THE PHASE TRANSITION IN INHOMOGENEOUS RANDOM GRAPHS

BÉLA BOLLOBÁS[*§‖], SVANTE JANSON[†¶], AND OLIVER RIORDAN[‡§]

Abstract. The 'classical' random graph models, in particular $G(n,p)$, are 'homogeneous', in the sense that the degrees (for example) tend to be concentrated around a typical value. Many graphs arising in the real world do not have this property, having, for example, power-law degree distributions. Thus there has been a lot of recent interest in defining and studying 'inhomogeneous' random graph models.

One of the most studied properties of these new models is their 'robustness', or, equivalently, the 'phase transition' as an edge density parameter is varied. For $G(n,p)$, $p = c/n$, the phase transition at $c = 1$ has been a central topic in the study of random graphs for well over 40 years.

Many of the new inhomogeneous models are rather complicated; although there are exceptions, in most cases precise questions such as determining exactly the critical point of the phase transition are approachable only when there is independence between the edges. Fortunately, some models studied have this property already, and others can be approximated by models with independence.

Here we introduce a very general model of an inhomogeneous random graph with (conditional) independence between the edges, which scales so that the number of edges is linear in the number of vertices. This scaling corresponds to the $p = c/n$ scaling for $G(n,p)$ used to study the phase transition; also, it seems to be a property of many large real-world graphs. Our model includes as special cases many models previously studied.

We show that, under one very weak assumption (that the expected number of edges is 'what it should be'), many properties of the model can be determined, in particular the critical point of the phase transition, and the size of the giant component above the transition. We do this by relating our random graphs to branching processes, which are much easier to analyze.

We also consider other properties of the model, showing, for example, that when there is a giant component, it is 'stable': for a typical random graph, no matter how we add or delete $o(n)$ edges, the size of the giant component does not change by more than $o(n)$.

Date: April 29, 2005; revised June 20, 2006.

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

[*]Department of Mathematical Sciences, University of Memphis, Memphis TN 38152, USA.

[†]Department of Mathematics, Uppsala University, PO Box 480, SE-751 06 Uppsala, Sweden.

[‡]Royal Society Research Fellow, Department of Pure Mathematics and Mathematical Statistics, University of Cambridge, UK.

[§]Trinity College, Cambridge CB2 1TQ, UK.

[¶]Churchill College, Cambridge CB3 0DS, UK.

[‖]Research supported by NSF grants CCR 0225610 and DMS 0505550.

Contents

1. Introduction 3
2. The model 6
   2.1. A branching process 13
   2.2. An integral operator 14
3. Main results 14
   3.1. Existence, size and uniqueness of the giant component 15
   3.2. Stability 18
   3.3. Bounds on the small components 19
   3.4. Degree sequence 19
   3.5. Distances between vertices 20
   3.6. The phase transition 22
4. Examples 24
5. Branching process lemmas 30
6. Branching process results 37
7. Approximation 41
8. The number of edges 45
   8.1. Generalized vertex spaces 50
9. The giant component 52
10. Edges in the giant component 64
11. Stability 68
    11.1. Counting cuts in the giant component 70
    11.2. Branching process analysis of the two-core 73
12. Bounds on the small components 83
13. Vertex degrees 88
14. Distances between vertices 90
    14.1. Typical distances 91
    14.2. The diameter 97
15. The phase transition 107
16. Applications and relationship to earlier results 112
    16.1. Dubins' model 112
    16.2. The mean-field scale-free model 113
    16.3. The CHKNS model 114
    16.4. The rank 1 case 116
    16.5. Turova's model 120
    16.6. Functions of $\max\{x, y\}$. 123
17. Paths and cycles 123
18. Further remarks 126

Appendix A. Probabilistic lemmas 129

References 131

1. Introduction

The theory of random graphs was founded in the late 1950s and early 1960s by Erdős and Rényi [47, 48], who started the systematic study of the space $\mathcal{G}(n,M)$ of all graphs with n labelled vertices and $M = M(n)$ edges, with all graphs equiprobable. (Usually, one writes $G(n,M)$ for a random element of $\mathcal{G}(n,M)$.) At about the same time, Gilbert [51] introduced the closely related model $\mathcal{G}(n,p)$ of random graphs on n labelled vertices: a random $G(n,p) \in \mathcal{G}(n,p)$ is obtained by selecting edges independently, each with probability $p = p(n)$. For many questions, such as those considered in this paper, the models are essentially equivalent (if $p = M/\binom{n}{2}$, say). As Erdős and Rényi are the founders of the theory of random graphs, it is not surprising that both $G(n,p)$ and $G(n,M)$ are now known as Erdős–Rényi random graphs.

In addition to these two 'classical' models, much attention has been paid to the space of random r-regular graphs, and to the space $\mathcal{G}(k - \mathrm{out})$ of random directed graphs where each vertex has out-degree k, and the undirected graphs underlying these. All these random graph models are 'homogeneous' in the sense that all vertices are exactly equivalent in the definition of the model. Furthermore, in a typical realization, most vertices are in some sense similar to most others. For example, the vertex degrees in $G(n,p)$ or in $G(n,M)$ do not vary very much: their distribution is close to a Poisson distribution.

In contrast, many large real-world graphs are highly inhomogeneous. One reason is that the vertices may have been 'born' at different times, with old and new vertices having very different properties. Experimentally, the spread of degrees is often very large. In particular, in many examples the degree distribution follows a power law. In the last few years, this has led to the introduction and analysis of many new random graph models designed to incorporate or explain these features. Recent work in this area perhaps started from the observations of Faloutsos, Faloutsos and Faloutsos [49] concerning scaling in real-life networks, in particular the power-law distribution of degrees in the 'internet graph', and similar observations concerning the 'web graph' made by Kleinberg, Kumar, Raghavan, Rajagopalan and Tomkins [66], and by Barabási and Albert [9], who also looked at several other real world graphs. The latter two groups introduced two of the first models to explain these observations, using the ideas of 'copying' and of 'growth with preferential attachment', respectively. Observations of and proposed models for other networks followed, including protein interaction networks, telephone call graphs, scientific collaboration graphs and many others. Extensive surveys of the mostly experimental or heuristic work have been written by Barabási and Albert [3] and Dorogovtsev and Mendes [42].

Some of the first rigorous mathematical results concerning (precisely defined variants of) these new models are those of Bollobás and Riordan [23], Bollobás, Riordan, Spencer and Tusnády [26], Buckley and Osthus [29] and

Cooper and Frieze [40]. For a partial survey of the rapidly growing body of rigorous work see Bollobás and Riordan [21]. Needless to say, surveys in an active area such as this quickly become dated, and there are already too many rigorous results in the field to list here.

Perhaps the most striking and important result of Erdős and Rényi concerns the sudden emergence of the 'giant component', the phase transition in $G(n, p)$ that occurs at $p = 1/n$: if $c > 0$ is a constant, then the largest component of $G(n, c/n)$ has order $O(\log n)$ whp if $c < 1$, and order $\Theta(n)$ whp if $c > 1$ (see Section 3 for the notation). In particular, a giant component of order $\Theta(n)$ exists (with high probability as $n \to \infty$) if and only if $c > 1$. Over twenty years later, Bollobás [15] and Łuczak [71] proved considerably sharper results about the exact nature of this phase transition: in particular, they determined the exact size of the 'window' in which the transition takes place. Further, very detailed results were proved by Janson, Knuth, Łuczak and Pittel [58]; see Bollobás [16] and Janson, Łuczak and Ruciński [59] for numerous results and references.

Our main purpose in this paper is to lay the foundations of a very general theory of inhomogeneous sparse random graphs. To this end

- we shall define a general model that is sufficiently flexible to include exactly many of the specific spaces of inhomogeneous random graphs that have been studied in recent years,
- we shall establish a close connection between the component structure of a random graph in this model, the survival probability of a related branching process, and the norm of a certain operator,
- we shall use these connections to study the phase transition in our model, examining especially the numbers of vertices and edges in the giant component, and
- we shall prove results concerning the stability of the giant component under the addition and deletion of edges.

In addition, we shall study various other properties of our model, including the degree distribution, the numbers of paths and cycles, and the typical distance between pairs of vertices in the giant component. Furthermore, we shall spell out what our results say about many specific models that have been studied previously.

Although we shall give many examples throughout the paper, to motivate the definitions it may help to bear in mind one particular example of the general class of models we shall study. This example is the *uniformly grown random graph*, or *c/j-graph*, $G_n^{1/j}(c)$. Here $c > 0$ is a parameter that will be kept constant as n varies, and the graph $G_n^{1/j}(c)$ is the graph on $[n] = \{1, 2, \dots, n\}$ in which edges are present independently, and the probability that for $i \neq j$ the edge ij is present is $p_{ij} = \min\{c/\max\{i, j\}, 1\}$, or simply c/j if $i < j$ and $c \leq 2$.

A natural generalization of $G(n,p)$ that includes the example above as a very special case is obtained by replacing the single parameter p by a symmetric $n \times n$ matrix (p_{ij}) with $0 \le p_{ij} \le 1$. We write $G(n,(p_{ij}))$ for the random graph with vertex set $[n]$ where i and j are connected by an edge with probability p_{ij}, and these events are independent for all pairs (i,j) with $i < j$; see [16, p. 35]. We are interested in asymptotics as $n \to \infty$, usually with $(p_{ij}) = \big(p_{ij}(n)\big)$ depending on n. It seems difficult to obtain substantial asymptotic results for $G(n,(p_{ij}))$ without further restrictions; the model is too general. (However, for connectedness Alon [7] proved a number of results.)

Here we are mainly interested in random graphs where the average degree is $\Theta(1)$; one of the main cases treated in this paper is $p_{ij} = \kappa(i/n, j/n)/n$ for a suitable function κ on $(0,1]^2$. Taking $\kappa(x,y) = c/\max\{x,y\}$, we obtain $G_n^{1/j}(c)$. Many other graphs studied earlier by different authors can also be obtained by choosing $\kappa(x,y)$ suitably; see Section 16, and the forthcoming papers [20, 57]. A precise definition of the random graphs treated here will be given in Section 2, and some simple examples in Section 4.

The rest of the paper is organized as follows. In Section 2 we define the model $G(n,\kappa)$ we shall study, along with the branching process $\mathfrak{X}_\kappa$ and integral operator T_κ to which we shall relate its component structure.

In Subsection 3.1 we present our main results on the giant component of $G(n,\kappa)$: under certain weak assumptions we obtain necessary and sufficient conditions for $G(n,\kappa)$ to have a giant component, show that when the giant component exists it is unique, and find its normalized size and number of edges. Further results are presented in the following subsections, on the 'stability' of the giant component in Subsection 3.2, on small components in Subsection 3.3, on the degree sequence in Subsection 3.4, and on the typical distance between vertices of the giant component in Subsection 3.5. In Subsection 3.6 we turn to the phase-transition in $G(n,\kappa)$; more precisely, we examine the growth rate of the giant component as it emerges.

Since our model is very general, and the definition rather lengthy, special cases of the model play an important role in the paper; we have described one, $G_n^{1/j}(c)$ already. In Section 4 we give several further simple examples, to illustrate the definitions and results of the previous sections. Towards the end of the paper, we shall discuss several other special cases more extensively, in particular describing the relationship to other models studied earlier; we consider these to be applications rather than illustrations of the model, and so present them after the proofs.

The next several sections are devoted to the proofs of the main results; the reader interested primarily in the applications may wish to skip straight to Section 16. We start by analyzing the branching process $\mathfrak{X}_\kappa$, proving results about this process that will help us relate $G(n,\kappa)$ to $\mathfrak{X}_\kappa$. The study of $\mathfrak{X}_\kappa$ itself is not one of our main aims. In Section 5 we prove various lemmas needed in Section 6 to prove the results about $\mathfrak{X}_\kappa$ that we shall use throughout the paper.

Next, we turn to preparatory results concerning $G(n,\kappa)$ itself, starting with simple approximation lemmas in Section 7; basic results on the number of edges of $G(n,\kappa)$ are given in Section 8.

Our main results about the existence and size of the giant component are proved in Section 9, using material from the previous sections; the reader who is interested only in the derivations of these results should read Sections 2 to 9. The number of edges in the giant component is determined in Section 10.

Sections 11 to 15 are devoted to the proofs of the results in Subsections 3.2 to 3.6: broadly speaking, these proofs rely on the results up to Section 9, but not on each other, so the reader may safely omit any subset of these sections. The stability result is proved in Section 11, the results on small components in Section 12, the vertex degrees are studied in Section 13, the distance between vertices in Section 14, and the phase transition in Section 15. The latter results may be viewed purely as statements about a branching process, in which case the proofs need only the results of Sections 5 and 6.

In Section 16 we apply our general model to deduce results about several specific models, in particular ones that have been studied in recent years, and discuss the relationship of our results to earlier work.

In Section 17 we give some simple results about paths and cycles in $G(n,\kappa)$, inspired by the work of Turova on a special case of the model (described in Section 16), and show that a conjecture of hers holds under mild conditions. In Section 18 we discuss several (at least superficially) related models as well as possible future work. Finally, in the appendix we give some basic results on random measures used throughout the paper.

2. The model

In this section we define the random graph model that we shall study throughout the paper, as well as a branching process and an integral operator that will be key to characterizing the component structure of this random graph. This section also includes various remarks on the definitions, including descriptions of several minor variants; the formal definitions may be understood without reference to these remarks. To make sense of the definitions, the reader may wish to keep in mind the model $G_n^{1/j}(c)$ defined in the introduction.

Our model is an extension of one defined by Söderberg [88]. Let $\mathcal{S}$ be a separable metric space equipped with a Borel probability measure μ. We shall often suppress the measure μ in our notation, writing, for example, $\|\ \|_p$ for the norm in $L^p(\mathcal{S}) = L^p(\mathcal{S},\mu)$, and 'a.e. on $\mathcal{S}$' for μ-a.e. on $\mathcal{S}$. Much of the time (for example, when studying $G_n^{1/j}(c)$), we shall take $\mathcal{S} = (0,1]$ with μ Lebesgue measure. Throughout the paper, the 'kernel' κ will be a symmetric non-negative function on $\mathcal{S} \times \mathcal{S}$. Further conditions on κ will be given in Definitions 2.7 and 2.10.

For each n we have a deterministic or random sequence $\mathbf{x}_n = (x_1, \dots, x_n)$ of points in $\mathcal{S}$. Formally, we should write $\mathbf{x}_n = (x_1^{(n)}, \dots, x_n^{(n)})$, say, as we assume no relationship between the ith elements of $\mathbf{x}_n$ and of $\mathbf{x}_{n'}$. However, this notation would be rather cumbersome, and it will always be clear which $\mathbf{x}_n$ an x_i is an element of. Writing δ_x for the measure consisting of a point mass of weight 1 at x, and

$$\nu_n := \frac{1}{n} \sum_{i=1}^{n} \delta_{x_i} \tag{2.1}$$

for the empirical distribution of $\mathbf{x}_n$, we shall assume that ν_n converges in probability to μ as $n \to \infty$, with convergence in the usual space of probability measures on $\mathcal{S}$ (see, e.g., [13]). This condition has a simple down-to-earth description in terms of the number of x_i in certain sets A: a set $A \subseteq \mathcal{S}$ is a *μ-continuity set* if A is (Borel) measurable and $\mu(\partial A) = 0$, where ∂A is the boundary of A. The convergence condition $\nu_n \xrightarrow{\mathrm{p}} \mu$ means exactly that for every μ-continuity set A,

$$\nu_n(A) := \#\{i : x_i \in A\}/n \xrightarrow{\mathrm{p}} \mu(A); \tag{2.2}$$

see Appendix A for technical details.

One example where (2.2) holds is the random case, where the x_i are independent and uniformly distributed on $\mathcal{S}$ with distribution μ (as in Söderberg [88]); then (2.2) holds by the law of large numbers.

We shall often consider $\mathcal{S} = (0, 1]$ with the Lebesgue measure μ; in this case, condition (2.2) has to be verified only for intervals (see Remark A.3). For this pair $(\mathcal{S}, \mu)$ we shall have two standard choices for the (x_i): the deterministic case $x_i = i/n$, and the random case where the x_i are independent and uniformly distributed on $(0, 1]$. To express $G_n^{1/j}(c)$ as a special case of our model, we shall take $x_i = i/n$.

For later formal statements, we gather the preceding assumptions into the following definitions.

Definition. A *ground space* is a pair $(\mathcal{S}, \mu)$, where $\mathcal{S}$ is a separable metric space and μ is a Borel probability measure on $\mathcal{S}$.

Definition. A *vertex space* $\mathcal{V}$ is a triple $(\mathcal{S}, \mu, (\mathbf{x}_n)_{n \geq 1})$, where $(\mathcal{S}, \mu)$ is a ground space and, for each $n \geq 1$, $\mathbf{x}_n$ is a random sequence $(x_1, x_2, \dots, x_n)$ of n points of $\mathcal{S}$, such that (2.2) holds.

Of course, we do not need $(\mathbf{x}_n)_{n \geq 1}$ to be defined for every n, but only for an infinite set of integers n.

Definition. A *kernel* κ on a ground space $(\mathcal{S}, \mu)$ is a symmetric non-negative (Borel) measurable function on $\mathcal{S} \times \mathcal{S}$. By a kernel on a vertex space $(\mathcal{S}, \mu, (\mathbf{x}_n)_{n \geq 1})$ we mean a kernel on $(\mathcal{S}, \mu)$.

From now on, unless otherwise stated, we shall always write a vertex space $\mathcal{V}$ as $(\mathcal{S}, \mu, (\mathbf{x}_n)_{n\ge1})$, and $\mathbf{x}_n$ as $(x_1, x_2, \dots, x_n)$. As noted above, the (distributions of) the individual x_i depend on n; in the notation we suppress this dependence as it will always be clear which $\mathbf{x}_n$ an x_i is a member of.

Let κ be a kernel on the vertex space $\mathcal{V}$. Given the (random) sequence $(x_1, \dots, x_n)$, we let $G^{\mathcal{V}}(n, \kappa)$ be the random graph $G^{\mathcal{V}}(n, (p_{ij}))$ with

$$p_{ij} := \min\{\kappa(x_i, x_j)/n, 1\}. \tag{2.3}$$

In other words, $G^{\mathcal{V}}(n, \kappa)$ has n vertices $\{1, \dots, n\}$ and, given $x_1, \dots, x_n$, an edge ij (with $i \ne j$) exists with probability p_{ij}, independently of all other (unordered) pairs ij. Often, we shall suppress the dependence on $\mathcal{V}$, writing $G(n, \kappa)$ for $G^{\mathcal{V}}(n, \kappa)$. We have described one example already: if we take $\kappa(x, y) = c/\max\{x, y\}$, with $\mathcal{S} = (0, 1]$, μ Lebesgue measure, and $x_i = x_i^{(n)} = i/n$, then (2.3) gives $p_{ij} = \min\{c/\max\{i, j\}, 1\}$, so $G^{\mathcal{V}}(n, \kappa)$ is exactly the uniformly grown random graph $G_n^{1/j}(c)$ described in the introduction. We shall discuss several other examples in Sections 4 and 16.

Remark 2.1. The random graph $G(n, \kappa)$=$G^{\mathcal{V}}(n, \kappa)$ depends not only on κ but also on the choice of $x_1, \dots, x_n$. Much of the time, our notation will not indicate how the points x_i are chosen, since this choice is irrelevant for our results as long as (2.2) holds and certain pathologies are excluded (see (2.9), Lemma 8.1 and Example 8.6). The freedom of choice of $x_1, \dots, x_n$ gives our model flexibility, as shown by Proposition 9.3, Theorem 12.1 and the examples in Sections 4 and 16, but does not affect the asymptotic behaviour. Of course, this asymptotic behaviour does depend very much on $\mathcal{S}$ and μ.

In order to make our results easy to apply, it will be convenient to extend the definitions above in two ways, by allowing $\mu(\mathcal{S})$ to take any value in $(0, \infty)$, and by allowing the number of vertices of $G^{\mathcal{V}}(n, \kappa)$ to be random, rather than exactly n. As we shall see later, this makes no essential difference, and we shall almost always work with the n vertex model in our arguments. We shall consider the 'generalized' model only for the convenience of a reader wishing to apply the results in the next section, obviating the need for a separate reduction to the n vertex model in each case. All other readers may safely ignore the 'generalized' model, including the formal definitions that we now state.

Definition. A *generalized ground space* is a pair $(\mathcal{S}, \mu)$, where $\mathcal{S}$ is a separable metric space and μ is a Borel measure on $\mathcal{S}$ with $0 < \mu(\mathcal{S}) < \infty$.

Let $I \subset (0, \infty)$ be any unbounded set, the *index set* parametrizing our model. Usually, I is the positive integers, or the positive reals. For compatibility with our earlier definitions, we write n for an element of I, even though this need not be an integer.

Definition. A *generalized vertex space* $\mathcal{V}$ is a triple $(\mathcal{S}, \mu, (\mathbf{x}_n)_{n\in I})$, where $(\mathcal{S}, \mu)$ is a generalized ground space and, for each $n \in I$, $\mathbf{x}_n$ is a random sequence $(x_1, x_2, \ldots, x_{v_n})$ of points of $\mathcal{S}$ of random length $v_n \geq 0$, such that (2.2) holds, i.e., such that

$$\nu_n(A) := \#\{i : x_i \in A\}/n \xrightarrow{\text{p}} \mu(A) \tag{2.4}$$

as $n \in I$ tends to infinity for every μ-continuity set A; equivalently, $\nu_n \xrightarrow{\text{p}} \mu$.

The definition of a *kernel* κ on a generalized ground space is exactly as before. Finally, given a kernel κ on a generalized vertex space, for $n \in I$ we let $G^{\mathcal{V}}(n, \kappa)$ be the random graph on $\{1, 2, \ldots, v_n\}$ in which, given $\mathbf{x}_n = (x_1, \ldots, x_{v_n})$, each possible edge ij, $1 \leq i < j \leq v_n$, is present with probability

$$p_{ij} := \min\bigl\{\kappa(x_i, x_j)/n, 1\bigr\}, \tag{2.5}$$

and the events that different edges are present are independent.

Note that if $\mathcal{V}$ is a generalized ground space, then, applying (2.4) with $A = \mathcal{S}$, we see that $G^{\mathcal{V}}(n, \kappa)$ has $\mu(\mathcal{S})n + o_p(n)$ vertices. In both (2.4) and (2.5) we divide by n, rather than by the actual number of vertices, or by $\mu(\mathcal{S})n$; this turns out to be most convenient normalization. Roughly speaking, by conditioning on the sequences $(\mathbf{x}_n)$, or by adding $o_p(n)$ isolated vertices, we may assume without loss of generality that the number of vertices is deterministic. Furthermore, multiplying κ and the index variable n by some constant factor, and dividing $\mu(\mathcal{S})$ by the same factor, leaves the edge probabilities p_{ij} unchanged. As the condition (2.4) is also unaffected by this transformation, the only effect on the model is to rescale the parameter n, and we may assume without loss of generality that we have a vertex space rather than a generalized vertex space; see Subsection 8.1.

Remark 2.2. We regard our random graphs as indexed by n, and consider what happens as $n \to \infty$. This is for notational convenience only; we could consider graphs indexed by some other (possibly continuous) parameter, m, say, such that the number of vertices v_m of the graph with parameter m tends to infinity. This superficial modification is covered by the definitions above: instead of considering graphs on 2^n vertices, say, one can always consider graphs on n vertices with n restricted to an 'index set' I consisting of the powers of 2.

The generalized vertex space setting also allows the number of vertices to be random. In other words, we may let $\{x_i\}$ be a point process on $\mathcal{S}$, for example, a Poisson process of intensity n; see Examples 4.9 and 4.11 and Subsection 16.5.

Remark 2.3. Changing κ on a set of measure zero may have a significant effect on the graph $G^{\mathcal{V}}(n, \kappa)$; see Example 8.6, for instance. Indeed, if the x_i are deterministic, then $G^{\mathcal{V}}(n, \kappa)$ depends only on the values of κ on a discrete set. This means that in our proofs we cannot just ignore measure zero sets in the usual way. Later we shall impose very weak conditions to control such effects; see Remark 2.8.

Before turning to the key definitions, giving conditions under which we can prove substantial results about $G^{\mathcal{V}}(n,\kappa)$, let us make some remarks about some minor variants of the model.

Remark 2.4. As an alternative to (2.3) (or (2.5)), we could use κ to define the intensities of Poisson processes of edges, and ignore multiple edges, so the probability p_{ij} that there is an edge between i and j would be given by

$$p_{ij} := 1 - \exp\bigl(-\kappa(x_i,x_j)/n\bigr), \tag{2.6}$$

rather than by (2.3). The results below are valid for this version too; this can be shown either by checking that all arguments hold with only trivial changes, or by defining $\kappa_n(x,y) := n\bigl(1-\exp(-\kappa(x,y)/n)\bigr)$ and using the setting in Definition 2.9.

Another alternative, studied by Britton, Deijfen and Martin-Löf [27] in a special case (see Subsection 16.4), is to let $p_{ij}/(1-p_{ij}) = \kappa(x_i,x_j)/n$, i.e., to take

$$p_{ij} := \kappa(x_i,x_j)/(n+\kappa(x_i,x_j)). \tag{2.7}$$

Again, the results below remain valid; we now define $\kappa_n(x,y) := \kappa(x,y)/(1+\kappa(x,y)/n)$.

Remark 2.5. In this paper we treat only simple graphs. One natural variation that yields a multigraph is to let the number of edges between i and j have a Poisson distribution with mean $\kappa(x_i,x_j)/n$. Under suitable conditions (e.g. that κ is bounded), it is easy to see that whp there are no triple edges, and that the number of double edges is $O_p(1)$; more precisely, it has an asymptotic Poisson distribution with mean $\frac{1}{4}\iint\kappa^2$, see Section 17. The underlying simple graph is just the graph defined in Remark 2.4.

Another variation (which can be combined with the previous one) is to permit loops by allowing $i=j$ in the definition above. These variations do not affect our results on component sizes.

Remark 2.6. Our model can be extended to a random graph process on a fixed vertex set describing an inhomogeneously growing random graph: Start without any edges and, given $\mathbf{x}_n$, add edges at random times given by independent Poisson processes with intensities $\kappa(x_i,x_j)/n$. (Ignore multiple edges.) At time t, we obtain the version of the random graph $G(n,t\kappa)$ given by (2.6); cf. Remark 2.4. Alternatively, we may add edges sequentially, with each new edge chosen at random with probabilities proportional to $\kappa(x_i,x_j)$; this gives the same process except for a (random) change of time scale.

Without further restrictions, the model $G^{\mathcal{V}}(n,\kappa)$ we have defined is too general for us to prove meaningful results. Indeed, the entire graph may be determined by the behaviour of κ on a measure zero set. Usually, κ is continuous, so this problem does not arise. However, there are natural examples with κ discontinuous, so we shall assume that κ is continuous a.e. rather than everywhere. With this weaker condition, to relate the behaviour of $G^{\mathcal{V}}(n,\kappa)$ to that of κ we shall need some extra assumptions. The behaviour of the total number of edges turns out to be the key to the elimination of pathologies.

As usual, we write $e(G)$ for the number of edges in a graph G. Note that $\mathbb{E}\, e\big(G(n,(p_{ij}))\big) = \sum_{i<j} p_{ij}$, so we have

$$\mathbb{E}\, e\big(G^{\mathcal{V}}(n,\kappa)\big) = \mathbb{E} \sum_{i<j} \min\big\{\kappa(x_i,x_j)/n, 1\big\}. \tag{2.8}$$

In well behaved cases, this expectation is asymptotically $n\frac{1}{2}\iint \kappa$; see, for example, Lemma 8.1.

Definition 2.7. A kernel κ is *graphical* on a (generalized) vertex space $\mathcal{V} = (\mathcal{S}, \mu, (\mathbf{x}_n))$ if the following conditions hold:

(i) κ is continuous a.e. on $\mathcal{S} \times \mathcal{S}$;

(ii) $\kappa \in L^1(\mathcal{S} \times \mathcal{S}, \mu \times \mu)$;

(iii)

$$\frac{1}{n}\,\mathbb{E}\, e\big(G^{\mathcal{V}}(n,\kappa)\big) \to \tfrac{1}{2} \iint_{\mathcal{S}^2} \kappa(x,y)\, d\mu(x)\, d\mu(y). \tag{2.9}$$

Note that whether κ is graphical on $\mathcal{V}$ depends on the sequences $\mathbf{x}_n$. Also, as we shall see in Remark 8.4, if κ is graphical on $\mathcal{V}$, then so is $c\kappa$ for any constant $c > 0$. (This statement would be trivial without the $\min\{\cdot, 1\}$ operation in the right-hand side of (2.8). With this, it is still not hard to check.)

Remark 2.8. Conditions (i) and (ii) are natural technical conditions; at first sight, condition (iii) is perhaps unexpected. As we shall see, some extra condition is needed to exclude various pathologies; see Example 8.6, for example. Condition (iii) is in fact extremely weak: the natural interpretation of κ is that it measures the density of edges, so the integral should be the expected number of edges, suitably normalized. Thus condition (iii) says that $G^{\mathcal{V}}(n,\kappa)$ has about the right number of edges, so if (iii) does not hold, κ has failed to capture even the most basic property of the graph. What is surprising, is that this condition is enough: we shall show that the assumptions above are enough for κ to capture many properties of the graph.

In fact, in many circumstances, condition (iii) is automatically satisfied. Indeed, one of the two inequalities implicit in this definition, namely

$$\liminf \frac{1}{n}\,\mathbb{E}\, e\big(G^{\mathcal{V}}(n,\kappa)\big) \geq \tfrac{1}{2} \iint_{\mathcal{S}^2} \kappa(x,y)\, d\mu(x)\, d\mu(y),$$

always holds; see Lemma 8.1. This lemma also shows that (iii) holds whenever κ is bounded and $\mathcal{V}$ is a vertex space. It also holds whenever $\mathcal{V}$ is a vertex space in which the x_i are (pairwise) independent and distributed according to μ. Moreover, condition (iii) is likely to hold, and to be easy to verify, for any particular model that is of interest. Proposition 8.9 shows that when (iii) does hold, so the normalized number of edges converges in expectation, then it also converges in probability. Note also that (2.9) holds if and only if the corresponding relation holds for the variants of $G^{\mathcal{V}}(n,\kappa)$ discussed in Remark 2.4; see Remark 8.4.

In conjunction with condition (iii), condition (ii) says that the expected number of edges is $O(n)$, so the (expected) average degree is $O(1)$. There are interesting cases with more edges, but they will not be treated here; cf. Section 18.

We can be somewhat more general and allow minor deviations in (2.3) by letting κ depend on n. This will ensure that our results apply directly to the various variations on the model discussed above. The conditions we shall need on a sequence of kernels are contained in the next definition.

Definition 2.9. Let $\mathcal{V} = (\mathcal{S}, \mu, (\mathbf{x}_n))$ be a (generalized) vertex space and let κ be a kernel on $\mathcal{V}$. A sequence (κ_n) of kernels on $(\mathcal{S}, \mu)$ is *graphical on $\mathcal{V}$ with limit κ* if, for a.e. $(y,z) \in \mathcal{S}^2$,

$$y_n \to y \text{ and } z_n \to z \text{ imply that } \kappa_n(y_n, z_n) \to \kappa(y,z), \tag{2.10}$$

κ satisfies conditions (i) and (ii) of Definition 2.7, and

$$\frac{1}{n}\,\mathbb{E}\, e(G^{\mathcal{V}}(n,\kappa_n)) \to \tfrac{1}{2}\iint_{\mathcal{S}^2} \kappa(x,y)\, d\mu(x)\, d\mu(y). \tag{2.11}$$

Note that if κ is a graphical kernel on $\mathcal{V}$, then the sequence κ_n with $\kappa_n = \kappa$ for every n is a graphical sequence on $\mathcal{V}$ with limit κ. Much of the time, members of a graphical sequence of kernels on $\mathcal{V}$ are themselves graphical on $\mathcal{V}$.

Much of the time, the conditions in Definition 2.9 will be all we shall need to prove results about $G^{\mathcal{V}}(n,\kappa_n)$. However, when we come to the size of the giant component, one additional condition will be needed.

Definition 2.10. A kernel κ on a (generalized) ground space $(\mathcal{S}, \mu)$ is *reducible* if

$$\exists A \subset \mathcal{S} \text{ with } 0 < \mu(A) < \mu(\mathcal{S}) \text{ such that } \kappa = 0 \text{ a.e. on } A \times (\mathcal{S} \setminus A);$$

otherwise κ is *irreducible*. Thus κ is irreducible if

$$A \subseteq \mathcal{S} \text{ and } \kappa = 0 \text{ a.e. on } A \times (\mathcal{S} \setminus A) \text{ implies } \mu(A) = 0 \text{ or } \mu(\mathcal{S} \setminus A) = 0. \tag{2.12}$$

Roughly speaking, κ is reducible if the vertex set of $G^{\mathcal{V}}(n,\kappa)$ can be split into two parts so that the probability of an edge from one part to the other is zero, and irreducible otherwise. For technical reasons, we consider a slight weakening of irreducibility.

Definition 2.11. A kernel κ on a (generalized) ground space $(\mathcal{S}, \mu)$ is *quasi-irreducible* if there is a μ-continuity set $\mathcal{S}' \subseteq \mathcal{S}$ with $\mu(\mathcal{S}') > 0$ such that the restriction of κ to $\mathcal{S}' \times \mathcal{S}'$ is irreducible, and $\kappa(x,y) = 0$ if $x \notin \mathcal{S}'$ or $y \notin \mathcal{S}'$.

Remark 2.12. Given a quasi-irreducible kernel κ and the associated graph $G_n = G^{\mathcal{V}}(n,\kappa)$, we can consider the irreducible restriction κ' of κ to $\mathcal{S}' \times \mathcal{S}'$, and the corresponding graph $G'_n = G^{\mathcal{V}'}(n,\kappa')$ obtained from G_n by deleting the vertices with types in $\mathcal{S} \setminus \mathcal{S}'$; these vertices are isolated in G_n. This graph is an instance of our model with a generalized vertex space $\mathcal{V}'$; note that the number N' of vertices of G'_n is random. Thus, we may reduce suitable questions about quasi-irreducible kernels to the irreducible case. In our main results, the reader will lose nothing by reading irreducible instead of quasi-irreducible. We state some of the results for the quasi-irreducible case because this is all we need in the proofs (even without removing isolated vertices as above), and we sometimes need the quasi-irreducible case of one result to prove the irreducible case of another.

2.1. **A branching process.** Let κ be a kernel on a (generalized) ground space $(\mathcal{S},\mu)$. To study the component structure of $G(n,\kappa)$, we shall use the multi-type Galton–Watson branching process with type space $\mathcal{S}$, where a particle of type $x \in \mathcal{S}$ is replaced in the next generation by a set of particles distributed as a Poisson process on $\mathcal{S}$ with intensity $\kappa(x,y)\,d\mu(y)$. (Thus, the number of children with types in a subset $A \subseteq \mathcal{S}$ has a Poisson distribution with mean $\int_A \kappa(x,y)\,d\mu(y)$, and these numbers are independent for disjoint sets A and for different particles; see, e.g., Kallenberg [61].) We denote this branching process, started with a single particle of type x, by $\mathfrak{X}_\kappa(x)$. When $\mu(\mathcal{S}) = 1$, so μ is a probability measure, we write $\mathfrak{X}_\kappa$ for the same process with the type of the initial particle random, distributed according to μ.

Let $\rho_k(\kappa;x)$ be the probability that the branching process $\mathfrak{X}_\kappa(x)$ has a total population of exactly k particles, and let $\rho_{\geq k}(\kappa;x)$ be the probability that the total population is at least k. Furthermore, let $\rho(\kappa;x)$ be the probability that the branching process survives for eternity. If the probability that a particle has infinitely many children is 0, then $\rho(\kappa;x)$ is equal to $\rho_\infty(\kappa;x)$, the probability that the total population is infinite; see Remark 5.2.

Set

$$\rho_k(\kappa) := \int_{\mathcal{S}} \rho_k(\kappa;x)\,d\mu(x), \qquad \rho(\kappa) := \int_{\mathcal{S}} \rho(\kappa;x)\,d\mu(x), \tag{2.13}$$

and define $\rho_{\geq k}(\kappa)$ analogously. Thus, if $\mu(\mathcal{S}) = 1$, then $\rho(\kappa)$ is the survival probability of the branching process $\mathfrak{X}_\kappa$. Note that multiplying κ by a constant factor c and dividing μ by the same factor leaves the branching process $\mathfrak{X}_\kappa(x)$, and hence $\rho(\kappa;x)$ and $\rho_k(\kappa;x)$, unchanged. However, $\rho(\kappa)$, for example, is decreased by a factor of c.

Remark 2.13. As we shall see later, the branching process $\mathfrak{X}_\kappa(x)$ arises naturally when exploring a component of $G(n,\kappa)$ starting at a vertex of type x; this is directly analogous to the use of the single-type Poisson branching process in the analysis of the Erdős-Rényi graph $G(n,c/n)$. In models with a fixed degree sequence, a related 'size-biased' branching process arises, as it matters how we reach a vertex. Here, due to the independence between edges, there is no size-biasing.

2.2. **An integral operator.** Given a kernel κ on a (generalized) ground space $(\mathcal{S}, \mu)$, let T_κ be the integral operator on $(\mathcal{S}, \mu)$ with kernel κ, defined by

$$(T_\kappa f)(x) = \int_{\mathcal{S}} \kappa(x, y) f(y) \, d\mu(y), \tag{2.14}$$

for any (measurable) function f such that this integral is defined (finite or $+\infty$) for a.e. x. As usual, we need never consider non-measurable functions; in future, we shall assume without comment that all functions considered are measurable. Note that $T_\kappa f$ is defined for every $f \geq 0$, with $0 \leq T_\kappa f \leq \infty$. If $\kappa \in L^1(\mathcal{S} \times \mathcal{S})$, as we shall assume throughout, then $T_\kappa f$ is also defined for every bounded f; in this case $T_\kappa f \in L^1(\mathcal{S})$ and thus $T_\kappa f$ is finite a.e.

We define

$$\|T_\kappa\| := \sup\bigl\{\|T_\kappa f\|_2 : f \geq 0, \, \|f\|_2 \leq 1\bigr\} \leq \infty. \tag{2.15}$$

When finite, $\|T_\kappa\|$ is the norm of T_κ as an operator in $L^2(\mathcal{S}, \mu)$; it is infinite if T_κ does not define a bounded operator in L^2. Trivially, $\|T_\kappa\|$ is at most the Hilbert–Schmidt norm of T_κ:

$$\|T_\kappa\| \leq \|T_\kappa\|_{HS} := \|\kappa\|_{L^2(\mathcal{S}\times\mathcal{S})} = \left(\iint_{\mathcal{S}^2} \kappa(x, y)^2 \, d\mu(x) \, d\mu(y) \right)^{1/2}. \tag{2.16}$$

We also define the non-linear operator Φ_κ by

$$\Phi_\kappa f := 1 - e^{-T_\kappa f} \tag{2.17}$$

for $f \geq 0$. Note that for such f we have $0 \leq T_\kappa f \leq \infty$, and thus $0 \leq \Phi_\kappa f \leq 1$. We shall characterize the survival probability $\rho(\kappa; x)$, and thus $\rho(\kappa)$, in terms of the non-linear operator Φ_κ, showing essentially that the function $x \mapsto \rho(\kappa; x)$ is the maximal fixed point of the non-linear operator Φ_κ; see Theorem 6.2.

3. Main results

In this section we present our main results describing various properties of the general model $G^{\mathcal{V}}(n, \kappa_n)$; some further general results will be given in the later sections devoted to individual properties. In Section 16, we shall present results for special cases of the model, including several that have been studied previously.

All our results are asymptotic, and all unspecified limits are taken as $n \to \infty$. We use the following standard notation: for (deterministic) functions $f = f(n)$ and $g = g(n)$, we write $f = O(g)$ if f/g is bounded, $f = \Omega(g)$ if f/g is bounded away from zero, i.e., if $g = O(f)$, and $f = \Theta(g)$ if $f = O(g)$ and $g = O(f)$. We write $f = o(g)$ if $f/g \to 0$.

Turning to sequences of events and random variables, we say that an event holds *with high probability* (whp), if it holds with probability tending to 1 as $n \to \infty$. (Formally, it is a *sequence* E_n of events that may hold whp, but the n is often suppressed in the notation.) We write $\xrightarrow{\mathrm{p}}$ for convergence in probability. Thus, for example, if $a \in \mathbb{R}$, then $X_n \xrightarrow{\mathrm{p}} a$ if and only if, for every $\varepsilon > 0$, the relations $X_n > a - \varepsilon$ and $X_n < a + \varepsilon$ hold whp.

We shall use O_p, o_p and Θ_p in the standard way (see e.g. Janson, Łuczak and Ruciński [59]); for example, if (X_n) is a sequence of random variables, then $X_n = O_p(1)$ means "X_n is bounded in probability" and $X_n = o_p(1)$ means that $X_n \xrightarrow{\mathrm{p}} 0$. Given a function $f(n) > 0$, we shall write $X_n = O(f(n))$ whp if there exists a constant $C < \infty$ such that $|X_n| \leq Cf(n)$ whp. (This is written $X_n = O_C(f(n))$ in [59].) Note that this is stronger than $X_n = O_p(f(n))$; the two statements can be written as $\exists C \forall \varepsilon \limsup_n \mathbb{P}(|X_n| > Cf(n)) < \varepsilon$ and $\forall \varepsilon \exists C \limsup_n \mathbb{P}(|X_n| > Cf(n)) < \varepsilon$, respectively. We shall use $X_n = \Theta(f(n))$ whp similarly.

We denote the orders of the components of a graph G by $C_1(G) \geq C_2(G) \geq \dots$, with $C_j(G) = 0$ if G has fewer than j components. We let $N_k(G)$ denote the total number of vertices in components of order k, and write $N_{\geq k}(G)$ for $\sum_{j \geq k} N_j(G)$, the number of vertices in components of order at least k.

We shall write $a \wedge b$ and $a \vee b$ for $\min\{a, b\}$ and $\max\{a, b\}$, and use the same notation for the pointwise minimum or maximum of two functions.

As noted in the previous section, a reader who wishes to understand the following results, rather than apply them to a specific model, may safely ignore all references to generalized vertex spaces.

3.1. **Existence, size and uniqueness of the giant component.** Our first result gives a necessary and sufficient condition for the existence of a giant component in our model.

Theorem 3.1. *Let (κ_n) be a graphical sequence of kernels on a (generalized) vertex space $\mathcal{V}$ with limit κ.*

(i) *If $\|T_\kappa\| \leq 1$, then $C_1\big(G^{\mathcal{V}}(n, \kappa_n)\big) = o_p(n)$, while if $\|T_\kappa\| > 1$, then $C_1\big(G^{\mathcal{V}}(n, \kappa_n)\big) = \Theta(n)$ whp.*

(ii) *For any $\varepsilon > 0$, whp we have*

$$\frac{1}{n} C_1\big(G^{\mathcal{V}}(n, \kappa_n)\big) \leq \rho(\kappa) + \varepsilon. \tag{3.1}$$

(iii) *If κ is quasi-irreducible, then*

$$\frac{1}{n} C_1(G^{\mathcal{V}}(n, \kappa_n)) \xrightarrow{\mathrm{p}} \rho(\kappa). \tag{3.2}$$

In all cases $\rho(\kappa) < \mu(\mathcal{S})$; furthermore, $\rho(\kappa) > 0$ if and only if $\|T_\kappa\| > 1$.

This result will be proved in Section 9, along with an additional result, Theorem 9.10, concerning the distribution of the types of the vertices in the giant component. We have included the final statement about $\rho(\kappa)$ for ease of future reference, even though it is purely a statement about the branching process $\mathfrak{X}_\kappa$. As remarked above, $\rho(\kappa)$ can be found from the solutions of the non-linear equation $f = \Phi_\kappa(f)$; see Theorem 6.2.

Theorem 3.1 has several immediate consequences. As customary, we say that a sequence of random graphs G_n (with $\Theta(n)$ vertices in G_n) has a *giant component* (whp) if $C_1(G_n) = \Theta(n)$ whp. For simplicity we state these results in the form where the kernel κ_n is independent of n.

Corollary 3.2. *Let κ be a graphical kernel on a (generalized) vertex space $\mathcal{V}$, and consider the random graphs $G^{\mathcal{V}}(n, c\kappa)$ where $c > 0$ is a constant. Then the threshold for the existence of a giant component is $c = \|T_\kappa\|^{-1}$. More precisely, if $c \le \|T_\kappa\|^{-1}$, then $C_1\big(G^{\mathcal{V}}(n, c\kappa)\big) = o_p(n)$, while if $c > \|T_\kappa\|^{-1}$ and κ is irreducible, then $C_1\big(G^{\mathcal{V}}(n, c\kappa)\big) = \rho(c\kappa)n + o_p(n) = \Theta_p(n)$.*

Corollary 3.3. *Let κ be a graphical kernel on a (generalized) vertex space $\mathcal{V}$. Then the property that $G^{\mathcal{V}}(n, c\kappa)$ has whp a giant component holds for every $c > 0$ if and only if $\|T_\kappa\| = \infty$. Otherwise it has a finite threshold $c_0 > 0$.*

The corollaries above are immediate from Theorem 3.1, the observation that $\|T_{c\kappa}\| = c\|T_\kappa\|$, and the fact that κ graphical on $\mathcal{V}$ implies $c\kappa$ graphical on $\mathcal{V}$ (see Remark 8.4). In the light of the results above, we say that a kernel κ is *subcritical* if $\|T_\kappa\| < 1$, *critical* if $\|T_\kappa\| = 1$, and *supercritical* if $\|T_\kappa\| > 1$. We use the same expressions for a random graph $G(n, \kappa)$ and a branching process $\mathfrak{X}_\kappa$.

The next result shows that the number of edges in the graph at the point where the giant component emerges is maximal in the classical Erdős–Rényi case, or the slightly more general 'homogeneous case' described in Example 4.6; see Section 15 for the proof. (In this result we do need $\mu(\mathcal{S}) = 1$ as a normalization.)

Proposition 3.4. *Let κ_n be a graphical sequence of kernels on a vertex space $\mathcal{V}$ with limit κ, and assume that κ is critical, i.e. $\|T_\kappa\| = 1$. Then $\frac{1}{n}e(G^{\mathcal{V}}(n, \kappa_n)) \xrightarrow{\mathrm{p}} \frac{1}{2}\iint \kappa \le 1/2$, with equality in the uniform case $\kappa = 1$; more precisely, equality holds if and only if $\int_{\mathcal{S}} \kappa(x, y)\, d\mu(y) = 1$ for a.e. x.*

We can also determine the asymptotic number of edges in the giant component. As this is not always uniquely defined, for any graph G, let $\mathcal{C}_1(G)$ be the *largest component* of G, i.e., the component with most vertices, chosen according to any fixed rule if there is a tie. In order to state the next result concisely, let

$$\zeta(\kappa) := \frac{1}{2}\iint_{\mathcal{S}^2} \kappa(x, y)\big(\rho(\kappa; x) + \rho(\kappa; y) - \rho(\kappa; x)\rho(\kappa; y)\big)\, d\mu(x)\, d\mu(y). \tag{3.3}$$

Note that the bracket above is the probability that, given independent branching processes $\mathfrak{X}_\kappa(x)$ and $\mathfrak{X}_\kappa(y)$, at least one survives. Intuitively, given that a certain edge is present in $G^{\mathcal{V}}(n, \kappa_n)$, this edge is in the giant component if, when exploring the rest of the graph from its end-vertices, there is at least one from which we can reach many vertices.

Theorem 3.5. *Let (κ_n) be a graphical sequence of kernels on a (generalized) vertex space $\mathcal{V}$ with quasi-irreducible limit κ. Then*

$$\frac{1}{n}e\big(\mathcal{C}_1(G^{\mathcal{V}}(n,\kappa_n))\big) \xrightarrow{\mathrm{p}} \zeta(\kappa). \tag{3.4}$$

This result will be proved in Section 10, together with some properties of $\zeta(\kappa)$.

Under our assumptions, the giant component is whp unique when it exists; the second largest component is much smaller. Indeed, as we shall show in Section 9, only $o_p(n)$ vertices are in 'large' components other than the largest.

Theorem 3.6. *Let (κ_n) be a graphical sequence of kernels on a (generalized) vertex space $\mathcal{V}$ with quasi-irreducible limit κ, and let $G_n = G^{\mathcal{V}}(n,\kappa_n)$. If $\omega(n) \to \infty$ and $\omega(n) = o(n)$, then*

$$\sum_{j\ge 2:\ C_j(G_n)\ge\omega(n)} C_j(G_n) = o_p(n). \tag{3.5}$$

In particular,

$$C_2(G_n) = o_p(n). \tag{3.6}$$

Remark 3.7. If κ is reducible and the x_i are (absolutely) continuous random variables, then $G^{\mathcal{V}}(n,\kappa)$ decomposes into two (or more) disjoint parts that can be regarded as $G^{\mathcal{V}_i}(n_i,\kappa_i)$, for suitable $\mathcal{V}_i$, n_i and κ_i. By considering the parts separately, many of our results for the irreducible case can be extended to the reducible case; note, however, that each of the parts may contain a giant component, so it is possible to have $C_2 = \Theta_p(n)$. The restriction to the case where the x_i are continuous, which includes the Poisson case of Example 4.9, is necessary unless we impose a further restriction on κ: in general there may be a subset $A \subset \mathcal{S}$ of measure zero which always contains some x_i, and this can link the subgraphs $G^{\mathcal{V}_i}(n_i,\kappa_i)$. Worse still, such an A may contain an x_i with probability bounded away from 0 and 1, so $\frac{1}{n}C_1(G^{\mathcal{V}}(n,\kappa))$ need not converge in probability.

Remark 3.8. If κ and κ' are two kernels on the same vertex space with $\kappa \le \kappa'$, then $G^{\mathcal{V}}(n,\kappa)$ and $G^{\mathcal{V}}(n,\kappa')$ are random graphs on the same vertex set, and there is a natural coupling between them in which $G^{\mathcal{V}}(n,\kappa)$ is always a subgraph of $G^{\mathcal{V}}(n,\kappa')$, i.e., a coupling with $G(n,\kappa) \subseteq G(n,\kappa')$. Similarly, one can couple the corresponding branching processes so that every particle present in one is present in the other, i.e., so that $\mathfrak{X}_\kappa \subseteq \mathfrak{X}_{\kappa'}$. Thus $\rho(\kappa) \le \rho(\kappa')$. If κ is irreducible and $\rho(\kappa) > 0$, then it follows from Theorem 6.2 and Lemma 5.12 that $\rho(\kappa') > \rho(\kappa)$ unless $\kappa' = \kappa$ a.e. Similarly, the threshold $c_0(\kappa') := \|T_{\kappa'}\|^{-1}$ is at most $c_0(\kappa) := \|T_\kappa\|^{-1}$. Here, however, somewhat surprisingly, we may have $c_0(\kappa') = c_0(\kappa)$ even if $\kappa' > \kappa$; see Subsection 16.3. On the other hand, it is easily seen that if T_κ is compact, κ is irreducible, and $\kappa' > \kappa$ on a set of positive measure, then $\|T_{\kappa'}\| > \|T_\kappa\|$ and thus $c_0(\kappa') < c_0(\kappa)$.

3.2. **Stability.** The giant component of $G_n = G^{\mathcal{V}}(n, \kappa_n)$ is stable in the sense that its size does not change much if we add or delete a few edges; this is made precise in the following theorem. Note that the edges added or deleted do not have to be random or independent of the existing graph; they can be chosen by an adversary after inspecting the whole of G_n. Also, we may delete vertices instead of (or as well as) edges.

Theorem 3.9. *Let (κ_n) be a graphical sequence of kernels on a (generalized) vertex space $\mathcal{V}$ with irreducible limit κ, and let $G_n = G^{\mathcal{V}}(n, \kappa_n)$. For every $\varepsilon > 0$ there is a $\delta > 0$ (depending on κ) such that, whp,*

$$(\rho(\kappa) - \varepsilon)n \leq C_1(G'_n) \leq (\rho(\kappa) + \varepsilon)n \tag{3.7}$$

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.

In particular, if G'_n is a graph on $V(G_n)$ with $e(G'_n \triangle G_n) = o_p(n)$ then

$$C_1(G'_n) = C_1(G_n) + o_p(n) = \rho(\kappa)n + o_p(n).$$

Theorem 3.9 is proved in Section 11. Clearly, in proving the first inequality in (3.7), we may assume that $G'_n \subseteq G_n$, and in proving the second that $G_n \subseteq G'_n$. The latter case will be easy to deal with using Theorem 3.6. Proving the first inequality amounts to showing that, whp, the giant component of G_n cannot be cut into two pieces of size at least $\Theta(n)$ by deleting $o(n)$ vertices and then $o(n)$ edges. For edge deletion, Luczak and McDiarmid [70] gave a very simple proof of this result in the Erdős-Rényi case, which adapts easily to the finite type case and hence (using our general results) to the full generality of Theorem 3.9. This proof is presented in Section 11.

Another approach to proving Theorem 3.9 involves reducing this statement to an equivalent statement about the two-core, which is very easy to prove in the uniform case. This reduction involves relating the two-core to the branching process, using results that we believe are of interest in their own right, presented in Section 11. Unfortunately, while the general case of the two-core result can be proved by branching process methods, the proof is very complicated, so we shall not give it.

Remark 3.10. Theorem 3.9 may be viewed as a statement about the vulnerability of large-scale networks to attack by an adversary who knows the detailed structure of the network, and attempts to disconnect the network into small pieces by deleting a small fraction of the vertices or edges. The vulnerability of 'scale-free' networks to such attacks has been considered by many people; see, for example, [5, 31, 38, 24]; it turns out that such networks are much more resilient to random failures than $G(n, c/n)$, but also more vulnerable to attack. In general, the flexibility available to the attacker makes rigorous analysis difficult, although a result for the Barabási-Albert model was given in [24]. Theorem 3.9 shows in particular that, for $G^{\mathcal{V}}(n, \kappa)$, the network is at most a constant factor more vulnerable than a homogeneous network: a constant fraction of the vertices or edges must be deleted to destroy (or significantly shrink) the giant component.

Remark 3.11. As pointed out by Britton and Martin-Löf [28], in the case of vertex deletion Theorem 3.9 also has the following interpretation: suppose that G_n represents the network of contacts that may allow the spread of an infectious disease from person to person, and that we wish to eliminate the possibility of an epidemic by vaccinating some of the population. Even if the entire network of contacts is known, if the source of the infection is not known, a significant (constant, as $n \to \infty$) proportion of the population must be vaccinated: otherwise, there is still a giant component in the graph on the unvaccinated people, and if the infection starts at one of its vertices, it spreads to $\Theta(n)$ people.

3.3. **Bounds on the small components.** For the classical random graph $G(n, c/n)$ it is well-known that in the subcritical ($c < 1$) case, $C_1 = O(\log n)$ whp, and that in the supercritical ($c > 1$) case, $C_2 = O(\log n)$ whp; see [16, 59], for example. These bounds do not always hold in the general framework we are considering here, but if we add some conditions, then we can improve the estimates $o_p(n)$ in Theorem 3.1 and (3.6) to $O(\log n)$ whp. As before, we write G_n for $G^{\mathcal{V}}(n, \kappa_n)$.

Theorem 3.12. *Let (κ_n) be a graphical sequence of kernels on a (generalized) vertex space $\mathcal{V}$ with limit κ.*

(i) *If κ is subcritical, i.e., $\|T_\kappa\| < 1$, and $\sup_{x,y,n} \kappa_n(x, y) < \infty$, then $C_1(G_n) = O(\log n)$ whp.*
(ii) *If κ is supercritical, i.e., $\|T_\kappa\| > 1$, κ is irreducible, and either $\inf_{x,y,n} \kappa_n(x, y) > 0$ or $\sup_{x,y,n} \kappa_n(x, y) < \infty$, then $C_2(G_n) = O(\log n)$ whp.*

Theorem 3.12 is proved in Section 12. Note that in part (ii) we draw the same conclusion from the very different assumptions $\inf_{x,y,n} \kappa_n(x, y) > 0$ and $\sup_{x,y,n} \kappa_n(x, y) < \infty$. There is no similar result for the subcritical case (part (i)) assuming only that $\inf_{x,y,n} \kappa_n(x, y) > 0$: [19, Theorems 1 and 2] show that the random graph $G_n^{1/j}(c)$ with $0 < c < 1/4$ is subcritical and satisfies $C_1(G_n^{1/j}(c)) = n^{\Theta(1)}$ whp.

3.4. **Degree sequence.** We next turn to the degrees of the vertices of $G_n = G^{\mathcal{V}}(n, \kappa_n)$, where $\kappa_n \to \kappa$. As we shall see, the degree of a vertex of a given type x is asymptotically Poisson with a mean

$$\lambda(x) := \int_{\mathcal{S}} \kappa(x, y)\, d\mu(y) \tag{3.8}$$

that depends on x. This leads to a mixed Poisson distribution for the degree D of a (uniformly chosen) random vertex of G_n. We write Z_k for the number of vertices of G_n with degree k.

Theorem 3.13. *Let (κ_n) be a graphical sequence of kernels on a (generalized) vertex space $\mathcal{V}$ with limit κ, and let $G_n = G^{\mathcal{V}}(n, \kappa_n)$. For any fixed $k \geq 0$,*

$$Z_k/n \xrightarrow{\mathrm{p}} \int_{\mathcal{S}} \frac{\lambda(x)^k}{k!} e^{-\lambda(x)} \, d\mu(x),$$

where $\lambda(x)$ is defined by (3.8)*. Equivalently,*

$$Z_k/|V(G_n)| \xrightarrow{\mathrm{p}} \mathbb{P}(\Xi = k),$$

where Ξ has the mixed Poisson distribution $\int_{\mathcal{S}} \mathrm{Po}(\lambda(x)) \, d\mu(x)/\mu(\mathcal{S})$.

In other words, if D is the degree of a random vertex of $G_n = G^{\mathcal{V}}(n, \kappa_n)$, and we normalize so that $\mu(\mathcal{S}) = 1$, then

$$\mathcal{L}(D \mid G_n) \xrightarrow{\mathrm{p}} \mathcal{L}(\Xi) = \int_{\mathcal{S}} \mathrm{Po}(\lambda(x)) \, d\mu(x).$$

As we shall show in Section 13 and Subsections 16.2 and 16.4, our model includes natural examples of 'scale-free' random graphs, where the degree distribution has a power-law tail. We believe that when it comes to modelling real-world graphs with, for example, observed power laws for vertex degrees, our model provides an interesting and flexible alternative to existing models based on generating graphs with a given degree sequence (e.g., Molloy and Reed [77, 78]), or given expected degrees (e.g., Aiello, Chung and Lu [1]).

3.5. **Distances between vertices.** Next, we consider the distances between vertices of $G_n = G^{\mathcal{V}}(n, \kappa_n)$ where, as usual, κ_n is a graphical sequence of kernels on $\mathcal{V}$ with limit κ. Let us write $d(v, w)$ for the graph distance between two vertices of G_n, which we take to be infinite if they lie in different components. Note that

$$\bigl|\{\{v, w\} : d(v, w) < \infty\}\bigr| = \sum_i \binom{C_i(G_n)}{2}, \tag{3.9}$$

where $\{v, w\}$ denotes an unordered pair of distinct vertices of G_n.

Under certain conditions, we can give upper and lower bounds on $d(v, w)$ for almost all pairs with $d(v, w) < \infty$.

Theorem 3.14. *Let κ_n be a graphical sequence of kernels on a (generalized) vertex space $\mathcal{V}$ with limit κ, with $\|T_\kappa\| > 1$. Let $G_n = G^{\mathcal{V}}(n, \kappa_n)$, and let $\varepsilon > 0$ be fixed.*

(i) *If κ is irreducible, then*

$$\bigl|\{\{v, w\} : d(v, w) < \infty\}\bigr| = \frac{C_1(G_n)^2}{2} + o_p(n^2) = \frac{\rho(\kappa)^2 n^2}{2} + o_p(n^2).$$

(ii) *If $\sup_{x,y,n} \kappa_n(x, y) < \infty$, then*

$$\bigl|\{\{v, w\} : d(v, w) \leq (1 - \varepsilon) \log n / \log \|T_\kappa\|\}\bigr| = o_p(n^2).$$

(iii) *If κ is irreducible and $\|T_\kappa\| < \infty$, then*

$$\bigl|\bigl\{\{v,w\} : d(v,w) \le (1+\varepsilon)\log n/\log\|T_\kappa\|\bigr\}\bigr| = \rho(\kappa)^2 n^2/2 + o_p(n^2).$$

(iv) *If κ is irreducible and $\|T_\kappa\| = \infty$, then there is a function $f(n) = o(\log n)$ such that*

$$\bigl|\bigl\{\{v,w\} : d(v,w) \le f(n)\bigr\}\bigr| = \rho(\kappa)^2 n^2/2 + o_p(n^2). \tag{3.10}$$

Note that part (i) is immediate from (3.9) and Theorems 3.1 and 3.6. Related earlier results are discussed briefly in Section 14.

In the finite-type non-critical case, we can give an asymptotic formula for the 'diameter' of G_n, i.e., for

$$\operatorname{diam}(G_n) := \max\{d(v,w) \ : \ v,w \in V(G),\ d(v,w) < \infty\},$$

the maximum of the diameters of the components of G_n. This turns out to depend not only on the norm of T_κ, but also on the norm of the operator associated to the 'dual kernel' $\hat\kappa$.

