Sparse random graphs with clustering

Béla Bollobás[*†‡] Svante Janson[§] Oliver Riordan[¶]

July 13, 2008

Abstract

In 2007 we introduced a general model of sparse random graphs with independence between the edges. The aim of this paper is to present an extension of this model in which the edges are far from independent, and to prove several results about this extension. The basic idea is to construct the random graph by adding not only edges but also other small graphs. In other words, we first construct an inhomogeneous random hypergraph with independent hyperedges, and then replace each hyperedge by a (perhaps complete) graph. Although flexible enough to produce graphs with significant dependence between edges, this model is nonetheless mathematically tractable. Indeed, we find the critical point where a giant component emerges in full generality, in terms of the norm of a certain integral operator, and relate the size of the giant component to the survival probability of a certain (non-Poisson) multi-type branching process. While our main focus is the phase transition, we also study the degree distribution and the numbers of small subgraphs. We illustrate the model with a simple special case that produces graphs with power-law degree sequences with a wide range of degree exponents and clustering coefficients.

1 Introduction and results

In [10], a very general model for sparse random graphs was introduced, corresponding to an inhomogeneous version of $G(n, c/n)$, and many properties of this model were determined, in particular, the critical point of the phase transition where the giant component emerges. Part of the motivation was to unify many of the new random graph models introduced as approximations to real-world networks. Indeed, the model of [10] includes many of these models as exact special cases, as well as the 'mean-field' simplified versions of many of the more

*Department of Mathematical Sciences, University of Memphis, Memphis TN 38152, USA
†Trinity College, Cambridge CB2 1TQ, UK
‡Research supported in part by NSF grants CCR-0225610, DMS-0505550 and W911NF-06-1-0076
§Department of Mathematics, Uppsala University, PO Box 480, SE-751 06 Uppsala, Sweden
¶Mathematical Institute, University of Oxford, 24–29 St Giles', Oxford OX1 3LB, UK

complicated models. (The original forms are frequently too complex for rigorous mathematical analysis, so such mean-field versions are often studied instead.) Unfortunately, there are many models with key features that are not captured by their mean-field versions, and hence not by the model of [10]. The main problem is that many real-world networks exhibit *clustering*: for example, while there are n vertices and only $5n$ edges, there may be $10n$ triangles, say. In contrast, the model of [10], like $G(n, c/n)$, produces graphs that contain essentially no triangles or short cycles.

Most models introduced to approximate particular real-world networks turn out to be mathematically intractable, due to the dependence between edges. Nevertheless, many such models have been studied; as this is not our main focus, let us just list a few examples of early work in this field. One of the starting points in this area was the (homogeneous) 'small-world' model of Watts and Strogatz [35]. Another was the observation of power-law degree sequences in various networks by Faloutsos, Faloutsos and Faloutsos [26], among others. Of the new inhomogeneous models, perhaps the most studied is the 'growth with preferential attachment' model introduced in an imprecise form by Barabási and Albert [5], later made precise as the 'LCD model' by Bollobás and Riordan [15]. Another is the 'copying' model of Kumar, Raghavan, Rajagopalan, Sivakumar, Tomkins and Upfal [31], generalized by Cooper and Frieze [22], among others. For (early) surveys of work in this field see, for example, Barabási and Albert [1], Dorogovtsev and Mendes [24], or Bollobás and Riordan [13].

Roughly speaking, any sparse model with clustering must include significant dependence, so one might expect it to be impossible to construct a general model of this type that is still mathematically tractable. However, it turns out that one can do this. The model that we shall define is essentially a generalization of that in [10], although we shall handle certain technicalities in a different way here.

1.1 The model

Let us set the scene for our model. By a *type space* we simply mean a probability space $(\mathcal{S}, \mu)$. Often, we shall take $\mathcal{S} = [0, 1]$ or $(0, 1]$ with μ Lebesgue measure. Sometimes we consider $\mathcal{S}$ finite. Of course, any model with $\mathcal{S}$ finite can be realized as a model with type space $[0, 1]$, but sometimes the notation will be simpler with $\mathcal{S}$ finite. More generally, as shown in [28], every instance of the random graph model we are going to describe can be realized as an equivalent model with type space $[0, 1]$. Hence, when it comes to proofs, we lose no generality by taking $\mathcal{S} = [0, 1]$, but we usually prefer allowing an arbitrary type space, which is more flexible for applications. For example, as with the model in [10], type spaces such as $\mathcal{S} = [0, 1]^2$ are likely to be useful for geometric applications, as in [11].

Let $\mathcal{F}$ consist of one representative of each isomorphism class of finite connected graphs, chosen so that if $F \in \mathcal{F}$ has r vertices then $V(F) = [r] = \{1, 2, \ldots, r\}$. Given $F \in \mathcal{F}$ with r vertices, let κ_F be a measurable function from $\mathcal{S}^r$ to $[0, \infty)$; we call κ_F the *kernel* corresponding to F. A sequence $\underset{\sim}{\kappa} = (\kappa_F)_{F \in \mathcal{F}}$

is a *kernel family*. In our results we shall impose an additional integrability condition on $\underset{\sim}{\kappa}$, but this is not needed to define the model.

Let $\underset{\sim}{\kappa}$ be a kernel family and n an integer; we shall define a random graph $G(n,\underset{\sim}{\kappa})$ with vertex set $[n] = \{1,2,\dots,n\}$. First let $x_1, x_2, \dots, x_n \in \mathcal{S}$ be iid (independent and identically distributed) with the distribution μ. Given $\mathbf{x} = (x_1,\dots,x_n)$, construct $G(n,\underset{\sim}{\kappa})$ as follows, starting with the empty graph. For each r and each $F \in \mathcal{F}$ with $|F| = r$, and for every r-tuple of distinct vertices $(v_1,\dots,v_r) \in [n]^r$, add a copy of F on the vertices $v_1,\dots,v_r$ (with vertex i of F mapped to v_i) with probability

$$p = p(v_1,\dots,v_r;F) = \frac{\kappa_F(x_{v_1},\dots,x_{v_r})}{n^{r-1}}, \tag{1}$$

all these choices being independent. If $p > 1$, then we simply add a copy with probability 1. We shall often call the added copies of the various F that together form $G(n,\underset{\sim}{\kappa})$ *atoms* since, in our construction of $G(n,\underset{\sim}{\kappa})$, they may be viewed as indivisible building blocks. Sometimes we refer to them as *small graphs*, although there is in general no bound on their sizes. Usually we think of $G(n,\underset{\sim}{\kappa})$ as a simple graph, in which case we simply replace any multiple edges by single edges. Typically there will be very few multiple edges, so this makes little difference.

Note that we assume that the atoms of $G(n,\underset{\sim}{\kappa})$ are connected. The extension to the case where some atoms may be disconnected is discussed in Subsection 4.1.

The reason for dividing by n^{r-1} in (1) is that we wish to consider sparse graphs; indeed, our main interest is the case when $G(n,\underset{\sim}{\kappa})$ has $O(n)$ edges. As it turns out, we can be slightly more general; however, when κ_F is integrable (which we shall always assume), the expected number of added copies of each graph F is $O(n)$.

Remark 1. There are several plausible choices for the normalization in (1). The one we have chosen means that if $\kappa_F = c$ is constant, then (asymptotically) there are on average cn copies of F in total, and each vertex is on average in rc copies of F. An alternative is to divide the expression in (1) by r; then (asymptotically) each vertex would on average be in c copies of F. Another alternative, natural when adding cliques only but less so in the general case, would be to divide by $r!$; this is equivalent to considering unordered sets of r vertices instead of ordered r-tuples. When there is only one kernel, corresponding to adding edges, this would correspond to the normalization used in [10], and in particular to that of the classical model $G(n,c/n)$; the normalization we use here differs from this by a factor of 2. Yet another normalization would be to divide by $\mathrm{aut}(F)$, the number of automorphisms of F; this is equivalent to considering the distinct copies of F in K_n, which is natural but leads to extra factors $\mathrm{aut}(F)$ in many formulae, and we do not find that the advantages outweigh the disadvantages.

As in [10], there are several minor variants of $G(n,\underset{\sim}{\kappa})$; perhaps the most important is the *Poisson multi-graph* version of $G(n,\underset{\sim}{\kappa})$. In this variant, for each F and each r-tuple, we add a Poisson $\mathrm{Po}(p)$ number of copies of F with this vertex set, where p is given by (1), and we keep multiple edges.

Alternatively, we could add a Poisson number of copies and delete multiple edges, which is the same as adding one copy with probability $1 - e^{-p}$ and no copy otherwise. More generally, we could add one copy of F with probability $p + o(p)$, and two or more copies with probability $o(p)$. As long as the error terms are uniform over graphs F and r-tuples $(v_1, \ldots, v_r)$, all our results will apply in this greater generality. Since this will follow by simple sandwiching arguments (after reducing to the 'bounded' case; see Definition 15), we shall consider whichever form of the model is most convenient; usually this turns out to be the Poisson multi-graph form.

Remark 2. Under certain mild conditions, the results of [29] imply a strong form of asymptotic equivalence between the various versions of the model. For example, if we add copies of F with probability $p + O(p^2)$, where the implied constant is uniform over F and $(v_1, \ldots, v_r)$, and

$$\mathbb{E} \sum_F \sum_{v_1, \ldots, v_{|F|}} p(v_1, \ldots, v_{|F|}; F)^3 = o(1), \tag{2}$$

then the resulting model is equivalent to that with probability p, in that the two random graphs can be coupled to agree whp; this is a straightforward modification of [29, Corollary 2.13(i)]. Extending the argument in [29, Example 3.2], it can be shown that (2) holds if

$$\sum_{F \in \mathcal{F}} \int_{\mathcal{S}^{|F|}} \kappa_F^{|F|/(|F|-1)} < \infty.$$

This certainly holds for the bounded kernel families (see Definition 15) that we consider in most of our proofs, although (2) is easy to verify directly for such kernel families.

In the special case where all κ_F are zero apart from κ_{K_2}, the kernel corresponding to an edge, we recover (essentially) a special case of the model of [10]; we call this the *edge-only* case, since we add only edges, not larger graphs. We write κ_2 for κ_{K_2}. Note that in the edge-only case, given $\mathbf{x}$, two vertices i and j are joined with probability

$$\frac{\kappa_2(x_i, x_j) + \kappa_2(x_j, x_i)}{n} + O\left(\frac{(\kappa_2(x_i, x_j) + \kappa_2(x_j, x_i))^2}{n^2} \right).$$

The correction term will never matter, so we may as well replace κ_2 by its symmetrized version. In fact, we shall always assume that κ_F is invariant under the action of the automorphism group $\mathrm{Aut}(F)$ of the graph F. In the Poisson version, or if we add copies of graphs F with probability $1 - e^{-p}$, there are no correction terms at all; in the edge-only case, given $\mathbf{x}$, vertices i and j are joined with probability $1 - \exp\big((\kappa_2(x_i, x_j) + \kappa_2(x_j, x_i))/n\big)$, and in general we obtain exactly the same random graph if we symmetrize each κ_F w.r.t. $\mathrm{Aut}(F)$.

For any kernel family $\underset{\sim}{\kappa}$, let κ_{e} be the corresponding *edge kernel*, defined by

$$\kappa_{\mathrm{e}}(x,y)=\sum_F\sum_{ij\in E(F)}\int_{\mathcal{S}^{V(F)\setminus\{i,j\}}}\kappa_F(x_1,\dots,x_{i-1},x,x_{i+1},\dots,x_{j-1},y,x_{j+1},\dots,x_{|F|}), \tag{3}$$

where the second sum runs over all $2e(F)$ ordered pairs (i,j) with $ij\in E(F)$, and we integrate over all variables apart from x and y. Note that the sum need not always converge; since every term is positive this causes no problems: we simply allow $\kappa_{\mathrm{e}}(x,y)=\infty$ for some x,y. Given x_i and x_j, the probability that i and j are joined in $G(n,\underset{\sim}{\kappa})$ is at most $\kappa_{\mathrm{e}}(x_i,x_j)/n$, and this upper bound is typically quite sharp. For example, if $\underset{\sim}{\kappa}$ is bounded in the sense of Definition 15 below, then the probability is $\kappa_{\mathrm{e}}(x_i,x_j)/n+O(1/n^2)$. In other words, κ_{e} captures the edge probabilities in $G(n,\underset{\sim}{\kappa})$, but not the correlations.

Before proceeding to deeper properties, let us note that the expected number of added copies of F is $(1+O(n^{-1}))n\int_{\mathcal{S}^{|F|}}\kappa_F$. Unsurprisingly, the actual number turns out to be concentrated about this mean. Let

$$\xi(\underset{\sim}{\kappa})=\sum_{F\in\mathcal{F}}e(F)\int_{\mathcal{S}^{|F|}}\kappa_F=\frac{1}{2}\int_{\mathcal{S}^2}\kappa_{\mathrm{e}}\le\infty$$

be the *asymptotic edge density* of $\underset{\sim}{\kappa}$. Since every copy of F contributes $e(F)$ edges, the following theorem is almost obvious, provided we can ignore overlapping edges. A formal proof will be given in Section 6. (A similar result for the total number of atoms is given in Lemma 40.)

Theorem 3. *As $n\to\infty$, $e(G(n,\underset{\sim}{\kappa}))/n$ converges in probability to the asymptotic edge density $\xi(\underset{\sim}{\kappa})$. In other words, if $\xi(\underset{\sim}{\kappa})<\infty$ then $e(G(n,\underset{\sim}{\kappa}))=\xi(\underset{\sim}{\kappa})n+o_{\mathrm{p}}(n)$, and if $\xi(\underset{\sim}{\kappa})=\infty$ then, for every constant C, we have $e(G(n,\underset{\sim}{\kappa}))>Cn$ whp. Moreover, $\mathbb{E}e(G(n,\underset{\sim}{\kappa}))/n\to\xi(\underset{\sim}{\kappa})\le\infty$*

As in [10], our main focus will be the emergence of the giant component. When studying the component structure of $G(n,\underset{\sim}{\kappa})$, the model can be simplified somewhat: when we add a connected graph F to a graph G, the effect on the component structure is simply to unite all components of G that meet the vertex set of F, so only the vertex set of F matters, not the graph structure. We say that $\underset{\sim}{\kappa}$ is a *clique kernel family* if the only non-zero kernels are those corresponding to complete graphs; the corresponding random graph model $G(n,\underset{\sim}{\kappa})$ is a *clique model*. For questions concerning component structure, it suffices to study clique models. For clique kernels we write κ_r for κ_{K_r}; as above, we always assume that κ_r is symmetric, here meaning invariant under all permutations of the coordinates of $\mathcal{S}^r$. Given a general kernel family $\underset{\sim}{\kappa}$, the corresponding (symmetrized) clique kernel family is given by $\bar{\underset{\sim}{\kappa}}=(\kappa_r)_{r\ge2}$ with

$$\kappa_r(x_1,\dots,x_r)=\sum_{F\in\mathcal{F}:|F|=r}\frac{1}{r!}\sum_{\pi\in\mathfrak{S}_r}\kappa_F(x_{\pi(1)},\dots,x_{\pi(r)}), \tag{4}$$

where $\mathfrak{S}_r$ denotes the symmetric group of permutations of $[r]$. (This is consistent with our notation $\kappa_2=\kappa_{K_2}$.) In the Poisson version, with or without merging of parallel edges, the probability of adding some connected graph F on a given set of r vertices is exactly the same in $G(n, \underset{\sim}{\kappa})$ and $G(n, \bar{\underset{\sim}{\kappa}})$, so there is a natural coupling of these random graphs in which they have exactly the same components. In the non-Poisson version, the probabilities are not quite the same, but close enough for our results to transfer from one to the other. Thus, when considering the size of the giant component in $G(n, \underset{\sim}{\kappa})$, we may always replace $\underset{\sim}{\kappa}$ by the corresponding clique kernel family.

It is often convenient to think of a clique model as a random hypergraph, with the cliques as the hyperedges; for this reason we call a clique kernel family a *hyperkernel*. Note that each unordered set of r vertices corresponds to $r!$ r-tuples, so the probability that we add a K_r on a given set of r vertices is $r!\kappa_r(x_{v_1}, \dots, x_{v_r})/n^{r-1}$. (More precisely, this is the expected number of K_rs added with this vertex set.)

1.2 A branching process

Associated to each hyperkernel $\underset{\sim}{\kappa} = (\kappa_r)_{r\geq 2}$, there is a branching process $\mathfrak{X}_{\underset{\sim}{\kappa}}$ with type space $\mathcal{S}$, defined as follows. We start with generation 0 consisting of a single particle whose type is chosen randomly from $\mathcal{S}$ according to the distribution μ. A particle P of type x gives rise to children in the next generation according to a two step process: first, for each $r \geq 2$, construct a Poisson process Z_r on $\mathcal{S}^{r-1}$ with intensity

$$r\kappa_r(x, x_2, \dots, x_r)\, d\mu(x_2) \cdots d\mu(x_r). \tag{5}$$

We call the points of $Z = \bigcup_{r\geq 2} Z_r$ the *child cliques* of P. There are $r-1$ children of P for each child clique $(x_2, \dots, x_r) \in \mathcal{S}^{r-1}$, one each of types $x_2, \dots, x_r$. Thus the children of P form a multiset on $\mathcal{S}$, with a certain compound Poisson distribution we have just described. As usual, the children of different particles are independent of each other, and of the history.

Considering the relationship to the graph $G(n, \underset{\sim}{\kappa})$, the initial factor r in (5) arises because a particular vertex v may be any one of the r vertices in an r-tuple $(v_1, \dots, v_r)$ on which we add a K_r.

We also consider the branching processes $\mathfrak{X}_{\underset{\sim}{\kappa}}(x)$, $x \in \mathcal{S}$, defined exactly as $\mathfrak{X}_{\underset{\sim}{\kappa}}$, except that we start with a single particle of the given type x.

1.3 Two integral operators

We shall consider two integral operators naturally associated to $\mathfrak{X}_{\underset{\sim}{\kappa}}$. Given any (measurable) $f : \mathcal{S} \to [0, 1]$, define $S_{\underset{\sim}{\kappa}}(f)$ by

$$\begin{aligned} &S_{\underset{\sim}{\kappa}}(f)(x) \\ &= \sum_{r=2}^{\infty} \int_{\mathcal{S}^{r-1}} r\kappa_r(x, x_2, x_3, \dots, x_r) \left(1 - \prod_{i=2}^{r}(1 - f(x_i))\right) d\mu(x_2) \cdots d\mu(x_r), \end{aligned} \tag{6}$$

and let

$$\Phi_{\underset{\sim}{\kappa}}(f)(x) = 1 - e^{-S_{\underset{\sim}{\kappa}}(f)(x)}.$$

(The factors r in (6) and in the definition of $\mathfrak{X}_{\underset{\sim}{\kappa}}$ are unfortunate consequences of our choice of normalization.)

Let P be a particle of $\mathfrak{X}_{\underset{\sim}{\kappa}}$ in generation t with type x, and suppose that each particle in generation $t+1$ of type y has some property $\mathcal{Q}$ with probability $f(y)$, independently of the other particles. Given a child clique $(x_2, \ldots, x_r)$ of P, the bracket in the definition of $S_{\underset{\sim}{\kappa}}$ expresses the probability that one or more of the $r-1$ corresponding child particles has property $\mathcal{Q}$. Hence $S_{\underset{\sim}{\kappa}}(f)(x)$ is the expected number of child cliques containing a particle with property $\mathcal{Q}$, and, from the Poisson distribution of the child cliques, $\Phi_{\underset{\sim}{\kappa}}(f)(x)$ is the probability that there is at least one such clique, i.e., the probability that at least one child of P has property $\mathcal{Q}$.

