/\xi))^{-1}$ and $\widehat{\chi} \sim c(\log(1/\xi))(\log \log(1/\xi))^2$, and thus

$$\widehat{\chi}(\lambda\kappa) = \exp\left(-\frac{c + o(1)}{\lambda - \lambda_{\text{cr}}}\right), \quad \lambda \searrow \lambda_{\text{cr}},$$

with an even more dramatic singularity. Of course, this sequence of examples can be continued to yield towers of exponents.

6.3 The CHKNS model

Consider the family of kernels $\lambda\kappa$, $\lambda > 0$, with

$$\kappa(x, y) := \frac{1}{x \vee y} - 1 \tag{6.26}$$

on $\mathcal{S} = (0, 1]$ with Lebesgue measure μ . We thus have

$$\begin{aligned} T_{\lambda\kappa}f(x) &= \lambda\left(\frac{1}{x} - 1\right) \int_0^x f(y) dy + \lambda \int_x^1 \left(\frac{1}{y} - 1\right) f(y) dy \\ &= \frac{\lambda}{x} \int_0^x f(y) dy + \lambda \int_x^1 \frac{f(y)}{y} dy - \lambda \int_0^1 f(y) dy. \end{aligned} \tag{6.27}$$

Remark 6.4. Equivalently, by a change of variable, we could consider the kernel $\lambda(e^{x \wedge y} - 1)$ on $\mathcal{S} = [0, \infty)$ with $d\mu = e^{-x} dx$; we leave it to the reader to reformulate results in this setting.

This kernel arises in connection with the CHKNS random graph model introduced by Callaway, Hopcroft, Kleinberg, Newman and Strogatz [13]. This graph grows from a single vertex; vertices are added one-by-one, and after each vertex is added, an edge is added with probability $\delta \in (0, 1)$; the endpoints are chosen uniformly among all existing vertices. Following Durrett [18; 19], we consider a modification where at each step a Poisson $\text{Po}(\delta)$ number of edges are added to the graph, again with endpoints chosen uniformly at random. As discussed in detail in [5, Section 16.3], this yields a random graph of the type $G^{\mathcal{V}}(n, \kappa_n)$ for a graphical sequence of kernels (κ_n) with limit $\lambda\kappa$, where $\lambda = 2\delta$, on a suitable vertex space \mathcal{V} (with \mathcal{S} and μ as above).

Let us begin by solving (3.5). If $f = T_{\lambda\kappa}f + 1$, then (6.27) implies first that $f \in C(0, 1)$ and then $f \in C^1(0, 1)$. Hence we can differentiate and find, using (6.27) again, that

$$f'(x) = (T_{\lambda\kappa}f)'(x) = -\frac{\lambda}{x^2} \int_0^x f(y) dy. \quad (6.28)$$

With $F(x) := \int_0^x f(y) dy$, this yields $F''(x) = -\lambda F(x)/x^2$, with the solution $F(x) = C_1x^{\alpha_+} + C_2x^{\alpha_-}$, where α_{\pm} are the roots of $\alpha(\alpha-1) = -\lambda$, i.e., $\alpha_{\pm} = \frac{1}{2} \pm \sqrt{\frac{1}{4} - \lambda}$; if $\lambda = 1/4$ we have a double root $\alpha_+ = \alpha_- = 1/2$ and the solution is $F(x) = C_1x^{1/2} + C_2x^{1/2} \log x$. Hence any integrable solution of (3.5) must be of the form $f(x) = C_+x^{\alpha_+-1} + C_-x^{\alpha_--1}$, or $f(x) = C_+x^{-1/2} + C_-x^{-1/2} \log x$ if $\lambda = 1/4$. Any such f satisfies (6.28), and since (6.27) yields $T_{\lambda\kappa}f(1) = 0$, it solves (3.5) if and only if $f(1) = 1$, i.e., if $C_+ + C_- = 1$ ($C_+ = 1$ if $\lambda = 1/4$).

If $0 < \lambda < 1/4$, then $0 < \alpha_- < 1/2 < \alpha_+ < 1$, so the solution $f(x) = x^{\alpha_+-1}$ is in $L^2(0, 1)$ and non-negative; by Corollary 3.6, this is the unique non-negative solution in L^2 , and

$$\chi(\lambda\kappa) = \int_0^1 x^{\alpha_+-1} dx = \frac{1}{\alpha_+} = \frac{2}{1 + \sqrt{1 - 4\lambda}} = \frac{1 - \sqrt{1 - 4\lambda}}{2\lambda}. \quad (6.29)$$

(If we are lucky, or with hindsight, we may observe directly that x^{α_+-1} is a solution of (3.5) by (6.31) below, and apply Corollary 3.6 directly, eliminating most of the analysis above.)

For $\lambda < 1/4$, we have shown that $\chi(\lambda\kappa)$ is finite, so $\lambda\kappa$ is subcritical; thus $\lambda_{\text{cr}} \geq 1/4$. Since the right-hand side in (6.29) has a singularity at $\lambda = 1/4$, Theorem 5.1 shows that $\lambda_{\text{cr}} > 1/4$ is impossible, so we conclude that $\lambda_{\text{cr}} = 1/4$. (Equivalently, $\|T_{\kappa}\| = 4$.) This critical value for the CHKNS model has earlier been found by Callaway, Hopcroft, Kleinberg, Newman and Strogatz [13] by a non-rigorous method, also using (6.29) which they found in a different way; another non-rigorous proof was given by Dorogovtsev, Mendes and Samukhin [16], and the first rigorous proof was given by Durrett [18; 19]. See also Bollobás, Janson and Riordan [4; 5], where different methods were used not involving the susceptibility. The argument above seems to be new.

By Theorem 5.1, we can let $\lambda \nearrow \lambda_{\text{cr}}$ in (6.29), and see that the equation holds for $\lambda = \lambda_{\text{cr}} = 1/4$ too; i.e., $\chi(\lambda_{\text{cr}}\kappa) = 2$.

We see also that in the (sub)critical case $\lambda \leq 1/4$, $\chi(\lambda\kappa; x) = x^{\alpha_+-1}$.

We have no need for the other solutions of (3.5), but note that our analysis shows that for $\lambda < \lambda_{\text{cr}}$, the other non-negative, integrable solutions of (3.5) are given by $x^{\alpha_+-1} + C(x^{\alpha_--1} - x^{\alpha_+-1})$, with $C > 0$. Similarly, although we have no need for the solutions of (3.5) for $\lambda \geq \lambda_{\text{cr}}$, let us note that for the critical case $\lambda = \lambda_{\text{cr}}$, the argument above shows that there is a minimal non-negative solution $x^{-1/2}$, which belongs to L^1 but not to L^2 ; there are further solutions $x^{-1/2} - Cx^{-1/2} \log x$, $C > 0$. For $\lambda > 1/4$, the roots α_{\pm} are complex, and the only real integrable solution to (3.5) is $\frac{1}{2}(x^{\alpha_+-1} + x^{\alpha_--1}) = \text{Re } x^{\alpha_+-1} = x^{-1/2} \cos((\lambda - \frac{1}{4})^{1/2} \log x)$, which oscillates; thus there is no finite non-negative solution at all.