Definition 3.15. Let κ be a supercritical kernel on a (generalized) ground space $(\mathcal{S},\mu)$. The *dual kernel* is the kernel $\hat\kappa$ on the generalized ground space $(\mathcal{S},\hat\mu)$ defined by $\hat\kappa(x,y) = \kappa(x,y)$, with $d\hat\mu(x) = (1-\rho(\kappa;x))\,d\mu(x)$.

Note that $\hat\kappa$ and κ are identical as functions on $\mathcal{S}\times\mathcal{S}$. However, they are defined on different generalized ground spaces. Hence, the operators T_κ and $T_{\hat\kappa}$ have (in general) different norms. If we wish to consider only ground spaces, we may renormalize, defining $\hat\kappa'$ on $(\mathcal{S},\hat\mu')$ by $\hat\kappa'(x,y) = (1-\rho(\kappa))\kappa(x,y)$ and $d\hat\mu'(x) = (1-\rho(\kappa;x))/(1-\rho(\kappa))\,d\mu(x)$. The choice of normalization does not affect the norm of the operator: $\|T_{\hat\kappa}\| = \|T_{\hat\kappa'}\|$.

The relevance of the dual kernel is that it describes the 'small' components of $G^{\mathcal{V}}(n,\kappa)$; see Section 12. The distribution of these small components is essentially the same as the distribution of trees hanging off the two-core of the giant component, which affects the diameter of $G^{\mathcal{V}}(n,\kappa)$.

Theorem 3.16. *Let κ be a kernel on a (generalized) vertex space $\mathcal{V} = (\mathcal{S},\mu,(\mathbf{x}_n))$, with $\mathcal{S} = \{1,2,\dots,r\}$ finite and $\mu(\{i\}) > 0$ for each i, and let $G_n = G^{\mathcal{V}}(n,\kappa)$. If $0 < \|T_\kappa\| < 1$, then*

$$\frac{\operatorname{diam}(G_n)}{\log n} \xrightarrow{\mathrm{p}} \frac{1}{\log\|T_\kappa\|^{-1}}$$

as $n\to\infty$. If $\|T_\kappa\| > 1$ and κ is irreducible, then

$$\frac{\operatorname{diam}(G_n)}{\log n} \xrightarrow{\mathrm{p}} \frac{2}{\log\|T_{\hat\kappa}\|^{-1}} + \frac{1}{\log\|T_\kappa\|},$$

where $\hat\kappa$ is the dual kernel to κ.

Note that we do not require κ to be graphical on $\mathcal{V}$: if $\mathcal{V}$ is a vertex space, then, as $\mathcal{S}$ is finite, any kernel κ on $\mathcal{V}$ is graphical; see Remark 4.5. If $\mathcal{V}$ is a generalized vertex space, then κ need not be graphical. However, by conditioning on the sequences $(\mathbf{x}_n)$, we can reduce to the vertex space case; see Subsection 8.1.

The assumptions of Theorem 3.16 are much more restrictive than those of our other results: we require the type space to be finite. Note, however, that even the single type case of this result, concerning the classical random graph $G(n, c/n)$, is non-trivial; it answers in the negative a question of Chung and Lu [33]. This special case of Theorem 3.16 was proved independently by Fernholz and Ramachandran [50], again as a special case of a result for a more general model. The nature of their model makes their proof much more difficult than that of Theorem 3.16; see Subsection 14.2.

3.6. **The phase transition.** Finally, we turn to the phase transition in $G^{\mathcal{V}}(n,\kappa)$, where the giant component first emerges. As usual, to study the transition, we should vary a single density parameter. Here, it is most natural to fix a graphical kernel κ on a vertex space $\mathcal{V}$, and to study $G^{\mathcal{V}}(n, c\kappa)$ for a real parameter $c > 0$, as in Corollary 3.2. Instead, we could consider random subgraphs of $G^{\mathcal{V}}(n,\kappa)$ obtained by keeping each edge, or edge vertex, independently with probability p, and use p as the parameter; as we shall see in Examples 4.10 and 4.11, all three approaches are equivalent, so we shall use the first.

By Theorem 3.1, the size of the largest component of $G^{\mathcal{V}}(n, c\kappa)$ is described by the function $\rho(c\kappa)$, which is 0 for $c \le c_0 := \|T_\kappa\|^{-1}$ and strictly positive for larger c. With $\mathcal{V}$ and κ fixed, let us denote this function by $\rho(c)$, $c > 0$. We shall see (from Theorem 6.4) that $\rho(c)$ is continuous on $(0,\infty)$.

Since $\rho(c) = 0$ for $c \le c_0$ but not for larger c, the function ρ is not analytic at c_0; in physical terminology, *there is a phase transition at* c_0.

For the classical Erdős–Rényi random graph $G(n, c/n)$ (obtained with $\kappa = 1$), it is well-known that ρ is continuous but the first derivative has a jump at $c_0 = 1$; more precisely, ρ' jumps from 0 to $\rho'_+(c_0) = 2$. For finite d, we shall say that the phase transition in $G^{\mathcal{V}}(n,\kappa)$ has *exponent* k if $\rho(c_0+\varepsilon) = \Theta(\varepsilon^k)$ as $\varepsilon \searrow 0$. As we have just noted, in $G(n, c/n)$ the phase transition has exponent 1. If $\rho(c_0 + \varepsilon) = o(\varepsilon^k)$ for all k, we say that the phase transition has *infinite exponent*. We are deliberately avoiding the physical term 'order', as it is not used in a consistent way in this context. In other contexts, discontinuous phase transitions are possible; see, for example, Aizenman, Chayes, Chayes and Newman [2].

It was shown in [19] (see also Dorogovtsev, Mendes and Samukhin [43] and Durrett [44]) that in the case $\mathcal{S} = (0,1]$ and $\kappa(x,y) = 1/(x \vee y)$, the phase transition 'is of infinite order', i.e., has infinite exponent (see Subsection 16.1 for more details). We shall see in Subsection 16.4 that it is also possible to have a phase transition with any finite exponent larger than 1 (including non-integer values).

The next theorem shows that the phase transition has exponent 1 for a wide class of kernels κ, including all bounded κ. We also prove that for this class there is no other phase transition: as $\rho = 0$ on $(0, c_0)$, it is trivially analytic there, and we shall prove that ρ is analytic on (c_0, ∞). As ρ is defined in terms of the branching process, rather than a graph, we do not need a vertex space for the statement of the next result; to deduce conclusions for graphs of the type we consider, we should let κ be an irreducible graphical kernel on a vertex space $\mathcal{V}$, satisfying the additional condition (3.11) below. (Also, there is no need to consider generalized ground spaces, as we may trivially normalize so that $\mu(\mathcal{S}) = 1$ by multiplying κ by $\mu(\mathcal{S})$ and dividing μ by the same factor – this leaves the branching process unchanged.) 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 3.17. *Let κ be a kernel on a ground space $(\mathcal{S}, \mu)$. Suppose that κ is irreducible, and that*

$$\sup_x \int_{\mathcal{S}} \kappa(x,y)^2 \, d\mu(y) < \infty. \tag{3.11}$$

(i) *The function $c \mapsto \rho(c) := \rho(c\kappa)$ is analytic except at $c_0 := \|T_\kappa\|^{-1}$.*
(ii) *The linear operator T_κ has an eigenfunction ψ of eigenvalue $\|T_\kappa\| < \infty$, and every such eigenfunction is bounded and satisfies*

$$\rho(c_0 + \varepsilon) = 2c_0^{-1} \frac{\int_{\mathcal{S}} \psi \int_{\mathcal{S}} \psi^2}{\int_{\mathcal{S}} \psi^3} \varepsilon + O(\varepsilon^2), \qquad \varepsilon > 0, \tag{3.12}$$

so $\rho'_+(c_0) = 2c_0^{-1} \int_{\mathcal{S}} \psi \int_{\mathcal{S}} \psi^2 / \int_{\mathcal{S}} \psi^3 > 0$ and ρ has a phase transition at c_0 with exponent 1.

The proof is given in Section 15. Note that (3.11) implies that $\kappa \in L^2 \subseteq L^1$. Theorem 3.17 has an easy consequence concerning the extremality of the Erdős–Rényi random graphs, also proved in Section 15.

Corollary 3.18. *Let κ be an irreducible kernel on a ground space $(\mathcal{S}, \mu)$ such that* (3.11) *holds, and let $c_0 := \|T_\kappa\|^{-1} > 0$. Then $c_0 \rho'_+(c_0) \le 2$, with equality in the classical Erdős–Rényi case; more precisely, equality holds if and only if $\int_{\mathcal{S}} \kappa(x,y)\, d\mu(y) = 1$ for a.e. x.*

Let κ be an irreducible graphical kernel on a vertex space $\mathcal{V}$; let us assume (3.11) and, as a normalization, that $c_0 = 1$. Letting c increase from the threshold c_0, Corollary 3.18 says that the giant component of $G^{\mathcal{V}}(n, c\kappa)$ has maximal growth-rate in the Erdős–Rényi case, and, more generally, in the 'homogeneous' case treated in Example 4.6 below. In this example, the vertex degrees are more or less the same, so there is no first-order inhomogeneity in the graph; any inhomogeneity in vertex degrees leads to a slower growth.

Remark 3.19. By Theorem 3.5, the number of edges in the giant component of $G(n, c\kappa)$ near the phase transition is asymptotically determined by the behaviour of the function $\zeta(c\kappa)$ as $c \searrow c_0$. As we shall show in Proposition 10.1, if $\|T_\kappa\| < \infty$, then $\zeta(c\kappa)/\rho(c\kappa) \to 1$ as $c \searrow c_0 := \|T_\kappa\|^{-1}$. In particular, under the conditions of Theorem 3.17, there is a phase transition of exponent 1 in ζ too. (In addition, the proof of Theorem 3.17 will show that $c \mapsto \zeta(c\kappa)$ is also analytic except at c_0.) In the case $\|T_\kappa\| = \infty$, when $c_0 = 0$, it is not always true that $\zeta(c\kappa) \sim \rho(c\kappa)$ as $c \searrow c_0$: this will be shown by Example 4.13. An important case when this does hold is described in Subsection 16.4.

4. Examples

In this section we give several simple examples of the random graph model we study; these examples are chosen to illustrate the definitions and the scope of the model, as well as various pathologies that may occur. In subsequent sections we shall refer to several of these examples; in particular, many of our proofs will be based on the 'finite-type' case. Further examples of interest in their own right are discussed at length in Section 16, as applications of our results. We often suppress the dependence on $\mathcal{V}$, writing $G(n,\kappa)$ for $G^{\mathcal{V}}(n,\kappa)$.

Example 4.1. The Erdős-Rényi random graph. If $\kappa = c$ is constant, then the edge probabilities p_{ij} given by (2.3) are all equal to c/n (for $n > c$). Thus any choice of vertex space gives the classical Erdős–Rényi random graph $G(n, c/n)$. The simplest choice is to let $\mathcal{S}$ consist of a single point. Then the operator T_κ is simply multiplication by c, so $\|T_\kappa\| = c$ and Corollary 3.2 yields the classical result that there is a phase transition at $c = 1$. Furthermore, the function $\rho(c; x)$ reduces to the single value $\rho(c)$, and the survival probability $\rho(c)$ of the branching process $\mathfrak{X}_c$ is given by the formula

$$\rho(c) = 1 - e^{-c\rho(c)}, \qquad \text{with } \rho(c) > 0 \text{ if } c > 1; \tag{4.1}$$

this classical branching process result is the simplest case of Theorem 6.2 below. Returning to the graph, in this case Theorem 3.1 reduces to the classical result of Erdős and Rényi [47].

Example 4.2. The homogeneous bipartite random graph. Set $\mathcal{S} = \{1, 2\}$, $\mu\{1\} = \mu\{2\} = 1$, and let $\mathcal{V} = (\mathcal{S}, \mu, (\mathbf{x}_n))$ be a generalized vertex space in which $\mathbf{x}_n$ consists of n vertices of type 1 and n vertices of type 2. Let κ be defined by $\kappa(1,1) = \kappa(2,2) = 0$ and $\kappa(1,2) = \kappa(2,1) = c$. Then $G^{\mathcal{V}}(n,\kappa)$ is the random bipartite graph $G(n, n; c/n)$ with n vertices in each class, where each possible edge between classes is present with probability c/n, independently of the other edges. While it is natural to use a generalized vertex space to describe this example, it is not necessary: the same graph can be written as $G^{\mathcal{V}}(n,\kappa)$ in another way: take $\mu\{1\} = \mu\{2\} = 1/2$, and let $\mathcal{V}$ be a vertex space where $\mathbf{x}_n$ is defined only for n even, and then consists of $m = n/2$ vertices of each type. Let $\kappa(1,1) = \kappa(2,2) = 0$ as before, and $\kappa(1,2) = \kappa(2,1) = 2c$, so the edge probabilities are $2c/n = c/m$.

Example 4.3. The finite-type case. Let $\mathcal{S} = \{s_1, \dots, s_r\}$ be finite. Then κ is an $r \times r$ matrix. In this case, $G(n, \kappa)$ has vertices of r different types (or colours), say n_i vertices of type i, with two vertices of types i and j joined by an edge with probability $n^{-1}\kappa(i,j)$ (for $n \ge \max \kappa$). The condition (2.2) means that $n_i/n \to \mu_i$ for each i (in probability if the n_i are random), where $\mu_i := \mu\{i\} \ge 0$.

This case has been studied by Söderberg [88, 89, 90, 91], who noted our Theorem 3.1 in this case (with $\kappa_n = \kappa$ for all n).

Most of our proofs will be based on a disguised form of this case, described by the following definition.

Definition 4.4. A kernel κ on a (generalized) ground space $(\mathcal{S}, \mu)$ is *regular finitary* if $\mathcal{S}$ has a finite partition into sets $S_1, \dots, S_r$ such that κ is constant on each $S_i \times S_j$, where each S_i is a μ-continuity set, i.e., is measurable and has $\mu(\partial S_i) = 0$.

Clearly, if κ is regular finitary on $(\mathcal{S}, \mu)$ then the random graph $G^{\mathcal{V}}(n, \kappa)$ has the same distribution as a finite-type graph $G^{\mathcal{V}'}(n, \kappa')$, $\mathcal{V}' = (\mathcal{S}', \mu', (\mathbf{y}_n))$: take $\mathcal{S}' = \{1, \dots, r\}$, let $y_k = i$ whenever $x_k \in S_i$, and define $\mu'\{i\}$ and $\kappa'(i,j)$ in the obvious way. Let $n_i = \#\{l : x_l \in S_i\} = n\nu_n(S_i)$, where ν_n is as in (2.1). The numbers n_i may be random, but since each S_i is a μ-continuity set, (2.2) yields $n_i/n = \nu_n(S_i) \xrightarrow{\mathrm{p}} \mu(S_i)$, so $\mathcal{V}'$ is a (generalized) vertex space.

Remark 4.5. Let us note for later that a finite-type or regular finitary kernel κ on a vertex space $\mathcal{V}$ is automatically graphical on $\mathcal{V}$; conditions (i) and (ii) of Definition 2.7 are trivial in this case, while condition (iii) holds in the much more general case of κ bounded; see Lemma 8.1. This observation does not extend to generalized vertex spaces: there may be a very large number of vertices with some small probability, so the expectation in (2.9) need not converge, or even exist; see Remark 8.2.

Example 4.6. The homogeneous case. Generalizing the Erdős-Rényi and homogeneous bipartite cases above, let $(\mathcal{S}, \mu)$ be an arbitrary (generalized) ground space, and let κ be such that $\int_{\mathcal{S}} \kappa(x,y)\,d\mu(y)$ is essentially independent of $x \in \mathcal{S}$, i.e., that

$$\int_{\mathcal{S}} \kappa(x,y)\,d\mu(y) = c \qquad \text{for a.e. } x, \tag{4.2}$$

for some constant c. (This says roughly that, asymptotically, all vertices have the same average degree.) Then $T_\kappa 1 = c$ a.e., so the constant function 1 is a positive eigenfunction with eigenvalue c, and thus $\|T_\kappa\| = c$, and by Theorem 3.1 there is a giant component (and $\rho(\kappa) > 0$) if and only if $c > 1$.

Normalizing (if necessary) so that $\mu(\mathcal{S}) = 1$, in the branching process, apart from particles with types in a measure zero set, which arise in $\mathfrak{X}_\kappa$ with probability 0, the number of children of each particle has a $\mathrm{Po}(c)$ distribution. Hence, ignoring the types of the particles, the distributions of the process $\mathfrak{X}_\kappa$ and the single-type process $\mathfrak{X}_c$ are the same. In particular, $\rho(\kappa) = \rho(c)$, so $\rho(\kappa) = \rho(c)$ is given by (4.1) in this case too. If κ is irreducible, the global behaviour of $G(n, \kappa)$ is thus exactly the same as that of $G(n, c/n)$, at least in terms of the size of the giant component. The local behaviour can be quite different, though. For example, $G(n, \kappa)$ may have many more triangles or other small cycles than $G(n, c/n)$; see Example 17.4. On the other hand, by Theorem 3.13, the vertex degrees have an asymptotic $\mathrm{Po}(c)$ distribution just as in $G(n, c/n)$.

A natural example of such a homogeneous κ is given by taking $\mathcal{S}$ as $(0, 1]$ (now better regarded as the circle $\mathbb{T}$), μ as Lebesgue measure, and $\kappa(x, y) = h(x - y)$ for an even function $h \geq 0$ of period 1. For example, h can be constant on a small interval $(-\delta, \delta)$ and vanish outside it; this gives a modification of $G(n, c/n)$ where only "short" edges are allowed.

More generally, $\mathcal{S}$ can be any compact homogeneous space, for example a sphere, with Haar measure μ and an invariant metric d, and $\kappa(x, y)$ a function of the distance $d(x, y)$.

Example 4.7. Take $\mathcal{S} = (0, 1]$ with μ the Lebesgue measure, and let $x_i = i/n$. Set $\kappa(x, y) = \mathbf{1}[x + y \leq 1]$ and consider the kernel $c\kappa$, so that

$$p_{ij} = \begin{cases} c/n, & i + j \leq n; \\ 0, & i + j > n. \end{cases}$$

Thus $G(n, c\kappa)$ can be obtained from the random graph $G(n, c/n)$ by deleting all edges ij with $i + j > n$.

The operator T_κ is compact, and it easy to see that it has eigenvalues $(-1)^k \omega_k^{-1}$ and eigenfunctions $\cos(\omega_k x)$, with $\omega_k = (k + 1/2)\pi$, $k = 0, 1, \dots$. Hence $\|T_\kappa\| = 2/\pi$ and the critical value is $c_0 = \pi/2$. Theorem 3.17 shows that at the critical value we have $c_0 \rho'_+(c_0) = 3/2$.

Example 4.8. I.i.d. vertices. For any ground space $(\mathcal{S}, \mu)$, we can obtain a vertex space by taking $x_1, \dots, x_n$ to be i.i.d. random points in $\mathcal{S}$ with distribution μ. (This has been proposed by Söderberg [88].) In this case

$$\begin{aligned} \mathbb{E}\, e\big(G(n, \kappa)\big) &= \frac{n(n-1)}{2} \iint_{\mathcal{S}^2} \frac{\kappa(x, y) \wedge n}{n} \, d\mu(x)\, d\mu(y) \\ &< \frac{n}{2} \iint_{\mathcal{S}^2} \kappa(x, y)\, d\mu(x)\, d\mu(y). \end{aligned}$$

Hence, by Lemma 8.1 below, (2.9) always holds, and to verify that a kernel κ is graphical, we only have to check conditions (i) and (ii) in Definition 2.7. Similarly, for a sequence of kernels, (2.11) holds provided $\iint \kappa_n \to \iint \kappa$.

Example 4.9. Poisson process graph. For any generalized ground space $(\mathcal{S}, \mu)$ and any $\lambda > 0$, let $\mathbf{x}_\lambda = (x_1, \dots, x_{v_\lambda})$ be the points of a Poisson process on $\mathcal{S}$ with intensity measure $\lambda\mu$. In other words, v_λ has a Poisson distribution $\mathrm{Po}(\lambda\mu(\mathcal{S}))$, and, given v_λ, the points x_i are i.i.d. as in Example 4.8. Then $(\mathcal{S}, \mu, (\mathbf{x}_\lambda))$ is a generalized vertex space. Here, it is natural to write λ rather than n for an element of the index set $I = (0, \infty)$. Note that (2.4) holds because $\lambda\nu_\lambda(A) \sim \mathrm{Po}(\lambda\mu(A))$. This is the canonical example of a generalized vertex space, and one of the main reasons for allowing a random number of vertices.

Let κ be a kernel on $(\mathcal{S}, \mu)$, so, given $\mathbf{x}_\lambda$, the edge probabilities in the graph $G^{\mathcal{V}}(\lambda, \kappa)$ are given by

$$p_{ij} = \min\bigl\{\kappa(x_i, x_j)/\lambda, 1\bigr\},$$

for $1 \le i < j \le v_\lambda$. As in Example 4.8, (2.9) always holds. To see this, let $\mathcal{V}' = (\mathcal{S}, \mu', (\mathbf{y}_n)_{n\ge1})$, where $\mu' = \mu/\mu(\mathcal{S})$ is the normalized version of μ, and $\mathbf{y}_n$ consists of n i.i.d.points of $\mathcal{S}$ chosen with distribution μ'. Given that $v_\lambda = n$, the distribution of $G^{\mathcal{V}}(\lambda, \kappa)$ is exactly that of $G^{\mathcal{V}'}(n, (n/\lambda)\kappa)$. In particular, as (2.9) holds for the latter graph,

$$\begin{aligned}\mathbb{E}\bigl(e(G^{\mathcal{V}}(\lambda, \kappa)) \mid v_\lambda = n\bigr) &\sim \frac{n}{2} \iint_{\mathcal{S}^2} \frac{n\kappa}{\lambda}\, d\mu'(x)\, d\mu'(y)\\ &= \frac{n^2}{2\lambda\mu(\mathcal{S})^2} \iint_{\mathcal{S}^2} \kappa(x, y)\, d\mu(x)\, d\mu(y).\end{aligned}$$

As $\lambda \to \infty$ we have $\mathbb{E}(v_\lambda^2) \sim (\lambda\mu(\mathcal{S}))^2$, and (2.9) follows. Hence, as in Example 4.8, a kernel κ on $(\mathcal{S}, \mu)$ is graphical on $\mathcal{V}$ if and only if conditions (i) and (ii) of Definition 2.7 hold.

In this Poisson process example, it is easy to see that allowing a random number of vertices makes the model only superficially more general. Indeed, renormalizing so that $\mu(\mathcal{S}) = 1$, since $v_\lambda \sim \mathrm{Po}(\lambda)$, we can regard v_λ as a random function of λ, which is increasing (a Poisson process), and then $v_\lambda/\lambda \xrightarrow{\text{a.s.}} 1$ as $\lambda \to \infty$. It follows that if we condition on the process $v_\lambda(\lambda)$, then Theorem 3.1 applies a.s. to the corresponding graphs $G^{\mathcal{V}'}(v_\lambda, (v_\lambda/\lambda)\kappa)$ on the (ungeneralized) vertex space $\mathcal{V}'$. Thus, conditioning on $v_\lambda(\lambda)$,

$$\lambda^{-1}C_1(G^{\mathcal{V}}(\lambda, \kappa)) \xrightarrow{\text{p}} \rho(\kappa) \qquad \text{as } \lambda \to \infty \tag{4.3}$$

holds a.s. It follows that (4.3) holds unconditionally too. Other properties can be treated similarly. We shall see later, in Subsection 8.1, that all our results for generalized vertex spaces can be reduced to the vertex space case.

Example 4.10. Edge percolation. Let κ be an irreducible graphical kernel on a (generalized) vertex space $\mathcal{V}$ with $\|T_\kappa\| > 1$, and let $0 < p \le 1$. Independently of everything else, keep each edge in $G(n, \kappa)$ with probability p and delete it with probability $1 - p$. Denote the resulting graph by $G^{\langle p\rangle}(n, \kappa)$.

This random graph $G^{\langle p\rangle}(n, \kappa)$ is nothing but $G(n, \tilde\kappa_n)$, where

$$\tilde\kappa_n(x, y) := p\bigl(\kappa(x, y) \wedge n\bigr).$$

Clearly, $x_n \to x$ and $y_n \to y$ imply $\tilde\kappa_n(x_n, y_n) \to p\kappa(x, y)$, provided (x, y) is a point of continuity of κ. Furthermore, $\frac{1}{n}\,\mathbb{E}\, e\bigl(G^{\langle p\rangle}(n, \kappa)\bigr) = \frac{p}{n}\,\mathbb{E}\, e\bigl(G(n, \kappa)\bigr) \to$

$p\frac{1}{2}\iint \kappa$. Hence, $(\tilde{\kappa}_n)$ is a graphical sequence with limit $p\kappa$, so Theorem 3.1 applies with κ replaced by $p\kappa$, and

$$n^{-1}C_1\big(G^{\langle p\rangle}(n,\kappa)\big) \xrightarrow{\mathrm{p}} \rho(p\kappa).$$

In particular, $G^{\langle p\rangle}(n,\kappa)$ has whp a component of order $\Theta(n)$ if and only if $\|T_{p\kappa}\| > 1$, i.e., if $p > \|T_\kappa\|^{-1}$. Thus, as expected, we obtain the same threshold for edge percolation in $G(n,\kappa)$ (meaning that there remains a giant component) as for the existence of a giant component in $G(n,p\kappa)$; see Corollary 3.2.

Of course, the same conclusions follow if we start with the more general setting of Definition 2.9.

Example 4.11. Vertex percolation. Again, let κ be an irreducible graphical kernel on a vertex space $\mathcal{V} = (\mathcal{S},\mu,(\mathbf{x}_n))$ with $\|T_\kappa\| > 1$, and let $0 < p \le 1$. Independently of everything else, keep each vertex in $G^{\mathcal{V}}(n,\kappa)$ with probability p and delete it with probability $1-p$. Denote the resulting graph by $G^{[p]}(n,\kappa)$. This graph is again an instance of our model with a generalized vertex space. Indeed, writing $\mathbf{y}_n$ for the subsequence of $\mathbf{x}_n$ corresponding to the vertices that were not deleted, $\mathcal{V}' = (\mathcal{S},p\mu,(\mathbf{y}_n))$ is a generalized vertex space, and $G^{[p]}(n,\kappa)$ has exactly the distribution of $G^{\mathcal{V}'}(n,\kappa)$. Since the kernel κ is graphical on $\mathcal{V}$, and $\mathbb{E}\big(e\big(G^{[p]}(n,\kappa)\big)\big) = p^2\,\mathbb{E}\big(e\big(G^{\mathcal{V}}(n,\kappa)\big)\big)$, the kernel κ is also graphical on $\mathcal{V}'$, so our results apply to $G^{\mathcal{V}'}(n,\kappa)$ and hence to $G^{[p]}(n,\kappa)$.

Here, one must be a little careful with the normalization: the norm of T_κ defined with respect to $(\mathcal{S},p\mu)$ is p times $\|T_\kappa\|$, the norm defined with respect to $(\mathcal{S},\mu)$. In particular, Theorem 3.1 tells us that $G^{[p]}(n,\kappa)$ has whp a component of order $\Theta(n)$ if and only if $p\|T_\kappa\| > 1$, i.e. if $p > \|T_\kappa\|^{-1}$. We thus obtain the same threshold for vertex percolation in $G(n,\kappa)$ as for edge percolation in Example 4.10.

Once again, we could have started with the setting of Definition 2.9; we could also have started with a generalized vertex space.

Note that we can obtain the Poisson graph $\widetilde{G}_\lambda(\kappa)$ in Example 4.9 as a limit of the vertex percolation model $G^{[p]}(n,\kappa)$ in Example 4.11 if we take $p = \lambda/n$ and let $n \to \infty$.

Our next example shows that even in the supercritical, irreducible case, the second largest component may be rather large – certainly much larger than $O(\log n)$ as in the Erdős–Rényi case.

Example 4.12. Large second component. Let $\mathcal{S} = \{1,2,3,\dots\}$ with $\mu\{k\} = 2^{-k}$, and let $x_1,\dots,x_n$ be i.i.d. random points in $\mathcal{S}$ with distribution μ. Let $(\varepsilon_k)_1^\infty$ be a sequence of positive numbers tending to zero, to be chosen below. Set $\kappa(k,k) = 2^{k+1}$ for $k \ge 1$, $\kappa(1,k) = \kappa(k,1) = \varepsilon_k$ for $k \ge 2$, and $\kappa(i,j) = 0$ otherwise. Note that $\kappa \in L^1(\mathcal{S}\times\mathcal{S},\mu\times\mu)$; as noted in Example 4.8, from our choice of x_i it follows that κ is graphical on $(\mathcal{S},\mu,(\mathbf{x}_n)_{n\ge1})$.

For each $k \geq 1$, the graph $G(n, \kappa)$ contains $n_k \sim \mathrm{Bi}(n, 2^{-k})$ vertices of type k, forming a random subgraph H_k which has the distribution of the Erdős–Rényi graph $G(n_k, 2^{k+1}/n)$. Each potential edge between H_1 and H_k is present with probability ε_k/n. Note that $n_k = \bar{n}_k + O_p(\bar{n}_k^{1/2})$, where $\bar{n}_k = \mathbb{E}\, n_k = n/2^k$, and thus each H_k is (whp) supercritical. In particular, whp $C_1(G(n, \kappa)) \geq C_1(H_1) \geq cn$ for some $c > 0$, so $G(n, \kappa)$ is supercritical.

Let $k_n \to \infty$ with $\log_2 n - k_n \to \infty$, so that $\bar{n}_{k_n} \to \infty$. Let us choose the ε_k so that $\varepsilon_{k_n} \leq n^{-2}$. Then the expected number of edges between H_1 and H_{k_n} is $\mathbb{E}(n_1 n_{k_n})\varepsilon_{k_n}/n \leq n\varepsilon_{k_n} \to 0$, so whp H_{k_n} is isolated in $G(n, \kappa)$. As $n_{k_n} = \bar{n}_{k_n} + O_p(\bar{n}_{k_n}^{1/2})$, we may couple the $G(n, \kappa)$ for different n so that

$$n_{k_n} = \bar{n}_{k_n} + O(\bar{n}_{k_n}^{1/2}) \tag{4.4}$$

holds a.s. (Here the implicit constant is random.) We may then condition on n_{k_n}, assuming that n_{k_n} is deterministic, and that (4.4) holds.

Clearly, H_{k_n} is a uniform Erdős–Rényi random graph $G(n_{k_n}, 2^{k_n+1}/n)$. As $2^{k_n+1}/n \sim 2/n_{k_n}$, this graph is supercritical (for large n), and has a largest component of order $(c + o_p(1))n_{k_n}$ for some constant c. Thus,

$$C_1\big(H_{k_n}\big) = (c + o_p(1))n_{k_n} = (c + o_p(1))\bar{n}_{k_n} = (c + o_p(1))n/2^{k_n}.$$

Given any function $\omega(n)$ with $\omega(n) = o(n)$, we can choose k_n so that $2^{k_n}\omega(n)/n \to 0$; it follows that whp $C_2(G(n, \kappa)) \geq C_1(H_{k_n}) > (c/2)n/2^{k_n} > \omega(n)$. Thus, the $o_p(n)$ bound in Theorem 3.6 is best possible.

The final example in this section shows that when $\|T_\kappa\| = \infty$, the ratio of the number of edges to the number of vertices in the giant component of $G(n, c\kappa)$ need not tend to 1 as $c \to 0$. In fact, it may even tend to ∞.

Example 4.13. Dense giant component. Let $\mathcal{S}$, μ and $\mathbf{x}_n$ be as in Example 4.12, and let $\kappa(1, k) = \kappa(k, 1) = 1$ for $k \geq 1$, $\kappa(k, k) = 4^k/k^2$ for $k \geq 2$, and $\kappa(i, j) = 0$ otherwise. Again $\kappa \in L^1$, so κ is graphical. Let $c > 0$ be small but fixed and consider $G^{\mathcal{V}}(n, c\kappa)$. Let k_0 be the smallest integer such that $2^{k_0}/k_0^2 > 1/c$; taking c small enough, we may assume that $k_0 \geq 10$.

Using the notation of Example 4.12, if $k \geq 2$, then H_k forms a random subgraph of the type $G(n_k, c4^k/(k^2 n))$. Since $n_k c4^k/(k^2 n) \approx c2^k/k^2$, this subgraph is a supercritical Erdős–Rényi graph if $k \geq k_0$, and if $k \geq k_0 + 1$, classical results show that whp H_k contains a component of order $\Theta(n_k) = \Theta(2^{-k}n)$ with $\Theta(n_k^2 c4^k/(k^2 n)) = \Theta(nc/k^2)$ edges; throughout this example the implicit constants in $\Theta(\cdot)$ and $O(\cdot)$ notation do not depend on c. Each of these components is whp of order n, so they are subsets of the giant component of $G^{\mathcal{V}}(n, \kappa)$. Summing over $k = k_0 + 1, \ldots, 2k_0$, the giant component thus has at least $\Theta(nc/k_0)$ edges, so $\zeta(c\kappa) = \Omega(c/k_0)$; see Theorem 3.5.

To bound the number of vertices in the giant component, condition on $x_1, \ldots, x_n$ and say that a vertex of type k is *light* if $k \leq k_0 - 3$, and *heavy* otherwise. The total number of heavy vertices is $O(n2^{-k_0})$ whp. Furthermore, it is easy to check that if c is small enough, then the expected number of edges to light vertices from each heavy vertex is at most $1/2$, as is the expected degree of each light vertex. Each light vertex in the giant component has to be connected to some heavy vertex by a path whose other vertices all are light. As the expected number of such paths starting at a given heavy vertex is at most $\sum_{l\geq 1}(1/2)^l = 1$, the expected number of light vertices in the giant component is $O(n2^{-k_0})$ too. Hence, the number of vertices in the giant component is $O_p(n2^{-k_0})$, so $\rho(c\kappa) = O(2^{-k_0}) = O(c/k_0^2)$. Consequently, $\zeta(c\kappa)/\rho(c\kappa) = \Omega(k_0) = \Omega(\log(1/c))$, as $c \to 0$. In particular, $\zeta(c\kappa)/\rho(c\kappa) \to \infty$ as $c \searrow c_0 = 0$; see Remark 3.19.

5. Branching process lemmas

In this section and the next we study the Poisson branching processes $\mathfrak{X}_\kappa(x)$ and $\mathfrak{X}_\kappa$ defined in Subsection 2.1, and their survival probabilities. These turn out to be given by the solutions to a certain non-linear functional equation (5.3). Let us briefly recall some definitions.

Let $(\mu, \mathcal{S})$ be a (generalized) ground space. The branching process $\mathfrak{X}_\kappa(x)$ is a multi-type Galton–Watson branching processes with type space $\mathcal{S}$: a particle of type $y \in \mathcal{S}$ is replaced in the next generation by its 'children', a set of particles whose types are distributed as a Poisson process on $\mathcal{S}$ with intensity $\kappa(y,z)\,d\mu(z)$. The zeroth generation of $\mathfrak{X}_\kappa(x)$ consists of a single particle of type x. Note that the distribution of $\mathfrak{X}_\kappa(x)$ is unaffected if we multiply κ by a constant and divide μ by the same constant; thus, we may assume without loss of generality that $\mu(\mathcal{S}) = 1$. We shall make this assumption throughout this section. In this normalized case, the branching process $\mathfrak{X}_\kappa$ is just the process $\mathfrak{X}_\kappa(x)$ started with a single particle whose (random) type is distributed according to the probability measure μ.

Here, we have no need for the metric or topological structure of $\mathcal{S}$; in this section and the next, $\mathcal{S}$ can be any measurable space equipped with a probability measure μ. We assume, as before, that the kernel κ is a measurable symmetric non-negative function on $\mathcal{S}^2$. We shall also assume that $\kappa \in L^1(\mathcal{S}\times\mathcal{S}, \mu\times\mu)$, i.e., that $\iint \kappa < \infty$.

Let us recall our notation for the survival probabilities of particles in $\mathfrak{X}_\kappa(x)$. We write $\rho_k(\kappa; x)$ for the probability that the total population consists of exactly k particles, and $\rho_{\geq k}(\kappa; x)$ for the probability that the total population contains at least k particles. Furthermore, $\rho(\kappa; x)$ is the probability that the branching process survives for eternity.

We write $\rho_k(\kappa)$, $\rho_{\geq k}(\kappa)$ and $\rho(\kappa)$ for the corresponding probabilities for $\mathfrak{X}_\kappa$, so that, e.g., $\rho_k(\kappa) = \int_{\mathcal{S}} \rho_k(\kappa; x)\,d\mu(x)$.

We start with a trivial observation that will enable us to eliminate certain pathologies.

Lemma 5.1. *If $\kappa = \kappa'$ a.e., then $\rho(\kappa; x) = \rho(\kappa'; x)$ and $\rho_{\geq k}(\kappa; x) = \rho_{\geq k}(\kappa'; x)$ hold for a.e. x; hence $\rho(\kappa) = \rho(\kappa')$ and $\rho_{\geq k}(\kappa) = \rho_{\geq k}(\kappa')$.*

Proof. There is a measure zero set $N \subset \mathcal{S}$ such that if $x \notin N$, then $\kappa(x,y) = \kappa'(x,y)$ for a.e. y. It follows that if we start the processes $\mathfrak{X}_\kappa$ and $\mathfrak{X}_{\kappa'}$ at the same $x \notin N$, the processes will be identical in distribution. Hence $\rho(\kappa; x) = \rho(\kappa'; x)$ and $\rho_{\geq k}(\kappa; x) = \rho_{\geq k}(\kappa'; x)$ for all $x \notin N$, and the result follows from (2.13). $\square$

For the sake of convenience, in this section we impose one more assumption on κ, namely that

$$\int_{\mathcal{S}} \kappa(x, y)\, d\mu(y) < \infty \tag{5.1}$$

for *every* $x \in \mathcal{S}$. This assumption loses no generality, as (5.1) holds for a.e. x, since $\iint \kappa < \infty$. Writing N for the measure zero set of x such that (5.1) does not hold, define $\bar{\kappa}$ by setting $\bar{\kappa}(x, y) = 0$ if $x \in N$ or $y \in N$, and $\bar{\kappa}(x, y) = \kappa(x, y)$ otherwise. Then $\kappa = \bar{\kappa}$ a.e., so by Lemma 5.1 we have $\rho(\bar{\kappa}) = \rho(\kappa)$ and so on.

All the assumptions above apply to all the kernels considered below, denoted κ_1, κ', etc. In this section, unless explicitly stated, we do *not* assume that κ is irreducible.

Remark 5.2. Condition (5.1) means that a particle of type x has a finite number of children in the branching process. As we are assuming (5.1) for all x, a particle survives for eternity (has descendants in all future generations) if and only if it has infinitely many descendants. In other words, $\rho(\kappa; x) = \rho_\infty(\kappa; x)$.

Remark 5.3. Our process is a very special branching process since we assume that the children of a particle are distributed according to a Poisson process; in particular, the number of children has a Poisson distribution. Other branching processes, and other functional equations, appear when studying random graphs with dependencies between edges, as in [22, 25, 84], but will not be considered here.

Note also that even with the Poisson assumption, our processes are special. For the branching process, there is no reason to assume κ to be symmetric; moreover, μ may be any σ-finite measure, and the hypothesis $\kappa \in L^1$ could be weakened to (5.1) for a.e. x (or perhaps removed completely). We shall, however, consider only the special case just defined; this will be useful in the proofs. We have not yet investigated to what extent the results generalize and remark only that in non-symmetric situations, the norm $\|T_\kappa\|$ should be replaced by the spectral radius.

There is an abundant literature on branching processes with different types; see, for example, the book by Mode [76]. However, we have not found the results we need in the generality required here, so for the sake of completeness we give full proofs, although the results are only minor extensions of known results; see, for example, [76, Chapter 6].

We start with the connection between our branching process and the operator Φ_κ defined in (2.17).

Lemma 5.4. *Consider the random offspring of a single particle of type x; let N be the number of children, and denote their types by $(\xi_i)_{i=1}^N$. If g is a measurable function on $\mathcal{S}$ with $0 \le g \le 1$, then*

$$\mathbb{E}\prod_{i=1}^{N}\bigl(1-g(\xi_i)\bigr) = e^{-(T_\kappa g)(x)} = 1-(\Phi_\kappa g)(x). \tag{5.2}$$

Proof. This is a standard formula for Poisson processes; see, for example, Kallenberg [61, Lemma 12.2(i)], taking $f = -\ln(1-g)$. For completeness, we include the simple proof. Indeed, let $\nu = \nu_x$ be the measure defined by $d\nu(y) = \kappa(x,y)\,d\mu(y)$. Then $N \sim \mathrm{Po}(\nu(\mathcal{S}))$ and, given N, the types ξ_i of the children are i.i.d. with the renormalized distribution $\nu' = \nu/\nu(\mathcal{S})$. Hence, given N, the conditional expectation of $\prod_{i=1}^N\bigl(1-g(\xi_i)\bigr)$ is just

$$\begin{aligned}\prod_{i=1}^{N}\mathbb{E}\bigl(1-g(\xi_i)\bigr) &= \bigl(\mathbb{E}(1-g(\xi_1))\bigr)^N \\ &= \left(1-\int_{\mathcal{S}} g(y)\,d\nu'(y)\right)^N = \bigl(1-(T_\kappa g)(x)/\nu(\mathcal{S})\bigr)^N.\end{aligned}$$

Using $\mathbb{P}(N=n) = e^{-\nu(\mathcal{S})}\nu(\mathcal{S})^n/n!$ and taking the expectation, the result follows. ∎

Our next aim is to study the fixed points of Φ_κ, i.e., the solutions of the equation

$$f = \Phi_\kappa f := 1-e^{-T_\kappa f}, \tag{5.3}$$

where f is a non-negative function on $\mathcal{S}$.

Remark 5.5. If $f = g$ a.e., then $\Phi_\kappa f = \Phi_\kappa g$. In particular, if $f = \Phi_\kappa f$ a.e., then $\Phi_\kappa f = \Phi_\kappa(\Phi_\kappa f)$; thus, if f satisfies (5.3) a.e., then there is a solution $\bar f$ to (5.3) (viz. $\Phi_\kappa f$) such that $f = \bar f$ a.e. This shows that it makes no essential difference if we require (5.3) to hold only a.e. (which might be natural from an L^2 perspective). We shall, however, find it convenient to interpret (5.3) and similar relations as holding everywhere unless we explicitly state otherwise. Similarly, if $\kappa = \bar\kappa$ a.e., then for any solution f to $f = \Phi_\kappa f$ there is a unique $\bar f$ with the properties that $\bar f = \Phi_{\bar\kappa}\bar f$ and $\bar f = f$ a.e.

Note that Φ_κ is monotone: if $0 \le f \le g$ a.e. then $T_\kappa f \le T_\kappa g$ and thus $\Phi_\kappa f \le \Phi_\kappa g$.

In the lemma below, 1 denotes the function with constant value 1.

Lemma 5.6. (i) *For $m \ge 0$ the probability that a particle of type x has descendants in at least m further generations is $(\Phi_\kappa^m 1)(x)$.*

(ii) *As $m \to \infty$, $(\Phi_\kappa^m 1)(x) \searrow \rho(\kappa; x)$.*

(iii) *The function $\rho_\kappa = \rho_\kappa(x) = \rho(\kappa; x)$ is a solution of* (5.3)*, i.e., satisfies $\Phi_\kappa\rho_\kappa = \rho_\kappa$.*

(iv) *The function ρ_κ is the maximum solution of* (5.3)*: if f is any other solution, then $\rho_\kappa(x) \ge f(x)$ for every x.*

Proof. (i) Let $g_m(x)$ be this probability. Then, with $g = g_m$, the left-hand side of (5.2) is the probability that none of the children of x has descendants in at least m generations, i.e., the probability $1 - g_{m+1}(x)$ that x does not have descendants in $m + 1$ generations. Thus $g_{m+1} = \Phi_\kappa g_m$, and the result follows by induction, since $g_0(x) = 1$.

(ii) An immediate consequence of (i).

(iii) This follows by the same argument as (i) (and is also a consequence of (ii) and dominated convergence).

(iv) Suppose that f is a solution of (5.3). Then $f = \Phi_\kappa f \leq 1$, and thus $f = \Phi_\kappa^m f \leq \Phi_\kappa^m 1$ for every m. Hence, $f \leq \rho_\kappa$ follows from (ii). □

Remark 5.7. If we do not impose (5.1) for all x, then (iii), i.e., $(\Phi_\kappa \rho_\kappa)(x) = \rho_\kappa(x)$, could fail for x in the measure zero set for which (5.1) does not hold. This is because a particle of type x has infinitely many children, which may have finite but unbounded lines of descendants; for an example, take $\mathcal{S} = (0, 1]$ and $\kappa = 1$ except that $\kappa(x, 1) = \kappa(1, x) = 1/x$.

We continue to study the functional equation (5.3).

Lemma 5.8. *Suppose that $f \geq 0$ with $f = \Phi_\kappa f$. Then*

(i) $0 \leq f < 1$;

(ii) $T_\kappa f \geq f$, *with strict inequality when* $f(x) > 0$;

(iii) $T_\kappa f \leq f/(1 - f)$, *with strict inequality when* $f(x) > 0$;

(iv) *if κ is irreducible, then either $f = 0$ everywhere or $f > 0$ a.e.*

Proof. (i) We have $f(x) = 1 - e^{-(T_\kappa f)(x)} \leq 1$. Hence, $(T_\kappa f)(x) \leq (T_\kappa 1)(x) = \int_{\mathcal{S}} \kappa(x, y)\, d\mu(y) < \infty$, where the second inequality is just our assumption (5.1). Therefore, $f(x) = 1 - e^{-(T_\kappa f)(x)} < 1$.

(ii) This is immediate from $f = 1 - e^{-T_\kappa f} \leq T_\kappa f$, with equality only when $T_\kappa f = 0$.

(iii) We have $e^{-T_\kappa f} = 1 - f$, and thus, as $f < 1$,

$$T_\kappa f \leq e^{T_\kappa f} - 1 = \frac{1}{1 - f} - 1 = \frac{f}{1 - f},$$

with equality only when $T_\kappa f = 0$.

(iv) Let $A := \{x \in \mathcal{S} : f(x) = 0\}$. For $x \in A$, $(\Phi_\kappa f)(x) = f(x) = 0$, and thus $(T_\kappa f)(x) = 0$. Hence $\kappa(x, y) = 0$ for a.e. $y \notin A$. Consequently, $\kappa = 0$ a.e. on $A \times (\mathcal{S} \setminus A)$, which by (2.12) implies $\mu(A) = 0$ or $\mu(A) = 1$. In the latter case, $f = 0$ a.e., and thus $f = \Phi_\kappa f = 0$. □

In the next two lemmas we consider irreducible κ.

Lemma 5.9. *Suppose that κ is irreducible. Suppose further that $f = \Phi_\kappa f$ and $g = \Phi_\kappa g$ with $0 \leq f \leq g$. Then either $f = 0$ or $f = g$.*

Proof. By Lemma 5.8(iv) we may assume that $f > 0$ a.e.

Let $h = (g - f)/2 \geq 0$; thus $f + h = (f + g)/2$. The function $t \mapsto 1 - e^{-t}$ is strictly concave; in particular, $1 - e^{-(t+u)/2} \geq \frac{1}{2}\big((1 - e^{-t}) + (1 - e^{-u})\big)$.

Hence,

$$\begin{aligned}\Phi_\kappa\Big(\frac{f+g}{2}\Big) &= 1-e^{-T_\kappa((f+g)/2)} = 1-e^{-(T_\kappa f+T_\kappa g)/2}\\ &\geq \tfrac12\Big(\big(1-e^{-T_\kappa f}\big)+\big(1-e^{-T_\kappa g}\big)\Big) = \tfrac12(f+g) = f+h,\end{aligned} \tag{5.4}$$

with strict inequality at every point where $f < g$ and thus $\Phi_\kappa f < \Phi_\kappa g$ and $T_\kappa f < T_\kappa g$. On the other hand,

$$\begin{aligned}1-\Phi_\kappa\Big(\frac{f+g}{2}\Big) &= e^{-T_\kappa(f+h)} = e^{-T_\kappa f}e^{-T_\kappa h} = (1-f)e^{-T_\kappa h}\\ &\geq (1-f)(1-T_\kappa h).\end{aligned} \tag{5.5}$$

Combining (5.4) and (5.5), we find that

$$(1-f)(1-T_\kappa h) \leq 1-(f+h) = 1-f-h$$

and thus

$$(1-f)T_\kappa h \geq h \tag{5.6}$$

with strict inequality when $g > f$.

Suppose now that $g > f$ on a set of positive measure. Then, inequality (5.6), the fact that $f > 0$ a.e., and Lemma 5.8(iii) imply that

$$\int_{\mathcal S} f\,T_\kappa h\,d\mu > \int_{\mathcal S} f\frac{h}{1-f}\,d\mu = \int_{\mathcal S} h\frac{f}{1-f}\,d\mu \geq \int_{\mathcal S} h\,T_\kappa f\,d\mu. \tag{5.7}$$

Note that the integrals above are finite because $\kappa \in L^1$ and $f, h \leq 1$. However, as κ is symmetric, T_κ is a symmetric operator, and so $\int_{\mathcal S} f\,T_\kappa h\,d\mu = \int_{\mathcal S} h\,T_\kappa f\,d\mu$, contradicting (5.7). This shows that $g = f$ a.e. and thus $f = \Phi_\kappa f = \Phi_\kappa g = g$. □

Lemma 5.10. *Suppose that κ is irreducible. Then $f = 0$ and $f = \rho_\kappa$ are the only solutions to* (5.3)*; these solutions may coincide.*

Proof. By parts (iii) and (iv) of Lemma 5.6, the function ρ_κ is a solution of (5.3), and $0 \leq f \leq \rho_\kappa$ for every solution f of (5.3). The result follows by Lemma 5.9. □

It remains to decide whether $\rho_\kappa = 0$ or not. Recall that $\|T_\kappa\|$ is defined in (2.15). We shall show that $\rho_\kappa = 0$ if and only if $\|T_\kappa\| \leq 1$.

Lemma 5.11. *If $\|T_\kappa\| \leq 1$, then $\rho_\kappa = 0$.*

Proof. Suppose that f is a solution of (5.3), and that we do not have $f = 0$ a.e. Lemma 5.8(ii) implies that $T_\kappa f \geq f$, with $T_\kappa f > f$ on a set of positive measure, and hence that $\|T_\kappa f\|_2 > \|f\|_2$, contradicting $\|T_\kappa\| \leq 1$. Consequently, if f is a solution of (5.3), then $f = 0$ a.e., and thus $f = \Phi_\kappa f = 0$, so the only solution is $f = 0$. In particular, $\rho_\kappa = 0$ since ρ_κ is a solution by Lemma 5.6. □

It remains to show that if $\|T_\kappa\| > 1$, then ρ_κ is not identically zero. We proceed in several steps.

Lemma 5.12. *If $f \geq 0$ and $\Phi_\kappa f \geq f$, then $\Phi_\kappa^m f \nearrow g$ as $m \to \infty$, for some $g \geq f \geq 0$ with $\Phi_\kappa g = g$.*

Proof. By induction, $f \leq \Phi_\kappa f \leq \Phi_\kappa^2 f \leq \dots$. Since $0 \leq \Phi_\kappa^m f \leq 1$, the limit $g(x) := \lim_{m\to\infty}(\Phi_\kappa^m f)(x)$ exists for every x, and $g \geq 0$. Monotone convergence yields

$$(T_\kappa g)(x) = \lim_{m\to\infty} \int_{\mathcal{S}} \kappa(x,y)(\Phi_\kappa^m f)(y)\, d\mu(y) = \lim_{m\to\infty}(T_\kappa(\Phi_\kappa^m f))(x)$$

and thus

$$(\Phi_\kappa g)(x) = \lim_{m\to\infty}(\Phi_\kappa(\Phi_\kappa^m f))(x) = g(x).$$

□

Lemma 5.13. *If there is a bounded function $f \geq 0$, not a.e. 0, such that $T_\kappa f \geq (1+\delta) f$ for some $\delta > 0$, then $\rho_\kappa > 0$ on a set of positive measure.*

Proof. Let $M = \sup f < \infty$. Fix $\varepsilon > 0$ with $(1 - M\varepsilon)(1+\delta) \geq 1$. Since $-\log(1-x) \leq x/(1-x)$ we have

$$-\ln(1-\varepsilon f) \leq \frac{1}{1-\varepsilon M}\varepsilon f \leq (1+\delta)\varepsilon f \leq \varepsilon T_\kappa f,$$

and thus

$$\Phi_\kappa(\varepsilon f) = 1 - e^{-\varepsilon T_\kappa f} \geq 1 - (1 - \varepsilon f) = \varepsilon f.$$

By Lemma 5.12, there exists a solution g to $\Phi_\kappa g = g$ with $g \geq \varepsilon f$, and thus g not a.e. 0. By part (iv) of Lemma 5.6, we have $\rho_\kappa \geq g$. □

Remark 5.14. The proof of Lemma 5.13 shows that $\rho_\kappa \geq \frac{\delta}{1+\delta}\frac{f}{\sup f}$. In particular, this immediately implies Theorem 10 of [19].

We should like to find an eigenfunction of T_κ with eigenvalue greater than 1, so that we can apply Lemma 5.13. If the Hilbert–Schmidt norm of T_κ (see (2.16)) is finite, then a standard result gives us such an eigenfunction.

Lemma 5.15. *If $\|T_\kappa\|_{HS} < \infty$, then T_κ is compact and has an eigenfunction $\psi \in L^2(\mathcal{S})$, $\psi \geq 0$, with eigenvalue $\|T_\kappa\|$.*

If, in addition, κ is irreducible, then $\psi > 0$ a.e., and every eigenfunction with eigenvalue $\|T_\kappa\|$ is a multiple of ψ.

Proof. Suppose that $\|T_\kappa\|_{HS} < \infty$. It is well-known (see e.g. [17, XIV.6, p. 202]) that T_κ is then compact, and so has an eigenfunction $g \in L^2$ with eigenvalue of modulus $\lambda := \|T_\kappa\|$. Then

$$T_\kappa|g| \geq |T_\kappa g| = \lambda|g| \qquad \text{a.e.},$$

and since $\|T_\kappa\| = \lambda$ we must have $T_\kappa|g| = \lambda|g|$ a.e. Hence $\psi := |g|$ is an eigenfunction with eigenvalue $\lambda = \|T_\kappa\|$.

Now suppose that κ is irreducible, with $\|T_\kappa\|_{HS} < \infty$, and let h be any (real) function in L^2 with $T_\kappa h = \lambda h$ a.e. By the argument above, $T_\kappa|h| = \lambda|h|$ a.e. holds as well. Let $A := \{|h| = 0\}$. Then $T_\kappa|h| = \lambda|h| = 0$ a.e. on A, so $\kappa = 0$ a.e. on $A \times (\mathcal{S} \setminus A)$ and (2.12) yields $\mu(A) = 0$ or 1. Hence either $h = 0$ a.e. or $h \neq 0$ a.e. In particular, taking $h = g$ we see that $\psi > 0$ a.e.

Returning to a general h satisfying $T_\kappa h = \lambda h$ a.e., as $T_\kappa(|h| + h) = \lambda(|h| + h)$ a.e. by linearity, we can apply the argument above to $|h| + h$, deducing that either $h > 0$ a.e. or $h \le 0$ a.e. Finally, applying this to $h - a\psi$, with a chosen such that $\int (h - a\psi)\, d\mu = 0$, we see that $h - a\psi = 0$ a.e. □

The second part of Lemma 5.15 will be needed only in Section 15.

After this preparation, it is easy to show that if $\|T_\kappa\| > 1$ then $\rho_\kappa > 0$ on a set of positive measure.

Lemma 5.16. *If $1 < \|T_\kappa\| \le \infty$, then $\rho_\kappa > 0$ on a set of positive measure. Thus* (5.3) *has at least one non-zero solution.*

Proof. Since $\|T_\kappa\| > 1$, there is function $f \in L^2$ with $\|f\|_2 = 1$ and $\|T_\kappa f\|_2 > 1$. As $T_\kappa |f| \ge |T_\kappa f|$, we may assume that $f \ge 0$. Let T_N be the integral operator on $\mathcal{S}$ with the truncated kernel $\kappa_N(x, y) := \kappa(x, y) \wedge N$, $N \ge 1$. By monotone convergence, $T_N f \nearrow T_\kappa f$ as $N \to \infty$, and thus $\|T_N f\|_2 \nearrow \|T_\kappa f\|_2 > 1$. We can thus choose an N such that $\|T_N f\|_2 > 1 = \|f\|_2$, and thus $\|T_N\| > 1$. Set $\delta = \|T_N\| - 1 > 0$.

Since the kernel κ_N is bounded and μ is a finite measure, by Lemma 5.15 T_N has an eigenfunction $\psi \in L^2(\mathcal{S})$ with $\psi \ge 0$ and

$$T_N \psi = \|T_N\| \psi = (1 + \delta)\psi. \tag{5.8}$$

Since the kernel κ_N is bounded, it follows that $T_N\psi$ is a bounded function. Indeed, $(T_N\psi)(x) \le N \int_S \psi\, d\mu = N\|\psi\|_1 \le N\|\psi\|_2 < \infty$. From (5.8) it follows that ψ is bounded.

Since $\kappa \ge \kappa_N \ge 0$, we have, using (5.8) again,

$$T_\kappa \psi \ge T_N \psi = (1 + \delta)\psi,$$

and the result follows by Lemma 5.13. □

The final lemma of this section will enable us to reduce the reducible case to the irreducible one.

Lemma 5.17. *Let κ be a symmetric measurable function on $\mathcal{S} \times \mathcal{S}$. Then there exists a partition $\mathcal{S} = \bigcup_{i=0}^N \mathcal{S}_i$ with $0 \le N \le \infty$ such that each $\mathcal{S}_i$ is measurable, $\mu(\mathcal{S}_i) > 0$ for $i \ge 1$, the restriction of κ to $\mathcal{S}_i \times \mathcal{S}_i$ is irreducible for each $i \ge 1$, and $\kappa = 0$ a.e. on $(\mathcal{S} \times \mathcal{S}) \setminus \bigcup_{i=1}^N (\mathcal{S}_i \times \mathcal{S}_i)$.*

Note that $\kappa = 0$ a.e. on $\mathcal{S}_0 \times \mathcal{S}_0$.

Proof. Let $\mathcal{G}$ be the family of all measurable subsets $A \subseteq \mathcal{S}$ such that $\kappa = 0$ a.e. on $A \times (\mathcal{S} \setminus A)$. It is easily verified that $\mathcal{G}$ is a σ-field; thus $(\mathcal{S}, \mathcal{G}, \mu)$ is a finite measure space. Hence there exists a partition $\mathcal{S} = \bigcup_{i=0}^N \mathcal{S}_i$ with $0 \le N \le \infty$ and each $\mathcal{S}_i \in \mathcal{G}$ such that for each $i \ge 1$ the set $\mathcal{S}_i$ is an atom in $(\mathcal{S}, \mathcal{G}, \mu)$ with positive measure, while $\mathcal{S}_0$ is non-atomic, i.e., contains no atoms with non-zero measure. (We allow $\mathcal{S}_0 = \emptyset$.) Here '$\mathcal{S}_i$ is an atom' means that if $A \subseteq \mathcal{S}_i$ with $A \in \mathcal{G}$, then $\mu(A) = 0$ or $\mu(A) = \mu(\mathcal{S}_i)$; this is equivalent to (2.12), so κ is irreducible on $\mathcal{S}_i \times \mathcal{S}_i$ for each $i \ge 1$.

Finally, since $\mathcal{S}_0$ is non-atomic, for every positive integer M there exists a partition $\mathcal{S}_0 = \bigcup_{j=1}^{M} T_j$ with $T_j \in \mathcal{G}$ and $\mu(T_j) = \mu(\mathcal{S}_0)/M$. Then $\kappa = 0$ a.e. on $T_i \times T_j$ when $i \neq j$, and thus

$$(\mu\times\mu)\big\{(x,y) \in \mathcal{S}_0\times\mathcal{S}_0 : \kappa(x,y) \neq 0\big\} \leq (\mu \times \mu)\Big(\bigcup_{j=1}^{M}(T_j\times T_j)\Big) = M\Big(\frac{\mu(\mathcal{S}_0)}{M}\Big)^2.$$

Letting $M \to \infty$, we see that $\kappa = 0$ a.e. on $\mathcal{S}_0 \times \mathcal{S}_0$. □

Remark 5.18. One application of Lemma 5.17 is a generalization of Lemma 5.10 to arbitrary κ. With $\mathcal{S}_i$ as in Lemma 5.17, let J be the set of indices i such that the restriction of the operator T_κ to $L^2(\mathcal{S}_i)$ has norm strictly greater than 1. Then there are $2^{|J|}$ solutions of (5.3), where $0 \leq |J| \leq \infty$: for every subset $J' \subseteq J$, there is exactly one solution that equals ρ_κ a.e. on $\bigcup_{i\in J'} \mathcal{S}_i$ and vanishes a.e. elsewhere. (This is easily seen using the argument in the proof of Theorem 6.4 below.)

6. Branching process results

In this section we collect the branching process results we shall use. These are all simple consequences of the lemmas in the previous section. In this section, κ will always be a kernel on a measure space $(\mathcal{S}, \mu)$, i.e., a symmetric non-negative measurable function on $\mathcal{S} \times \mathcal{S}$. Unless explicitly stated otherwise, μ will be a probability measure, i.e., $\mu(\mathcal{S}) = 1$. We shall assume that $\kappa \in L^1$; as noted in the previous section, it follows that (5.1) holds a.e. x. We do not assume that (5.1) holds for every x except when explicitly stated.