Let $\rho(\underset{\sim}{\kappa})$ denote the survival probability of the branching process $\mathfrak{X}_{\underset{\sim}{\kappa}}$, and $\rho_{\underset{\sim}{\kappa}}(x)$ the survival probability of $\mathfrak{X}_{\underset{\sim}{\kappa}}(x)$. From the comments above we see that the function $\rho_{\underset{\sim}{\kappa}}$ satisfies

$$\rho_{\underset{\sim}{\kappa}} = \Phi_{\underset{\sim}{\kappa}}(\rho_{\underset{\sim}{\kappa}}).$$

Using simple standard arguments as in [10], for example, it is easy to check that $\rho_{\underset{\sim}{\kappa}}$ is given by the maximum solution to this equation, i.e., the pointwise supremum of all solutions $f : \mathcal{S} \to [0, 1]$ to

$$f = 1 - e^{-S_{\underset{\sim}{\kappa}}(f)}; \tag{7}$$

see Lemma 7 below. From the definitions of $\mathfrak{X}_{\underset{\sim}{\kappa}}$ and $\mathfrak{X}_{\underset{\sim}{\kappa}}(x)$, it is immediate that

$$\rho(\underset{\sim}{\kappa}) = \int_{\mathcal{S}} \rho_{\underset{\sim}{\kappa}}(x)\, d\mu(x).$$

In our analysis, we shall also consider the linear operator $T_{\kappa_{\mathrm{e}}}$ defined by

$$T_{\kappa_{\mathrm{e}}}(f)(x) = \int_{\mathcal{S}} \kappa_{\mathrm{e}}(x, y) f(y)\, d\mu(y). \tag{8}$$

For a hyperkernel $\underset{\sim}{\kappa}$,

$$\kappa_{\mathrm{e}}(x, y) = \sum_{r \geq 2} r(r-1) \int_{\mathcal{S}^{r-2}} \kappa_r(x, y, x_3, x_4, \ldots, x_r)\, d\mu(x_3) \cdots d\mu(x_r), \tag{9}$$

from which it is easy to check that $T_{\kappa_{\mathrm{e}}}$ is the linearized form of $S_{\underset{\sim}{\kappa}}$: more precisely, $T_{\kappa_{\mathrm{e}}}$ is obtained by replacing $1 - \prod_{i=2}^{r}(1 - f(x_i))$ by $\sum_{i=2}^{r} f(x_i)$ in the definition (6) of $S_{\underset{\sim}{\kappa}}$.

Let us note two simple consequences of this fact. For any sequence $(y_i)_i$ in $[0, 1]$ we have $1 - \prod_i (1 - y_i) \leq \sum_i y_i$, so

$$0 \leq S_{\underset{\sim}{\kappa}}(f) \leq T_{\kappa_{\mathrm{e}}}(f) \tag{10}$$

for any $f : \mathcal{S} \to [0,1]$. Also, $1 - \prod_i (1 - y_i) > 0$ if and only if $\sum_i y_i > 0$. Since the integral of a non-negative function is positive if and only if the function is positive on a set of positive measure, it follows that for any $f : \mathcal{S} \to [0,1]$ we have

$$S_{\underset{\sim}{\kappa}}(f)(x) > 0 \iff T_{\kappa_{\mathrm{e}}}(f)(x) > 0. \tag{11}$$

In the edge-only case, when only κ_2 is non-zero, $\kappa_{\mathrm{e}} = 2\kappa_2$ and $T_{\kappa_{\mathrm{e}}} = S_{\underset{\sim}{\kappa}}$. When translating results from [10], it is sometimes $T_{\kappa_{\mathrm{e}}}$ and sometimes $S_{\underset{\sim}{\kappa}}$ that plays the role of the linear operator T_κ appearing there.

1.4 Main results

In most of our results we shall need to impose some sort of integrability condition on our kernel family; the exact condition depends on the context.

Definition 4. (i) A kernel family $\underset{\sim}{\kappa} = (\kappa_F)_{F \in \mathcal{F}}$ is *integrable* if

$$\int \underset{\sim}{\kappa} = \sum_{F \in \mathcal{F}} |F| \int_{\mathcal{S}^{|F|}} \kappa_F < \infty. \tag{12}$$

This means that the expected number of atoms containing a given vertex is bounded.

(ii) A kernel family $\underset{\sim}{\kappa} = (\kappa_F)_{F \in \mathcal{F}}$ is *edge integrable* if

$$\sum_{F \in \mathcal{F}} e(F) \int_{\mathcal{S}^{|F|}} \kappa_F < \infty;$$

equivalently, $\xi(\underset{\sim}{\kappa}) < \infty$ or $\int_{\mathcal{S}^2} \kappa_{\mathrm{e}} < \infty$. This means that the expected number of edges in $G(n, \underset{\sim}{\kappa})$ is $O(n)$, see Theorem 3, and thus the expected degree of a given vertex is bounded.

Note that a hyperkernel (κ_r) is integrable if and only if $\sum_{r \ge 2} r \int_{\mathcal{S}^r} \kappa_r < \infty$, and edge integrable if and only if $\sum_{r \ge 2} r^2 \int_{\mathcal{S}^r} \kappa_r < \infty$.

Since we only consider connected atoms F, it is clear that

$$\text{edge integrable} \implies \text{integrable}.$$

Our main result is that, if $\underset{\sim}{\kappa}$ is an integrable kernel family satisfying a certain extra assumption, then the normalized size of the giant component in $G(n, \underset{\sim}{\kappa})$ is simply $\rho(\underset{\sim}{\kappa})$. The extra assumption is essentially that the graph does not split into two pieces. As in [10], we say that a symmetric kernel $\kappa_{\mathrm{e}} : \mathcal{S}^2 \to [0, \infty)$ is *reducible* if

$$\exists A \subset \mathcal{S} \text{ with } 0 < \mu(A) < 1 \text{ such that } \kappa_{\mathrm{e}} = 0 \text{ a.e. on } A \times (\mathcal{S} \setminus A);$$

otherwise κ_{e} is *irreducible.* Thus κ_{e} is irreducible if

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

A kernel family $(\kappa_F)_{F\in\mathcal{F}}$ or $(\kappa_r)_{r\geq 2}$ is *irreducible* if the corresponding edge kernel κ_{e} is irreducible. It is easy to check that a kernel family $(\kappa_F)_{F\in\mathcal{F}}$ is irreducible if and only if for every $A \subset \mathcal{S}$ with $0 < \mu(A) < 1$ there exists an $F \in \mathcal{F}$ such that, with $r = |F|$, if $x_1, \ldots, x_r$ are chosen iid at random in $\mathcal{S}$ with distribution μ, then there is a positive probability that $\{x_i\} \cap A \neq \emptyset$, $\{x_i\} \cap (\mathcal{S} \setminus A) \neq \emptyset$ and $\kappa_F(x_1, \ldots, x_r) > 0$. Informally, $(\kappa_F)_{F\in\mathcal{F}}$ is irreducible if, whenever we partition the type space into two non-trivial parts, edges between vertices with types in the two parts are possible.

Note that a kernel family $\underset{\sim}{\kappa}'$ and the corresponding hyperkernel $\underset{\sim}{\kappa}$ do *not* have the same edge kernel: replacing each atom by a clique in general adds edges, so $\kappa'_{\mathrm{e}} \leq \kappa_{\mathrm{e}}$ with strict inequality possible. If κ'_{e} is irreducible, then so is κ_{e}; using the characterization of irreducibility above, it is easy to check that the reverse implication also holds.

We are now ready to state our main result; we write C_i for the number of vertices in the ith largest component of a graph G.

Theorem 5. *Let $\underset{\sim}{\kappa}' = (\kappa'_F)_{F\in\mathcal{F}}$ be an irreducible, integrable kernel family, and let $\underset{\sim}{\kappa} = (\kappa_r)_{r\geq 2}$ be the corresponding hyperkernel, given by* (4)*. Then*

$$C_1\big(G(n, \underset{\sim}{\kappa}')\big) = \rho(\underset{\sim}{\kappa})n + o_{\mathrm{p}}(n),$$

and $C_2\big(G(n, \underset{\sim}{\kappa}')\big) = o_{\mathrm{p}}(n)$.

The reducible case reduces to the irreducible one; see Remark 26.

Of course, for Theorem 5 to be useful we would like to know something about the survival probability $\rho(\underset{\sim}{\kappa})$; as noted earlier, $\rho(\underset{\sim}{\kappa})$ can be calculated from $\rho_{\underset{\sim}{\kappa}}$, which is in turn the largest solution to a certain equation (7). Of course, the main thing we would like to know is when $\rho(\underset{\sim}{\kappa})$ is positive; as in [10], it turns out that the answer depends on the L^2-norm $\|T_{\kappa_{\mathrm{e}}}\| \leq \infty$ of the operator $T_{\kappa_{\mathrm{e}}}$ defined by (8). (Since this operator is symmetric, its L^2-norm is the same as its spectral radius. In other contexts, it may be better to work with the latter.)

Theorem 6. *Let $\underset{\sim}{\kappa}$ be an integrable hyperkernel. Then $\rho(\underset{\sim}{\kappa}) > 0$ if and only if $\|T_{\kappa_{\mathrm{e}}}\| > 1$. Furthermore, if $\underset{\sim}{\kappa}$ is irreducible and $\|T_{\kappa_{\mathrm{e}}}\| > 1$, then $\rho_{\underset{\sim}{\kappa}}(x)$ is the unique non-zero solution to the functional equation* (7)*, and $\rho_{\underset{\sim}{\kappa}}(x) > 0$ holds for a.e. x.*

In general, $\|T_{\kappa_{\mathrm{e}}}\|$ may be rather hard to calculate; a non-trivial example where we can calculate the norm easily is given in Subsection 7.2. Let us give a trivial example here: suppose that each κ_r is constant, say $\kappa_r = c_r$. Then $\kappa_{\mathrm{e}}(x, y) = \sum_r r(r-1)c_r = 2\xi(\kappa)$ for all x and y, so

$$\|T_{\kappa_{\mathrm{e}}}\| = 2\xi(\kappa). \tag{13}$$

This is perhaps surprising: it tells us that for such uniform hyperkernels, the critical point where a giant component emerges is determined only by the total number of edges added; it does not matter what size cliques they lie in. This is not true for arbitrary kernel families: we must first replace each atom by a clique.

Note that for any hyperkernel,

$$\|T_{\kappa_{\mathrm{e}}}\| \geq \langle 1, T_{\kappa_{\mathrm{e}}} 1 \rangle = \iint \kappa_{\mathrm{e}} = 2\xi(\kappa),$$

with equality if and only if 1 is an eigenfunction, i.e., if the asymptotic expected degrees $\lambda(x) = \int_{\mathcal{S}} \kappa_{\mathrm{e}}(x, y)\, d\mu(y)$ are the same (ignoring sets of measure 0); c.f. [10, Proposition 3.4].

Before turning to the proofs, let us remark on the relationship of the results above to the main results of [10]. In the edge-only case, the present results are almost (see below) special cases of those [10]. The set-up here is much simpler, as we choose to insist that the vertex types $x_1, \ldots, x_n$ are iid. This avoids many of the complications arising in [10]. In one way, the present set-up is, even in the edge-only case, more general than that considered in [10]: with the types iid, there is no need to restrict the kernels other than to assume integrability (in [10] we needed them continuous a.e.), and one does not need to impose the 'graphicality' assumption of [10]. Thus the edge-only case here actually complements the results in [10]. We could form a common generalization, but we shall not do this in detail; we believe that it is just a question of combining the various technicalities here and in [10], and that no interesting new difficulties arise. Of course, these technicalities are rather beside the point of the present paper; our interest is the extension from kernels to hyperkernels. This turns out not to be quite as straightforward as one might perhaps expect. The problem is that the correlation between edges forces us to deal with a non-linear operator, namely $S_{\underset{\sim}{\kappa}}$.

The rest of the paper is organized as follows. In the next section we prove the results about the non-Poisson branching process $\mathfrak{X}_{\underset{\sim}{\kappa}}$ that we shall need later, the most important of which is Theorem 6. In Section 3 we consider the local coupling between the graph and the branching process, showing in particular that the 'right' number of vertices are in components of any fixed size. In Section 4 we complete the proof of Theorem 5, showing that whp there is at most one 'large' component, which is then a 'giant' component of the right size. In Sections 5 and 6 we consider simpler properties of $G(n, \underset{\sim}{\kappa})$, namely the asymptotic degree distribution and the number of subgraphs isomorphic to a given graph. Our results in Section 6 include Theorem 3 as a simple special case. In Section 7 we illustrate the flexibility of the model by carrying out explicit calculations for a special case, giving graphs with power-law degree sequences with a range of exponents and a range of clustering and mixing coefficients; see Section 7 for the definitions of these coefficients. Finally, in Section 8 we discuss connections between our model and various notions of graph limit, and state two open questions.

2 Analysis of the branching process

In this section, which is the heart of the paper, we forget about graphs, and study the (compound Poisson) branching process $\mathfrak{X}_{\underset{\sim}{\kappa}}$. One might expect the arguments of [10] to carry over *mutatis mutandis* to the present context, but in the branching process analysis this is very far from the truth; this applies especially to the proof of Theorem 10.

Throughout the section we work with an integrable hyperkernel $\underset{\sim}{\kappa} = (\kappa_r)_{r\geq 2}$, i.e., we assume that $\int \underset{\sim}{\kappa} = \sum_r r \int \kappa_r < \infty$. Our main aim in this section is to prove Theorem 6.

For $x \in \mathcal{S}$ let

$$\lambda(x) = (S_{\underset{\sim}{\kappa}}(1))(x) = \sum_{r=2}^{\infty} \int_{\mathcal{S}^{r-1}} r\kappa_r(x, x_2, x_3, \ldots, x_r)\, d\mu(x_2) \cdots d\mu(x_r),$$

so $\lambda(x)$ is the expected number of child cliques of a particle of type x. We have

$$\int_{\mathcal{S}} \lambda(x)\, d\mu(x) = \sum_{r\geq 2} \int_{\mathcal{S}^r} r\kappa_r = \int \underset{\sim}{\kappa},$$

which is finite by our assumption (12). It follows that $\lambda(x) < \infty$ holds almost everywhere. Changing each kernel κ_r on a set of measure zero, we may assume that $\lambda(x)$ is finite for all x. (Such a change is irrelevant for the branching process and for the graph.) From now on, we thus assume that $\lambda(x) < \infty$ holds for all x, for any hyperkernel $\underset{\sim}{\kappa}$ we consider.

Since a Poisson random variable with finite mean is always finite, any particle in $\mathfrak{X}_{\underset{\sim}{\kappa}}$ has a finite number of child cliques, and hence a finite number of children, even though the expected number of children may perhaps be infinite. Hence, the event that the branching process dies out (i.e., that some generation is empty) coincides with the event that it is finite.

Using this fact, we have the following, standard result. Recall that $\rho_{\underset{\sim}{\kappa}}(x)$ denotes the survival probability of the branching process $\mathfrak{X}_{\underset{\sim}{\kappa}}(x)$ that starts with a single particle of type x, and $\rho_{\underset{\sim}{\kappa}}$ the function $x \mapsto \rho_{\underset{\sim}{\kappa}}(x)$.

Lemma 7. *The function $\rho_{\underset{\sim}{\kappa}}$ satisfies the functional equation* (7)*. Furthermore, if $f : \mathcal{S} \to [0, 1]$ is any other solution to* (7)*, then $0 \leq f(x) \leq \rho_{\underset{\sim}{\kappa}}(x) < 1$ holds for every x.*

Proof. Let $\rho_t(x)$ be the probability that $\mathfrak{X}_{\underset{\sim}{\kappa}}(x)$ survives for at least t generations, so ρ_0 is identically 1. Then $\rho_{t+1} = \Phi_{\underset{\sim}{\kappa}}(\rho_t)$. The result follows from the monotonicity of $\Phi_{\underset{\sim}{\kappa}}$ and the fact that $\rho_t(x) \searrow \rho_{\underset{\sim}{\kappa}}(x)$, noting that $\Phi_{\underset{\sim}{\kappa}}(1)(x) = 1 - e^{-\lambda(x)} < 1$ for the strict inequality. $\square$

We next turn to the uniqueness of the non-zero solution (if any) to (7). The key ingredient in establishing this is the following simple inequality concerning the non-linear operator $S_{\underset{\sim}{\kappa}}$.

Lemma 8. *Let* $\underset{\sim}{\kappa}$ *be an integrable hyperkernel, and let* f *and* g *be measurable functions on* $\mathcal{S}$ *with* $0 \leq f \leq g \leq 1$*. Then*

$$\int_{\mathcal{S}} f S_{\underset{\sim}{\kappa}} g \leq \int_{\mathcal{S}} g S_{\underset{\sim}{\kappa}} f.$$

Proof. We may write $S_{\underset{\sim}{\kappa}}$ as $\sum_{r \geq 2} S_r$, where S_r is the nonlinear operator corresponding to the single kernel κ_r, so $S_r(f)$ is defined by the summand in (6). Thus it suffices to prove that

$$\int_{\mathcal{S}} f S_r g \leq \int_{\mathcal{S}} g S_r f. \tag{14}$$

We shall in fact show that for any (distinct) $x_1, \ldots, x_r \in \mathcal{S}$ we have

$$\sum_{\pi \in \mathfrak{S}_r} f(x_{\pi(1)}) \left(1 - \prod_{i=2}^{r} \big(1 - g(x_{\pi(i)})\big) \right) \leq \sum_{\pi \in \mathfrak{S}_r} g(x_{\pi(1)}) \left(1 - \prod_{i=2}^{r} \big(1 - f(x_{\pi(i)})\big) \right) \tag{15}$$

Since κ_r is symmetric, (14) follows. (In fact, (14) can be true in general only if (15) always holds, considering the symmetrization of a delta function.) Now (15) can be viewed as an inequality in $2r$ variables $f(x_1), \ldots, f(x_r), g(x_1), \ldots, g(x_r)$. This inequality is linear in each variable. Furthermore, it is linear in each pair $(f(x_i),\, g(x_i))$. In proving (15) for any $0 \leq f \leq g \leq 1$, we may thus assume that for each i one of three possibilities holds: $0 = f(x_i) = g(x_i)$, $f(x_i) = g(x_i) = 1$, or $f(x_i) = 0$ and $g(x_i) = 1$. In other words, we may assume that f and g are $\{0,1\}$-valued.