Before proceeding to $\widehat{\chi}$ in the supercritical case, let us calculate ρ_k for small k . We begin by observing, from (6.27), that $T_{\lambda\kappa}1(x) = -\lambda \log x$. Hence (6.1) yields

$$\rho_1(\lambda\kappa; x) = e^{\lambda \log x} = x^\lambda. \quad (6.30)$$

Further, by (6.27), for every non-zero $\gamma > -1$,

$$T_{\lambda\kappa}(x^\gamma) = \frac{\lambda}{\gamma(\gamma+1)}(1-x^\gamma). \quad (6.31)$$

Hence (6.2) yields

$$\rho_2(\lambda\kappa; x) = x^\lambda T_{\lambda\kappa}(x^\lambda) = \frac{1}{1+\lambda}(x^\lambda - x^{2\lambda}). \quad (6.32)$$

Similarly, (6.3) and (6.4) yield

$$\rho_3(\lambda\kappa; x) = \frac{(2+3\lambda)x^{3\lambda} - 4(1+2\lambda)x^{2\lambda} + (2+5\lambda)x^\lambda}{2(1+\lambda)^2(1+2\lambda)}, \quad (6.33)$$

and a formula for $\rho_4(\lambda\kappa; x)$ that we omit, and so on. By integration we then obtain

$$\rho_1(\lambda\kappa) = \frac{1}{1+\lambda}, \quad (6.34)$$

$$\rho_2(\lambda\kappa) = \frac{\lambda}{(1+\lambda)^2(1+2\lambda)}, \quad (6.35)$$

$$\rho_3(\lambda\kappa) = \frac{3\lambda^2}{(1+\lambda)^3(1+2\lambda)(1+3\lambda)}, \quad (6.36)$$

$$\rho_4(\lambda\kappa) = \frac{2\lambda^3(7+15\lambda)}{(1+\lambda)^4(1+2\lambda)^2(1+3\lambda)(1+4\lambda)}. \quad (6.37)$$

It is obvious that each $\rho_k(\lambda\kappa; x)$ is a polynomial in x^λ with coefficients that are rational functions in λ , with only factors $1+j\lambda$, $j=1, \dots, k$, in the denominator. Hence, each $\rho(\lambda\kappa)$ is a rational function of the same type.

There is no obvious general formula for the numbers $\rho_k(\lambda\kappa)$, but, surprisingly, they satisfy a simple quadratic recursion, given in the following theorem. This recursion was found by Callaway, Hopcroft, Kleinberg, Newman and Strogatz [13], using their recursive construction of the graph, see also [19, Chapter 7.1]. (The argument in [13] is non-rigorous, but as pointed out by Durrett [18; 19], it is not hard to make it rigorous.) We give here a proof that instead uses the branching process, which gives more detailed information about the distribution of the ‘locations’ of the components.

Theorem 6.5. *For the CHKNS kernel (6.26), $\rho_k(\lambda\kappa)$ satisfies the recursion*

$$\rho_k(\lambda\kappa) = \frac{k\lambda}{2(1+k\lambda)} \sum_{j=1}^{k-1} \rho_{k-j}(\lambda\kappa)\rho_j(\lambda\kappa), \quad k \geq 2, \quad (6.38)$$

with $\rho_1(\lambda\kappa) = 1/(1 + \lambda)$. Hence, for each $k \geq 1$, $\rho_k(\lambda\kappa)$ is a rational function of λ , with poles only at $-1/j$, $j = 1 \dots, k$.

Moreover, each function $\rho_k(x) = \rho_k(\lambda\kappa; x)$ is a polynomial in x^λ , with coefficients that are rational functions of λ , which can be calculated recursively by

$$x \frac{d}{dx} \rho_k(\lambda\kappa; x) = k\lambda \rho_k(\lambda\kappa; x) - \sum_{j=1}^{k-1} j\lambda \rho_{k-j}(\lambda\kappa) \rho_j(\lambda\kappa; x), \quad k \geq 1, \quad (6.39)$$

together with the boundary conditions $\rho_1(\lambda\kappa; 1) = 1$ and $\rho_k(\lambda\kappa; 1) = 0$, $k \geq 2$.

Proof. Fix $\lambda > 0$. To simplify the notation, throughout this proof we write κ for the kernel so far denoted $\lambda\kappa$. Let $\varepsilon \in (0, 1/2)$, say, and let \mathfrak{X}'_κ be \mathfrak{X}_κ with all points scaled by the factor $(1 - \varepsilon)$; this is the branching process defined by $\mathcal{S}' := (0, 1 - \varepsilon]$, $d\mu' := (1 - \varepsilon)^{-1} dx$ and $\kappa'(x, y) := \lambda \left(\frac{1-\varepsilon}{x \vee y} - 1 \right)$. In \mathfrak{X}'_κ , the offspring process of an individual of type x has intensity

$$\kappa'(x, y) d\mu'(y) = \lambda \left(\frac{1}{x \vee y} - \frac{1}{1 - \varepsilon} \right) dy = \kappa(x, y) dy - \frac{\varepsilon\lambda}{1 - \varepsilon} dy, \quad y \leq 1 - \varepsilon.$$

This is less than the intensity in \mathfrak{X}_κ . We let $\kappa'(x, y) = 0$ if $x > 1 - \varepsilon$ or $y > 1 - \varepsilon$, and define $\kappa''(x, y) = \kappa(x, y) - \kappa'(x, y) \geq 0$. More precisely, for $0 < x \leq 1 - \varepsilon$ and $0 < y \leq 1$,

$$\kappa''(x, y) = \begin{cases} \frac{\varepsilon\lambda}{1 - \varepsilon}, & 0 < y \leq 1 - \varepsilon, \\ \lambda \left(\frac{1}{y} - 1 \right) \leq \frac{\varepsilon\lambda}{1 - \varepsilon}, & 1 - \varepsilon < y \leq 1. \end{cases} \quad (6.40)$$

Thus $\mathfrak{X}_\kappa(x)$ and $\mathfrak{X}'_\kappa(x)$ may be coupled in the natural way so that $\mathfrak{X}'_\kappa(x) \subseteq \mathfrak{X}_\kappa(x)$ in the sense that an individual in $\mathfrak{X}'_\kappa(x)$, of type z say, also belongs to $\mathfrak{X}_\kappa(x)$, and its children in $\mathfrak{X}_\kappa(x)$ are its children in $\mathfrak{X}'_\kappa(x)$ plus some children born according to an independent Poisson process with intensity $\kappa''(z, y) dy$; we call the latter children (if any) *adopted*. An adopted child of type y gets children and further descendants according to a copy of $\mathfrak{X}_\kappa(y)$, independent of everything else. Note that this adoption intensity $\kappa''(x, y)$ is independent of $x \in \mathcal{S}'$, and that the total adoption intensity is $\int_0^1 \kappa''(x, y) dy = \varepsilon\lambda + O(\varepsilon^2)$.

Fix $k \geq 1$. If $|\mathfrak{X}_\kappa(x)| = k$, then either $|\mathfrak{X}'_\kappa(x)| = k$ and there are no adoptions, or $|\mathfrak{X}'_\kappa(x)| = j$ for some $j < k$ and there are one or more adoptions, with a total family size of $k - j$. If $|\mathfrak{X}'_\kappa(x)| = k$, then the probability of some adoption is $k\varepsilon\lambda + O(\varepsilon^2)$, and thus

$$\mathbb{P}(|\mathfrak{X}_\kappa(x)| = k \mid |\mathfrak{X}'_\kappa(x)| = k) = 1 - k\varepsilon\lambda + O(\varepsilon^2). \quad (6.41)$$

Now, suppose that $|\mathfrak{X}'_\kappa(x)| = j < k$. The probability of two or more adoptions is $O(\varepsilon^2)$. Suppose that there is a single adoption. If the adopted child has type y , the probability that this leads to an adopted branch of size $k - j$, and thus to $|\mathfrak{X}_\kappa(x)| = k$, is $\rho_{k-j}(\kappa; y)$. By (6.40), the adoption intensity $\kappa''(z, y)$ is independent of z as remarked above, and is almost uniform on $(0, 1]$; it follows that the probability that $|\mathfrak{X}_\kappa(x)| = k$, given $|\mathfrak{X}'_\kappa(x)| = j$ and that there is a single adoption, by some individual of type z in $\mathfrak{X}'_\kappa(x)$, equals

$$\frac{\int_0^1 \kappa''(z, y) \rho_{k-j}(\kappa; y) dy}{\int_0^1 \kappa''(z, y) dy} = \int_0^1 \rho_{k-j}(\kappa; y) dy + O(\varepsilon) = \rho_{k-j}(\kappa) + O(\varepsilon). \quad (6.42)$$