Theorem 6.1. *Suppose that κ is a kernel on the space $(\mathcal{S}, \mu)$, that $\kappa \in L^1$, and that* (5.1) *holds for every x. Then the function ρ_κ defined by $\rho_\kappa(x) = \rho(\kappa; x)$ is the maximum solution of* (5.3)*. Furthermore:*

(i) *If $\|T_\kappa\| \leq 1$, then $\rho(\kappa; x) = 0$ for every x, and* (5.3) *has only the zero solution.*
(ii) *If $1 < \|T_\kappa\| \leq \infty$, then $\rho(\kappa; x) > 0$ on a set of positive measure. If, in addition, κ is irreducible, then $\rho(\kappa; x) > 0$ for a.e. x, and $\rho(\kappa; x)$ is the only non-zero solution of* (5.3).

In particular, $\rho(\kappa) > 0$ if and only if $\|T_\kappa\| > 1$.

Proof. The first statement is just part (iv) of Lemma 5.6. The remaining statements follow directly from Lemmas 5.11, 5.16, Lemma 5.8(iv) and 5.10. □

The next result is essentially a restatement of Theorem 6.1, in the setting of the results in Section 3. Thus, μ will not necessarily be a probability measure, and we shall not require that (5.1) holds; this makes very little difference. This result gives the promised characterization of $\rho(\kappa; x)$ and $\rho(\kappa)$ in terms of a functional equation, in the full generality of the setting of Theorem 3.1.

Recall that (2.17) defines Φ_κ only for non-negative functions; we thus consider only non-negative solutions to (6.1) below.

Theorem 6.2. *Let κ be a kernel on a (generalized) ground space $(\mathcal{S},\mu)$, with $\kappa \in L^1(\mathcal{S}\times\mathcal{S},\mu\times\mu)$. There is a (necessarily unique) maximum solution $\tilde\rho_\kappa$ to*

$$\Phi_\kappa(\tilde\rho_\kappa) = \tilde\rho_\kappa, \tag{6.1}$$

i.e., a solution that pointwise dominates all other solutions. Furthermore, $\rho(\kappa;x) = \tilde\rho_\kappa(x)$ for a.e. x, and

$$\Phi_\kappa(\rho_\kappa) = \rho_\kappa \quad a.e., \tag{6.2}$$

where the function ρ_κ is defined by $\rho_\kappa(x) := \rho(\kappa;x)$.

If $\|T_\kappa\| \le 1$, then $\tilde\rho_\kappa$ is identically zero, and this is thus the only solution to (6.1). If $\|T_\kappa\| > 1$, then $\tilde\rho_\kappa$ is positive on a set of positive measure. Thus $\rho(\kappa) > 0$ if and only if $\|T_\kappa\| > 1$.

If $\|T_\kappa\| > 1$ and κ is irreducible, then $\tilde\rho_\kappa$ is the unique non-zero solution to (6.1), and $\tilde\rho_\kappa = \rho_\kappa > 0$ a.e.

Theorem 6.2 follows almost immediately from Theorem 6.1 and Lemma 5.1.

Proof. Multiplying κ by a constant factor and dividing μ by the same constant factor does not affect the definition of the branching process $\mathfrak{X}_\kappa(x)$. Hence, the function ρ_κ is not affected by this rescaling. As the operators T_κ and Φ_κ are also unchanged, we may assume without loss of generality that $\mu(\mathcal{S}) = 1$. As noted in Section 5, since $\kappa \in L^1$ there is a kernel $\bar\kappa$ with $\bar\kappa = \kappa$ a.e., such that (5.1) holds for $\bar\kappa$ for every x. Applying Theorem 6.1 to the kernel $\bar\kappa$, the result follows by Lemma 5.1 and Remark 5.5. □

We now study monotonicity and continuity properties of $\rho(\kappa;x)$ and $\rho(\kappa)$ when κ is varied. For the rest of the section, we assume that $\mu(\mathcal{S}) = 1$. As usual, we say that a sequence of functions f_n *increases (a.e.)* to a function f if for every x (a.e. x) the sequence $f_n(x)$ is monotone increasing and converges to $f(x)$. As before, we write ρ_κ for the function given by $\rho_\kappa(x) := \rho(\kappa;x)$. We start with a trivial lemma.

Lemma 6.3. *If $\kappa_1 \le \kappa_2$, then $\rho_{\kappa_1} \le \rho_{\kappa_2}$.*

Proof. Immediate by coupling the branching processes. □

Theorem 6.4. (i) *Let $(\kappa_n)_1^\infty$ be a sequence of kernels on $(\mathcal{S},\mu)$ increasing a.e. to κ. Then $\rho_{\kappa_n} \nearrow \rho_\kappa$ for a.e. x and $\rho(\kappa_n) \nearrow \rho(\kappa)$.*

(ii) *Let $(\kappa_n)_1^\infty$ be a sequence of kernels on $(\mathcal{S},\mu)$ decreasing a.e. to κ. Then $\rho_{\kappa_n} \searrow \rho_\kappa$ for a.e. x and $\rho(\kappa_n) \searrow \rho(\kappa)$.*

Proof. On the measure zero set where $\kappa_n \not\to \kappa$, redefine all κ_n and κ to be 0. By Lemma 5.1, this does not affect the conclusions, so we may assume $\kappa_n \nearrow \kappa$ or $\kappa_n \searrow \kappa$ everywhere. Similarly, we may assume that (5.1) holds for every x, for each κ_n and for κ. It suffices to prove the conclusions for ρ_κ: the conclusions for $\rho(\kappa)$ follow from (2.13) and dominated convergence.

(i) We choose a partition $\mathcal{S} = \bigcup_{i=0}^{N} \mathcal{S}_i$ as in Lemma 5.17, and redefine κ and all κ_n to be 0 on $(\mathcal{S}\times\mathcal{S})\setminus\bigcup_{i=1}^{N}(\mathcal{S}_i\times\mathcal{S}_i)$; this only changes the kernels on a set of measure zero, so we may again apply Lemma 5.1. Now $\rho_{\kappa_n} = \rho_\kappa = 0$ on $\mathcal{S}_0$. We may consider each $\mathcal{S}_i$, $i \geq 1$, separately, and we may thus assume without loss of generality that κ is irreducible. The only problem is that the restriction of μ to $\mathcal{S}_i$ does not have total mass 1, but this is not a real problem, and can be handled by renormalizing, i.e., dividing the measure by $\mu(\mathcal{S}_i)$ and multiplying all kernels by the same factor; as remarked earlier, this operation does not affect the branching process.

We have shown that we may assume that κ is irreducible; let us do so. By Lemma 6.3, if $m \leq n$, then $\rho_{\kappa_m} \leq \rho_{\kappa_n}$. Thus (ρ_{κ_n}) is an increasing sequence of functions, all bounded by 1, so the limit $\rho^*_\kappa(x) := \lim_{n\to\infty}\rho_{\kappa_n}(x)$ exists everywhere. By monotone convergence,

$$(T_\kappa\rho^*_\kappa)(x) = \int_{\mathcal{S}} \kappa(x,y)\rho^*_\kappa(y)\,d\mu(y) = \lim_{n\to\infty}\int_{\mathcal{S}} \kappa_n(x,y)\rho_{\kappa_n}(y)\,d\mu(y)$$
$$= \lim_{n\to\infty}(T_{\kappa_n}\rho_{\kappa_n})(x),$$

so $\Phi_\kappa\rho^*_\kappa = \lim_{n\to\infty}\Phi_{\kappa_n}\rho_{\kappa_n} = \lim_{n\to\infty}\rho_{\kappa_n} = \rho^*_\kappa$. Hence, by Lemma 5.10, either $\rho^*_\kappa = \rho_\kappa$, and we are done, or $\rho^*_\kappa = 0$. In the latter case, each $\rho_{\kappa_n} = 0$, and thus, by Lemma 5.16, $\|T_{\kappa_n}\| \leq 1$.

Hence, if $f \in L^2$ with $f \geq 0$ and $\|f\|_2 \leq 1$, then $\|T_{\kappa_n}f\|_2 \leq 1$. Monotone convergence shows that, as $n\to\infty$, $T_{\kappa_n}f \nearrow T_\kappa f$ and $\|T_{\kappa_n}f\|_2 \nearrow \|T_\kappa f\|_2$. Consequently, $\|T_\kappa f\|_2 \leq 1$ for each such f, and thus $\|T_\kappa\| \leq 1$. By Theorem 6.1, $\rho_\kappa = 0$ in this case, so $\rho_\kappa = \rho^*_\kappa$ in this case too.

(ii) This is similar. Now (ρ_{κ_n}) is a decreasing sequence of functions, and $\rho^*_\kappa(x) := \lim_{n\to\infty}\rho_{\kappa_n}(x)$ still exists everywhere. By dominated convergence, $(T_\kappa\rho^*_\kappa)(x) = \lim_{n\to\infty}(T_{\kappa_n}\rho_{\kappa_n})(x)$, so $\Phi_\kappa\rho^*_\kappa = \lim_{n\to\infty}\Phi_{\kappa_n}\rho_{\kappa_n} = \lim_{n\to\infty}\rho_{\kappa_n} = \rho^*_\kappa$. In other words, ρ^*_κ satisfies (5.3). Furthermore, by Lemma 6.3 again, $\rho_{\kappa_n} \geq \rho_\kappa$, so $\rho^*_\kappa \geq \rho_\kappa$. Since ρ_κ is a maximal solution to (5.3) by Lemma 5.6, $\rho^*_\kappa = \rho_\kappa$. □

Theorem 6.5. (i) *Let $(\kappa_n)_1^\infty$ be a sequence of kernels on $(\mathcal{S},\mu)$ increasing a.e. to κ. Then, for every $k \geq 1$, $\rho_{\geq k}(\kappa_n;x) \nearrow \rho_{\geq k}(\kappa;x)$ for a.e. x and $\rho_{\geq k}(\kappa_n) \nearrow \rho_{\geq k}(\kappa)$.*

(ii) *Let $(\kappa_n)_1^\infty$ be a sequence of kernels on $(\mathcal{S},\mu)$ decreasing a.e. to κ. Then, for every $k \geq 1$, $\rho_{\geq k}(\kappa_n;x) \searrow \rho_{\geq k}(\kappa;x)$ for a.e. x and $\rho_{\geq k}(\kappa_n) \searrow \rho_{\geq k}(\kappa)$.*

Proof. As in the proof of Theorem 6.4, we may assume that $\kappa_n \nearrow \kappa$ or $\kappa_n \searrow \kappa$ everywhere, and that (5.1) always holds.

(i) Let $\kappa_0 := 0$ and $\Delta\kappa_n := \kappa_n - \kappa_{n-1}$, $n \geq 1$. The children of a particle of type x are given by a Poisson process with intensity $\kappa(x,y)\,d\mu(y) = \sum_n \Delta\kappa_n(x,y)\,d\mu(y)$, which can be represented as the sum of independent Poisson processes with intensities $\Delta\kappa_n(x,y)\,d\mu(y)$. We label the children in the nth of these processes by n, and give the initial 'root' vertex label 0. This gives a labelling of all particles in the branching process $\mathfrak{X}_\kappa(x)$ (which starts with a single particle of type x) such that the subset of all particles that, together with all their ancestors, have labels at most n gives the branching process $\mathfrak{X}_{\kappa_n}(x)$. Consequently (using this coupling of the processes), the family tree of the initial particle in $\mathfrak{X}_{\kappa_n}(x)$ will grow to its family tree in $\mathfrak{X}_\kappa(x)$ as $n \to \infty$. Hence $\rho_{\geq k}(\kappa_n; x) \nearrow \rho_{\geq k}(\kappa; x)$ and $\rho_{\geq k}(\kappa_n) \nearrow \rho_{\geq k}(\kappa)$.

(ii) We may similarly label all particles in $\mathfrak{X}_{\kappa_1}(x)$ with labels $\{1, 2, \dots, \infty\}$ such that $\mathfrak{X}_{\kappa_n}(x)$ $[\mathfrak{X}_\kappa(x)]$ consists of all particles that, together with their ancestors, have labels at least n $[\infty]$. By Remark 5.2, a particle always has a finite number of children, so a particle survives for eternity if and only if it has infinitely many descendants. By Theorem 6.4 we have

$$\rho(\kappa_m; x) \searrow \rho(\kappa; x) \tag{6.3}$$

for a.e. x. Fix any x for which (6.3) holds. Writing $|\mathfrak{X}_{\kappa_n}(x)|$ for the total population of the branching process $\mathfrak{X}_{\kappa_n}(x)$, whenever $|\mathfrak{X}_{\kappa_m}(x)| < \infty$ for some m, we have $|\mathfrak{X}_{\kappa_n}(x)| \searrow |\mathfrak{X}_\kappa(x)|$ as $n \to \infty$; indeed, for large n the entire processes $\mathfrak{X}_{\kappa_n}(x)$ and $\mathfrak{X}_\kappa(x)$ coincide. From (6.3), with probability 1 either $|\mathfrak{X}_\kappa(x)| = \infty$, in which case $|\mathfrak{X}_{\kappa_n}(x)| \geq |\mathfrak{X}_\kappa(x)| = \infty$ for all n, or there is an m with $|\mathfrak{X}_{\kappa_m}(x)| < \infty$, in which case $|\mathfrak{X}_{\kappa_n}(x)| = |\mathfrak{X}_\kappa(x)|$ for all large enough n. Thus, the events $|\mathfrak{X}_{\kappa_n}(x)| \geq k$ converge a.e. to $|\mathfrak{X}_\kappa(x)| \geq k$, and $\rho_{\geq k}(\kappa_n; x) \searrow \rho_{\geq k}(\kappa; x)$. □

Suppose that κ is supercritical (i.e., that $\|T_\kappa\| > 1$), and assume for simplicity that (5.1) holds for every x. Consider the branching process $\mathfrak{X}_\kappa(x)$ starting with a particle of type x, and classify its children in the first generation according to whether they have infinitely many descendants or not. By the properties of Poisson processes, this exhibits the children as the union of two independent Poisson processes with intensities $\kappa(x, y)\rho_\kappa(y)\, d\mu(y)$ and $\kappa(x, y)(1 - \rho_\kappa(y))\, d\mu(y)$ respectively, where the first litter consists of the children with infinitely many descendants, or, equivalently, those whose descendants live for ever.

The process $\mathfrak{X}_\kappa(x)$ eventually becomes extinct if and only if the first litter is empty. It follows that if $\widehat{\mathfrak{X}}_\kappa(x)$ denotes the branching process $\mathfrak{X}_\kappa(x)$ conditioned on extinction, then $\widehat{\mathfrak{X}}_\kappa(x)$ is itself a multi-type Galton–Watson branching process, where the set of children of a particle of type z is given by a Poisson process with intensity $\kappa(z, y)(1 - \rho_\kappa(y))\, d\mu(y)$. This is another instance of the situation studied here, with μ replaced by $\widehat{\mu}$ defined by $d\widehat{\mu}(y) := (1 - \rho_\kappa(y))\, d\mu(y)$, except that $\widehat{\mu}$ is not a probability measure – this is unimportant since we can normalize and consider $\widehat{\kappa}' := (1 - \rho(\kappa))\kappa$ and $\widehat{\mu}' := (1 - \rho(\kappa))^{-1}\widehat{\mu}$; see Definition 3.15 and the discussion following.

The process $\widehat{\mathfrak{X}}_\kappa(x)$ dies out by construction, and is thus subcritical or critical. Example 12.4 shows that it can be critical (even when κ is irreducible). In many cases, however, $\widehat{\mathfrak{X}}_\kappa(x)$ is subcritical; we give one simple criterion.

Lemma 6.6. *Suppose that κ is irreducible and that $\|T_\kappa\| > 1$. If $g \geq 0$ is integrable and such that $T_\kappa\big((1 - \rho_\kappa)g\big) \geq g$ a.e., then $g = 0$ a.e.*

Proof. We may assume that (5.1) holds for every x. By Theorem 6.1 and Lemma 5.8(iii), $(1-\rho_\kappa)T_\kappa\rho_\kappa < \rho_\kappa$ a.e. If $g > 0$ on a set of positive measure, then

$$\int_{\mathcal{S}} g\rho_\kappa \, d\mu > \int_{\mathcal{S}} g(1-\rho_\kappa)T_\kappa\rho_\kappa \, d\mu = \int_{\mathcal{S}} \rho_\kappa T_\kappa\bigl(g(1-\rho_\kappa)\bigr) \, d\mu \geq \int_{\mathcal{S}} \rho_\kappa g \, d\mu,$$

a contradiction. □

Theorem 6.7. *Suppose that κ is a quasi-irreducible kernel on $(\mathcal{S},\mu)$, and that $\|T_\kappa\| > 1$. Let $\widehat{\mu}$ be the measure defined by $d\widehat{\mu}(y) = (1-\rho_\kappa(y))\, d\mu(y)$, and let $\widehat{T}_\kappa$ be the corresponding integral operator*

$$\widehat{T}_\kappa g := \int_{\mathcal{S}} \kappa(x,y)g(y)\, d\widehat{\mu}(y) = T_\kappa\bigl((1-\rho_\kappa)g\bigr).$$

Then $\|\widehat{T}_\kappa\|_{L^2(\widehat{\mu})} \leq 1$.

If, in addition, $\iint_{\mathcal{S}^2} \kappa(x,y)^2 \, d\mu(x)\, d\mu(y) < \infty$, then $\|\widehat{T}_\kappa\|_{L^2(\widehat{\mu})} < 1$.

Note that with $\widehat{\kappa}' = (1-\rho(\kappa))\kappa$ and $\widehat{\mu}' = (1-\rho(\kappa))^{-1}\widehat{\mu}$ as above, we have $\widehat{T}_\kappa = T_{\widehat{\kappa}'}$, where $T_{\widehat{\kappa}'}$ is defined by $T_{\widehat{\kappa}'}g(x) := \int_{\mathcal{S}} \widehat{\kappa}'(x,y)g(y)\, d\widehat{\mu}'(y)$. Thus $\|\widehat{T}_\kappa\|_{L^2(\widehat{\mu})} = \|T_{\widehat{\kappa}'}\|_{L^2(\widehat{\mu}')}$.

Proof. We may assume (5.1) and that κ is irreducible. The discussion above and Theorem 6.1 show that $\|T_{\widehat{\kappa}'}\|_{L^2(\widehat{\mu}')} \leq 1$, as $\widehat{\mathfrak{X}}_\kappa(x)$ dies out by construction. (An analytic proof is easily given too, using a truncation of κ and the argument below for the second part.)

For the second part, the additional assumption implies that $\iint_{\mathcal{S}^2} \widehat{\kappa}'(x,y)^2 \, d\widehat{\mu}'(x)\, d\widehat{\mu}'(y) < \infty$, so $\|T_{\widehat{\kappa}'}\|_{HS} < \infty$. Lemma 5.15 shows that $\widehat{T}_\kappa = T_{\widehat{\kappa}'}$ has an eigenfunction $g \geq 0$ with eigenvalue $\|\widehat{T}_\kappa\|$, and thus

$$\|\widehat{T}_\kappa\| g = \widehat{T}_\kappa g = T_\kappa\bigl((1-\rho_\kappa)g\bigr) \qquad \text{a.e.}$$

If $\|\widehat{T}_\kappa\| \geq 1$, this contradicts Lemma 6.6. □

With a few exceptions, in the rest of the paper we shall not refer directly to the lemmas in Section 5; the results in this section describe the properties of the branching process we shall use.

7. Approximation

In this section we introduce certain upper and lower approximations to a kernel κ on a (generalized) ground space $(\mathcal{S},\mu)$, in preparation for the study of the random graph $G^{\mathcal{V}}(n,\kappa_n)$. Recall that $\mathcal{S}$ is a separable metric space, and that μ is a Borel measure on $\mathcal{S}$ with $0 < \mu(\mathcal{S}) < \infty$. We usually assume that $\mu(\mathcal{S}) = 1$; in this section, this makes no difference. Here the metric and topological structure of $\mathcal{S}$ will be important.

Given a sequence of finite partitions $\mathcal{P}_m = \{A_{m1},\dots,A_{mM_m}\}$, $m \geq 1$, of $\mathcal{S}$ and an $x \in \mathcal{S}$, we define $i_m(x)$ by

$$x \in A_{m,i_m(x)}. \tag{7.1}$$

As usual, for $A \subset \mathcal{S}$ we write $\operatorname{diam}(A)$ for $\sup\{d(x,y) : x, y \in A\}$, where d is the metric on our metric space $\mathcal{S}$.

Lemma 7.1. *Let $(\mathcal{S}, \mu)$ be a (generalized) ground space. There exists a sequence of finite partitions $\mathcal{P}_m = \{A_{m1}, \dots, A_{mM_m}\}$, $m \geq 1$, of $\mathcal{S}$ such that*

(i) *each A_{mi} is measurable and $\mu(\partial A_{mi}) = 0$;*
(ii) *for each m, $\mathcal{P}_{m+1}$ refines $\mathcal{P}_m$, i.e., each A_{mi} is a union $\bigcup_{j \in J_{mi}} A_{m+1,j}$ for some set J_{mi};*
(iii) *for a.e. $x \in \mathcal{S}$, $\operatorname{diam}(A_{m,i_m(x)}) \to 0$ as $m \to \infty$, where $i_m(x)$ is defined by (7.1).*

Proof. If $\mathcal{S} = (0,1]$ and μ is continuous, e.g., μ is the Lebesgue measure, we can take $\mathcal{P}_m$ as the dyadic partition into intervals of length 2^{-m}. If $\mathcal{S} = (0,1]$ and μ is arbitrary, we can do almost the same; we only shift the endpoints of the intervals a little when necessary to avoid point masses of μ.

In general, we can proceed as follows. Let $z_1, z_2, \dots$ be a dense sequence of points in $\mathcal{S}$. For any z_i, the balls $B(z_i, r)$, $r > 0$, have disjoint boundaries, and thus all except at most a countable number of them are μ-continuity sets. Consequently, for every $m \geq 1$ we may choose balls $B_{mi} = B(z_i, r_{mi})$ that are μ-continuity sets and have radii satisfying $1/m < r_{mi} < 2/m$. Then, $\bigcup_i B_{mi} = \mathcal{S}$, and if we define $B'_{mi} := B_{mi} \setminus \bigcup_{j<i} B_{mj}$, we obtain for each m an infinite partition $\{B'_{mi}\}_1^\infty$ of $\mathcal{S}$ into μ-continuity sets, each with diameter at most $4/m$. To get a finite partition, we choose N_m large enough to ensure that, with $B'_{m0} := \bigcup_{i>N_m} B'_{mi}$, we have $\mu(B'_{m0}) < 2^{-m}$; then $\{B'_{mi}\}_{i=0}^{N_m}$ is a partition of $\mathcal{S}$ for each m, with $\operatorname{diam}(B'_{mi}) \leq 4/m$ for $i \geq 1$.

Finally, we let $\mathcal{P}_m$ consist of all intersections $\bigcap_{l=1}^m B'_{li_l}$ with $0 \leq i_l \leq N_l$; then conditions (i) and (ii) are satisfied. Condition (iii) follows from the Borel–Cantelli Lemma: as $\sum_m \mu(B'_{m0})$ is finite, a.e. x is in finitely many of the sets B'_{m0}. For any such x, if m is large enough then $x \in B'_{mi}$ for some $i \geq 1$, so the part of $\mathcal{P}_m$ containing x has diameter at most $\operatorname{diam}(B'_{mi}) \leq 4/m$. □

Recall that a kernel κ on $(\mathcal{S}, \mu)$ is a symmetric measurable function on $\mathcal{S} \times \mathcal{S}$. Fixing a sequence of partitions with the properties described in Lemma 7.1, we can define sequences of lower and upper approximations to κ by

$$\kappa_m^-(x,y) := \inf\{\kappa(x', y') : x' \in A_{m,i_m(x)},\ y' \in A_{m,i_m(y)}\}, \tag{7.2}$$

$$\kappa_m^+(x,y) := \sup\{\kappa(x', y') : x' \in A_{m,i_m(x)},\ y' \in A_{m,i_m(y)}\}. \tag{7.3}$$

We thus replace κ by its infimum or supremum on each $A_{mi} \times A_{mj}$. As κ_m^+ might be $+\infty$, we shall use it only for bounded κ.

By Lemma 7.1(ii),

$$\kappa_m^- \leq \kappa_{m+1}^- \quad \text{and} \quad \kappa_m^+ \geq \kappa_{m+1}^+.$$

Furthermore, if κ is continuous a.e. then, by Lemma 7.1(iii),

$$\kappa_m^-(x,y) \to \kappa(x,y) \text{ and } \kappa_m^+(x,y) \to \kappa(x,y) \text{ for a.e. } (x,y) \in \mathcal{S}^2. \tag{7.4}$$

Since $\kappa_m^- \le \kappa$, we can obviously construct our random graphs so that $G(n,\kappa_m^-) \subseteq G(n,\kappa)$; in the sequel we shall assume this. Similarly, we shall assume that $G(n,\kappa_m^+) \supseteq G(n,\kappa)$ when κ is bounded.

If (κ_n) is a graphical sequence of kernels with limit κ, we define instead

$$\kappa_m^-(x,y) := \inf\{(\kappa \wedge \kappa_n)(x',y') : x' \in A_{m,i_m(x)},\ y' \in A_{m,i_m(y)},\ n \ge m\}. \tag{7.5}$$

By Lemma 7.1(ii), we have $\kappa_m^- \le \kappa_{m+1}^-$, and from Lemma 7.1(iii) and (2.10) we see that

$$\kappa_m^-(x,y) \nearrow \kappa(x,y) \quad \text{as } m \to \infty, \text{ for a.e. } (x,y) \in \mathcal{S}^2. \tag{7.6}$$

Moreover, when $n \ge m$ we have

$$\kappa_n \ge \kappa_m^-, \tag{7.7}$$

and we may assume that $G(n,\kappa_m^-) \subseteq G(n,\kappa_n)$.

For a uniformly bounded graphical sequence (κ_n) of kernels with limit κ, we similarly define

$$\kappa_m^+(x,y) := \sup\{(\kappa \vee \kappa_n)(x',y') : x' \in A_{m,i_m(x)},\ y' \in A_{m,i_m(y)},\ n \ge m\} < \infty. \tag{7.8}$$

Relations corresponding to (7.6) and (7.7) hold for κ_m^+; we collect these and an additional result in the following lemma.

Lemma 7.2. *Let $(\kappa_n)_{n\in I}$ be a graphical sequence of kernels on a (generalized) vertex space $\mathcal{V}$ with limit κ, and suppose that $\sup_{x,y,n} \kappa_n(x,y) < \infty$. Then there is a sequence κ_m^+, $m = 1,2,\dots,$ of regular finitary kernels on $\mathcal{V}$ with the following properties.*

(i) *We have $\kappa_m^+(x,y) \searrow \kappa(x,y)$ as $m \to \infty$ for a.e. $(x,y) \in \mathcal{S}^2$.*
(ii) *Whenever $n \ge m$ we have $\kappa_m^+(x,y) \ge \kappa_n(x,y)$ for* every *$(x,y) \in \mathcal{S}^2$.*
(iii) *$\|T_{\kappa_m^+}\| \searrow \|T_\kappa\|$ as $m \to \infty$.*

Proof. Let $\mathcal{P}_m = \{A_{m1},\dots,A_{mM_m}\}$, $m \ge 1$, be a sequence of partitions with the properties described in Lemma 7.1, and define $\kappa_m^+(x,y)$ by (7.8). (If $\kappa_n = \kappa$ for all n, this is just (7.3).) Then (ii) holds trivially. By Lemma 7.1(iii) and (2.10), $\kappa_m^+ \searrow \kappa$ a.e., proving (i). Finally, by dominated convergence, $\|T_{\kappa_m^+} - T_\kappa\|_{HS} \to 0$. Hence,

$$\|T_\kappa\| \le \|T_{\kappa_m^+}\| \le \|T_\kappa\| + \|T_{\kappa_m^+} - T_\kappa\| \le \|T_\kappa\| + \|T_{\kappa_m^+} - T_\kappa\|_{HS} \searrow \|T_\kappa\|,$$

proving (iii). ☐

We finish this section with a result for lower approximations corresponding to Lemma 7.2, but with one additional ingredient: for lower approximations to be useful we shall often need them to be quasi-irreducible.

Lemma 7.3. *If $(\kappa_n)_{n\in I}$ is a graphical sequence of kernels on a (generalized) vertex space $\mathcal{V}$ with limit κ, there is a sequence $\hat\kappa_m^-$, $m = 1, 2, \ldots$, of regular finitary kernels on $\mathcal{V}$ with the following properties.*

(i) *If κ is quasi-irreducible, then so is $\hat\kappa_m^-$ for all large m.*
(ii) *We have $\hat\kappa_m^-(x, y) \nearrow \kappa(x, y)$ as $m \to \infty$ for a.e. $(x, y) \in \mathcal{S}^2$.*
(iii) *Whenever $n \geq m$ we have $\hat\kappa_m^-(x, y) \leq \kappa_n(x, y)$ for* every $(x, y) \in \mathcal{S}^2$.

Before turning to the proof, let us note that the conclusions of the lemma are obvious for suitably 'nice' kernels κ (or sequences $\kappa_n \to \kappa$); for example if κ is continuous, $\mathcal{S}$ is compact and $\kappa > 0$. Indeed, if we partition $\mathcal{S}$ into finitely many pieces $\mathcal{S}_i$ in a suitable way, we may then set $\hat\kappa_m^-(x, y) = \inf\{\kappa(x', y') : x' \in \mathcal{S}_i,\ y' \in \mathcal{S}_j\}$ whenever $x \in \mathcal{S}_i$ and $y \in \mathcal{S}_j$. Note also that in the application we shall need condition (iii) for *every* $(x, y) \in \mathcal{S}^2$: while changes in a kernel κ on a set of measure zero do not affect the branching process $\mathfrak{X}_\kappa$, they can affect the graph $G^{\mathcal{V}}(n, \kappa)$.

Proof of Lemma 7.3. We may assume that $\kappa > 0$ on a set of positive measure, as otherwise we may take $\hat\kappa_m^- = 0$ for every m and there is nothing to prove. We shall construct the sequence $\hat\kappa_m^-$ in two stages.

Let $\mathcal{P}_m = \{A_{m1}, \ldots, A_{mM_m}\}$, $m \geq 1$, be a sequence of partitions with the properties described in Lemma 7.1. If $\kappa_n = \kappa$ for all n, we start with κ_m^- defined in (7.2). In general, with a sequence κ_n, we use instead the definition (7.5).

Each κ_m^- is of the regular finitary type treated above, and the κ_m^- have two of the properties required for the $\hat\kappa_m^-$, namely (ii) and (iii), by (7.6) and (7.7), respectively. However, (i) may fail, as some κ_m^- may be reducible. From now on we shall assume that κ is quasi-irreducible, as otherwise we may take $\hat\kappa_m^- = \kappa_m^-$. In fact, without loss of generality we may assume that κ is irreducible. Indeed, it suffices to prove this case as, given a quasi-irreducible κ, we may then apply the result to the irreducible restriction to $\mathcal{S}' \times \mathcal{S}'$, and extend the approximating $\hat\kappa_m^-$ obtained to $\mathcal{S} \times \mathcal{S}$ by taking them to be zero off $\mathcal{S}' \times \mathcal{S}'$. We shall thus assume that κ is irreducible.

If $\kappa_m^- = 0$ a.e. for every m, then $\kappa = 0$ a.e. by (7.6), contradicting our assumption. We may thus assume that there exists an m_0 such that $\kappa_{m_0}^- > 0$ on a set of positive measure. We consider only $m \geq m_0$, and assume for notational convenience that $m_0 = 1$. Thus there exist i and j (possibly equal) with $\mu(A_{1i}), \mu(A_{1j}) > 0$ and $\kappa_1^- > 0$ on $A_{1i} \times A_{1j}$. From now on we fix such a pair i and j.

For $m \geq 1$, let $E_m := \bigcup\{A_{mi} : \mu(A_{mi}) = 0\}$, noting that $\mu(E_m) = 0$, and let B_m be the set of all $x \in \mathcal{S}$ such that for some $k \geq 1$ there exists a sequence $x_0, \ldots, x_k$ with $x_0 = x$, $x_k \in A_{1i}$, $\kappa_m^-(x_{l-1}, x_l) > 0$, and $x_l \notin E_m$ for $l = 1, \ldots, k$. (Note that $x = x_0$ may belong to E_m.) Since κ_m^- is constant on each $A_{mp} \times A_{mq}$, B_m is a union of some of the sets A_{mp}. It is easily seen that $B_m \subseteq B_{m+1}$, that $B_m \supseteq B_1 \supseteq A_{1j}$, that the restriction of κ_m^- to B_m is irreducible and that $\kappa_m^- = 0$ on $(B_m \setminus E_m) \times (\mathcal{S} \setminus B_m)$ and thus a.e. on $B_m \times (\mathcal{S} \setminus B_m)$.

Let $B := \bigcup_1^\infty B_m$. If $n \geq m$, then $B_m \subseteq B_n$ and thus $\kappa_n^- = 0$ a.e. on $B_m \times (\mathcal{S} \setminus B) \subseteq B_n \times (\mathcal{S} \setminus B_n)$. Letting $n \to \infty$, (7.6) shows that $\kappa = 0$ a.e. on $B_m \times (\mathcal{S} \setminus B)$. Letting now $m \to \infty$ (taking the union) yields $\kappa = 0$ a.e. on $B \times (\mathcal{S} \setminus B)$. Since κ is irreducible, it follows by (2.12) that $\mu(B) = 0$ or $\mu(\mathcal{S} \setminus B) = 0$. As $B \supseteq B_1 \supseteq A_{1j}$, we have $\mu(B) > 0$, so $\mu(\mathcal{S} \setminus B) = 0$. In other words, a.e. $x \in B = \bigcup_m B_m$.

Now define

$$\hat{\kappa}_m^-(x, y) = \kappa_m^-(x, y)\mathbf{1}[x \in B_m]\mathbf{1}[y \in B_m].$$

Thus $\hat{\kappa}_m^-$ is 0 off $B_m \times B_m$, and the restriction to B_m is by construction irreducible and of the regular finitary type, so condition (i) of the lemma is satisfied. Furthermore, by (7.6) and the fact that $B_m \nearrow B$ with $\mu(\mathcal{S} \setminus B) = 0$, we have $\hat{\kappa}_m^-(x, y) \nearrow \kappa(x, y)$ as $m \to \infty$ for a.e. $(x, y) \in \mathcal{S}^2$, so (ii) holds.

Finally, if $n \geq m$, then $\hat{\kappa}_m^- \leq \kappa_m^- \leq \kappa_n$, so (iii) holds. □

8. The number of edges

In this section we consider circumstances in which the condition (2.9) or (2.11) on the convergence of the number of edges in $G^{\mathcal{V}}(n, \kappa)$ does, or does not, hold. In doing so, we shall make frequent use of the approximating kernels κ_m^- and κ_m^+ defined for a single kernel κ by (7.2) and (7.3), and for a sequence by (7.5) and (7.8). As before, we shall always write the (generalized) vertex space $\mathcal{V}$ under consideration as $(\mathcal{S}, \mu, (\mathbf{x}_n))$, unless otherwise specified.

Lemma 8.1. *Let κ be an a.e. continuous kernel on a (generalized) vertex space $\mathcal{V}$. Then*

$$\liminf \frac{1}{n} \mathbb{E}\, e\big(G^{\mathcal{V}}(n, \kappa)\big) \geq \tfrac{1}{2} \iint_{\mathcal{S}^2} \kappa(x, y)\, d\mu(x)\, d\mu(y). \tag{8.1}$$

If κ is a bounded a.e. continuous kernel on a vertex space $\mathcal{V}$, then

$$\lim \frac{1}{n} \mathbb{E}\, e\big(G^{\mathcal{V}}(n, \kappa)\big) = \tfrac{1}{2} \iint_{\mathcal{S}^2} \kappa(x, y)\, d\mu(x)\, d\mu(y). \tag{8.2}$$

Proof. Write $G(n, \kappa)$ for $G^{\mathcal{V}}(n, \kappa)$. Consider first the regular finitary case defined in Definition 4.4. For $n \geq \max \kappa$, conditioning on $n_1, \ldots, n_r$ we have

$$\begin{aligned}\frac{1}{n} \mathbb{E}\, e\big(G(n, \kappa) \mid n_1, \ldots, n_r\big) &= \frac{1}{2n} \sum_{i,j=1}^{r} (n_i n_j - n_i \delta_{ij}) \frac{1}{n} \kappa(i, j) \\ &\xrightarrow{\mathrm{p}} \tfrac{1}{2} \sum_{i,j=1}^{r} \kappa(i, j) \mu(S_i) \mu(S_j) = \tfrac{1}{2} \iint_{\mathcal{S}^2} \kappa(x, y)\, d\mu(x)\, d\mu(y). \end{aligned} \tag{8.3}$$

Taking expectations and applying Fatou's Lemma, it follows that (8.1) holds in this case.

In general, to prove (8.1) we use Lemma 7.1 and the approximation (7.2). For every m, by the case just treated,

$$\liminf_{n\to\infty} \frac{1}{n}\,\mathbb{E}\, e\big(G(n,\kappa)\big) \geq \liminf_{n\to\infty} \frac{1}{n}\,\mathbb{E}\, e\big(G(n,\kappa_m^-)\big) = \tfrac{1}{2}\iint_{\mathcal{S}^2} \kappa_m^-.$$

As $m\to\infty$, the monotone convergence theorem implies that $\iint \kappa_m^- \to \iint \kappa$, and (8.1) follows.

If $\mathcal{V}$ is a vertex space and κ is regular finitary, then the left-hand side of (8.3) is bounded by $\max\kappa/2$, so by the dominated convergence theorem we have $\frac{1}{n}\,\mathbb{E}\, e\big(G(n,\kappa)\big) \to \frac{1}{2}\iint\kappa$. In general, if $\mathcal{V}$ is a vertex space and κ is bounded, we can use κ_m^+ in place of κ_m^- to show that $\limsup_{n\to\infty}\frac{1}{n}\,\mathbb{E}\, e\big(G(n,\kappa)\big) \leq \frac{1}{2}\iint\kappa$. □

Remark 8.2. Condition (8.2) may fail for a generalized vertex space $\mathcal{V}$, even if κ is constant. The problem is that the definition of a generalized vertex space only imposes 'whp conditions' on the number of vertices, giving no control on the distribution in the $o(1)$ probability case that these conditions fail, and hence giving no control on expectations. In particular, with κ identically 1, the expected number of edges is essentially $\frac{1}{2n}\,\mathbb{E}(|V(G^{\mathcal{V}}(n,\kappa))|^2)$, and we have no control over this expectation – it can even be infinite.

When the number v_n of vertices is sufficiently concentrated (for example Poisson), this problem does not arise. Indeed, (8.2) holds whenever κ is bounded and $\mathrm{Var}(v_n/n)\to 0$; since $v_n/n \xrightarrow{\mathrm{p}} \mu(\mathcal{S})$ by assumption, the variance condition is easily shown to be equivalent to $\mathbb{E}(v_n/n)^2 \to \mu(\mathcal{S})^2$, and to imply uniform integrability of $(v_n/n)^2$; see e.g. [61, Proposition 4.12]. (If the parameter n is not restricted to integers, we may have to consider a sequence of indices n.) Since the left-hand side of (8.3) is bounded by $\max\kappa\,(v_n/n)^2$, which is also uniformly integrable, we may take the expectation in (8.3) and obtain (8.2).

Our main results concern statements that hold whp, and convergence in probability of various quantities. For such statements, a small chance of a very large number of vertices is not a problem.

The following lemma shows that the condition to be graphical is essentially equivalent to a statement about approximations with bounded kernels.

Lemma 8.3. *Let κ be a bounded a.e. continuous kernel on a (generalized) vertex space $\mathcal{V}$.*

If $\mathcal{V}$ is a vertex space, then κ is graphical if and only if

(i) *for every $\varepsilon > 0$ there exists an $M < \infty$ such that*

$$\limsup \frac{1}{n}\,\mathbb{E}\, e\big(G^{\mathcal{V}}(n,\kappa)\setminus G^{\mathcal{V}}(n,\kappa\wedge M)\big) \leq \varepsilon.$$

In general, κ is graphical if and only if (i) *holds together with*

(ii) *for every $M<\infty$,*

$$\limsup \frac{1}{n}\,\mathbb{E}\, e\big(G^{\mathcal{V}}(n,\kappa\wedge M)\big) \leq \tfrac{1}{2}\iint_{\mathcal{S}^2} \kappa(x,y)\,d\mu(x)\,d\mu(y).$$

Proof. It is obvious that (i) and (ii) imply that $\limsup \frac{1}{n} \mathbb{E}\, e\big(G^{\mathcal{V}}(n,\kappa)\big) \le \frac{1}{2}\iint_{\mathcal{S}^2} \kappa(x,y)\, d\mu(x)\, d\mu(y)$, which together with Lemma 8.1 shows that κ is graphical.

Conversely, suppose that κ is graphical on $\mathcal{V}$. Then, from the definition of graphicality (see (2.9)), $\mathbb{E}\, e\big(G^{\mathcal{V}}(n,\kappa)\big) \to \frac{1}{2}\iint \kappa$. Applying Lemma 8.1 to $\kappa \wedge M$, it follows that

$$\begin{aligned}\limsup \frac{1}{n} &\mathbb{E}\, e\big(G^{\mathcal{V}}(n,\kappa) \setminus G^{\mathcal{V}}(n,\kappa\wedge M)\big) \\ &= \tfrac{1}{2}\iint \kappa - \liminf \frac{1}{n}\mathbb{E}\, e\big(G^{\mathcal{V}}(n,\kappa\wedge M)\big) \le \tfrac{1}{2}\iint (\kappa - \kappa\wedge M) < \varepsilon\end{aligned}$$

if M is large enough. ☐

Note that (ii) almost always holds by Lemma 8.1 and Remark 8.2. Arguing as in the proof of Lemma 8.3, one can show that (ii) can be replaced by the condition that each $\kappa \wedge M$ be graphical; we omit the details.

Remark 8.4. Lemma 8.3 implies that, if κ is a graphical kernel on a (generalized) vertex space $\mathcal{V}$ and $0 < c < \infty$, then $c\kappa$ is also graphical on $\mathcal{V}$. Indeed, it suffices to check condition (iii) of Definition 2.7, namely that $\frac{1}{n}\mathbb{E}\, e\big(G(n,c\kappa)\big) \to \frac{1}{2}\iint c\kappa(x,y)$. Without the $\min\{\cdot,1\}$ in the formula (2.8), this would be immediate from the same condition for κ; indeed, the claim that $c\kappa$ is graphical is equivalent to the claim that replacing this 1 with $1/c$ does not affect $\mathbb{E}\, e\big(G(n,\kappa)\big)$ by more than $o(n)$.

Since $c\kappa \wedge cM = c(\kappa\wedge M)$, it is obvious that condition (ii) of Lemma 8.3 holds for $c\kappa$ if and only if it holds for κ. Moreover, if $n \ge M$,

$$\tfrac{1}{n}\mathbb{E}\big(e(G^{\mathcal{V}}(n,\kappa)\setminus G^{\mathcal{V}}(n,\kappa\wedge M)) \mid \mathbf{x}_n\big) = n^{-2}\sum_{i<j}\big(\kappa(x_i,x_j)\wedge n - \kappa(x_i,x_j)\wedge M\big).$$

It is clear that if we replace κ by $c\kappa$ and M by cM, and assume $n \ge 2(1\vee c)M$, then this sum changes by at most a constant factor. Hence condition (i) of Lemma 8.3 also holds for $c\kappa$ if and only if it holds for κ.

Lemmas 8.1 and 8.3 hold also for the variants of $G^{\mathcal{V}}(n,\kappa)$ defined in Remark 2.4, by the same proofs. Moreover, it is easily seen that conditions (i) and (ii) of Lemma 8.3 hold for one of these versions if and only if they hold for $G^{\mathcal{V}}(n,\kappa)$. Hence κ is graphical if and only if the analogue of (2.9) for one of these variants holds.

The results above can be extended to sequences (κ_n) satisfying (2.10), using the approximations κ_m^- defined by (7.5). In particular, a similar argument shows that if (κ_n) is a graphical sequence of kernels on a (generalized) vertex space $\mathcal{V}$ with limit κ, and (c_n) is a sequence of positive reals with $c_n \to c > 0$, then $(c_n\kappa_n)$ is graphical on $\mathcal{V}$ with limit $c\kappa$.

Let us emphasize that relation (8.2), i.e., condition (2.9) from the definition of graphicality, often holds for unbounded κ too, and for generalized vertex spaces $\mathcal{V}$. One example is when the x_i are random as in Example 4.8;

another is the Poisson process case in Example 4.9. A rather different example is the following.

Example 8.5. Suppose that $\mathcal{S} = (0,1]$, μ is the Lebesgue measure and $x_i = i/n$; this vertex space was considered in Example 4.7 and will be used in several further examples in Section 16, with several different kernels. Suppose that κ is decreasing in each variable, so $\kappa(x,y) \geq \kappa(x',y')$ when $x \leq x'$ and $y \leq y'$. Then

$$\frac{1}{n}\,\mathbb{E}\, e\big(G(n,\kappa)\big) = \frac{1}{2n^2}\sum_{i\neq j}\kappa(i/n,j/n)\wedge n \leq \frac{1}{2}\int_0^1\int_0^1 \kappa(x,y)\,dx\,dy.$$

Hence (8.1) implies that (8.2) holds in this case. Note that this includes both (16.1) and (16.4).

We next give a simple example where (2.9), i.e., (8.2), fails.

Example 8.6. Take again $\mathcal{S} = (0,1]$, let μ be the Lebesgue measure and set $x_i = i/n$. Let $0 < \delta < 1$ be constant and define κ by

$$\kappa(x,y) = \begin{cases} m & \text{if } x \wedge y = 1/m \text{ and } m \geq 1; \\ \delta & \text{otherwise.} \end{cases}$$

Note that $\kappa = \delta$ a.e., and hence $\rho(\kappa) = \rho(\delta) = 0$ as for $G(n,\delta/n)$ in Example 4.1; furthermore, κ is continuous a.e.

Now, $\kappa(1/n, j/n) = n$ for every $j \leq n$. Hence, $G(n,\kappa)$ contains the star consisting of all edges $1j$, $1 < j \leq n$, so $G(n,\kappa)$ is connected and $C_1(G(n,\kappa)) = n$, although, as remarked above, $\rho(\kappa) = 0$. Consequently, (3.2) fails in this case. Note that all assumptions of Theorem 3.1 are satisfied except (2.9); indeed, $e(G(n,\kappa)) \geq n-1$ while $\iint_{\mathcal{S}^2}\kappa(x,y)\,d\mu(x)\,d\mu(y) = \delta$.

We can modify this example to make κ continuous on $(0,1]^2$: for $0 < \varepsilon < 1/4$, let $\kappa_\varepsilon(x,y) = \phi(x\wedge y)$ with $\phi(1/m) = m$, $\phi(1/m \pm \varepsilon m^{-4}) = \delta$, and ϕ linear in between. If ε is small enough, then $\|T_{\kappa_\varepsilon}\| < 1$ (because the Hilbert–Schmidt norm satisfies $\|T_{\kappa_\varepsilon}\|_{HS} \to \delta$ as $\varepsilon \to 0$ by dominated convergence); thus $\rho(\kappa_\varepsilon) = 0$, although $C_1(G(n,\kappa_\varepsilon)) = n$.

We next give a result on the number of edges conditioned on $\mathbf{x}_n$; this time we consider a sequence (κ_n) of kernels.

Lemma 8.7. *Let (κ_n) be a graphical sequence of kernels on a (generalized) vertex space $\mathcal{V}$ with limit κ. Then*

$$\frac{1}{n}\,\mathbb{E}\Big(e\big(G^{\mathcal{V}}(n,\kappa_n)\big)\;\Big|\;\mathbf{x}_n\Big) \xrightarrow{\text{p}} \tfrac{1}{2}\iint_{\mathcal{S}^2}\kappa(x,y)\,d\mu(x)\,d\mu(y).$$

Proof. Let $W_n := \mathbb{E}\big(e\big(G^{\mathcal{V}}(n,\kappa_n)\big) \mid \mathbf{x}_n\big)/n$ and $w := \frac{1}{2}\iint_{\mathcal{S}^2}\kappa(x,y)\,d\mu(x)\,d\mu(y)$. By our assumption (2.11), we have $\mathbb{E}\,W_n \to w$.

Define κ_m^- by (7.5). By (8.3), applied to κ_m^-,

$$W_n^{(m)} := \mathbb{E}\big(e\big(G^{\mathcal{V}}(n,\kappa_m^-)\big) \mid \mathbf{x}_n\big)/n \xrightarrow{\text{p}} w_m := \tfrac{1}{2}\iint_{\mathcal{S}^2}\kappa_m^-(x,y)\,d\mu(x)\,d\mu(y).$$

Let $\varepsilon > 0$ be given. By (7.6) and monotone convergence, $w_m \to w$ as $m \to \infty$, so we may choose m such that $w_m > w - \varepsilon$. For $n \ge m$ we have $W_n \ge W_n^{(m)}$, and hence

$$\mathbb{P}\big(W_n < w - 2\varepsilon\big) \le \mathbb{P}\big(W_n^{(m)} < w_m - \varepsilon\big) \to 0 \qquad \text{as } n \to \infty.$$

Hence, writing f_- for $-(f \wedge 0)$, we have $(W_n - w)_- \xrightarrow{\mathrm{p}} 0$ and, by dominated convergence, $\mathbb{E}(W_n - w)_- \to 0$. Consequently, $\mathbb{E}\,|W_n - w| = 2\,\mathbb{E}(W_n - w)_- + \mathbb{E}(W_n - w) \to 0$. □

Remark 8.8. Recalling (2.2) or (2.4), the convergence condition for the empirical distribution ν_n of the types of the vertices in a (generalized) vertex space, we have $\nu_n \xrightarrow{\mathrm{p}} \mu$ and $W_n \xrightarrow{\mathrm{p}} w$ (in the notation of the proof above), where ν_n and W_n are functions of $\mathbf{x}_n$. Coupling the $\mathbf{x}_n$ for different n appropriately (a simple application of the Skorohod coupling theorem [61, Theorem 4.30]), or considering appropriate subsequences, we may assume that $\nu_n \to \mu$ and $W_n \to w$ a.s. Consequently, we may condition on $\mathbf{x}_n$ and assume that (2.2) and (2.11) still hold. In other words, after conditioning on $\mathbf{x}_n$, $\mathcal{V}$ is still a (generalized) vertex space and (κ_n) is still graphical with limit κ. By conditioning in this way we may thus assume that $\mathbf{x}_n$ is deterministic; see Subsection 8.1.

Our next result shows that the number of edges is concentrated, so that the actual number converges as well as its mean.

Proposition 8.9. *Let (κ_n) be a graphical sequence of kernels on a (generalized) vertex space $\mathcal{V}$ with limit κ. Then*

$$\frac{1}{n} e\big(G^{\mathcal{V}}(n, \kappa_n)\big) \xrightarrow{\mathrm{p}} \tfrac{1}{2} \iint_{\mathcal{S}^2} \kappa(x,y)\, d\mu(x)\, d\mu(y).$$

Proof. Let $G_n = G^{\mathcal{V}}(n, \kappa_n)$ and, as above, $W_n = \mathbb{E}(e(G_n) \mid \mathbf{x}_n)/n$. Conditioned on $\mathbf{x}_n$, the number $e(G_n)$ of edges is a sum of independent $\mathrm{Be}(p_{ij})$ variables, and thus $\mathrm{Var}(e(G_n) \mid \mathbf{x}_n) \le \mathbb{E}(e(G_n) \mid \mathbf{x}_n)$. Hence, using (2.11),

$$\mathbb{E}\big(e(G_n)/n - W_n\big)^2 = n^{-2}\,\mathbb{E}\big(\mathrm{Var}(e(G_n) \mid \mathbf{x}_n)\big) \le n^{-2}\,\mathbb{E}\big(e(G_n)\big) \to 0.$$

Consequently, $e(G_n)/n - W_n \xrightarrow{\mathrm{p}} 0$, and the result follows by Lemma 8.7. □

Finally, we note that small sets of vertices do not connect to too many edges. For this we need a simple lemma.

Lemma 8.10. *Let κ be a bounded kernel on a (generalized) vertex space $\mathcal{V}$, and let $G_n = G^{\mathcal{V}}(n, \kappa)$. Then $\sum_i d_{G_n}^2(i) = O(n)$ whp.*

Proof. Let $a := 2\mu(\mathcal{S})$ and $b := \sup \kappa(x,y) < \infty$. Then whp $v_n \le an$, and thus, in the natural coupling, $G^{\mathcal{V}}(n, \kappa) \subseteq G(\lfloor an \rfloor, b/n)$ whp. Consequently, it suffices to prove the result for $G(\lfloor an \rfloor, b/n)$ or, changing the notation slightly, for $G(n, c/n)$ for every fixed $c > 0$. However, for any graph G, $\sum_i d_G(i) = 2e(G)$ and $\sum_i d_G(i)(d_G(i) - 1) = 2P_2(G)$, twice the number of paths of length 2 (cf. Section 17). It is well-known, and easy to prove, that

$e(G(n,c/n))/n \xrightarrow{\mathrm{p}} \alpha_1$ and $P_2(G(n,c/n))/n \xrightarrow{\mathrm{p}} \alpha_2$ for some constants α_1, α_2 (depending on c), see e.g. [59, Chapter 3 and Theorem 6.5]. Consequently, with $C := 2\alpha_1 + 2\alpha_2 + 1$, $\sum_i d^2_{G(n,c/n)}(i) < Cn$ whp. (Alternatively, we may use Theorem 17.1.) □

Proposition 8.11. *Let (κ_n) be a graphical sequence of kernels on a (generalized) vertex space $\mathcal{V}$ with limit κ. Given $\varepsilon > 0$, there is a $\delta > 0$ so that whp the sum of the degrees of any set of at most δn vertices of $G_n = G^{\mathcal{V}}(n, \kappa_n)$ is at most εn. In particular, any set of $o_p(n)$ vertices of G_n has $o_p(n)$ neighbours.*

Proof. Let $d_G(v)$ denote the degree of vertex v in a graph G, and let $\hat{\kappa}^-_m$ be as in Lemma 7.3. Since $\hat{\kappa}^-_m \nearrow \kappa$ a.e., $\iint \hat{\kappa}^-_m \to \iint \kappa$, and thus we can choose m such that $\iint \hat{\kappa}^-_m > \iint \kappa - \varepsilon$. Let $G'_n := G^{\mathcal{V}}(n, \hat{\kappa}^-_m)$. For $n \ge m$ we have $\hat{\kappa}^-_m \le \kappa_n$, and we may as usual assume that $G'_n \subseteq G_n$. Moreover, Proposition 8.9 applies to both G_n and G'_n, so

$$\begin{aligned} \frac{1}{2n} \sum_{i \in V(G_n)} \bigl(d_{G_n}(i) - d_{G'_n}(i)\bigr) &= \frac{1}{n} e\bigl(G^{\mathcal{V}}(n, \kappa_n) \setminus G^{\mathcal{V}}(n, \hat{\kappa}^-_m)\bigr) \\ &\xrightarrow{\mathrm{p}} \tfrac{1}{2} \iint_{\mathcal{S}^2} \kappa - \tfrac{1}{2} \iint_{\mathcal{S}^2} \hat{\kappa}^-_m < \varepsilon. \end{aligned}$$

Hence, whp

$$\sum_{i \in V(G_n)} \bigl(d_{G_n}(i) - d_{G'_n}(i)\bigr) < 2\varepsilon n. \tag{8.4}$$

By Lemma 8.10, applied to G'_n, there is a constant $C < \infty$ such that whp $\sum_i d^2_{G'_n}(i) < Cn$. Hence, if $\delta = \varepsilon^2/C$, the Cauchy–Schwarz inequality shows that whp for every set $A \subseteq V(G_n)$ with $|A| \le \delta n$,

$$\sum_{i \in A} d_{G'_n}(i) \le \Bigl(|A| \sum_{i \in A} d^2_{G'_n}(i)\Bigr)^{1/2} < \bigl(|A| Cn\bigr)^{1/2} \le \varepsilon n.$$

Combining this with (8.4), we obtain $\sum_{i \in A} d_{G_n}(i) < 3\varepsilon n$, whp for all such A, and the result follows by replacing ε by $\varepsilon/3$. □

8.1. **Generalized vertex spaces.** Our main results concern graphical sequences of kernels on generalized vertex spaces, expressing properties of the graphs $G^{\mathcal{V}}(n, \kappa_n)$ in terms of the limiting kernel κ. As noted earlier, it is intuitively clear that we lose no generality by restricting our attention to vertex spaces. Furthermore, as noted in Remark 8.8, we may assume that the vertex types $\mathbf{x}_n$ are deterministic. As we shall now see, Lemma 8.7 and a simple probabilistic lemma given in the appendix imply precise forms of these assertions. We start by showing that we may take the sequences $\mathbf{x}_n$ to be deterministic.

Let κ_n be a graphical sequence of kernels on a generalized vertex space $\mathcal{V}$ with limit κ. As noted in Remark 8.8, by coupling appropriately we may assume that, after conditioning on $(\mathbf{x}_n)$, the triple $(\mathcal{S}, \mu, (\mathbf{x}_n))$, in which the sequences $\mathbf{x}_n$ are now deterministic, is (a.s.) still a generalized vertex space, and that κ_n is (a.s.) graphical on this space with limit κ. Almost all our results assert that (given some $\varepsilon > 0$) a certain event $\mathcal{E}_n$ holds whp; recall that statements like $X_n \xrightarrow{\mathrm{p}} a$ and $X_n = o_p(a_n)$ can be expressed in this form. The $\mathbf{x}_n$ deterministic case of such a result then implies that (a.s.) the conditional probabilities $\mathbb{P}(\mathcal{E}_n \mid \mathbf{x}_n)$ tend to 1. Taking expectation with respect to the random sequences $\mathbf{x}_n$, it follows by dominated convergence that $\mathbb{P}(\mathcal{E}_n) \to 1$, i.e., the result holds also for random $\mathbf{x}_n$.

Some of our results are of the form $X_n = O(a_n)$ whp. Again, it suffices to prove such a result for deterministic $\mathbf{x}_n$; the general case then follows by Lemma A.5, with $Y_n = \nu_n$, and $\mathcal{M}_n$ the set of all measures of the form $n^{-1}\sum_1^N \delta_{x_i}$, a subset of the metric space $\mathcal{M}$ of all finite Borel measures on $\mathcal{S}$. The key point is that ν_n determines $\mathbf{x}_n$ up to relabelling the vertices, and that the conditional distribution of the unlabelled graph $G^{\mathcal{V}}(n, \kappa_n)$ given $\mathbf{x}_n$ does not depend on the labelling, or on $\mathbf{x}_{n'}$, $n' \neq n$.

We now turn to the simple reduction from generalized vertex spaces to vertex spaces. Although the arguments apply to all our main results, for definiteness, we shall illustrate them with one particular example: we shall show that statement (iii) of Theorem 3.1, namely

$$\frac{1}{n} C_1(G^{\mathcal{V}}(n, \kappa_n)) \xrightarrow{\mathrm{p}} \rho(\kappa), \tag{8.5}$$

follows from the same statement restricted to the case that $\mathcal{V}$ is a vertex space.

Let $\mathcal{V} = (\mathcal{S}, \mu, (\mathbf{x}_n)_{n\in I})$ be a generalized vertex space, and let κ_n, $n \in I$, be a graphical sequence of kernels on $\mathcal{V}$ with limit κ. As noted in Section 2, purely formal manipulations show that taking $\mu(\mathcal{S}) = 1$ loses no generality, although one must be a little careful with the introduction of normalizing factors. To spell this out pedantically, let $I' = \mu(\mathcal{S})I = \{\mu(\mathcal{S})n : n \in I\}$, let $\mu' = \mu/\mu(\mathcal{S})$ be the normalized version of the measure μ, and let $\mathcal{V}'$ be the generalized vertex space $(\mathcal{S}, \mu', (\mathbf{y}_m)_{m\in I'})$ defined by $\mathbf{y}_m = \mathbf{x}_{m/\mu(\mathcal{S})}$, so the sequences $(\mathbf{x}_n)$ and $(\mathbf{y}_m)$ are identical except for our rescaling of the index set. Writing κ' for $\mu(\mathcal{S})\kappa$ and κ'_m for $\mu(\mathcal{S})\kappa_{m/\mu(\mathcal{S})}$, for $n \in I$ the graphs $G^{\mathcal{V}}(n, \kappa_n)$ and $G^{\mathcal{V}'}(m, \kappa'_m)$, $m = \mu(\mathcal{S})n$, have exactly the same distribution. Also (as a consequence), the sequence κ'_n is graphical on $\mathcal{V}'$ with limit κ', so our main results, in particular Theorem 3.1, apply to the model $G^{\mathcal{V}'}(m, \kappa'_m)$.

Multiplying κ by the constant factor $\mu(\mathcal{S})$ and dividing μ by the same factor leaves the branching process Φ_κ, and hence the survival probability $\rho(\kappa; x)$, unchanged, and so divides $\rho(\kappa)$ by a factor $\mu(\mathcal{S})$. Multiplying the index variable n by $\mu(\mathcal{S})$ divides the left-hand side of (8.5) by the same factor, so this relation for $G^{\mathcal{V}}(n, \kappa_n)$ follows from the same relation for the model $G^{\mathcal{V}'}(m, \kappa'_m)$.