Suppose then for a contradiction that (15) fails for some $\{0,1\}$-valued f and g with $f \leq g$. Then there must be some permutation π such that

$$f(x_{\pi(1)}) \left(1 - \prod_{i=2}^{r} \big(1 - g(x_{\pi(i)})\big) \right) > g(x_{\pi(1)}) \left(1 - \prod_{i=2}^{r} \big(1 - f(x_{\pi(i)})\big) \right), \tag{16}$$

which we may take without loss of generality to be the identity permutation. Since both sides of (16) are $\{0,1\}$-valued, the left must be 1 and the right 0. Since the left is 1, we have $f(x_1) = 1$, so, using $f \leq g$, $g(x_1) = 1$. But now for the right hand side of (16) to be 0 the final product in (16) must be 1, so $f(x_i) = 0$ for $i = 2, \ldots, r$, i.e., f takes the value 1 only once. Of course, g must take the value 1 at least twice, otherwise we have equality. But now the left hand side of (15) is exactly $(r-1)!$, coming from terms with $\pi(1) = 1$ and hence $f(x_{\pi(1)}) = 1$. The right hand side is at least $(r-1)!$, from any π mapping 1 to some $j \neq 1$ with $g(x_j) = 1$. Hence (15) holds after all, giving a contradiction and completing the proof. $\square$

If $\underset{\sim}{\kappa}$ is reducible, then (7) may in general have several non-zero solutions. To prove uniqueness in the irreducible case we need to know what irreducibility tells us about $S_{\underset{\sim}{\kappa}}$.

Lemma 9. *If there exists a measurable* $f : \mathcal{S} \to [0,1]$ *with* $0 < \mu\{f > 0\} < 1$ *and* $\{S_{\underset{\sim}{\kappa}} f > 0\} \subseteq \{f > 0\}$*, then* $\underset{\sim}{\kappa}$ *is reducible.*

Proof. Let $A = \{f > 0\}$, so by assumption $S_{\underset{\sim}{\kappa}} f = 0$ on $A^{\mathrm{c}} = \mathcal{S} \setminus A$. From (11) we have $\{T_{\kappa_{\mathrm{e}}} f = 0\} = \{S_{\underset{\sim}{\kappa}} f = 0\}$, so $T_{\kappa_{\mathrm{e}}} f = 0$ on A^{c}. From the definition of $T_{\kappa_{\mathrm{e}}}$ it follows that $\kappa_{\mathrm{e}} = 0$ a.e. on $A^{\mathrm{c}} \times A$, so κ_{e} is reducible. But this is what it means for $\underset{\sim}{\kappa}$ to be reducible. $\square$

In fact, taking f to be a suitable characteristic function, one can check that the converse of Lemma 9 also holds.

Using Lemmas 8 and 9 it is easy to deduce uniqueness of any non-zero solution to (7).

Theorem 10. *Let* $\underset{\sim}{\kappa}$ *be an irreducible, integrable hyperkernel, and let* f *and* g *be solutions to* (7) *with* $0 \le f(x) \le g(x) \le 1$ *for every* x*. Then either* $f = 0$ *or* $f = g$*. In particular, the only solutions to* (7) *are* $\rho_{\underset{\sim}{\kappa}}$ *and the zero function, which may or may not coincide.*

Proof. We may suppose that f is not 0 a.e.; otherwise, $f = \Phi_{\underset{\sim}{\kappa}}(f)$ would be identically zero. Since $\underset{\sim}{\kappa}$ is irreducible and f solves (7), which implies $\{f = 0\} = \{S_{\underset{\sim}{\kappa}} f = 0\}$, it follows from Lemma 9 that $f > 0$ a.e. Note that, since $\underset{\sim}{\kappa}$ is integrable, $S_{\underset{\sim}{\kappa}}(g)(x) \le S_{\underset{\sim}{\kappa}}(1)(x) = \lambda(x) < \infty$ for a.e. x, and thus $g = \Phi_{\underset{\sim}{\kappa}}(g) < 1$ a.e.

Since f and g solve (7), we have $S_{\underset{\sim}{\kappa}}(f)(x) = -\log(1 - f(x))$ and $S_{\underset{\sim}{\kappa}}(g)(x) = -\log(1 - g(x))$. Hence,

$$
\begin{aligned}
f S_{\underset{\sim}{\kappa}}(g) = -f \log(1-g) &= f(g + g^2/2 + g^3/3 + \cdots) \\
&\ge g(f + f^2/2 + f^2/3 + \cdots) = g S_{\underset{\sim}{\kappa}}(f)
\end{aligned}
$$

whenever $0 \le f \le g \le 1$, with strict inequality whenever $0 < f < g$. Since $\underset{\sim}{\kappa}$ is integrable, it is immediate from the definition (6) that $S_{\underset{\sim}{\kappa}} f$ and $S_{\underset{\sim}{\kappa}} g$ are integrable, and it follows that

$$
\int_{\mathcal{S}} f S_{\underset{\sim}{\kappa}} g \ge \int_{\mathcal{S}} g S_{\underset{\sim}{\kappa}} f,
$$

with strict inequality unless $f = g$ a.e. Since Lemma 8 gives the reverse inequality, we have $f = g$ a.e., and thus $f = \Phi_{\underset{\sim}{\kappa}} f = \Phi_{\underset{\sim}{\kappa}} g = g$. The second statement then follows from Lemma 7. $\square$

Theorem 10 generalizes the corresponding result in [10], namely Lemma 5.9. Indeed, in the edge-only case (when only κ_2 is non-zero), the operators $S_{\underset{\sim}{\kappa}}$ and $T_{\kappa_{\mathrm{e}}}$ coincide, and Lemma 8 holds trivially, using the symmetry of $T_{\kappa_{\mathrm{e}}}$. This shows that, with hindsight, the proof of Lemma 5.9 in [10] may be simplified considerably, by considering $\int_{\mathcal{S}} fTg$ instead of $\int_{\mathcal{S}} fTh$, $h = (g-f)/2$. This is significant, since the proof in [10] does not adapt readily to the present context.

Although simple, the proof of Theorem 10 above is a little mysterious from a branching process point of view. It is tempting to think that the result is 'obvious', and indeed that a corresponding result should hold for any Galton–Watson process. However, some conditions are certainly necessary, and it is not clear what the right conditions are for a general process. (Irreducibility is always needed, of course.) In [34], a corresponding result is proved for a general branching process satisfying a certain continuity assumption; the proof uses the convexity property $\Phi(\lambda f) \geq \lambda \Phi(f)$ for any function $0 \leq f \leq 1$ and any $0 \leq \lambda \leq 1$, which holds for all branching processes. In Theorem 10, continuity is not needed, but some kind of symmetry is; there does not seem to be an obvious common generalization of these results.

Indeed, the next example shows that the situation is not that simple: in the compound Poisson case (as opposed to the simple Poisson case), symmetry of the relevant linear operator is not enough.

Example 11. Let $\mathcal{S} = \{1, 2, 3, \ldots\}$ with $\mu\{i\} = 2^{-i}$ for each i, and consider the branching process $\mathfrak{X} = \mathfrak{X}(x)$ with type space $(\mathcal{S}, \mu)$ defined as follows. Start with a single particle of some given type x. Each particle of type i has a Poisson number of children of type $i+1$ with mean $2 = 2^{i+2}\mu\{i+1\}$; we call these 'forward children'. Also, for $i \geq 2$, a particle of type i has 'backward children' of type $i-1$: the number of these is 4^{i+1} times a Poisson with mean 4^{-i}. Note that the expected number of backward children is $4 = 2^{i+1}\mu\{i-1\}$. Defining the 'edge-kernel' κ_{e} so that the expected number of children of type j that each particle of type i has is given by $\kappa_{\mathrm{e}}(i,j)\mu\{j\}$, we have $\kappa_{\mathrm{e}}(i,j) = 2^{1+\max\{i,j\}}$ if $|i-j| = 1$ and $\kappa_{\mathrm{e}}(i,j) = 0$ otherwise, so κ_{e} is symmetric and irreducible.

Define the non-linear operator Φ associated to $\mathfrak{X}$ in the natural way, so $\Phi(f)(x)$ is the probability that at least one child of the root of type x has a certain property, if each child of type y has this property with probability $f(y)$. As before, the survival probability $\rho(x)$ satisfies $\rho = \Phi(\rho)$.

Let $\tau(x)$ denote the probability that the process *survives transiently*, i.e., survives forever, but, for each i, contains in total only finitely many particles of type i. Consider the 'forward process' given by ignoring backward children. This is simply a Poisson Galton–Watson process with on average 2 offspring, and so survives with some positive probability. Also, given that it survives, there is a positive probability that for every t, generation t contains at most 3^t particles, say. But since the particles in generation t have type $x+t$, the expected number of *sets* of backwards children of all particles in the forward process is at most $\sum_{t\geq 0} 3^t 4^{-t-1} < \infty$, and with positive probability the particles in the forward process have no backwards children. But in this case, the forward process is the whole process, and the process survives transiently. Hence $\tau(x) > 0$ for every x.

Let $\sigma(x) = \rho(x) - \tau(x)$ be the probability that the process survives recurrently. Considering the children of the initial particle, we see that $\sigma = \Phi(\sigma)$. The process restricted to any two consecutive types is already supercritical, and so has positive probability of surviving by alternating between these types. Thus $\sigma(x) > 0$ for all x. We showed above that $\tau(x) = \rho(x) - \sigma(x) > 0$ for all x, so $0 < \sigma(x) < \rho(x)$, and $f = \Phi(f)$ has (at least) two non-zero solutions, namely σ and ρ.

Let us turn to the analysis of the solution $\rho_{\underset{\sim}{\kappa}}$ to (7), and in particular the question of when $\rho > 0$, i.e., when the branching process $\mathfrak{X}_{\underset{\sim}{\kappa}}$ is supercritical. Throughout we consider an integrable hyperkernel $\underset{\sim}{\kappa}$, with corresponding edge kernel κ_{e}.

Recall that we may assume that $\lambda(x) = S_{\underset{\sim}{\kappa}}(1)(x)$ is finite everywhere. Hence, for any f satisfying (7), we have $f(x) < 1$ for all x. On the other hand, we cannot assume that κ_{e} is integrable, or indeed finite. For one natural example, consider the integrable hyperkernel with each κ_r constant, and $\kappa_r = 1/r^3$. In this case $\kappa_{\mathrm{e}}(x, y) = \infty$ for all x and y. If κ_{e} is infinite on a set of positive measure, then we take $\|T_{\kappa_{\mathrm{e}}}\|$ to be infinite.

Lemma 12. *If $\|T_{\kappa_{\mathrm{e}}}\| \leq 1$, then $\rho(\underset{\sim}{\kappa}) = 0$.*

Proof. Suppose that f is a solution to (7) that is not 0 a.e. Since $-\log(1-t) > t$ for $0 < t < 1$, we have $S_{\underset{\sim}{\kappa}}(f)(x) \geq f(x)$, with strict inequality on a set of positive measure. But $T_{\kappa_{\mathrm{e}}}(f)(x) \geq S_{\underset{\sim}{\kappa}}(f)(x)$ by (10), so $T_{\kappa_{\mathrm{e}}}(f)(x) \geq f(x)$, with strict inequality on a set of positive measure. Hence $\|T_{\kappa_{\mathrm{e}}} f\|_2 > \|f\|_2$, so $\|T_{\kappa_{\mathrm{e}}}\| > 1$. $\square$

Lemmas 5.12 and 5.13 of [10] carry over to the present context, with only minor modifications.

Lemma 13. *If $0 \leq f \leq 1$ and $\Phi_{\underset{\sim}{\kappa}}(f) \geq f$, then $\Phi^m_{\underset{\sim}{\kappa}}(f) \nearrow g$ as $m \to \infty$, for some $1 \geq g \geq f$ with $\Phi_{\underset{\sim}{\kappa}}(g) = g$.*

Proof. By induction $f \leq \Phi_{\underset{\sim}{\kappa}}(f) \leq \Phi^2_{\underset{\sim}{\kappa}}(f) \leq \cdots$. Since $0 \leq \Phi^m_{\underset{\sim}{\kappa}}(f) \leq 1$, it follows that $g(x) = \lim_{m\to\infty} \Phi^m_{\underset{\sim}{\kappa}}(f)(x)$ exists for every x, and $0 \leq g \leq 1$. From monotone convergence we have $S_{\underset{\sim}{\kappa}}(g) = \lim_{m\to\infty} S_{\underset{\sim}{\kappa}}(\Phi^m_{\underset{\sim}{k}}(f))$, from which it follows that $\Phi_{\underset{\sim}{\kappa}}(g) = g$. $\square$

Lemma 14. *If there is a function $f : \mathcal{S} \to [0, 1]$, not a.e. 0, such that $S_{\underset{\sim}{\kappa}}(f) \geq (1+\delta)f$ for some $\delta > 0$, then $\rho(\underset{\sim}{\kappa}) > 0$.*

Proof. The proof is the same as that of Lemma 5.13 in [10], using $S_{\underset{\sim}{\kappa}}$ in place of T_κ. $\square$

The next step is to show that if $\|T_{\kappa_{\mathrm{e}}}\| > 1$, then there is a function f with the property described in Lemma 14. In [10] we did this by considering a bounded kernel. Here we have to be a little more careful, as we are working with the non-linear operator $S_{\underset{\sim}{\kappa}}$ rather than with $T_{\kappa_{\mathrm{e}}}$; this is no problem if we truncate our kernels suitably.

Definition 15. We call a hyperkernel $\underset{\sim}{\kappa} = (\kappa_r)_{r\geq 2}$ *bounded* if two conditions hold: only finitely many of the κ_r are non-zero, and each κ_r is bounded.

Similarly (for later use), a general kernel family $(\kappa_F)_{F\in\mathcal{F}}$ is *bounded* if only finitely many of the κ_F are non-zero, and each κ_F is bounded.

In other words, $\underset{\sim}{\kappa}$ is bounded if there are constants R and M such that $\kappa_r = 0$ for $r > R$, and κ_r is pointwise bounded by M for $r \leq R$. Note that if $\underset{\sim}{\kappa}$ is bounded, then the corresponding edge kernel κ_{e} is bounded in the usual sense.

Given a hyperkernel $\underset{\sim}{\kappa} = (\kappa_r)$, for each $M > 0$ we let $\underset{\sim}{\kappa}^M$ be the bounded hyperkernel obtained from κ by truncating each κ_r, $r \leq M$, at M, and replacing κ_r by a zero kernel for $r > M$. I.e.,

$$\kappa_r^M = \begin{cases} \kappa_r \wedge M, & r \leq M, \\ 0, & r > M. \end{cases} \tag{17}$$

The truncation $\underset{\sim}{\kappa}^M = (\kappa_F^M)_{F\in\mathcal{F}}$ of a general kernel family $(\kappa_F)_{F\in\mathcal{F}}$ is defined similarly.

Lemma 16. *If $\|T_{\kappa_{\mathrm{e}}}\| > 1$ then there is a $\delta > 0$ and an $f : \mathcal{S} \to [0, 1]$, not a.e. 0, such that $S_{\underset{\sim}{\kappa}}(f) \geq (1 + \delta)f$.*

Proof. We slightly modify the proof of Lemma 5.16 of [10].

Consider the truncated hyperkernels $\underset{\sim}{\kappa}^M$ defined in (17). From (9) and monotone convergence, the corresponding edge kernels κ_{e}^M tend up to κ_{e} (which may be infinite in some places) pointwise. Arguing as in the proof of Lemma 5.16 of [10], since $\|T_{\kappa_{\mathrm{e}}}\| > 1$ there is some positive f with $\|f\|_2 = 1$ and $1 < \|T_{\kappa_{\mathrm{e}}} f\|_2 \leq \infty$. By monotone convergence, $T_{\kappa_{\mathrm{e}}^M} f \nearrow T_{\kappa_{\mathrm{e}}} f$, so $\|T_{\kappa_{\mathrm{e}}^M} f\|_2 \nearrow \|T_{\kappa_{\mathrm{e}}} f\|_2$, and there is some M with $\|T_{\kappa_{\mathrm{e}}^M}\| \geq \|T_{\kappa_{\mathrm{e}}^M} f\|_2 > 1$.

Since κ_{e}^M is bounded, by Lemma 5.15 of [10] it follows that there is a bounded $f \geq 0$ with f not 0 a.e. such that

$$T_{\kappa_{\mathrm{e}}^M} f = \|T_{\kappa_{\mathrm{e}}^M}\| f = (1 + 2\delta)f,$$

where $\delta = (\|T_{\kappa_{\mathrm{e}}^M}\| - 1)/2 > 0$. We may assume that $0 \leq f \leq 1$. If $0 \leq y_i \leq \gamma < 1$, $i = 1, \ldots, r$, then (by induction) $1 - \prod_{i=1}^r (1 - y_i) \geq (1 - \gamma)^{r-1} \sum_1^r y_i$, and it follows that if $\gamma > 0$ is chosen small enough, then

$$S_{\underset{\sim}{\kappa}^M}(\gamma f) \geq (1 - \gamma)^{M-1} T_{\kappa_{\mathrm{e}}^M}(\gamma f) \geq (1 + \delta)(\gamma f).$$

Since $S_{\underset{\sim}{\kappa}}(\gamma f) \geq S_{\underset{\sim}{\kappa}^M}(\gamma f)$, the result follows. □

Theorem 6 follows by combining the results above.

Proof of Theorem 6. Together Lemmas 12, 14 and 16 show that $\rho(\underset{\sim}{\kappa}) > 0$ if and only if $\|T_{\kappa_{\mathrm{e}}}\| > 1$. Uniqueness is given by Theorem 10. The final statement is immediate from Lemma 9. □

Having proved Theorem 6, our next aim is to prove Theorem 5. The basic strategy will involve comparing the neighbourhoods of a vertex in the random graph $G(n,\underset{\sim}{\kappa})$ with the branching process $\mathfrak{X}_{\underset{\sim}{\kappa}}$. As in [10], it will be convenient to carry out the comparison only for certain restricted hyperkernels. In order to deduce results about $G(n,\underset{\sim}{\kappa})$ in general, one needs approximation results both for the graph and for the branching process. We now turn to such results for branching processes.

Lemma 6.3 and Theorems 6.4 and 6.5 of [10] carry over to the present context, *mutatis mutandis*, using the results above about $\rho(\underset{\sim}{\kappa})$ instead of the equivalents in [10], and replacing T_κ by $S_{\underset{\sim}{\kappa}}$ or T_{κ_e} as appropriate: $S_{\underset{\sim}{\kappa}}$ when considering Φ_κ, and T_{κ_e} when arguing using L^2-norms. In these results $\rho_{\underset{\sim}{\kappa}}$ denotes the function $x \mapsto \rho_{\underset{\sim}{\kappa}}(x)$, and $\rho_{\ge k}(\underset{\sim}{\kappa}, x)$ and $\rho_{\ge k}(\underset{\sim}{\kappa})$ denote respectively the probabilities that $\mathfrak{X}_{\underset{\sim}{\kappa}}(x)$ and $\mathfrak{X}_{\underset{\sim}{\kappa}}$ have total size at least k.