Since the probability of an adoption at all is $j\varepsilon\lambda + O(\varepsilon^2)$, we obtain

$$\mathbb{P}(|\mathfrak{X}_\kappa(x)| = k \mid |\mathfrak{X}'_\kappa(x)| = j) = j\lambda\rho_{k-j}(\kappa)\varepsilon + O(\varepsilon^2). \quad (6.43)$$

Consequently, for every $k \geq 1$ and $x \in (0, 1 - \varepsilon]$,

$$\rho_k(\kappa; x) = (1 - k\lambda\varepsilon)\rho_k(\kappa'; x) + \sum_{j=1}^{k-1} j\lambda\rho_{k-j}(\kappa)\rho_j(\kappa'; x)\varepsilon + O(\varepsilon^2). \quad (6.44)$$

(The implicit constant in O here and below may depend on k but not on x or ε .) Replace x by $(1 - \varepsilon)x$ and observe that, by definition, $|\mathfrak{X}'_\kappa((1 - \varepsilon)x)| \stackrel{d}{=} |\mathfrak{X}_\kappa(x)|$ and thus $\rho_j(\kappa'; (1 - \varepsilon)x) = \rho_j(\kappa; x)$. This yields

$$\rho_k(\kappa; (1 - \varepsilon)x) = (1 - k\lambda\varepsilon)\rho_k(\kappa; x) + \sum_{j=1}^{k-1} j\lambda\rho_{k-j}(\kappa)\rho_j(\kappa; x)\varepsilon + O(\varepsilon^2). \quad (6.45)$$

Letting $\varepsilon \searrow 0$ we see first that $\rho_k(\kappa; x)$ is Lipschitz continuous in $(0, 1)$, and then that it is differentiable with

$$x \frac{d}{dx} \rho_k(\kappa; x) = k\lambda\rho_k(\kappa; x) - \sum_{j=1}^{k-1} j\lambda\rho_{k-j}(\kappa)\rho_j(\kappa; x), \quad k \geq 1, \quad (6.46)$$

which is (6.39) in the present notation.

For $k = 1$, (6.46) gives $\rho_1(\kappa; x) = Cx^\lambda$, for some constant C . For $x = 1$ we have $\kappa(1, y) = 0$, so the branching process $\mathfrak{X}_\kappa(x)$ dies immediately, and $\rho_1(\kappa; x) = 1$. Thus $\rho_1(\kappa; x) = x^\lambda$ as shown in (6.30). For $k \geq 2$, we note that $x\rho_k(\kappa; x) \rightarrow 0$ as $x \rightarrow 0$ or $x \rightarrow 1$, because $\rho_k(\kappa; x) \leq 1 - \rho_1(\kappa; x) = 1 - x^\lambda$, and thus, integrating by parts,

$$\int_0^1 x \frac{d}{dx} \rho_k(\kappa; x) = [x\rho_k(\kappa; x)]_0^1 - \int_0^1 \rho_k(\kappa; x) dx = 0 - \rho_k(\kappa).$$

Hence, integration of (6.46) yields the recursion formula

$$(1 + k\lambda)\rho_k(\kappa) = \sum_{j=1}^{k-1} j\lambda\rho_{k-j}(\kappa)\rho_j(\kappa), \quad k \geq 2. \quad (6.47)$$

Replacing j by $k - j$ in the right-hand side of (6.47) and summing the two equations, we find that

$$2(1 + k\lambda)\rho_k(\kappa) = \sum_{j=1}^{k-1} (j + k - j)\lambda\rho_{k-j}(\kappa)\rho_j(\kappa), \quad k \geq 2, \quad (6.48)$$

which is (6.38). □

The susceptibility $\widehat{\chi}$ was calculated for all λ by Callaway, Hopcroft, Kleinberg, Newman and Strogatz [13] using the recursion formula (6.38), see also Durrett [18; 19]. We repeat their argument for completeness.

Let $G(z) := \sum_{k=1}^{\infty} \rho_k(\lambda\kappa)z^k$ be the probability generating function of $|\mathfrak{X}_{\lambda\kappa}|$, defined at least for $|z| \leq 1$. Note that in the supercritical case, $|\mathfrak{X}_{\lambda\kappa}|$ is a defective random variable which may be ∞ ; we have $G(1) = 1 - \mathbb{P}(|\mathfrak{X}_{\lambda\kappa}| = \infty) = 1 - \rho(\lambda\kappa)$. Further, $G'(1) = \widehat{\chi}(\lambda\kappa) \leq \infty$.

The recursion (6.38) yields, most easily from the version (6.47),

$$G(z) + \lambda z G'(z) = \lambda z G'(z) G(z) + (1 + \lambda) \rho_1(\lambda\kappa) z = \lambda z G'(z) G(z) + z, \quad (6.49)$$

and thus

$$G'(z) = \frac{z - G(z)}{\lambda z(1 - G(z))}, \quad |z| < 1. \quad (6.50)$$

In the supercritical case, $G(1) < 1$, and we can let $z \nearrow 1$ in (6.50), yielding $\widehat{\chi}(\lambda\kappa) = G'(1) = 1/\lambda$. (In the subcritical case, l'Hôpital's rule, or differentiation of (6.49), yields a quadratic equation for $G'(1)$, with (6.29) as a solution; this is the method by which (6.29) was found in [13].)

Summarizing, we have rigorously verified the explicit formula by Callaway, Hopcroft, Kleinberg, Newman and Strogatz [13]:

$$\widehat{\chi}(\lambda\kappa) = \begin{cases} \frac{1 - \sqrt{1 - 4\lambda}}{2\lambda}, & \lambda \leq \frac{1}{4}, \\ \frac{1}{\lambda}, & \lambda > \frac{1}{4}. \end{cases} \quad (6.51)$$

Note that there is a singularity at $\lambda = 1/4$ with a finite jump from 2 to 4, with infinite derivative on the left side and finite derivative on the right side. It is striking that there is a simple explicit formula for $\widehat{\chi}(\lambda\kappa) = G'(1)$, while no formula is known for $G(1) = 1 - \rho(\lambda\kappa)$. This is presumably related to the fact that $\widehat{\chi}(\lambda\kappa)$ may be found by solving the linear equation (3.5), whereas $\rho(\lambda\kappa)$ is related to the non-linear equation (2.17). As $\lambda = 1/4 + \varepsilon \searrow 1/4$, $\rho(\lambda\kappa)$ approaches 0 extremely rapidly, as $\exp(-(\pi/2\sqrt{2})\varepsilon^{-1/2} + O(\log \varepsilon))$ [16; 5]; the behaviour at the singularity is thus very different for $G(1)$ and $G'(1)$.

Note also that, by (2.11), the discontinuous function $\widehat{\chi}(\lambda\kappa)$ is the pointwise sum of the analytic functions $k\rho_k(\kappa)$.