Apart from the rather trivial normalization above, there are two further differences between vertex spaces and generalized vertex spaces. One is that, in the former, the index set is discrete, indeed a subset of the integers. This makes very little difference: for any result of the form $f(G_n) \xrightarrow{\mathrm{p}} a$, $n \in I$, it suffices to consider 'thin' index sets I, say discrete sets $\{i_1, i_2, \dots, \}$ with $i_1 \geq 100$ and $i_{t+1} \geq 2i_t$. Indeed, if $f(G_n) \xrightarrow{\mathrm{p}} a$ fails, there is an $\varepsilon > 0$ and an unbounded set $I' \subset I$ with $\mathbb{P}(|f(G_n) - a| \geq \varepsilon) \geq \varepsilon$ for every $n \in I'$, and then $f(G_n) \xrightarrow{\mathrm{p}} a$ fails along any subsequence of I', and hence along at least one thin sequence. Thus, in all our main results we need only consider 'thin' index sets.

The final extension allowed by generalized vertex spaces is a that the number of vertices in $\mathbf{x}_n$ may be random, rather than exactly n. As noted at the start of the section, we may assume that each $\mathbf{x}_n$ is deterministic, and in particular that the number v_n of vertices is deterministic. This does not quite give a vertex space, as we need not have $v_n = n$: instead, taking $A = \mathcal{S}$ in (2.4), we have $v_n/n \to 1$. Rescaling the indexing parameter as above, replacing n by v_n (after taking a subsequence if necessary) and multiplying κ_n by a factor v_n/n, does not affect the distribution of the graph, so the resulting kernels are still graphical with limit κ. Hence, our results for vertex spaces apply. In particular, using (8.5) for vertex spaces, we find that

$$\frac{1}{v_n} C_1(G^{\mathcal{V}}(n, \kappa_n)) \xrightarrow{\mathrm{p}} \rho(\kappa).$$

As $n \sim v_n$, this implies (8.5).

We have shown that it suffices to prove (8.5), i.e., part (iii) of Theorem 3.1, for vertex spaces in which the sequences $\mathbf{x}_n$ are deterministic; this was our aim in this subsection. Similar comments apply to all our results.

9. The giant component

In this section we prove our main results, Theorems 3.1 and 3.6 of Subsection 3.1, concerning the existence, size and uniqueness of the giant component in the random graph $G^{\mathcal{V}}(n, \kappa_n)$. The basic strategy will be to relate the neighbourhoods of a vertex of $G^{\mathcal{V}}(n, \kappa_n)$ to the branching process, by exploring these neighbourhoods step by step. In the context of random graphs, this step-by-step exploration and comparison with a branching process, which now is standard, was perhaps first used by Karp [62], who applied it to study the size of the giant component in random directed graphs; similar ideas were used earlier in other contexts, for example by Kendall [63] in the study of epidemics.

Let us first recall some notation. We shall work with the branching process $\mathfrak{X}_\kappa$ defined in Subsection 2.1 and studied in Sections 5 and 6. As before, when the branching process is started with a single particle of type x we denote it $\mathfrak{X}_\kappa(x)$. Unless explicitly stated otherwise, κ will be a kernel on a vertex space $\mathcal{V} = (\mathcal{S}, \mu, (\mathbf{x}_n)_{n \geq 1})$; most of the time we shall not consider generalized vertex spaces. We shall assume that $\kappa \in L^1$, i.e., that $\iint \kappa < \infty$. Any additional assumptions on κ (such as irreducibility) will be stated.

Recall that $\rho_{\geq k}(\kappa;x)$ is the probability that $\mathfrak{X}_\kappa(x)$ contains at least k particles in total (in all generations taken together), and $\rho_k(\kappa;x)$ is the probability that $\mathfrak{X}_\kappa(x)$ contains exactly k particles in total, while $\rho(\kappa;x)$ is the probability that $\mathfrak{X}_\kappa(x)$ survives for eternity, i.e., for infinitely many generations. Starting the process with a particle of random type with distribution μ, the corresponding probabilities for $\mathfrak{X}_\kappa$ are $\rho_{\geq k}(\kappa)$, $\rho_k(\kappa)$, and $\rho(\kappa)$.

A key step in our proofs will be an additional result, relating the fixed-size components of $G^{\mathcal{V}}(n,\kappa_n)$ to the branching process $\mathfrak{X}_\kappa$. As before, we write $N_k(G)$ for the number of vertices of a graph G in components of order k, and $N_{\geq k}(G)$ for $\sum_{j\geq k} N_j(G)$, the number of vertices in components of order at least k.

Theorem 9.1. *Let (κ_n) be a graphical sequence of kernels on a vertex space $\mathcal{V}$ with limit κ. If $k \geq 1$ is fixed, then $N_k\big(G^{\mathcal{V}}(n,\kappa_n)\big)/n \xrightarrow{\mathrm{p}} \rho_k(\kappa)$.*

Remark 9.2. In [25, 84], results similar to Theorem 3.1 were proved (for special κ but with more complicated dependencies) using a careful coupling of the discovery process of the random graph and the limiting branching process; here we shall do this coupling only in the simple case of finitely many types (Example 4.3); the general case will then follow by approximation and monotonicity arguments. In particular, we shall show that any $G_n = G^{\mathcal{V}}(n,\kappa)$ contains a $G'_n = G^{\mathcal{V}}(n,\kappa')$, where κ' may be regarded as a kernel defined on a finite set $\mathcal{S}$, such that $C_1(G'_n)$ is no more than $o_p(n)$ smaller than $C_1(G_n)$; a formal statement is given below. This reduces many questions concerning the very general model $G^{\mathcal{V}}(n,\kappa)$ to the much simpler 'finite-type' case.

Proposition 9.3. *Let (κ_n) be a graphical sequence of kernels on a vertex space $\mathcal{V}$ with quasi-irreducible limit κ. Given any $\varepsilon > 0$, there is a vertex space $\mathcal{V}' = (\mathcal{S}',\mu',(\mathbf{y}_n)_{n\geq 1})$ with $\mathcal{S}'$ finite and a quasi-irreducible kernel κ' on $\mathcal{S}'\times\mathcal{S}'$ with the following properties: $\rho(\kappa') \geq \rho(\kappa) - \varepsilon$, the graphs $G_n = G^{\mathcal{V}}(n,\kappa_n)$ and $G'_n = G^{\mathcal{V}'}(n,\kappa')$ can be coupled so that $G'_n \subseteq G_n$ for sufficiently large n, and $C_1(G'_n)/n \xrightarrow{\mathrm{p}} \rho(\kappa')$.*

The assertion concerning $C_1(G'_n)$ will follow from the other assertions and Theorem 3.1. However, we shall prove Proposition 9.3 as a step towards the proof of Theorem 3.1. This is an example where quasi-irreducibility is forced on us: if we assume κ is irreducible, we still cannot insist that κ' is irreducible.

We now turn to the proofs. We start by giving two elementary results that will be useful below. The first concerns $N_{\geq k}(G)$, the number of vertices of a graph G that are in components of order at least k. Note that for any graph G and any $k \geq 1$,

$$C_1(G) \leq \max\big\{k, N_{\geq k}(G)\big\}, \tag{9.1}$$

since if $C_1(G) \geq k$ then $N_{\geq k}(G) \geq C_1(G)$.

Lemma 9.4. *If $k \geq 2$ and G, G' are two graphs with $G \subseteq G'$, then*

$$N_{\geq k}(G) \leq N_{\geq k}(G') \leq N_{\geq k}(G) + 2k\bigl(e(G') - e(G)\bigr).$$

Proof. If we add a single edge to G, the set of vertices belonging to components of orders $\geq k$ will either remain the same or increase by the inclusion of one or two smaller components; hence $N_{\geq k}(G)$ will increase by at most $2(k-1)$. The result follows by iterating $e(G') - e(G)$ times. □

Lemma 9.5. *As $k \to \infty$, $\rho_{\geq k}(\kappa; x) \searrow \rho(\kappa; x)$ a.e. x, and $\rho_{\geq k}(\kappa) \searrow \rho(\kappa)$.*

Proof. As $\kappa \in L^1$, (5.1) holds a.e. x. By Lemma 5.1, we may assume that (5.1) holds for every x. Then every particle in the branching process $\mathfrak{X}_\kappa$ has a finite number of children, so a particle survives for eternity if and only if it has infinitely many descendants, and the result follows. □

Now we turn to the main part of this section, which concerns the connection between the order of the giant component of $G^{\mathcal{V}}(n, \kappa)$ and the survival probability $\rho(\kappa)$.

We begin by studying the case when $\mathcal{S}$ is finite. It will turn out that this case gives essentially everything, using our monotonicity results and Lemma 9.4. We use the notation in Example 4.3. We shall assume that we have a fixed κ, as in Definition 2.7, rather than a convergent sequence κ_n as in Definition 2.9 and Theorem 3.1. In addition, we shall assume that the matrix κ is irreducible and that $\mu(\{i\}) > 0$ for every i. As observed by Söderberg [88], we then can adapt the standard branching process argument for the classical random graph $G(n, c/n)$, see, e.g., [59, Section 5.2]. The details are as follows.

Lemma 9.6. *Let κ be a kernel on the vertex space $\mathcal{V} = (\mathcal{S}, \mu, (\mathbf{x}_n)_{n\geq 1})$, where $\mathcal{S} = \{1, 2, \dots, r\}$, and suppose that $\mu(\{i\}) > 0$ for every i. Writing G_n for $G^{\mathcal{V}}(n, \kappa)$, if κ is irreducible we have*

$$C_1(G_n)/n \xrightarrow{\mathrm{p}} \rho(\kappa). \tag{9.2}$$

Whether or not κ is irreducible, for any fixed k we have

$$N_{\geq k}\bigl(G_n\bigr)/n \xrightarrow{\mathrm{p}} \rho_{\geq k}(\kappa). \tag{9.3}$$

Proof. Recall that we have n_i vertices of type i, $i = 1, \dots, r$, and that $n_i/n \xrightarrow{\mathrm{p}} \mu_i = \mu\{i\}$. Coupling the graphs (or just the $\mathbf{x}_n$) for different n appropriately, we may of course assume that $n_i/n \to \mu_i$ a.s. From now on we condition on $n_1, \dots, n_r$; we may thus assume that $n_1, \dots, n_r$ are deterministic with $n_i/n \to \mu_i$.

Let $\omega(n)$ be any function such that $\omega(n) \to \infty$ and $\omega(n)/n \to 0$. (Although it might seem more natural to fix $\omega(n) = \log\log n$, say, we shall need this flexibility in the choice of $\omega(n)$ later.) We call a component of $G_n := G^{\mathcal{V}}(n, \kappa)$ *big* if it has at least $\omega(n)$ vertices. Let B be the union of the big components, so $|B| = N_{\geq\omega(n)}(G_n)$.

Fix $\varepsilon > 0$. We may assume that n is so large that $\omega(n)/n < \varepsilon\mu_i$ and $|n_i/n - \mu_i| < \varepsilon\mu_i$ for every i; thus $(1-\varepsilon)\mu_i n < n_i < (1+\varepsilon)\mu_i n$. We may also assume that $n > \max\kappa$, as κ is a function on the finite set $\mathcal{S}\times\mathcal{S}$.

Select a vertex and explore its component in the usual way, one vertex at a time. We first reveal all edges from the initial vertex, and put all neighbours that we find in a list of unexplored vertices; we then choose one of these and reveal its entire neighbourhood, and so on. Stop when we have found at least $\omega(n)$ vertices (so $x \in B$), or when there are no unexplored vertices left (so we have found the entire component and $x \notin B$).

Consider one step in this exploration, and assume that we are about to reveal the neighbourhood of a vertex x of type i. Let us write n'_j for the number of unused vertices of type j remaining. Note that $n_j \ge n'_j \ge n_j - \omega(n)$, so

$$(1-2\varepsilon)\mu_j < n'_j/n < (1+\varepsilon)\mu_j. \tag{9.4}$$

The number of new neighbours of x of type j has a binomial $\mathrm{Bi}\big(n'_j, \kappa(i,j)/n\big)$ distribution, and the numbers for different j are independent. The total variation distance between a binomial $\mathrm{Bi}(n,p)$ distribution and the Poisson distribution with the same mean is at most p, see, e.g., the first inequality in Barbour, Holst and Janson [11, (1.23)]. Hence the total variation distance between the binomial distribution above and the Poisson distribution $\mathrm{Po}\big(\kappa(i,j)n'_j/n\big)$ is at most $\kappa(i,j)/n = O(1/n)$. Also, by (9.4),

$$(1-2\varepsilon)\kappa(i,j)\mu_j \le \kappa(i,j)n'_j/n \le (1+\varepsilon)\kappa(i,j)\mu_j.$$

Since we perform at most $\omega(n)$ steps in the exploration, we may, with an error probability of $O(\omega(n)/n) = o(1)$, couple the exploration with two multi-type branching processes $\mathfrak{X}_{(1-2\varepsilon)\kappa}$ and $\mathfrak{X}_{(1+\varepsilon)\kappa}$ such that the first process always finds at most as many new vertices of each type as the exploration, and the second process finds at least as many. Consequently, for a vertex x of type i,

$$\rho_{\ge\omega(n)}\big((1-2\varepsilon)\kappa; i\big) + o(1) \le \mathbb{P}(x \in B) \le \rho_{\ge\omega(n)}\big((1+\varepsilon)\kappa; i\big) + o(1). \tag{9.5}$$

Note for later (after (9.8)) that, as for any constant C the Poisson distribution with mean C has probability $o(1/n)$ of exceeding $\log n$, the probability that we find more than $\log n$ new neighbours in one step is $O(1/n)$. It follows that the probability that we reach more than $\omega(n) + \log n$ vertices during the exploration is $o(1)$. (Informally, we cannot 'overshoot' by more than $\log n$.)

Since $\omega(n) \to \infty$, by Lemma 9.5 we have $\rho_{\ge\omega(n)}(\kappa'; i) \to \rho(\kappa'; i)$ for every kernel $\kappa' \in L^1$, so we can rewrite (9.5) as

$$\rho\big((1-2\varepsilon)\kappa; i\big) + o(1) \le \mathbb{P}(x \in B) \le \rho\big((1+\varepsilon)\kappa; i\big) + o(1).$$

Letting $\varepsilon \to 0$ we find, using Theorem 6.4, that if x is of type i, then the probability that the component containing x is big satisfies

$$\mathbb{P}(x \in B) \to \rho(\kappa; i). \tag{9.6}$$

(Recall that we are conditioning on the types of the vertices, treating the numbers n_i of vertices of type i as deterministic, and assuming that $n_i/n \to \mu_i$.) Summing over all vertices x we find

$$\begin{aligned} \frac{1}{n}\,\mathbb{E}\,|B| = \frac{1}{n}\sum_x \mathbb{P}(x \in B) &= \frac{1}{n}\sum_{i=1}^r n_i\,\mathbb{P}(x \in B \mid x \text{ is of type } i) \\ &\to \sum_{i=1}^r \mu_i \rho(\kappa;i) = \rho(\kappa). \end{aligned} \tag{9.7}$$

Note that this limit is independent of the choice of $\omega(n)$ in the definition of B. Hence, if we define B' using another such function $\omega'(n)$, it follows from (9.7) (considering $\omega \wedge \omega'$ and $\omega \vee \omega'$) that

$$\mathbb{E}\,|B \triangle B'|/n \to 0. \tag{9.8}$$

Next, start with two distinct vertices x and y, of types i and j, say, and explore their components as above, again stopping each exploration if we find $\omega(n)$ vertices. Assume for the moment that $\omega(n)$ is small, say $\omega(n) = \log n$. The probability that during the truncated exploration we find a connection between the two components is $O(\omega(n)^2/n) + o(1) = o(1)$. (Here we use the fact noted after (9.2), that we are not likely to overshoot: with probability $1 - o(1)$, at every stage, even after stopping the exploration of one component because it has become too large, the explored parts of the components contain at most $\omega(n) + \log n$ vertices.) As before, fix $\varepsilon > 0$. For n large enough, ignoring the possibility of joining the truncated components of x and y, we can couple the two explorations as above with independent branching processes (with $(1 - 3\varepsilon)\kappa$ for the lower bound) to obtain

$$\begin{aligned} &\rho_{\geq\omega(n)}\big((1-3\varepsilon)\kappa;i\big)\rho_{\geq\omega(n)}\big((1-3\varepsilon)\kappa;j\big) + o(1) \\ &\qquad \leq \mathbb{P}(x,y \in B) \leq \rho_{\geq\omega(n)}\big((1+\varepsilon)\kappa;i\big)\rho_{\geq\omega(n)}\big((1+\varepsilon)\kappa;j\big) + o(1). \end{aligned}$$

Letting $\varepsilon \to 0$, it follows, as above, that

$$\mathbb{P}(x,y \in B) \to \rho(\kappa;i)\rho(\kappa;j); \tag{9.9}$$

therefore, summing over all pairs of vertices x, y, we find that

$$\frac{1}{n^2}\,\mathbb{E}\,|B|^2 = \frac{1}{n^2}\sum_{x \neq y} \mathbb{P}(x,y \in B) + \frac{1}{n^2}\,\mathbb{E}\,|B| \to \sum_{i,j=1}^r \mu_i\mu_j\rho(\kappa;i)\rho(\kappa;j) = \rho(\kappa)^2.$$

Combining this and (9.7), we see that $\mathrm{Var}(|B|/n) \to 0$, and thus that

$$|B|/n \xrightarrow{\mathrm{p}} \rho(\kappa). \tag{9.10}$$

So far, we have assumed that $\omega(n)$ was small. However, by (9.8), having proved (9.10) for one choice of $\omega(n)$ it follows that (9.10) holds for every choice of $\omega(n)$ satisfying $\omega(n) \to \infty$ and $\omega(n) = o(n)$.

For any choice of $\omega(n)$ with $\omega(n) \to \infty$ and $\omega(n) = o(n)$, equation (9.10) gives the upper bound on the size $C_1(G_n)$ of the largest component claimed in (9.2), since $C_1(G_n) \le \max\{\omega(n), |B|\}$ by (9.1). In other words, for any $\varepsilon > 0$,

$$|C_1(G_n))|/n \le \rho(\kappa) + \varepsilon \tag{9.11}$$

holds whp.

To obtain the matching lower bound, it remains to show that all but $o_p(n)$ vertices in B belong to a single component. (We note that this is the only place where the irreducibility of κ is needed.) We first consider the simpler case where $\kappa(i,j) > 0$ for every i and j; we shall return to the general case afterwards. We shall reveal the edges in G_n in two rounds: given $0 < \varepsilon < 1$, we may take independent graphs $G_{n,0}$ and $G_{n,1}$ on the same vertex set, with the distributions of $G(n,(1-\varepsilon)\kappa)$ and $G(n,\varepsilon\kappa)$ respectively, so that $G_{n,0} \cup G_{n,1} \subseteq G_n$. We shall think of $G_{n,0}$ as containing almost all the edges of G_n, and $G_{n,1}$ as containing a few edges we initially keep in reserve.

Recalling that $|\mathcal{S}| = r$, set $\omega(n) = rn^{2/3}$, and let B_0 be the union of the big components in $G_{n,0}$. From (9.10), applied with $(1-\varepsilon)\kappa$ in place of κ, whp we have

$$|B_0|/n \ge \rho((1-\varepsilon)\kappa) - \varepsilon. \tag{9.12}$$

We claim that whp all vertices of B_0 lie in a single component in G_n. To see this, we condition on $G_{n,0}$ and use the random graph $G_{n,1}$: let $x, y \in B_0$ be vertices in distinct components C_x, C_y of $G_{n,0}$. As $\omega = rn^{2/3}$, there are $1 \le i, j \le r$ such that C_x contains a set C_x^i of at least $n^{2/3}$ vertices of type i, and C_y a set C_y^j of at least $n^{2/3}$ vertices of type j. Now the probability that $G_{n,1}$ does not contain a $C_x^i - C_y^j$ edge is $(1-\kappa(i,j)/n)^{|C_x^i||C_y^j|} = \exp(-\Omega(n^{1/3})) = o(n^{-2})$. As there are at most n^2 pairs to consider, it follows that whp all vertices of B_0 lie in a single component of G_n, and hence, from (9.12), that

$$|C_1(G_n)|/n \ge \rho((1-\varepsilon)\kappa) - \varepsilon \tag{9.13}$$

holds whp.

The case when some $\kappa(i,j)$ may be zero is only slightly more complicated. This time, we replace $G_{n,1}$ by r independent graphs $G_{n,l}$ with the distribution of $G(n,\varepsilon\kappa/r)$. Given C_x^i and C_y^j as above, the irreducibility of κ implies that there is a sequence of types, $i = i_1, i_2, \dots, i_t = j$, such that $\kappa(i_l, i_{l+1}) > 0$ for all l. As there are only r types, we may suppose that $t \le r+1$ (note that we may have $i = j$). Let $A_1 = C_x^i$, and, for $2 \le l \le t-1$, let A_l be the set of vertices of type i_l adjacent to A_{l-1} in $G_{n,l-1}$. As $|A_1| = \Omega(n^{2/3})$ and $\kappa(i_1,i_2) > 0$, the expected size of A_2 is $\Omega(n^{2/3})$; furthermore, from a standard Chernoff bound, with probability $1 - \exp(-\Omega(n^{2/3})) = 1 - o(n^{-2})$ we have $|A_2| \ge \mathbb{E}\,|A_2|/2$, say. Iterating, we see that for some $c > 0$ we have $|A_{t-1}| \ge cn^{2/3}$ with probability $1 - o(n^{-2})$. Finally, we find an edge in $G(n, t-1)$ from A_{t-1} to C_y^j with very high probability, as above, establishing (9.13) in this case as well.

Letting $\varepsilon \to 0$ and using Theorem 6.4, the right-hand side of (9.13) tends to $\rho(\kappa)$, so (9.13) proves the lower bound on $C_1(G_n)$ claimed in (9.2). Combining this with the upper bound (9.11), equation (9.2) follows.

To prove (9.3), observe that if we replace $\omega(n)$ by a fixed number k in the argument leading to (9.10) above, and use Theorem 6.5 instead of Theorem 6.4, we obtain (9.3) instead of (9.10). Note that this argument has not made use of the irreducibility of κ either. □

Note that the first part of Lemma 9.6 and Theorem 6.2 imply Theorem 3.1 in the case when $\mathcal{S}$ is finite, $\mu\{i\} > 0$ for every $i \in \mathcal{S}$, $\kappa_n = \kappa$ for every n, and κ is irreducible.

We next consider the regular finitary case in Definition 4.4; let us recall the definition. A kernel κ on a vertex space $\mathcal{V}$ is *regular finitary* if $\mathcal{S}$ may be partitioned into a finite number r of μ-continuity sets $S_1, \dots, S_r$ so that κ is constant on each $S_i \times S_j$. A μ-continuity set is a measurable set $A \subseteq \mathcal{S}$ with $\mu(\partial A) = 0$. We next prove an extension of Lemma 9.6 to this regular finitary case.

Lemma 9.7. *Let κ be a regular finitary kernel on a vertex space $\mathcal{V}$, and let $G_n = G^{\mathcal{V}}(n, \kappa)$. Then* (9.3) *holds. If κ is irreducible, then* (9.2) *holds.*

Proof. As noted in Example 4.3, the regular finitary case differs only in notation from the finite case, so it suffices to prove that the conclusions of Lemma 9.6 hold without the assumption that each $\mu\{i\} > 0$. Due to the generality of our model, we cannot just ignore sets of measure zero; see Remark 2.3.

Using the notation of Lemma 9.6, let us say that a type i is *bad* if $\mu_i = 0$, and let $\mathcal{S}' := \{i \in \mathcal{S} : i \text{ is not bad}\}$. Conditioning on the sequences n_i as in the proof of Lemma 9.6, if i is a bad type then $n_i/n \to 0$. Hence, if we eliminate all vertices of bad type, we are left with a random graph $G'_n = G(n', (n'/n)\kappa')$, where κ' is the restriction of κ to $\mathcal{S}' \times \mathcal{S}'$ and $n'/n \to 1$. It is easily seen that $\rho(\kappa') = \rho(\kappa)$, and $\rho_{\geq k}(\kappa') = \rho_{\geq k}(\kappa)$. The expected degree of any vertex is at most $\max \kappa < \infty$, so the expected number of edges with at least one bad endpoint is $o(n)$. Hence, Lemma 9.4 shows that for each fixed k, $\mathbb{E}\big(N_{\geq k}(G_n) - N_{\geq k}(G'_n)\big) = o(n)$. Consequently, (9.3) holds for G_n because it holds for G'_n.

Similarly, applying (9.2) to G'_n, we see that if $\varepsilon > 0$ then $C_1(G_n)/n \geq C_1(G'_n)/n > \rho(\kappa) - \varepsilon$ whp. In the opposite direction, (9.3) yields that for every $\varepsilon > 0$ and $k \geq 1$, whp $N_{\geq k}(G_n)/n \leq \rho_{\geq k}(\kappa) + \varepsilon$, and (9.1) implies $C_1(G_n)/n \leq \rho_{\geq k}(\kappa) + \varepsilon$ whp. Further, by Lemma 9.5, we have $\rho_{\geq k}(\kappa) \searrow \rho(\kappa)$ as $k \to \infty$. Taking k large enough, we find that $C_1(G_n)/n \leq \rho(\kappa) + 2\varepsilon$ whp, and (9.2) follows. □

For technical reasons, we prove a slight extension of Lemma 9.7 to the quasi-irreducible regular finitary case; cf. Remark 2.12.

Lemma 9.8. *Let κ be a regular finitary kernel on a vertex space $\mathcal{V}$. Suppose that κ is quasi-irreducible, i.e., that there is a μ-continuity set $\mathcal{S}' \subseteq \mathcal{S}$ such that κ restricted to $\mathcal{S}'$ is irreducible and $\kappa = 0$ off $\mathcal{S}' \times \mathcal{S}'$. Then* (9.2) *holds for $G_n = G^{\mathcal{V}}(n, \kappa)$.*

Proof. We may ignore all vertices with types not in $\mathcal{S}'$, since they will be isolated, and consider the restriction of our model to $\mathcal{S}'$. Note that we now have n' vertices, with $n'/n \xrightarrow{\mathrm{p}} \mu(\mathcal{S}')$. The case $\mu(\mathcal{S}') = 0$ is trivial, and otherwise we can consider the normalized measure $\mu/\mu(\mathcal{S}')$ on $\mathcal{S}'$ and the kernel $\kappa' = \mu(\mathcal{S}')\kappa$ on $\mathcal{S}' \times \mathcal{S}'$. It is easily checked that Lemma 9.7 implies that (9.2) holds for G_n in this case as well. □

It turns out that most of the work is behind us; roughly speaking, to prove Theorem 3.1 we shall approximate with the regular finitary case and use Lemma 9.7. There are some complications, as we must ensure irreducibility of the approximations, but these have already been dealt with in Section 7: when κ_n is a graphical sequence of kernels with quasi-irreducible limit κ, Lemma 7.3 gives us a sequence of quasi-irreducible regular finitary kernels $\hat{\kappa}_m^-$ approaching κ from below. Furthermore, $\hat{\kappa}_m^- \leq \kappa_n$ when $n \geq m$, so we may and shall assume that

$$G(n, \hat{\kappa}_m^-) \subseteq G(n, \kappa_n) \tag{9.14}$$

for $n \geq m$. This will allow us to apply Lemma 9.4.

We are now in a position to prove our main results. We start with the approximation result Proposition 9.3, which shows that for many purposes we need only consider the finite-type case.

Proof of Proposition 9.3. We use the kernels $\hat{\kappa}_m^-$ constructed in Lemma 7.3. From Lemma 7.3(i)(ii) and Theorem 6.4, if m is large enough then $\hat{\kappa}_m^-$ is quasi-irreducible and $\rho(\hat{\kappa}_m^-) \geq \rho(\kappa) - \varepsilon$. Fix such an m. We may regard the regular finitary kernel $\hat{\kappa}_m^-$ as a kernel on a finite set $\mathcal{S}'$, so the graph $G^{\mathcal{V}}(n, \hat{\kappa}_m^-)$ has the required distribution for G'_n. From Lemma 7.3(iii) and (9.14) we can couple G'_n and G_n so that $G'_n \subseteq G_n$ whenever $n \geq m$. Finally, from Lemma 9.8, we have $C_1(G'_n)/n \xrightarrow{\mathrm{p}} \rho(\kappa')$ as required. □

Next, it will be convenient to prove a restatement of Theorem 9.1.

Lemma 9.9. *Let (κ_n) be a graphical sequence of kernels on a vertex space $\mathcal{V}$ with limit κ, and let $k \geq 1$ be fixed. Then*

$$N_{\geq k}\big(G^{\mathcal{V}}(n, \kappa_n)\big)/n \xrightarrow{\mathrm{p}} \rho_{\geq k}(\kappa). \tag{9.15}$$

Note that Lemma 9.9 immediately implies Theorem 9.1, as $N_k = N_{\geq k} - N_{\geq k+1}$.

Proof. As before, to avoid clutter we suppress the dependence on $\mathcal{V}$, writing $G(n, \cdot)$ for $G^{\mathcal{V}}(n, \cdot)$. We shall also write G_n for $G(n, \kappa_n) = G^{\mathcal{V}}(n, \kappa_n)$. We may assume that $k \geq 2$, since the case $k = 1$ is trivial. We use the $\hat{\kappa}_m^-$ constructed in Lemma 7.3.

For each m, by Lemma 9.7 we have

$$N_{\geq k}\big(G(n,\hat{\kappa}_m^-)\big)/n \xrightarrow{\mathrm{p}} \rho_{\geq k}(\hat{\kappa}_m^-). \tag{9.16}$$

Let $\varepsilon > 0$. Since $\rho_{\geq k}(\hat{\kappa}_m^-) \to \rho_{\geq k}(\kappa)$ as $m \to \infty$ by Theorem 6.5, we can choose m such that $\rho_{\geq k}(\hat{\kappa}_m^-) > \rho_{\geq k}(\kappa) - \varepsilon$. Using (9.14), it follows from (9.16) that whp

$$N_{\geq k}\big(G(n,\kappa_n)\big)/n \geq N_{\geq k}\big(G(n,\hat{\kappa}_m^-)\big)/n > \rho_{\geq k}(\kappa) - \varepsilon, \tag{9.17}$$

proving the lower bound claimed in (9.15).

To prove the upper bound, consider any $\eta > 0$. By monotone convergence,

$$\iint_{\mathcal{S}^2} \hat{\kappa}_m^-(x,y)\,d\mu(x)\,d\mu(y) \to \iint_{\mathcal{S}^2} \kappa(x,y)\,d\mu(x)\,d\mu(y)$$

as $m \to \infty$. Hence we may choose m such that $\iint \kappa - \iint \hat{\kappa}_m^- < \eta/k$. We now fix this m.

By (2.11) and Lemma 8.1 (applied to the bounded kernel $\hat{\kappa}_m^-$), we have

$$\begin{aligned}&\mathbb{E}\big(e(G_n) - e\big(G(n,\hat{\kappa}_m^-)\big)\big)/n \\ &\quad\to \tfrac{1}{2}\iint_{\mathcal{S}^2} \kappa(x,y)\,d\mu(x)\,d\mu(y) - \tfrac{1}{2}\iint_{\mathcal{S}^2} \hat{\kappa}_m^-(x,y)\,d\mu(x)\,d\mu(y) < \eta/2k.\end{aligned}$$

Hence, for n large,

$$\mathbb{E}\big(e(G_n) - e\big(G(n,\hat{\kappa}_m^-)\big)\big)/n < \eta/2k, \tag{9.18}$$

and by (9.14) and Lemma 9.4, $\mathbb{E}\big(N_{\geq k}(G_n) - N_{\geq k}\big(G(n,\hat{\kappa}_m^-)\big)\big)/n < \eta$. Hence, for large n, using also (9.16),

$$\begin{aligned}\mathbb{P}\big(N_{\geq k}(G_n)/n > \rho_{\geq k}(\kappa) + 2\varepsilon\big) &\leq \mathbb{P}\big(N_{\geq k}(G_n)/n > \rho_{\geq k}(\hat{\kappa}_m^-) + 2\varepsilon\big) \\ &\leq \mathbb{P}\big(N_{\geq k}(G(n,\hat{\kappa}_m^-))/n > \rho_{\geq k}(\hat{\kappa}_m^-) + \varepsilon\big) \\ &\qquad + \mathbb{P}\big(\big(N_{\geq k}(G_n) - N_{\geq k}\big(G(n,\hat{\kappa}_m^-)\big)\big)/n > \varepsilon\big) \\ &< \eta + \eta/\varepsilon.\end{aligned}$$

Letting $\eta \to 0$, we find $N_{\geq k}(G_n) \leq \rho_{\geq k}(\kappa) + 2\varepsilon$ whp, which together with (9.17) completes the proof of the lemma. □

As noted above, Theorem 9.1 is just a reformulation of Lemma 9.9. We are now ready to prove Theorem 3.1.

Proof of Theorem 3.1. As noted in Subsection 8.1, without loss of generality we may assume that $\mathcal{V}$ is a vertex space, rather than a generalized vertex space. As above we write G_n for $G^{\mathcal{V}}(n,\kappa_n)$, and consider the approximating kernels $\hat{\kappa}_m^-$ constructed in Lemma 7.3.

We first observe that (2.13), Theorem 6.2 and Lemma 5.8(i) imply that $\rho(\kappa) < 1$, and that $\rho(\kappa) > 0$ if and only if $\|T_\kappa\| > 1$.

Next we prove the upper bound (3.1) on the size of the giant component of G_n. Fix $\varepsilon > 0$. By (9.1) and Lemma 9.9, for every fixed $k \geq 1$, whp

$$C_1(G_n)/n \leq k/n + N_{\geq k}(G_n)/n < \varepsilon + \rho_{\geq k}(\kappa) + \varepsilon. \tag{9.19}$$

By Lemma 9.5, as $k \to \infty$, $\rho_{\geq k}(\kappa) \searrow \rho(\kappa)$. Hence we may choose k so large that $\rho_{\geq k}(\kappa) < \rho(\kappa) + \varepsilon$, and (9.19) yields $C_1(G_n)/n < \rho_{\geq k}(\kappa) + 3\varepsilon$ whp, proving (3.1).

For quasi-irreducible κ, the lower bound on the size of the giant component claimed in (3.2) follows from Proposition 9.3. Alternatively, we may argue as in the proof of Lemma 9.9. Fix $\varepsilon > 0$. By Theorem 6.4, $\rho(\hat{\kappa}_m^-) \to \rho(\kappa)$ as $m \to \infty$, so we can choose m such that $\rho(\hat{\kappa}_m^-) > \rho(\kappa) - \varepsilon$, and then, by Lemma 9.8 and (9.14), whp

$$C_1\bigl(G(n, \kappa_n)\bigr)/n \geq C_1\bigl(G(n, \hat{\kappa}_m^-)\bigr)/n > \rho(\kappa) - \varepsilon.$$

Together with (3.1), this proves the convergence claimed in (3.2).

It remains to prove part (i) of Theorem 3.1. When $\|T_\kappa\| \leq 1$ we have $\rho(\kappa) = 0$, so (3.1) yields $C_1(G_n) = o_p(n)$, as required. Suppose that $\|T_\kappa\| > 1$. Recall that $\hat{\kappa}_m^- \nearrow \kappa$ a.e., by Lemma 7.3(ii). It follows as in the proof of Lemma 5.16 that $\|T_{\hat{\kappa}_m^-}\| > 1$ if m is large enough. Let us fix such an m. As $\hat{\kappa}_m^-$ is of the regular finitary type, there is a finite partition $\mathcal{S} = \bigcup_{i=0}^r \mathcal{S}_i$ of $\mathcal{S}$ into μ-continuity sets such that the restriction κ_i' of $\hat{\kappa}_m^-$ to $\mathcal{S}_i \times \mathcal{S}_i$ is irreducible for $1 \leq i \leq r$, and $\hat{\kappa}_m^-$ is zero a.e. off $\bigcup_{i=1}^r \mathcal{S}_i \times \mathcal{S}_i$. (This can be regarded as an application of Lemma 5.17 with $\mathcal{S}$ finite. However, the lemma is trivial in this case.) As $T_{\hat{\kappa}_m^-}$ operates separately on each $\mathcal{S}_i$, we have $\|T_{\hat{\kappa}_m^-}\| = \max_i \|T_{\kappa_i'}\|$, where $\|T_{\kappa_i'}\|$ is defined either on the generalized ground space $(\mathcal{S}_i, \mu|_{\mathcal{S}_i})$, or, equivalently, by extending κ_i' to $\mathcal{S} \times \mathcal{S}$ by setting $\kappa_i'(x, y) = 0$ if $x \notin \mathcal{S}_i$ or $y \notin \mathcal{S}_i$. In particular, there is an i with $\|T_{\kappa_i'}\| > 1$. Extending κ_i' to $\mathcal{S} \times \mathcal{S}$ as above, κ_i' is a supercritical quasi-irreducible kernel on $\mathcal{S}$ of the regular finitary type, with $\kappa_n \geq \hat{\kappa}_m^- \geq \kappa_i'$ for large n. Hence, by (9.14) and Lemma 9.8, we have

$$C_1(G_n)/n \geq C_1(G(n, \kappa_i'))/n \xrightarrow{\mathrm{p}} \rho(\kappa_i') > 0,$$

and $C_1(G_n) = \Theta(n)$ whp follows, completing the proof. □

We next prove Theorem 3.6, showing that the second largest component has size $o_p(n)$.

Proof of Theorem 3.6. Let (κ_n) be a graphical sequence of kernels on a (generalized) vertex space $\mathcal{V}$ with quasi-irreducible limit κ, and $\omega(n)$ a function satisfying $\omega(n) \to \infty$ and $\omega(n) = o(n)$. Our task is to show that

$$N_{\geq \omega(n)}(G_n) := \sum_{j \geq 1:\, C_j(G_n) \geq \omega(n)} C_j(G_n) = n\rho(\kappa) + o_p(n).$$

Then (3.5) follows by Theorem 3.1. In turn, (3.6) follows immediately, taking $\omega(n) = \log n$, say. As before, we may assume that $\mathcal{V}$ is a vertex space.

Let $\varepsilon > 0$. For an upper bound on $N_{\geq \omega(n)}(G_n)$, fix a large k such that $\rho_{\geq k}(\kappa) < \rho(\kappa) + \varepsilon$. For large n we have $\omega(n) > k$ and thus by Lemma 9.9 whp $N_{\geq \omega(n)}(G_n)/n \leq N_{\geq k}(G_n)/n < \rho(\kappa) + 2\varepsilon$.

For a lower bound, assume that $\rho(\kappa) > 0$. Then, by Theorem 3.1, whp $C_1(G_n) > \frac{1}{2}\rho(\kappa)n > \omega(n)$, so by Theorem 3.1 again, whp $N_{\geq\omega(n)}(G_n)/n \geq C_1(G_n)/n > \rho(\kappa) - \varepsilon$. This is trivially true if $\rho(\kappa) = 0$ too.

Since ε was arbitrary, the proof is complete. □

We now turn to a result giving the distribution of the types of the vertices making up the giant component; to state this, we need some more definitions.

Let $\mathcal{C}_1(G_n)$ be the largest component of $G_n = G^{\mathcal{V}}(n, \kappa_n)$, i.e., the component with most vertices, chosen by any rule if there is a tie. (Thus, if a sequence (G_n) has a unique giant component, then $\mathcal{C}_1(G_n)$ is this giant component.) Let $\nu_n^1 := \frac{1}{n}\sum_{i\in\mathcal{C}_1(G_n)} \delta_{x_i}$ be the random measure with total mass $C_1(G_n)/n$ that describes the distribution of the points x_i corresponding to the vertices in the largest component. We equip the space of finite positive Borel measures on $\mathcal{S}$ with the weak topology; see Appendix A. As before, we write ρ_κ for the function defined by $\rho_\kappa(x) = \rho(\kappa; x)$.

Theorem 9.10. *Let (κ_n) be a graphical sequence of kernels on a (generalized) vertex space $\mathcal{V}$ with quasi-irreducible limit κ. Then $\nu_n^1 \xrightarrow{\text{p}} \mu_\kappa$ in the space of finite measures on $\mathcal{S}$ with the weak topology, where μ_κ is the measure on $\mathcal{S}$ defined by $d\mu_\kappa = \rho_\kappa\, d\mu$. In other words, for every μ-continuity set A,*

$$\nu_n^1(A) = \frac{1}{n}\#\{i \in \mathcal{C}_1(G_n) : x_i \in A\} \xrightarrow{\text{p}} \mu_\kappa(A) = \int_A \rho(\kappa; x)\, d\mu(x), \tag{9.20}$$

where $G_n = G^{\mathcal{V}}(n, \kappa_n)$. Furthermore, if $f : \mathcal{S} \to \mathbb{R}$ is continuous μ-a.e. and satisfies

$$\frac{1}{n}\sum_{i\in V(G_n)} |f(x_i)| \xrightarrow{\text{p}} \int_{\mathcal{S}} |f|\, d\mu < \infty, \tag{9.21}$$

then

$$\frac{1}{n}\sum_{i\in\mathcal{C}_1(G_n)} f(x_i) \xrightarrow{\text{p}} \int_{\mathcal{S}} f\, d\mu_\kappa = \int_{\mathcal{S}} f(x)\rho(\kappa; x)\, d\mu(x). \tag{9.22}$$

In particular, (9.22) *holds for every bounded and μ-a.e. continuous $f : \mathcal{S} \to \mathbb{R}$.*

Condition (9.21) is very natural and often easy to verify; for example, if $\mathcal{V}$ is a vertex space in which the x_i are i.i.d., as in Example 4.8, or a generalized vertex space in which $\mathbf{x}_n$ is a Poisson process, as in Example 4.9, then (9.21) holds for every integrable f by the law of large numbers. Similarly, if $\mathcal{S} = (0, 1]$ and $x_i = i/n$, then (9.21) holds for every decreasing integrable positive f. Note that some restriction on f is needed for (9.22); it is not hard to construct an example where (9.22) fails, and so does (9.21).

Proof. We begin by proving the first statement. We proceed in several steps, as before. Arguing as in Subsection 8.1, we may assume without loss of generality that $\mathcal{V}$ is a vertex space.

First we assume that the conditions of Lemma 9.6 are satisfied: $\mathcal{S}$ is finite, κ is fixed and irreducible, and $\mu_i = \mu(\{i\}) > 0$ for every i. We use the notation of the proof of Lemma 9.6; in particular, ω is some function with $\omega(n) \to \infty$ and $\omega(n) = o(n)$, and B is the set of vertices of $G_n = G^{\mathcal{V}}(n,\kappa)$ in 'big' components, i.e., components of order at least $\omega(n)$.

Let V_i be the set of vertices of type i. The arguments leading to (9.10) in the proof of Lemma 9.6 yield also $|B \cap V_i|/n \xrightarrow{\mathrm{p}} \rho(\kappa;i)\mu_i$; see (9.6) and (9.9).

If $\rho(\kappa) > 0$, then the conclusion (9.2) of Lemma 9.6 implies that whp $\mathcal{C}_1(G_n) \subseteq B$, and thus (9.2) and (9.10) imply that $|B \triangle \mathcal{C}_1(G_n)|/n \xrightarrow{\mathrm{p}} 0$. This is clearly true when $\rho(\kappa) = 0$ too, and implies that $|\mathcal{C}_1(G_n) \cap V_i|/n \xrightarrow{\mathrm{p}} \rho(\kappa;i)\mu_i$ for every i, which is exactly (9.20).

The result extends to the case when some $\mu_i = 0$ as before. Thus (9.20) holds in the irreducible regular finitary case considered in Lemma 9.7, provided A is one of the sets $\mathcal{S}_i$ in the partition or a union of such sets. In fact, A may be any μ-continuity set, since we may replace the partition $\{\mathcal{S}_i\}$ by $\{\mathcal{S}_i \cap A,\ \mathcal{S}_i \setminus A\}$, noting that all parts are μ-continuity sets. Similarly, the extension to the quasi-irreducible case is immediate, as in Lemma 9.8.

We now turn to the general case. Note that $\mu_\kappa(\mathcal{S}) = \rho(\kappa)$. Assume that $\rho(\kappa) > 0$; otherwise the result is trivial (with $\mu_\kappa = 0$) by Theorem 3.1.

Fix a μ-continuity set A. Use a sequence of partitions $\mathcal{P}_m$ as in Lemma 7.1, and consider the finitary approximation $\hat{\kappa}_m^-$ given by Lemma 7.3 for some m. Let ν_{nm}^1 be the random measure ν_n^1 defined for $G(n,\hat{\kappa}_m^-)$. (As before, we suppress the dependence on $\mathcal{V}$.) By the finitary case completed above,

$$\nu_{nm}^1(A) := \frac{1}{n}\#\bigl\{i \in \mathcal{C}_1\bigl(G(n,\hat{\kappa}_m^-)\bigr) : x_i \in A\bigr\} \xrightarrow{\mathrm{p}} \mu_{\hat{\kappa}_m^-}(A) \tag{9.23}$$

for every fixed m.

Let $\varepsilon > 0$ and choose m so large that $\rho(\hat{\kappa}_m^-) > \rho(\kappa) - \varepsilon$ and $\rho(\hat{\kappa}_m^-) > 0$ (see Theorem 6.4). Then, applying Theorem 3.1 to $\hat{\kappa}_m^-$ and (3.6) of Theorem 3.6 to κ, whp $C_1(G(n,\hat{\kappa}_m^-)) > \frac{1}{2}\rho(\hat{\kappa}_m^-)n > C_2(G_n)$. Recalling the coupling (9.14), it follows that the largest component of $G(n,\hat{\kappa}_m^-)$ is contained in the largest component of G_n, i.e., $\mathcal{C}_1(G(n,\hat{\kappa}_m^-)) \subseteq \mathcal{C}_1(G_n)$, and thus $\nu_{nm}^1 \le \nu_n^1$. Consequently, from (9.23), whp

$$\nu_n^1(A) \ge \nu_{nm}^1(A) \ge \mu_{\hat{\kappa}_m^-}(A) - \varepsilon \ge \mu_\kappa(A) - 2\varepsilon, \tag{9.24}$$

because $\mu_\kappa(A) - \mu_{\hat{\kappa}_m^-}(A) \le \mu_\kappa(\mathcal{S}) - \mu_{\hat{\kappa}_m^-}(\mathcal{S}) = \rho(\kappa) - \rho(\hat{\kappa}_m^-) < \varepsilon$.

Since $\mathcal{S} \setminus A$ also is a μ-continuity set, we may replace A by $\mathcal{S} \setminus A$ in (9.24) and obtain that whp

$$\nu_n^1(\mathcal{S} \setminus A) \ge \mu_\kappa(\mathcal{S} \setminus A) - 2\varepsilon.$$

Since $\nu_n^1(\mathcal{S}) = C_1(G_n)/n$ and $\mu_\kappa(\mathcal{S}) = \rho(\kappa)$, this and Theorem 3.1 show that whp

$$\nu_n^1(A) = C_1(G_n)/n - \nu_n^1(\mathcal{S} \setminus A) \le \rho(\kappa) + \varepsilon - \mu_\kappa(\mathcal{S} \setminus A) + 2\varepsilon = \mu_\kappa(A) + 3\varepsilon.$$

This and (9.24) yield $\nu_n^1(A) \xrightarrow{\mathrm{p}} \mu_\kappa(A)$, so we have shown that (9.20) holds for this A. We have shown that (9.20) holds for an arbitrary μ-continuity set A, which yields $\nu_n^1 \xrightarrow{\mathrm{p}} \mu_\kappa$ by Lemma A.2.

Turning to the second part of the lemma, note that the left-hand sides of (9.21) and (9.22) are equal to $\int |f|\,d\nu_n$ and $\int f\,d\nu_n^1$, respectively. If f is bounded and μ-a.e. continuous, Lemma A.2 thus shows that these relations follow from (2.4) and (9.20), respectively.

To deduce (9.22) from (9.21) for unbounded f, we use the truncations $f_M := (|f| \wedge M)\,\mathrm{sign}(f)$. Let $\varepsilon > 0$. By monotone convergence, $\int |f_M|\,d\mu \to \int |f|\,d\mu$ as $M \to \infty$. Thus, we can choose M such that $\int |f_M|\,d\mu > \int |f|\,d\mu - \varepsilon$. Since (9.21) holds for bounded μ-a.e. continuous functions, it holds for f_M, so

$$\frac{1}{n}\sum_{i\in V(G_n)} \big(|f(x_i)| - |f_M(x_i)|\big) \xrightarrow{\mathrm{p}} \int_{\mathcal{S}} |f|\,d\mu - \int_{\mathcal{S}} |f_M|\,d\mu < \varepsilon.$$

Hence the left-hand side is at most ε whp. Consequently, whp

$$\begin{aligned}\left|\frac{1}{n}\sum_{i\in\mathcal{C}_1(G_n)} f(x_i) - \frac{1}{n}\sum_{i\in\mathcal{C}_1(G_n)} f_M(x_i)\right| &\le \frac{1}{n}\sum_{i\in\mathcal{C}_1(G_n)} \big|f(x_i) - f_M(x_i)\big| \\ \le \frac{1}{n}\sum_{i\in V(G_n)} \big|f(x_i) - f_M(x_i)\big| &= \frac{1}{n}\sum_{i\in V(G_n)} \big(|f(x_i)| - |f_M(x_i)|\big) \le \varepsilon,\end{aligned}$$

with the first two inequalities holding unconditionally. Note that

$$\left|\int_{\mathcal{S}} f\,d\mu_\kappa - \int_{\mathcal{S}} f_M\,d\mu_\kappa\right| \le \int_{\mathcal{S}} |f - f_M|\,d\mu = \int_{\mathcal{S}} |f|\,d\mu - \int_{\mathcal{S}} |f_M|\,d\mu < \varepsilon.$$

Since $\varepsilon > 0$ is arbitrary, and (9.22) holds for each f_M, relation (9.22) for f follows by a standard 3ε-argument. ☐

10. Edges in the giant component

The main aim of this section is to prove Theorem 3.5, which claims that if κ_n is a graphical sequence of kernels on a (generalized) vertex space $\mathcal{V}$ with quasi-irreducible limit κ, then $\frac{1}{n}e(\mathcal{C}_1(G^{\mathcal{V}}(n,\kappa_n))) \xrightarrow{\mathrm{p}} \zeta(\kappa)$, where $\zeta(\kappa)$ is defined in (3.3) as

$$\zeta(\kappa) := \frac{1}{2}\iint_{\mathcal{S}^2} \kappa(x,y)\big(\rho(\kappa;x) + \rho(\kappa;y) - \rho(\kappa;x)\rho(\kappa;y)\big)\,d\mu(x)\,d\mu(y). \quad (10.1)$$

Before turning to the proof, we briefly examine the behaviour of $\zeta(\kappa)$, giving two alternative formulae for $\zeta(\kappa)$, together with upper and lower bounds in terms of ρ_κ.

From the symmetry of κ and the definition (2.14) of T_κ, (10.1) is equivalent to

$$\zeta(\kappa) = \int_{\mathcal{S}} (1 - \rho_\kappa/2) T_\kappa \rho_\kappa\,d\mu,$$

where, as usual, ρ_κ is the function defined by $\rho_\kappa(x) := \rho(\kappa; x)$. By relation (6.2) of Theorem 6.2 and the definition of Φ_κ in (2.17), it follows that

$$\zeta(\kappa) = \int_{\mathcal{S}} \Big(1 - \frac{\rho(\kappa;x)}{2}\Big) \ln\Big(\frac{1}{1-\rho(\kappa;x)}\Big)\, d\mu(x). \tag{10.2}$$

Writing G_n for $G^{\mathcal{V}}(n, \kappa_n)$, note that the assumptions of Theorem 3.5 include convergence of the expectation of $e(G_n)/n$. As shown in Proposition 8.9, an easy consequence of these assumptions is that

$$e(G_n)/n \xrightarrow{\mathrm{p}} \frac{1}{2} \iint_{\mathcal{S}^2} \kappa. \tag{10.3}$$

In the light of (10.3), relation (3.4) is equivalent to the assertion that number of edges not in the giant component is

$$\frac{n}{2} \iint_{\mathcal{S}^2} (1-\rho(\kappa;x))\kappa(x,y)(1-\rho(\kappa;y))\, d\mu(x)\, d\mu(y) + o_p(n).$$

In any connected graph, the number of edges is at least the number of vertices minus 1; hence $\zeta(\kappa) \ge \rho(\kappa)$. In fact, Theorem 6.2 has the following simple consequence.

Proposition 10.1. *Let κ be a kernel on a (generalized) ground space $(\mathcal{S}, \mu)$. Then*

$$\rho(\kappa) \le \zeta(\kappa) \le \tfrac{1}{2}(\|T_\kappa\| + 1)\rho(\kappa) \le \|T_\kappa\|\rho(\kappa).$$

Furthermore, the first two inequalities are strict when $\rho(\kappa) > 0$.

Proof. If $0 < s < 1$, then $s < (1 - s/2)\ln(1/(1-s))$, as is easily verified by computing the Taylor series. Thus,

$$\rho(\kappa;x) \le \Big(1 - \frac{\rho(\kappa;x)}{2}\Big) \ln\Big(\frac{1}{1-\rho(\kappa;x)}\Big),$$

with strict inequality when $\rho(\kappa;x) > 0$. Integrating with respect to μ, the left-hand side becomes $\rho(\kappa)$, while, from (10.2), the right-hand side becomes $\zeta(\kappa)$. Thus $\rho(\kappa) \le \zeta(\kappa)$, with strict inequality if $\rho(\kappa) > 0$.

In the other direction, if $0 < s < 1$, then $(1-s)\ln(1/(1-s)) < s - s^2/2$, as can again be verified by computing the Taylor series. Hence,

$$\Big(1 - \frac{s}{2}\Big) \ln\Big(\frac{1}{1-s}\Big) < s - \frac{s^2}{2} + \frac{s}{2}\ln\Big(\frac{1}{1-s}\Big).$$

Substituting $s = \rho(\kappa;x)$ and integrating, it follows that

$$\zeta(\kappa) \le \int_{\mathcal{S}} \bigg(\rho(\kappa;x) - \frac{1}{2}\rho(\kappa;x)^2 + \frac{1}{2}\rho(\kappa;x)\ln\Big(\frac{1}{1-\rho(\kappa;x)}\Big)\bigg)\, d\mu(x),$$

with strict inequality when $\rho(\kappa;x) > 0$. Writing ρ_κ for the function defined by $\rho_\kappa(x) := \rho(\kappa;x)$, from (6.2) and the definition (2.17) of Φ_κ, we have

$$\ln\Big(\frac{1}{1-\rho(\kappa;x)}\Big) = (T_\kappa \rho_\kappa)(x).$$

It follows that

$$\begin{aligned}\zeta(\kappa) &\le \rho(\kappa) - \tfrac{1}{2}\langle \rho_\kappa, \rho_\kappa\rangle + \tfrac{1}{2}\langle \rho_\kappa, T_\kappa \rho_\kappa\rangle \\ &\le \rho(\kappa) + \tfrac{1}{2}\bigl(\|T_\kappa\| - 1\bigr)\int_{\mathcal{S}} \rho_\kappa^2 \, d\mu(x) \le \tfrac{1}{2}\bigl(\|T_\kappa\|+1\bigr)\rho(\kappa),\end{aligned}$$

with strict inequality unless $\zeta(\kappa) = \rho(\kappa) = 0$. □

Our proof of Theorem 3.5 will be very similar to that of Theorem 3.1, except that we need to consider certain branching process expectations $\sigma(\kappa)$ and $\sigma_{\ge k}(\kappa)$ in place of $\rho(\kappa)$ and $\rho_{\ge k}(\kappa)$. In preparation for the proof, we shall relate $\zeta(\kappa)$ to the branching process $\mathfrak{X}_\kappa$ via $\sigma(\kappa)$. As before, we assume that κ is a kernel on $(\mathcal{S}, \mu)$ with $\kappa \in L^1$; in particular, it is convenient here to normalize so that $\mu(\mathcal{S}) = 1$.

Let A be a Poisson process on $\mathcal{S}$, with intensity given by a finite measure λ, so that A is a random multi-set on $\mathcal{S}$. If g is a bounded measurable function on multi-sets on $\mathcal{S}$, it is easy to see that

$$\mathbb{E}\left(|A| g(A)\right) = \int_{\mathcal{S}} \mathbb{E}\, g(A \cup \{y\}) \, d\lambda(y). \tag{10.4}$$

(This is a simple consequence of the well-known fact that the Palm distribution equals the distribution of $A \cup \{y\}$. To show (10.4) directly, note that we may construct A as follows: first decide the total number N of points in A, according to a Poisson $\mathrm{Po}(c)$ distribution with mean $c = \lambda(\mathcal{S})$. Then let $(a_i)_{i=1}^N$ be a sequence of i.i.d. random points of $\mathcal{S}$, each distributed according to the normalized form λ/c of λ, and take $A = \{a_1, \dots, a_N\}$. Let ν be the measure (on finite sequences of points in $\mathcal{S}$) associated to $(a_i)_{i=1}^N$, and let ν' be the measure with density $N\, d\nu$. Recalling that if Z has a $\mathrm{Po}(c)$ distribution, then $k\,\mathbb{P}(Z = k) = c\,\mathbb{P}(Z-1 = k)$, we find that ν'/c may be constructed by taking $N-1$ to have a $\mathrm{Po}(c)$ distribution, and then taking the a_i i.i.d. as before, or, equivalently, by constructing a sequence according to ν and appending a new random point with the distribution λ/c. Neglecting the order of the points, (10.4) follows.)

Let $X(x)$ denote the first generation of the branching process $\mathfrak{X}_\kappa(x)$. Thus $X(x)$ is given by a Poisson process on $\mathcal{S}$ with intensity $\kappa(x,y)\, d\mu(y)$. Suppose that (5.1) holds, so $X(x)$ is finite. Let $\sigma(\kappa; x)$ denote the expectation of $|X(x)| \mathbf{1}[|\mathfrak{X}_\kappa(x)| = \infty]$, recalling that under the assumption (5.1), the branching process $\mathfrak{X}_\kappa(x)$ dies out if and only if $|\mathfrak{X}_\kappa(x)| < \infty$. Then

$$\begin{aligned}\int_{\mathcal{S}} \kappa(x,y)\,d\mu(y) - \sigma(\kappa;x) &= \mathbb{E}\left(|X(x)|\mathbf{1}[|\mathfrak{X}_\kappa(x)| < \infty]\right) \\ &= \mathbb{E}\left(|X(x)| \prod_{z\in X(x)} (1-\rho(\kappa;z))\right) \\ &= \int_{\mathcal{S}} \kappa(x,y)(1-\rho(\kappa;y))\,\mathbb{E}\left(\prod_{z\in X(x)} (1-\rho(\kappa;z))\right) d\mu(y) \\ &= \int_{\mathcal{S}} \kappa(x,y)(1-\rho(\kappa;y))(1-\rho(\kappa;x))\,d\mu(y).\end{aligned}$$

Here the penultimate step is from (10.4); the last step uses the fact that the branching process dies out if and only if none of the children of the initial particle survives. Writing X for the first generation of $\mathfrak{X}_\kappa$, let

$$\sigma(\kappa) := \mathbb{E}\left(|X|\mathbf{1}[|\mathfrak{X}_\kappa| = \infty]\right) = \int_{\mathcal{S}} \sigma(\kappa;x)\,d\mu(x).$$

Then, integrating over x and subtracting from $\iint \kappa(x,y)$, we obtain

$$\sigma(\kappa) = \iint_{\mathcal{S}^2} \kappa(x,y)\bigl(1-(1-\rho(\kappa;x))(1-\rho(\kappa;y))\bigr)\,d\mu(y)\,d\mu(x), \tag{10.5}$$

i.e., $\sigma(\kappa) = 2\zeta(\kappa)$, where $\zeta(\kappa)$ is defined in (3.3).

Lemma 10.2. *Let κ be a quasi-irreducible kernel on a ground space $(\mathcal{S},\mu)$, with $\kappa \in L^1$. If $(\kappa_n)_1^\infty$ is a sequence of kernels that increase to κ a.e., then $\sigma(\kappa_n) \to \sigma(\kappa) < \infty$.*

Proof. This is immediate from Theorem 6.4(i), (10.5), the fact that $\iint \kappa < \infty$, and dominated convergence. ☐

As we shall see next, $\sigma(\kappa)$ is the limit of the expectations

$$\sigma_{\geq k}(\kappa) := \mathbb{E}\left(|X|\mathbf{1}[|\mathfrak{X}_\kappa| \geq k]\right).$$

Lemma 10.3. *With $\kappa \in L^1$ fixed,*

$$\sigma_{\geq k}(\kappa) \searrow \sigma(\kappa) \quad \textit{as } k \to \infty. \tag{10.6}$$

Proof. We have $|X| \geq |X|\mathbf{1}[|\mathfrak{X}_\kappa| \geq k] \searrow |X|\mathbf{1}[|\mathfrak{X}_\kappa| = \infty]$. As $\mathbb{E}\,|X| = \iint \kappa(x,y) < \infty$, the result follows by dominated convergence. ☐

Using the above lemmas we can prove Theorem 3.5. As the argument is very similar to that for Theorem 3.1, we give only an outline.

Proof of Theorem 3.5. As usual, we may assume without loss of generality that $\mathcal{V}$ is a vertex space. Let $M_{\geq k}(G)$ denote the number of edges of a graph G that lie in components of order at least k.