Lemma 17. *If $\underset{\sim}{\kappa} \le \underset{\sim}{\kappa}'$, then $\rho(\underset{\sim}{\kappa}) \le \rho(\underset{\sim}{\kappa}')$.* □

Theorem 18. (i) *Let $\underset{\sim}{\kappa}_n$, $n = 1, 2, \ldots$, be a sequence of hyperkernels on $(\mathcal{S}, \mu)$ increasing a.e. to an integrable hyperkernel $\underset{\sim}{\kappa}$. Then $\rho_{\underset{\sim}{\kappa}_n} \nearrow \rho_{\underset{\sim}{\kappa}}$ a.e. and $\rho(\underset{\sim}{\kappa}_n) \nearrow \rho(\underset{\sim}{\kappa})$.*

(ii) *Let $\underset{\sim}{\kappa}_n$, $n = 1, 2, \ldots$, be a sequence of integrable hyperkernels on $(\mathcal{S}, \mu)$ decreasing a.e. to $\underset{\sim}{\kappa}$. Then $\rho_{\underset{\sim}{\kappa}_n} \searrow \rho_{\underset{\sim}{\kappa}}$ a.e. and $\rho(\underset{\sim}{\kappa}_n) \searrow \rho(\underset{\sim}{\kappa})$.* □

Theorem 19. (i) *Let $\underset{\sim}{\kappa}_n$, $n = 1, 2, \ldots$, be a sequence of hyperkernels on $(\mathcal{S}, \mu)$ increasing a.e. to a hyperkernel $\underset{\sim}{\kappa}$. Then, for every $k \ge 1$, $\rho_{\ge k}(\underset{\sim}{\kappa}_n; x) \nearrow \rho_{\ge k}(\underset{\sim}{\kappa}; x)$ for a.e. x and $\rho_{\ge k}(\underset{\sim}{\kappa}_n) \nearrow \rho_{\ge k}(\underset{\sim}{\kappa})$.*

(ii) *Let $\underset{\sim}{\kappa}_n$, $n = 1, 2, \ldots$, be a sequence of integrable hyperkernels on $(\mathcal{S}, \mu)$ decreasing a.e. to $\underset{\sim}{\kappa}$. Then, for every $k \ge 1$, $\rho_{\ge k}(\underset{\sim}{\kappa}_n; x) \searrow \rho_{\ge k}(\underset{\sim}{\kappa}; x)$ for a.e. x and $\rho_{\ge k}(\underset{\sim}{\kappa}_n) \searrow \rho_{\ge k}(\underset{\sim}{\kappa})$.* □

Remark 20. The assumption that $\underset{\sim}{\kappa}_n$ be integrable in Theorems 18(ii) and 19(ii) can be weakened to $\lambda_{\underset{\sim}{\kappa}_n}(x) < \infty$ for a.e. x, where $\lambda_{\underset{\sim}{\kappa}_n}(x)$ is the expected number of child cliques in $\mathfrak{X}_{\underset{\sim}{\kappa}_n}$ of a particle of type x; see [10].

3 Local coupling

We now turn to the local coupling between our random graph and the corresponding branching process, relating the distribution of small components in $G(n,\underset{\sim}{\kappa})$ to the branching process $\mathfrak{X}_{\underset{\sim}{\kappa}}$. In [10], we were essentially forced to condition on the vertex types, since these were allowed to be deterministic to start with. Here, with iid vertex types, there is no need to do so. This allows us to couple directly for all bounded hyperkernels, rather than simply for finite type ones.

We shall consider a variant of the usual component exploration process, designed to get around the following problem. When we test edges from a given vertex v to all other vertices, the probability of finding a given edge vw depends on the type of w as well as that of v. Hence, *not* finding such an edge changes the conditional distribution of the type of w. If the kernel is well behaved, it is easy to see that this is a small effect. Rather than quantify this, it is easier to embed $G(n, \underset{\sim}{\kappa})$ inside a larger random graph with uniform kernels. Testing edges in the larger graph does not affect the conditional distribution of the vertex types; we make this precise below. In doing so, it will be useful to take the hypergraph viewpoint: given a hyperkernel $\underset{\sim}{\kappa}$, let $H(n, \underset{\sim}{\kappa})$ be the hypergraph on $[n]$ constructed according to the same rules as $G(n, \underset{\sim}{\kappa})$, except that instead of adding a K_r we add a hyperedge with r vertices. In fact, we consider the Poisson version of the model, allowing multiple copies of the same hyperedge.

Let $\underset{\sim}{\kappa}$ be a bounded hyperkernel, and let $\underset{\sim}{\kappa}^+$ be a corresponding upper bound, so κ_r^+ is the constant kernel M for $r \le R$, and zero for $r > R$, while $\kappa_r \le \kappa_r^+$ holds pointwise for all r.

We construct coupled random (multi-)hypergraphs H_n and H_n^+ on $[n]$ as follows: first construct $H_n^+ = H(n, \underset{\sim}{\kappa}^+)$ by taking for $2 \le r \le R$ a Poisson $\mathrm{Po}(r!M/n^{r-1})$ number of copies of each possible r-element hyperedge, with all these numbers independent. Although in our formal definition of H_n^+ we first decide the vertex types, H_n^+ is clearly independent of these types. Hence, given H_n^+, the types are (still) iid with distribution μ.

Given H_n^+ and the iid types $x_1, \ldots, x_n$ of the vertices, we may form H_n by selecting each hyperedge $\{v_1, \ldots, v_r\}$ of H_n^+ to be a hyperedge of H_n with probability $\kappa_r(x_{v_1}, \ldots, x_{v_r})/M$, independently of all other hyperedges. It is easy to see that this gives the right distribution for $H_n = H(n, \underset{\sim}{\kappa})$. (If we disallowed multiple copies of an edge, there would be an irrelevant small correction here.)

Turning to the branching processes, there is an analogous coupling of $\mathfrak{X}_{\underset{\sim}{\kappa}}$ and $\mathfrak{X}_{\underset{\sim}{\kappa}^+}$: first construct $\mathfrak{X}_{\underset{\sim}{\kappa}^+}$, which may be viewed as a single-type process, according to our two-step construction via child cliques. Then assign each particle a type according to the distribution μ, independently of the other particles and of the branching process. Then form the child cliques in $\mathfrak{X}_{\underset{\sim}{\kappa}}$ by keeping each child clique in $\mathfrak{X}_{\underset{\sim}{\kappa}^+}$ with an appropriate probability depending on the types, deleting not only the children corresponding to deleted child cliques, but also all their descendants.

Let $v \in [n]$ be chosen uniformly at random, independently of H_n and H_n^+. Let Γ_d denote the d-neighbourhood of v in H_n, and Γ_d^+ that in H_n^+. Counting the expected number of cycles shows that for any fixed d, the hypergraph Γ_d^+ is whp treelike. Furthermore, standard arguments as for $G(n, c/n)$ show that one may couple Γ_d^+ and the first d generations of $\mathfrak{X}_{\underset{\sim}{\kappa}^+}$ so as to agree in the natural sense whp. When Γ_d^+ is treelike, then $\Gamma_d \subset \Gamma_d^+$ may be constructed using exactly the same random deletion process that gives (the first d generations of) $\mathfrak{X}_{\underset{\sim}{\kappa}}$ as a subset of $\mathfrak{X}_{\underset{\sim}{\kappa}^+}$. It follows that Γ_d and the first d generations of $\mathfrak{X}_{\underset{\sim}{\kappa}}$ may be coupled to agree whp.

Recalling that $G(n, \underset{\sim}{\kappa})$ and H_n have the same components, for any fixed $k \ge 1$ one can determine whether the component containing v has exactly k

vertices by examining Γ_{k+1}. It follows that

$$\mathbb{E}N_k(G(n,\underset{\sim}{\kappa})) = n\mathbb{P}(|\mathfrak{X}_{\underset{\sim}{\kappa}}| = k) + o(1).$$

As in [10], starting from two random vertices easily gives a corresponding second moment bound, giving the following result.

Lemma 21. *Let $\underset{\sim}{\kappa}$ be a bounded hyperkernel. Then*

$$\frac{1}{n}N_k(G(n,\underset{\sim}{\kappa})) \xrightarrow{\mathrm{p}} \mathbb{P}(|\mathfrak{X}_{\underset{\sim}{\kappa}}| = k)$$

for any fixed k. □

Of course it makes no difference whether we work with N_k or $N_{\geq k} = n - \sum_{j=1}^{k-1} N_j$: Lemma 21 also tells us that

$$\frac{1}{n}N_{\geq k}(G(n,\underset{\sim}{\kappa})) \xrightarrow{\mathrm{p}} \mathbb{P}(|\mathfrak{X}_{\underset{\sim}{\kappa}}| \geq k). \tag{18}$$

The extension to arbitrary hyperkernels is easy from Theorem 19.

Lemma 22. *Let $\underset{\sim}{\kappa}$ be an integrable hyperkernel. Then for each fixed k we have*

$$\frac{1}{n}N_{\geq k}(G(n,\underset{\sim}{\kappa})) \xrightarrow{\mathrm{p}} \rho_{\geq k}(\underset{\sim}{\kappa}).$$

Proof. As in [10], we simply approximate $\underset{\sim}{\kappa}$ by bounded hyperkernels. For $M > 0$ let $\underset{\sim}{\kappa}^M$ be the truncated hyperkernel defined by (17).

Let $k \geq 1$ be fixed, and let $\varepsilon > 0$ be arbitrary. From monotone convergence and integrability,

$$\lim_{M\to\infty} \sum_{r\geq 2} \int_{\mathcal{S}^r} r\kappa_r^M = \sum_{r\geq 2} \int_{\mathcal{S}^r} r\kappa_r < \infty,$$

so for M large enough we have

$$\Delta = \sum_{r\geq 2} \int_{\mathcal{S}^r} r(\kappa_r - \kappa_r^M) \leq \varepsilon^2/(6k),$$

say. By Theorem 19(i), increasing M if necessary, we may also assume that

$$\rho_{\geq k}(\underset{\sim}{\kappa}^M) \geq \rho_{\geq k}(\underset{\sim}{\kappa}) - \varepsilon/3. \tag{19}$$

Since $\underset{\sim}{\kappa}^M \leq \underset{\sim}{\kappa}$ holds pointwise, we may couple the hypergraphs H_n' and H_n associated to $G(n,\underset{\sim}{\kappa}^M)$ and $G(n,\underset{\sim}{\kappa})$ so that $H_n' \subseteq H_n$. Recall that $G(n,\underset{\sim}{\kappa})$ is produced from H_n by replacing each hyperedge E with r vertices by an r-clique. However, as noted earlier, if we form G_n from H_n by replacing each E by any connected simple graph on the same set of vertices, then G_n and $G(n,\underset{\sim}{\kappa})$ will have exactly the same component structure, and in particular $N_{\geq k}(G_n) =$

$N_{\geq k}(G(n, \underset{\sim}{\kappa}))$. Let us form G_n and G_n' in this way from H_n and H_n', replacing any hyperedge with r vertices by some tree on the same set of vertices. Recalling that $H_n' \subseteq H_n$, we may of course assume that $G_n' \subseteq G_n$.

Writing $E_r(H)$ for the number of r-vertex hyperedges in a hypergraph H,

$$\begin{aligned}\mathbb{E}\big(E(G_n) \setminus E(G_n')\big) &\leq \sum_{r\geq 2}(r-1)\mathbb{E}\big(E_r(H_n) - E_r(H_n')\big) \\ &\leq \sum_{r\geq 2}(r-1)n\int_{\mathcal{S}^r}(\kappa_r - \kappa_r^M) \leq n\Delta.\end{aligned}$$

Hence,

$$\mathbb{P}\big(|E(G_n) \setminus E(G_n')| \geq \varepsilon n/6k\big) \leq n\Delta/(\varepsilon n/6k) \leq \varepsilon.$$

Recalling that $G_n' \subseteq G_n$ and noting that adding one edge to a graph cannot change $N_{\geq k}$ by more than $2k$, we see that with probability at least $1-\varepsilon$ we have

$$\big|N_{\geq k}(G(n, \underset{\sim}{\kappa}^M)) - N_{\geq k}(G(n, \underset{\sim}{\kappa}))\big| = \big|N_{\geq k}(G_n') - N_{\geq k}(G_n)\big| \leq 2k(\varepsilon n/6k) = \varepsilon n/3.$$

Applying Lemma 21 (or rather (18)) to the bounded hyperkernel $\underset{\sim}{\kappa}^M$, we have $\frac{1}{n}N_{\geq k}(G(n, \underset{\sim}{\kappa}^M)) \xrightarrow{\mathrm{p}} \rho_{\geq k}(\underset{\sim}{\kappa}^M)$. Using (19) it follows that with probability at least $1-2\varepsilon$, say, we have $|\frac{1}{n}N_{\geq k}(G(n, \underset{\sim}{\kappa})) - \rho_{\geq k}(\underset{\sim}{\kappa})| \leq \varepsilon$. Since $\varepsilon > 0$ was arbitrary, we thus have $\frac{1}{n}N_{\geq k}(G(n, \underset{\sim}{\kappa})) \xrightarrow{\mathrm{p}} \rho_{\geq k}(\underset{\sim}{\kappa})$ as required. □

4 The giant component

The local coupling results of the previous section easily give us the 'right' number of vertices in large components. As usual, we will pass from this to a giant component by using the 'sprinkling' method of Erdős and Rényi [25], first uncovering the bulk of the edges, and then using the remaining 'sprinkled' edges to join up the large components. The following lemma gathers together the relevant consequences of the results in the previous section.

Lemma 23. *Let $\underset{\sim}{\kappa} = (\kappa_r)$ be an integrable hyperkernel, and let $G_n = G(n, \underset{\sim}{\kappa})$. Then $C_1(G_n) \leq \rho(\underset{\sim}{\kappa})n + o_{\mathrm{p}}(n)$. Furthermore, given any $\varepsilon > 0$, there is a $\delta > 0$ and $\omega = \omega(n) \to \infty$ such that*

$$N_{\geq\omega}(G_n') \geq (\rho(\underset{\sim}{\kappa}) - \varepsilon)n \tag{20}$$

holds whp, where $G_n' = G(n, (1-\delta)\underset{\sim}{\kappa})$.

Proof. From Lemma 22 we have $\frac{1}{n}N_{\geq k}(G_n) \xrightarrow{\mathrm{p}} \rho_{\geq k}(\underset{\sim}{\kappa})$ for each fixed k. Since $\rho_{\geq k}(\underset{\sim}{\kappa}) \to \rho(\underset{\sim}{\kappa})$ as $k \to \infty$, it follows that for some $\omega = \omega(n) \to \infty$ we have

$$\frac{1}{n}N_{\geq\omega}(G_n) \xrightarrow{\mathrm{p}} \rho(\underset{\sim}{\kappa}); \tag{21}$$

we may and shall assume that $\omega = o(n)$. Since $C_1(G_n) \le \max\{\omega, N_{\ge\omega}(G_n)\}$, the first statement of the lemma follows.

For the second, we may of course assume that $\rho(\underset{\sim}{\kappa}) > \varepsilon$; otherwise, there is nothing to prove. As $\delta \to 0$, from Theorem 18(i) we have $\rho((1-\delta)\underset{\sim}{\kappa}) \to \rho(\underset{\sim}{\kappa})$. Fix $\delta > 0$ with $\rho((1-\delta)\underset{\sim}{\kappa}) > \rho(\underset{\sim}{\kappa}) - \varepsilon/2$, and let $G'_n = G(n, (1-\delta)\underset{\sim}{\kappa})$. Applying (21) to G'_n, there is some $\omega = \omega(n) \to \infty$ such that

$$N = N_{\ge\omega}(G'_n) = \rho((1-\delta)\underset{\sim}{\kappa})n + o_{\mathrm{p}}(n) \ge (\rho(\underset{\sim}{\kappa}) - \varepsilon/2)n + o_{\mathrm{p}}(n),$$

which implies (20). □

In the light of Lemma 23, to prove Theorem 5 it suffices to show that if $\underset{\sim}{\kappa}$ is irreducible, then for any $\varepsilon > 0$ we have

$$C_1(G_n) \ge (\rho(\underset{\sim}{\kappa}) - 2\varepsilon)n \tag{22}$$

whp; then $C_1(G_n)/n \xrightarrow{\mathrm{p}} \rho(\underset{\sim}{\kappa})$ as required. Also, from (21) and the fact that $C_1(G_n) + C_2(G_n) \le \max\{2\omega, N_{\ge\omega}(G_n) + \omega\}$, we obtain $C_2(G_n) = o_{\mathrm{p}}(n)$ as claimed.

Since $(1-\delta)\underset{\sim}{\kappa} \le \underset{\sim}{\kappa}$, there is a natural coupling of the graphs G'_n and G_n appearing in Lemma 23 in which $G'_n \subseteq G_n$ always holds. Our aim is to show that, whp, in passing from G'_n to G_n, the extra 'sprinkled' edges join up almost all of the N vertices of G'_n in 'large' components (those of size at least ω) into a single component.

Unfortunately, we have to uncover the vertex types before sprinkling, so we do not have the usual independence between the bulk and sprinkled edges. A similar problem arose in Bollobás, Borgs, Chayes and Riordan [9] in the graph context, as opposed to the present hypergraph context. It turns out that we can easily reduce to the graph case, and thus apply a lemma from [9]. This needs a little setting up, however. Here it will be convenient to take $\mathcal{S} = [0,1]$ with μ Lebesgue measure; as noted in Section 1, this loses no generality.

Let κ be a bounded symmetric measurable function $\kappa : [0,1]^2 \to \mathbb{R}$. Following Frieze and Kannan [27], the *cut norm* $\|\kappa\|_\square$ of κ is defined by

$$\|\kappa\|_\square = \sup_{S,T\subseteq[0,1]} \left| \int_{S\times T} \kappa(x,y)\,dx\,dy \right|,$$

where the supremum is taken over all pairs of measurable sets.

Given a kernel κ on $[0,1]$ and a measurable function $\varphi : [0,1] \to [0,1]$, let $\kappa^{(\varphi)}$ be the kernel defined by

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

If φ is a measure-preserving bijection, then $\kappa^{(\varphi)}$ is a *rearrangement* of κ. (One can also consider measure-preserving bijections between subsets of $[0,1]$ with full measure; it makes no difference.) We write $\kappa \sim \kappa'$ if κ' is a rearrangement of κ.

Given two kernels κ, κ' on $[0,1]$, the *cut metric* of Borgs, Chayes, Lovász, Sós and Vesztergombi [17] is defined by

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

Note that this is a pseudo-metric rather than a metric, as we can have $\delta_\square(\kappa,\kappa') = 0$ for different kernels. (Probabilistically, it is probably more natural to consider couplings between kernels as in [17], rather than rearrangements, but this is harder to describe briefly and turns out to make no difference.)

Let A_n be a symmetric n-by-n matrix with non-negative entries a_{ij}, which we may think of as a (dense) weighted graph. There is a piecewise-constant kernel κ_{A_n} associated to A_n; this simply takes the value a_{ij} on the square $((i-1)/n, i/n] \times ((j-1)/n, j/n]$, $1 \le i,j \le n$. There is also a sparse random graph $G(A_n)$ associated to A_n; this is the graph on $[n]$ in which edges are present independently, and the probability that ij is an edge is a_{ij}/n. (If A_n has non-zero diagonal entries then $G(A_n)$ may contain loops. These are irrelevant here.)

The main result of Bollobás, Borgs, Chayes and Riordan [9] is that if κ is an irreducible bounded kernel and (A_n) is a sequence of matrices with uniformly bounded entries such that $\delta_\square(\kappa_{A_n},\kappa)\to 0$, then the normalized size of the giant component in $G(A_n)$ converges in probability to $\rho(\kappa)$. The sprinkling argument there relies on the following lemma.