Remark 6.6. We can obtain higher moments of the distribution $(\rho_k(\lambda\kappa))_{k \geq 1}$ of $|\mathfrak{X}_{\lambda\kappa}|$ by repeatedly differentiating the differential equation (6.50) for its probability generating function and then letting $z \nearrow 1$. In the supercritical case, this yields the moments of $|\mathfrak{X}_{\lambda\kappa}| 1_{\{|\mathfrak{X}_{\lambda\kappa}| < \infty\}}$ (or, equivalently, the moments of $|\mathfrak{X}_{\lambda\kappa}|$ conditioned on $|\mathfrak{X}_{\lambda\kappa}| < \infty$); it follows that all these moments are finite, and we can obtain explicit formulae for them one-by-one. For example, with $\rho = \rho(\lambda\kappa)$,

$$\mathbb{E}(|\mathfrak{X}_{\lambda\kappa}|^2 1_{\{|\mathfrak{X}_{\lambda\kappa}| < \infty\}}) = G''(1) + G'(1) = \frac{1 - \rho}{\lambda\rho} + \frac{1}{\lambda} = \frac{1}{\lambda\rho}, \quad (6.52)$$

$$\mathbb{E}(|\mathfrak{X}_{\lambda\kappa}|^3 1_{\{|\mathfrak{X}_{\lambda\kappa}| < \infty\}}) = G'''(1) + 3G''(1) + G'(1) = \frac{2}{\lambda^2\rho^2} + \frac{1}{\lambda\rho}. \quad (6.53)$$

It can be seen that for each $m \geq 1$, as $\lambda \searrow \lambda_{\text{cr}}$, and thus $\rho \rightarrow 0$, we have

$$\mathbb{E}(|\mathfrak{X}_{\lambda\kappa}|^m 1_{\{|\mathfrak{X}_{\lambda\kappa}| < \infty\}}) \sim c_m \rho^{1-m} \quad (6.54)$$

for some constant $c_m > 0$; we do not know any general formula for c_m . For any $\lambda > \lambda_{\text{cr}} = \frac{1}{4}$ and $a, b > 0$, writing $\hat{X} := |\mathfrak{X}_{\lambda\kappa}| 1_{\{|\mathfrak{X}_{\lambda\kappa}| < \infty\}}$, from (6.51) and (6.52)–(6.53) we obtain

$$\begin{aligned} \mathbb{E}\left(\hat{X}^2 1_{\{\hat{X} \leq a/\rho\}}\right) &\leq \frac{a}{\rho} \mathbb{E} \hat{X} = \frac{a}{\rho} \hat{\chi}(\lambda\kappa) = \frac{a}{\lambda\rho}, \\ \mathbb{E}\left(\hat{X}^2 1_{\{\hat{X} \geq b/\rho\}}\right) &\leq \frac{\rho}{b} \mathbb{E} \hat{X}^3 = \frac{2}{b\lambda^2\rho} + \frac{1}{b\lambda}, \end{aligned}$$

and hence

$$\mathbb{E}\left(\hat{X}^2 1_{\{a/\rho \leq \hat{X} \leq b/\rho\}}\right) \geq \frac{1}{\lambda\rho} - \frac{a}{\lambda\rho} - \frac{2}{b\lambda^2\rho} - \frac{1}{b\lambda} = \frac{1}{\lambda\rho} \left(1 - a - \frac{2}{b\lambda} - \frac{\rho}{b}\right).$$

Choosing, for example, $a = 1/4$ and $b = 32$, so $b\lambda > 8$, the last quantity is at least $1/(3\lambda\rho) > 1.3/\rho$ if λ is close to λ_{cr} , and thus, for such λ at least,

$$\mathbb{P}\left(\frac{1}{4\rho} \leq |\mathfrak{X}_{\lambda\kappa}| \leq \frac{32}{\rho}\right) \geq \frac{1.3}{\rho} \left(\frac{\rho}{b}\right)^2 > \frac{\rho}{1000}.$$

Hence, $|\mathfrak{X}_{\lambda\kappa}|$ may be as large as about ρ^{-1} with probability about ρ , as suggested by (6.54).

Note that each $\rho_k(\lambda\kappa)$ is a continuous function of λ , so as $\lambda \searrow \lambda_{\text{cr}}$, the (defective) distribution of $|\mathfrak{X}_{\lambda\kappa}|$ converges to the distribution of the critical $|\mathfrak{X}_{\lambda_{\text{cr}}\kappa}|$, which has mean $\chi(\lambda_{\text{cr}}\kappa) = 2$ and $\mathbb{P}(|\mathfrak{X}_{\lambda_{\text{cr}}\kappa}| = k) \sim 2/(k^2 \log k)$ as $k \rightarrow \infty$, see [19, Section 7.3].

In the subcritical case, $\rho_k(\lambda\kappa)$ decreases as a power of k , see [19, Section 7.3] for details.

We have so far studied $\chi(\lambda\kappa)$ and $\hat{\chi}(\lambda\kappa)$, or, equivalently, the cluster size in the branching process $\mathfrak{X}_{\lambda\kappa}$. Let us now return to the random graphs; we then have to be careful with the precise definitions. The Poisson version of the CHKNS model mentioned above can be described as the random multigraph where the number of edges between vertices i and j is $\text{Po}(\lambda_{ij})$ with intensity $\lambda_{ij} := \lambda(1/(j-1) - 1/n)$, for $1 \leq i < j \leq n$, independently for all such pairs i, j , see [18; 19; 5]. For the moment, let us call this random graph G_n^I . Let G_n^{II} be defined similarly, but with $\lambda_{ij} := \lambda(1/j - 1/n)$, and let G_n^{III} be defined similarly with $\lambda_{ij} := \lambda(1/j - 1/(n+1))$, for $1 \leq i < j \leq n$. Since multiple edges do not matter for the components, we may as well consider the corresponding simple graphs with multiple edges coalesced; then the probability of an edge between i and j , $i < j$, is $p_{ij} := 1 - \exp(-\lambda_{ij})$. (If, for simplicity, we consider $\lambda \leq 1$ only, it is easy to see that the results below hold also if we instead let the edges appear with probabilities $p_{ij} = \lambda_{ij}$; this follows by the same arguments or by contiguity and [22, Corollary 2.12(iii)].)

We first consider G_n^{II} ; note that this is exactly (the Poisson version of) $G^{\mathcal{V}}(n, \lambda\kappa)$ with κ defined in (6.26) and the vertex space \mathcal{V} given by $\mathcal{S} = (0, 1]$ with μ Lebesgue measure

as above, and the deterministic sequence $\mathbf{x}_n = (x_1, \dots, x_n)$ with $x_i = i/n$. Arguing as in the proof of Theorem 4.7, summing over distinct indices only, and using the fact that κ is non-increasing in each variable, we find that the expected number $\mathbb{E} P_\ell(G_n^{II})$ of paths of length ℓ is

$$\begin{aligned} \mathbb{E} P_\ell(G_n^{II}) &\leq \sum_{j_0, \dots, j_\ell=1}^n \prod_{i=1}^{\ell} \frac{\lambda \kappa(j_{i-1}, j_i)}{n} \\ &\leq \sum_{j_0, \dots, j_\ell=1}^n n \int_{\prod_i ((j_{i-1}-1)/n, j_i/n]} \prod_{i=1}^{\ell} \lambda \kappa(x_{i-1}, x_i) dx_0 \cdots dx_\ell \\ &\leq n \int_{\mathcal{S}^{\ell+1}} \prod_{i=1}^{\ell} \lambda \kappa(x_{i-1}, x_i) dx_0 \cdots dx_\ell = n \langle T_{\lambda \kappa}^\ell 1, 1 \rangle. \end{aligned}$$

Hence Lemmas 4.6 and 4.3 imply that (4.5) holds and $\chi(G_n^{II}) \xrightarrow{p} \chi(\lambda \kappa)$.

For G_n^{III} , we observe that G_n^{III} can be seen as an induced subgraph of G_{n+1}^{II} , and thus

$$\mathbb{E} \sum_{\ell} P_\ell(G_n^{III}) \leq \mathbb{E} \sum_{\ell} P_\ell(G_{n+1}^{II}) \leq (n+1) \chi(\lambda \kappa). \quad (6.55)$$

Hence Lemma 4.3 implies that $\chi(G_n^{III}) \xrightarrow{p} \chi(\lambda \kappa)$.

Finally, it is easily checked that G_n^I and G_n^{III} satisfy the conditions of [22, Corollary 2.12(iii)], and thus are contiguous. Hence $\chi(G_n^I) \xrightarrow{p} \chi(\lambda \kappa)$ too. (One can also compare G_n^I and G_n^{III} as in [4, Lemma 11].)