We start with the case when $\mathcal{S}$ is finite and κ is irreducible, writing G_n for $G^{\mathcal{V}}(n,\kappa_n)$. Let $d(x)$ denote the degree of a vertex x of G_n. Using the local coupling of the neighbourhood of a random vertex x to the branching process $\mathfrak{X}_\kappa$ described in the proof of Lemma 9.6, considering $\mathbb{E}(d(x)\mathbf{1}[x \in B])$ in place of $\mathbb{P}(x \in B)$, the proof of Lemma 9.6 yields the relations

$$2e(\mathcal{C}_1(G_n))/n \xrightarrow{\mathrm{p}} \sigma(\kappa) \tag{10.7}$$

and

$$2M_{\geq k}(G_n)/n \xrightarrow{\mathrm{p}} \sigma_{\geq k}(\kappa), \tag{10.8}$$

corresponding to (9.2) and (9.3). As before, the same formulae in the quasi-irreducible regular finitary setting of Lemma 9.8 follow.

To complete the proof, we consider the approximating kernels $\hat{\kappa}_m^-$ constructed in Lemma 7.3. By Lemma 7.3(ii) and Lemma 10.2 we have $\sigma(\hat{\kappa}_m^-) \to \sigma(\kappa)$. Applying (10.7) to $\hat{\kappa}_m^-$ and using the coupling $G^{\mathcal{V}}(n, \hat{\kappa}_m^-) \subseteq G^{\mathcal{V}}(n, \kappa_n)$, $n \geq m$, it follows that for any $\varepsilon > 0$,

$$e(\mathcal{C}_1(G_n))/n \geq \sigma(\kappa)/2 - \varepsilon \tag{10.9}$$

holds whp. This is exactly the lower bound claimed in (3.4).

For the upper bound, we claim first that, for each fixed k,

$$2M_{\geq k}(G^{\mathcal{V}}(n, \kappa_n))/n \xrightarrow{\mathrm{p}} \sigma_{\geq k}(\kappa).$$

The argument is exactly as for (9.15), except that in place of (9.18) we show that there is an m for which $\mathbb{E}(e(G_n) - e(G(n, \hat{\kappa}_m^-)))/n < \eta/(2k^2)$, and in place of Lemma 9.4 we use the fact that, for $k \geq 1$, adding an edge to a graph cannot change $M_{\geq k}$ by more than $2\binom{k-1}{2} + 1 \leq k^2$. The rest of the proof is as for Theorem 3.1, using

$$e(\mathcal{C}_1(G_n))/n \leq k^2/n + M_{\geq k}(G_n)/n$$

in place of (9.19) and Lemma 10.3 instead of Lemma 9.5. □

11. Stability

This section is devoted to the proof of the 'stability' result, Theorem 3.9, which states that deleting a few vertices and their incident edges, and then adding or deleting a few edges, does not change the size of the giant component of $G_n = G^{\mathcal{V}}(n, \kappa_n)$ significantly. As usual, without loss of generality we may restrict our attention to the case where $\mathcal{V}$ is a vertex space; we shall return to this later. For the moment, we shall ignore vertex deletion; our aim is thus to prove the following special case of Theorem 3.9.

Theorem 11.1. *Let (κ_n) be a graphical sequence of kernels on a vertex space $\mathcal{V}$ with irreducible limit κ, and let $G_n = G^{\mathcal{V}}(n, \kappa_n)$. For every $\varepsilon > 0$ there is a $\delta > 0$ (depending on κ) such that, whp,*

$$(\rho(\kappa) - \varepsilon)n \leq C_1(G_n') \leq (\rho(\kappa) + \varepsilon)n \tag{11.1}$$

for every graph G_n' on $V(G_n) = [n]$ with $e(G_n' \mathbin{\triangle} G_n) \leq \delta n$.

We shall see later (at the end of Subsection 11.1) that Theorem 3.9 follows. As noted in Subsection 3.2, to prove Theorem 11.1 it suffices to consider separately the cases where edges are added and where edges are deleted. More precisely, as $G'_n \cap G_n \subseteq G'_n \subseteq G'_n \cup G_n$, it suffices to prove the upper bound in (11.1) for $G'_n \supseteq G_n$, and the lower bound for $G'_n \subseteq G_n$.

The upper bound is easy. Indeed, by Lemma 9.5, $\rho_{\geq k}(\kappa) \searrow \rho(\kappa)$ as $k \to \infty$. Thus, given $\varepsilon > 0$, we may choose k such that $\rho_{\geq k}(\kappa) \leq \rho(\kappa) + \varepsilon/3$. By Lemma 9.9, whp $N_{\geq k}(G_n) \leq (\rho(\kappa) + \varepsilon/2)n$. Taking $\delta = \varepsilon/4k$, it follows by Lemma 9.4 that whp $N_{\geq k}(G'_n) \leq (\rho(\kappa) + \varepsilon)n$, which implies the upper bound in (11.1).

For the lower bound, our aim is to show that whp

$$C_1(G_n - E) \geq (\rho(\kappa) - \varepsilon)n \tag{11.2}$$

for every $E \subseteq E(G_n)$ with $|E| \leq \delta n$.

We may assume that $\rho(\kappa) > 0$, as otherwise there is nothing to prove. As in the proof of Theorem 3.1, it suffices to consider the regular finitary case; in fact, given $\varepsilon > 0$, by Proposition 9.3 there is a vertex space $\mathcal{V}'$ with finite type space and a quasi-irreducible kernel κ' on $\mathcal{V}'$ with $\rho(\kappa') > \rho(\kappa) - \varepsilon/2$ such that we may consider $G^{\mathcal{V}'}(n, \kappa')$ as a subgraph of G_n. It suffices to prove that there is a $\delta > 0$ such that removing at most δn edges from $G^{\mathcal{V}'}(n, \kappa')$ leaves whp a graph with a component of order at least $(\rho(\kappa') - \varepsilon/2)n$. Replacing ε by 2ε, this is exactly (11.2), but with $G^{\mathcal{V}}(n, \kappa_n)$ replaced by $G^{\mathcal{V}'}(n, \kappa')$. Thus we may assume that $G_n = G^{\mathcal{V}}(n, \kappa)$, where κ is a quasi-irreducible kernel on a finite set $\mathcal{S} = \{1, 2, \dots, r\}$. In fact, by rescaling, as in the proof of Lemma 9.8, we may assume that κ is irreducible. Finally, as in the proof of Lemma 9.7, we may assume that $\mu\{i\} > 0$ for every i, as there are $o_p(n)$ edges incident with types i with $\mu\{i\} = 0$. In other words, we may assume the setting of Lemma 9.6. We shall do so for the rest of this section; thus $G_n = G^{\mathcal{V}}(n, \kappa)$, where $\mathcal{V} = (\mathcal{S}, \mu, (\mathbf{x}_n)_{n \geq 1})$ is a vertex space, and

$$\mathcal{S} = \{1, 2, \dots, r\}, \ \mu\{i\} > 0 \ \forall i, \ \kappa \text{ is irreducible, and } \|T_\kappa\| > 1. \tag{11.3}$$

In a paper studying the bisection width of sparse random graphs, Luczak and McDiarmid [70] proved (11.2) for the Erdős-Rényi case, where $|\mathcal{S}| = 1$ or κ is constant. Their proof adapts easily to the finite-type case, from which, as shown above, Theorem 11.1 follows. We present this proof in Subsection 11.1.

A different, perhaps more natural, approach to proving (11.2) is to work with the branching process $\mathfrak{X}_\kappa$, using the coupling of vertex neighbourhoods in G_n with $\mathfrak{X}_\kappa$ to reduce (11.2) to an equivalent statement for the two-core, Lemma 11.10 below. The latter statement has a very simple proof in the uniform case. We present this approach here, in Subsection 11.2 below, because the intermediate results, relating properties of the two-core to the branching process, are likely to be of interest in their own right. Unfortunately, while Lemma 11.10 can be proved in the general case by branching process methods, our proof is rather complicated. As the result follows from Theorem 11.1, which can be proved more simply by the method of Luczak and McDiarmid, we omit the proof. A reader interested only in the proof of Theorem 11.1 can safely omit Subsection 11.2.

11.1. **Counting cuts in the giant component.** In this subsection we prove Theorem 11.1, and then deduce Theorem 3.9. Apart from the straightforward adaptations to non-constant κ, the argument for Theorem 11.1 is that of Luczak and McDiarmid [70]. We start with a deterministic lemma whose statement and proof are taken verbatim from [70].

Lemma 11.2. *For any $\varepsilon > 0$, there exist $\eta_0 = \eta_0(\varepsilon) > 0$ and n_0 such that the following holds. For all $n \geq n_0$, and for all connected graphs G with n vertices, there are at most $(1+\varepsilon)^n$ bipartitions of G with at most $\eta_0 n$ cross edges.*

Proof. Let T be an arbitrary spanning tree of G. Any 2-partition S, $\overline{S}$ of T is determined uniquely by the corresponding set of cross edges, together with the specification for each cross edge of which of its endpoints is in S. For as T is connected, the cross edges specify a nonempty subset S^* of S, and then S is the set of vertices v such that there is a path from v to one of the vertices in S^* where this path does not use any of the cross edges. (If $v \in \overline{S}$ then no path from v to S^* can avoid the cross edges, and if $v \in S$ then any shortest path from v to S^* avoids the cross edges.) Hence, since T has $n-1$ edges, the number of 2-partitions of T (and hence also of G) with at most ηn cross edges is no more than

$$\sum_{j \leq \eta n} 2^j \binom{n}{j} = O\left(n 2^{\eta n} \left((1-\eta)^{1-\eta}\eta^\eta\right)^{-n}\right),$$

assuming $\eta \leq 1/2$. Now let $\varepsilon > 0$. As $\eta \to 0$, $2^\eta/((1-\eta)^{1-\eta}\eta^\eta) \to 1$. Hence, for η sufficiently small and n sufficiently large, there are at most $(1+\varepsilon)^n$ partitions with at most ηn cross edges. □

Recall our assumptions (11.3), that $\mathcal{S} = \{1, 2, \ldots, r\}$, $\mu\{i\} > 0$ for every i, κ is irreducible and $\|T_\kappa\| > 1$. As usual, we condition on $\mathbf{x}_n$, so we may assume that $\mathbf{x}_n$ is deterministic for every n, so there are n_i vertices of type i, with $n_i/n \to \mu\{i\} > 0$ as $n \to \infty$.

The main additional ingredient needed to adapt the proof of [70] to non-constant kernels is the following simple lemma.

Lemma 11.3. *Suppose that the assumptions* (11.3) *hold. For any $\varepsilon > 0$ there is a $\theta = \theta(\kappa, \varepsilon) > 0$ with the following property. If n is large enough then, whenever V_1, V_2 are disjoint sets of at least εn vertices of $G_n = G(n, \kappa)$ such that $V_1 \cup V_2$ contains at least εn vertices of each type, the expected number of edges from V_1 to V_2 in G_n is at least θn.*

Proof. We assume that $n \geq \max \kappa$. Let $\varepsilon' = \min\{\varepsilon/r, \varepsilon/2\}$, and let

$$\theta = (\varepsilon')^2 \min\{\kappa(i,j) : \kappa(i,j) > 0\} > 0.$$

There are types i and j such that V_1 contains at least $\varepsilon' n$ vertices of type i, and V_2 at least $\varepsilon' n$ vertices of type j. As κ is irreducible, there is a sequence $i = i_0, i_1, \dots, i_t = j$ such that $\kappa(i_s, i_{s+1}) > 0$ for each s. For each i_s, from our condition on $V_1 \cup V_2$, one or both of V_1 and V_2 must contain at least $\varepsilon' n$ vertices of type i_s. It follows that for some s, V_1 contains at least $\varepsilon' n$ vertices of type i_s, and V_2 contains at least $\varepsilon' n$ vertices of type i_{s+1}. But then the expected number of edges from V_1 to V_2 is at least $(\varepsilon' n)^2 \kappa(i_s, i_{s+1})/n \geq \theta n$, as required. □

Using Lemma 11.3, the proof of Lemma 2 in [70] adapts immediately to our setting. Note that we use different notation (in particular, Greek letters) from [70], for consistency with the rest of the present paper.

Proof of Theorem 11.1. As noted at the start of the section, it suffices to prove (11.2), assuming that (11.3) holds.

Given $\delta, \varepsilon > 0$, by an (ε, δ)-*cut* in a graph G we shall mean a partition $(W, \overline{W})$ of the vertex set of G with $|W|$, $|\overline{W}| \geq \varepsilon |G|$, such that G contains at most $\delta |G|$ edges from W to $\overline{W}$. We know from Theorem 3.1 that $\frac{1}{n} C_1(G_n) \xrightarrow{\mathrm{p}} \rho(\kappa) > 0$, so proving (11.2) is equivalent to showing that for any $\varepsilon > 0$ there is a $\delta = \delta(\varepsilon) > 0$ such that whp the giant component of G_n has no (ε, δ)-cut.

Given $0 < \gamma < 1$, let G_1, G_2 be independent graphs with the distributions of $G(n, (1-\gamma)\kappa)$ and $G(n, \gamma\kappa)$, respectively. We may and shall couple the pair (G_1, G_2) with $G_n \sim G(n, \kappa)$ so that $G_1 \cup G_2 \subseteq G_n$. (The union has almost the distribution of G_n; the only difference arises from the possibility of G_1 and G_2 sharing edges.)

Fix $\varepsilon > 0$. Recall that κ is supercritical, so $\rho(\kappa) > 0$. Furthermore, κ is irreducible, so $\rho(\kappa; i) > 0$ for each i. By Theorem 6.4, for each i we have $\rho((1-\gamma)\kappa; i) \nearrow \rho(\kappa; i)$ as $\gamma \to 0$. Let us fix a γ such that

$$\rho((1-\gamma)\kappa; i) \geq (1 - \varepsilon/3)\rho(\kappa; i)$$

holds for every i. Thus, $\rho((1-\gamma)\kappa) \geq (1-\varepsilon/3)\rho(\kappa)$.

Following (in this respect) the notation of [70], let U and U_1 denote the largest components of G_n and G_1 respectively, chosen according to any rule if there is a tie. Then, by Theorems 3.1 and 3.6, the events

$$A_1 := \{|U_1| \geq (1-\varepsilon/2)\rho(\kappa)n\}$$

and

$$A_2 := \{U_1 \subseteq U \text{ and } |U_1| \geq (1-\varepsilon/2)|U|\}$$

hold whp; for the condition $U_1 \subseteq U$, note that U_1 must be contained in some component of G_n, and whp only U is large enough.

Let $\varepsilon_1 = \min\{\rho(\kappa; i)\mu\{i\} : i \in \mathcal{S}\}/2 > 0$. By Theorem 9.10, the event

$$A_1' := \{U_1 \text{ contains at least } \varepsilon_1 n \text{ vertices of each type } i\}$$

holds whp. Without loss of generality, we may assume that $\varepsilon < \varepsilon_1$. Let $\nu = \gamma\theta(\kappa, \varepsilon\rho(\kappa)/2)$, where θ is the function appearing in Lemma 11.3. If n is large enough then, by Lemma 11.3, whenever A_1' holds, if we partition the vertex set of U_1 into two parts V_1, V_2 each of size at least $\varepsilon\rho(\kappa)n/2 < \varepsilon_1 n$, then the expected number of edges in G_2 from V_1 to V_2 is at least νn.

Continuing exactly as in [70], but keeping our notation for the relevant constants, let $\eta > 0$ satisfy

$$1 + 2\eta \le \exp(\nu/8),$$

and let $\delta > 0$ be the minimum of $\nu/4$ and $\frac{1}{2}\eta_0(\eta)$ (from Lemma 11.2). Let

$$A_3 := \{U \text{ has an } (\varepsilon, \delta)\text{-cut in } G_n\},$$

and

$$A_4 := \{U_1 \text{ has an } (\varepsilon/2, 2\delta)\text{-cut in } G_n\}.$$

We claim that $A_2 \cap A_3 \subseteq A_4$. Indeed, suppose that A_2 holds and that U has an (ε, δ)-cut into $B \cup C$. Let $B_1 = B \cap U_1$ and $C_1 = C \cap U_1$. Then U_1 has a partition into $B_1 \cup C_1$, both $|B_1|$ and $|C_1|$ are at least

$$\varepsilon|U| - (|U| - |U_1|) \ge \varepsilon|U_1|/2,$$

and the number of cross edges is at most $\delta|U| \le 2\delta|U_1|$, so A_4 holds, proving the claim. As A_2 holds whp, and our aim is to show that $\mathbb{P}(A_3) \to 0$, it thus suffices to show that $\mathbb{P}(A_4) \to 0$.

Let us condition on G_1, assuming that $A = A_1 \cap A_1'$ holds. By Lemma 11.2 and our choice of δ, there are at most $(1+\eta)^n$ $(\varepsilon/2, 2\delta)$-cuts of U_1 in G_1. Consider any one such cut, partitioning U_1 into $B \cup C$, say. Let X_2 be the number of edges of G_2 from B to C. Recalling that G_1 and G_2 are independent, as noted above, $\mathbb{E}(X_2) \ge \nu n \ge 4\delta n$. As X_2 has a binomial distribution, a standard Chernoff estimate implies that

$$\mathbb{P}(X_2 \le 2\delta n) \le \mathbb{P}(X_2 \le \mathbb{E}(X_2)/2) \le \exp(-\mathbb{E}(X_2)/8) \le \exp(-\nu n/8).$$

As $G_2 \subseteq G_n$, the probability that $B \cup C$ is an $(\varepsilon/2, 2\delta)$-cut of U_1 in G_n is at most $\exp(-\nu n/8)$. Hence, conditional on G_1 and assuming that A holds,

$$\mathbb{P}(A_4 \mid G_1) \le (1+\eta)^n \exp(-\nu n/8) \le (1+\eta)^n (1+2\eta)^{-n} = o(1).$$

As the estimate above holds uniformly for all G_1 such that A holds, it follows that $\mathbb{P}(A_4 \mid A) = o(1)$. As A holds whp, this shows that $\mathbb{P}(A_4) \to 0$, as required. □

As noted earlier, it is easy to deduce Theorem 3.9 from Theorem 11.1. Recall that the only differences between these results are that in Theorem 3.9 we allow $\mathcal{V}$ to be a generalized vertex space, and we allow the deletion of vertices as well as the addition and deletion of edges.

Proof of Theorem 3.9. We first show that, as usual, we lose no generality by assuming that $\mathcal{V}$ is a vertex space. Although this is not obvious at first sight, the general arguments in Subsection 8.1 apply. Indeed, the only potential problem arises when we condition on the sequences $(\mathbf{x}_n)$, since δ might depend on $(\mathbf{x}_n)$. However, fixing ε and defining X_n as the smallest number of changes (edge/vertex deletions or edge additions) that can be made to G_n to obtain a graph G_n' for which (3.7) fails, then Theorem 3.9 states exactly that, for any $\varepsilon > 0$, we have $n/X_n = O(1)$ whp. As noted in Subsection 8.1, in proving that any function of G_n is $O(1)$ whp, we may assume that the sequences $(\mathbf{x}_n)$ are deterministic, by conditioning and applying Lemma A.5.

From now on we assume that $\mathcal{V}$ is a vertex space. Turning to vertex deletion, given an $\varepsilon > 0$, let $\delta > 0$ be such that the conclusion of Theorem 11.1 holds. By Proposition 8.11, there is a $\delta' > 0$ such that the event $\mathcal{E}$ that any $\delta' n$ vertices of G_n are incident with at most $\delta n/2$ edges holds whp. Set $\delta'' = \min\{\delta', \delta/2\}$.

Let G_n' be any graph obtained from G_n by deleting at most $\delta'' n \leq \delta' n$ vertices, and then adding and deleting at most $\delta'' n \leq \delta n/2$ edges. If $\mathcal{E}$ holds, then replacing the deleted vertices as isolated vertices to obtain a graph G_n'' on $V(G_n)$, we have

$$|E(G_n'') \triangle E(G_n)| \leq \delta n/2 + \delta n/2 = \delta n.$$

Hence, by Theorem 11.1, whp every such G_n' satisfies (11.1), which is exactly (3.7). This completes the proof of Theorem 3.9. $\square$

The above proof of Theorem 11.1 is much simpler than any proof we have been able to find based directly on branching process methods. However, the branching process approach does give additional insight into the relationship between the giant component and two-core of G_n and the branching process $\mathfrak{X}_\kappa$.

11.2. **Branching process analysis of the two-core.** Throughout this subsection we work with a kernel κ on a vertex space $\mathcal{V}$ satisfying the assumptions (11.3). As usual, we assume without loss of generality that the number n_i of vertices of each type i is deterministic, with $n_i/n \to \mu(\{i\})$ as $n \to \infty$. The cornerstone of the branching process approach is the following form of the coupling between the neighbourhood exploration process in G_n and the branching process $\mathfrak{X}_\kappa$.

Lemma 11.4. *There is a function $L_0 = L_0(n) \to \infty$ such that we may couple the neighbourhood exploration process of a random vertex v of $G_n = G^{\mathcal{V}}(n, \kappa)$ with the branching process $\mathfrak{X}_\kappa$ so that whp they agree for the first L_0 generations.*

The sense of agreement is that there is a bijection between the vertices of G_n at distance at most L_0 from v and the first L_0 generations of $\mathfrak{X}_\kappa$ mapping v to the initial particle and preserving type and adjacency, where particles in the branching process are adjacent if one is a child of the other.

Proof. The argument is the same as the proof of (9.5), except for the error bounds. Note that it suffices to consider the case L_0 fixed. With L_0 fixed, the total number of vertices encountered has bounded expectation, so we may abandon the coupling if we reach more than $\log n$ vertices, say, in the neighbourhood exploration. At every step, the number of unused vertices of type j is $\mu\{j\}n + o(n)$. Using this estimate in place of (9.4), we may couple the number of new neighbours of each type found with a corresponding

$\mathrm{Po}(\kappa(i,j)\mu\{j\})$ random variable so as to agree with probability $1 - o(1)$. As the expected total number of steps is $O(1)$, the total error probability is $o(1)$. □

If we have $\mu\{j\}n + O(1)$ vertices of type j, Lemma 11.4 holds for any $L_0 = o(\log n)$.

As in the proof of Lemma 9.6, the coupling easily extends to the L_0-neighbourhoods of two vertices. Given G_n, let v and w be chosen independently and uniformly at random from the vertices of G_n.

Lemma 11.5. *There is an $L_0(n) \to \infty$ such that we may couple (G_n, v, w) with two independent copies $\mathfrak{X}_\kappa$, $\mathfrak{X}'_\kappa$ of the branching process $\mathfrak{X}_\kappa$ so that whp the first L_0 neighbourhoods of v and of w agree with the first L_0 generations of $\mathfrak{X}_\kappa$ and of $\mathfrak{X}'_\kappa$, respectively.*

We omit the proof, noting only that for L_0 fixed, the probability that v and w are within graph distance $2L_0$ is $o(1)$.

The next step is to find a way of applying the coupling results above to expectations of functions of the neighbourhoods. This will require some care, due to the possible large contribution to an expectation from the low probability event that the coupling fails.

We consider functions $f(v, G)$ defined on a pair (v, G), where G is a graph in which each vertex has a type from $\mathcal{S} = \{1, 2, \ldots, r\}$, and v is a distinguished vertex of G, the *root*. We call such a function an *L-neighbourhood function* if it is invariant under type preserving rooted-graph isomorphisms and depends only on the subgraph of G induced by vertices within a fixed distance L of v. We define $f(\mathfrak{X}_\kappa)$ by evaluating f on the branching process in the natural way: form a graph from the branching process as above, and take the initial particle as the root. Thus Lemma 11.4 implies that we can couple (G_n, v) with $\mathfrak{X}_\kappa$ so that $f(v, G_n) = f(\mathfrak{X}_\kappa)$ whp for every L-neighbourhood function f.

Given an L-neighbourhood function f, let

$$S_n := \frac{1}{n} \sum_{v \in V(G_n)} f(v, G_n).$$

Also, for v and w independent random vertices of G_n, let $X_n = f(v, G_n)$ and $Y_n = f(w, G_n)$. Note that $\mathbb{E}(S_n) = \mathbb{E}(X_n) = \mathbb{E}(Y_n)$.

Theorem 11.6. *Let $\mathcal{V}$ be a vertex space with finite type space $\mathcal{S}$, and let κ be a kernel on $\mathcal{V}$. If f is an L-neighbourhood function such that $\sup_n \mathbb{E}(X_n^4) < \infty$, then $S_n \xrightarrow{\mathrm{p}} \mathbb{E}(f(\mathfrak{X}_\kappa))$.*

Proof. Let $X = f(\mathfrak{X}_\kappa)$. By Lemma 11.4 we may couple X_n and $\mathfrak{X}_\kappa$ so that $\mathbb{P}(X_n \neq X) \to 0$, and hence $X_n \xrightarrow{\mathrm{p}} X$. Since $\sup_n \mathbb{E}(X_n^4) < \infty$ implies that the variables X_n are uniformly integrable, it follows that

$$\mathbb{E}(S_n) = \mathbb{E}(X_n) \to \mathbb{E}(X); \tag{11.4}$$

see [61, Lemma 4.11], for example.

Let Y be an independent copy of X. From Lemma 11.5 we may couple (X_n, Y_n) with (X, Y) so that $\mathbb{P}((X_n, Y_n) \neq (X, Y)) \to 0$. In particular, $X_n Y_n \xrightarrow{\mathrm{p}} XY$. As

$$\mathbb{E}((X_nY_n)^2) = \mathbb{E}(X_n^2Y_n^2) \leq \sqrt{\mathbb{E}(X_n^4)\,\mathbb{E}(Y_n^4)} = \mathbb{E}(X_n^4) \leq C,$$

for some $C < \infty$, the variables X_nY_n are also uniformly integrable, so $\mathbb{E}(X_nY_n) \to \mathbb{E}(XY)$. But $\mathbb{E}(S_n^2) = \mathbb{E}(X_nY_n)$ by linearity of expectation, while X and Y are independent and have the same distribution. Thus $\mathbb{E}(S_n^2) \to \mathbb{E}(XY) = \mathbb{E}(X)^2$. Together with (11.4), this proves the result. □

Remark 11.7. Theorem 11.6 can be applied to any L-neighbourhood function f bounded by a polynomial of the number of vertices within distance L of v. Indeed, the number of vertices at distance t from v in G_n is stochastically dominated by the number N_t of particles in generation t of a Galton-Watson branching process in which the number of children of each particle has a $\mathrm{Bi}(n, \max \kappa/n)$ distribution. As a $\mathrm{Bi}(1, p)$ distribution is stochastically dominated by a $\mathrm{Po}(1, -\log(1-p))$ distribution, if n is large enough then N_t is dominated by N_t', the number of particles in generation t of the single-type Poisson branching process $\mathfrak{X}_{2\max\kappa}$. The probability generating function of N_t' is obtained by iterating that of the Poisson distribution t times. As all moments of a Poisson distribution are finite, it follows that all moments of N_t' are finite, so the fourth moment of any power of N_t' is finite.

Proposition 8.11 states that, given $\varepsilon > 0$, there is a $\delta > 0$ such that whp any set of at most δn vertices of G_n is incident with at most εn edges. A very special case of Theorem 11.6 gives an alternative proof of this result (under the more restrictive assumptions of the present section). Indeed, writing X for the number of particles in the first generation of $\mathfrak{X}_\kappa$, since $\mathbb{E}\, X$ is finite, we have $\mathbb{E}(X\mathbf{1}[X > M]) \to 0$ as $M \to \infty$. Given $\varepsilon > 0$ there is thus an M for which $\mathbb{E}(X\mathbf{1}[X > M]) < \varepsilon/3$. Writing $d_G(v)$ for the degree of a vertex v in the graph G, let $f(v, G) = d_G(v)\mathbf{1}[d_G(v) > M]$; clearly, f is a 1-neighbourhood function. By Remark 11.7, Theorem 11.6 applies to f, so

$$S_n = \frac{1}{n} \sum_{v \in V(G_n)\,:\, d(v) > M} d(v) \xrightarrow{\mathrm{p}} \mathbb{E}(X\mathbf{1}[X > M]) < \varepsilon/3.$$

Hence $S_n \leq \varepsilon/2$ whp. Set $\delta = \varepsilon/(2M)$, and let W be a set of at most δn vertices of G_n. Then

$$\begin{aligned} \sum_{w \in W} d(w) \quad &\leq \sum_{w \in W\,:\, d(w) > M} d(w) + M|W| \\ &\leq nS_n + \varepsilon n/2 \leq \varepsilon n \end{aligned}$$

whenever $S_n \leq \varepsilon/2$ holds, so whp any set of at most βn vertices of G_n are incident with at most εn edges.

Our next result is a simple observation concerning short cycles. As before, we assume throughout that (11.3) holds.

Lemma 11.8. *Let $L = L(n) = o(\log n)$. The probability that a random vertex v of G_n is within distance L of a cycle of length at most L is $o(1)$.*

Proof. As all edge probabilities are bounded by $p = \max \kappa/n$, the expected number of vertices v at distance $d \geq 0$ from a cycle of length $l \geq 3$ is at most $n^{l+d}p^{l+d} \leq (\max \kappa)^{l+d}$. Summing over $l, d \leq L$, the expectation is $o(n)$. □

The *two-core* $C^2(G)$ of a graph G is the maximal subgraph of G with minimum degree at least 2. Equivalently, $C^2(G)$ consists of those vertices and edges of G that lie in some cycle in G, or on a path joining two vertex-disjoint cycles. We shall work with the two-core $G^2 := C^2(G_n)$ of G_n. To do so, we need to relate certain properties of G^2 to the branching process $\mathfrak{X}_\kappa$. In the light of Lemma 11.4, it will be useful to have a reasonably accurate ($o(1)$ error probability) 'local' characterization of when a vertex v is in the two-core. We shall need similar results for vertices not in the two-core, but connected to it by short paths. Note that Theorem 3.6 gives us a corresponding characterization for the giant component: for a suitable $L(n) \to \infty$, up to an error probability of $o(1)$, a vertex v is in the giant component if and only if it is in a component of size at least L, and using Lemma 11.4, it is easy to check that whp when this condition holds the *L-distance set*, the set of vertices at graph distance exactly L from v, is non-empty, so there is a path of length L starting at v. (We omit the details as we use this statement only to motivate what follows, not in the proof.) For the two-core, we need two vertex-disjoint paths.

Let $L = L(n)$ be a function tending to infinity slowly, to be chosen below. For a vertex v of G_n and an integer $d \geq 0$, let $TC_d(v)$ be the event that v is at graph distance at most d from the two-core G^2 of G_n. Thus $TC_0(v)$ is the event $v \in G^2$. Let $LTC_d(v)$ be the 'localized' event that there is a vertex w at distance $d' \leq d$ from v joined by two vertex-disjoint paths of length L to vertices at distance $d' + L$ from v. Thus, as we explore the neighbourhoods of v successively, $LTC_d(v)$ is the event that after $d' \leq d$ steps we reach a vertex w (which we expect to be the closest vertex of the two-core to v) with two neighbours in the next generation each of which has neighbours for at least $d - 1$ further generations.

Lemma 11.9. *Let $d \geq 0$ be fixed, and let v be a random vertex of G_n. Provided $L(n)$ tends to infinity sufficiently slowly, the event $TC_d(v) \triangle LTC_d(v)$, i.e., the event that one of $TC_d(v)$ and $LTC_d(v)$ holds but not the other, has probability $o(1)$.*

Proof. Assume, as we may, that $L = o(\log n)$. We start with the case $d = 0$. Let us say that a cycle is *short* if it has length at most $2L$. By Lemma 11.8, the probability that v is within distance L of a short cycle is $o(1)$. If $TC_0(v)$ holds, i.e., v is in the two-core, then v is in a cycle, or on a path joining two vertex-disjoint cycles. Assuming that v is not close to a short cycle, in either case we can find two vertex-disjoint paths of length L starting from v, so $LTC_0(v)$ holds. Hence $\mathbb{P}(TC_0(v) \setminus LTC_0(v)) = o(1)$.

The reverse bound is more difficult, as what we need is an equivalent for the two-core of Theorem 3.6, which states that almost all vertices in largish components are in a single giant component. In fact, we can use Theorem 3.6. Suppose that $LTC_0(v) \setminus TC_0(v)$ holds. Note that v is not in a cycle by definition of the two-core. Let w_1, w_2 be two neighbours of v joined by vertex-disjoint paths to vertices x_1, x_2 at distance L from v. In $G_n - v$, there is no path from w_1 to w_2; otherwise, there would be a cycle in G_n containing v. Hence, at least one of w_1 and w_2, let us say w_1, is not in the giant component of $G_n - v$. (Here, by the giant component we mean the largest component, chosen according to any fixed rule if there is a tie.) But w_1 is in a component of size at least L, as witnessed by the path w_1x_1. In summary, if $LTC_0(v) \setminus TC_0(v)$ holds, so does the event $E(v)$ that v is adjacent in G_n to a vertex w in an intermediate component of $G_n - v$, i.e., a component other than the largest having size at least L. As the random vertex v is chosen independently of G_n, the graph $G_n - v$ is an $(n-1)$-vertex graph to which Theorem 3.6 applies. Hence, taking $\omega(n) = L(n)$, by Theorem 3.6 the number of vertices w of $G_n - v$ in intermediate components is $o_p(n)$. Conditioning on $G_n - v$ tells us nothing about the edges from v to $G_n - v$. As κ is bounded, it follows that $E(v)$ has probability $o(1)$. Thus $\mathbb{P}(LTC_0(v) \setminus TC_0(v)) = o(1)$, completing the proof in the case $d = 0$.

The general case follows using Proposition 8.11. If $TC_d(v) \setminus LTC_d(v)$ holds, then v is within distance $d + L$ of a vertex on a short cycle. Hence, by Lemma 11.8, $\mathbb{P}(TC_d(v) \setminus LTC_d(v)) = o(1)$. If $LTC_d(v) \setminus TC_d(v)$ holds, then v is within distance d of a vertex v' for which $LTC_0(v') \setminus TC_0(v')$ holds. By the case $d = 0$ above, $o_p(n)$ vertices v' have this property, and the result follows by applying Proposition 8.11 d times. □

We now turn to the branching process equivalents of the events TC_d and LTC_d. Considering the branching process $\mathfrak{X}_\kappa$ (started with a single particle of random type), let DS_d be the event that there is a particle in some generation $d' \leq d$ which has at least two children with descendants in all future generations. Similarly, let $LDS_{d,L}$ be the event that there is a particle x in generation up to d, say in generation d', such that x has two children each of which has one or more descendants in generation $d' + L$, i.e., L generations after x. Note that $LDS_{d,L}$ depends only on the first $d+L$ generations of the branching process. Suppose that $L(n)$ grows slowly enough that Lemma 11.4 applies with $2L$ in place of L. Then for any fixed d we have $d + L \leq 2L$ for large enough n, and, with v a random vertex of G_n as before, from Lemma 11.4 we have

$$\mathbb{P}(LTC_d(v)) = \mathbb{P}(LDS_{d,L(n)}) + o(1) \quad \text{as } n \to \infty. \tag{11.5}$$

Note that for d and L fixed, the event $LDS_{d,L}$, which is defined in terms of the branching process, does not depend on n, so $\mathbb{P}(LDS_{d,L})$ is a constant. For each d, as L increases the events $LDS_{d,L}$ decrease to the event DS_d.

Hence,

$$\lim_{L\to\infty} \mathbb{P}(LDS_{d,L}) = \mathbb{P}(DS_d). \tag{11.6}$$

Suppose now that $L(n)$ tends to infinity sufficiently slowly that (11.5) and Lemma 11.9 hold. Then, from (11.6), $\mathbb{P}(LDS_{d,L(n)}) \to \mathbb{P}(DS_d)$ as $n \to \infty$. Hence, from (11.5), $\mathbb{P}(LTC_d(v)) = \mathbb{P}(DS_d)+o(1)$. Finally, using Lemma 11.9 we obtain

$$\mathbb{P}(TC_d(v)) = \mathbb{P}(DS_d) + o(1).$$

Considering two random vertices v, w of G_n and using Lemma 11.5 instead of Lemma 11.4, we obtain $\mathbb{P}(TC_d(v) \cap TC_d(w)) = \mathbb{P}(DS_d)^2 + o(1)$ similarly.

Writing TC_d for the set of vertices of G_n for which $TC_d(v)$ holds, i.e., for the set of vertices within distance d of the two-core, it follows that $\mathbb{E}\,|TC_d|/n \to \mathbb{P}(DS_d)$ and $\mathbb{E}(|TC_d|/n)^2 \to (\mathbb{P}(DS_d))^2$, and thus

$$|TC_d|/n \xrightarrow{\mathrm{p}} \mathbb{P}(DS_d). \tag{11.7}$$

As $\|T_\kappa\| > 1$, the branching process $\mathfrak{X}_\kappa$ is supercritical, so $\mathbb{P}(DS_0) > 0$. Hence, taking $d = 0$ in (11.7),

$$|V(G^2)|/n \xrightarrow{\mathrm{p}} \mathbb{P}(DS_0) > 0. \tag{11.8}$$

The reason for considering the two-core G^2 of G_n is that Theorem 3.9 boils down to a statement about G^2. Roughly speaking, the largest component of G_n consists of the two-core with some trees hanging off it, and it is easy to see what effect deleting edges from the trees has on the size of the largest component. The question is what happens when edges are deleted from the two-core.

Lemma 11.10. *Suppose that* (11.3) *holds, i.e.,* $\mathcal{S} = \{1, 2, \dots, r\}$, $\mu\{i\} > 0$ *for each* i, κ *is irreducible, and* $\|T_\kappa\| > 1$. *Let* G^2 *be the two-core of* $G^{\mathcal{V}}(n,\kappa)$. *For any* $\varepsilon > 0$ *there is a* $\delta > 0$ *such that the following statement holds whp: for any set* $W \subset V(G^2)$ *with* $|W| \ge \varepsilon n$ *and* $|V(G^2)\setminus W| \ge \varepsilon n$ *there are more than* δn *edges of* G^2 *joining* W *to* $V(G^2) \setminus W$.

In other words, under the assumptions of Lemma 9.6, if $\|T_\kappa\| > 1$ then the two-core G^2 cannot be cut into two large (size $\Theta(n)$) pieces by a small set of edges. Note that the two-core itself is large by (11.8). As the proof of Lemma 11.10 is rather long, we first show that it implies Theorem 3.9.

Deduction of Theorem 3.9 from Lemma 11.10. We have already shown (at the end of Subsection 11.1) that, using Proposition 8.11, Theorem 3.9 can be deduced from Theorem 11.1. As noted at the start of the section, in proving Theorem 11.1 we may assume that (11.3) holds, and it suffices to prove (11.2). From now on, let us fix the quantity $\varepsilon > 0$ appearing in (11.2).

The events DS_d form an increasing sequence, and their union is contained in the event S that the branching process $\mathfrak{X}_\kappa$ survives (contains points in all generations). Also, $S \setminus \bigcup_d DS_d$ is the event that the process survives, but with only a single infinite line of descent. From basic properties of Poisson processes, starting from a particle of type x, the types its surviving children, i.e., its children that have descendants in all later generations, form a Poisson process on $\mathcal{S}$ with intensity $\kappa(x,y)\rho(\kappa;y)\,d\mu(y)$. In particular, the number of such children is Poisson with some mean $\lambda(x) > 0$. It follows that, conditional on a particle surviving, the probability that it has at least two surviving children is positive, and hence, as the type space $\mathcal{S}$ is finite, bounded way from zero. Hence $\mathbb{P}(S \setminus \bigcup_d DS_d) = 0$, so $\mathbb{P}(DS_d) \nearrow \mathbb{P}(S) = \rho(\kappa)$, and there is a constant D such that $\mathbb{P}(DS_D) \ge \rho(\kappa) - \varepsilon/3$. From (11.7), for any fixed D the set TC_D of vertices within distance D of the two-core has size $\mathbb{P}(DS_D)n + o_p(n)$, so whp

$$|TC_D| \ge (\rho(\kappa) - \varepsilon/2)n. \tag{11.9}$$

Let $\varepsilon' < \mathbb{P}(DS_0)/3$ be a small positive constant to be chosen later, and let $\delta = \min\{\varepsilon', \delta(\varepsilon')\}$, where $\delta(\cdot)$ is the function appearing in Lemma 11.10. Let us delete an arbitrary set E of at most δn edges from G_n, leaving a graph G_n'. Let $G^{2-} \subseteq G^2$ be the largest remaining connected part of G^2. We claim that

$$|V(G^2) \setminus V(G^{2-})| \le \varepsilon' n$$

holds whp. Note that $|V(G^2)| \ge 3\varepsilon' n$ whp by (11.8). If this inequality holds and every component of $G^2 \setminus E$ has size at most $|V(G^2)| - \varepsilon' n$, then there is a union H of components of $G^2 \setminus E$ with between $\varepsilon' n$ and $|V(G^2)| - \varepsilon' n$ vertices: indeed, if the largest component has at least $\varepsilon' n$ vertices, this will do as H. Otherwise, every component has at most $\varepsilon' n$ vertices, and the smallest union H with at least $\varepsilon' n$ vertices will do. The existence of an H with the stated properties has probability $o(1)$ by Lemma 11.10, proving the claim.

Let X be the component of G_n' containing G^{2-}. If $v \in TC_D \setminus X$ then, as $v \in TC_D$, there is a path in G_n of length at most D from v to a vertex of G^2. Taking any such path, as $v \notin X$, either the path ends in a vertex of $G^2 \setminus G^{2-}$, of which there are whp at most $\varepsilon' n$, or it contains an edge of E, and hence contains an endvertex of such an edge; there are at most $2\delta n \le 2\varepsilon' n$ such endvertices. In particular, whp all $v \in TC_D \setminus X$ are within distance D (in G_n) of some set of at most $3\varepsilon' n$ vertices. Applying Proposition 8.11 D times, it follows that if we choose ε' small enough, then whp $|TC_D \setminus X| \le \varepsilon n/2$. Using (11.9) it follows that whp

$$C_1(G_n') \ge |X| \ge (\rho(\kappa) - \varepsilon)n,$$

completing the proof of Theorem 3.9. □

It remains only to prove Lemma 11.10. The uniform case (κ constant) has a simple proof, presented below.

Proof of Lemma 11.10, uniform case. In a moment we shall restrict to the uniform case; for now, we assume (11.3).

As $\|T_\kappa\| > 1$, the branching process $\mathfrak{X}_\kappa$ is supercritical, so $\rho(\kappa) > 0$. From irreducibility, it follows that $\rho(\kappa; i) > 0$ for every i. Consider the event TS that the initial particle has exactly three children that survive. As the initial particle has positive probability of having exactly three children, $\mathbb{P}(TS) > 0$. Arguing as for (11.8), one can show that the number of vertices of degree exactly 3 in G^2 is $\mathbb{P}(TS)n + o_p(n)$; in fact, both statements are special cases of Lemma 11.11 below.

We shall condition on the vertex set and (labelled) degree sequence of G^2. In other words, we shall condition on the sequence $\mathbf{d} = (d(1), \dots, d(n))$, where $d(i)$ is the degree in G^2 of the vertex i and $d(i) = 0$ if $i \notin V(G^2)$. Let us write n_2 for $|V(G^2)| = |\{i : d(i) > 0\}|$, m_2 for $e(G^2) = \frac{1}{2}\sum d(i)$, and $n_{\geq 3}$ for the number of i for which $d(i) \geq 3$. Note that whp $n_{\geq 3} \geq \mathbb{P}(TS)n/2 > 0$. Also, $e(G_n) = O(n)$ whp (for example, by Proposition 8.9), and G_n has maximum degree $O(\log n)$ whp. Thus there are positive constants ε_1 and C, depending only on κ, such that

$$n_{\geq 3} \geq \varepsilon_1 n, \quad m_2 \leq C n_2, \quad \text{and } \max_i d(i) \leq C \log n, \tag{11.10}$$

hold whp.

From now on we consider the uniform case, where κ is constant, or, equivalently, $|\mathcal{S}| = 1$. This is just the usual Erdős–Rényi random graph $G_n = G(n, c/n)$, with $c > 1$. We condition on $\mathbf{d}$, assuming, as we may, that the conditions (11.10) hold. In the uniform case it is easy to see that (given $\mathbf{d}$) the graph G^2 is uniformly distributed among all graphs with degree sequence $\mathbf{d}$. This is because any graph can be decomposed into its two-core and a collection of vertex-disjoint trees, each sharing at most one vertex with the two-core. Hence, any two graphs H_1, H_2 with degree sequence $\mathbf{d}$ can be extended in exactly the same ways to graphs H_1', H_2' with vertex set $[n]$ so that H_i' has two-core H_i. As corresponding graphs H_1', H_2' have the same number of edges, they are equally likely in the model $G(n, c/n)$. Summing over the possible extensions, H_1 and H_2 are equally likely to arise as G^2.

Let H be the random multigraph with degree sequence $\mathbf{d}$ generated by the configuration model of [14]. In other words, for each vertex i we take $d(i)$ 'stubs', and we pair the $2m_2$ stubs randomly, with all $(2m_2 - 1)!!$ pairings equally likely. For every pair in this pairing, we take an edge between the corresponding vertices. This generates a multigraph H with degree sequence $\mathbf{d}$, where H may contain loops and multiple edges. Let p_{simple} be the probability that H is simple. From our assumptions on $\mathbf{d}$ it is easy to check that $p_{\mathrm{simple}} = \exp(-o(n))$. This very crude lower bound is all that we shall need; the much stronger bound $\exp(-O((\log n)^2))$ follows from (11.10) and the general results in McKay [75]. In fact, using the fact that $\sum_i d(i)^2 = O(n)$ whp, one can show that $p_{\mathrm{simple}} = \Theta(1)$. Given that H is simple, it is uniformly distributed among all simple graphs with degree sequence $\mathbf{d}$, i.e., H has the distribution of G^2. Hence, to show that G^2 has a certain property whp it suffices to show that H has the property with probability $1 - o(p_{\mathrm{simple}})$.

For $W \subseteq V(G^2)$ let $\overline{W} = V(G^2) \setminus W$. Let $S(W)$ denote the set of stubs associated to W, so $|S(W)| = \sum_{i \in W} d(i)$. For a set S of stubs, let

$\overline{S} = S(V(G^2)) \setminus S$, and let $p_{cut}(S, \overline{S})$ denote the probability that, in the random pairing, every stub in S is paired with another stub in S. If $V(G^2)$ has a partition W, $\overline{W}$ with at most δn edges between W and $\overline{W}$, then there is an $S \subseteq S(W)$ with $|S(W) \setminus S| \le \delta n$ so that every stub in S is paired with another stub in S. Hence, the expected number of such partitions with W, $\overline{W}$ large is at most

$$EC := \sum_{W \subset V(G^2):|W|,|\overline{W}| \ge \varepsilon n} \; \sum_{S \subseteq S(W):|S(W) \setminus S| \le \delta n} p_{cut}(S, \overline{S}), \tag{11.11}$$

and it suffices to show that EC is $o(p_{\mathrm{simple}})$ if we choose δ small enough. Now $p_{cut}(S, \overline{S}) = 0$ if $|S|$ is odd, and otherwise

$$p_{cut}(S, \overline{S}) = \frac{(|S|-1)!!\,(|\overline{S}|-1)!!}{(2m_2-1)!!} \le \binom{m_2}{|S|/2}^{-1}. \tag{11.12}$$

Every vertex of G^2 has degree at least 2, so for any W we have $|S(W)|/2 \ge |W|$. As, from (11.10), there are at least $\varepsilon_1 n$ vertices of degree 3 in G^2, either $|S(W)|/2 \ge |W| + \varepsilon_1 n/4$ or $|S(\overline{W})|/2 \ge |\overline{W}| + \varepsilon_1 n/4$. It follows that whenever $|W|$, $|\overline{W}| \ge \varepsilon n$ we have

$$\binom{m_2}{|S(W)|/2} \ge \exp(4an)\binom{n_2}{|W|},$$

where $a > 0$ is a constant depending only on ε_1 and ε, and hence only on κ and ε. Choosing δ small enough, it follows that

$$\binom{m_2}{|S|/2} \ge \exp(3an)\binom{n_2}{|W|}, \tag{11.13}$$

whenever $S \subseteq S(W)$ with $|S| \ge |S(W)| - \delta n$. Given $w = |W|$, there are at most $\binom{n_2}{w}$ choices for W. Given W, there are, crudely, at most $m_2\binom{2m_2}{\delta n}$ choices for $S \subseteq S(W)$ with $|S(W) \setminus S| \le \delta n$. Hence, from (11.11), (11.12) and (11.13),

$$EC \le \sum_{\varepsilon n \le w \le n_2 - \varepsilon n} \binom{n_2}{w} m_2 \binom{2m_2}{\delta n} \exp(-3an) \binom{n_2}{w}^{-1}.$$

Choosing δ small enough, it follows that $EC = O(n^2 \exp(-2an))$, and hence that $EC \le \exp(-an)$ for n large enough. As $p_{\mathrm{simple}} = \exp(-o(n))$, this completes the proof in the uniform ($|\mathcal{S}| = 1$) case. □

One might hope that any argument for the uniform case would adapt easily to the finite-type case. However, we have been unable to find a simple extension of the argument above. Our branching-process based argument for the general case is somewhat involved and rather lengthy, and we shall not present it. This is because adapting a proof due to Luczak and McDiarmid [70] gives the much simpler proof of Theorem 11.1 in Subsection 11.1, which in turn immediately implies the general case of Lemma 11.10. We believe, however, that the results in this subsection are likely to be useful for determining other properties of the two-core.

We close this section with a final result, Lemma 11.11 below, relating any 'local' property of the two-core to the branching process $\mathfrak{X}_\kappa$. This will require a little introduction.

Recall that, by Lemma 11.4, if $L \to \infty$ sufficiently slowly then we may couple the $2L$-neighbourhood of a random vertex v of G_n with the first $2L$ generations of the branching process $\mathfrak{X}_\kappa$ so that they agree with probability $1 - o(1)$. By Lemma 11.9, for almost every vertex v, v is in the two-core if and only if there are two disjoint paths of length L starting at v. This allows us to adapt the coupling, and hence Theorem 11.6, to the two-core.

For $v \in G^2$ and $t \geq 0$, let $\Gamma_t(v, G^2)$ be the set of vertices of G^2 at graph distance t from v, and set $\Gamma_t(v, G^2) = \emptyset$ if $v \notin G^2$. Note that any vertex on a path joining two vertices of G^2 is in G^2, so $\Gamma_t(v, G^2)$ is the t-distance set of v in the graph G^2. Let $\overline{\mathfrak{X}}_\kappa$ be obtained from $\mathfrak{X}_\kappa$ in two steps: first, delete any particle that does not have descendants in all future generations. Then, if the initial particle has only one remaining child, delete everything; we write $\overline{\mathfrak{X}}_\kappa = \emptyset$ in this case. We obtain a certain branching process $\overline{\mathfrak{X}}_\kappa$ having the following properties whenever $\overline{\mathfrak{X}}_\kappa \neq \emptyset$: the first particle has at least two children, and every later particle at least one child.

For constant D, if $L = L(n) \to \infty$ then up to an $o(1)$ error probability, the first $D + L$ generations of $\mathfrak{X}_\kappa$ determine the first D generations of $\overline{\mathfrak{X}}_\kappa$: consider surviving to generation $D+L$ instead of surviving forever. Let v be a random vertex of G_n. If $L \to \infty$ sufficiently slowly then, by Lemma 11.9 and Proposition 8.11 (applied D times), the probability that v is within distance D of a vertex w for which one of $TC_0(w)$ and $LTC_0(w)$ holds but not the other is $o(1)$. Using Lemma 11.4, it follows that the G^2-neighbourhoods $(\Gamma_t(v, G^2))_{t=0}^D$ of a random vertex $v \in G_n$ can be coupled with the first D generations of $\overline{\mathfrak{X}}_\kappa$ so as to agree with probability $1 - o(1)$. Similarly, Lemma 11.5 implies its equivalent for G^2 and $\overline{\mathfrak{X}}_\kappa$. Using these two results, an analogue of Theorem 11.6 follows. In the result below we take $f(\overline{\mathfrak{X}}_\kappa)$ to be zero when $\overline{\mathfrak{X}}_\kappa$ is empty. The proof follows exactly that of Theorem 11.6, so we omit it.

Lemma 11.11. *Let D be fixed, and let $f = f(v, G)$ be a D-neighbourhood function bounded by a polynomial of the number of vertices within distance D of v. Then*

$$S_n := \frac{1}{n} \sum_{v \in G^2} f(v, G^2) \xrightarrow{\mathrm{p}} \mathbb{E}(f(\overline{\mathfrak{X}}_\kappa)).$$

Of course, the condition on f could be replaced by a fourth-moment condition as in Theorem 11.6. As an immediate consequence of Lemma 11.11, we can describe, for example, the typed degree sequence of G^2, taking $f_{i,d}(v, G)$ to be the 1-neighbourhood function taking the value 1 when v has degree d and type i, and 0 otherwise. Lemma 11.11 can be used as the basis of a

proof of Lemma 11.10, but as noted above, the details are rather involved; see the first version of this paper, at
`http://arXiv.org/math.PR/0504589v1` .

Remark 11.12. Recently, Riordan [85] proved an analogue of Lemma 11.11 for the k-core, again using a direct coupling of the neighbourhood exploration process with a branching process.

12. Bounds on the small components

In this section we prove Theorem 3.12, i.e., that the sizes of the small components of $G_n = G^{\mathcal{V}}(n,\kappa)$ are $O(\log n)$ whp under certain assumptions. As before, by the *giant component* in a graph G_n we mean the unique largest component, provided it has $\Theta(n)$ vertices – all other components are *small*. Thus, in the supercritical case ($\|T_\kappa\| > 1$), where there is a giant component, a small component is any component other than the largest, so our aim is to prove an upper bound on $C_2(G_n)$. In the strictly subcritical case ($\|T_\kappa\| < 1$), all components are small, and our aim is to show that $C_1(G_n) = O(\log n)$ whp.

We shall prove three results that together imply Theorem 3.12, namely Theorems 12.5, 12.6 and 12.7 below. In the supercritical case, we shall also prove a more general result, Theorem 12.1 below, describing the distribution of the graph formed from G_n by deleting the giant (more precisely, largest) component, $\mathcal{C}_1(G_n)$. This description is in terms of another instance of our general model, involving the dual kernel $\hat{\kappa}$ defined in Definition 3.15.

Theorem 12.1. *Let (κ_n) be a graphical sequence of kernels on a (generalized) vertex space $\mathcal{V}$ with quasi-irreducible limit κ, with $\|T_\kappa\| > 1$. Let $G_n = G^{\mathcal{V}}(n,\kappa_n)$, and let G'_n be the graph obtained from G_n by deleting all vertices in the largest component $\mathcal{C}_1(G_n)$. There is a generalized vertex space $\widehat{\mathcal{V}} = (\mathcal{S}, \widehat{\mu}, (\mathbf{y}_n))$ with $\widehat{\mu}$ given by $d\widehat{\mu}(x) = (1-\rho(\kappa;x))\,d\mu(x)$, such that G'_n and $G^{\widehat{\mathcal{V}}}(n,\kappa_n)$ can be coupled to agree whp. Furthermore, the sequence (κ_n) is graphical on $\widehat{\mathcal{V}}$ with quasi-irreducible limit κ.*

If we wish, we can renormalize so that $(\mathcal{S},\widehat{\mu})$ becomes a ground space; see the comment after Definition 3.15. However, the resulting graph still has a random number of vertices, so we cannot insist that $\widehat{\mathcal{V}}$ is a vertex space.

Theorem 12.1 is the natural generalization to our context of the old 'duality result' of Bollobás [15] for the Erdős–Rényi model $G(n,c/n)$ that was the basis of the study of the phase transition there (see also Łuczak [71], Janson, Knuth, Łuczak and Pittel [58], and the books [16, 59]).

Remark 12.2. We know that the random graph $G^{\widehat{\mathcal{V}}}(n,\kappa_n)$ in Theorem 12.1 cannot be supercritical, since otherwise G_n would have a second giant component. It may be critical, see Example 12.4, but is typically subcritical; one sufficient condition for subcriticality is given in the next result.

Theorem 12.3. *Under the assumptions of Theorem 12.1, if, in addition,* $\iint_{\mathcal{S}^2} \kappa(x,y)^2 \, d\mu(x) \, d\mu(y) < \infty$*, then* $G^{\widehat{\mathcal{V}}}(n, \kappa_n)$ *is subcritical.*

Proof. As usual, we may normalize so that $\mu(\mathcal{S}) = 1$. The result then follows immediately from Theorem 6.7. $\square$

Example 12.4. As in Example 4.12, let $\mathcal{S} = \{1, 2, 3, \dots\}$ with $\mu\{k\} = 2^{-k}$, and let $x_1, \dots, x_n$ be i.i.d. random points in $\mathcal{S}$ with distribution μ. Let $(\varepsilon_k)_1^\infty$ be a sequence of positive numbers tending to zero, to be chosen below. Set $\kappa(1,1) = 4$, $\kappa(k,k) = 2^k$ for $k \geq 2$ (instead of 2^{k+1} in Example 4.12), $\kappa(1,k) = \kappa(k,1) = \varepsilon_k$ for $k \geq 2$, and $\kappa(i,j) = 0$ otherwise.

Furthermore, again as in Example 4.12, let H_k be the subgraph of $G(n,\kappa)$ induced by the $n_k \sim \mathrm{Bi}(n, 2^{-k})$ vertices of type k. Then, conditional on n_k, each H_k has the distribution of the Erdős–Rényi graph $G(n_k, \kappa(k,k)/n)$. As before, H_1 is supercritical, which implies that $G(n,\kappa)$ is supercritical.

Let $k_n \to \infty$ slowly, and choose ε_k such that $\varepsilon_{k_n} \leq n^{-2}$; to be specific, set $k_n := \lceil \log_2 \log n \rceil$ and $\varepsilon_k := \exp(-2^{k+2})$. Then (4.4) implies that H_{k_n} is a critical Erdős–Rényi graph: the edge probability is not exactly one over the number n_{k_n} of vertices, but is $(1 + o(n_{k_n}^{-1/3}))/n_{k_n}$, which is within the "scaling window". It follows, as in Example 4.12, that whp $C_2(G(n,\kappa)) \geq C_1(H_{k_n}) = \Theta_p\big(\bar{n}_{k_n}^{2/3}\big) > n^{2/3}/\log n$.

It is easy to see, analytically from Theorem 6.2 or probabilistically from Theorem 9.10, that $\rho(k) \to 0$ as $k \to \infty$. Consider the graph $G^{\widehat{\mathcal{V}}}(n, \kappa_n)$ in Theorem 12.1; the norm of the corresponding integral operator $\widehat{T}_\kappa$ on $L^2(\widehat{\mu})$ is at least the norm when restricted to $\{k\}$, which is exactly $1 - \rho(k)$, $k \geq 2$. Hence the norm is at least 1. By Remark 12.2, it follows that the norm of $\widehat{T}_\kappa$ is exactly 1, i.e., that $G^{\widehat{\mathcal{V}}}(n, \kappa_n)$ is critical.

We now turn to the proofs of Theorems 3.12 and 12.1, starting with the subcritical case of Theorem 3.12, which we restate below.

Theorem 12.5. *Let* (κ_n) *be a graphical sequence of kernels on a (generalized) vertex space* $\mathcal{V}$ *with limit* κ*. If* κ *is subcritical, i.e.,* $\|T_\kappa\| < 1$*, and* $\sup_{x,y,n} \kappa_n(x,y) < \infty$*, then* $C_1(G_n) = O(\log n)$ *whp.*

Proof. As usual, we may assume that $\mathcal{V}$ is a vertex space; see Subsection 8.1. Consider first the case when $\kappa_n = \kappa$ for all n and $\mathcal{S}$ is finite, or, equivalently, the regular finitary case. In this case, the result follows by comparing the neighbourhood exploration process of a vertex in the graph to a subcritical branching process: this comparison is similar to that made in the proof of Lemma 9.6. This time, setting $\omega(n) = A \log n$, where A is a (large) constant to be chosen below, instead of the upper bound in (9.5) we claim that

$$\mathbb{P}(x \in B) \leq \rho_{\geq \omega(n)}\big((1 + 2\varepsilon)\kappa; i\big), \tag{12.1}$$

for all sufficiently large n.

This can be proved by the comparison argument used for (9.5), except that in the final step, instead of using the total variation distance between the binomial and Poisson distributions, we note that if $\preceq$ denotes stochastic domination and $p' = -\log(1-p)$, so that $p' = p + O(p^2)$ as $p \to 0$, then $\mathrm{Bi}(1,p) \preceq \mathrm{Po}(p')$, and thus $\mathrm{Bi}(m,p) \preceq \mathrm{Po}(mp')$ for every m. Hence, for n large enough, $\mathrm{Bi}(n_j', \kappa(i,j)/n) \preceq \mathrm{Bi}(n_j, \kappa(i,j)/n) \preceq \mathrm{Po}\big((1+2\varepsilon)\kappa(i,j)\mu_j\big)$. Note that in the argument leading to (9.5), and hence in our proof of (12.1), we do not assume that κ is irreducible.

For ε small enough, the branching process $\mathfrak{X}_{(1+2\varepsilon)\kappa}$ is subcritical. Therefore, there is an $a > 0$ such that

$$\rho_{\ge k}((1+2\varepsilon)\kappa; i) \le e^{-ak} \tag{12.2}$$

holds for all i and all $k \ge 1$. Inequality (12.2) is undoubtedly well known for finite-type Galton-Watson processes, but for the sake of completeness we sketch a proof. For $z > 0$ let $g_z(x) := \mathbb{E}\, z^{|\mathfrak{X}_\kappa(x)|}$, where $|\mathfrak{X}_\kappa(x)|$ denotes the total population of the branching process. Using $\mathbb{E}\, w^{\mathrm{Po}(\lambda)} = e^{\lambda(w-1)}$ and independence of the Poisson numbers of particles of each type in the first generation, we have $g_z = z e^{T_\kappa(g_z - 1)}$. If κ is subcritical, then by the implicit function theorem this functional equation has a finite solution for z in a neighbourhood of 1, say for $|z-1| < \delta$, and it follows by an argument similar to the proof of Lemma 5.6 that indeed $g_z(x) < \infty$ for all x when $z < 1 + \delta$.