Lemma 24. *Let κ be an irreducible bounded kernel on $[0,1]$, and δ and $\beta_{\max}$ positive constants. There is a constant $c = c(\kappa,\beta_{\max},\delta) > 0$ such that whenever A_n is a sequence of symmetric matrices with entries in $[0,\beta_{\max}]$ with $\delta_\square(\kappa_{A_n},\kappa)\to 0$, then for sufficiently large n we have*

$$\mathbb{P}(V_n \sim_{G(A_n)} V_n') \ge 1 - \exp(-cn)$$

for all disjoint V_n, $V_n' \subset [n]$ with $|V_n|, |V_n'| \ge \delta n$, where $V_n \sim_{G(A_n)} V_n'$ denotes the event that $G(A_n)$ contains a path starting in V_n and ending in V_n'. □

In fact, this lemma is not stated explicitly in [9], but this is exactly the content of the end of Section 3. (For a stronger version of this lemma see [12].)

We shall apply Lemma 24 to graphs $G(A_n)$ corresponding to (subgraphs of) $G(n,\delta\underset{\sim}{\kappa})$, where δ is as in Lemma 23. To achieve independence between edges, we shall simply take only one edge from each hyperedge. Unfortunately, the problem of conditioning on the x_i still remains; we shall return to this shortly.

Let $\underset{\sim}{\kappa}$ be an integrable hyperkernel and H_n the Poisson (multi-)hypergraph corresponding to $G(n,\underset{\sim}{\kappa})$. Given the sequence $\mathbf{x} = (x_1,\dots,x_n)$, let $\widetilde{G}(n,\underset{\sim}{\kappa},\mathbf{x})$ be the random (multi-)graph formed from H_n by replacing each r-vertex hyperedge E by a single edge, chosen uniformly at random from the $\binom{r}{2}$ edges corresponding to E. With $\mathbf{x}$ fixed, the numbers of copies of each edge E in H_n are independent Poisson random variables. From basic properties of Poisson processes, it follows that the number of copies of each edge ij in $\widetilde{G}(n,\underset{\sim}{\kappa},\mathbf{x})$ are also independent Poisson random variables. Our next aim is to calculate the edge probabilities in $\widetilde{G}(n,\underset{\sim}{\kappa},\mathbf{x})$.

Given $x_1, \dots, x_n$ and distinct $i, j \in [n]$, for $r \geq 2$ let

$$a_{r,i,j} = n^{-(r-2)} \sum \kappa_r(x_i, x_j, x_{k_3}, \dots, x_{k_r}),$$

where the sum runs over all $(n-2)_{(r-2)}$ sequences $k_3, \dots, k_r$ of distinct indices in $[n] \setminus \{i, j\}$, and let A be the n-by-n matrix with entries

$$a_{ij} = 2 \sum_{r \geq 2} a_{r,i,j} \tag{23}$$

for $i \neq j$ and $a_{ij} = 0$ if $i = j$.

With $\mathbf{x}$ given, the expected number of r-vertex hyperedges in H_n containing ij is $r(r-1)a_{r,i,j}/n$. Hence the expected number of ij edges in $\widetilde{G}(n, \underset{\sim}{\kappa}, \mathbf{x})$ is exactly a_{ij}/n. Now a_{ij} clearly depends on x_i and x_j. Unfortunately, it also depends on all the other x_k. The next lemma will show that the latter dependence can be neglected.

Set

$$\kappa_r(x, y, *) = \int_{\mathcal{S}^{r-2}} \kappa_r(x, y, x_3, x_4, \dots, x_r)\, d\mu(x_3) \cdots d\mu(x_r),$$

and let τ be the 're-scaled' edge kernel defined by

$$\tau(x, y) = 2 \sum_{r \geq 2} \kappa_r(x, y, *). \tag{24}$$

Comparing with the formula (9) for $\kappa_{\mathrm{e}}(x, y)$, note that we have divided each term in the sum in (9) by $\binom{r}{2}$, the number of edges in K_r. Note that

$$\tau(x, y) = 0 \iff \kappa_{\mathrm{e}}(x, y) = 0. \tag{25}$$

Lemma 25. *Let $\underset{\sim}{\kappa} = (\kappa_r)_{r \geq 2}$ be an integrable hyperkernel. Then*

$$\mathbb{E} \frac{1}{n^2} \sum_{i \neq j} |a_{r,i,j} - \kappa_r(x_i, x_j, *)| = o(1) \tag{26}$$

for every r, and

$$\mathbb{E} \frac{1}{n^2} \sum_{i \neq j} |a_{ij} - \tau(x_i, x_j)| = o(1). \tag{27}$$

Proof. We have

$$\mathbb{E}\big(a_{r,i,j} \mid x_i, x_j\big) = (n-2)_{(r-2)}\, n^{-(r-2)} \kappa_r(x_i, x_j, *).$$

Suppose first that κ_r is bounded. Let

$$\begin{aligned} Y_{ij} &= a_{r,i,j} - (n-2)_{(r-2)}\, n^{-(r-2)} \kappa_r(x_i, x_j, *) \\ &= n^{-(r-2)} \sum \big(\kappa_r(x_i, x_j, x_{k_3}, \dots, x_{k_r}) - \kappa_r(x_i, x_j, *)\big), \end{aligned}$$

where the sum again runs over all $(n-2)_{(r-2)}$ sequences $k_3, \dots, k_r$ of distinct indices in $[n] \setminus \{i, j\}$. Given x_i and x_j, each term in the sum has mean 0, and any two terms with disjoint index sets $\{k_3, \dots, k_r\}$ are independent. Since there are $O(n^{2r-5})$ pairs of terms with overlapping index sets, and κ_r is bounded, we have

$$\mathbb{E}(Y_{ij}^2 \mid x_i, x_j) = O(n^{2r-5-2(r-2)}) = O(n^{-1}).$$

Thus $\mathbb{E} Y_{ij}^2 = O(n^{-1})$ and $\mathbb{E} \frac{1}{n^2} \sum_{i \neq j} Y_{ij}^2 = O(n^{-1})$, and the Cauchy–Schwarz inequality yields

$$\begin{aligned} \mathbb{E} \frac{1}{n^2} \sum_{i \neq j} |a_{r,i,j} - \kappa_r(x_i, x_j, *)| &= \mathbb{E} \frac{1}{n^2} \sum_{i \neq j} |Y_{ij}| + O(1/n) \\ &\leq \left(\mathbb{E} \frac{1}{n^2} \sum_{i \neq j} Y_{ij}^2 \right)^{1/2} + O(1/n) = O(n^{-1/2}). \end{aligned}$$

This proves (26) and thus (27) for bounded hyperkernels.

For general hyperkernels, we use truncation and define κ_r^M by (17). For the corresponding $a_{r,i,j}^{(M)}$, $A^{(M)}$ and τ^M,

$$\mathbb{E} \frac{1}{n^2} \sum_{i \neq j} |a_{r,i,j} - a_{r,i,j}^{(M)}| \leq \int (\kappa_r - \kappa_r^M),$$

$$\mathbb{E} \frac{1}{n^2} \sum_{i \neq j} |\kappa_r(x_i, x_j, *) - \kappa_r^M(x_i, x_j, *)| \leq \int (\kappa_r - \kappa_r^M),$$

and thus

$$\mathbb{E} \frac{1}{n^2} \sum_{i \neq j} |a_{ij} - a_{ij}^{(M)}| \leq 2 \sum_r \int (\kappa_r - \kappa_r^M),$$

$$\mathbb{E} \frac{1}{n^2} \sum_{i \neq j} |\tau(x_i, x_j) - \tau^M(x_i, x_j)| \leq 2 \sum_r \int (\kappa_r - \kappa_r^M).$$

Since (κ_r) is integrable, given any $\varepsilon > 0$ we can make these expected differences less than ε by choosing M large enough, and the result follows from the bounded case. $\square$

With the preparation above we are now ready to prove Theorem 5.

Proof of Theorem 5. We assume without loss of generality that $\mathcal{S} = [0, 1]$, with μ Lebesgue measure.

Let $\underset{\sim}{\kappa}' = (\kappa'_F)_{F \in \mathcal{F}}$ be an irreducible, integrable kernel family, let $\underset{\sim}{\kappa} = (\kappa_r)_{r \geq 2}$ be the corresponding hyperkernel, given by (4), and let $\varepsilon > 0$. As noted after Lemma 23, in the light of this lemma, it suffices to prove the lower bound (22) on $C_1(G(n, \underset{\sim}{\kappa}))$. We may and shall assume that $\rho(\underset{\sim}{\kappa}) > 0$ and $\varepsilon < \rho(\underset{\sim}{\kappa})/10$, say.

Let $\delta > 0$ and $\omega = \omega(n)$ be as in Lemma 23, and let H_n, H'_n and $\widetilde{H}_n$ be the Poisson multi-hypergraphs associated to the hyperkernels $\underset{\sim}{\kappa}$, $(1-\delta)\underset{\sim}{\kappa}$ and $\delta\underset{\sim}{\kappa}$, respectively. Using the same vertex types $\mathbf{x} = (x_1, \ldots, x_n)$ for all three hypergraphs, there is a natural coupling in which $H_n = H'_n \cup \widetilde{H}_n$, with H'_n and $\widetilde{H}_n$ *conditionally* independent given $\mathbf{x}$.

Define A and τ by (23) and (24), respectively, starting from the integrable hyperkernel $\delta\underset{\sim}{\kappa}$. Note that τ is a kernel on $[0,1]$, while A is an n-by-n matrix that depends on $\mathbf{x}$. Recall from (25) that $\tau(x,y) = 0$ if and only if $\kappa_{\mathrm{e}}(x,y) = 0$, so τ is irreducible. In order to be able to apply Lemma 24, we would like to work with a bounded kernel and matrices that are bounded uniformly in n. We achieve this simply by considering $\overline{A} = (\overline{a}_{ij})$ and $\overline{\tau}$ defined by

$$\overline{a}_{ij} = \min\{a_{ij}, 1\} \quad \text{and} \quad \overline{\tau}(x,y) = \min\{\tau(x,y), 1\}.$$

Let B be the 'sampled' matrix corresponding to τ, defined by

$$b_{ij} = \tau(x_i, x_j),$$

and $\overline{B}$ the corresponding matrix associated to $\overline{\tau}$. The second statement of Lemma 25 tells us that exactly

$$\mathbb{E}\frac{1}{n^2}\sum_{i \neq j} |a_{ij} - b_{ij}| = o(1).$$

Since $|\overline{a}_{ij} - \overline{b}_{ij}| \leq |a_{ij} - b_{ij}|$ for $i \neq j$, while $|\overline{a}_{ii} - \overline{b}_{ii}| \leq 1$, it follows that

$$\mathbb{E}\frac{1}{n^2}\sum_{i,j} |\overline{a}_{ij} - \overline{b}_{ij}| = o(1),$$

or, equivalently, that $\mathbb{E}\|\kappa_{\overline{A}} - \kappa_{\overline{B}}\|_1 = o(1)$, where we write κ_M for the piecewise constant kernel associated to a matrix M.

Since $\delta_\square(\kappa_1, \kappa_2) \leq \|\kappa_1 - \kappa_2\|_\square \leq \|\kappa_1 - \kappa_2\|_1$, it follows that $\mathbb{E}\delta_\square(\kappa_{\overline{A}}, \kappa_{\overline{B}}) \to 0$, and hence that $\delta_\square(\kappa_{\overline{A}}, \kappa_{\overline{B}}) \xrightarrow{\mathrm{p}} 0$. Coupling the random sequences $\mathbf{x}$ for different n appropriately, we may and shall assume that

$$\delta_\square(\kappa_{\overline{A}}, \kappa_{\overline{B}}) \to 0 \tag{28}$$

almost surely.

Since $\overline{\tau}$ is a bounded kernel on $[0,1]$, i.e., a 'graphon' in the terminology of [17], Theorem 4.7 of Borgs, Chayes, Lovász, Sós and Vesztergombi [17] tells us that with probability at least $1 - e^{-n^2/(2\log_2 n)} = 1 - o(1)$, we have $\delta_\square(\kappa_{\overline{B}}, \overline{\tau}) \leq 10 \sup \overline{\tau}/\sqrt{\log_2 n} = o(1)$. It follows that $\delta_\square(\kappa_{\overline{B}}, \overline{\tau}) \to 0$ both in probability and almost surely. Using (28), we see that

$$\delta_\square(\kappa_{\overline{A}}, \overline{\tau}) \to 0 \tag{29}$$

almost surely. Note that $\kappa_{\overline{A}}$ depends on the sequences $\mathbf{x}$.

Let G'_n and G_n be the simple graphs underlying H'_n and H_n. From Lemma 23, (20) holds whp. Let us now condition on H'_n, and in particular on the vertex types $\mathbf{x}$ and on G'_n. We assume, as we may, that (20) holds for all large enough n, and that (29) holds. It suffices to show that with *conditional* probability $1 - o(1)$ we have $C_1(G_n) \geq (\rho(\underset{\sim}{\kappa}) - 2\varepsilon)n$.

Let $\widetilde{G}_n = \widetilde{G}(n, \delta\underset{\sim}{\kappa}, \mathbf{x})$ be the random (multi-)graph defined from $\widetilde{H}_n$ by taking one edge from each hyperedge as before, noting that $G'_n \cup \widetilde{G}_n \subseteq G_n$. As noted earlier, given $\mathbf{x}$ (and G'_n), each possible edge ij is present in $\widetilde{G}_n$ independently. In the multi-graph version, the number of copies of ij is Poisson with mean a_{ij}/n. Passing to a subgraph, we shall take instead the number of copies to be Poisson with mean $\overline{a}_{ij}/n$. Since this mean is $O(1/n)$, the probability that one or more copies of ij is present is a'_{ij}/n, where $a'_{ij} = \overline{a}_{ij} + O(1/n)$. Since $\delta_\square(\kappa_{A'}, \kappa_{\overline{A}}) = O(1/n) = o(1)$, we have $\delta_\square(\kappa_{A'}, \overline{\tau}) \to 0$. Since $\overline{\tau}$ is an irreducible bounded kernel, the (simple graphs underlying) $\widetilde{G}_n$ satisfy the assumptions of Lemma 24, so there is a constant $c > 0$ such that for any two set V_n, V'_n of at least $\varepsilon n/2$ vertices of $\widetilde{G}_n$, the probability that V_n and V'_n are *not* joined by a path in $\widetilde{G}_n$ is at most e^{-cn}.

Recall that we have conditioned on G'_n, assuming (20). Suppose also that $C_1(G_n) \leq (\rho(\underset{\sim}{\kappa}) - 2\varepsilon)n$. Then there is a partition (V_1, V_2) of the set of vertices of G'_n in large components in G'_n with $|V_1|, |V_2| \geq \varepsilon n$ such that there is no path in G_n from V_1 to V_2. Let us call such partition (V_1, V_2) a *bad partition*. Having conditioned on G'_n, noting that in any potential bad partition V_1 must be a union of large components of G'_n, the number of possible choices for (V_1, V_2) is at most $2^{n/\omega} = e^{o(n)}$. On the other hand, since $\widetilde{G}_n \subseteq G_n$, the probability that any given partition is bad is at most e^{-cn}, so the expected number of bad partitions is $o(1)$, and whp there is no bad partition. Thus $C_1(G_n) > (\rho(\underset{\sim}{\kappa}) - 2\varepsilon)n$ holds whp, as required. □

Remark 26. The restriction to irreducible kernel families in Theorem 5 is of course necessary; roughly speaking, if $\underset{\sim}{\kappa}$ is reducible, then our graph $G(n, \underset{\sim}{\kappa})$ falls into two or more parts. Lemma 23 still applies to show that we have $\rho(\underset{\sim}{\kappa})n + o_{\mathrm{p}}(n)$ vertices in large components, but it may be that two or more parts have giant components, each of smaller order than $\rho(\underset{\sim}{\kappa})n$.

More precisely, let $\underset{\sim}{\kappa}$ be a reducible, integrable kernel family. Thus the edge kernel κ_{e} is reducible. By Lemma 5.17 of [10], there is a partition $\mathcal{S} = \bigcup_{i=0}^N \mathcal{S}_i$, $N \leq \infty$, of our ground space $\mathcal{S}$ (usually $[0,1]$) such that each $\mathcal{S}_i$ is measurable, the restriction of κ_{e} to $\mathcal{S}_i$ is irreducible (in the natural sense), and, apart from a measure zero set, κ_{e} is zero off $\bigcup_{i=1}^N \mathcal{S}_i \times \mathcal{S}_i$.

Suppressing the dependence on n, let G_i be the subgraph of $G(n, \underset{\sim}{\kappa})$ induced by the vertices with types in $\mathcal{S}_i$. Since the vertex types are iid, the probability that $G(n, \underset{\sim}{\kappa})$ contains any edges other than those of $\bigcup_{i \geq 1} G_i$ is 0. Now G_i has a random number n_i of vertices, with a binomial $\mathrm{Bi}(n, \mu(\mathcal{S}_i))$ distribution, which is concentrated around its mean. Given n_i, the graph G_i is another instance of our model.

Let $a_i = \int_{\mathcal{S}_i} \rho_{\underset{\sim}{\kappa}}(x)\, d\mu(x)$, so that $\sum_i a_i = \rho(\underset{\sim}{\kappa}) < 1$. From the remarks above it is easy to check that Theorem 5 gives $C_1(G_i)/n \xrightarrow{\mathrm{p}} a_i$ and $C_2(G_i) = o_{\mathrm{p}}(n)$; we omit the details. Sorting the a_i into decreasing order $\hat{a}_1, \hat{a}_2, \ldots$, it follows that $C_i(G(n, \underset{\sim}{\kappa})) = \hat{a}_i n + o_{\mathrm{p}}(n)$ for each fixed (finite) $1 \le i \le N$, in particular, for $i = 1$ and $i = 2$.

4.1 Disconnected atoms and percolation

One of the most studied features of the various inhomogeneous network models is their 'robustness' under random failures, and in particular, the critical point for site or bond percolation on these random graphs. For example, this property of the Barabási–Albert [5] model was studied experimentally by Barabási, Albert and Jeong [2], heuristically by Callaway, Newman, Strogatz and Watts [20] (see also [1]) and Cohen, Erez, ben-Avraham and Havlin [21], and rigorously in [14, 34]. In the present context, given $0 < p < 1$, we would like to study the random subgraphs $G^{\langle p \rangle}(n, \underset{\sim}{\kappa})$ and $G^{[p]}(n, \underset{\sim}{\kappa})$ of $G(n, \underset{\sim}{\kappa})$ obtained by deleting edges or vertices respectively, keeping each edge or vertex with probability p, independently of the others. In the edge-only model of [10], these graphs were essentially equivalent to other instances of the same model: roughly speaking, $G^{\langle p \rangle}(n, \kappa) \cong G(n, p\kappa)$ and $G^{[p]}(n, \kappa) \cong G(pn, p\kappa)$. (For precise statements, see [10, Section 4].)

Here, the situation is a little more complex. When we delete edges randomly from $G(n, \underset{\sim}{\kappa})$, it may be that what is left of a particular atom F is disconnected. This forces us to consider *generalized kernel families* $(\kappa_F)_{F \in \mathcal{G}}$ with one kernel κ_F for each $F \in \mathcal{G}$, where the set $\mathcal{G}$ consists of one representative of each isomorphism class of finite (not necessarily connected) graphs.