It turns out that in probability bounds such as the one we have just proved do not obviously transfer from G_n^I to the original CHKNS model. On the other hand (as we shall see below), bounds on the expected number of paths do. Hence, in order to analyze the original CHKNS model, we shall need to show that

$$\limsup \mathbb{E} n^{-1} \sum_{\ell} P_\ell(G_n^I) \leq \chi(\lambda \kappa). \quad (6.56)$$

If $\lambda > 1/4$, then $\lambda \kappa$ supercritical, so $\chi(\lambda \kappa) = \infty$ and there is nothing to prove. Suppose then that $\lambda \leq 1/4$. We may regard G_n^I with the vertex 1 deleted as G_{n-1}^{III} . Writing $P(G)$ for the total number of paths in a graph G , and P^* for the number involving the vertex 1, by (6.55) we thus have

$$\mathbb{E} P(G_n^I) - \mathbb{E} P^*(G_n^I) = \mathbb{E} P(G_{n-1}^{III}) \leq n \chi(\lambda \kappa),$$

so to prove (6.56) it suffices to show that $\mathbb{E} P^*(G_n^I) = o(n)$.

Let $S(G_n^I)$ denote the number of paths in G_n^I starting at vertex 1. Since a path visiting vertex 1 may be viewed as the edge disjoint union of two paths starting there, and edges are present independently, we have $\mathbb{E} P^*(G_n^I) \leq (\mathbb{E} S(G_n^I))^2$. Now $\mathbb{E} S(G_n^I)$ is given by 1

plus the sum over i of $1/i$ times the expected number of paths in G_{n-1}^{III} starting at vertex i . Durrett [18, Theorem 6] proved the upper bound

$$\frac{3}{8} \frac{1}{\sqrt{ij}} \frac{(\log i + 2)(\log n - \log j + 2)}{\log n + 4}$$

on the expected number of paths between vertices i and j in the graph H on $[n]$ in which edges are present independently and the probability of an edge ij , $i < j$, is $1/(4j)$ (a form of Dubins' model; see the next section). In fact, his result is stated for the probability that a path is present, but the proof bounds the expected number of paths. (The factor $1/\sqrt{ij}$ is omitted in [18, Theorem 6]; this is simply a typographical error.) This bound carries over to G_{n-1}^{III} , which we may regard as a subgraph of H . Multiplying by $1/i$ and summing, a little calculation shows that this bound implies that $\mathbb{E} S(G_n^I) = O(n^{1/2}/\log n)$ for $\lambda = 1/4$, and hence for any $\lambda \leq 1/4$. From the comments above, (6.56) follows, and for any $\lambda > 0$ we have $\chi(G_n^I) \xrightarrow{P} \chi(\lambda\kappa)$.

Recall that the original CHKNS model G_n has the same expected edge densities as G_n^I , but the mode of addition is slightly different, with 0 or 1 edges added at each step, rather than a Poisson number; this introduces some dependence between edges. However, as noted in [4], the form of this dependence is such that conditioning on a certain set of edges being present can only reduce the probability that another given edge is present. Thus, any given path is at most as likely in G_n as in G_n^I , and (6.56) carries over to the CHKNS model. On the other hand, the effect of this dependence is small except for the first few vertices, and it is easy to see that $N_k(G_n)$ has almost the same distribution as $N_k(G_n^I)$. In particular, $N_k(G_n)/n \xrightarrow{P} \rho_k(\lambda\kappa)$, so the proof of Theorem 4.1 goes through. Using Lemma 4.2 it follows that $\chi(G_n) \xrightarrow{P} \chi(\lambda\kappa)$.

Turning to the supercritical case, let $M_k(G)$ denote the number of components of a graph G , other than \mathcal{C}_1 , that have order k . We claim that, in all variants G_n^I , G_n^{II} , G_n^{III} or the original CHKNS model, for fixed $\lambda > \lambda_{cr}$ there is some sequence of events \mathcal{E}_n that holds whp, and some $\eta > 0$ such that

$$n^{-1} \mathbb{E}(M_k(G_n) \mid \mathcal{E}_n) \leq 100e^{-\eta k^{1/5}}, \tag{6.57}$$

say, for all $n, k \geq 1$. Suppose for the moment that (6.57) holds. Then

$$\mathbb{E}(\widehat{\chi}(G_n) \mid \mathcal{E}_n) = n^{-1} \sum_{k \geq 1} k^2 \mathbb{E}(M_k(G_n) \mid \mathcal{E}_n) \leq \sum_k 100k^2 e^{-\eta k^{1/5}} < \infty.$$

For each fixed k we have $n^{-1} \mathbb{E} k^2 M_k(G_n) = n^{-1} \mathbb{E}(kN_k(G_n) - O(k)) \rightarrow k\rho_k(\lambda\kappa)$. Since \mathcal{E}_n holds whp and $n^{-1}k^2 M_k(G_n)$ is bounded it follows that $n^{-1}k^2 \mathbb{E}(M_k(G_n) \mid \mathcal{E}_n) \rightarrow k\rho_k(\lambda\kappa)$. Hence, by dominated convergence, $\mathbb{E}(\widehat{\chi}(G_n) \mid \mathcal{E}_n) \rightarrow \sum k\rho_k(\lambda\kappa) = \widehat{\chi}(\lambda\kappa)$, and (which we know already in this case), $\widehat{\chi}(\lambda\kappa)$ is finite. By Lemma 4.4(ii), it then follows that $\widehat{\chi}(G_n) \xrightarrow{P} \widehat{\chi}(\lambda\kappa)$.

To prove (6.57) we use an idea from [4]; with an eye to the next subsection, in the proof we shall not rely on the exact values of the edge probabilities, only on certain bounds.

Fix $\lambda > \lambda_{\text{cr}}$. Choosing η small, in proving (6.57) we may and shall assume that k is at least some constant that may depend on λ . Set $\delta = k^{-1/100}$, and let G'_n be the subgraph of G_n induced by the first $n' = (1 - \delta)n$ vertices. (We ignore the irrelevant rounding to integers.) In all variants $G'_n, G_n^I, G_n^{II}, G_n^{III}$, the distribution of G'_n stochastically dominates that of $G_{n'}$, so whp G'_n contains a component C of order at least $3\rho(\lambda\kappa)n'/4 \geq \rho(\lambda\kappa)n/2$. Let us condition on G'_n , assuming that this holds. Note that whp the largest component of G_n will contain C , so it suffices to bound the expectation of M'_k , the number of k -vertex components of G_n not containing C . To adapt what follows to the original CHKNS model, we should instead condition on the edges added by time n' as the graph grows; we omit the details.

Suppose that C' is a component of G'_n other than C . Consider some vertex v , $n' < v \leq (1 - \delta/2)n$. Then v has probability at least $\lambda(1/v - 1/n) \geq \lambda\delta/(2n) \geq \delta/(8n)$ of being joined to any given vertex, and hence probability at least $\delta|S|/(9n)$ of having at least one neighbour in any given set S of vertices. Hence with probability at least $\delta^2\rho(\lambda\kappa)|C'|/(200n)$, v has neighbours in C and in C' . Since these events are independent for different v , the probability that C' is not part of the same component of G_n as C is at most

$$(1 - \delta^2\rho(\lambda\kappa)|C'|/(200n))^{\delta n/2} \leq \exp(-\delta^3\rho(\lambda\kappa)|C'|/400) = \exp(-a\delta^3|C'|),$$

for some $a > 0$ independent of k .

Let A be the number of components of G'_n of size at least $k^{1/4}$ that are not joined to C in G_n . Then it follows that $\mathbb{E}A \leq ne^{-ak^{1/5}}$.