Taking $A = 2/a$ for some a for which inequality (12.2) holds, and recalling (12.1), $\mathbb{P}(x \in B) \le n^{-2}$ follows, and so whp there are no vertices in B, i.e., there is no component of order $A \log n$ or larger.

For the general case, it suffices to bound κ_n for large n from above by a subcritical finitary κ'. Recalling that $\sup_{x,y,n} \kappa_n(x,y) < \infty$, let κ_m^+ be defined as in Lemma 7.2. As $\|T_\kappa\| < 1$, by Lemma 7.2 we have $\|T_{\kappa_m^+}\| < 1$ for m large enough, and we can take $\kappa' = \kappa_m^+$. ∎

We now turn to the supercritical case, proving two results that together imply part (ii) of Theorem 3.12.

Theorem 12.6. *Let (κ_n) be a graphical sequence of kernels on a (generalized) vertex space $\mathcal{V}$ with irreducible limit κ. If κ is supercritical, i.e., $\|T_\kappa\| > 1$, and $\inf_{x,y,n} \kappa_n(x,y) > 0$, then $C_2(G_n) = O(\log n)$ whp.*

Proof. As usual, we may assume that $\mathcal{V}$ is a vertex space. As the kernel κ is supercritical, we have $\|T_\kappa\| > 1$, so there is an integer k such that $(1 - 1/k)\|T_\kappa\| > 1$. Fix such a k throughout the proof.

Recall that $G_n = G^{\mathcal{V}}(n, \kappa_n)$ is constructed by choosing in an appropriate manner a (deterministic or random) sequence $\mathbf{x}_n = (x_1, \dots, x_n)$ giving the types of the vertices, and then constructing the edges using the kernel κ_n. Independently of G_n, let us partition $V(G_n) = [n]$ into k subsets S_i in a random way, by independently assigning each vertex to a random subset. In other words, we construct G_n (which has vertex set $[n]$) and then partition its vertex set into k classes S_i. Let $S_i^c := [n] \setminus S_i$.

Let A be a very large constant, to be chosen later. We aim to prove that the event E that G_n contains a component with more than $A \log n$ and at most n/A vertices has probability $o(1)$. Every component $\mathcal{C}$ of G_n meets some S_i in at least $|\mathcal{C}|/k$ vertices. Let us say that a component $\mathcal{C}$ of G_n is *bad* if it has at least $A\log n$ and at most n/A vertices, and meets S_1 in at least $A\log n/k$ vertices. Then, as all S_i are equivalent, it suffices to prove that whp G_n has no bad component.

To this end, consider the subgraph G'_n of G_n induced by the vertices in S_1^c. Let $\mu' = (1-1/k)\mu$, and let $\mathbf{y}_n$ be the (random) subsequence of $\mathbf{x}_n$ corresponding to those vertices i with $i \notin S_1$. Note that $\mathcal{V}' = (\mathcal{S}, \mu', (\mathbf{y}_n)_{n\geq 1})$ is a generalized vertex space: condition (2.2) for $\mathcal{V}'$ follows from the same condition for $\mathcal{V}$ and the random choice of S_1. Also, G'_n has exactly the distribution of $G^{\mathcal{V}'}(n, \kappa_n)$.

Since $\mathbb{E}\, e(G'_n) = (1-1/k)^2\, \mathbb{E}\, e(G_n)$, and the sequence (κ_n) is graphical on $\mathcal{V}$ with limit κ, it is graphical on $\mathcal{V}'$ with the same limit κ. As κ is irreducible on $(\mathcal{S}, \mu)$, it is irreducible on $(\mathcal{S}, \mu')$. Let us write κ' for the kernel κ when viewed as a kernel on $(\mathcal{S}, \mu')$. Thus $T_{\kappa'}$ is an operator on $L^2(\mathcal{S}, \mu')$, and, by assumption, $\|T_{\kappa'}\| = (1-1/k)\|T_\kappa\| > 1$. Let $\mathcal{C}_1$ be the largest component of G'_n (chosen according to any rule if there is a tie). Then, by Theorem 3.1, whp $\mathcal{C}_1$ contains at least $\rho(\kappa')n/2$ vertices, and $\rho(\kappa') > 0$.

From now on, we shall assume that $\mathcal{C}_1$ has at least $\rho(\kappa')n/2$ vertices, and choose A so that $1/A < \rho(\kappa')/2$. If a component $\mathcal{C}$ of G_n is bad, then it cannot contain $\mathcal{C}_1$, and hence sends no edges to $\mathcal{C}_1$. Let G''_n be the spanning subgraph of G_n obtained from G_n by deleting all edges between $\mathcal{C}_1$ and vertices in S_1. If $\mathcal{C}$ is a bad component of G_n, then all edges of $\mathcal{C}$ are present in G''_n, so $\mathcal{C}$ is a component of G''_n. Hence, the probability that G_n has a bad component is bounded by $o(1)$ (the probability that $\mathcal{C}_1$ is too small) plus the probability that some component $\mathcal{C} \neq \mathcal{C}_1$ of G''_n containing at least $A\log n/k$ vertices of S_1 sends no edges to $\mathcal{C}_1$ in G_n. Conditioning on $\mathbf{x}_n$, S_1 and G''_n, we have not tested any edges between $\mathcal{C}_1$ and S_1, so these edges are present independently, each with its original probability. These individual probabilities are all at least $(\inf \kappa)/n$, where $\inf\kappa := \inf_{x,y,n} \kappa_n(x,y) > 0$ by assumption.

Thus, for any of the at most n components $\mathcal{C} \neq \mathcal{C}_1$ of G''_n with at least $M = A\log n/k$ vertices in S_1, the probability that $\mathcal{C}$ sends no edges to $\mathcal{C}_1$ is at most

$$(1 - (\inf\kappa)/n)^{M|\mathcal{C}_1|}.$$

As $\inf\kappa > 0$, $|\mathcal{C}_1| \geq \rho(\kappa')n/2$ and $M = A\log n/k$, we can make this probability $o(n^{-100})$ by choosing A large enough. □

Next, we turn to our general result Theorem 12.1 on the distribution of the small components of a supercritical graph $G_n = G^{\mathcal{V}}(n, \kappa_n)$. This will then be used to prove the final statement in Theorem 3.12, restated as Theorem 12.7 below.

Proof of Theorem 12.1. The result is a simple consequence of Theorems 3.1, 9.10, 3.5 and 3.6. Indeed, given $G_n = G^{\mathcal{V}}(n, \kappa_n)$, let $\mathcal{C}_1$ be the largest component of G_n (chosen according to any fixed rule if there is a tie), and, for each n, let $\mathbf{y}_n$ be the subsequence of $\mathbf{x}_n$ consisting of those x_i for which $i \notin \mathcal{C}_1$. Then, by Theorem 9.10, $\widehat{\mathcal{V}} = (\mathcal{S}, \widehat{\mu}, (\mathbf{y}_n))$ is a generalized vertex space, where $\widehat{\mu}$ is defined by $d\widehat{\mu}(x) = (1 - \rho(\kappa; x))\, d\mu(x)$; indeed, the only non-trivial condition to verify is (2.4), which is immediate from the same condition for $\mathcal{V}$ and (9.20).

Next, we must show that the graphs G_n' and $G^{\widehat{\mathcal{V}}}(n, \kappa_n)$ may be coupled so that their edge-sets agree whp. This is easy to see: as usual, we condition throughout on $\mathbf{x}_n$, treating $\mathbf{x}_n$ as deterministic. If we condition also on the vertex set V of $\mathcal{C}_1$, the only information this gives about edges of G_n inside $V^c = V(G_n) \setminus V$ is that $G_n' = G_n[V^c]$ contains no component larger than $|V|$ (and that certain components of order exactly $|V|$ are ruled out). Without this condition, $G_n[V^c]$ would have exactly the distribution of $G^{\widehat{\mathcal{V}}}(n, \kappa_n)$, so it suffices to prove that $C_1\bigl(G^{\widehat{\mathcal{V}}}(n, \kappa_n)\bigr) < |V|$ holds whp. In fact, as $|V| \ge \rho(\kappa)n/2$ whp, it suffices to prove that

$$\eta := \mathbb{P}\left(C_1\bigl(G^{\widehat{\mathcal{V}}}(n, \kappa_n)\bigr) \ge \rho(\kappa)n/2\right)$$

tends to 0 as $n \to \infty$.

Recall that the sequence $\mathbf{x}_n$ is deterministic. Pick a vertex j of G_n at random, and explore its component in G_n in the usual way, by finding the neighbours of j, then the neighbours of the neighbours, and so on. Let V_j be the vertex set of this component. The nature of the exploration process ensures that, given V_j, the edges of $G_n[V_j^c]$ are present independently with their unconditional probabilities. In other words, writing $\mathbf{y}_n'$ for the subsequence of $\mathbf{x}_n$ consisting of those x_i with $i \notin V_j$, and setting $\widehat{\mathcal{V}}' = (\mathcal{S}, \widehat{\mu}, (\mathbf{y}_n'))$, the edge sets of $G_n[V_j^c]$ and $G^{\widehat{\mathcal{V}}'}(n, \kappa_n)$ have the same distribution.

Let us condition on the event E that $|V_j| \ge \rho(\kappa)n/2$. As Theorem 3.1 can be applied to G_n, and j was chosen at random, we see that $\mathbb{P}(E)$ is bounded away from 0. Also, appealing to Theorem 3.6, we see that whp G_n has a unique component of size at least $\rho(\kappa)n/2$. Thus, given E, we have $\mathbf{y}_n = \mathbf{y}_n'$ whp. Hence, with probability $\mathbb{P}(E)(o(1) + \eta)$, the graph G_n contains two components of order at least $\rho(\kappa)n/2$. By Theorem 3.6, this probability is $o(1)$, so we have $\eta = o(1)$, as required. Finally, as κ is quasi-irreducible on $(\mathcal{S}, \mu)$, it is quasi-irreducible on $(\mathcal{S}, \widehat{\mu})$.

It remains only to show that the sequence (κ_n) is graphical on $\widehat{\mathcal{V}}$ with limit κ. Now (κ_n) is graphical on $\mathcal{V}$ with limit κ, and all conditions of Definition 2.9 for $\widehat{\mathcal{V}}$ apart from the last, (2.11), follow immediately from the corresponding conditions for $\mathcal{V}$. In other words, we must show that

$$\begin{aligned}\mathbb{E}\, e\bigl(G^{\widehat{\mathcal{V}}}(n, \kappa_n)\bigr)/n \to & \iint_{\mathcal{S}^2} \kappa(x,y)\, d\widehat{\mu}(x)\, d\widehat{\mu}(y) \\ &= \iint_{\mathcal{S}^2} \kappa(x,y)(1 - \rho(\kappa; x))(1 - \rho(\kappa; y))\, d\mu(x)\, d\mu(y).\end{aligned}$$

As we may couple G'_n and $G^{\widehat{\mathcal{V}}}(n, \kappa_n)$ to agree whp, and as the number of edges in either graph is bounded by that in G_n, and thus, divided by n, is uniformly integrable, it suffices to prove the same limiting formula for $\mathbb{E}\, e(G'_n)/n$. This is immediate from the definition of G'_n, condition (2.11) for G_n, and Theorem 3.5. □

Theorem 12.7. *Let (κ_n) be a graphical sequence of kernels on a (generalized) vertex space $\mathcal{V}$ with irreducible limit κ. If κ is supercritical, i.e., $\|T_\kappa\| > 1$, and $\sup_{x,y,n} \kappa_n(x, y) < \infty$, then $C_2(G_n) = O(\log n)$ whp.*

Proof. By Theorem 12.1, whp $C_2(G_n) = C_1(G^{\widehat{\mathcal{V}}}(n, \kappa_n))$, and the result follows from Theorems 12.3 and 12.5. □

Remark 12.8. One might think that $C_2(G_n) = O(\log n)$ would hold whp for any supercritical kernel κ, i.e., that Theorems 12.6 and 12.7 would hold without the condition $\inf_{x,y,n} \kappa_n(x, y) > 0$ or $\sup_{x,y,n} \kappa_n(x, y) < \infty$, at least for $\kappa_n = \kappa$, say. However, it is easy to construct counterexamples, similar to Example 8.6, by taking $x_i = i/n \in \mathcal{S} = (0, 1]$ and modifying a suitable kernel to introduce a largish star with centre 1 not joined to the giant component.

Such pathologies are not the only counterexamples: in the case where the x_i are uniformly distributed, for any function $\omega(n) = o(n)$, Example 4.12 provides an example of a supercritical random graph with $C_2(G_n) > \omega(n)$ whp; thus the $o_p(n)$ bound in Theorem 3.6 is best possible.

13. Vertex degrees

In this section we turn to the vertex degrees, proving Theorem 3.13: if κ_n is a graphical sequence of kernels on a vertex space $\mathcal{V}$ with limit κ, and we define $\lambda(x)$ by

$$\lambda(x) := \int_{\mathcal{S}} \kappa(x, y)\, d\mu(y),$$

then, writing Z_k for the number of vertices of degree k in $G^{\mathcal{V}}(n, \kappa_n)$, our aim is to show that

$$Z_k/n \xrightarrow{\mathrm{p}} \mathbb{P}(\Xi = k) = \int_{\mathcal{S}} \frac{\lambda(x)^k}{k!} e^{-\lambda(x)}\, d\mu(x),$$

where Ξ has the mixed Poisson distribution $\int_{\mathcal{S}} \mathrm{Po}(\lambda(x))\, d\mu(x)$.

In fact, Theorem 3.13 is stated for a generalized vertex space, and includes limiting results both for Z_k/n and for $Z_k/|V(G_n)|$. Since $|V(G_n)|/n \xrightarrow{\mathrm{p}} \mu(\mathcal{S})$, these results are equivalent. As usual, the statement for generalized vertex spaces reduces to that for vertex spaces (see Subsection 8.1), and what we must prove is exactly the statement above.

Proof of Theorem 3.13. Consider first the regular finitary case in Definition 4.4. Take a vertex v of type i, let D_v be its degree, and let $D_{v,j}$ be the number of edges from v to vertices of type j, $j = 1, \dots, r$; thus $D_v = \sum_j D_{v,j}$. Assume that $n \geq \max \kappa$ and condition on $n_1, \dots, n_r$. Then the

$D_{v,j}$ are independent for $j = 1, \dots, r$, and $D_{v,j} \sim \mathrm{Bi}\big(n_j - \delta_{ij}, \kappa(i,j)/n\big) \xrightarrow{\mathrm{d}} \mathrm{Po}\big(\mu_j \kappa(i,j)\big)$; hence

$$D_v \xrightarrow{\mathrm{d}} \mathrm{Po}\Big(\sum_j \mu_j \kappa(i,j)\Big) = \mathrm{Po}(\lambda(i)),$$

as $\lambda(i) = \int \kappa(i,j)\, d\mu(j) = \sum_j \kappa(i,j)\mu_j$. Consequently,

$$\mathbb{P}(D_v = k) \to \mathbb{P}\big(\mathrm{Po}(\lambda(i)) = k\big) = \frac{\lambda(i)^k}{k!} e^{-\lambda(i)}.$$

Let $Z_{k,i}$ be the number of vertices in $G^{\mathcal{V}}(n,\kappa)$ of type i with degree k. Then, still conditioning on $n_1, \dots, n_r$,

$$\frac{1}{n} \mathbb{E}\, Z_{k,i} = \frac{1}{n} n_i\, \mathbb{P}(D_v = k) \to \mu_i\, \mathbb{P}\big(\mathrm{Po}(\lambda(i)) = k\big).$$

It is easily checked that $\mathrm{Var}(Z_{k,i} \mid n_1, \dots, n_r) = O(n)$. Hence

$$\frac{1}{n} Z_{k,i} \xrightarrow{\mathrm{p}} \mathbb{P}\big(\mathrm{Po}(\lambda(i)) = k\big)\mu_i,$$

and thus, summing over i,

$$\frac{1}{n} Z_k = \sum_i \frac{1}{n} Z_{k,i} \xrightarrow{\mathrm{p}} \sum_i \mathbb{P}\big(\mathrm{Po}(\lambda(i)) = k\big)\mu_i = \mathbb{P}(\Xi = k).$$

This proves the theorem in the regular finitary case. In general, define κ_m^- by (7.5). Let $\varepsilon > 0$ be given. From (7.6) and monotone convergence, there is an m such that

$$\iint_{\mathcal{S}^2} \kappa_m^-(x,y)\, d\mu(x)\, d\mu(y) > \iint_{\mathcal{S}^2} \kappa(x,y)\, d\mu(x)\, d\mu(y) - \varepsilon. \tag{13.1}$$

For $n \ge m$ we have $\kappa_m^- \le \kappa_n$ by (7.7), so we may assume that $G(n,\kappa_m^-) \subseteq G(n,\kappa_n)$. (Here, as usual, we suppress the dependence on $\mathcal{V}$.) Then, using Proposition 8.9 twice and (13.1),

$$\begin{aligned}\frac{1}{n} e\big(G(n,\kappa_n) \setminus G(n,\kappa_m^-)\big) &= \frac{1}{n} e\big(G(n,\kappa_n)\big) - \frac{1}{n} e\big(G(n,\kappa_m^-)\big) \\ &\xrightarrow{\mathrm{p}} \tfrac{1}{2} \iint_{\mathcal{S}^2} \kappa(x,y)\, d\mu(x)\, d\mu(y) - \tfrac{1}{2} \iint_{\mathcal{S}^2} \kappa_m^-(x,y)\, d\mu(x)\, d\mu(y) < \frac{\varepsilon}{2},\end{aligned}$$

so whp $e\big(G(n,\kappa_n) \setminus G(n,\kappa_m^-)\big) < \varepsilon n$. Let us write $Z_k^{(m)}$ for the number of vertices of degree k in $G(n,\kappa_m^-)$. It follows that whp

$$|Z_k^{(m)} - Z_k| < 2\varepsilon n. \tag{13.2}$$

Writing $\Xi^{(m)}$ for the equivalent of Ξ defined using κ_m^- in place of κ, by the first part of the proof, $Z_k^{(m)}/n \xrightarrow{\mathrm{p}} \mathbb{P}(\Xi^{(m)} = k)$. Thus whp

$$|Z_k^{(m)}/n - \mathbb{P}(\Xi^{(m)} = k)| < \varepsilon. \tag{13.3}$$

Finally, we have $\mathbb{E}\,\Xi = \int_{\mathcal{S}} \lambda(x)\,d\mu(x) = \iint_{\mathcal{S}^2} \kappa(x,y)\,d\mu(x)\,d\mu(y)$. Since $\lambda^{(m)}(x) \le \lambda(x)$, we can assume that $\Xi^{(m)} \le \Xi$, and thus

$$\mathbb{P}(\Xi \ne \Xi^{(m)}) = \mathbb{P}(\Xi - \Xi^{(m)} \ge 1) \le \mathbb{E}(\Xi - \Xi^{(m)}) = \iint_{\mathcal{S}^2} \kappa - \iint_{\mathcal{S}^2} \kappa_m^- < \varepsilon. \tag{13.4}$$

Combining (13.2), (13.3) and (13.4), we see that $|Z_k/n - \mathbb{P}(\Xi = k)| < 4\varepsilon$ whp. □

Let Λ be the random variable $\lambda(\xi)$, where ξ is a random point in $\mathcal{S}$ with distribution μ. Then we can also describe the mixed Poisson distribution of Ξ as $\mathrm{Po}(\Lambda)$. Under mild conditions, the tail probabilities $\mathbb{P}(\Xi > t)$ and $\mathbb{P}(\Lambda > t)$ are similar for large t. We state this for the case of power-law tails; the result generalizes to regularly varying tails. As above, let D be the degree of a random vertex in $G^{\mathcal{V}}(n, \kappa_n)$. Let $Z_{\ge k}$ be the number of vertices with degree $\ge k$.

Corollary 13.1. *Let (κ_n) be a graphical sequence of kernels on a vertex space $\mathcal{V}$ with limit κ. Suppose that $\mathbb{P}(\Lambda > t) = \mu\{x : \lambda(x) > t\} \sim at^{-\alpha}$ as $t \to \infty$ for some $a > 0$ and $\alpha > 1$. Then*

$$Z_{\ge k}/n \xrightarrow{\mathrm{p}} \mathbb{P}(\Xi \ge k) \sim ak^{-\alpha},$$

where the first limit is for k fixed and $n \to \infty$, and the second for $k \to \infty$. In particular, $\lim_{n\to\infty} \mathbb{P}(D \ge k) \sim ak^{-\alpha}$ as $k \to \infty$.

Proof. It suffices to show that $\mathbb{P}(\Xi \ge k) \sim ak^{-\alpha}$; the remaining conclusions then follow from Theorem 3.13. For any $\varepsilon > 0$, $\mathbb{P}\big(\mathrm{Po}(\Lambda) > t \mid \Lambda > (1+\varepsilon)t\big) \to 1$ and $\mathbb{P}\big(\mathrm{Po}(\Lambda) > t \mid \Lambda < (1-\varepsilon)t\big) = o(t^{-\alpha})$ as $t \to \infty$, for example by standard Chernoff estimates [59, Remark 2.6]. It follows that $\mathbb{P}(\Xi > t) = \mathbb{P}\big(\mathrm{Po}(\Lambda) > t\big) \sim at^{-\alpha}$ as $t \to \infty$. □

This result shows that our model does include natural cases with power-law degree distributions. For example, taking $\mathcal{S} = (0,1]$ with the Lebesgue measure, and $\kappa_n(x,y) = \kappa(x,y) = c/\sqrt{xy}$ for $c > 0$ constant, we have $\lambda(x) = 2c/\sqrt{x}$, so $\mathbb{P}(\Lambda > t) = 4c^2/t^2$ for $t \ge 2c$. Thus, by Corollary 13.1, $\mathbb{P}(D \ge k) \sim 4c^2/k^2$ as $k \to \infty$. In fact, in this case $G^{\mathcal{V}}(n,\kappa)$ is the 'mean-field' version of the Barabási–Albert scale-free model; see Subsection 16.2. For other power laws, see Subsection 16.4.

14. Distances between vertices

One of the properties of inhomogeneous graphs that has received much attention is their 'diameter'. For example, considering the scale-free model of Barabási and Albert [9], the diameter was determined heuristically and experimentally to be $\Theta(\log n)$ in [4, 10, 81]; for a precise version of this model, the LCD model, the value $(1 + o(1)) \log n/ \log\log n$ was found rigorously in [23]; later, this value was also found heuristically in [39].

Often, the diameter is taken to mean the average distance between a random pair of vertices, or perhaps the 'typical' distance, although the usual graph theoretic definition (the maximum distance between a pair of vertices) is also used. Here we shall consider both interpretations.

14.1. **Typical distances.** In this subsection we study the 'typical' distance between vertices; our aim is to prove Theorem 3.14, giving upper and lower bounds on the distances between almost all pairs of vertices, showing that almost all pairs of vertices in the giant component are at distance roughly $\log n/\log \|T_\kappa\|$.

Many related results have been published, concerning random graphs with a fixed degree sequence, or random graphs with a given expected degree sequence; we shall only describe a few here. These models are similar to (and in some cases special cases of) the rank 1 case of our model; see Subsection 16.4. For example, Chung and Lu [35, 36] studied distances in a 'random graph with given expected degrees'. For power-law degrees with exponent $\beta > 3$, where their model is a special case of ours, they obtained an asymptotic diameter of $\log n$ over the log of the average of the squares of the degrees, a special case of Theorem 3.14; see Subsection 16.4 for the connection to our model.

Van der Hofstad, Hooghiemstra and Van Mieghem [53] (see also [54, 55, 56]) studied a model where the vertex degrees are i.i.d. with a certain distribution, and the graph is chosen uniformly among all graphs with these degrees. They analyze the growth of vertex neighbourhoods in this model by using a branching process; this process is single type, but the number of children of a particle is not Poisson, so it is rather different from the one considered here. They obtain very precise results on the distances between a random pair of vertices, showing that it is $\log n/\log c + O_p(1)$, where c is the expectation of $d(v)(d(v) - 1)$. (The -1 here comes from the degrees being (conditionally) fixed, rather than essentially Poisson.)

There are many other papers in this area, both heuristic and mathematical; we shall not attempt to list them. Let us mention only that Fernholz and Ramachandran [50], while mainly focussing on the diameter (see Subsection 14.2) also treat the typical distance between vertices. For further references we refer the reader to the discussion of related work in [53].

Let us now begin our preparation for the proof of Theorem 3.14. Let κ_n be a graphical sequence of kernels on a vertex space $\mathcal{V}$ with limit κ, and let $G_n = G^{\mathcal{V}}(n, \kappa_n)$. Note that we do not consider generalized vertex spaces; arguing as in Subsection 8.1, to prove Theorem 3.14 it suffices to consider vertex spaces. (For part (iv), we use also the fact that, by standard arguments, the conclusion holds if and only if (3.10) holds for $f(n) = \eta \log n$, for every $\eta > 0$.) We shall write $d_G(v, w)$ for the graph distance between two vertices v, w of a graph G, taking $d_G(v, w) = \infty$ if v and w are not in the same component of G. When the graph G is not specified, $G_n = G^{\mathcal{V}}(n, \kappa_n)$ is to be understood.

Lemma 14.1. *If κ is quasi-irreducible, then*

$$\big|\big\{\{v,w\} : d(v,w) < \infty\big\}\big| = \frac{C_1(G_n)^2}{2} + o_p(n^2) = \frac{\rho(\kappa)^2 n^2}{2} + o_p(n^2). \quad (14.1)$$

Proof. As noted in Subsection 3.5, by (3.9) this is immediate from Theorems 3.1 and 3.6. □

In particular, almost all pairs with either or both vertices outside the giant component are not connected at all, so we shall study the typical distance only in the supercritical case $\|T_\kappa\| > 1$.

Lemma 14.2. *Let κ be a regular finitary kernel on a vertex space $\mathcal{V}$ with $\|T_\kappa\| > 1$. For any $\varepsilon > 0$,*

$$\mathbb{E}\big|\big\{\{v,w\} : d(v,w) \le (1-\varepsilon)\log n/\log\|T_\kappa\|\big\}\big| = o(n^2).$$

Proof. Changing only the notation, we may assume that the type space $\mathcal{S}$ is finite, say $\mathcal{S} = \{1, 2, \dots, r\}$. It turns out that, as usual, we may assume that $\mu(\{i\}) > 0$ for every i; however, here we cannot simply ignore $o_p(n)$ edges, so an argument is needed. Suppose that $\mu(\{i\}) = 0$ for some i. Taking $\kappa'(i,j) = \kappa'(j,i) = \max\kappa$ for all j, and $\kappa'(j,k) = \kappa(j,k)$ for $j,k \ne i$, we have $\kappa \le \kappa'$, so we may couple $G_n := G^{\mathcal{V}}(n,\kappa)$ with $G'_n = G^{\mathcal{V}}(n,\kappa')$ so that $G_n \subseteq G'_n$. Note that $\|T_\kappa\| = \|T_{\kappa'}\|$, as $\kappa = \kappa'$ a.e.

For any $\eta > 0$, define μ' by $\mu'\{i\} = \eta$ and $\mu'\{j\} = (1-\eta)\mu(\{j\})$, $j \ne i$. Thus μ' is obtained from μ by shifting some measure from types other than i to type i. Changing the types of some vertices correspondingly, we obtain a vertex space $\mathcal{V}' = (\mathcal{S}, \mu', (\mathbf{x}'_n)_{n\ge1})$ such that whenever a vertex has type j in $\mathcal{V}$, it has either type j or type i in $\mathcal{V}'$. As $\kappa'(j,k)$ is maximal when one or both of j and k is equal to i, it follows that we can couple G'_n and $G''_n = G^{\mathcal{V}'}(n,\kappa')$ so that $G'_n \subseteq G''_n$. As $\eta \to 0$, the norm of $T_{\kappa'}$ defined with respect to μ' tends to the norm defined with respect to μ. Since $G_n \subseteq G''_n$, to prove Lemma 14.2 for G_n, it thus suffices to prove the same result for G''_n, defined on a vertex space with $\mu'\{i\} > 0$. Iterating, it suffices to prove Lemma 14.2 in the case where $\mu(\{i\}) > 0$ for every i.

Let $\Gamma_d(v) = \Gamma_d(v, G_n)$ denote the *d-distance set* of v in G_n, i.e., the set of vertices of G_n at graph distance exactly d from v, and let $\Gamma_{\le d}(v) = \Gamma_{\le d}(v, G_n)$ denote the d-neighbourhood $\bigcup_{d'\le d}\Gamma_{d'}(v)$ of v.

Let $0 < \varepsilon < 1/10$ be arbitrary. The proof of (12.1) involved first showing that, for n large enough, the neighbourhood exploration process starting at a given vertex v of G_n with type i (chosen without inspecting G_n) could be coupled with the branching process $\mathfrak{X}_{(1+2\varepsilon)\kappa}(i)$ so that the branching process dominates. In particular, the two processes can be coupled so that for every d, $|\Gamma_d(v)|$ is at most the number N_d of particles in generation d of $\mathfrak{X}_{(1+2\varepsilon)\kappa}(i)$. Elementary properties of the branching process imply that $\mathbb{E}\, N_d = O(\|T_{(1+2\varepsilon)\kappa}\|^d) = O(((1+2\varepsilon)\lambda)^d)$, where $\lambda = \|T_\kappa\| > 1$.

Set $D = (1-10\varepsilon)\log n/\log\lambda$. Then $D < (1-\varepsilon)\log n/\log((1+2\varepsilon)\lambda)$ if ε is small enough, which we shall assume. Thus,

$$\mathbb{E}\,|\Gamma_{\le D}(v)| \le \mathbb{E}\sum_{d=0}^{D} N_d = O(((1+2\varepsilon)\lambda)^D) = O(n^{1-\varepsilon}) = o(n).$$

Summing over v, the expected number of pairs of vertices within distance D is $o(n^2)$, and the result follows. □

We now turn to the reverse bound, showing that most vertices in the giant component are within distance roughly $\log n/\log\|T_\kappa\|$. First we consider two random vertices.

Lemma 14.3. *Let κ be a quasi-irreducible regular finitary kernel on a vertex space $\mathcal{V}$ with $\|T_\kappa\| > 1$, and let v and w be two vertices of $G^{\mathcal{V}}(n,\kappa)$ chosen independently and uniformly. Then, for any $\varepsilon > 0$,*

$$\mathbb{P}\big(d(v,w) < (1+\varepsilon)\log n/\log\|T_\kappa\|\big) \to \rho(\kappa)^2$$

as $n\to\infty$.

Proof. Note that an upper bound $\rho(\kappa)^2 + o(1)$ follows from Lemma 14.1, so it suffices to prove a corresponding lower bound. This time we may simply ignore types i with $\mu(\{i\}) = 0$, working entirely within the subgraph G_n' of $G_n = G^{\mathcal{V}}(n,\kappa)$ induced by vertices of the remaining types. As there $o_p(n)$ vertices of types i with $\mu(\{i\}) = 0$, changing ε slightly it suffices to prove the result for G_n'. Thus we shall assume that $\mu(\{i\}) > 0$ for every i. Also, restricting to a suitable subset of the types and renormalizing, we may and shall assume that κ is irreducible.

Fix $0 < \eta < 1/10$. We shall assume that η is small enough that $(1-2\eta)\lambda > 1$, where $\lambda = \|T_\kappa\|$. In the argument leading to (9.5) in proof of Lemma 9.6, we showed that, given $\omega(n)$ with $\omega(n) = o(n)$ and a vertex v of type i, the neighbourhood exploration process of v in G_n could be coupled with the branching process $\mathfrak{X}_{(1-2\eta)\kappa}(i)$ so that whp the former dominates until it reaches size $\omega(n)$. More precisely, writing $N_{d,k}$ for the number of particles of type k in generation d of $\mathfrak{X}_{(1-2\eta)\kappa}(i)$, and $\Gamma_{d,k}(v)$ for the set of type-k vertices at graph distance d from v, whp

$$|\Gamma_{d,k}(v)| \ge N_{d,k},\ k = 1,\dots,r,\ \text{ for all } d \text{ s.t. } |\Gamma_{\le d}(v)| < \omega(n). \tag{14.2}$$

The key point is that this coupling works because we have only 'looked at' $o(n)$ vertices at each step.

Let us call a kernel κ *bipartite* if

$$\mathcal{S} = L \cup R,\ \text{ with } \kappa(i,j) = 0 \text{ whenever } i,j\in L \text{ or } i,j \in R, \tag{14.3}$$

in which case the graph G_n is bipartite. For the moment, let us suppose that κ is not bipartite. Let $N_t(i)$ be the number of particles of type i in the tth generation of $\mathfrak{X}_\kappa$, and let $\mathbf{N}_t$ be the vector $N_t(1),\dots,N_t(r)$. Also, let $\boldsymbol{\nu} = (\nu_1,\dots,\nu_r)$ be the eigenvector of κ with eigenvalue λ (unique, up to normalization, as κ is irreducible). From standard branching process results, for example, [8, Theorems V.6.1 and V.6.2], we have

$$\mathbf{N}_t/\lambda^t \to X\boldsymbol{\nu} \quad \text{a.s.}, \tag{14.4}$$

where $X \geq 0$ is a real-valued random variable, X is continuous except that it has some mass at 0, and $X = 0$ if and only if the branching process eventually dies out.

Let D be the integer part of $\log(n^{1/2+2\eta})/\log((1-2\eta)\lambda)$. From (14.4), whp either $N_D = 0$, or $N_{D,k} \geq n^{1/2+\eta}$ for each k. Furthermore, as $\lim_{d\to\infty} \mathbb{P}(N_d \neq 0) = \rho((1-2\eta)\kappa)$ and $D \to \infty$, we have $\mathbb{P}(N_D \neq 0) \to \rho((1-2\eta)\kappa)$. Thus, if n is large enough,

$$\mathbb{P}\left(\forall k : N_{D,k} \geq n^{1/2+\eta}\right) \geq \rho((1-2\eta)\kappa) - \eta.$$

By Theorem 6.4, the right-hand side tends to $\rho(\kappa)$ as $\eta \to 0$. Hence, given any fixed $\gamma > 0$, if we choose η small enough we have

$$\mathbb{P}\left(\forall k : N_{D,k} \geq n^{1/2+\eta}\right) \geq \rho(\kappa) - \gamma \tag{14.5}$$

for n large enough. It is easy to check that $\mathbb{E}(|\Gamma_{\leq D}(v)|) = o(n^{2/3})$ if η is small enough; for example, we may argue as in the proof of Lemma 14.2. Hence,

$$|\Gamma_{\leq D}(v)| \leq n^{2/3} \quad \text{whp}, \tag{14.6}$$

and whp the coupling described in (14.2) extends at least to the D-neighbourhood.

Now let v and w be two fixed vertices of $G^{\mathcal{V}}(n,\kappa)$, of types i and j respectively. We explore both their neighbourhoods at the same time, stopping either when we reach distance D in both neighbourhoods, or we find an edge from one to the other, in which case v and w are within graph distance $2D+1$. We consider two independent branching processes $\mathfrak{X}_{(1-2\eta)\kappa}(i)$, $\mathfrak{X}'_{(1-2\eta)\kappa}(j)$, with $N_{d,k}$ and $N'_{d,k}$ vertices of type k in generation d respectively. By (14.6), whp we encounter $o(n)$ vertices in the explorations so, by the argument leading to (14.2), whp either the explorations meet, or

$$|\Gamma_{D,k}(v)| \geq N_{D,k} \quad \text{and} \quad |\Gamma_{D,k}(w)| \geq N'_{D,k}, \quad k = 1, \dots, r.$$

Using (14.5) and the independence of the branching processes, it follows that

$$\mathbb{P}\left(d(v,w) \leq 2D+1 \text{ or } \forall k : |\Gamma_{D,k}(v)|, |\Gamma_{D,k}(w)| \geq n^{1/2+\eta}\right) \\ \geq (\rho(\kappa) - \gamma)^2 - o(1). \tag{14.7}$$

Conditional on the second event in (14.7) holding and not the first, we have not examined any edges from $\Gamma_D(v)$ to $\Gamma_D(w)$, so these edges are present independently with their original unconditioned probabilities. For any i', j', the expected number of these edges is at least $|\Gamma_{D,i'}(v)||\Gamma_{D,j'}(v)|\kappa(i',j')/n$. Choosing i', j' such that $\kappa(i',j') > 0$, this expectation is $\Omega((n^{1/2+\eta})^2/n) =$

$\Omega(n^{2\eta})$. It follows that at least one edge is present with probability $1 - \exp(-\Omega(n^{2\eta})) = 1 - o(1)$. If such an edge is present, then $d(v, w) \leq 2D + 1$. Thus, (14.7) implies that

$$\mathbb{P}(d(v, w) \leq 2D + 1) \geq (\rho(\kappa) - \gamma)^2 - o(1) \geq \rho(\kappa)^2 - 2\gamma - o(1).$$

Choosing η small enough, we have $2D + 1 \leq (1 + \varepsilon) \log n / \log \lambda$. As γ is arbitrary, we have

$$\mathbb{P}(d(v, w) \leq (1 + \varepsilon) \log n / \log \lambda) \geq \rho(\kappa)^2 - o(1),$$

and the lemma follows.

The argument for the bipartite case is essentially the same, except that if v and w are of types in the same class of the bipartition, we should look for an edge between $\Gamma_D(v)$ and $\Gamma_{D-1}(w)$. □

Lemmas 14.1 and 14.3 have the following immediate consequence.

Corollary 14.4. *Let κ be a quasi-irreducible regular finitary kernel on a vertex space $\mathcal{V}$ with $\|T_\kappa\| > 1$. For any $\varepsilon > 0$,*

$$\big|\big\{\{v, w\} : d(v, w) \leq (1 + \varepsilon) \log n / \log \|T_\kappa\|\big\}\big| = \rho(\kappa)^2 n^2/2 + o_p(n^2). \quad (14.8)$$

Proof. It follows from Lemma 14.1 that the expected number of vertex pairs $\{v, w\}$ with $d(v, w) < \infty$ is $\rho(\kappa)^2 n^2/2 + o(n^2)$.

Fix $\varepsilon > 0$. From Lemma 14.3, the expected number of pairs of vertices at distance less than $d = (1 + \varepsilon) \log n / \log \|T_\kappa\|$ is $\rho(\kappa)^2 n^2/2 + o(n^2)$. Hence, the expected number of pairs with $d \leq d(v, w) < \infty$ is $o(n^2)$, so there are $o_p(n^2)$ such pairs. Using (14.1) again, (14.8) follows. □

After this preparation it is easy to deduce Theorem 3.14. As noted earlier, it suffices to consider vertex spaces, rather than generalized vertex spaces.

Proof of Theorem 3.14. Let κ_n be a graphical sequence of kernels on a vertex space $\mathcal{V}$ with limit κ, let $G_n = G^{\mathcal{V}}(n, \kappa_n)$, and let $\varepsilon > 0$ be fixed. We must prove four statements, which we recall separately below.

(i) The first part of Theorem 3.14 is exactly Lemma 14.1 (but with quasi-irreducible replaced by irreducible), which we have already proved.

(ii) We must show that if $\sup_{x,y,n} \kappa_n(x, y) < \infty$, then only $o_p(n^2)$ vertices of G_n are within distance $(1 - \varepsilon) \log n / \log \|T_\kappa\|$. As usual, we approximate with the regular finitary case. Let κ_m^+ be a sequence of regular finitary kernels on $\mathcal{V}$ with the properties guaranteed by Lemma 7.2. By Lemma 7.2(iii), $\|T_{\kappa_m^+}\| \to \|T_\kappa\| > 1$, so there is an m such that $(1 - \varepsilon/2)/\log \|T_{\kappa_m^+}\| \geq (1 - \varepsilon)/\log \|T_\kappa\|$. Fixing such an m, we may couple G_n and $G^{\mathcal{V}}(n, \kappa_m^+)$ so that $G_n \subseteq G^{\mathcal{V}}(n, \kappa_m^+)$ for $n \geq m$, and the result follows by applying Lemma 14.2 to $G^{\mathcal{V}}(n, \kappa_m^+)$ with $\varepsilon/2$ in place of ε.

(iii) This time we must show that if κ is irreducible and $\|T_\kappa\| < \infty$, then $\rho(\kappa)^2 n^2/2 + o_p(n^2)$ pairs of vertices of G_n are within distance $(1 + \varepsilon) \log n / \log \|T_\kappa\|$. By (14.1), it suffices to prove the lower bound. Again, we approximate with the regular finitary case, this time working with a graph $G'_n = G^{\mathcal{V}}(n, \hat{\kappa}_m^-) \subseteq G_n$. The argument is as above, but using the approximating kernels $\hat{\kappa}_m^-$ given by Lemma 7.3 instead of κ_m^+, and applying Corollary 14.4 with $\varepsilon/2$ in place of ε: by Lemma 7.3(i), the quasi-irreducibility condition of Corollary 14.4 is satisfied, while Lemma 7.3(ii) implies $\|T_{\hat{\kappa}_m^-}\| \nearrow \|T_\kappa\|$.

(iv) This time we must show essentially that if κ is irreducible and $\|T_\kappa\| = \infty$ then almost all pairs of vertices in the giant component are within distance $o(\log n)$. First, let $\eta > 0$. By the same proof as for part (iii) above, except that we have $\|T_{\hat{\kappa}_m^-}\| \to \infty$, we see that $\rho(\kappa)^2 n^2/2 + o_p(n^2)$ pairs of vertices of G_n are within distance $\eta \log n$. Since η is arbitrary, a standard argument shows that we can replace $\eta \log n$ by some function $f(n) = o(\log n)$. □

Remark 14.5. Part (ii) of Theorem 3.14 does not hold if we omit the condition that $\sup_{x,y,n} \kappa_n(x, y) < \infty$, even if $\kappa_n = \kappa$ for all n, with $\|T_\kappa\| < \infty$. To see this, let $\mathcal{S} = [0, 1]$ with μ the Lebesgue measure, and let $(\mathbf{x}_n)_{n\geq1}$ be disjoint deterministic sequences such that $\mathcal{V} = (\mathcal{S}, \mu, (\mathbf{x}_n)_{n\geq1})$ is a vertex space. We shall write $\mathbf{x}_n$ as $(x_1^{(n)}, \dots, x_n^{(n)})$ to emphasize the dependence of the terms on n; for example, we may take $x_i^{(n)} = (i - \sqrt{2}/2)/n$. Taking $\kappa(x, y) = 2$ for all x, y, the graph $G^{\mathcal{V}}(n, \kappa)$ is a supercritical Erdős–Rényi random graph.

Forming κ' by modifying κ on a set of measure zero, we can effectively add or delete $o(n)$ given edges to/from $G^{\mathcal{V}}(n, \kappa)$ whilst keeping κ' graphical on $\mathcal{V}$ with $\kappa' = \kappa = 2$ a.e. In particular, given $f(n) = o(n)$, taking $\kappa'(x_1^{(n)}, x_i^{(n)}) = n^2$, say, for $1 \leq i \leq f(n)$, we may ensure that in the graph $G'_n = G^{\mathcal{V}}(n, \kappa')$, whp the vertex 1 is joined to all of the vertices $2, 3, \dots, f(n)$. By Theorem 3.1, the giant component still has $\rho(2)n + o_p(n)$ vertices.

For any $\omega(n) \to \infty$ it is easy to check that if we choose $f(n)$ large enough, all but $o_p(n)$ vertices in the giant component are within distance $\omega(n)$ of one of the vertices $1, 2, \dots, f(n)$, and thus, all but $o_p(n^2)$ pairs of vertices in the giant component are within distance $2\omega(n) + 2$. Hence, even if $\|T_\kappa\|$ is bounded, the typical distance between vertices may be smaller than any given function tending to infinity.

Even if we allow $\|T_\kappa\| = \infty$, the typical distance cannot be as small as a constant: one can check that when κ is irreducible, for any C there are whp $\Theta(n^2)$ pairs of vertices in the giant component at distance at least C. In fact, there are $\Theta(n)$ vertices in the giant component whose C-neighbourhood is a path. This can be proved using a combination of the arguments leading to Theorems 3.1 and 9.1.

Remark 14.6. Using the same vertex space as in Remark 14.5, for any $f(n) = o(n)$, we can define a graphical sequence of kernels κ''_n with (irreducible) limit $\kappa = 2$ such that whp $G^{\mathcal{V}}(n, \kappa''_n)$ is obtained from the Erdős-Rényi graph $G(n, 2/n)$ by deleting $f(n)+1$ vertices and replacing them with a path of length $f(n)$ not joined to the rest of the graph. In this graph there are at least $(f(n)/3)^2$ pairs of vertices at distance at least $f(n)/3$. Hence the average distance between vertices (counting only pairs at finite distance) is at least $\Omega(f(n)^3)/n^2$, which may be much larger than $O(\log n)$: in fact, it may be larger than any given function that is $o(n)$. This is the reason for considering the distances between almost all pairs rather than the average distance in Theorem 3.14(ii).

Remark 14.7. In certain cases, we know better bounds than $o(\log n)$ on the typical distances between vertices. For example, the $(1+o(1))\log n/\log\log n$ formula for the diameter of the $m = 2$ LCD model proved by Bollobás and Riordan [23] certainly holds as a bound on the typical distances in the much simpler 'mean-field' case described in Subsection 16.2, where $\kappa(x, y) = 1/\sqrt{xy}$ and $x_i = i/n$.

However, without further restrictions (which could be on κ, or on the distributions of the x_i), we cannot strengthen the $o(\log n)$ bound in part (iv) of Theorem 3.14. Indeed, given a graphical kernel κ on a vertex space $\mathcal{V} = (\mathcal{S}, \mu, (\mathbf{x}_n)_{n\geq 1})$ with $\|T_\kappa\| = \infty$, and any function $g(n) \to \infty$, setting $\kappa_n = \kappa \wedge g(n)$ we have $\mathbb{E}\,|\Gamma_d(v)| \leq g(n)^d$ for a given vertex v and every $d \geq 0$. Indeed, we have $G(n, \kappa_n) \subseteq G(n, g(n)/n)$. The argument in the proof of Lemma 14.2 shows that the typical distance is at least $(1+o(1))\log n/\log g(n)$. Hence the $o(\log n)$ bound in part (iv) is best possible. A similar example may be constructed with a fixed kernel κ by modifying the sequences $(\mathbf{x}_n)_{n\geq 1}$ appropriately; take for example $\kappa(x, y) = 1/\sqrt{xy}$ and $x_i = \max\big\{i/n, 1/g(n)\big\}$.

14.2. **The diameter.** Let κ be a kernel on a (generalized) vertex space $\mathcal{V}$ in which the set of types is finite. In this subsection we study the diameter of $G_n = G^{\mathcal{V}}(n, \kappa)$, measured in the usual graph theoretical sense for disconnected graphs:

$$\operatorname{diam}(G_n) := \max\{d(v, w) \;:\; v, w \in V(G),\ d(v, w) < \infty\},$$

where $d(v, w)$ is the graph distance between v and w in G_n.

The following is a partial list of existing work on the diameter of sparse random graphs: Bollobás and Fernandez de la Vega [18] found the asymptotic diameter of random r-regular graphs, Łuczak [73] obtained detailed results for $G(n, c/n)$ with $c < 1$, Chung and Lu [33] studied $G(n, c/n)$, $c > 1$, and Fernholz and Ramachandran [50] obtained a precise result for random graphs with i.i.d. degrees (see below).

When G_n has finite-type, provided κ is not critical we can easily find the diameter of G_n in the form $(c + o(1))\log n$. The constant $c = c(\kappa)$ will be obtained from the branching process $\mathfrak{X}_\kappa$ in a simple way, different in the sub- and super-critical cases; see Theorems 14.8 and 14.11 below. Together, these results, which we shall prove separately, constitute Theorem 3.16.

The results in this subsection correspond to those of Fernholz and Ramachandran [50] for a different model, where the distribution of the vertex degrees is fixed, the vertex degrees are sampled independently from this distribution and, conditional on the degree sum being even, the graph is then chosen uniformly at random from all graphs with the given degree sequence. The special case of our model where κ has rank one is a special case of this model; see Subsection 16.4. In general, the two models are different. The proofs in [50] are much more complicated than those we shall present here, because their model does not have independence built in. Thus, roughly speaking, Fernholz and Ramachandran have to work to get the branching process approximation that we have here as our starting point. Also, here we keep things simple by considering only the finite type case; as noted in Subsection 3.5, even the single type case is non-trivial.

Throughout this subsection, when exploring the neighbourhoods of a vertex v in $G_n = G^{\mathcal{V}}(n,\kappa)$, we fix in advance an arbitrary order on the vertices of G_n. At each step in the exploration, among unexplored vertices at minimal distance from v, we choose the first vertex w in this order, and reveal all edges from w to vertices not yet reached by the exploration. In this way we reveal the vertex sets of the neighbourhoods $\Gamma_t(v)$, $t = 1, 2, \ldots$ successively. Furthermore, the graph we reveal is always a tree, rooted at v. We shall denote this graph by $T(v)$, and call it the *reduced component* of v. For $t \geq 0$ we write $T_t(v)$ for $T(v) \cap \Gamma_{\leq t}(v)$, the *reduced t-neighbourhood* of v.

Specifying which unexplored vertex to choose next does not affect the coupling arguments leading to (9.5), for example, where any unexplored vertex could be chosen at each step. The advantage is that the tree $T(v)$ is uniquely specified even if the component containing v has cycles; below we shall sum the probability that the *reduced* component of a vertex is a particular tree over all trees. Using the reduced component guarantees that the corresponding events are disjoint. Note that if v and w lie in the same component, then $d(v,w)$ is the same as the graph distance in $T(v)$ between the root, v, and w.

We shall first prove the subcritical case of Theorem 3.16, restated below.

Theorem 14.8. *Let κ be a kernel on a (generalized) vertex space $\mathcal{V} = (\mathcal{S}, \mu, (\mathbf{x}_n))$, with $\mathcal{S} = \{1, 2, \ldots, r\}$ finite and $\mu(\{i\}) > 0$ for each i. If $0 < \|T_\kappa\| < 1$, then*
$$\frac{\operatorname{diam}(G_n)}{\log n} \xrightarrow{\mathrm{p}} \frac{1}{\log \|T_\kappa\|^{-1}}$$
as $n \to \infty$, where $G_n = G^{\mathcal{V}}(n,\kappa)$.

Proof. We may assume without loss of generality that κ is irreducible. Also, by conditioning on the sequences $(\mathbf{x}_n)$, we may assume that $\mathcal{V}$ is a vertex space, and that the number n_i of vertices of type i is deterministic (see Subsection 8.1).

Let $p_{d,i}$ be the probability that the branching process $\mathfrak{X}_\kappa(i)$ survives for at least d generations. As the number of particles in the first generation that have descendants in generation $d+1$ has a Poisson distribution, we have
$$p_{d+1,i} = 1 - \exp\Big(-\sum_j \kappa(i,j)\mu(\{j\})p_{d,j}\Big).$$

Recalling that $p_{d,i} \searrow \rho(\kappa, i) = 0$ as $d \to \infty$, and using $1 - \exp(-x) = x + O(x^2)$, it follows easily that, for each i,

$$p_{d,i} = (\|T_\kappa\| + o(1))^d \quad \text{as} \quad d \to \infty. \tag{14.9}$$

Let $\omega = \omega(n) = A \log n$, where A is a constant, chosen large enough that the estimates below hold. As $\mathfrak{X}_\kappa$ is subcritical, $\rho_{\geq k}(\kappa)$ decays exponentially with k; see (12.2). Thus $\rho_{\geq \omega}(\kappa) = o(n^{-2})$, say. Let $d = (1 \pm \varepsilon) \log n / \log \|T_\kappa\|^{-1}$, where ε is a small positive constant; we shall consider both choices of sign below.

Let T be a rooted tree where each vertex has a type from $\mathcal{S} = \{1, 2, \dots, r\}$. We shall say that a tree T is *relevant* if it has height at least d and contains at most ω vertices. Let $\pi(T)$ be the probability that $\mathfrak{X}_\kappa$ is isomorphic to T, in the natural sense.

The sum of $\pi(T)$ over relevant T is the probability that $\mathfrak{X}_\kappa$ survives at least d generations and contains at most ω particles in total, which is $(\|T_\kappa\| + o(1))^d + o(n^{-2}) = n^{-1 \mp \varepsilon + o(1)}$.

Let $p(T)$ be the probability that the reduced component of a random vertex v of G_n is isomorphic to T in the natural sense. From the step-by-step exploration, one can check that

$$p(T) = (1 + o(1))^{|T|} \pi(T) = n^{o(1)} \pi(T)$$

for any relevant T: the proof is similar to that of Lemma 11.4, but one shows that at each step the conditional probability of finding the right number a of new neighbours of a particular type in the graph is within a factor $(1 + o(1))^{a+1}$ of the corresponding Poisson probability, as long as both a and the number of previously uncovered vertices are $o(n)$. Let σ be the sum of $p(T)$ over relevant T. Then it follows that $\sigma = n^{-1 \mp \varepsilon + o(1)}$. In particular, $n\sigma = o(1)$ if we take the plus sign in $d = (1 \pm \varepsilon) \log n / \log \|T_\kappa\|^{-1}$, and $n\sigma \to \infty$ if we take the minus sign.

Using (12.2) again, the expected number of vertices in G_n with more than $\omega(n)$ vertices in their (reduced or unreduced) component is $o(1)$ (in fact, $o(n^{-100})$), so whp there no such vertices. Taking the plus sign in d, the expected number of vertices in G_n whose reduced component is a relevant tree is $n\sigma = o(1)$. Together, these bounds show that whp every vertex v is such that all w with $d(v, w) < \infty$ have $d(v, w) < d$. Thus, whp $\operatorname{diam}(G_n) < d = (1 + \varepsilon) \log n / \log \|T_\kappa\|^{-1}$. As $\varepsilon > 0$ was arbitrary, this proves the upper bound in Theorem 14.8.

For the lower bound we take the minus sign in d, so $n\sigma = n^{\varepsilon + o(1)} \to \infty$, and use the second moment method. The key point is that if T, T' are relevant trees, and $p(T, T')$ is the probability that independently chosen random vertices v, w have *vertex-disjoint* reduced neighbourhoods isomorphic to T, T' respectively, then

$$p(T, T') = (1 + O(\omega/n))^\omega p(T) p(T') \sim p(T) p(T'). \tag{14.10}$$

This again follows from the step-by-step exploration, as finding one tree uses up at most ω vertices. (Note that we do not have $p(T) \sim \pi(T)$, because all we know about n_i is that $n_i \sim n\mu(\{i\})$.)

Let X_v be the indicator function of the event that the reduced neighbourhood of v is a relevant tree, and let $N = \sum_{v=1}^{n} X_v$, so $\mathbb{E}(N) = n\sigma$. Expanding $\mathbb{E}(N^2) = \sum_v \sum_w \mathbb{E}(X_v X_w)$, the contribution from pairs v, w in the same component is at most $\sum_v \mathbb{E}(X_v \omega) = n\sigma\omega$: if the component containing v is relevant, then by definition it contains at most ω vertices w. Using (14.10) above, it follows that

$$\mathbb{E}(N^2) = (1 + o(1))n^2\sigma^2 + O(n\sigma\omega).$$

Now $\mathbb{E}(N) = n\sigma = n^{\varepsilon+o(1)}$, which is much larger than ω, so $\mathbb{E}(N^2) \sim \mathbb{E}(N)^2$, and whp $N > 0$. So whp there is a vertex v whose neighbourhood is a relevant tree, and thus includes a vertex w at distance at least $d = (1 - \varepsilon)\log n/\log \|T_\kappa\|^{-1}$, completing the proof of Theorem 14.8. □

Remark 14.9. Recall that $p(T)$ was defined as the probability that the *reduced* neighbourhoods of a random vertex v are isomorphic to T, allowing the possibility that there are some edges within each $\Gamma_d(v)$. The reason was that, in proving the upper bound on the diameter, we must rule out components of large diameter that contain cycles, as well as components that are trees. If we redefine $p(T)$ to exclude edges within each $\Gamma_d(v)$, then $p(T)$ changes by a factor $(1 + O(|T|/n))^{|T|}$. As we only ever consider T with $|T| = O(\log n)$, this factor is $1 + o(1)$, and all our estimates go through. In particular, whp G_n contains a tree component of diameter at least $(1 - \varepsilon)\log n/\log \|T_\kappa\|^{-1}$.

Remark 14.10. One might expect Theorem 14.8 to generalize immediately from $\mathcal{S}$ finite to (at least) the case κ bounded. However, $\operatorname{diam}(G)$ does *not* always decrease when an edge is added to G, as the new edge might join two components. Thus one cannot just sandwich $G^{\mathcal{V}}(n,\kappa)$ between finite-type graphs and apply Theorem 14.8. In the unbounded case, the construction described in Remark 14.6 shows that for any $\omega(n) = o(n)$ one can construct a graph $G^{\mathcal{V}}(n,\kappa)$ with κ supercritical, such that $\operatorname{diam}(G^{\mathcal{V}}(n,\kappa)) \geq \omega(n)$ whp. Modifying the construction by starting with a subcritical Erdős-Rényi graph gives an example with κ subcritical.

We now turn to the supercritical case of Theorem 3.16, restated as Theorem 14.11 below. Recall that, given a supercritical kernel κ on a ground space $(\mathcal{S},\mu)$, there is a 'dual' kernel $\hat\kappa$ on a ground space $(\mathcal{S},\hat\mu)$, defined as follows: as a function on $\mathcal{S}\times\mathcal{S}$, $\hat\kappa = (1-\rho(\kappa))\kappa$, while $d\hat\mu(x) = (1-\rho(k;x))/(1-\rho(\kappa))\,d\mu(x)$. In particular, in the finite-type case, $\hat\mu(\{i\}) = (1-\rho(k;i))/(1-\rho(\kappa))\mu(\{i\})$. Note that we have chosen to renormalize the dual kernel defined in Definition 3.15 so that $\hat\mu$ is a probability measure. As discussed after Definition 3.15, this makes essentially no difference; however, it allows us to speak of the branching process $\mathfrak{X}_{\hat\kappa}$ started with a particle whose type is chosen according to $\hat\mu$. (In the remark after Definition 3.15

we wrote $\hat{\kappa}'$ and $\hat{\mu}'$ for the renormalized dual kernel and associated measure; here we write $\hat{\kappa}$ and $\hat{\mu}$ for notational convenience.) When we write $\|T_{\hat{\kappa}}\|$, we mean the norm of $T_{\hat{\kappa}}$ defined with respect to $(\mathcal{S}, \hat{\mu})$.

Theorem 14.11. *Let κ be an irreducible kernel on a (generalized) vertex space $\mathcal{V} = (\mathcal{S}, \mu, (\mathbf{x}_n))$, with $\mathcal{S} = \{1, 2, \ldots, r\}$ finite and $\mu(\{i\}) > 0$ for each i. If $\|T_\kappa\| > 1$, then*

$$\frac{\mathrm{diam}(G_n)}{\log n} \xrightarrow{\mathrm{p}} \frac{2}{\log \|T_{\hat{\kappa}}\|^{-1}} + \frac{1}{\log \|T_\kappa\|},$$

where $G_n = G^{\mathcal{V}}(n, \kappa)$.

The relevance of the dual kernel is that it describes components other than the giant component. In particular, it follows from Theorem 12.1 and Theorem 14.8 that the diameter of the largest 'small' component of G_n will be $(1 + o(1)) \log n / \log \|T_{\hat{\kappa}}\|^{-1}$. As we shall see, the same quantity will give the height of the tallest tree attached to the two-core. The diameter will be given by two such trees attached to vertices of the two-core at typical distance, $(1 + o(1)) \log n / \log \|T_\kappa\|$.

The idea of the proof is as follows: instead of considering the event that the neighbourhoods of a vertex v form a tree of height at least d, we consider the event that the neighbourhoods are *thin* for d generations, meaning that each generation has size at most ω, with $\omega = A \log n$ as before. For the upper bound, we will show that whp no vertex has neighbourhoods that are thin for more than $d = (1 + \varepsilon) \log n / \log \alpha^{-1}$ generations, where $\alpha = \|T_{\hat{\kappa}}\|$. For the lower bound, we will find two trees of height roughly d attached to typical vertices of the two-core.

The reason for considering thin neighbourhoods is that, once $\Gamma_t(v)$ is larger than $A \log n$ for some t, the neighbourhoods $\Gamma_s(v)$, $s \geq t$, grow reasonably rapidly.

From now on, we assume that κ is an irreducible kernel on a finite ground space $(\mathcal{S}, \mu)$, with $\mu(\{i\}) > 0$ for each i, and that the number n_i of vertices of each type i is deterministic, with $n_i/n \to \mu(\{i\})$. As before, it suffices to prove Theorem 14.11 under these assumptions. Let

$$t(v) := \min\{r : |\Gamma_r(v)| \geq \omega\}$$

denote the index of the first *thick* neighbourhood of a vertex v, when there is one.