Rather than present a formal statement, let us consider a particular example. Suppose that $\underset{\sim}{\kappa}$ is the generalized kernel family with only one kernel κ_F, corresponding to the disjoint union F of K_3 and K_2. Let $\underset{\sim}{\kappa}'$ be the kernel family with two kernels,

$$\kappa_3(x, y, z) = \int_{\mathcal{S}^2} \kappa_F(x, y, z, u, v)\, d\mu(u)\, d\mu(v),$$

corresponding to K_3 and

$$\kappa_2(u, v) = \int_{\mathcal{S}^3} \kappa_F(x, y, z, u, v)\, d\mu(x)\, d\mu(y)\, d\mu(z)$$

for K_2. Then $G(n, \underset{\sim}{\kappa})$ and $G(n, \underset{\sim}{\kappa}')$ are clearly very similar; the main differences are that $G(n, \underset{\sim}{\kappa})$ contains exactly the same number of added triangles and K_2s, whereas in $G(n, \underset{\sim}{\kappa}')$ the numbers are only asymptotically equal, and that in $G(n, \underset{\sim}{\kappa})$ a triangle and a K_2 added in one step are necessarily disjoint. Since almost all pairs of triangles and K_2s in $G(n, \underset{\sim}{\kappa}')$ are disjoint anyway, it is not hard to check that $G(n, \underset{\sim}{\kappa})$ and $G(n, \underset{\sim}{\kappa}')$ are 'locally equivalent', in that the neighbourhoods of a random vertex in the two graphs can be coupled to agree up to a fixed size whp.

More generally, given a generalized kernel family $\underset{\sim}{\kappa} = (\kappa_F)_{F\in\mathcal{G}}$, let $\underset{\sim}{\kappa}'$ be the kernel family obtained by replacing each kernel κ_F by one kernel for each component F' of F, obtained by integrating over variables corresponding to vertices of $F \setminus F'$ as above. This may produce several new kernels for a given connected F'; we of course simply add these together to produce a single kernel $\kappa'_{F'}$. Note that

$$\sum_{F'} |F'| \int_{\mathcal{S}^{|F'|}} \kappa'_{F'} = \sum_{F} |F| \int_{\mathcal{S}^{|F|}} \kappa_F,$$

so if $\underset{\sim}{\kappa}$ is integrable, then so is $\underset{\sim}{\kappa}'$. Although $G(n, \underset{\sim}{\kappa})$ and $G(n, \underset{\sim}{\kappa}')$ are not exactly equivalent, the truncation and local approximation arguments used to prove Theorem 5 carry over easily, and we find that whenever $\underset{\sim}{\kappa}'$ is irreducible,

$$C_1(G(n, \underset{\sim}{\kappa})) = \rho(\underset{\sim}{\kappa}'')n + o_{\mathrm{p}}(n), \tag{30}$$

where $\underset{\sim}{\kappa}''$ is the hyperkernel corresponding to $\underset{\sim}{\kappa}'$, obtained by replacing each atom F by a clique as before. Note that this corresponds to replacing each *component* of an atom F in $G(n, \underset{\sim}{\kappa})$ by a clique.

Turning to bond percolation on $G(n, \underset{\sim}{\kappa})$, i.e., to the study of the random subgraph $G^{\langle p\rangle}(n, \underset{\sim}{\kappa})$ of $G(n, \underset{\sim}{\kappa})$, let $\underset{\sim}{\kappa}^{\langle p\rangle}$ be the kernel family obtained by replacing each kernel κ_F by $2^{e(F)}$ kernels $\kappa_{F'} = p^{e(F')}(1-p)^{e(F)-e(F')}\kappa_F$, one for each spanning subgraph of F. (As before, we then combine kernels corresponding to isomorphic graphs F'.) Working work with the Poisson multigraph formulation of our model, the graphs $G^{\langle p\rangle}(n, \underset{\sim}{\kappa})$ and $G(n, \underset{\sim}{\kappa}^{\langle p\rangle})$ have exactly the same distribution. This observation and (30) allow us (in principle, at least) to decide whether $G^{\langle p\rangle}(n, \underset{\sim}{\kappa})$ has a giant component, i.e., to find the critical point for bond percolation on $G(n, \underset{\sim}{\kappa})$.

Let us illustrate this with the very simple special case in which each kernel κ_F, $F \in \mathcal{G}$, is constant, say $\kappa_F = c_F$. We assume that $\underset{\sim}{\kappa}$ is integrable, i.e., that $\sum_F |F|c_F < \infty$. In this case each kernel $\kappa_F^{\langle p\rangle}$ making up $\underset{\sim}{\kappa}^{\langle p\rangle}$ is also constant, and the same applies to the hyperkernel $\underset{\sim}{\kappa}''$ corresponding to $\underset{\sim}{\kappa}^{\langle p\rangle}$. Hence, from the remarks above and (13), $G^{\langle p\rangle}(n, \underset{\sim}{\kappa})$ has a giant component if and only if the asymptotic edge density $\xi(\underset{\sim}{\kappa}'')$ of the hyperkernel $\underset{\sim}{\kappa}''$ is at least $1/2$. Since we obtain $\underset{\sim}{\kappa}''$ by first taking random subgraphs of our original atoms F, and then replacing each component by a clique, we see that

$$\xi(\underset{\sim}{\kappa}'') = \sum_{F\in\mathcal{F}} c_F \theta_F(p),$$

where $\theta_F(p)$ is the expected number of unordered pairs of distinct vertices of F that lie in the same component of the random subgraph $F^{\langle p\rangle}$ of F obtained by keeping each edge with probability p, independently of the others. Alternatively,

$$2\xi(\underset{\sim}{\kappa}'') = \sum_{F\in\mathcal{F}} c_F |F| (\chi(F^{\langle p\rangle}) - 1),$$

where $\chi(F^{\langle p\rangle})$ is the *susceptibility* of $F^{\langle p\rangle}$, i.e., the expected size of the component of a random vertex of $F^{\langle p\rangle}$. If we have only a finite number of non-zero c_F, then $\xi(\underset{\sim}{\kappa}'')$ may be evaluated as a polynomial in p, and the critical point found exactly.

Turning to site percolation, there is a similar reduction to another instance of our model, most easily described by modifying the type space. Indeed, we add a new type $\star$ corresponding to deleted vertices, and set $\mu'(\star) = 1 - p$. Setting $\mu'(A) = p\mu(A)$ for $A \subset \mathcal{S}$, we obtain a probability measure μ' on $\mathcal{S}' = \mathcal{S} \cup \{\star\}$. Replacing each kernel κ_F by $2^{|F|}$ kernels $\kappa_{F'}$ on $\mathcal{S}'$ defined appropriately (with F' corresponding to the subgraph of F spanned by the non-deleted vertices), one can show that $G^{[p]}(n, \underset{\sim}{\kappa})$ is very close to (in the Poisson version, identical to) a suitable instance $G(n', \underset{\sim}{\kappa}')$ of our model, where n' is now random but concentrated around its mean pn. In the first instance $\underset{\sim}{\kappa}'$ may include kernels for disconnected graphs, but as above we can find an asymptotically equivalent kernel family involving only connected graphs. In this way one can find the asymptotic size of any giant component in $G^{[p]}(n, \underset{\sim}{\kappa})$; we omit the mathematically straightforward but notationally complex details.

5 Vertex degrees

Heuristically, the vertex degrees in $G(n, \underset{\sim}{\kappa})$ can be described as follows. Consider a vertex v and condition on its type x_v. The number of atoms that contain v then is asymptotically Poisson with a certain mean depending on $\underset{\sim}{\kappa}$ and x_v. However, each atom may add several edges to the vertex v, and thus the asymptotic distribution of the vertex degree is compound Poisson (see below for a definition). Moreover, this compound Poisson distribution typically depends on the type x_v, so the final result is that, asymptotically, the vertex degrees have a mixed compound Poisson distribution. In this section we shall make this precise and rigorous.

We begin with some definitions. If λ is a finite measure on $\mathbb{N}$, then $\mathrm{CPo}(\lambda)$, the *compound Poisson distribution with intensity* λ, is defined as the distribution of $\sum_{j=1}^{\infty} jX_j$, where $X_j \sim \mathrm{Po}(\lambda\{j\})$ are independent Poisson random variables. Equivalently, $\mathrm{CPo}(\lambda)$ is the distribution of the sum $\sum_\nu \xi_\nu$ of the points of a Poisson process $\{\xi_\nu\}$ on $\mathbb{N}$ with intensity λ, regarded as a multiset. (The latter definition generalizes to arbitrary measures λ on $(0, \infty)$ such that $\int_0^\infty t \wedge 1 \, d\lambda(t) < \infty$, but we consider in this paper only the integer case.) Since X_j has probability generating function $\mathbb{E} z^{X_j} = e^{\lambda\{j\}(z-1)}$, $\mathrm{CPo}(\lambda)$ has probability generating function

$$\varphi_{\mathrm{CPo}(\lambda)}(z) = \mathbb{E} z^{\sum_{j=1}^{\infty} jX_j} = \mathbb{E} \prod_{j=1}^{\infty} z^{jX_j} = \prod_{j=1}^{\infty} e^{\lambda\{j\}(z^j-1)} = e^{\sum_{j=1}^{\infty} \lambda\{j\}(z^j-1)},$$

whenever this is defined, which it certainly is for $|z| \le 1$.

If Λ is a random finite measure on $\mathbb{N}$, then $\mathrm{MCPo}(\Lambda)$ denotes the corresponding *mixed compound Poisson distribution*. From now on, for each $x \in \mathcal{S}$, λ_x will be a finite measure on $\mathbb{N}$, depending measurably on x. We shall write Λ for the corresponding random measure on $\mathbb{N}$, obtained by choosing x from $\mathcal{S}$

according to the distribution μ and then taking λ_x. Thus $\mathrm{MCPo}(\Lambda)$ is defined by the point probabilities

$$\mathrm{MCPo}(\Lambda)\{i\} = \int_{\mathcal{S}} \mathrm{CPo}(\lambda_x)\{i\}\, d\mu(x)$$

or, equivalently, the probability generating function

$$\varphi_{\mathrm{MCPo}(\Lambda)}(z) = \int_{\mathcal{S}} \varphi_{\mathrm{CPo}(\lambda_x)}(z)\, d\mu(x) = \int_{\mathcal{S}} e^{\sum_{j=1}^{\infty} \lambda_x\{j\}(z^j-1)}\, d\mu(x).$$

Remark 27. Since we have assumed that λ is a finite measure, $\mathbb{E}\sum_j X_j = \lambda(\mathbb{N}) < \infty$; thus a.s. $\sum_j X_j < \infty$ and only finitely many X_j are non-zero, whence $\sum_j jX_j < \infty$ a.s. This verifies that $\mathrm{CPo}(\lambda)$ is a proper probability distribution. On the other hand, the mean of $\mathrm{CPo}(\lambda)$ is

$$\mathbb{E}\,\mathrm{CPo}(\lambda) = \sum_j j\mathbb{E}X_j = \sum_j j\lambda\{j\} = \int_0^\infty t\, d\lambda(t),$$

which may be infinite. As a consequence,

$$\mathbb{E}\,\mathrm{MCPo}(\Lambda) = \int_{\mathcal{S}} \int_0^\infty t\, d\lambda_x(t)\, d\mu(x) \le \infty. \tag{31}$$

Let d_{TV} denote the total variation distance between two random variables, or rather their probability distributions, defined by

$$d_{\mathrm{TV}}(X,Y) = \sup_A |\mathbb{P}(X \in A) - \mathbb{P}(Y \in A)|, \tag{32}$$

where the supremum is taken over all measurable sets $A \subseteq \mathbb{R}$. We shall use the following trivial upper bound on the total variation distance between two compound Poisson distributions.

Lemma 28. *If λ and λ' are two finite measures on $\mathbb{N}$, then*

$$d_{\mathrm{TV}}\big(\mathrm{CPo}(\lambda), \mathrm{CPo}(\lambda')\big) \le \|\lambda - \lambda'\| = \sum_j |\lambda\{j\} - \lambda'\{j\}|.$$

Proof. Let $X_j \sim \mathrm{Po}(\lambda\{j\})$ be as above and let $X'_j \sim \mathrm{Po}(\lambda'\{j\})$ be another family of independent Poisson variables. We can easily couple the families so that $\mathbb{P}(X_j \ne X'_j) \le |\lambda\{j\} - \lambda'\{j\}|$ for every j.

Then

$$\begin{aligned} d_{\mathrm{TV}}\big(\mathrm{CPo}(\lambda), \mathrm{CPo}(\lambda')\big) &\le \mathbb{P}\Big(\sum_j jX_j \ne \sum_j jX'_j\Big) \le \sum_j \mathbb{P}(X_j \ne X'_j) \\ &= \sum_j |\lambda\{j\} - \lambda'\{j\}|. \end{aligned}$$

$\square$

Given an integrable kernel family $\underset{\sim}{\kappa}$ and $x \in \mathcal{S}$, $F \in \mathcal{F}$ and $j \in V(F) = [|F|]$, let

$$\lambda_{F,j}(x) = \int_{\mathcal{S}^{|F|-1}} \kappa_F(x_1, \dots, x_{j-1}, x, x_{j+1}, \dots, x_{|F|}) \, d\mu(x_1) \cdots d\mu(x_{j-1}) \, d\mu(x_{j+1}) \cdots d\mu(x_{|F|}) \tag{33}$$

be the (asymptotic) expected number of added copies of F containing a given vertex of type x in which the given vertex corresponds to vertex j in F. Let $d_F(j)$ be the degree of vertex j in F, and define the measure

$$\lambda_x = \sum_{F \in \mathcal{F}} \sum_{j \in V(F)} \lambda_{F,j}(x) \delta_{d_F(j)}, \tag{34}$$

where, as usual, δ_d denotes the probability measure assigning mass 1 to d. Thus λ_x is a measure on $\mathbb{N}$, with point masses given by

$$\lambda_x\{d\} = \sum_{F \in \mathcal{F}} \sum_{j : d_F(j) = d} \lambda_{F,j}(x), \tag{35}$$

the (asymptotic) expected number of atoms containing a given vertex of type x and having degree d there. From (33), $\int_{\mathcal{S}} \lambda_{F,j}(x) \, d\mu(x) = \int_{\mathcal{S}^{|F|}} \kappa_F$, and thus by (34)

$$\int_{\mathcal{S}} \|\lambda_x\| \, d\mu(x) = \sum_{F,j} \int_{\mathcal{S}} \lambda_{F,j}(x) \, d\mu(x) = \sum_F |F| \int_{\mathcal{S}^{|F|}} \kappa_F < \infty.$$

Consequently, λ_x is a finite measure on $\mathbb{N}$ for a.e. x, and the mixed compound Poisson distribution $\mathrm{MCPo}(\Lambda)$ is defined.

Let the random variable $D = D_n$ be the degree of any fixed vertex in $G(n, \underset{\sim}{\kappa})$. Equivalently, by symmetry, we can take D_n to be the degree of a uniformly random vertex. Furthermore, for $\ell \geq 0$, let n_ℓ be the number of vertices with degree ℓ in $G(n, \underset{\sim}{\kappa})$. Then the random sequence $(n_\ell/n)_{\ell=0}^\infty$ can be regarded as a (random) probability distribution, viz., the conditional distribution of the degree of a random vertex in $G(n, \underset{\sim}{\kappa})$, given this random graph. Note that $\mathbb{P}(D_n = \ell) = \mathbb{E} n_\ell / n$.

Theorem 29. *Suppose that $\underset{\sim}{\kappa} = (\kappa_F)_{F \in \mathcal{F}}$ is an integrable kernel family. Then, as $n \to \infty$,*

(i) $D_n \xrightarrow{\mathrm{d}} \mathrm{MCPo}(\Lambda)$, *and*

(ii) $\mathbb{E} D_n \to \mathbb{E}\, \mathrm{MCPo}(\Lambda) = \sum_{F \in \mathcal{F}} 2e(F) \int_{\mathcal{S}^{|F|}} \kappa_F = 2\xi(\underset{\sim}{\kappa}) \leq \infty.$

(iii) *Moreover, for every fixed ℓ,*

$$n_\ell = \mathrm{MCPo}(\Lambda)\{\ell\} n + o_{\mathrm{p}}(n) \tag{36}$$

and thus $(n_\ell/n)_{\ell=0}^\infty \xrightarrow{\mathrm{d}} \mathrm{MCPo}(\Lambda)$ in the space of probability measures on $\mathbb{N}$.

Note that the limit distribution exists for every integrable kernel family, but has finite expectation only if the kernel family is edge integrable.

As usual, Theorem 29 applies to the variants of the model $G(n, \underset{\sim}{\kappa})$ discussed in Section 1. In the proof, we shall mostly work with the (non-Poisson) multi-graph form, where we add at most one copy of a certain small graph F with a particular vertex set, but keep any resulting multiple edges.

Proof. Assume first that $\underset{\sim}{\kappa}$ is a bounded kernel family, with $\kappa_F \leq M$ and $\kappa_F = 0$ if $|F| > M$. Fix a vertex $v \in [n]$, and let D be the degree of v. For $F \in \mathcal{F}$ with $|F| \leq M$ and $j \in V(F)$, let $N_{F,j}$ be the number of added copies of F that contain v with v corresponding to vertex j in F. Let

$$D' = \sum_{F,j} N_{F,j} d_F(j); \tag{37}$$

this is the number of edges added to v, including possible repetitions. Thus $D = D'$ unless two added edges with endpoint v coincide. For any other vertex w, conditioned on the types $\mathbf{x} = (x_1, \dots, x_n)$, the number of atoms containing both v and w is a sum $\sum_\nu I_\nu$ of independent Bernoulli variables $I_\nu \sim \mathrm{Be}(p_\nu)$, for ν in some index set. For each $r = 2, \dots, M$ there are $O(n^{r-2})$ such variables, each with $p_\nu = O(n^{1-r})$. Hence,

$$\mathbb{P}\Big(\sum_\nu I_\nu \geq 2 \;\Big|\; \mathbf{x}\Big) \leq \sum_{\nu_1 \neq \nu_2} p_{\nu_1} p_{\nu_2} \leq \Big(\sum_\nu p_\nu\Big)^2 = O(n^{-2}).$$

Since there are $n-1$ possible choices for w, it follows that

$$d_{\mathrm{TV}}\big((D \mid \mathbf{x}),\, (D' \mid \mathbf{x})\big) \leq \mathbb{P}(D \neq D' \mid \mathbf{x}) = O(n^{-1}). \tag{38}$$

Hence, in proving (i), it makes no difference whether we work with D' or with D, i.e., with the multi-graph or simple graph version of $G(n, \underset{\sim}{\kappa})$.

Conditioned on $\mathbf{x}$, $N_{F,j}$ is a sum of independent Bernoulli variables $\mathrm{Be}(p_{F,j,\alpha}(\mathbf{x}))$ for α in some index set $\mathcal{A}_{F,j}$, with $p_{F,j,\alpha}(\mathbf{x}) = O(n^{1-|F|})$ given by (1) and $|\mathcal{A}_{F,j}| = O(n^{|F|-1})$.