For any $v \leq n'$, the expected number of edges from ‘late’ vertices $w > n'$ to v is at most $1/2$, say. (We may assume δ is small if λ is large.) Let B be the number of vertices receiving at least $k^{1/4}$ edges from late vertices. Then it is easy to check (using a Chernoff bound or directly) that $\mathbb{E}B \leq ne^{-bk^{1/4}}$ for some $b > 0$. The subgraph of G_n induced by the late vertices is dominated by an Erdős–Rényi random graph with average degree at most $1/2$. Let N be the number of components of this subgraph with size at least $k^{1/4}$. Then, since the component exploration process is dominated by a subcritical branching process, we have $\mathbb{E}N \leq ne^{-ck^{1/4}}$ for some $c > 0$.

Let M''_k be the number of k -vertex components of G_n other than that containing C that do not contain any of the components/vertices counted by A, B or N . Since $\mathbb{E}(M'_k - M''_k) \leq \mathbb{E}(A + B + N) \leq ne^{-dk^{1/5}}$ for some $d > 0$, it suffices to bound $\mathbb{E}M''_k$. Condition on G'_n and explore from some vertex not in C . To uncover a component counted by M''_k , this exploration must cross from late to early vertices at least $k^{1/4}$ times – each time we reach a component of size at most $k^{1/4}$, and from each of these vertices we get back to at most $k^{1/4}$ late vertices, and from each of those to at most $k^{1/4}$ other late vertices before we next cross over to early vertices. However, every time we find an edge from a late to an early vertex (conditioning on the presence of such an edge but not its destination early vertex), we have probability at least $\rho(\lambda\kappa)/2$ of hitting C . It follows that $\mathbb{E}M''_k \leq n(1 - \rho(\lambda\kappa)/2)^{k^{1/4}}$, and (6.57) follows.

Note that since $\widehat{\chi}(\lambda\kappa)$ is a discontinuous function of λ , we cannot obtain convergence to $\widehat{\chi}(\lambda\kappa)$ for an arbitrary sequence $\lambda_n \rightarrow \lambda$, as in Theorem 4.7 and Section 6.1. In fact, it

follows easily from Theorem 4.1 that if $\lambda_n \searrow \lambda_{\text{cr}}$ slowly enough, then $\chi(G^{\mathcal{V}}(n, \lambda_n \kappa)) \xrightarrow{\text{P}} \infty > \chi(\lambda_{\text{cr}} \kappa)$ and $\widehat{\chi}(G^{\mathcal{V}}(n, \lambda_n \kappa)) > \lim_{\lambda \searrow \lambda_{\text{cr}}} \widehat{\chi}(\lambda \kappa) - \varepsilon = 4 - \varepsilon > \widehat{\chi}(\lambda_{\text{cr}} \kappa)$ whp for every $\varepsilon \in (0, 2)$, for any vertex space \mathcal{V} (with \mathcal{S} and μ as above), and thus in particular for G_n^{II} .

6.4 Dubins' model

A random graph closely related to the CHKNS model is the graph $G^{\mathcal{V}}(n, \lambda \kappa)$ with kernel

$$\kappa(x, y) := \frac{1}{x \vee y} \tag{6.58}$$

on $\mathcal{S} = (0, 1]$, where the vertex space \mathcal{V} is as in Section 6.3, so $\mathbf{x}_n = (x_1, \dots, x_n)$. In this case, the probability p_{ij} of an edge between i and j is given (for $\lambda \leq 1$) by $p_{ij} = \lambda \kappa(i/n, j/n)/n = \lambda/(i \vee j)$. Note that this is independent of n , so we may regard $G^{\mathcal{V}}(n, \lambda \kappa)$ as an induced subgraph of an infinite random graph with vertex set \mathbb{N} and these edge probabilities, with independent edges.

This infinite random graph was introduced by Dubins, who asked when it is a.s. connected. Shepp [32] proved that this holds if and only if $\lambda > 1/4$. The finite random graph $G^{\mathcal{V}}(n, \lambda \kappa)$ was studied by Durrett [18, 19], who showed that $\lambda_{\text{cr}} = 1/4$; thus the critical value for the emergence of a giant component in the finite version coincides with the critical value for connectedness of the infinite version. See also [4; 31; 5].

We have

$$T_{\lambda \kappa} f(x) = \frac{\lambda}{x} \int_0^x f(y) dy + \lambda \int_x^1 \frac{f(y)}{y} dy. \tag{6.59}$$

We can solve (3.5) as in Section 6.3; we get the same equation (6.28) and thus the same solutions $f(x) = C_+ x^{\alpha_+ - 1} + C_- x^{\alpha_- - 1}$ (unless $\lambda = 1/4$ when we also get a logarithmic term), and substitution into (6.59) shows that this is a solution of (3.5) if and only if $C_+ \alpha_+ + C_- \alpha_- = 1$, see (6.62) below. If $0 < \lambda < 1/4$, so $\alpha_+ > 1/2 > \alpha_-$, there is thus a positive solution $f(x) = \alpha_+^{-1} x^{\alpha_+ - 1}$ in L^2 . (This is the unique solution in L^2 , by a direct check or by Corollary 3.6.) Hence, Corollary 3.6 yields

$$\chi(\lambda \kappa) = \int_0^1 f(x) dx = \alpha_+^{-2} = \frac{1 - 2\lambda - \sqrt{1 - 4\lambda}}{2\lambda^2}, \quad 0 < \lambda < 1/4. \tag{6.60}$$

Since this function is analytic on $(0, 1/4)$ but has a singularity at $\lambda = 1/4$ (although it remains finite there), Theorem 5.1 shows that $\lambda_{\text{cr}} = 1/4$, which gives a new proof of this result by Durrett [18]. Note that $\chi(\lambda_{\text{cr}} \kappa) = 4$ is finite.

We can estimate the expected number of paths as in Section 6.3, and show by Lemmas 4.6 and 4.3 that $\chi(G^{\mathcal{V}}(n, \lambda \kappa)) \xrightarrow{\text{P}} \chi(\lambda \kappa)$ for any $\lambda > 0$.

In the supercritical case, the tail bound (6.57) goes through, showing that for any $\lambda > \lambda_{\text{cr}}$ we have $\widehat{\chi}(\lambda \kappa) < \infty$, and $\widehat{\chi}(G^{\mathcal{V}}(n, \lambda \kappa)) \xrightarrow{\text{P}} \widehat{\chi}(\lambda \kappa)$. Unfortunately, while the argument gives a tail bound on the sum $\sum_k k \rho_k(\lambda \kappa)$ for each fixed $\lambda > \lambda_{\text{cr}}$, the dependence on λ is rather bad, so it does not seem to tell us anything about the behaviour of $\widehat{\chi}(\lambda \kappa)$ as λ approaches the critical point.

We can easily calculate ρ_k for small k . First, by (6.59), $T_{\lambda\kappa}1(x) = \lambda - \lambda \log x$. Hence (6.1) yields

$$\rho_1(\lambda\kappa; x) = e^{-\lambda + \lambda \log x} = e^{-\lambda} x^\lambda. \quad (6.61)$$

Further, instead of (6.31) we now have, for every non-zero $\gamma > -1$,

$$T_{\lambda\kappa}(x^\gamma) = \frac{\lambda}{\gamma} - \frac{\lambda}{\gamma(\gamma+1)} x^\gamma. \quad (6.62)$$

Hence (6.2) yields

$$\rho_2(\lambda\kappa; x) = e^{-\lambda} x^\lambda T_{\lambda\kappa}(e^{-\lambda} x^\lambda) = e^{-2\lambda} x^\lambda \left(1 - \frac{x^\lambda}{\lambda+1}\right). \quad (6.63)$$

Similarly, by (6.3) and some calculations,

$$\rho_3(\lambda\kappa; x) = \frac{e^{-3\lambda}}{2(1+\lambda)^2(1+2\lambda)} \left((2+3\lambda)x^{3\lambda} - 4(1+2\lambda)(1+\lambda)x^{2\lambda} + (2+3\lambda)(1+2\lambda)(1+\lambda)x^\lambda \right),$$

and so on. By integration we then obtain

$$\rho_1(\lambda\kappa) = \frac{e^{-\lambda}}{1+\lambda}, \quad (6.64)$$

$$\rho_2(\lambda\kappa) = \frac{2\lambda e^{-2\lambda}}{(1+\lambda)(1+2\lambda)}, \quad (6.65)$$

$$\rho_3(\lambda\kappa) = \frac{(15\lambda^2 + 18\lambda^3)e^{-3\lambda}}{2(1+\lambda)^2(1+2\lambda)(1+3\lambda)}. \quad (6.66)$$

It is clear that each $\rho_k(\lambda\kappa)$ is $e^{-k\lambda}$ times a rational function of λ , but we do not know any general formula or a recursion that enables us to calculate $\widehat{\chi}(\lambda\kappa)$ in the supercritical case as in Section 6.3.