Lemma 14.12. *For any $\varepsilon > 0$, whp the graph G_n does not contain two vertices v, w with the properties that $t(v)$, $t(w)$ are defined, $t(v), t(w) \leq n^{1/2}$, and $d(v, w) \geq t(v) + t(w) + (1 + \varepsilon) \log n / \log \|T_\kappa\|$.*

Proof. We show that the expected number of pairs is $o(1)$, by showing that the probability that a random pair v, w has the properties is $o(n^{-2})$. Explore the neighbourhoods of v and w simultaneously, stopping at the first thick neighbourhood of each, if there is one. Suppose, as we may, that $t(v)$ and

$t(w)$ are defined and at most $n^{1/2}$. Then with very high probability we have seen $o(n)$ vertices (the neighbourhoods can't have grown too much in the last step). If the neighbourhoods have already joined, we are happy. Otherwise, continue exploring. Simple Chernoff bounds show that for any $\eta > 0$, if A is chosen large enough, with probability $1 - o(n^{-100})$ the number of vertices of each type found at each subsequent step is within a factor $1 \pm \eta$ of its expectation. It follows that the neighbourhoods grow by a factor of $(1 \pm 2\eta)\|T_\kappa\|$ at each step, after a few steps to allow the distribution of types to converge to the relevant eigenvector of κ. Once both neighbourhoods reach size $n^{1/2+\eta}$, they join at the next step with very high probability. □

For the rest of this section, let $\alpha = \|T_{\hat\kappa}\|$ denote the norm of the dual kernel $\hat\kappa$. As $\mathcal{S}$ is finite, $\alpha < 1$ by Theorem 12.3. Recall, from the discussion before Lemma 6.6, that $\mathfrak{X}_{\hat\kappa}$ has the same distribution as $\mathfrak{X}_\kappa$ conditioned on extinction. Let $t_{d,n}$ be the probability that $\mathfrak{X}_\kappa$ stays alive but thin for d generations:

$$t_{d,n} := \mathbb{P}\big(1 \le |X_t| \le \omega \ : \ 1 \le t \le d\big),$$

where X_t is generation t of $\mathfrak{X}_\kappa$. Note that $t_{d,n}$ depends on n, via the definition of ω.

Lemma 14.13. *For any $\eta > 0$, if n is large enough, then*

$$(\alpha - \eta)^d \le t_{d,n} \le (\alpha + \eta)^d \tag{14.11}$$

holds for all d in the range $\frac{1}{2}\log n/\log\alpha^{-1} \le d \le 2\log n/\log\alpha^{-1}$.

Proof. We start with the lower bound.

Let $p'_{r,i}$ be the probability that the branching process $\mathfrak{X}_{\hat\kappa}(i)$, started with a particle of type i, survives for at least r generations. By (14.9), we have $p'_{r,i} = (\alpha + o(1))^r$ as $r \to \infty$. Let $p''_{r,j,i}$ be the probability that generation r of $\mathfrak{X}_{\hat\kappa}(j)$ consists of a single particle of type i. As the branching process is subcritical, one can check that $p''_{r,i,i} = \Theta(p'_{r,i})$ for r even. (The restriction r even is only needed if the kernel $\hat\kappa$ is bipartite, in the sense of (14.3).) We shall need only the much weaker statement that $p''_{r,i,i} = (\alpha + o(1))^r$ as $r \to \infty$ with r even; this can be proved along the same lines as (14.9): let N_1 be the number of particles x in the first generation of $\mathfrak{X}_{\hat\kappa}(j)$ with the property that the descendants of x in generation $d+1$ consist of a single particle of type i. Let N_2 be the number of particles x in the first generation of $\mathfrak{X}_{\hat\kappa}(j)$ with more than one descendant in generation $d+1$ of type i, or any descendants of types other than i. Note that N_1 and N_2 have independent Poisson distributions, with

$$\mathbb{E}(N_1) = \sum_k \hat\kappa(j,k)\hat\mu\{k\}p''_{d,k,i},$$

and

$$\mathbb{E}(N_2) = \sum_k \hat\kappa(j,k)\hat\mu\{k\}(p'_{d,k} - p''_{d,k,i}).$$

Since $p''_{d,k,i} \le p'_{d,k} \le \rho_{\ge d}(\hat{\kappa}; k) \to 0$, we have $\mathbb{E}(N_1)$, $\mathbb{E}(N_2) \to 0$. By definition,

$$p''_{d+1,j,i} = \mathbb{P}(N_1 = 1, N_2 = 0) = \mathbb{E}(N_1)\exp(-\mathbb{E}(N_1) - \mathbb{E}(N_2)) \sim \mathbb{E}(N_1).$$

In other words,

$$p''_{d+1,j,i} = (1 + o(1))\sum_k \hat{\kappa}(j,k)\hat{\mu}\{k\}p''_{d,k,i}.$$

Recalling that α is the norm of $T_{\hat{\kappa}}$ defined with respect to $\hat{\mu}$, it follows that $p''_{r,i,i} = (\alpha + o(1))^r$ as $r \to \infty$ with r even.

The probability that $\mathfrak{X}_{\hat{\kappa}}$ has any thick generation at all is at most the expected total size of $\mathfrak{X}_{\hat{\kappa}}$ divided by ω, which is $O(1/\omega) = O(1/\log n)$. Taking $r \to \infty$ slowly enough, for example $r = 2\lceil \log\log\log n \rceil$, this probability is much smaller than $p''_{r,i,i}$. Hence, with probability $p'''_{r,i} = (\alpha + o(1))^r$ the branching process $\mathfrak{X}_{\hat{\kappa}}(i)$ remains thin for r generations, and the rth generation is a single particle of type i. Restarting, we see that for any d we have $t_{d,n} \ge \mu(\{i\})(p'''_{r,i})^{\lceil d/r \rceil}$, and the lower bound in (14.11) follows.

Let $c > 0$ be a (small) constant. Simple Chernoff bounds show that, given that a certain generation t of $\mathfrak{X}_\kappa$ has size at least $c\log n$ (and given the numbers of particles of each type), generation $t + c_1$ has size at least ω with probability at least $1 - n^{-c_2}$, for some constants c_1, c_2 depending on c and κ. Hence, if $L = L(\kappa, c)$ is chosen large enough, the probability that $\mathfrak{X}_\kappa$ stays thin for d generations and has size at least $c\log n$ for the last L of these is at most n^{-100}: given that generation $d - L$ has size at least $c\log n$, the probability that generation $d - L + c_1$ is still thin is at most n^{-c_2}. If this generation is thin but also has size larger than $c\log n$, generation $d - L + 2c_1$ is unlikely to be thin, and so on.

Hence, $t_{d,n}$ is within n^{-100} of the probability that $\mathfrak{X}_\kappa$ stays thin for d generations and one of the last L of the first d generations has size at most $c\log n$. Let $q_r = q_{r,c}$ be the probability of the event that $\mathfrak{X}_\kappa$ survives for at least r generations, that the first r generations are thin, and that generation r has size at most $c\log n$. We have shown that

$$t_{d,n} \le \sum_{d-L \le r \le d} q_r + n^{-100}. \tag{14.12}$$

As a generation of size $c\log n$ has probability $n^{-O(c)}$ of dying out immediately, the probability that $\mathfrak{X}_\kappa$ survives for *exactly* r generations is at least $q_r n^{-O(c)}$. Hence, $p'_{r,i}$, the probability that $\mathfrak{X}_{\hat{\kappa}}$, which is just $\mathfrak{X}_\kappa$ conditioned on dying out, survives for r generations is also at least $q_r n^{-O(c)}$. Using $p'_{r,i} = (\alpha + o(1))^r$, it follows that $q_r \le n^{O(c)}(\alpha + o(1))^r$. In particular, choosing c small enough (depending on κ and η), for $r \ge \frac{1}{3}\log n/\log\alpha^{-1}$, say, we have $q_r \le (\alpha + \eta/2)^r$. Using (14.12), it follows that

$$t_{d,n} \le L(\alpha + \eta/2)^{d-L} + n^{-100} \le (\alpha + \eta)^d$$

for n large enough and d in the range considered in (14.11), completing the proof of Lemma 14.13. □

Let T be a rooted tree in which each vertex has a type from $\mathcal{S}$. Let $h(T)$ be the *height* of T, i.e., the maximal distance of a vertex form the root. We shall say that $\mathfrak{X}_\kappa$ is *consistent with* T if the first $h(T)$ generations of $\mathfrak{X}_\kappa$ are isomorphic to T, and write $\pi'(T)$ for the probability of this event. Analogously, we say that the reduced neighbourhoods of a vertex v of G_n are *consistent with* T if $T_{h(T)}(v)$, the tree formed by the first $h(T)$ such neighbourhoods, is isomorphic to T. We write $p'(T)$ for the probability that the reduced neighbourhoods of a random vertex of G_n are consistent with T. The definitions are almost the same as those of $\pi(T)$, $p(T)$, except that we do not care what happens after the first $h(T)$ generations. (We could have used these definitions in the subcritical case – there it was not essential that we explored the whole component.)

We are now ready to prove Theorem 14.11.

Proof of Theorem 14.11. As before, we may assume that $\mathcal{V}$ is a vertex space, and that the number n_i of vertices of type i is deterministic.

Let $\varepsilon > 0$ be fixed, and let $d = (1 \pm \varepsilon)\log n/\log \alpha^{-1}$, where $\alpha = \|T_{\hat\kappa}\|$, as above. As before, let $\omega = A\log n$, where A is a constant chosen large enough for our bounds to hold. Having chosen A, let A' be another constant chosen large enough that the bounds below hold.

Let $t^+_{h,n}$ be the probability that the branching process $\mathfrak{X}_\kappa$ is alive and thin for h generations (as in the definition of $t_{d,n}$), but that these first h generations contain more than $A'\log n$ particles. The proof of Lemma 14.13 shows that if A' is chosen large enough, then $t^+_{h,n} \le n^{-99}$ for any h. Indeed, the argument leading to (14.12) gives a corresponding bound on $t^+_{h,n}$ with q_r replaced by the probability q^+_r of an appropriate event, defined as q_r, but with the extra condition that there are at least $A'\log n/2$ particles in the first r generations. (We take A' large enough that $L\omega \le A'\log n/2$.) Then, as before, q^+_r can be bounded by $n^{O(1)}$ times the probability that $\mathfrak{X}_{\hat\kappa}$ survives for exactly r generations and contains at least $A'\log n/2$ particles. Using only the exponential decay of $\rho_{\ge k}(\hat\kappa)$, this is at most n^{-100} if A' is large enough, for any r.

Let us say that a tree T is *relevant* if T has height d, is thin, and contains at most $A'\log n$ vertices. As T has $O(\log n)$ vertices, we have $p'(T) = n^{o(1)}\pi'(T)$, as before. The sum of $\pi'(T)$ over all relevant T is exactly

$$t_{d,n} - t^+_{d,n} = (\alpha + o(1))^d - O(n^{-99}) = n^{-1\mp\varepsilon+o(1)},$$

by Lemma 14.13 and our bound on $t^+_{d,n}$ above. As before, the sign above is the opposite of the sign we choose in $d = (1 \pm \varepsilon)\log n/\log \alpha^{-1}$. It follows that the sum σ of $p'(T)$ over relevant trees is $n^{-1\mp\varepsilon+o(1)}$.

To prove the upper bound in Theorem 14.11, let $d = (1+\varepsilon)\log n/\log \alpha^{-1}$. Then we have $n\sigma = o(1)$, so whp no vertex of G_n has neighbourhoods consistent with a relevant tree. It is easy to check that whp no vertex has neighbourhoods consistent with a thin tree T' of height d that is not relevant (because it contains more than $A'\log n$ vertices): any such T' contains a subtree T'' given by the first h generations of T' for some h, such that T'' is thin and has between $A'\log n$ and $A'\log n+\omega=O(\log n)$ vertices. But then $p'(T'')=n^{o(1)}\pi'(T'')$, and $\sum_{T''}\pi'(T'')$ is at most $\sum_{h\le n}t^+_{h,n}\le n^{-98}$. It follows that whp no vertex of v has neighbourhoods consistent with any thin tree of height d. Thus, for every v, either $\Gamma_d(v)$ is empty, or $t(v)\le d$. Applying Lemma 14.12, the upper bound on $\operatorname{diam}(G_n)$ claimed in Theorem 14.11 follows.

For the lower bound, we aim to find many tall thin trees attached to the two-core. As in Remark 14.9, for this part of the proof we modify our notion of the consistency of the neighbourhoods of a vertex v of G_n with a tree T of height d, by disallowing edges within each $\Gamma_t(v)$, $t\le d$. We redefine $p'(T)$ correspondingly; as in Remark 14.9, this changes $p'(T)$ by a factor $1+o(1)$ for trees of the size we consider, so our estimate $p'(T)=n^{o(1)}\pi'(T)$ goes through.

A *good* tree T will be a relevant tree with height $d=(1-\varepsilon)\log n/\log\alpha^{-1}$ in which generation d consists of a single vertex a. Changing ε slightly, we shall assume that d is a multiple of the quantity r considered in the proof of Lemma 14.13. Then this proof shows that with probability at least $(\alpha+o(1))^d=n^{-1+\varepsilon+o(1)}$ the first d generations of $\mathfrak{X}_\kappa$ are thin and generation d consists of a single particle. With our bound on $t^+_{d,n}$, it follows that $\sum_T\pi'(T)=n^{-1+\varepsilon+o(1)}$, where the sum is over good trees. As $p'(T)=n^{o(1)}\pi'(T)$ for each good T, we have $\sum_T p'(T)=n^{-1+\varepsilon+o(1)}$. As in the subcritical case, the second moment method gives us many good trees in the graph, but this is not enough – they might not be attached to the two-core.

For $v\in V(G)$ and T a good tree, let $E_2(v,T)$ be the event that the following all hold, where $d_1=C\log\log n$, and C, C_1 are constants to be chosen below: the first d neighbourhoods of v form the tree T, the single $a\in\Gamma_d(v)$ has two neighbours b_1, b_2 in $\Gamma_{d+1}(v)$ each of which has at least ω 'descendants' in $\Gamma_{d+d_1}(v)$, and $|\Gamma_{\le d+d_1}(v)|\le(\log n)^{C_1}$. We claim that if the constants C and C_1 are chosen large enough, and v is a random vertex of G_n, then $\mathbb{P}(E_2(v,T))=\Omega(p'(T))$. To see this, note that $p'(T)$ is exactly the probability that the first condition is satisfied. Conditional on this happening, bounding the neighbourhood exploration below by a supercritical branching process shows that the existence of b_1, b_2 with the required properties has probability bounded away from zero. Finally, given that $|\Gamma_d(v)|=1$, the expected size of the next d_1 generations is at most $(2\sup\kappa)^{d_1}=(\log n)^{O(1)}$, so the claim follows.

Let $E_2(v)$ be the event that $E_2(v,T)$ holds for some good T. As $\sum_T p'(T) = n^{-1+\varepsilon+o(1)}$, we have

$$\mathbb{P}(E_2(v)) = \sum_T \mathbb{P}(E_2(v,T)) = n^{-1+\varepsilon+o(1)}, \tag{14.13}$$

where the sum is over good T.

We would like to show that whp there are two (in fact, many) vertices v, w for which $E_2(v)$, $E_2(w)$ hold. We could use the second moment method, but as we shall need the relevant neighbourhoods of v and w to be disjoint, it turns out to be easier to test vertices one by one.

Whether $E_2(v,T)$ holds can be determined by exploring the neighbourhoods of v, stopping when at most $M = (\log n)^{C_1}$ vertices have been uncovered. Let us construct a sequence of t tests, $t = n/M^3$, as follows. Each test starts from a vertex v_i, where the v_i are chosen independently and uniformly at random from $V(G_n)$. In the ith test, we explore the neighbourhoods of v_i, uncovering at most M vertices, and attempting to verify that $E_2(v_i)$ holds. We abort the attempt if we reach a vertex uncovered in a previous attempt. As at most $tM = n/M^2$ vertices have previously been uncovered, for each vertex we reach, the probability that it was previously uncovered is at most $O(M^{-2})$. As the ith test involves examining at most M vertices, conditional on everything so far, the ith test succeeds with probability

$$(1-O(M^{-2}))^M\, \mathbb{P}(E_2(v)) \sim \mathbb{P}(E_2(v)) = n^{-1+\varepsilon+o(1)}.$$

The number of tests that succeed dominates a binomial random variable with mean $tn^{-1+\varepsilon+o(1)} = n^{\varepsilon+o(1)} \to \infty$, so whp at least two tests succeed. Hence, whp there are vertices v, w in G_n for which $E_2(v)$, $E_2(w)$ hold, with the relevant neighbourhoods disjoint.

Now whp G_n has the property that whenever $E_2(v)$ holds, the corresponding $a \in \Gamma_d(v)$ is in the two-core of G_n. The argument is as for Lemma 14.12: we may continue expanding the large neighbourhoods of b_1, b_2 until they meet. Hence, whp G_n contains two vertices v_1, v_2 belonging to separate trees T_1, T_2 of height d, each attached to the two-core by the single vertex at distance d from the root. We shall need only this last fact, basic properties of the model, and Theorem 3.14.

Recall that G_n is a graph on labelled vertices $\{1,2,\ldots,n\}$, each of which has a type in $\mathcal{S}$. Given G_n and the vertex types, let us separate G_n into the two-core G^2, a list of trees T_i each attached to the two-core at some attachment vertex a_i, and the rest of G_n. So far, we remember the label of each vertex. Now, however, let us forget the labels of the attachment vertices a_i of T_i, while remembering their types. To reconstruct G_n, we should identify each attachment vertex with a vertex of G^2 of the same type. Moreover, we may pick these vertices independently and uniformly at random from the allowed vertices of G^2; this is because all possible (labelled) graphs formed in this way have the same number of vertices of each type, and the same number of edges between vertices of each pair of types, and hence the same probability in our model.

We have shown above that whp our list T_i contains two trees, say T_1 and T_2, each of which has a vertex v_i at distance d from the corresponding attachment vertex a_i. By (11.8), whp G^2 contains $\Theta(n)$ vertices. Hence, by Theorem 3.14(ii), whp almost all pairs of vertices of G^2 are at distance at least $d' = (1-\varepsilon)\log n/\log\|T_\kappa\|$. Hence, whp the vertices of G^2 at which we reattach the a_i are at distance at least d'. Thus, whp, $\operatorname{diam}(G_n) \ge d(v_1, v_2) \ge 2d + d'$, completing the proof of Theorem 14.11. □

15. The phase transition

Our main aim in this section is to prove Theorem 3.17, which claims that if a kernel κ on a ground space $(\mathcal{S}, \mu)$ is irreducible and satisfies (3.11), i.e.,

$$\sup_x \int_{\mathcal{S}} \kappa(x,y)^2 \, d\mu(y) < \infty, \tag{15.1}$$

then the function $c \mapsto \rho(c) := \rho(c\kappa)$ is analytic except at $c = c_0 := \|T_\kappa\|^{-1}$, that T_κ has an eigenfunction ψ of eigenvalue $\|T_\kappa\|$, and that every such eigenfunction is bounded and satisfies (3.12).

Proof of Theorem 3.17. Note that our assumption (15.1) on κ implies that T_κ is a Hilbert–Schmidt operator, and thus compact in L^2; see Lemma 5.15. It further implies, by the Cauchy–Schwarz inequality, that T_κ is bounded $L^2 \to L^\infty$, and that (5.1) holds for all x.

(i) It is trivial that the function $c \mapsto \rho(c) = \rho(c\kappa)$ is analytic for $c < c_0$, so we shall assume that $c > c_0$. We shall show that we can extend this function to a suitable neighbourhood of c in the complex plane, and that this extension is (complex) analytic; this implies that ρ is a real analytic function at c. Actually, we will show that there is an analytic map $z \mapsto \rho_z^+$ into $L^2(\mu)$, defined in a neighbourhood of c, such that $\rho_z^+ = \rho_{z\kappa}$ when z is real. Here, as before, $\rho_{z\kappa}$ is the function defined by $\rho_{z\kappa}(x) = \rho(z\kappa; x)$, the survival probability of the branching process $\mathfrak{X}_{z\kappa}(x)$, which starts with a particle of type x. We may then take $\rho^+(z) := \int \rho_z^+(x)\, d\mu(x)$ as the extension of ρ.

To show that the claimed extension exists, we will use the implicit function theorem for complex analytic functions in the Banach space L^2. Of course, in this proof we use the complex version of L^2. Recall that a function f mapping an open subset U of a complex Banach space E into another Banach space F is *analytic* if and only if it is differentiable (in the Fréchet sense) at every point in U; the derivative $f'(x)$ at a point $x \in U$ then is a continuous linear operator $E \to F$, see, e.g., Hervé [52, Section 3.1]. For background on differentiable functions in (real or complex) Banach spaces, see Cartan [32]; in particular, as a special case of [32, Theorème 4.7.1] (which holds in both the real and complex cases), we have the following.

Lemma 15.1 (The implicit function theorem)**.** *Let B be a complex Banach space and let $f_0 \in B$, $z_0 \in \mathbb{C}$. Let $\Psi : \mathbb{C} \times B \to B$ be an analytic function, and denote by $D_2\Psi(z, f)$ the partial derivative of Ψ with respect to the second variable, i.e., let $D_2\Psi(z, f)$ be the derivative of $f \mapsto \Psi(z, f)$. Suppose that $\Psi(z_0, f_0) = 0$ and that $D_2\Psi(z_0, f_0)$ is invertible. Then there exists a neighbourhood U of z_0 and an analytic function $z \mapsto f(z)$ defined in U such that $f(z_0) = f_0$ and $\Psi(z, f(z)) = 0$, $z \in U$.* □

For convenience, note that by replacing κ by $c\kappa$, we may assume that $c = 1$ (and thus $c_0 < 1$). We then apply Lemma 15.1 with $B = L^2(\mu)$, $z_0 = c = 1$, $f_0 = \rho_\kappa$ and $\Psi(z, f) = \Phi_\kappa(zf) - f$, where $\Phi_\kappa(f) = 1 - e^{-T_\kappa f}$ as above. Since T_κ is a bounded linear map $L^2 \to L^\infty$, and $g \mapsto e^g$ is analytic $L^\infty \to L^\infty$, Φ is analytic $L^2 \to L^\infty \subseteq L^2$ and thus Ψ is analytic. It is also easily seen that $D_2\Psi(z, f) = z\Phi'_\kappa(zf) - I$. It remains to show that that the partial derivative at $(1, \rho_\kappa)$ is invertible; we state this as another lemma.

Lemma 15.2. *Assume that κ is irreducible,* (15.1) *holds and* $c_0 = \|T_\kappa\|^{-1} < 1$*. Let $\tilde\mu$ be the measure $d\tilde\mu = e^{T_\kappa\rho_\kappa}\, d\mu$ on $\mathcal{S}$. Then $\|\Phi'_\kappa(\rho_\kappa)\|_{L^2(\tilde\mu)} < 1$, and hence $D_2\Psi(1, \rho_\kappa) = \Phi'_\kappa(\rho_\kappa) - I$ is invertible in $L^2(\tilde\mu)$ and in $L^2(\mu)$.*

Proof. Since T_κ maps L^2 into L^∞, we have $T_\kappa\rho_\kappa \in L^\infty$. Hence $L^2(\tilde\mu) = L^2(\mu)$ with equivalent norms.

We have

$$\Phi'_\kappa(f)(g) = e^{-T_\kappa f} T_\kappa g. \tag{15.2}$$

Hence, for any $g, h \in L^2(\tilde\mu) = L^2(\mu)$, writing $\bar h$ for the complex conjugate of h, we see that

$$\langle \Phi'_\kappa(\rho_\kappa)(g), h\rangle_{L^2(\tilde\mu)} = \int_{\mathcal{S}} e^{-T_\kappa\rho_\kappa}(T_\kappa g)\bar h e^{T_\kappa\rho_\kappa}\, d\mu = \int_{\mathcal{S}} (T_\kappa g)\bar h\, d\mu$$

is a hermitian form in g and h, because T_κ is a symmetric operator in $L^2(\mu)$. Hence $\Phi'_\kappa(\rho_\kappa)$ is a symmetric operator in $L^2(\tilde\mu)$. Furthermore, as remarked above, T_κ is compact in $L^2(\mu)$ and thus also in $L^2(\tilde\mu)$.

If we had $\|\Phi'_\kappa(\rho_\kappa)\|_{L^2(\tilde\mu)} \ge 1$, there would thus be an eigenfunction g with an eigenvalue λ with $|\lambda| \ge 1$. Let $h := |g| \ge 0$, so $\int h > 0$. Since $\Phi'_\kappa(\rho_\kappa)$ has a non-negative kernel by (15.2), we obtain

$$h \le |\lambda g| = |\Phi'_\kappa(\rho_\kappa)g| \le \Phi'_\kappa(\rho_\kappa)h = e^{-T_\kappa\rho_\kappa} T_\kappa h. \tag{15.3}$$

Moreover, $\rho_\kappa > 0$ a.e. and thus, by part (iii) of Lemma 5.6, $T_\kappa\rho_\kappa > 0$ a.e. and

$$T_\kappa\rho_\kappa < e^{T_\kappa\rho_\kappa} - 1 = e^{T_\kappa\rho_\kappa}\Phi_\kappa\rho_\kappa = e^{T_\kappa\rho_\kappa}\rho_\kappa \qquad \text{a.e.} \tag{15.4}$$

Multiplying (15.3) and (15.4) and integrating, recalling that $\int h > 0$, we obtain

$$\int hT_\kappa\rho_\kappa\, d\mu < \int \rho_\kappa T_\kappa h\, d\mu,$$

which contradicts the symmetry of T_κ.

This contradiction shows that $\|\Phi'_\kappa(\rho_\kappa)\|_{L^2(\tilde\mu)} < 1$. Hence $I - \Phi'_\kappa(\rho_\kappa)$ is invertible in $L^2(\tilde\mu) = L^2(\mu)$, completing the proof of Lemma 15.2. □

Continuing the proof of Theorem 3.17, we can now apply the implicit function theorem (Lemma 15.1) to conclude the existence of an analytic function $z \mapsto \rho_z^+ \in L^2(\mu)$ with $\Phi_\kappa(z\rho_z^+) = \rho_z^+$, defined in a complex neighbourhood U of 1. We may further (by continuity) assume that U is so small that $\|\rho_z^+\|_2 > 0$. For real $z \in U$, we thus have $\Phi_{z\kappa}(\rho_z^+) = \Phi_\kappa(z\rho_z^+) = \rho_z^+$ (in L^2, i.e., a.e.), so $\rho_z^+ = \rho_{z\kappa}$ a.e. by Theorem 6.1 and Remark 5.5, which completes the proof of (i).

(ii) This time we scale so that $\|T_\kappa\| = 1$, and thus $c_0 = 1$, and write $\rho_{1+\varepsilon}$ for $\rho_{(1+\varepsilon)\kappa}$; we assume below that $0 \le \varepsilon < 1$. Thus, by Theorem 6.1,

$$1 - e^{-(1+\varepsilon)T_\kappa \rho_{1+\varepsilon}} = \Phi_{(1+\varepsilon)\kappa}(\rho_{1+\varepsilon}) = \rho_{1+\varepsilon},$$

i.e.,

$$(1+\varepsilon)T_\kappa \rho_{1+\varepsilon} = -\ln(1-\rho_{1+\varepsilon}) = \rho_{1+\varepsilon} + R(\rho_{1+\varepsilon}) \tag{15.5}$$

where

$$R(f) := \frac{f^2}{2} + \frac{f^3}{3} + \dots \tag{15.6}$$

By Theorem 6.1, $\rho_{1+\varepsilon} > 0$ a.e. when $\varepsilon > 0$, but $\rho_1 = 0$. By Lemma 5.15, there exists an eigenfunction $\psi \in L^2$ with $\psi = T_\kappa \psi$, which now implies $\psi \in L^\infty$. Furthermore, ψ is determined up to a constant factor, so the coefficient $\int \psi \int \psi^2 / \int \psi^3$ does not depend on the choice of ψ. We will for convenience assume that ψ is chosen with $\psi \ge 0$ and $\int \psi^2 \, d\mu = \|\psi\|_2^2 = 1$.

The operator T_κ maps the subspace $\psi^\perp \subset L^2(\mu)$ into itself; let T'_κ denote the restriction of T_κ to this subspace. Then 1 is not an eigenvalue of T'_κ, and thus (since T_κ is compact), 1 does not belong to the spectrum of T'_κ, i.e., $I - T'_\kappa$ is invertible. By continuity, $I - (1+\varepsilon)T'_\kappa$ is also invertible for small ε, and there exists $\delta > 0$ and $C < \infty$ such that $\|\bigl(I - (1+\varepsilon)T'_\kappa\bigr)^{-1}\|_{\psi^\perp} \le C$ for $0 \le \varepsilon < \delta$, i.e.,

$$\|f\|_2 \le C \|\bigl(I - (1+\varepsilon)T_\kappa\bigr) f\|_2, \qquad 0 \le \varepsilon < \delta, \ f \in \psi^\perp. \tag{15.7}$$

Theorem 6.4 implies that $\rho_{1+\varepsilon} \searrow \rho_1$ a.e. as $\varepsilon \searrow 0$, and thus, by dominated convergence,

$$\|\rho_{1+\varepsilon}\|_2 \to 0. \tag{15.8}$$

We also have, by (15.5), $\rho_{1+\varepsilon} \le (1+\varepsilon)T_\kappa \rho_{1+\varepsilon}$ and thus, as T_κ is bounded from L^2 to L^∞,

$$\|\rho_{1+\varepsilon}\|_\infty \le (1+\varepsilon)\|T_\kappa \rho_{1+\varepsilon}\|_\infty \le C_1 \|\rho_{1+\varepsilon}\|_2. \tag{15.9}$$

In particular, by (15.8), $\|\rho_{1+\varepsilon}\|_\infty \to 0$ as $\varepsilon \to 0$.

Assume that $\varepsilon > 0$ is small enough to ensure that $\|\rho_{1+\varepsilon}\|_\infty < 1/2$. Then, by (15.6), $R(\rho_{1+\varepsilon}) \le \rho_{1+\varepsilon}^2$, and thus, using (15.9),

$$\|R(\rho_{1+\varepsilon})\|_2 \le \|\rho_{1+\varepsilon}^2\|_\infty = \|\rho_{1+\varepsilon}\|_\infty^2 \le C_2 \|\rho_{1+\varepsilon}\|_2^2. \tag{15.10}$$

Let Q be the orthogonal projection onto $\psi^\perp$ and let $\rho_{1+\varepsilon}^* := Q\rho_{1+\varepsilon}$. We thus have the orthogonal decomposition

$$\rho_{1+\varepsilon} = a_\varepsilon \psi + \rho_{1+\varepsilon}^*, \tag{15.11}$$

where

$$a_\varepsilon = \langle \rho_{1+\varepsilon}, \psi \rangle = \int_{\mathcal{S}} \rho_{1+\varepsilon} \psi \, d\mu. \tag{15.12}$$

Hence $0 \le a_\varepsilon \le \|\rho_{1+\varepsilon}\|_2$.

Applying the projection Q to (15.5) we find, since $QT_\kappa = T_\kappa Q$,

$$(1+\varepsilon) T_\kappa \rho^*_{1+\varepsilon} = \rho^*_{1+\varepsilon} + Q(R(\rho_{1+\varepsilon}))$$

and thus by (15.7) and (15.10), for $\varepsilon < \delta$,

$$\|\rho^*_{1+\varepsilon}\|_2 \le C \big\| \big(I - (1+\varepsilon)T_\kappa\big) \rho^*_{1+\varepsilon} \|_2 = C \| Q(R(\rho_{1+\varepsilon})) \|_2 \le C_3 \|\rho_{1+\varepsilon}\|_2^2. \tag{15.13}$$

Consequently, by (15.11) and (15.8), as $\varepsilon \to 0$,

$$\begin{aligned} a_\varepsilon &= \|a_\varepsilon \psi\|_2 = \|\rho_{1+\varepsilon} - \rho^*_{1+\varepsilon}\|_2 \\ &= \|\rho_{1+\varepsilon}\|_2 + O(\|\rho^*_{1+\varepsilon}\|_2) = \|\rho_{1+\varepsilon}\|_2 + O(\|\rho_{1+\varepsilon}\|_2^2) \sim \|\rho_{1+\varepsilon}\|_2. \end{aligned} \tag{15.14}$$

Furthermore, recalling (15.12) (twice), $\psi = T_\kappa \psi$, and (15.5),

$$\begin{aligned} (1+\varepsilon) a_\varepsilon = (1+\varepsilon) \langle T_\kappa \psi, \rho_{1+\varepsilon} \rangle &= \langle \psi, (1+\varepsilon) T_\kappa \rho_{1+\varepsilon} \rangle \\ &= \langle \psi, \rho_{1+\varepsilon} \rangle + \langle \psi, R(\rho_{1+\varepsilon}) \rangle = a_\varepsilon + \langle \psi, R(\rho_{1+\varepsilon}) \rangle. \end{aligned}$$

Therefore, appealing to (15.6), (15.11), (15.9), (15.13) and (15.14), we find that

$$\begin{aligned} \varepsilon a_\varepsilon &= \langle \psi, R(\rho_{1+\varepsilon}) \rangle = \langle \psi, \tfrac12 \rho^2_{1+\varepsilon} \rangle + O(\|\rho^3_{1+\varepsilon}\|_\infty) \\ &= \langle \psi, \tfrac12 a_\varepsilon^2 \psi^2) \rangle + \langle \psi, a_\varepsilon \psi \rho^*_{1+\varepsilon} \rangle + \langle \psi, \tfrac12 (\rho^*_{1+\varepsilon})^2 \rangle + O(\|\rho_{1+\varepsilon}\|^3_\infty) \\ &= \tfrac12 a_\varepsilon^2 \int \psi^3 \, d\mu + O(a_\varepsilon \|\rho^*_{1+\varepsilon}\|_2) + O(\|\rho^*_{1+\varepsilon}\|_2^2) + O(\|\rho_{1+\varepsilon}\|^3_\infty) \\ &= \tfrac12 a_\varepsilon^2 \int \psi^3 \, d\mu + O(a_\varepsilon \|\rho_{1+\varepsilon}\|_2^2) + O(\|\rho_{1+\varepsilon}\|_2^4) + O(\|\rho_{1+\varepsilon}\|^3_\infty) \\ &= \tfrac12 a_\varepsilon^2 \int \psi^3 \, d\mu + O(a_\varepsilon^3). \end{aligned}$$

If $\varepsilon > 0$ is small enough then, by (15.14), we have $a_\varepsilon > 0$, and so we can conclude that

$$\varepsilon = \tfrac12 a_\varepsilon \int \psi^3 \, d\mu + O(a_\varepsilon^2). \tag{15.15}$$

Finally, let $\varepsilon \searrow 0$. Then (15.14) and (15.8) imply that $a_\varepsilon \to 0$, and so from (15.15) we see that $\varepsilon \sim \frac12 a_\varepsilon \int \psi^3 \, d\mu = \Theta(a_\varepsilon)$, and, more precisely,

$$a_\varepsilon = \tfrac{2}{\int \psi^3 \, d\mu} \varepsilon + O(\varepsilon^2). \tag{15.16}$$

Consequently, using (15.11), (15.13), (15.14) and (15.16),

$$\begin{aligned} \rho(1+\varepsilon) = \int_{\mathcal{S}} \rho_{1+\varepsilon} \, d\mu &= a_\varepsilon \int_{\mathcal{S}} \psi \, d\mu + \int_{\mathcal{S}} \rho^*_{1+\varepsilon} \, d\mu \\ &= a_\varepsilon \int_{\mathcal{S}} \psi \, d\mu + O(\|\rho^*_{1+\varepsilon}\|_2) = 2 \tfrac{\int \psi \, d\mu}{\int \psi^3 \, d\mu} \varepsilon + O(\varepsilon^2), \end{aligned}$$

proving (3.12), and so completing the proof of Theorem 3.17. ☐

As noted in Subsection 3.6, Theorem 3.17 has a simple consequence, Corollary 3.18, showing that the rate $c_0\rho'_+(c_0)$ of emergence of the giant component at the phase transition is maximal in the Erdős–Rényi case, and, more generally, when (4.2) holds; see Example 4.6.

Proof of Corollary 3.18. Our aim is to show that if κ is an irreducible kernel on a ground space $(\mathcal{S}, \mu)$ for which (15.1) holds, and $c_0 := \|T_\kappa\|^{-1} > 0$, then $c_0\rho'_+(c_0) \leq 2$.

By Theorem 3.17, $c_0\rho'_+(c_0) = 2\int_\mathcal{S}\psi\int_\mathcal{S}\psi^2/\int_\mathcal{S}\psi^3$. By Lemma 5.15, we may assume that $\psi \geq 0$. Then, by Hölder's inequality, $\int_\mathcal{S}\psi \leq \bigl(\int_\mathcal{S}\psi^3\bigr)^{1/3}$ and $\int_\mathcal{S}\psi^2 \leq \bigl(\int_\mathcal{S}\psi^3\bigr)^{2/3}$, with equality if and only if ψ is a.e. constant, i.e., if and only if the constant function 1 is an eigenfunction of T_κ, which is equivalent to (4.2). □

Turning to the number of edges at the phase transition, we shall next prove Proposition 3.4, which says that if κ_n is a graphical sequence of kernels on a vertex space $\mathcal{V}$ with limit κ, and $\|T_\kappa\| = 1$, then $\frac{1}{n}e(G^\mathcal{V}(n,\kappa_n)) \xrightarrow{\mathrm{p}} \frac{1}{2}\iint\kappa \leq 1/2$, with equality if and only if (4.2) holds.

Proof of Proposition 3.4. By Proposition 8.9 we have $e\bigl(G^\mathcal{V}(n,\kappa_n)\bigr)/n \xrightarrow{\mathrm{p}} \gamma$, where $\gamma := \frac{1}{2}\iint\kappa$. If $\|T_\kappa\| = 1$, then

$$\frac{1}{2}\iint\kappa = \frac{1}{2}\langle 1, T_\kappa 1\rangle \leq \frac{1}{2}\|T_\kappa\| = \frac{1}{2},$$

with equality if and only if the constant function 1 is an eigenfunction, i.e., if and only if (4.2) holds. □

Remark 15.3. Proposition 3.4 says that the number of edges at the phase transition is *largest* in the classical Erdős–Rényi case, and is *strictly smaller* in all other cases except some very homogeneous ones. Together, Corollary 3.18 and Proposition 3.4 say roughly that inhomogeneities make the giant component appear sooner, but grow more slowly (initially, at least).

Remark 15.4. Another way to study the number of edges when the giant component is born is to consider the graph process in Remark 2.6. Let us stop the growth when the largest component first has at least $\omega(n)$ vertices, where $\omega(n)$ is a function chosen in advance, with $\omega(n) = o(n)$, and $\omega(n)$ increasing sufficiently rapidly with n. (If κ is bounded, we can take any ω with $\log n \ll \omega(n) \ll n$, see Theorem 3.12.) Then, for any $\varepsilon > 0$, whp we stop at a time between $(c_0-\varepsilon)/n$ and $(c_0+\varepsilon)/n$, where $c_0 := \|T_\kappa\|^{-1}$, and it follows easily from Proposition 8.9 that if N is the number of edges when we stop, then $N/n \xrightarrow{\mathrm{p}} \frac{1}{2}\iint\kappa$. Again we see that the number of edges required for a giant component is largest in the homogeneous case.

Remark 15.5. The proof of part (i) of Theorem 3.17 shows that the function $\rho(c\kappa;\cdot)$ defined before (2.13) depends analytically on $c \neq c_0$ as an element of L^2. Hence ν_n^1 in Theorem 9.10 also depends analytically on $c \neq c_0$.

We expect the following extensions of Theorem 3.17 to hold.

Conjecture 15.6. *Theorem 3.17(i) holds without the condition* (3.11)*, i.e., $\rho(c)$ is always analytic except at c_0.*

Conjecture 15.7. *Let κ be an irreducible kernel on a ground space $(\mathcal{S}, \mu)$. Then equation* (3.12) *holds with the larger error term $o(\varepsilon)$ whenever T_κ has an eigenfunction ψ of eigenvalue $\|T_\kappa\|$ with $\int_{\mathcal{S}} \psi^3 < \infty$; conversely, $\rho'_+(c_0) = 0$ if $c_0 > 0$ but no such ψ exists or ψ exists with $\int_{\mathcal{S}} \psi^3 = \infty$. Cf. Subsection 16.4.*

16. Applications and relationship to earlier results

In this section we apply our general results to several specific models that have been studied in recent years, and describe the relationships between our results and various earlier results.

16.1. **Dubins' model.** A common setting is the following: the vertex space $\mathcal{V}$ is $(\mathcal{S}, \mu, (\mathbf{x}_n)_{n\geq 1})$, where $\mathcal{S} = (0, 1]$, μ is the Lebesgue measure, and $\mathbf{x}_n = (x_1, \dots, x_n)$ with $x_i = i/n$. In this case, (2.3) gives $p_{ij} = \kappa(i/n, j/n)/n \wedge 1$ for the probability of an edge between vertices i and j. We shall consider several choices of κ in some detail.

Observe first that if κ is a positive function on $(0, \infty)^2$ that is homogeneous of degree -1, then (2.3) yields $p_{ij} = \kappa(i, j) \wedge 1$. Since this does not depend on n, in this case we can also consider the infinite graph $G(\infty, \kappa)$, defined in the same way as $G_n = G^{\mathcal{V}}(n, \kappa)$ but on the vertex set $\{1, 2, \dots\}$. Note that the graphs $G^{\mathcal{V}}(n, \kappa)$ are induced subgraphs of $G(\infty, \kappa)$ and that we can construct them by successively adding new vertices, and for each new vertex an appropriate random set of edges to earlier vertices.

We first consider $\kappa(x, y) = c/(x \vee y)$ with $c > 0$, so that if $j \geq c$ then

$$p_{ij} = c/j \qquad \text{for } i < j. \tag{16.1}$$

In this case we can regard $G^{\mathcal{V}}(n, \kappa)$ as a sequence of graphs grown by adding new vertices one at a time where, when vertex k is added, it gets $\mathrm{Bi}(k-1, c/k)$ edges, whose other endpoints are chosen uniformly among the other vertices. (We might instead take $\mathrm{Po}(c) \wedge (k-1)$ new edges, without any difference in the asymptotic results below.)

This infinite graph $G(\infty, \kappa)$ was considered by Dubins in 1984, who asked when $G(\infty, \kappa)$ is a.s. connected. Dubins' question was answered partially by Kalikow and Weiss [60]. A little later Shepp [86] proved that $G(\infty, \kappa)$ is a.s. connected if and only if $c > 1/4$. This result was generalized to more general homogeneous kernels by Durrett and Kesten [46].

The finite random graph $G^{\mathcal{V}}(n, \kappa)$ with this κ, i.e., with edge probabilities given by (16.1), has been studied by Durrett [44], who points out that it has the same critical value $c = 1/4$ for the emergence of a giant component as the infinite version has for connectedness, and by Bollobás, Janson and Riordan [19] who rigorously show that this example has a phase transition with infinite exponent. More precisely, denoting $\rho(\kappa)$ by $\rho(c)$, it was shown by Riordan [84] that

$$\rho(1/4+\varepsilon) = \exp\bigl(-\tfrac{\pi}{2}\varepsilon^{-1/2} + O(\log\varepsilon)\bigr). \tag{16.2}$$

A similar formula for the closely related CHKNS model (see Subsection 16.3), introduced by Callaway, Hopcroft, Kleinberg, Newman and Strogatz [30], had been given earlier by Dorogovtsev, Mendes and Samukhin [43] using non-rigorous methods.

To find the critical value by our methods, we have to find the norm of T_κ on $L^2(0,1)$. Using the isometry $U : f \mapsto e^{-x/2}f(e^{-x})$ of $L^2(0,1)$ onto $L^2(0,\infty)$, we may instead consider $\widetilde{T}_\kappa := UT_\kappa U^{-1}$, which by a simple calculation is the integral operator on $L^2(0,\infty)$ with kernel

$$\tilde\kappa(x,y) = e^{-x/2}\kappa(e^{-x},e^{-y})e^{-y/2} = ce^{-x/2-y/2+x\wedge y} = ce^{-|x-y|/2}. \tag{16.3}$$

Hence $\widetilde{T}_\kappa$ is the restriction to $(0,\infty)$ of the convolution with $h(x) := ce^{-|x|/2}$. Because of translation invariance, it is easily seen that $\widetilde{T}_\kappa$ has the same norm as convolution with h on $L^2(-\infty,\infty)$, and taking the Fourier transform we find

$$\|T_\kappa\| = \|\widetilde{T}_\kappa\| = \|f \mapsto h*f\|_{L^2(-\infty,\infty)} = \sup_{\xi\in\mathbb{R}} |\hat h(\xi)| = \int_{-\infty}^{\infty} h(x)\,dx = 4c.$$

Thus, Theorem 3.1 shows that there is a giant component if and only if $c > 1/4$, as shown in Durrett [44] and [19].

To find the size of the giant component is more challenging, and we refer to Riordan [84] for a proof of (16.2). Note that the hypothesis (3.11) of Theorem 3.17 fails, as do the conclusions in part (ii). Indeed, it is easy to see that T_κ is a non-compact operator, and that it has no eigenfunctions at all in L^2. We suspect that this is connected to the fact that the phase transition has infinite exponent.

16.2. **The mean-field scale-free model.** Another interesting case with a homogeneous kernel as in Subsection 16.1 is $\kappa(x,y) = c/\sqrt{xy}$ with $c > 0$; then, for $ij \ge c^2$, we have

$$p_{ij} = c/\sqrt{ij}. \tag{16.4}$$

This model has been studied in detail by Riordan [84]. Considering the sequence $G^{\mathcal{V}}(n,\kappa)$ as a growing graph, in this case, together with each new vertex we add a number of edges that has approximately a Poisson $\mathrm{Po}(2c)$ distribution; the other endpoint of each edge is chosen with probability proportional to $i^{-1/2}$, which is approximately proportional to the degree of vertex i. Hence, this random graph model resembles the growth with preferential attachment model of Barabási and Albert [9], which was made precise as the *LCD model* by Bollobás and Riordan [23]; see also [84]. In fact, up to a factor of $1+o(i^{-1})$ in the edge probabilities, the model defined by (16.4) is the so called 'mean-field' version of the Barabási–Albert model, having the same individual edge probabilities, but with edges present independently. This (by now common) use of 'mean-field' is not the standard one in physics, where it normally means that all vertices interact equally. (So the mean-field random graph model is $G(n, p)$.)

In this case, T_κ is an unbounded operator, because $x^{-1/2} \notin L^2(0,1)$, and thus there is no threshold. In other words, $\rho(c) := \rho(\kappa) > 0$ for every $c > 0$.

As shown by Riordan [84], $\rho(c)$ grows very slowly at first in this case too; more precisely,

$$\rho(c) \sim 2e^{1-\gamma} \exp\bigl(-1/(2c)\bigr) \qquad \text{as } c \to 0, \tag{16.5}$$

where γ is Euler's constant; see also Subsection 16.4 below. The result in [84] for the Barabási–Albert model is different, showing that in this model the dependence between edges is important.

Remark 16.1. Random graphs related to the ones defined here and in Subsection 16.1 but with some dependence between edges (and thus not covered by the present paper) can be obtained by adding at each new vertex a number of edges with some other distribution, for example $\mathrm{Bi}(m, p)$ for some fixed m and p. Such random graphs have been considered in [22, 25, 41, 84], and these papers show that not only the expected numbers of edges added at each step are important, but also the variances; the edge dependencies shift the threshold.

16.3. **The CHKNS model.** We next consider the CHKNS model of Callaway, Hopcroft, Kleinberg, Newman and Strogatz [30]. Here, the graph grows from a single vertex; vertices are added one by one, and after each vertex is added, an edge is added with probability δ; the endpoints are chosen uniformly among all existing vertices. (Multiple edges are allowed; this does not matter for the asymptotics.)

Following Durrett [44], we consider a modification (which is perhaps at least as natural): after adding each vertex, add a Poisson $\mathrm{Po}(\delta)$ number of edges to the graph, again choosing the endpoints of these edges uniformly at random. Thus, when vertex k is added, each existing pair of vertices acquires $\mathrm{Po}\bigl(\delta/\binom{k}{2}\bigr)$ new edges, and these numbers are independent. When we have reached n vertices, the number of edges between vertices i and j, with $1 \le i \le j \le n$, is thus Poisson with mean

$$e_{ij} := \sum_{k=j}^{n} \frac{\delta}{\binom{k}{2}} = 2\delta \sum_{k=j}^{n} \frac{1}{k(k-1)} = 2\delta\Bigl(\frac{1}{j-1} - \frac{1}{n}\Bigr), \tag{16.6}$$

and the probability that there is one or more edges between i and j is $p_{ij} := 1 - \exp(-e_{ij})$.

Hence, ignoring multiple edges, we have a graph G_n of our type, with $\mathcal{S} = (0, 1]$, μ Lebesgue measure, $x_i = i/n$ and

$$\begin{aligned}\kappa_n(x, y) &:= n\Big(1 - \exp\Big(-2\delta\Big(\frac{1}{n(x \vee y) - 1} - \frac{1}{n}\Big)\Big)\Big)\\ &\to \kappa(x, y) := 2\delta\Big(\frac{1}{x \vee y} - 1\Big).\end{aligned} \tag{16.7}$$

The conditions of Theorem 3.1 are immediately verified, and thus $C_1(G_n)/n \xrightarrow{\mathrm{p}} \rho(\kappa)$.

Instead of adding a Poisson number of edges at each step, we could add a binomial number by adding, after vertex k, each possible edge with probability $\delta/\binom{k}{2}$. We obtain the same results with slightly different κ_n but the same κ.

The original CHKNS model, $\widetilde{G}_n$, say, can be treated by the argument in [19]. It follows that $C_1(\widetilde{G}_n)/n \xrightarrow{\mathrm{p}} \rho(\kappa)$ holds for the CHKNS model too.

In particular, the threshold for the CHKNS model, as well as for Durrett's modification, is given by $\|T_\kappa\| = 1$, or $2\delta = \|T\|^{-1}$, where T is the integral operator with kernel $1/(x \vee y) - 1$ on $L^2(0, 1)$. This kernel is strictly smaller that the kernel $1/(x \vee y)$ considered in Subsection 16.1. However, changing variables as in (16.3), we see that T is equivalent to the operator on $L^2(0, \infty)$ with kernel $e^{-|x-y|/2} - e^{-(x+y)/2}$. Using translational invariance of the operator with kernel $e^{-|x-y|/2}$ considered in Subsection 16.1, considering functions supported in (R, ∞) and letting $R \to \infty$, it is easily seen that T has the same norm as this operator, namely 4.

Thus the thresholds for the CHKNS model and Durrett's modification are both given by $2\delta = 1/4$, i.e. $\delta = 1/8$, as was found by non-rigorous arguments by Callaway, Hopcroft, Kleinberg, Newman and Strogatz [30] and Dorogovtsev, Mendes and Samukhin [43], and first proved rigorously by Durrett [44]; see also [19].

To study the size of the giant component in these models, let us write $\kappa_0(x, y) := 1/(x \vee y)$ and $\kappa_1(x, y) := 1/(x \vee y) - 1$. Then $\kappa_1 < \kappa_0$, and thus $\rho(c\kappa_1) \le \rho(c\kappa_0)$ for each $c > 0$. (We have strict inequality for $c > 1/4$, see Remark 3.8; note that, as pointed out by Durrett [44], we have the same threshold 1/4 for both kernels although we have twice as many edges in $G(n, c\kappa_0)$ as in $G(n, c\kappa_1)$.) On the other hand, let $\eta > 0$ and consider only vertices $i \le j \le \eta n$. Then

$$n^{-1} c\kappa_1(i/n, j/n) = c\Big(\frac{1}{j} - \frac{1}{n}\Big) \ge \frac{(1 - \eta)c}{j}.$$

Hence, cf. (16.1), $G(n, c\kappa_1) \supseteq G(\eta n, (1 - \eta)c\kappa_0)$. (Note that these graphs have different numbers of vertices.) Thus, for every η with $0 < \eta < 1$,

$$\rho(c\kappa_0) \ge \rho(c\kappa_1) \ge \eta\rho\big((1 - \eta)c\kappa_0\big).$$

Taking $c = 1/4 + \varepsilon$ and $\eta = \varepsilon^2$, relation (16.2) for $\rho(c\kappa_0)$ implies the same estimate for $\rho(c\kappa_1)$; in other words, if $\delta = 1/8 + \varepsilon$, then the size of the giant component is given by

$$\rho(\kappa) = \rho(2\delta\kappa_1) = \exp\Big(-\frac{\pi}{2\sqrt{2}}\varepsilon^{-1/2} + O(\log\varepsilon)\Big).$$

As noted in Subsection 16.1, a similar formula (with no error term, and a particular constant in front of the exponential) was given by Dorogovtsev, Mendes and Samukhin [43], with a derivation part of which can be made rigorous; see Durrett [44].

16.4. **The rank 1 case.** In this subsection we consider a special case of our general model that, while very restrictive, is also very natural, and includes or is closely related to many random graph models considered by other authors. This is the *rank 1* case, where the kernel κ has the form $\kappa(x,y) = \psi(x)\psi(y)$ for some function $\psi > 0$ on $\mathcal{S}$. We shall assume that the kernel is graphical; in particular we assume $\int \psi\, d\mu < \infty$, but not necessarily that $\int \psi^2\, d\mu < \infty$.

The function $\psi(x)$ can be interpreted as the "activity" of a vertex at x, with the probability of an edge between two vertices proportional to the product of their activities. In the rank 1 case, $T_\kappa f = \big(\int f\psi\big)\psi$, so

$$\|T_\kappa\| = \|\psi\|_2^2 = \int \psi^2\, d\mu \le \infty. \tag{16.8}$$

Thus T_κ is bounded if and only if $\psi \in L^2$, in which case T_κ has rank 1, so it is compact, and ψ is the unique (up to multiplication by constants) eigenfunction with non-zero eigenvalue.

By Theorem 3.13, the distribution of vertex degrees is governed by the distribution of the function $\lambda(x) = (\int \psi\, d\mu)\psi(x)$ on $(\mathcal{S}, \mu)$. In particular, by Corollary 13.1, the degree sequence will (asymptotically) have a power-law tail if the distribution of $\lambda(x)$ has; for example, if $\mathcal{S} = (0,1]$ with μ Lebesgue measure, and $\psi(x) = cx^{-1/p}$. (Another, perhaps more canonical, version is to take $\psi(x) = x$ on $\mathcal{S} = [0,\infty)$, with a suitable finite Borel measure μ. Note that every random graph considered in this example may be defined in this way, since we may map $\mathcal{S}$ to $[0,\infty)$ by $x \mapsto \psi(x)$. Alternatively, we may map by $x \mapsto \lambda(x)$ and have $\psi(x) = cx$ with $c > 0$ and $\lambda(x) = x$.)

Random graphs of this type have been studied in several papers; we shall not attempt a complete list, mentioning only several examples. Chung and Lu [34] and Norros and Reittu [82] give results on the existence and size of a giant component. Britton, Deijfen and Martin-Löf [27] use (2.7) with $\kappa(x,y) = \psi(x)\psi(y)$ to define a random graph, and observe that conditioned on the vertex degrees, the resulting graph is uniformly distributed over all graphs with the given degree sequence; they further prove a version of Theorem 3.13 for this case.

Actually, in [34] and [82] the edge probabilities p_{ij} are given by $p_{ij} := w_i w_j / \sum_{i=1}^n w_i$, with w_i deterministic in [34] and random in [82]. Under suitable conditions on the w_i, these examples are also special cases of our general model. For suitable deterministic sequences $(w_i)_1^n$, we use Definition 2.9; we omit the details. For random i.i.d. w_i, as in [82], if we further assume $\mathbb{E}\, w_i = \omega < \infty$, we can, for example, let $\mathcal{S} = [0,\infty)$, $\mu = \mathcal{L}(w_1)$, $x_i = w_i\bigl(\sum_j w_j/n\omega\bigr)^{-1/2}$ and $\psi(x) = \omega^{-1/2}x$. Then $\kappa(x,y) = xy/\omega$, and we have $\lambda(x) = x$ in Theorem 3.13, and thus $\Lambda = w_1$ in Corollary 13.1 and $\Xi \sim \mathrm{Po}(w_1)$. Furthermore, from (16.8) the norm of T_κ is (essentially) the 'second order average degree' $\overline{d} = \sum w_i^2/\sum w_i$. Thus, for example, the result of Chung and Lu [35, 36] that, under certain assumptions, the typical distance between two vertices of the model $G(\mathbf{w})$ studied in [34] is $\log n/\log(\overline{d})$ corresponds to Theorem 3.14. (Chung and Lu also study sequences w_i falling outside the scope of our model.)

Chung and Lu [37] give a result for the 'volume' $\sum_{i\in\mathcal{C}_1} w_i$ of the giant component $\mathcal{C}_1$ of $G(\mathbf{w})$. This result corresponds to Theorem 9.10 with $f(x) = \lambda(x)$; indeed, under certain assumptions on the w_i, it is implied by Theorem 9.10. Unfortunately, the statement of the result in [37] is incomplete, as no conditions on the w_i are given. It is not clear what the right conditions are; certainly some restrictions are needed.

The random graphs $G(n,\kappa)$ obtained from rank 1 kernels should be compared to the random graphs with a given (suitably chosen) degree sequence $(d_i)_1^n$, studied by, for example, Luczak [72], Molloy and Reed [77, 78] and (in the power-law case) Aiello, Chung and Lu [1]. Note that in this model, the probability of an edge between i and j is roughly d_id_j/n, but there are dependencies between the edges. It was shown by Molloy and Reed [77] that, under some conditions, the threshold for the existence of a giant component in this model is $\sum_i d_i(d_i-2) = 0$. This fits well with our result, although we see no strict implication: Theorem 3.13 shows that, for our model, the degree of a random vertex converges in distribution to a random variable Ξ with the mixed Poisson distribution $\int_{\mathcal{S}} \mathrm{Po}(\lambda(x))\, d\mu(x)$. If $X \sim \mathrm{Po}(\lambda)$, then $\mathbb{E}(X(X-2)) = \mathbb{E}(X(X-1)-X) = \lambda^2 - \lambda$, so

$$\mathbb{E}\bigl(\Xi(\Xi-2)\bigr) = \int_{\mathcal{S}} \bigl(\lambda(x)^2 - \lambda(x)\bigr)\, d\mu(x) = \Bigl(\int_{\mathcal{S}} \psi\, d\mu\Bigr)^2 \Bigl(\int_{\mathcal{S}} \psi^2\, d\mu - 1\Bigr),$$

which vanishes when $\int_{\mathcal{S}} \psi^2 = 1$. As $\|T_\kappa\| = \int_{\mathcal{S}} \psi^2$, this is indeed the threshold for the emergence of a giant component in our model. The result of Molloy and Reed [77] that, in the supercritical case, the second largest component has size $O(\log n)$ corresponds to Theorem 3.12(ii); again there is no strict implication, but the kernels κ corresponding to the graphs studied by Molloy and Reed satisfy $\inf \kappa(x,y) > 0$. In a subsequent paper, Molloy and Reed [78] gave further results on the size of the giant component and on the structure of the remainder of the graph, corresponding to our Theorems 3.1 and 12.1.

The following variant of this model has also been studied: the degrees are first chosen according to some distribution, and then the graph is chosen uniformly among all graphs with the resulting degree sequence; see, for example, the results of Van der Hofstad, Hooghiemstra and Van Mieghem [53] and of Fernholz and Ramachandran [50] on distances and diameter, respectively, mentioned in Section 14.

Yet another variant of the rank 1 case of $G^{\mathcal{V}}(n,\kappa)$ was studied rather earlier by Khokhlov and Kolchin [64, 65], who proved results about the number of cycles; see Section 17.

In the rank 1 case, the size of the giant component (if any) of $G^{\mathcal{V}}(n,\kappa)$ can be found rather easily. In order to study the phase transition, let us consider the kernel $c\kappa(x,y) = c\psi(x)\psi(y)$, with $c > 0$ a parameter. By Corollary 3.2, the threshold for c is $c_0 = \|T_\kappa\|^{-1} = \left(\int_{\mathcal{S}} \psi^2\right)^{-1}$. For $c \ge c_0$, let

$$\alpha(c) := c \int \psi \rho_{c\kappa} \, d\mu, \tag{16.9}$$

where, as before, $\rho_{c\kappa}(x) = \rho(c\kappa; x)$ is the survival probability of the branching process $\mathfrak{X}_{c\kappa}(x)$.

We have $T_{c\kappa}\rho_{c\kappa} = cT_\kappa\rho_{c\kappa} = \alpha(c)\psi$. Thus, by Theorem 6.1 and (16.9),

$$\rho_{c\kappa} = \Phi_{c\kappa}(\rho_{c\kappa}) = 1 - e^{-T_{c\kappa}\rho_{c\kappa}} = 1 - e^{-\alpha(c)\psi}. \tag{16.10}$$

(The condition (5.1) holds for every x.) Let

$$\beta(t) := \int_{\mathcal{S}} \Big(1 - e^{-t\psi(x)}\Big) \psi(x) \, d\mu(x), \qquad t \ge 0. \tag{16.11}$$

Then, by (16.9) and (16.10),

$$\alpha(c) = c \int_{\mathcal{S}} \rho_{c\kappa}\psi \, d\mu = c\beta\big(\alpha(c)\big), \tag{16.12}$$

so $c = \alpha(c)/\beta\big(\alpha(c)\big)$, i.e., α is the inverse function to $t \mapsto \gamma(t) := t/\beta(t)$. Since β is explicitly given by (16.11), for any $c > c_0$ this gives (at least in principle) $\alpha(c)$, and hence, by (16.10), the function $\rho_{c\kappa}$. Then $\rho(c\kappa) = \int_{\mathcal{S}} \rho_{c\kappa}(x)\, d\mu(x)$ determines the asymptotic number of vertices in the giant component. Similarly, by Theorem 3.5, the asymptotic number of edges in the giant component is determined by $\zeta(c\kappa)$, which by the definition (3.3) and (16.12) is given by

$$\zeta(c\kappa) = c \int_{\mathcal{S}} \psi\rho_{c\kappa}\, d\mu \int_{\mathcal{S}} \psi \, d\mu - \frac{c}{2}\Big(\int_{\mathcal{S}} \psi\rho_{c\kappa}\, d\mu\Big)^2 = \alpha(c) \int_{\mathcal{S}} \psi \, d\mu - \frac{\alpha(c)^2}{2c}. \tag{16.13}$$

Moreover, the asymptotic value of $\sum_{i\in\mathcal{C}_1} f(x_i)/n$ is given by Theorem 9.10 for suitable functions f.