Let $\widehat{\lambda}_{F,j}(\mathbf{x}) = \mathbb{E}(N_{F,j} \mid \mathbf{x}) = \sum_\alpha p_{F,j,\alpha}(\mathbf{x})$. By a classical Poisson approximation theorem (see [6, (1.8)]),

$$d_{\mathrm{TV}}\big((N_{F,j} \mid \mathbf{x}),\, \mathrm{Po}(\widehat{\lambda}_{F,j}(\mathbf{x}))\big) \leq \sum_\alpha p_{F,j,\alpha}(\mathbf{x})^2 = O(n^{1-|F|}) = O(n^{-1}). \tag{39}$$

(This follows easily from the elementary $d_{\mathrm{TV}}(\mathrm{Be}(p), \mathrm{Po}(p)) \leq p^2$; see e.g. [6, page 4 and Theorem 2.M] for history and further results.) Furthermore, given $\mathbf{x}$, the random variables $N_{F,j}$ are independent, and thus (37) and (39) imply that if $\widehat{X}_{F,j} \sim \mathrm{Po}(\widehat{\lambda}_{F,j}(\mathbf{x}))$ are independent, then

$$d_{\mathrm{TV}}\Big((D' \mid \mathbf{x}),\, \sum_{F,j} d_F(j) \widehat{X}_{F,j}\Big) = O(n^{-1}).$$

Since $\sum_{F,j} d_F(j)\widehat{X}_{F,j}$ has a compound Poisson distribution $\mathrm{CPo}(\widehat{\lambda}(\mathbf{x}))$ with intensity $\widehat{\lambda}(\mathbf{x}) = \sum_{F,j} \widehat{\lambda}_{F,j}(\mathbf{x})\delta_{d_F(j)}$, we have

$$d_{\mathrm{TV}}\bigl((D' \mid \mathbf{x}),\, \mathrm{CPo}(\widehat{\lambda}(\mathbf{x}))\bigr) = O(n^{-1}).$$

By (38) and Lemma 28, this yields

$$d_{\mathrm{TV}}\bigl((D \mid \mathbf{x}),\, \mathrm{CPo}(\lambda_{x_v})\bigr) \le O(n^{-1}) + \|\widehat{\lambda}(\mathbf{x}) - \lambda_{x_v}\|.$$

In particular, for every $\ell \in \mathbb{N}$, taking $A = \{\ell\}$ in (32),

$$\bigl|\mathbb{P}(D = \ell \mid \mathbf{x}) - \mathrm{CPo}(\lambda_{x_v})\{\ell\}\bigr| \le O(n^{-1}) + \|\widehat{\lambda}(\mathbf{x}) - \lambda_{x_v}\|. \tag{40}$$

Taking the expectation of both sides, and noting that $\mathbb{E}\mathbb{P}(D = \ell \mid \mathbf{x}) = \mathbb{P}(D = \ell)$ and $\mathbb{E}\,\mathrm{CPo}(\lambda_{x_v})\{\ell\} = \mathrm{MCPo}(\Lambda)\{\ell\}$, we find that

$$\bigl|\mathbb{P}(D = \ell) - \mathrm{MCPo}(\Lambda)\{\ell\}\bigr| \le O(n^{-1}) + \mathbb{E}\|\widehat{\lambda}(\mathbf{x}) - \lambda_{x_v}\|. \tag{41}$$

We shall show that the final term is small.

By (1), with $r = |F|$,

$$\widehat{\lambda}_{F,j}(\mathbf{x}) = n^{1-r} \sum \kappa_F(x_{v_1}, \dots, x_{v_r}),$$

where the sum runs over all $(n-1)_{(r-1)}$ sequences $v_1, \dots, v_r$ of distinct elements in $[n]$ with $v_j = v$. Consequently, by (33),

$$\mathbb{E}\bigl(\widehat{\lambda}_{F,j}(\mathbf{x}) \mid x_v\bigr) = \bigl(1 - O(n^{-1})\bigr)\lambda_{F,j}(x_v). \tag{42}$$

Recalling that $\underset{\sim}{\kappa}$ is bounded, it is easy to check (as in the similar argument in the proof of Lemma 25) that

$$\mathrm{Var}\bigl(\widehat{\lambda}_{F,j}(\mathbf{x}) \mid x_v\bigr) = \mathbb{E}\bigl((\widehat{\lambda}_{F,j}(\mathbf{x}) - \mathbb{E}(\widehat{\lambda}_{F,j}(\mathbf{x}) \mid x_v))^2 \mid x_v\bigr) = O(n^{-1})$$

and thus, by the Cauchy–Schwarz inequality and (42),

$$\mathbb{E}\bigl(|\widehat{\lambda}_{F,j}(\mathbf{x}) - \lambda_{F,j}(x_v)| \bigm| x_v\bigr) = O(n^{-1/2}).$$

Consequently, using again that $\underset{\sim}{\kappa}$ is bounded,

$$\mathbb{E}\|\widehat{\lambda}(\mathbf{x}) - \lambda_{x_v}\| \le \mathbb{E}\sum_{F,j} |\widehat{\lambda}_{F,j}(\mathbf{x}) - \lambda_{F,j}(x_v)| = O(n^{-1/2}) = o(1), \tag{43}$$

so

$$\|\widehat{\lambda}(\mathbf{x}) - \lambda_{x_v}\| \xrightarrow{\mathrm{p}} 0.$$

Combining (43) and (41) we see that $\mathbb{P}(D = \ell) \to \mathrm{MCPo}(\Lambda)\{\ell\}$ for every ℓ, i.e., $D \xrightarrow{\mathrm{d}} \mathrm{MCPo}(\Lambda)$, which proves (i) for bounded $\underset{\sim}{\kappa}$.

Next we turn to the proof of (iii), assuming still that $\underset{\sim}{\kappa}$ is bounded. Fix a number $\ell \in \mathbb{N}$, and for $v \in [n]$ let D_v be the degree of v in $G(n, \underset{\sim}{\kappa})$, and I_v the indicator $\mathbf{1}[D_v = \ell]$.

Fix two distinct vertices v and w, let $\mathcal{G}$ be the set of atoms that contain both v and w, and let $\widetilde{D}_v$ and $\widetilde{D}_w$ be the degrees of the vertices if we delete (or ignore) the atoms in $\mathcal{G}$. Since $\underset{\sim}{\kappa}$ bounded, the expected number $\mathbb{E}|\mathcal{G}|$ of such exceptional atoms is $O(n^{-1})$, and thus

$$\mathbb{P}(D_v \neq \widetilde{D}_v),\, \mathbb{P}(D_w \neq \widetilde{D}_w) \leq \mathbb{P}(\mathcal{G} \neq \emptyset) \leq \mathbb{E}|\mathcal{G}| = O(n^{-1}).$$

Moreover, these bounds hold conditional on $\mathbf{x}$. Furthermore, given $\mathbf{x}$, $\widetilde{D}_v$ and $\widetilde{D}_w$ are independent. Consequently, for any $\ell \in \mathbb{N}$,

$$\begin{aligned}
\mathbb{E}(I_v I_w \mid \mathbf{x}) &= \mathbb{P}(D_v = D_w = \ell \mid \mathbf{x}) = \mathbb{P}(\widetilde{D}_v = \widetilde{D}_w = \ell \mid \mathbf{x}) + o(1) \\
&= \mathbb{P}(\widetilde{D}_v = \ell \mid \mathbf{x})\mathbb{P}(\widetilde{D}_w = \ell \mid \mathbf{x}) + o(1) \\
&= \mathbb{P}(D_v = \ell \mid \mathbf{x})\mathbb{P}(D_w = \ell \mid \mathbf{x}) + o(1) \\
&= \mathbb{E}(I_v \mid \mathbf{x})\mathbb{E}(I_w \mid \mathbf{x}) + o(1),
\end{aligned}$$

and thus $\mathrm{Cov}(I_v, I_w \mid \mathbf{x}) = o(1)$. Since $n_\ell = \sum_v I_v$, it follows that $\mathrm{Var}(n_\ell \mid \mathbf{x}) = o(n^2)$ and thus

$$n_\ell = \mathbb{E}(n_\ell \mid \mathbf{x}) + o_{\mathrm{p}}(n). \tag{44}$$

Further, if we write $h(x) = \mathrm{CPo}(\lambda_x)\{l\}$ and sum (40) (where $D = D_v$) over v, we obtain

$$\Bigl|\mathbb{E}(n_\ell \mid \mathbf{x}) - \sum_{v=1}^{n} h(x_v)\Bigr| = \Bigl|\sum_{v=1}^{n}\bigl(\mathbb{P}(D_v = \ell \mid \mathbf{x}) - h(x_v)\bigr)\Bigr| \leq O(1) + \sum_{v=1}^{n} \mathbb{E}\|\widehat{\lambda}(\mathbf{x}) - \lambda_{x_v}\|.$$

By (43), the right-hand side has expectation $o(n)$ and thus

$$\mathbb{E}(n_\ell \mid \mathbf{x}) = \sum_{v=1}^{n} h(x_v) + o_{\mathrm{p}}(n). \tag{45}$$

Now $h(x_1), \dots, h(x_n)$ are iid random variables with mean

$$\mathbb{E}h(x_v) = \int_{\mathcal{S}} h(x)\, d\mu(x) = \int_{\mathcal{S}} \mathrm{CPo}(\lambda_x)\{l\}\, d\mu(x) = \mathrm{MCPo}(\Lambda)\{l\}.$$

Hence, by the law of large numbers, $\frac{1}{n}\sum_{v=1}^{n} h(x_v) \xrightarrow{\mathrm{p}} \mathrm{MCPo}(\Lambda)\{l\}$, which is the same as

$$\sum_{v=1}^{n} h(x_v) = \mathrm{MCPo}(\Lambda)\{l\}n + o_{\mathrm{p}}(n). \tag{46}$$

The result (36) follows from (44), (45), (46).

Furthermore, (36) says that $(n_\ell/n)_\ell \xrightarrow{\mathrm{p}} \mathrm{MCPo}(\Lambda)$ in the space $\mathbb{R}^\infty$ of sequences, equipped with the product topology, which is the same as separate convergence of the components. However, it is well-known, and easy to see (e.g. by compactness) that restricted to the set of probability distributions, this equals the standard topology there.

We have proved (i) and (iii) for bounded $\underset{\sim}{\kappa}$. For general $\underset{\sim}{\kappa}$ we use truncations: define κ_F^M in analogy with (17), setting $\kappa_F^M = \kappa_F \wedge M$ for $|F| \le M$ and $\kappa_F^M = 0$ for $|F| > M$. We use λ^M, n_ℓ^M and so on to denote quantities defined for $G(n, \underset{\sim}{\kappa}^M)$. For fixed M, applying (36) for the bounded kernel family $\underset{\sim}{\kappa}^M$, we have $n_\ell^M/n \xrightarrow{\mathrm{p}} \mathrm{MCPo}(\Lambda^M)\{\ell\}$ as $n \to \infty$, and thus by dominated convergence

$$\mathbb{E}\big|n_\ell^M/n - \mathrm{MCPo}(\Lambda^M)\{\ell\}\big| \to 0. \tag{47}$$

Furthermore, for every $x \in \mathcal{S}$ and $d \ge 1$, the intensities $\lambda_x^M\{d\}$ converge to $\lambda_x\{d\}$ as $M \to \infty$, by (35), (33) and monotone convergence. Thus a simple coupling shows that $\mathrm{MCPo}(\Lambda^M) \xrightarrow{\mathrm{d}} \mathrm{MCPo}(\Lambda)$ as $M \to \infty$. We may couple $G(n, \underset{\sim}{\kappa})$ and $G(n, \underset{\sim}{\kappa}^M)$ in the obvious way so that $G(n, \underset{\sim}{\kappa})$ is obtained from $G(n, \underset{\sim}{\kappa}^M)$ by adding further atoms, say N_F^M copies of each $F \in \mathcal{F}$. Then $\mathbb{E}N_F^M \le n \int_{\mathcal{S}^{|F|}} (\kappa_F - \kappa_F^M)$, and since at most $\sum_F |F| N_F^M$ vertices are affected by the extra additions,

$$\mathbb{E}\Big|\frac{n_\ell}{n} - \frac{n_\ell^M}{n}\Big| \le \frac{1}{n}\mathbb{E}\sum_F |F| N_F^M \le \sum_F |F| \int_{\mathcal{S}^{|F|}} (\kappa_F - \kappa_F^M). \tag{48}$$

The right hand side is independent of n, and tends to 0 as $M \to \infty$ by dominated convergence and our assumption that $\underset{\sim}{\kappa}$ is integrable. For any $\varepsilon > 0$, we may thus choose M so large that the right hand side of (48) is less than ε, and also so that $|\mathrm{MCPo}(\Lambda^M)\{\ell\} - \mathrm{MCPo}(\Lambda)\{\ell\}| < \varepsilon$; then by (47), for large enough n,

$$\mathbb{E}\big|n_\ell/n - \mathrm{MCPo}(\Lambda)\{\ell\}\big| < 3\varepsilon,$$

which proves (36) and thus (iii). Further, (36) and dominated convergence yields $\mathbb{P}(D_n = \ell) = \mathbb{E}(n_\ell/n) \to \mathrm{MCPo}(\Lambda)\{\ell\}$, which proves (i).

Finally we prove (ii). (This could also easily be done directly in a fairly straightforward way.) First, (31) and (35) yield

$$\mathbb{E}\,\mathrm{MCPo}(\Lambda) = \int_{\mathcal{S}} \sum_{F,j} \lambda_{F,j}(x) d_F(j)\, d\mu(x) = \sum_{F \in \mathcal{F}} \sum_{j \in V(F)} d_F(j) \int_{\mathcal{S}^{|F|}} \kappa_F,$$

which yields the formula for $\mathbb{E}\,\mathrm{MCPo}(\Lambda)$ claimed in the theorem, since $\sum_j d_F(j) = 2e(F)$.

Next, the convergence in distribution (i) yields (by a version of Fatou's Lemma) the inequality $\liminf_{n\to\infty} \mathbb{E}D_n \ge \mathbb{E}\,\mathrm{MCPo}(\Lambda)$. Finally, recalling the definition (37) of D_n' (denoted D' in (37)), we have $D_n \le D_n'$ and thus

$$\begin{aligned} \mathbb{E}D_n \le \mathbb{E}D_n' &= \sum_{F,j} d_F(j)\mathbb{E}N_{F,j} = \sum_{F \in \mathcal{F}} \sum_{j \in V(F)} d_F(j) \frac{(n-1)_{(|F|-1)}}{n^{|F|-1}} \int_{\mathcal{S}^{|F|}} \kappa_F \\ &\le \sum_{F \in \mathcal{F}} 2e(F) \int_{\mathcal{S}^{|F|}} \kappa_F = \mathbb{E}\,\mathrm{MCPo}(\Lambda), \end{aligned}$$

yielding the opposite inequality $\limsup_{n\to\infty} \mathbb{E}D_n \le \mathbb{E}\,\mathrm{MCPo}(\Lambda)$. $\square$

Part (ii) of Theorem 29 is not surprising. Also, since by symmetry $\mathbb{E}D_n = \frac{2}{n}\mathbb{E}e(G(n, \underset{\sim}{\kappa}))$, it follows from Theorem 3 (which we shall not prove until the next section). For bounded kernel families, it is easy to see that also higher moments of D_n converge to the corresponding moments of $\mathrm{MCPo}(\Lambda)$, for example by first showing that $\mathbb{E}D_n^m = O(1)$ for every fixed m and combining this with (i). This extends to certain unbounded kernel families, but somewhat surprisingly not to all integrable kernel families, as the following example shows.

Example 30. Let $\mathcal{S} = [0, 1)$ with Lebesgue measure, and regard $\mathcal{S}$ as a circle with the usual metric $d(x, y) = \min(|x - y|, 1 - |x - y|)$. We construct our random graph by adding triangles only; thus $\kappa_F = 0$ for $F \neq K_3$, and we take

$$\kappa_3(x, y, z) = d(x, y)^{\varepsilon-1} + d(x, z)^{\varepsilon-1} + d(y, z)^{\varepsilon-1} \tag{49}$$

for some small $\varepsilon > 0$, for example $\varepsilon = 1/10$. Clearly, $\underset{\sim}{\kappa}$ is an integrable kernel family (and a hyperkernel).

Let $\Delta = \min_{1 \leq i < j \leq n} d(x_i, x_j)$ be the minimal spacing between the n independent uniformly distributed random points x_i, $1 \leq i \leq n$. It is well-known that this minimal spacing is of order n^{-2}; in fact, it is easy to see that for $0 \leq s \leq 1/n$ we have $\mathbb{P}(\Delta > s) = (1 - sn)^{n-1} \leq e^{-sn(n-1)}$, and in particular $\Delta \leq n^{\varepsilon-2}$ whp. Hence there exist whp two distinct indices i and j with $d(x_i, x_j) \leq n^{\varepsilon-2}$, and thus, for large n and every x_k, $\kappa_3(x_i, x_j, x_k) \geq n^{(\varepsilon-2)(\varepsilon-1)} \geq 2n^{2-3\varepsilon}$. If i and j are chosen such that this holds, then from (1) we have $p(i, j, k; K_3) \geq 2n^{-3\varepsilon}$ for all $k \neq i, j$, and thus the number of k such that the triangle ijk is an atom stochastically dominates the binomial distribution $\mathrm{Bi}(n - 2, 2n^{-3\varepsilon})$; hence this number is whp at least $n^{1-3\varepsilon}$.

We have shown that whp there are at least two vertices i and j with degrees $\geq n^{1-3\varepsilon}$, and thus, for large n, $\mathbb{P}(D_n \geq n^{1-3\varepsilon}) \geq (1-o(1))\frac{2}{n} \geq \frac{1}{n}$. Consequently, for large n,

$$\mathbb{E}D_n^2 \geq \frac{1}{n} n^{2(1-3\varepsilon)} = n^{1-6\varepsilon} \to \infty.$$

On the other hand, for some finite $c = \int_{\mathcal{S}^3} \kappa_3$, by symmetry, $\lambda_{K_3,j}(x) = c$ and $\lambda_x = 3c\delta_2$. Hence $\mathrm{MCPo}(\Lambda) = \mathrm{CPo}(3c\delta_2)$, which is the distribution of $2X$ with $X \sim \mathrm{Po}(3c)$, which has all moments finite.

As we shall see in Theorem 34, this situation cannot arise in the edge-only version of the model, i.e., the model in [10]; in the terminology of the next section, all copies of P_2 are then 'regular'.

In Section 7 we shall illustrate Theorem 29 by giving a natural family of examples with degree distributions with power-law tails.

6 Small subgraphs

In this section we turn to the final general property of $G(n, \underset{\sim}{\kappa})$ we shall study, the asymptotic number of copies of a fixed graph F in $G(n, \underset{\sim}{\kappa})$; throughout this section, $\underset{\sim}{\kappa}$ denotes a kernel family $(\kappa_F)_{F\in\mathcal{F}}$, rather than a hyperkernel. We work with the multi-graph version of the model.

Although mathematically not as interesting as the phase transition, the number of small graphs in $G(n, \underset{\sim}{\kappa})$ is important as it is directly related to the original motivation for the model. Indeed, recall that perhaps the main defect of the model of [10], i.e., the edge-only case of the present model, is that it produces graphs with very few ($O_{\mathrm{p}}(1)$) triangles, i.e., graphs with clustering coefficients that are essentially zero. This contrasts strongly with many of the real-world networks we wish to model.