6.5 Functions of $\max\{x, y\}$

The examples in Sections 6.3 and 6.4 are both of the type $\kappa(x, y) = \varphi(x \vee y)$ for some function φ on $(0, 1]$. It is known that if, for example, $\varphi(x) = O(1/x)$, then T_κ is bounded on L^2 , and thus there exists a positive $\lambda_{\text{cr}} > 0$; see [27; 1] and [5, Section 16.6].

We have

$$T_{\lambda\kappa}f(x) = \lambda\varphi(x) \int_0^x f(y) dy + \lambda \int_x^1 \varphi(y) f(y) dy. \quad (6.67)$$

If $\varphi \in C^1(0, 1]$, then any integrable solution of (3.5) must be in $C^1(0, 1]$ too, and differentiation yields $f' = \lambda\varphi'F$, where $F(x) := \int_0^x f(y) dy$ is the primitive function of f ;

furthermore, we have $f(1) = 1 + T_{\lambda\kappa}f(1) = 1 + \lambda\varphi(1)F(1)$. Hence, solving (3.5) is equivalent to solving the Sturm–Liouville problem

$$F''(x) = \lambda\varphi'(x)F(x) \tag{6.68}$$

with the boundary conditions

$$F(0) = 0 \quad \text{and} \quad F'(1) = \lambda\varphi(1)F(1) + 1. \tag{6.69}$$

If there is a solution to (6.68) and (6.69) with $F' \geq 0$ and $F' \in L^2$, then Corollary 3.6 shows that

$$\chi(\lambda\kappa) = \int_0^1 F'(x) \, dx = F(1). \tag{6.70}$$

The examples in Sections 6.3 and 6.4 are examples of this, as is the Erdős–Rényi case in Section 6.1 ($\varphi = 1$). We consider one more simple explicit example.

Example 6.7. Let $\varphi(x) = 1 - x$. Then (6.68) becomes $F'' = -\lambda F$, with the solution, using (6.69), $F(x) = A \sin(\sqrt{\lambda}x)$ with $A\sqrt{\lambda} \cos(\sqrt{\lambda}) = 1$. This solution satisfies $F' \geq 0$ if $\sqrt{\lambda} < \pi/2$, so we find $\lambda_{\text{cr}} = \pi^2/4$ and, by (6.70),

$$\chi(\lambda\kappa) = \frac{\tan(\sqrt{\lambda})}{\sqrt{\lambda}}, \quad \lambda < \lambda_{\text{cr}} = \pi^2/4. \tag{6.71}$$

6.6 Further examples

We give also a couple of counterexamples.

Example 6.8. Let $\mathcal{S} = \{1, 2\}$, with $\mu\{1\} = \mu\{2\} = 1/2$, and let $\kappa_\varepsilon(1, 1) = 2$, $\kappa_\varepsilon(2, 2) = 1$ and $\kappa_\varepsilon(1, 2) = \kappa_\varepsilon(2, 1) = \varepsilon$ for $\varepsilon \geq 0$.

For $\varepsilon = 0$, κ_0 is reducible; given the numbers n_1 and n_2 of vertices of the two types, the random graph $G^\nu(n, \lambda\kappa_0)$ consists of two disjoint independent random graphs $G(n_1, 2\lambda/n)$ and $G(n_2, \lambda/n)$; since $n_1/n, n_2/n \xrightarrow{P} 1/2$, the first part has a threshold at $\lambda = 1$ and the second a threshold at $\lambda = 2$. Similarly, the branching process $\mathfrak{X}_{\lambda\kappa_0}(x)$ is a single-type Galton–Watson process with offspring distribution $\text{Po}(\lambda)$ if $x = 1$ and $\text{Po}(\lambda/2)$ if $x = 2$, so $\mathfrak{X}_{\lambda\kappa_0}$ is a mixture of these. Hence, if $\widehat{\chi}_1(\lambda)$ denotes the (modified) susceptibility in the Erdős–Rényi case, given by (6.5) for $\lambda < 1$ and (6.6) for $\lambda \geq 1$, then

$$\widehat{\chi}(\lambda\kappa_0) = \frac{1}{2}\widehat{\chi}_1(\lambda) + \frac{1}{2}\widehat{\chi}_1(\lambda/2), \tag{6.72}$$

so $\widehat{\chi}(\lambda\kappa_0)$ has two singularities, at $\lambda = 1$ and $\lambda = 2$. Clearly, $\lambda_{\text{cr}} = 1$.

Now consider $\varepsilon > 0$ and let $\varepsilon \searrow 0$. Then $\lambda_{\text{cr}}(\kappa_\varepsilon) \leq \lambda_{\text{cr}}(\kappa_0) = 1$. Furthermore, for any fixed λ , $\rho(\lambda\kappa_\varepsilon; x) \rightarrow \rho(\lambda\kappa_0; x)$ by [5, Theorem 6.4(ii)], and hence $T_{\lambda\kappa_\varepsilon} \rightarrow T_{\lambda\kappa_0}$ (we may regard the operators as 2×2 matrices). Consequently, if $\lambda > 1$ with $\lambda \neq 2$ and thus $\|T_{\lambda\kappa_0}\| < 1$, then $(I - T_{\lambda\kappa_\varepsilon})^{-1} \rightarrow (I - T_{\lambda\kappa_0})^{-1}$, and thus $\widehat{\chi}(\lambda\kappa_\varepsilon) \rightarrow \widehat{\chi}(\lambda\kappa_0)$ by Theorem 3.3. This holds for $\lambda = 2$ also, with the limit $\widehat{\chi}(2\kappa_0) = \infty$, for example by (3.4) and Fatou’s lemma.

Since $\widehat{\chi}(\lambda\kappa_0)$ has singularities both at 1 and 2, we may choose $\delta \in (0, 1/2)$ such that $\widehat{\chi}((1 + \delta)\kappa_0) > \widehat{\chi}(\frac{3}{2}\kappa_0)$ and $\widehat{\chi}((2 - \delta)\kappa_0) > \widehat{\chi}(\frac{3}{2}\kappa_0)$, and then choose $\varepsilon > 0$ such that $\widehat{\chi}((1 + \delta)\kappa_\varepsilon) > \widehat{\chi}(\frac{3}{2}\kappa_\varepsilon)$ and $\widehat{\chi}((2 - \delta)\kappa_\varepsilon) > \widehat{\chi}(\frac{3}{2}\kappa_\varepsilon)$. This yields an example of an irreducible kernel κ such that $\widehat{\chi}(\lambda\kappa)$ is not monotone decreasing on $(\lambda_{\text{cr}}, \infty)$.

Example 6.9. Theorem 4.7 shows convergence of $\chi(G^\mathcal{V}(n, \kappa))$ to $\chi(\kappa)$ for any vertex space \mathcal{V} when κ is bounded. For unbounded κ , some restriction on the vertex space is necessary. (Cf. Theorem 4.8 with a very strong condition on \mathcal{V} and none on κ .) The reason is that our conditions on \mathcal{V} are weak and do not notice sets of vertices of order $o(n)$, but such sets can mess up χ .