We now turn to asymptotics as $c \searrow c_0$, in order to study the phase transition more closely. Note that $\rho_{c\kappa} \searrow 0$ a.e. as $c \searrow c_0$ by Theorem 6.4 so, by dominated convergence,

$$\alpha(c)/c \searrow 0 \quad \text{as} \quad c \searrow c_0. \tag{16.14}$$

Further, by (16.10) and dominated convergence,

$$\frac{\rho(c)}{\alpha(c)} = \int_{\mathcal{S}} \frac{\rho_{c\kappa}(x)}{\alpha(c)}\, d\mu(x) \to \int_{\mathcal{S}} \psi(x)\, d\mu(x) > 0 \qquad \text{as } c \searrow c_0. \tag{16.15}$$

Consequently, the behaviour of β at 0 determines, through γ and $\alpha = \gamma^{-1}$, the behaviour of $\rho(c)$ as $c \searrow c_0$. Note that, by (16.11), β is continuous with $\beta(0) = 0$ and

$$\beta'(t) = \int_{\mathcal{S}} e^{-t\psi(x)} \psi^2(x)\, d\mu(x), \qquad t > 0. \tag{16.16}$$

Moreover, from (16.13) we have

$$\zeta(c\kappa) = \alpha(c) \int_{\mathcal{S}} \psi\, d\mu - \frac{\alpha(c)^2}{2c} \sim \alpha(c) \int_{\mathcal{S}} \psi\, d\mu \sim \rho(c)$$

as $c \searrow c_0$, where the first $\sim$ is from (16.14) and the second from (16.15). Hence the asymptotics are the same as for the number of vertices; see Remark 3.19.

Let us consider some concrete examples. Once again, we take $\mathcal{S} = (0, 1]$ with μ Lebesgue measure, and let $\psi(x) = x^{-1/p}$ where $1 < p \le \infty$. We shall use C, C_1, etc. to denote various positive constants that depend on p.

Case 1: $1 < p < 2$. In this case, $\|\psi\|_2 = \infty$, so $c_0 = 0$. As $t \to 0$, by (16.16),

$$\beta'(t) := \int_0^1 e^{-tx^{-1/p}} x^{-2/p}\, dx = p \int_t^\infty e^{-y} t^{-2+p} y^{1-p}\, dy \sim C t^{p-2}, \tag{16.17}$$

noting for the last step that the integral $\int_0^\infty e^{-y} y^{1-p}\, dy$ is convergent. Thus $\beta(t) \sim C_1 t^{p-1}$ and $\gamma(t) = t/\beta(t) \sim C_2 t^{2-p}$. Consequently, using (16.15),

$$\rho(c) \sim C_3 \alpha(c) = C_3 \gamma^{-1}(c) \sim C_4 c^{1/(2-p)} \qquad \text{as } c \to 0.$$

Note that this exponent $1/(2-p)$ may be any real number in $(1, \infty)$.

Case 2: $p = 2$. This is the case (16.4) studied in Subsection 16.2 and [84]. We still have $\|\psi\|_2 = \infty$ and thus $c_0 = 0$. In analogy with (16.17) we now find that $\beta'(t) \sim 2\ln(1/t)$ as $t \to 0$. This yields $\beta(t) \sim 2t\ln(1/t)$ and $\gamma \sim 1/(2\ln(1/t))$ as $t \to 0$, and thus $\alpha(c) = \gamma^{-1}(c) = e^{-(1+o(1))/2c}$ and

$$\rho(c) = e^{-(1+o(1))/2c} \qquad \text{as } c \to 0.$$

More refined estimates can be obtained in the same way, see (16.5) and [84].

Case 3: $2 < p < 3$. For $p > 2$ we have $\int \psi^2\, d\mu < \infty$, and thus $c_0 > 0$, so we have a phase transition. (In fact, $c_0 = 1 - 2/p$.) By (16.16), $\beta'(t)$ is continuous for $t \ge 0$ with $\beta'(0) = \int \psi^2\, d\mu = c_0^{-1}$. Differentiating once more we obtain as $t \to 0$

$$\beta''(t) = -\int_0^1 e^{-tx^{-1/p}} x^{-3/p}\, d\mu(x) = -p \int_t^\infty e^{-y} t^{-3+p} y^{2-p}\, dy \sim -C t^{p-3}$$

and thus $\beta'(t) = c_0^{-1} - (C_1 + o(1))t^{p-2}$ and $\beta(t) = c_0^{-1}t - (C_2 + o(1))t^{p-1}$. Hence $\gamma(t) = t/\beta(t) = c_0 + (C_3 + o(1))t^{p-2}$. Consequently, using (16.15),

$$\rho(c_0 + \varepsilon) \sim C_4\alpha(c_0 + \varepsilon) \sim C_5\varepsilon^{1/(p-2)} \qquad \text{as } \varepsilon \searrow 0.$$

We thus have a phase transition at c_0 with exponent $1/(p-2)$. Note that this exponent may be any real number in $(1, \infty)$. (Taking instead e.g. $\psi(x) = x^{-1/2} \ln^{-1}(e^3/x)$, it is similarly seen that there is a phase transition with infinite exponent.)

Case 4: $p = 3$. Similar calculations show that, as $t \to 0$, $\beta''(t) \sim 3\ln t$, $\beta'(t) = 3 - (3 + o(1))t\ln(1/t)$, $\beta(t) = 3t - (3/2 + o(1))t^2 \ln(1/t)$, and $\gamma(t) = 1/3 + (1/6 + o(1))t \ln 1/t$. Consequently, with $c_0 = 1/3$,

$$\rho(c_0 + \varepsilon) \sim C\alpha(c_0 + \varepsilon) \sim C_1\varepsilon/\ln(1/\varepsilon) \qquad \text{as } \varepsilon \searrow 0,$$

so $\rho'(c_0) = 0$.

Case 5: $3 < p \leq \infty$. In this case, $\int \psi^3 \, d\mu < \infty$ and we find as $t \to 0$, $\beta''(t) \sim -C$, $\beta'(t) = c_0^{-1} - (C + o(1))t$, $\beta(t) = c_0^{-1}t - (C + o(1))t^2/2$, and $\gamma(t) = c_0 + (C_1 + o(1))t$. Consequently, $\rho(c_0 + \varepsilon) \sim C_2\alpha(c_0 + \varepsilon) \sim C_3\varepsilon$, so we have a phase transition with exponent 1. This is similar to Theorem 3.17, although (3.11) is not satisfied (except in the classical case $p = \infty$). Indeed, it can be checked that (3.12) holds, except that the error term may be larger (it is $\Theta(\varepsilon^{p-2})$ for $3 < p < 4$).

More generally, the same argument shows that (3.12) holds for any rank 1 kernel $\psi(x)\psi(y)$ with $\int \psi^3 \, d\mu < \infty$, provided the error term is weakened to $o(\varepsilon)$. (The error term is $O(\varepsilon^2)$ if $\int \psi^4 \, d\mu < \infty$.)

16.5. **Turova's model.** Turova [93, 94, 95, 96] has studied a dynamical random graph $G(t)$, $t \geq 0$, defined as follows, using three parameters $\gamma > 0$, $\lambda > 0$ and $\delta \geq 0$. The graph starts with a single vertex at time $t = 0$. Each existing vertex produces new, initially isolated, vertices according to a Poisson process with intensity γ. As soon as there are at least two vertices, each vertex sends out edges according to another Poisson process with intensity λ; the other endpoint is chosen uniformly among all other existing vertices. (Multiple edges are allowed, but this makes little difference.) Vertices live for ever, but edges die with intensity δ, i.e., the lifetime of an edge has an exponential distribution with mean $1/\delta$. (All these random processes and variables are independent. We use δ for Turova's μ to avoid conflicts with our notation.)

By homogeneity we may assume $\gamma = 1$; the general case follows by replacing λ and δ by λ/γ and δ/γ and changing the time scale.

Our analysis of the random graph $G(t)$ is very similar to that of Söderberg [88]; our theorems enable us to add technical rigour to his calculations. The vertices proliferate according to a Yule process (binary fission process): writing $N(t)$ for the number of vertices at time t, the probability that a new vertex is added in the infinitesimal time interval $[t, t + dt]$ is $N(t)\,dt$. It is well-known (see, e.g., Athreya and Ney [8, Theorems III.7.1–2]) that

$$e^{-t}N(t) \xrightarrow{\text{a.s.}} W \qquad \text{as } t \to \infty \tag{16.18}$$

for a random variable W with $W > 0$ a.s. (In fact, $W \sim \operatorname{Exp}(1)$, but we do not need this.)

We condition on the vertex process, and assume, as we may by (16.18), that

$$e^{-t}N(t) \to w \qquad \text{as } t \to \infty \tag{16.19}$$

for some $w > 0$. We take $\mathcal{S} = [0, \infty)$ and let $x_1, \dots, x_{N(t)}$ be the ages of the particles existing at time t. For any fixed $s > 0$, by (16.19) we have

$$\nu_t[s, \infty) := \frac{1}{N(t)}\#\{i : x_i \geq s\} = \frac{1}{N(t)}N(t - s) \to e^{-s},$$

as $t \to \infty$. This means (see Remark A.3) that $\nu_t \to \mu$, where μ is the measure on $[0, \infty)$ given by $d\mu/dx = e^{-x}$ (the exponential distribution).

If $x_i \leq x_j$, the number of edges at time t between two vertices of ages x_i and x_j has a Poisson distribution with mean

$$e_{ij} := \int_{t-x_i}^{t} e^{-\delta(t-s)}\frac{2\lambda}{N(s) - 1}\,ds = 2\lambda\int_0^{x_i} e^{-\delta s}\frac{ds}{N(t - s) - 1}.$$

Set

$$\kappa_t^*(x, y) := 2\lambda\int_0^{x\wedge y} e^{-\delta s}\frac{N(t)}{N(t - s) - 1}\,ds,$$

and

$$\kappa_t(x, y) = N(t)\bigl(1 - \exp(-\kappa_t^*(x, y)/N(t))\bigr).$$

Thus $e_{ij} = \kappa_t^*(x_i, x_j)/N(t)$, and the probability p_{ij} that there is at least one edge between i and j is given by $p_{ij} = 1 - e^{-e_{ij}} = \kappa_t(x_i, x_j)/N(t)$.

By (16.19), $N(t)/\bigl(N(t-s)-1\bigr) \to e^s$ as $t \to \infty$ for every s, and dominated convergence shows that if $\delta \neq 1$ and $x_t \to x$, $y_t \to y$, then

$$\kappa_t^*(x_t, y_t) \to \kappa_\delta(x, y) := 2\lambda\int_0^{x\wedge y} e^{-\delta s+s}\,ds = \frac{2\lambda}{1 - \delta}\bigl(e^{(1-\delta)(x\wedge y)} - 1\bigr) \tag{16.20}$$

For $\delta = 1$, corresponding to $\delta = \gamma$ in the non-rescaled model, let $\kappa_1(x, y) := 2\lambda(x \wedge y)$. Then $\kappa_t^*(x_t, y_t) \to \kappa_\delta(x, y)$ in this case also.

Theorem 3.1 thus applies to $G(t)$ conditioned on the process $(N(t))_{t\geq0}$, and we find (conditioned on $(N(t))_{t\geq0}$, and thus also unconditionally) that

$$\frac{C_1\bigl(G(t)\bigr)}{N(t)} \xrightarrow{\text{p}} \rho(\kappa_\delta),$$

with κ_δ given by (16.20).

To study $\rho(\kappa_\delta)$ further, and in particular to investigate the threshold as we vary λ keeping $\mu \geq 0$ fixed, we thus have to investigate the integral operator T_{κ_δ} with kernel κ_δ given by (16.20). The change of variables $x \to e^{-x}$

transforms $\mathcal{S}$ and μ to the standard setting $(0,1]$ with Lebesgue measure, and the kernel (16.20) becomes

$$\tilde{\kappa}_\delta(x,y) := \frac{2\lambda}{1-\delta}\Big((x \vee y)^{\delta-1} - 1\Big), \tag{16.21}$$

with $\tilde{\kappa}_1 := 2\lambda \ln(1/(x \vee y))$.

In the case $\delta = 0$, this is the same as (16.7); hence we have the same critical value 1/8 (for λ) as for the CHKNS model and the same $\rho(\kappa)$ giving the size of the giant component; in particular, the phase transition has infinite exponent. (Indeed, with $\delta = 0$ the model is very similar to (Durrett's form of) the CHKNS model discussed in Subsection 16.3; now a geometric number of edges between random vertices is added at each step, rather than a Poisson number.) For $\delta > 0$, the kernel $\tilde{\kappa}_\delta$ is in $L^2((0,1]^2)$, so $T_{\tilde{\kappa}_\delta}$ is compact (see Lemma 5.15) and its norm can be found by finding its eigenvalues. By the discussion in Subsection 16.6 below, this is equivalent to solving (16.23) with the given boundary values. In our case, denoting the eigenvalue by α, this means solving $\alpha G''(x) = -2\lambda x^{\delta-2} G(x)$ with boundary values $G(0) = G'(1) = 0$ (since $\phi(1) = 0$).

The general solution is easily written down as a linear combination of two hypergeometric series, and $G(0) = 0$ yields (up to a constant factor)

$$\begin{aligned} g(x) = G'(x) &= \sum_{n=0}^{\infty} \frac{1}{n!\,\Gamma(n+1/\delta)}\Big(-\frac{2\lambda}{\alpha\delta^2}x^\delta\Big)^n \\ &= \Big(\frac{2\lambda}{\alpha\delta^2}x^\delta\Big)^{-(1/\delta-1)/2} J_{1/\delta-1}\Big(2\Big(\frac{2\lambda}{\alpha\delta^2}x^\delta\Big)^{1/2}\Big), \end{aligned}$$

where J_ν is a Bessel function.

The condition $g(1) = G'(1) = 0$ (which gives the formula in Turova [93, Corollary 4.1] and [94]) thus leads to $J_{1/\delta-1}\Big(\big(\frac{8\lambda}{\alpha\delta^2}\big)^{1/2}\Big) = 0$, so if z_ν is the first positive zero of J_ν, then

$$\|T_{\kappa_\delta}\| = \|T_{\tilde{\kappa}_\delta}\| = \alpha_1 = \frac{8\lambda}{\delta^2 z_{1/\delta-1}^2}.$$

In other words, the critical value of λ is $\lambda_{\mathrm{cr}}(\delta) = \delta^2 z_{1/\delta-1}^2/8$, as given by a related argument by Söderberg [88].

Theorem 3.17 applies only when $\delta > 1/2$, but the eigenfunctions are continuous and bounded for every $\delta > 0$, and we believe that the phase transition has exponent 1, and that (3.12) holds, for every $\delta > 0$.

We can easily find the asymptotics of $\lambda_{\mathrm{cr}}(\delta)$ as $\delta \to 0$ or ∞; see Turova [93]. If $\lambda, \delta \to \infty$ with $\lambda/\delta \to c > 0$, then $\tilde{\kappa}_\delta(x,y) \to 2c$, pointwise and in $L^2((0,1]^2)$, and thus $\|T_{\tilde{\kappa}_\delta} - T_{2c}\| \le \|T_{\tilde{\kappa}_\delta} - T_{2c}\|_{HS} \to 0$. It follows that for large δ, the graph is subcritical if $2c < 1$ and supercritical if $2c > 1$. In other words, $\lambda_{\mathrm{cr}}/\delta \to 1/2$ as $\delta \to \infty$. Similarly, if $\delta \searrow 0$, then $\kappa_\delta \nearrow \kappa_0$ and it follows easily, e.g. by Theorem 6.4, that $\|T_{\kappa_\delta}\| \to \|T_{\kappa_0}\|$, and thus

$\lambda_{\mathrm{cr}}(\delta) \to \lambda_{\mathrm{cr}}(0) = 1/8$. (The contrary assertion in [93] is incorrect; see the erratum.)

16.6. **Functions of** $\max\{x, y\}$**.** In several of the examples above (see Subsections 16.1, 16.3, and 16.5) we have $\mathcal{S} = (0, 1]$, μ is the Lebesgue measure and $\kappa(x, y) = \phi(x \vee y)$ for some function $\phi \geq 0$ on $(0, 1]$. The integral operators T_κ with such kernels have been studied by Maz'ya and Verbitsky [74] and Aleksandrov, Janson, Peller and Rochberg [6]. In particular, these papers prove that T_κ is bounded if and only if $\sup_{x>0} x \int_x^1 \phi(y)^2\, dy < \infty$, and that T_κ is compact if and only if $x \int_x^1 \phi(y)^2\, dy \to 0$ as $x \to 0$.

In the case when ϕ is decreasing (as in the examples above), these criteria simplify to $\phi(x) = O(x^{-1})$ and $\phi(x) = o(x^{-1})$ as $x \to 0$, respectively.

Unfortunately, there is no general formula known for the norm of T_κ. (However, the criteria just given extend to estimates within constant factors; for example, if ϕ is decreasing, then $\sup(x\phi(x)) \leq \|T_\kappa\| \leq 4 \sup(x\phi(x))$.) In the compact case, at least if ϕ has a continuous derivative on $(0, 1]$, the eigenvalues, and thus the norm, can be found by studying a Sturm–Liouville equation. In fact, g is an eigenfunction with eigenvalue λ if

$$\lambda g(x) = \phi(x) \int_0^x g(y)\, dy + \int_x^1 \phi(y) g(y)\, dy. \tag{16.22}$$

If $\lambda \neq 0$, it is easily seen that then $g \in C^1(0, 1]$ and, by differentiating, that (16.22) is equivalent to

$$g(x) = G'(x), \qquad \lambda G''(x) = \phi'(x) G(x), \tag{16.23}$$

with the boundary conditions $G(0) = 0$, $G'(1) = \lambda^{-1}\phi(1)G(1)$; see [6, Section 9] and the example in Subsection 16.5 above.

17. Paths and cycles

Let $P_k(G)$ and $Q_k(G)$ be the numbers of paths and cycles, respectively, of length k (i.e., with k edges) in a graph G. Note that $P_1(G) = e(G)$ is the number of edges, and that $Q_1 = Q_2 = 0$ for simple graphs. (If we allow multiple edges and loops as in Remark 2.5, the results below extend to Q_2 and, under an additional continuity assumption on κ, to Q_1.)

In this section we briefly study the numbers $P_k = P_k(G_n)$ and $Q_k = Q_k(G_n)$, where $G_n = G^{\mathcal{V}}(n, \kappa)$. The results are easily extended to a sequence κ_n as in Definition 2.9 under appropriate conditions, but we leave the details to the reader.

For $k \geq 1$ let

$$\alpha_k(\kappa) := \frac{1}{2} \int_{\mathcal{S}^{k+1}} \kappa(x_0, x_1)\kappa(x_1, x_2) \cdots \kappa(x_{k-1}, x_k)\, d\mu(x_0) \cdots d\mu(x_k),$$

$$\beta_k(\kappa) := \frac{1}{2k} \int_{\mathcal{S}^k} \kappa(x_1, x_2) \cdots \kappa(x_{k-1}, x_k)\kappa(x_k, x_1)\, d\mu(x_1) \cdots d\mu(x_k).$$

Note that $\alpha_k(\kappa) = \frac{1}{2}\langle 1, T_\kappa^k 1\rangle$. Clearly, $\alpha_k(\kappa)$ and $\beta_k(\kappa)$ may be infinite. In this case, the limiting statements in the result below have their natural interpretations.

Theorem 17.1. *Let κ be an a.e. continuous kernel on a (generalized) vertex space $\mathcal{V}$, and let $G_n = G^{\mathcal{V}}(n,\kappa)$.*

(i) *For k fixed,*

$$\liminf_{n\to\infty} \mathbb{E}\, P_k(G_n)/n \ge \alpha_k(\kappa), \qquad k \ge 1,$$
$$\liminf_{n\to\infty} \mathbb{E}\, Q_k(G_n) \ge \beta_k(\kappa), \qquad k \ge 3.$$

(ii) *Suppose further that $\mathcal{V}$ is a vertex space. If κ is bounded, or if $x_1,\dots,x_n$ are i.i.d. random points with the distribution μ, then*

$$\mathbb{E}\, P_k(G_n)/n \to \alpha_k(\kappa), \qquad k \ge 1, \tag{17.1}$$
$$\mathbb{E}\, Q_k(G_n) \to \beta_k(\kappa), \qquad k \ge 3. \tag{17.2}$$

Moreover, whenever (17.1) *holds and $\alpha_k(\kappa)$ is finite,*

$$P_k(G_n)/n \xrightarrow{\mathrm{p}} \alpha_k(\kappa). \tag{17.3}$$

Similarly, whenever (17.2) *holds and the $\beta_k(\kappa)$ are finite,*

$$Q_k(G_n) \xrightarrow{\mathrm{d}} \mathrm{Po}\big(\beta_k(\kappa)\big), \qquad k \ge 3, \tag{17.4}$$

jointly for all $k \ge 3$ with independent limits.

Proof. The argument for parts (i) and (ii) is as in the proof of Lemma 8.1 (a special case), considering first the regular finitary case and then approximating with κ_m^- and κ_m^+ defined in (7.2) and (7.3); we omit the details. The case of i.i.d. x_i with the distribution μ is immediate.

The convergence (17.3) and the asymptotic (joint) Poisson distribution (17.4) of $Q_k(G_n)$ follow easily in the regular finitary case (first conditioning on $\mathbf{x}_n$ as in Remark 8.8 if $\mathcal{V}$ is a generalized vertex space), for example by the method of moments as for $G(n,p)$, cf. [16, 59]. The general cases then follow by appealing to Billingsley [13, Theorem 4.2], noting that if κ_m^- is defined by (7.2), then $\alpha_k(\kappa_m^-) \to \alpha_k(\kappa)$ and $\beta_k(\kappa_m^-) \to \beta_k(\kappa)$ by the monotone convergence theorem, while, from the assumption (17.2) and part (i) (applied to κ_m^-),

$$\begin{aligned}
&\limsup_{n\to\infty} \mathbb{E}\,\big|Q_k\big(G(n,\kappa)\big) - Q_k\big(G(n,\kappa_m^-)\big)\big| \\
&= \lim_{n\to\infty} \mathbb{E}\, Q_k\big(G(n,\kappa)\big) - \liminf_{n\to\infty} \mathbb{E}\, Q_k\big(G(n,\kappa_m^-)\big) \le \beta_k(\kappa) - \beta_k(\kappa_m^-) \to 0
\end{aligned}$$

as $m \to \infty$; an analogous bound holds for P_k/n. □

Part (ii) holds for many generalized vertex spaces too, but not for all. Indeed, Remark 8.2 extends easily to the present situation, although, writing v_n for the number of vertices of $G^{\mathcal{V}}(n,\kappa)$, the variance condition $\mathrm{Var}(v_n/n) \to 0$ should be replaced by a higher moment condition $\mathbb{E}(v_n^\ell/n^\ell) \to \mu(\mathcal{S})^\ell$, with $\ell = k+1$ for (17.1) and $\ell = k$ for (17.2). It is easily seen that (17.1) and

(17.2) hold in the situation in Example 8.5 too. However, these relations may fail for the counterexample in Example 8.6.

Khokhlov and Kolchin [64, 65] studied a model closely related to the rank 1 case of $G^{\mathcal{V}}(n,\kappa)$: each vertex has an activity a_i, and edges are added one by one, with the endpoints of the edge chosen independently, and the probability that a vertex is chosen proportional to its activity. They proved results about the distribution of the numbers of short cycles in this model corresponding to the last part of Theorem 17.1.

Proposition 17.2. (i) *If* $\|T_\kappa\| \le 1$*, then* $\alpha_k(\kappa) \le 1/2$ *for every* k*.*

(ii) *If* $\|T_\kappa\| > 1$*, then* $\alpha_k(\kappa) \to \infty$ *as* $k \to \infty$*.*

Consequently, if κ *is a graphical kernel on a vertex space* $\mathcal{V}$*, then* $G^{\mathcal{V}}(n,\kappa)$ *has a giant component if and only if* $\sup_k \alpha_k(\kappa) = \infty$*.*

Proof. The first statement is immediate, as $\langle 1, T_\kappa^k 1\rangle \le \|T_\kappa\|^k$. For the second statement, we argue as in the proof of Lemma 5.16: there is a bounded kernel $\kappa_N \le \kappa$ with $\|T_{\kappa_N}\| > 1$, and T_{κ_N} has a bounded eigenfunction ψ with eigenvalue $\lambda > 1$. Taking $\|\psi\|_\infty = 1$ we have $\langle 1, T_\kappa^k 1\rangle \ge \langle 1, T_{\kappa_N}^k 1\rangle \ge \langle 1, T_{\kappa_N}^k \psi\rangle = \lambda^k \langle 1, \psi\rangle \to \infty$. The final statement follows by Theorem 3.1(i), which states that $G^{\mathcal{V}}(n,\kappa)$ has a giant component if and only if $\|T_\kappa\| > 1$. □

Similarly, at least when $\iint \kappa^2 < \infty$, we have the following consequence of well-known properties of Hilbert–Schmidt operators (cf. Lemma 5.15).

Proposition 17.3. *If* $\iint \kappa^2 < \infty$*, then* T_κ *is compact and self-adjoint, and if* λ_i *are its (real) eigenvalues (counted with multiplicities), then*

$$\beta_k(\kappa) = \frac{1}{2k}\operatorname{Tr}(T_\kappa^k) = \frac{1}{2k}\sum_i \lambda_i^k < \infty, \qquad k \ge 2.$$

Example 17.4. As in Example 4.6, let $\mathcal{S} = (0,1]$ (regarded as a circle) with μ Lebesgue measure, and $\kappa(x,y) = h(x-y)$, with $h \ge 0$ an even periodic function that is integrable over $(0,1]$. Then T_κ is the convolution operator $f \mapsto h * f$ with eigenvalues $\widehat{h}(j) = \int_0^1 e^{-2\pi i j x} h(x)\,dx$, $j \in \mathbb{Z}$, so $\beta_k = \sum_{-\infty}^{\infty} \widehat{h}(j)^k$.

Considering functions h with small support, we can obtain arbitrarily large β_k with κ bounded and $\iint \kappa = 1$. Alternatively, we can take $\widehat{h}(j) = 1/\ln(2+|j|)$, for example; this defines an integrable function $h > 0$ which is continuous except at 0 [97, Theorems V.(1.5) and V.(1.8)], and thus a kernel κ with $\|T_\kappa\| = \int_0^1 h < \infty$ but $\beta_k(c\kappa) = \infty$ for every $k \ge 2$ and $c > 0$.

Example 17.5. Let $\kappa(x,y) = c/(x \vee y)$ on $\mathcal{S} = (0,1]$ as in Subsection 16.1. Then $\beta_k(\kappa) = \infty$ for every k and every $c > 0$; indeed, if, say, $x_i = i/n$, the expected number of k-cycles with vertices in $(2^{-m-1}n, 2^{-m}n)$ tends to a positive constant independent of $m \ge 0$, and thus $\mathbb{E}\, Q_k \to \infty$.

The same holds for $\kappa(x,y) = c(1/(x \vee y) - 1)$ as in (16.7) and the $\delta = 0$ case of (16.21).

Example 17.6. Let $\kappa(x,y) = \phi(x \vee y)$ with $\mathcal{S} = (0,1]$, as in Subsection 16.6, and assume that $\phi \geq 0$ is non-increasing with $x\phi(x)$ bounded. Assume that $x_1, \dots, x_n$ are i.i.d. and uniformly distributed on $(0,1]$. (It can be checked that the same conclusions hold for $x_i = i/n$.) Then, results of Aleksandrov, Janson, Peller and Rochberg [6, Theorems 4.1 and 4.6] imply that T_κ is a bounded positive operator, and it is compact with eigenvalues λ_i satisfying $\sum_i \lambda_i^k < \infty$ (which means that T_κ belongs to the Schatten–von Neumann class S_k) if and only if $\int_0^1 \bigl(x\phi(x)\bigr)^k / x \, dx < \infty$.

By Theorem 17.1, we have $\mathbb{E}\, Q_k \to \beta_k(\kappa) \leq \infty$ for $k \geq 3$. Proposition 17.3 assumes $\iint \kappa^2 < \infty$, but it can be shown (using truncations of ϕ) that the result extends to the present situation; hence $\beta_k(\kappa) < \infty$ if and only if $\int_0^1 \bigl(x\phi(x)\bigr)^k / x \, dx < \infty$.

Consequently, we may for any given ℓ choose ϕ such that $\beta_k = \infty$ for $3 \leq k \leq \ell$ but $\beta_k < \infty$ for $k > \ell$.

Under suitable conditions, the expected total number of cycles converges to $\sum_{k=3}^\infty \beta_k$; we omit the details. By Proposition 17.3 this sum is given by the following formula.

Corollary 17.7. *If* $\iint \kappa^2 < \infty$, *then*

$$\sum_{k=3}^{\infty} \beta_k(\kappa) = \begin{cases} \sum_i \bigl(-\frac{1}{2}\ln(1-\lambda_i) - \frac{1}{2}\lambda_i - \frac{1}{4}\lambda_i^2\bigr) < \infty, & \textit{if } \|T_\kappa\| < 1, \\ \infty, & \textit{if } \|T_\kappa\| \geq 1. \end{cases}$$

The sum on the right-hand side can be written as $-\frac{1}{2}\ln\bigl(\det_3(I - T_\kappa)\bigr)$, where $\det_3$ is a renormalized Fredholm determinant [87, §9].

Remark 17.8. Turova [94] studies the number of cycles in the random graph discussed in Subsection 16.5, including a formula for $\lim \mathbb{E}\, Q_k = \beta_k$. She conjectures that the threshold for the existence of a giant component is the same as the threshold for $\sum_k \beta_k = \infty$. (This conjecture inspired the present section.) We now see from Corollary 17.7 that this is true in great generality; for example, if κ is bounded, then Theorem 17.1 and Proposition 17.3 imply that the threshold $c_0 = \|T_\kappa\|^{-1}$ in Corollary 3.2 may be written as

$$c_0 = \sup\Bigl\{c : \sum_{k=3}^{\infty} \beta_k(c\kappa) < \infty\Bigr\}. \tag{17.5}$$

Note, however, that exactly at the threshold, i.e., for $G(n,\kappa)$ with $\|T_\kappa\| = 1$, there is no giant component although $\sum_{k=3}^\infty \beta_k(\kappa) = \infty$. Moreover, the relation (17.5) may fail for unbounded κ, see the examples above. In Turova's case (16.21), the relation (17.5) holds for $\delta > 0$ (when T_κ is Hilbert–Schmidt), but not for $\delta = 0$ (when T_κ is not compact), see Example 17.5.

18. Further remarks

Random graphs defined via kernels appear in various other contexts. One natural example is the 'dense' case: let κ be a symmetric function from

$[0,1]^2$ to $[0,1]$ with some suitable 'smoothness' property, and form a graph on $[n]$ by taking the probability p_{ij} of the edge ij to be $\kappa(x_i,x_j)$, where x_i is the type of vertex i (e.g., $x_i = i/n$), and different edges are present independently. Thus, when $\kappa = p$ is constant, one recovers the dense Erdős–Rényi graph $G(n,p)$. The study of this inhomogeneous dense model is as far from the concerns of the present paper as the study of $G(n,1/2)$, say, is from the study of $G(n,c/n)$.

Another case, dense but not so dense, is obtained from our model if we omit the restriction that $\kappa \in L^1(\mathcal{S}\times\mathcal{S},\mu\times\mu)$. One particular example that might have interesting behaviour is $\kappa(x,y) = 1/|x-y|$, with $x_i = i/n$, say, so $p_{ij} = 1/|i-j|$, for $i \neq j$. A similar model (in the rank 1 case of Subsection 16.4) has been studied by Norros and Reittu [82]. Newman and Schulman [79] studied percolation in a closely related infinite random graph: two vertices $i,j \in \mathbb{Z}$ are joined with probability $1-\exp(-\beta|i-j|^{-s})$, where β and s are parameters of the model.

Models with intermediate density (a number of edges that is more than linear but less than quadratic in the number n of vertices) could be obtained by defining the edge probabilities p_{ij} in terms of a kernel κ but with different scaling to that in (2.3). For example, we could take $p_{ij} = \min\{\kappa(x_i,x_j)/n^\alpha, 1\}$, where $0 \le \alpha \le 1$ is a fixed number, or $p_{ij} = n^{-\kappa(x_i,x_j)}$, say. Although these definitions bear a formal resemblance to the one we have used, they lead to very different models. Nevertheless, these models may also repay close attention. In some cases such models might correspond to, or resemble, graphs growing in time by the addition of vertices, with the addition of an increasing number of edges at each step: the case $p_{ij} = j^{-\alpha}$ for $i < j$ is one particular example.

A different connection between graphs and symmetric functions W from $[0,1]^2$ to $[0,1]$ arises in the work of Lovász and Szegedy [69], where the limit of a sequence of dense graphs is defined by considering the number of subgraphs isomorphic to each fixed graph.

Another natural model is the following: take the type space as $[0,1]^2$, say, with the Lebesgue measure, and take the types of the vertices to be independent. (Or, more naturally, generate the vertex types by a Poisson process of intensity n, so the total number of vertices is random.) Join two vertices with a probability $p(d,n)$ that is a function of d, the Euclidean distance between the (types of the) vertices, and n. Since the typical distances are order $n^{-1/2}$, the natural normalization is $p = f(dn^{1/2})$, for example, $p(d,n) = c_1\exp(-c_2 d^2 n)$.

In many ways, a model defined in this way is similar to that considered in this paper: if f decays sufficiently fast, the expected degrees are of order 1, and the degree distribution will be asymptotically Poisson. However, in other ways this graph is very different from the ones we have been studying: in particular, it has many small cycles. Determining the threshold for the emergence of the giant component in this model is likely to be as hard as finding the critical probability for a planar percolation model (indeed, it is essentially the same task), and is thus likely to be impossible except perhaps in very special cases.

Another interesting property of a graph is the behaviour of the contact process on the graph. Suppose that each vertex is either *susceptible*, or *infected*: infected vertices infect their susceptible neighbours with rate λ, and recover with rate 1, returning to the susceptible state. The process starts with a single randomly chosen infected vertex. When the average degree is of order 1, one might expect that there is a critical value λ_c such that for $\lambda < \lambda_c$ constant, the expected number of vertices ever infected is $O(1)$, while for $\lambda > \lambda_c$ constant, with probability bounded away from zero almost all (perhaps even all) vertices in the giant component become infected at some point, and the infection lasts an exponentially long time. This is the case for the 4-regular grid graph on the torus, say; see Liggett [68].

The behaviour of the contact process on the scale-free LCD graph has been studied by Berger, Borgs, Chayes and Saberi [12], who gave detailed results showing in particular that there is no threshold (i.e., $\lambda_c = 0$).

For $G^{\mathcal{V}}(n,\kappa)$, one might expect a positive threshold if and only if $\|T_\kappa\| < \infty$. (Perhaps an additional condition would be needed, such as κ bounded.) In fact, one can make a more detailed prediction based on the contact process on infinite trees: the threshold λ_c should be the same as the threshold for the process to continue forever on an infinite tree generated by the branching process $\mathfrak{X}_\kappa(x)$. Note that it is likely that there are *two* distinct thresholds for the behaviour of the contact process on such trees (this is known only for certain classes of trees, including regular trees; see [83, 67, 92]): a lower threshold λ_1 above which the process has positive probability of never dying out, and an upper threshold λ_2 above which a given vertex has a positive probability of becoming reinfected infinitely often. (In both cases, we start with a single infected vertex.) In the graph, λ_1 should be relevant: if the process survives with drift in the infinite tree, it will eventually revisit a given vertex of $G^{\mathcal{V}}(n,\kappa)$, as the neighbourhoods of a vertex are only locally treelike.

Related results have been proved by Durrett and Jung [45] for a d-dimensional version of the small-world model of Bollobás and Chung: the vertex set is a discrete torus, each vertex is connected to all vertices within a fixed distance, and then all pairs in a random matching of the vertices are added as 'long-range' edges. Durrett and Jung prove separation of λ_1 and λ_2 for an infinite version of this graph. Also, they show that for $\lambda > \lambda_1$, a modified contact process on the finite graph survives for an exponential time; the modification is to allow an infected vertex to infect a randomly chosen other vertex, at an arbitrarily small but positive rate γ. The result is likely to hold with $\gamma = 0$, since the 'long-range' edges already provide sufficient global randomness.

If the definition of the contact process on a graph G is modified so that when a vertex recovers it cannot be reinfected, one might expect that, starting with a single infected vertex v, the set of vertices eventually infected have a 'similar' distribution to the component of $G[p]$ containing v, where $G[p]$ is formed from G by keeping each edge independently with probability $p = \lambda/(1 + \lambda)$. Roughly speaking, for each edge ww' of G, we may declare the edge ww' to be open if whichever of w and w' is first infected will try to infect the other before it recovers, an event with probability $\lambda/(1 + \lambda)$. The set of infected vertices is the component of v in the graph formed by the open edges. Unfortunately, since the probabilities of infection from w to w' and from w to w'' both depend on the random time that w remains infected, the events that different edges are open are not independent. This fact is missed by Newman [80], who states that this modified contact process is equivalent to percolation on G; we should like to thank an anonymous referee for drawing this paper to our attention. Nevertheless, it may still be true that the threshold in this modified contact process is close to the percolation threshold on G, at least under certain conditions.

Appendix A. Probabilistic lemmas

In this appendix we prove three simple technical results concerning sequences of random variables. The first and third are used in the main body of the paper; the second is needed to prove the third. The first concerns random Borel measures.

Let $\mathcal{S}$ be a separable metric space, let $M(\mathcal{S})$ be the space of all finite (positive) Borel measures on $\mathcal{S}$, and let $P(\mathcal{S})$ be the subspace of all Borel probability measures on $\mathcal{S}$. We equip $M(\mathcal{S})$ and $P(\mathcal{S})$ with the usual (weak) topology: $\mu_n \to \mu$ if and only if $\int f\, d\mu_n \to \int f\, d\mu$ for every function f in the space $C_b(\mathcal{S})$ of bounded continuous functions on $\mathcal{S}$. Alternatively, as is well known, $\mu_n \to \mu$ if and only if $\mu_n(A) \to \mu(A)$ for every μ-continuity set A, i.e., every measurable set A with $\mu(\partial A) = 0$.

Remark A.1. The case of probability measures is perhaps better known, and is treated in detail in, for example, Billingsley [13]. Many results extend immediately to $M(\mathcal{S})$, either by inspecting the proof, or because $\mu_n \to \mu$ in $M(\mathcal{S})$ if and only if $\mu_n(\mathcal{S}) \to \mu(\mathcal{S})$ and either $\mu(\mathcal{S}) = 0$ or $\mu_n/\mu_n(\mathcal{S}) \to \mu/\mu(\mathcal{S})$ in $P(\mathcal{S})$.

The spaces $P(\mathcal{S})$ and $M(\mathcal{S})$ are themselves separable metric spaces; for $P(\mathcal{S})$, see [13, Appendix III].

The characterizations above of convergence in $P(\mathcal{S})$ and $M(\mathcal{S})$ extend to random measures and convergence in probability as follows; see Kallenberg [61, Theorem 16.16] for a similar (but stronger) theorem under a stronger hypothesis on $\mathcal{S}$. Note that both (ii) and (iii) are special cases of (iv).

Lemma A.2. *Let $\mathcal{S}$ be a separable metric space, and suppose that ν_n, $n \geq 1$, are random measures in $M(\mathcal{S})$. Then the following assertions are equivalent:*

(i) $\nu_n \xrightarrow{\text{p}} \mu$;
(ii) $\nu_n(A) \xrightarrow{\text{p}} \mu(A)$ *for every* μ*-continuity set* A;
(iii) $\int f\,d\nu_n \xrightarrow{\text{p}} \int f\,d\mu$ *for every bounded continuous function* $f : \mathcal{S} \to \mathbb{R}$.
(iv) $\int f\,d\nu_n \xrightarrow{\text{p}} \int f\,d\mu$ *for every bounded* μ*-a.e. continuous function* $f : \mathcal{S} \to \mathbb{R}$.

Proof. (i) $\Longrightarrow$ (ii) If A is a μ-continuity set then $\nu \mapsto \nu(A)$ defines a measurable function $M(\mathcal{S}) \to \mathbb{R}$ which is continuous at μ.

(ii) $\Longrightarrow$ (iv) We may suppose that $0 \le f \le 1$; the general case follows by linearity. Let N be the μ-null set consisting of points at which f is discontinuous, and let $A_t := f^{-1}(t,\infty) = \{x : f(x) > t\}$. If $x \in \overline{A_t} \setminus A_t$ and $x \notin N$, then $f(x) = t$ by continuity. Thus $\partial A_t \subseteq N \cup f^{-1}\{t\}$. Hence, the sets $\partial A_t \setminus N$ are disjoint, and $\mu(\partial A_t) = 0$ except for at most countably many t. When $\mu(\partial A_t) = 0$, we have $\nu_n(A_t) \xrightarrow{\text{p}} \mu(A_t)$ by (ii), so $\mathbb{E}\,|\nu_n(A_t) - \mu(A_t)| \to 0$ by dominated convergence. By dominated convergence again,

$$\begin{aligned}
\mathbb{E}\left|\int_{\mathcal{S}} f\,d\nu_n - \int_{\mathcal{S}} f\,d\mu\right| &= \mathbb{E}\left|\int_0^1 \nu_n(A_t)\,dt - \int_0^1 \mu(A_t)\,dt\right| \\
&\le \mathbb{E}\int_0^1 \bigl|\nu_n(A_t) - \mu(A_t)\bigr|\,dt = \int_0^1 \mathbb{E}\bigl|\nu_n(A_t) - \mu(A_t)\bigr|\,dt \\
&\to 0.
\end{aligned}$$

(iv) $\Longrightarrow$ (iii) Trivial.

(iii) $\Longrightarrow$ (i) The topological space $M(\mathcal{S})$ is metrizable, but the topology is also defined by the functionals $\mu \mapsto \int f\,d\mu$, $f \in C_b(\mathcal{S})$. Hence, if U is a neighbourhood of μ in $M(\mathcal{S})$, there is a finite set of functions $f_1, \dots, f_N \in C_b(\mathcal{S})$ and $\varepsilon > 0$ such that if $\left|\int f_i\,d\nu - \int f_i\,d\mu\right| < \varepsilon$, $i = 1, \dots, N$, then $\nu \in U$. Consequently,

$$\mathbb{P}(\nu_n \notin U) \le \sum_{i=1}^N \mathbb{P}\left(\left|\int f_i\,d\nu - \int f_i\,d\mu\right| \ge \varepsilon\right) \to 0.$$

$\square$

Remark A.3. To verify condition (ii), it often suffices to consider A in a suitably selected family of subsets. For example, it is well-known that on $\mathbb{R}$, it suffices to consider μ-continuity sets of the form $(-\infty, x]$ and, for $M(\mathcal{S})$, $\mathbb{R}$ itself; see [13, Section 3].

Recall that if X_n is a sequence of random variables and a_n a sequence of positive real numbers, then $X_n = O(a_n)$ whp means that there is a constant C such that $|X_n| \le Ca_n$ whp. Our final technical result (Lemma A.5 below) is simple, but perhaps a little surprising: we shall show that under suitable assumptions, if $X_n = O(a_n)$ holds conditionally (after conditioning on the sequences $\mathbf{x}_n$ in our model), then it holds unconditionally, i.e., that the implicit constant may be assumed to be deterministic. We start with a preparatory lemma.

Lemma A.4. *Let $\mathcal{A}_1, \mathcal{A}_2, \dots,$ be non-empty families of random variables such that for any sequence $X_n \in \mathcal{A}_n$ we have $X_n = O(a_n)$ whp. Then there is a constant C such that $\sup_{X \in \mathcal{A}_n} \mathbb{P}(|X| > Ca_n) \to 0$ as $n \to \infty$. In other words, the implicit constant in $X_n = O(a_n)$ whp can be chosen uniformly for $X_n \in \mathcal{A}_n$.*

Proof. Replacing X_n by X_n/a_n, we may assume that $a_n = 1$. Suppose the conclusion fails. Then, for every m there is an $\varepsilon_m > 0$ such that there are arbitrarily large n for which there is an $X_n \in \mathcal{A}_n$ with $\mathbb{P}(|X_n| > m) > \varepsilon_m$. Let (m_k) be a sequence of integers where each positive integer appears infinitely many times. Select inductively an increasing sequence (n_k) and $X_{n_k} \in \mathcal{A}_{n_k}$ such that $\mathbb{P}(|X_{n_k}| > m_k) > \varepsilon_{m_k}$. For $n \notin \{n_k\}$, choose X_n from $\mathcal{A}_n$ arbitrarily.

For any positive integer m, there are infinitely many k such that $m_k = m$, and thus infinitely many n such that $\mathbb{P}(|X_n| > m) > \varepsilon_m$. Hence (X_n) is not $O(1)$ whp, which contradicts our assumption. □

The next lemma can be stated in terms of families $\mu_n(y)$ of probability distributions (on $\mathbb{R}$) and mixtures $\mathbb{E}\,\mu_n(Y_n)$ of them, but we prefer a statement in terms of random variables $X_n(y) \sim \mu_n(y)$; we consider a sequence of families $X_n(y)$ of random variables defined for y in a subset $\mathcal{M}_n$ of a certain space $\mathcal{M}$ as this is convenient when we apply the lemma to $G^{\mathcal{V}}(n, \kappa)$.

Lemma A.5. *Let $\mathcal{M}$ be a metric space, and, for each $n \geq 1$, let $X_n(y)$, $y \in \mathcal{M}_n \subseteq \mathcal{M}$, be a (measurable) family of real-valued random variables. Let $y_0 \in \mathcal{M}$, and suppose that for every sequence (y_n) with $y_n \in \mathcal{M}_n$ and $y_n \to y_0$ we have $X_n(y_n) = O(a_n)$ whp. Then, if (Y_n) is a sequence of $\mathcal{M}_n$-valued random variables, independent of all $X_n(y)$, with $Y_n \xrightarrow{\mathrm{p}} y_0$, we have $X_n(Y_n) = O(a_n)$ whp.*

Proof. Since $Y_n \xrightarrow{\mathrm{p}} y_0$, there is a sequence $\delta_n \to 0$ such that $\mathbb{P}(d(Y_n, y_0) < \delta_n) \to 1$. Set $U_n = \{y \in \mathcal{M}_n : d(y, y_0) < \delta_n\}$, so $\mathbb{P}(Y_n \in U_n) \to 1$. Note that $y_n \in U_n$ implies $y_n \to y_0$, and thus $X_n(y_n) = O(a_n)$ whp. Let $\mathcal{A}_n = \{X_n(y) : y \in U_n\}$. By Lemma A.4, there exists C such that $\varepsilon_n := \sup_{y \in U_n} \mathbb{P}(|X_n(y)| > Ca_n) \to 0$. Finally, $\mathbb{P}(|X_n(Y_n)| > Ca_n) \leq \varepsilon_n + \mathbb{P}(Y_n \notin U_n) \to 0$. □

Acknowledgement. Part of this research was done during visits of S.J. to Cambridge, supported by the Swedish Royal Academy of Sciences and the London Mathematical Society. The paper was revised during a visit of all three authors to the Institute for Mathematical Sciences, National University of Singapore.

References

[1] W. Aiello, F. Chung & L. Lu, A random graph model for power law graphs, *Experiment. Math.* **10** (2001), 53–66.

[2] M. Aizenman, J.T. Chayes, L. Chayes & C. Newman, Discontinuity of the magnetization in one-dimensional $1/|x-y|^2$ Ising and Potts models, *J. Statist. Phys.* **50** (1988), 1–20.
[3] R. Albert & A.-L. Barabási, Statistical mechanics of complex networks, *Rev. Mod. Phys.* **74** (2002), 47–97.
[4] R. Albert, H. Jeong & A.-L. Barabási, Diameter of the world-wide web, *Nature* **401** (1999), 130–131.
[5] R. Albert, H. Jeong & A.-L. Barabási, Error and attack tolerance of complex networks, *Nature* **406** (2000), 378–382.
[6] 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, 61–150.
[7] N. Alon, A note on network reliability, *Discrete probability and algorithms (Minneapolis, MN, 1993)*, IMA Vol. Math. Appl., 72, Springer, New York, 1995, 11–14.
[8] K.B. Athreya & P.E. Ney, *Branching Processes*, Springer, Berlin, 1972.
[9] A.-L. Barabási & R. Albert, Emergence of scaling in random networks, *Science* **286** (1999), 509–512.
[10] A.-L. Barabási, R. Albert & H. Jeong, Scale-free characteristics of random networks: the topology of the world-wide web, *Physica A* **281** (2000), 69–77.
[11] A.D. Barbour, L. Holst & S. Janson, *Poisson Approximation*, Oxford Univ. Press, Oxford, 1992.
[12] N. Berger, C. Borgs, J.T. Chayes & A. Saberi, On the spread of viruses on the Internet, *Proceedings of the 16th ACM-SIAM Symposium on Discrete Algorithms (SODA)*, SIAM, 2005, 301–310.
[13] P. Billingsley, *Convergence of Probability Measures*, John Wiley & Sons, New York, 1968.
[14] B. Bollobás, A probabilistic proof of an asymptotic formula for the number of labelled regular graphs, *European J. Combinatorics* **1** (1980), 311–316.
[15] B. Bollobás, The evolution of random graphs, *Trans. Amer. Math. Soc.* **286** (1984), 257–274.
[16] B. Bollobás, *Random Graphs*, 2nd ed., Cambridge Univ. Press, Cambridge, 2001.
[17] B. Bollobás, *Linear Analysis*, 2nd ed., Cambridge Univ. Press, Cambridge, 1999.
[18] B. Bollobás & W. Fernandez de la Vega, The diameter of random regular graphs, *Combinatorica* **2** (1982), 125–134.
[19] B. Bollobás, S. Janson & O. Riordan, The phase transition in a uniformly grown random graph has infinite order, *Random Struct. Alg.* **26** (2005), 1–36.
[20] B. Bollobás, S. Janson & O. Riordan, Long-range percolation in $\mathbb{R}^d$, preprint available from `http://arXiv.org/math.PR/0508430` .
[21] B. Bollobás & O. Riordan, Mathematical results on scale-free random graphs, *Handbook of Graphs and Networks*, eds. S. Bornholdt & H.G. Schuster, Wiley-VCH, Weinheim (2003), 1–34.
[22] B. Bollobás & O. Riordan, Robustness and vulnerability of scale-free random graphs, *Internet Mathematics* **1** (2003), 1–35.
[23] B. Bollobás & O. Riordan, The diameter of a scale-free random graph, *Combinatorica* **24** (2004), 5–34.
[24] B. Bollobás & O. Riordan, Coupling scale-free and classical random graphs, *Internet Mathematics* **1** (2004), 215–225.
[25] B. Bollobás & O. Riordan, Slow emergence of the giant component in the growing m-out graph, *Random Struct. Alg.* **27** (2005), 1–24.
[26] B. Bollobás, O. Riordan, J. Spencer & G.Tusnády, The degree sequence of a scale-free random graph process, *Random Struct. Alg.* **18** (2001), 279–290.

[27] T. Britton, M. Deijfen & A. Martin-Löf, Generating simple random graphs with prescribed degree distribution, *J. Statist. Phys.*, to appear.
[28] T. Britton & A. Martin-Löf, private communication (2005).
[29] P.G. Buckley & D. Osthus, Popularity based random graph models leading to a scale-free degree sequence, *Discrete Math* **282** (2004), 53–68.
[30] 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.
[31] D. S. Callaway, M. E. J. Newman, S. H. Strogatz & D. J. Watts, Network robustness and fragility: percolation on random graphs, *Phys. Rev. Lett.* **85** (2000), 5468–5471.
[32] H. Cartan, *Calcul Différentiel*, Hermann, Paris, 1967.
[33] F. Chung & L. Lu, The diameter of sparse random graphs, *Adv. in Appl. Math.* **26** (2001), 257–279.
[34] F. Chung & L. Lu, Connected components in random graphs with given expected degree sequences, *Ann. Comb.* **6** (2002), 125–145.
[35] F. Chung & L. Lu, The average distances in random graphs with given expected degrees, *Proc. Natl. Acad. Sci. USA* **99** (2002), 15879–15882 (electronic).
[36] F. Chung & L. Lu, The average distance in a random graph with given expected degrees, *Internet Math.* **1** (2003), 91–113.
[37] F. Chung & L. Lu, The volume of the giant component of a random graph with given expected degrees, *SIAM J. Discrete Math.* **20** (2006), 395–411.
[38] R. Cohen, K. Erez, D. ben-Avraham & S. Havlin, Breakdown of the Internet under intentional attack, *Phys. Rev. Lett.* **86** (2001), 3682–3685.
[39] R. Cohen & S. Havlin, Ultra small world in scale-free networks, *Phys. Rev. Lett.* **90** (2003), 058701.
[40] C. Cooper & A. Frieze, A general model of web graphs, *Random Struct. Alg.* **22** (2003), 311–335.
[41] L. Devroye, C. McDiarmid, & B. Reed, Giant components for two expanding graph processes, *Mathematics and computer science II (Versailles, 2002)*, eds. B. Chauvin, P. Flajolet, D. Gardy & A. Mokkadem, Birkhäuser, Basel, 2002, 161–173.
[42] S.N. Dorogovtsev & J.F.F. Mendes, Evolution of networks, *Adv. Phys.* **51** (2002), 1079–1187.
[43] S.N. Dorogovtsev, J.F.F. Mendes & A.N. Samukhin, Anomalous percolation properties of growing networks, *Phys. Rev. E* **64** (2001), 066110.
[44] 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), 95–104, `http://dmtcs.loria.fr/proceedings/` .
[45] R. Durrett & P. Jung, Two phase transitions for the contact process on small worlds, preprint available from `http://arXiv.org/math.PR/0501481` .
[46] R. Durrett & H. Kesten, The critical parameter for connectedness of some random graphs, *A Tribute to Paul Erdős*, eds. A. Baker, B. Bollobás & A. Hajnal, Cambridge Univ. Press, Cambridge, 1990, 161–176.
[47] P. Erdős & A. Rényi, On the evolution of random graphs, *Magyar Tud. Akad. Mat. Kutató Int. Közl.* **5** (1960), 17–61.
[48] P. Erdős & A. Rényi, On the evolution of random graphs, *Bull. Inst. Internat. Statist.* **38** (1961), 343–347.
[49] M. Faloutsos, P. Faloutsos & C. Faloutsos, On power-law relationships of the internet topology, SIGCOMM 1999, *Comput. Commun. Rev.* **29** (1999), 251.
[50] D. Fernholz & V. Ramachandran, The diameter of sparse random graphs, preprint available from `http://www.cs.utexas.edu/~vlr/pubs.html` .
[51] E.N. Gilbert, Random graphs, *Ann. Math. Statist.* **30** (1959), 1141–1144.
[52] M. Hervé, *Analyticity in Infinite Dimensional Spaces*, Walter de Gruyter, Berlin, 1989.

[53] R. van der Hofstad, G. Hooghiemstra & P. Van Mieghem, Distances in random graphs with finite variance degrees, *Random Struct. Alg.* **27** (2005), 76–123.

[54] R. van der Hofstad, G. Hooghiemstra & D. Znamenski, Distances in random graphs with infinite mean degrees, preprint available from `http://arXiv.org/math.PR/0407091` .

[55] R. van der Hofstad, G. Hooghiemstra & D. Znamenski, Random graphs with arbitrary i.i.d. degrees, preprint available from `http://arXiv.org/math.PR/0502580` .

[56] R. van der Hofstad, G. Hooghiemstra & D. Znamenski, Distances in random graphs with finite mean and infinite variance degrees, preprint available from `http://arXiv.org/math.PR/0502581` .

[57] S. Janson, On a random graph related to quantum theory, preprint available from `http://arXiv.org/math.PR/0606454` .

[58] S. Janson, D. Knuth, T. Łuczak & B. Pittel, The birth of the giant component, with an introduction by the editors, *Random Struct. Alg.* **4** (1994), 231–358.

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

[60] S. Kalikow & B. Weiss, When are random graphs connected? *Israel J. Math.* **62** (1988), 257–268.

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

[62] R.M. Karp, The transitive closure of a random digraph, *Random Struct. Alg.* **1** (1990), 73–93.

[63] D.G. Kendall, Deterministic and stochastic epidemics in closed populations, *Proceedings of the Third Berkeley Symposium on Mathematical Statistics and Probability, 1954–1955, vol. IV*, University of California Press, Berkeley and Los Angeles, 1956 149–165.

[64] V.I. Khokhlov & V.F. Kolchin, On the number of cycles in a random non-equiprobable graph (Russian), *Diskretnaya Matematika* **2** (1990), 137–145. English transl. *Discrete Math. Appl.* **2** (1992), 109–118.

[65] V.I. Khokhlov & V.F. Kolchin, On the structure of a random graph with nonuniform distribution, *New trends in probability and statistics, Vol. 1 (Bakuriani, 1990)*, VSP, Utrecht, 1991, 445–456.

[66] J.M. Kleinberg, R. Kumar, P. Raghavan, S. Rajagopalan & A. Tomkins, The web as a graph: Measurements, models, and methods, *Proceedings of COCOON '99*, Lecture Notes in Computer Science **1627** (1999), 1–17.

[67] T.M. Liggett, Multiple transition points for the contact process on the binary tree, *Annals of Probability* **24** (1996), 1675–1710.

[68] T.M. Liggett, *Stochastic interacting systems: Contact, voter and exclusion processes,* Springer-Verlag, New York-Berlin, 1999.

[69] L. Lovász & B. Szegedy, Limits of dense graph sequences, Microsoft Technical Report TR-2004-79, `http://research.microsoft.com/users/lovasz/limits.pdf` .

[70] M.J. Luczak & C. McDiarmid, Bisecting sparse random graphs, *Random Struct. Alg.* **18** (2001), 31–38.

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

[72] T. Łuczak, Sparse random graphs with a given degree sequence, *Random graphs, Vol. 2 (Poznań, 1989)*, Wiley, New York, 1992, 165–182.

[73] T. Łuczak, Random trees and random graphs, *Random Struct. Alg.* **13** (1998), 485–500.

[74] V.G. Maz'ya & I.E. Verbitsky, The Schrödinger operator on the energy space: boundedness and compactness criteria, *Acta Math.* **188** (2002), 263–302.

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

[76] C. Mode, *Multitype Branching Processes: Theory and Applications*, Elsevier, New York, 1971.
[77] M. Molloy, & B. Reed, A critical point for random graphs with a given degree sequence, *Random Struct. Alg.* **6** (1995), 161–179.
[78] M. Molloy & B. Reed, The size of the giant component of a random graph with a given degree sequence, *Combin. Probab. Comput.* **7** (1998), 295–305.
[79] C.M. Newman & L.S. Schulman, One-dimensional $1/|j-i|^s$ percolation models: the existence of a transition for $s \le 2$, *Comm. Math. Phys.* **104** (1986), 547–571.
[80] M.E.J. Newman, Spread of epidemic disease on networks, *Phys. Rev. E* **66** (2002), 016128.
[81] M.E.J. Newman, S.H. Strogatz & D.J. Watts, Random graphs with arbitrary degree distribution and their applications, *Phys. Rev. E* **64** (2001), 026118.
[82] I. Norros & H. Reittu, On a conditionally Poissonian graph process, *Adv. Appl. Probab.* **38** (2006), 59–75.
[83] R. Pemantle, The contact process on trees, *Ann. Probab.* **20** (1992), 2089–2116.
[84] O. Riordan, The small giant component in scale-free random graphs, *Combin. Probab. Comput.* **14** (2005), 897–938.
[85] O. Riordan, The k-core and branching processes, preprint available from `http://arXiv.org/math.CO/0511093` .
[86] L.A. Shepp, Connectedness of certain random graphs, *Israel J. Math.* **67** (1989), 23–33.
[87] B. Simon, *Trace Ideals and their Applications*, LMS Lecture Notes Series **35**, Cambridge Univ. Press, Cambridge, 1979.
[88] B. Söderberg, General formalism for inhomogeneous random graphs, *Phys. Rev. E* **66** (2002), 066121.
[89] B. Söderberg, Random graphs with hidden color, *Phys. Rev. E* **68** (2003), 015102(R).
[90] B. Söderberg, Properties of random graphs with hidden color, *Phys. Rev. E* **68** (2003), 026107.
[91] B. Söderberg, Random graph models with hidden color, *Acta Physica Polonica B* **34** (2003), 5085–5102.
[92] A.M. Stacey, The existence of an intermediate phase for the contact process on trees, *Ann. Probab.* **24** (1996), 1711–1726.
[93] T.S. Turova, Dynamical random graphs with memory, *Phys. Rev. E* **65** (2002), 066102. Erratum: *Phys. Rev. E* **70** (2004), 059902(E).
[94] T.S. Turova, Long paths and cycles in dynamical graphs, *J. Statist. Phys.* **110** (2003), 385–417.
[95] T.S. Turova, Continuity of the percolation threshold in randomly grown graphs, preprint, 2005.
[96] T.S. Turova, Phase transitions in dynamical random graphs, in preparation.
[97] A. Zygmund, *Trigonometric Series*, 2nd ed., Cambridge Univ. Press, Cambridge, 1959.