In fact, assume that κ is unbounded. For each $n \geq 16$, find $(a_n, b_n) \in \mathcal{S}^2$ with $\kappa(a_n, b_n) > n$. Define \mathbf{x}_n by taking $\lfloor n^{3/4} \rfloor$ points $x_i = a_n$, $\lfloor n^{3/4} \rfloor$ points $x_i = b_n$, and the remaining $n - 2\lfloor n^{3/4} \rfloor$ points i.i.d. with distribution μ . It is easily seen that this yields a vertex space \mathcal{V} , and that we have created a component with at least $2\lfloor n^{3/4} \rfloor$ vertices. Consequently, $|\mathcal{C}_1| > n^{3/4}$, and by (1.2), $\chi(G^\mathcal{V}(n, \kappa)) \geq |\mathcal{C}_1|^2/n > n^{1/2}$, so $\chi(G^\mathcal{V}(n, \kappa)) \rightarrow \infty$, even if κ is subcritical and thus $\chi(\kappa) < \infty$.

Using a similar construction (but this time for more specific kernels κ), it is easy to give examples of unbounded supercritical κ with $\widehat{\chi}(\kappa) < \infty$ but $\widehat{\chi}(G^\mathcal{V}(n, \kappa)) \rightarrow \infty$ for suitable vertex spaces \mathcal{V} .

References

- [1] A.B. Aleksandrov, S. Janson, V.V. Peller & R. Rochberg, An interesting class of operators with unusual Schatten–von Neumann behavior. *Function Spaces, Interpolation Theory and Related Topics (Proceedings of the International Conference in honour of Jaak Peetre on his 65th birthday, Lund 2000)*, eds. M. Cwikel, M. Englis, A. Kufner, L.-E. Persson & G. Sparr, Walter de Gruyter, Berlin, 2002, pp. 61–150.
- [2] B. Bollobás, *Modern Graph Theory*, Graduate Texts in Mathematics **184**, Springer, New York, 1998, xiv + 394 pp.
- [3] B. Bollobás, C. Borgs, J. Chayes and O. Riordan, Percolation on dense graph sequences. *Annals of Probability* **38** (2010), 150–183.
- [4] B. Bollobás, S. Janson and O. Riordan, The phase transition in the uniformly grown random graph has infinite order. *Random Struct. Alg.* **26** (2005), 1–36.
- [5] B. Bollobás, S. Janson and O. Riordan, The phase transition in inhomogeneous random graphs. *Random Struct. Alg.* **31** (2007), 3–122.
- [6] B. Bollobás, S. Janson and O. Riordan, The cut metric, random graphs, and branching processes. *J. Statistical Physics* **140** (2010), 289–335.
- [7] B. Bollobás, S. Janson and O. Riordan, Sparse random graphs with clustering. *Random Struct. Alg.* **38** (2011), 269–323.
- [8] B. Bollobás O. Riordan, *Percolation*. Cambridge University Press, Cambridge, 2006, x + 323 pp.

- [9] B. Bollobás and O. Riordan, Metrics for sparse graphs. In *Surveys in Combinatorics 2009, London Math. Soc. Lecture Note Series* **365**, S. Huczynska, J.D. Mitchell and C.M. Roney-Dougal eds, CUP (2009), pp. 211–287.
- [10] É. Borel, Sur l’emploi du théorème de Bernoulli pour faciliter le calcul d’une infinité de coefficients. Application au problème de l’attente à un guichet. *C. R. Acad. Sci. Paris* **214** (1942), 452–456.
- [11] C. Borgs, J.T. Chayes, L. Lovász, V.T. Sós and K. Vesztergombi, Convergent sequences of dense graphs I: Subgraph frequencies, metric properties and testing. *Advances in Math.* **219** (2008), 1801–1851.
- [12] C. Borgs, J.T. Chayes, L. Lovász, V.T. Sós and K. Vesztergombi: Convergent sequences of dense graphs II. Multiway cuts and statistical physics. Preprint (2007). <http://research.microsoft.com/~borgs/Papers/ConRight.pdf>.
- [13] D.S. Callaway, J.E. Hopcroft, J.M. Kleinberg, M.E.J. Newman & S.H. Strogatz, Are randomly grown graphs really random? *Phys. Rev. E* **64** (2001), 041902.
- [14] L. Chayes and E.A. Smith, Large deviation analysis for layered percolation problems on the complete graph, to appear in *Random Struct. Alg.* DOI: 10.1002/rsa.20387
- [15] J. B. Conway, *A Course in Functional Analysis*. 2nd ed., Springer-Verlag, New York, 1990.
- [16] S.N. Dorogovtsev, J.F.F. Mendes & A.N. Samukhin, Anomalous percolation properties of growing networks. *Phys. Rev. E* **64** (2001), 066110.
- [17] N. Dunford & J.T. Schwartz, *Linear Operators. I. General Theory*. Interscience Publishers, New York, 1958. xiv+858 pp.
- [18] R. Durrett, Rigorous result for the CHKNS random graph model. *Proceedings, Discrete Random Walks 2003 (Paris, 2003)*, eds. C. Banderier & Chr. Krattenthaler, Discrete Mathematics and Theoretical Computer Science **AC** (2003), pp. 95–104, <http://www.dmtcs.org/proceedings/html/dmAC0109.abs.html>
- [19] R. Durrett, *Random Graph Dynamics*. Cambridge University Press, Cambridge, 2007.
- [20] M. Dwass, The total progeny in a branching process and a related random walk. *J. Appl. Probab.* **6** (1969), 682–686.
- [21] A. Frieze and R. Kannan, Quick approximation to matrices and applications. *Combinatorica* **19** (1999), 175–220.
- [22] S. Janson, Asymptotic equivalence and contiguity of some random graphs. *Random Struct. Alg.* **36** (2010), 26–45.
- [23] S. Janson, D.E. Knuth, T. Łuczak & B. Pittel, The birth of the giant component. *Random Struct. Alg.* **3** (1993), 233–358.
- [24] S. Janson & M. Łuczak, Susceptibility in subcritical random graphs. *J. Math. Phys.* **49** (2008), 125207, 23pp.
- [25] S. Janson & O. Riordan, Duality in inhomogeneous random graphs, and the cut metric. *Random Struct. Alg.* **39** (2011), 399–411.

- [26] T. Kato, *Perturbation Theory for Linear Operators*. 2nd ed. Springer-Verlag, Berlin, 1976. xxi+619 pp.
- [27] V.G. Maz'ya & I.E. Verbitsky, The Schrödinger operator on the energy space: boundedness and compactness criteria. *Acta Math.* **188** (2002), 263–302.
- [28] C. McDiarmid, On the method of bounded differences. *Surveys in Combinatorics, 1989*. LMS Lecture Note Series **141**, Cambridge Univ. Press (1989), 148–188.
- [29] R. Otter, The multiplicative process. *Ann. Math. Statist.* **20** (1949), 206–224.
- [30] J. Pitman, Enumerations of trees and forests related to branching processes and random walks. *Microsurveys in Discrete Probability (Princeton, NJ, 1997)*, 163–180, DIMACS Ser. Discrete Math. Theoret. Comput. Sci., 41, Amer. Math. Soc., Providence, RI, 1998.
- [31] O. Riordan, The small giant component in scale-free random graphs, *Combin. Probab. Comput.* **14** (2005), 897–938.
- [32] L.A. Shepp, Connectedness of certain random graphs. *Israel J. Math.* **67** (1989), 23–33.
- [33] J. Spencer & N. Wormald, Birth control for giants. *Combinatorica* **27** (2007), 587–628.
- [34] L. Takács, Ballots, queues and random graphs. *J. Appl. Probab.* **26** (1989), 103–112.
- [35] J.C. Tanner, A derivation of the Borel distribution. *Biometrika* **48** (1961), 222–224.