The simplest way that a copy of some graph F may arise in $G(n, \underset{\sim}{\kappa})$ is as an atom. The expected number of such copies is simply

$$\frac{n_{(|F|)}}{n^{|F|-1}} \int_{\mathcal{S}^{|F|}} \kappa_F \leq n \int_{\mathcal{S}^{|F|}} \kappa_F. \tag{50}$$

The next simplest way that a copy of F may arise is as a subgraph of some atom F' of $G(n, \underset{\sim}{\kappa})$. Let us call such copies of F *direct*. Let $n(F, F')$ denote the number of subgraphs of F' isomorphic to F, so $n(K_3, K_4) = 4$, for example. Set

$$\tilde{t}_1(F, \underset{\sim}{\kappa}) = \sum_{F' \in \mathcal{F}} n(F, F') \int \kappa_{F'} \leq \infty,$$

and let $n_{\mathrm{d}}(F, G(n, \underset{\sim}{\kappa}))$ denote the number of direct copies of F in $G(n, \underset{\sim}{\kappa})$. Then from (50) we see that

$$\mathbb{E} n_{\mathrm{d}}(F, G(n, \underset{\sim}{\kappa})) \leq \tilde{t}_1(F, \underset{\sim}{\kappa}) n,$$

and that if $\underset{\sim}{\kappa}$ is bounded, then

$$\mathbb{E} n_{\mathrm{d}}(F, G(n, \underset{\sim}{\kappa})) = \tilde{t}_1(F, \underset{\sim}{\kappa}) n + O(1).$$

The reason for the somewhat peculiar notation $\tilde{t}_1$ is as follows: the subscript 1 indicates direct copies (arising from only one atom). The tilde will be useful later to differentiate from standard notation $t(F, \kappa)$ in other contexts.

It will turn out that in well behaved cases (for example for all bounded kernel families), essentially all copies of any 2-connected graph F in $G(n, \underset{\sim}{\kappa})$ arise directly. Unfortunately, this is not the case for general F. Perhaps the main special cases we are interested in are stars; the number of copies of the star $K_{1,2}$ (i.e., the path P_2) is needed to calculate the clustering coefficient, for example. Note that the number of copies of the star $K_{1,k}$ ($k \geq 2$) in any graph G is simply $|G|/k!$ times the kth factorial moment of the degree of a random vertex; hence counting stars is closely related to studying the degree distribution, which we did for $G(n, \underset{\sim}{\kappa})$ in Section 5.

Let us say that a copy of F in $G(n, \underset{\sim}{\kappa})$ arises *indirectly* if it contains edges of at least two of the atoms making up $G(n, \underset{\sim}{\kappa})$. To understand the expected number of such copies we first need to understand the probability that a certain set of vertices form a copy of F *given the types of the vertices*. More precisely, we consider the expectation of the number of copies of F with a given vertex set, even though this number is highly unlikely to exceed 1.

Let F be a connected graph with r vertices. Let $\operatorname{emb}(F,F')$ denote the number of embeddings of F into F', i.e., the number of injective homomorphisms from F to F', so $\operatorname{emb}(F,F') = n(F,F')\operatorname{aut}(F)$. Fixing a labelling of F, let $X_F^0(G)$ denote the number of copies of F in a multigraph G with vertex i of F corresponding to vertex i of G. (Thus $X_F^0(G)$ is 0 or 1 if G is simple.) The contribution to $\mathbb{E}X_F^0(G(n,\underset{\sim}{\kappa}))$ from copies of F arising as subgraphs of atoms isomorphic to a given F' with r' vertices is exactly

$$\sum_{\varphi:F\to F'} \frac{(n-r)_{(r'-r)}}{n^{r'-1}} \int_{\mathcal{S}^{r'-r}} \kappa_{F'}(y_1,\dots,y_{r'}),$$

where φ runs over all $\operatorname{emb}(F,F')$ embeddings of F into F', we take $y_j = x_i$ if $\varphi(i) = j$, and we integrate over the remaining $r'-r$ variables y_j.

Set

$$\sigma_F(x_1,\dots,x_r) = \sigma_F(x_1,\dots,x_r;\underset{\sim}{\kappa}) = \sum_{F'} \sum_{\varphi:F\to F'} \int_{\mathcal{S}^{r'-r}} \kappa_{F'}(y_1,\dots,y_{r'}). \tag{51}$$

Then we have

$$\mathbb{E}\big(X_F^0(G(n,\underset{\sim}{\kappa})) \bigm| x_1,\dots,x_r\big) \le n^{-(r-1)}\sigma_F(x_1,\dots,x_r;\underset{\sim}{\kappa}), \tag{52}$$

and if $\underset{\sim}{\kappa}$ is bounded then the relative error is $O(n^{-1})$.

Comparing (51) and (3), note that if $F = K_2$, then $\sigma_F = \kappa_{\mathrm{e}}$. Before continuing, let us comment on the normalization. Recall that in defining $G(n,\underset{\sim}{\kappa})$, we consider all $r!$ possible ways of adding a (labelled) copy of F on vertex set $\{1,2,\dots,r\}$, say, adding each copy with probability $\kappa_F(x_1,\dots,x_r)/n^{r-1}$. This means that the contribution from κ_F to $X_F^0(G(n,\underset{\sim}{\kappa}))$ is $\operatorname{aut}(F)\kappa_F(x_1,\dots,x_r)/n^{r-1}$, and, correspondingly, the contribution from κ_F to σ_F is $\operatorname{aut}(F)\kappa_F(x_1,\dots,x_r)$. In other words, while having the same form as a kernel, σ_F is normalized differently. This situation arises already in the edge-only case, where $\kappa_{\mathrm{e}}(x,y) = 2\kappa_2(x,y)$. It turns out that here the normalization of σ_F, giving directly the probability that a certain set of edges forming a copy of F is present, is the natural one. Note that if we had used this normalization from the beginning, then formulae such as (50) would have extra factors.

Let F be a connected graph with vertex set $[r]$. We say that a set $F_1,\dots,F_a$ of connected graphs forms a *tree decomposition* of F if each F_i is connected, the union of the F_i is exactly F, any two of the F_i share at most one vertex, and the F_i intersect in a tree-like structure. The last condition may be expressed by saying that the F_i may be ordered so that each F_j other than the first meets the union of the previous ones in exactly one vertex. Equivalently, the intersection is tree-like if $|F| = 1 + \sum_i(|F_i|-1)$. Equivalently, defining (as usual) a *block* of a graph G to be either a maximal 2-connected subgraph of G or a bridge in G, $F_1,\dots,F_a$ forms a tree composition of F if each F_i is a connected union of one or more blocks of F, with each block contained in exactly one F_i. (Cf. [8, p. 74].)

Note that we allow $a = 1$, in which case $F_1 = F$. For $a \geq 2$, the order of the factors is irrelevant, so, for example, $K_{1,2}$ has a unique non-trivial tree decomposition, into two edges. Note also that if F is 2-connected, then it has only the trivial tree decomposition.

Let us say that a copy of F in $G(n, \kappa)$ if *regular* if it is the union of graphs $F_1, \ldots, F_a$ forming a tree decomposition of F, where each F_i arises directly as a subgraph of some atom F_i', and $V(F_i') \cap V(F_j') = V(F_i) \cap V(F_j)$ for all $i \neq j$ (with this intersection containing at most one vertex). We can write down exactly the probability that $G(n, \underset{\sim}{\kappa})$ contains a regular copy of F with vertex set $1, \ldots, r$ in terms of certain integrals of products of conditional expectations $\mathbb{E}(X^0_{F_i}(G) \mid x_1, \ldots, x_s)$. We shall not do so. Instead, let

$$t(F, \underset{\sim}{\kappa}) = \sum_{\{F_1, \ldots, F_a\}} \int_{\mathcal{S}^r} \sigma_{F_1} \sigma_{F_2} \cdots \sigma_{F_a} \, d\mu(x_1) \cdots d\mu(x_r), \tag{53}$$

where the sum runs over all tree decompositions of F and each term σ_{F_i} is evaluated at the subset of $x_1, \ldots, x_r$ corresponding to the vertices of $F_i \subset F$, and set

$$\tilde{t}(F, \underset{\sim}{\kappa}) = \mathrm{aut}(F)^{-1} t(F, \underset{\sim}{\kappa}). \tag{54}$$

Note that these definitions extend to disconnected graphs F, taking the sum over all combinations of one tree decomposition for each component of F.

The upper bound (52) easily implies that the expected number of regular copies of F in $G(n, \underset{\sim}{\kappa})$ is at most $\tilde{t}(F, \underset{\sim}{\kappa})n$, and furthermore this bound is correct within a factor $1 + O(n^{-1})$ if $\underset{\sim}{\kappa}$ is bounded; the factor $\mathrm{aut}(F)^{-1}$ appears because there are $n_{(r)} / \mathrm{aut}(F)$ potential copies of F. Note that the number $\mathrm{emb}(F, G)$ of embeddings of a graph F into a graph G, i.e., the number of injective homomorphisms from F to G, is simply $\mathrm{aut}(F) n(F, G)$. Hence $t(F, \underset{\sim}{\kappa})$ is the appropriate normalization for counting embeddings of F into $G(n, \underset{\sim}{\kappa})$ rather than copies of F. In other contexts, when dealing with dense graphs, it turns out to be most natural to consider homomorphisms from F to G, the number of which will be very close to $\mathrm{emb}(F, G)$. Thus the normalization in (53) is standard in related contexts. (See, for example, Lovász and Szegedy [32].)

Let us illustrate the definitions above with two simple examples.

Example 31. The simplest case is $F = K_2$. In this case, there is only the trivial tree decomposition, and (53) and (54) yield

$$\tilde{t}(K_2, \underset{\sim}{\kappa}) = \frac{1}{2} \int_{\mathcal{S}^2} \sigma_{K_2}(x, y) = \frac{1}{2} \int_{\mathcal{S}^2} \kappa_{\mathrm{e}}(x, y) = \xi(\underset{\sim}{\kappa}). \tag{55}$$

Example 32. Suppose that $\underset{\sim}{\kappa}$ contains only two non-zero kernels, κ_2, corresponding to an edge, and κ_3, corresponding to a triangle; our aim is to calculate $\tilde{t}(P_2, \underset{\sim}{\kappa})$ in this case, where P_2 is the path with 2 edges. Using symmetry of κ_2 and κ_3,

$$\sigma_{K_2}(x, y) = 2\kappa_2(x, y) + 6 \int_{\mathcal{S}} \kappa_3(x, y, z) \, d\mu(z), \tag{56}$$

while

$$\sigma_{P_2}(x, y, z) = 6\kappa_3(x, y, z), \tag{57}$$

reflecting the fact the P_2 ijk appears directly in $G(n, \underset{\sim}{\kappa})$ if and only if we added a triangle with vertex set $\{i, j, k\}$, and this vertex set corresponds to 6 3-tuples.

Since $\mathrm{aut}(P_2) = 2$, it follows that

$$\tilde{t}(P_2, \underset{\sim}{\kappa}) = \frac{1}{2} \int \big(\sigma_{P_2}(x, y, z) + \sigma_{K_2}(x, y)\sigma_{K_2}(y, z)\big)\, d\mu(x)\, d\mu(y)\, d\mu(z).$$

More generally, let F be any (simple) subgraph of $G_n = G(n, \underset{\sim}{\kappa})$ with k components. (We abuse notation by now writing F for a specific subgraph of G_n, rather than an isomorphism class of graphs.) Let $F_1', \ldots, F_a'$ list all atoms contributing edges of F, and let $F_i = F_i' \cap F$, where we take the intersection in the multigraph sense, i.e., intersect the edge sets. For example, if e_1 and e_2 are parallel edges in G_n and $e_1 \in E(F)$, $e_2 \in E(F_1')$, then $F_1 = F_1' \cap F$ contains no ij edge, even though F_1' and F each do so. Thus each F_i contains at least one edge, and F is the edge-disjoint union of the F_i. Since F has k components, when adding the F_i one by one, at least $a - k$ times a new component is *not* created, so at least $a - k$ times at least one vertex of F_i, and hence of F_i', is repeated. It follows that

$$\sum_i (|F_i'| - 1) \geq \Big|\bigcup_i F_i'\Big| - k. \tag{58}$$

Extending our earlier definition, we call F *regular* if equality holds in (58), and *exceptional* otherwise. Note that if any F_i is disconnected, then F is exceptional.

Let $n_{\mathrm{r}}(F, G_n)$ denote the number of regular copies of F in $G_n = G(n, \underset{\sim}{\kappa})$, and $n_{\mathrm{x}}(F, G_n)$ the number of exceptional copies.

Theorem 33. *Let $G_n = G(n, \underset{\sim}{\kappa})$, where $\underset{\sim}{\kappa}$ is a kernel family, and let F be a graph with k components. Then*

$$\mathbb{E} n_{\mathrm{r}}(F, G_n) \leq n^k \tilde{t}(F, \underset{\sim}{\kappa}).$$

If $\underset{\sim}{\kappa}$ is bounded, then

$$\mathbb{E} n_{\mathrm{x}}(F, G_n) = O(n^{k-1}),$$
$$\mathrm{Var}(n(F, G_n)) = O(n^{2k-1}),$$

and

$$n(F, G_n) = n_{\mathrm{r}}(F, G_n) + n_{\mathrm{x}}(F, G_n) = n^k \tilde{t}(F, \underset{\sim}{\kappa})(1 + O_{\mathrm{p}}(n^{-1/2})).$$

Proof. We have essentially given the proof of the first statement, so let us just outline it. To construct a regular copy of F in G_n we must first choose graphs $F_1, \ldots, F_a$ on $V(F)$ forming a tree decomposition of each component of F. Then we must choose a graph F_i' containing each F_i to be the atom that will contain F_i. Then we must choose $s = |\cup_i F_i'|$ distinct vertices $v_1, \ldots, v_s$ from $1, \ldots, n$ to be the vertices of the F_i', where (since F is regular), we have $s = k + \sum_i |F_i' - 1|$.

Note that there are $n_{(s)} \le n^s$ choices for the vertices v_i. (We are glossing over the details of the counting, and in particular various factors $\mathrm{aut}(H)$ for various graphs H. It should be clear comparing the definition of $\tilde{t}(F, \underset{\sim}{\kappa})$ with what follows that these are in the end accounted for correctly.)

Given the vertex types, the probability that these particular graphs F_i' arise is then (up to certain factors $\mathrm{aut}(F_i')$) a product of factors of the form $\kappa_{F_i'}/n^{|F_i'|-1}$, where the kernel is evaluated at an appropriate subset of $x_{v_1}, \ldots, x_{v_s}$. Note that the overall power of n in the denominator is $\sum_i (|F_i'| - 1) = s - k$. Integrating out over the variables x_j corresponding to $V(F_i') \setminus V(F_i)$, and summing over all $F_i' \supset F_i$, the factor $\kappa_{F_i'}$ becomes a factor σ_F. Finally, integrating out over the remaining variables, corresponding to vertices of F, and summing over decompositions, we obtain $n^k \tilde{t}(F, \underset{\sim}{\kappa})$ as an upper bound.

If $\underset{\sim}{\kappa}$ is bounded, then the number s of vertices appearing above is bounded, so $n_{(s)}/n^s = 1 - O(n^{-1})$, where the error term is uniform over all choices for $F_1', \ldots, F_a'$. It follows that in this case,

$$\mathbb{E} n_{\mathrm{r}}(F, G_n) = \tilde{t}(F, \underset{\sim}{\kappa}) n^k (1 - O(n^{-1})).$$

Arguing similarly for exceptional copies, the power $\sum_i (|F_i'| - 1)$ of n in the denominator is now at least $s - k + 1$, and it follows that if $\underset{\sim}{\kappa}$ is bounded, then $\mathbb{E} n_{\mathrm{x}}(F, G_n) = O(n^{k-1})$ as claimed. It follows that

$$\mathbb{E} n(F, G_n) = \tilde{t}(F, \underset{\sim}{\kappa}) n^k + O(n^{k-1}). \tag{59}$$

Finally, for the variance we simply note that $\mathbb{E} n(F, G_n)^2$ is the expected number of ordered pairs (F_1, F_2) of not necessarily disjoint copies of F in G_n. If F_1 and F_2 share one or more vertices, then $F_1 \cup F_2$ has at most $2k - 1$ components. From (59), the expected number of such pairs is $O(n^{2k-1})$. The expected number of pairs with F_1 and F_2 disjoint is simply $N \mathbb{E} n(2F, G_n)$, where $2F$ is the disjoint union of two copies of F and $N = \mathrm{aut}(2F)/\mathrm{aut}(F)^2$ is a symmetry factor, the number of ways $2F$ can be divided into 2 copies of F. (If F is connected then simply $N = 2$ and in general, if F has distinct components F_j with multiplicities m_j, then $N = \prod_j \binom{2m_j}{m_j}$.) Since $t(2F, \underset{\sim}{\kappa}) = t(F, \underset{\sim}{\kappa})^2$, we have $\tilde{t}(2F, \underset{\sim}{\kappa}) = \tilde{t}(F, \underset{\sim}{\kappa})^2/N$, so (59) gives

$$\mathbb{E} n(F, G_n)^2 = \tilde{t}(F, \underset{\sim}{\kappa})^2 n^{2k} + O(n^{2k-1}),$$

from which the variance bound follows. The final bound follows by Chebyshev's inequality. $\square$

For bounded kernel families, Theorem 33 is more or less the end of the story, although one can of course prove more precise results. For unbounded kernel families the situation is much more complicated. Let us first note that regular copies of F do not give rise to any problems.

Theorem 34. *Let $\underset{\sim}{\kappa}$ be a kernel family and F a connected graph, and let $G_n = G(n, \underset{\sim}{\kappa})$. Then $n_{\mathrm{r}}(F, G_n)/n \xrightarrow{\mathrm{p}} \tilde{t}(F, \underset{\sim}{\kappa}) \le \infty$. In other words, if $\tilde{t}(F, \underset{\sim}{\kappa}) = \infty$,* *then for any constant C, whp $n_{\mathrm{r}}(F, G_n) \geq Cn$, while if $\tilde{t}(F, \underset{\sim}{\kappa}) < \infty$, then $n_{\mathrm{r}}(F, G_n) = \tilde{t}(F, \underset{\sim}{\kappa})n + o_{\mathrm{p}}(n)$.*

Proof. We consider the truncated kernel families $\underset{\sim}{\kappa}^M$. Since $t(F, \underset{\sim}{\kappa})$ is a sum of integrals of products of sums of integrals of the kernels $\kappa_{F'}$, by monotone convergence we have $t(F, \underset{\sim}{\kappa}^M) \to t(F, \underset{\sim}{\kappa}) \leq \infty$ as $M \to \infty$, and hence $\tilde{t}(F, \underset{\sim}{\kappa}^M) \to \tilde{t}(F, \underset{\sim}{\kappa})$.

If $\tilde{t}(F, \underset{\sim}{\kappa}) = \infty$, choose M so that $\tilde{t}(F, \underset{\sim}{\kappa}^M) > C$, and couple $G'_n = G(n, \underset{\sim}{\kappa}^M)$ and G_n in the natural way so that $G'_n \subseteq G_n$. Since $\underset{\sim}{\kappa}^M$ is bounded, Theorem 33 implies that $n_{\mathrm{r}}(F, G'_n) \geq Cn$ whp. Since $n_{\mathrm{r}}(F, G_n) \geq n_{\mathrm{r}}(F, G'_n)$, the result follows.

If $\tilde{t}(F, \underset{\sim}{\kappa}) < \infty$, then given $\varepsilon > 0$, the truncation argument above shows that $n_{\mathrm{r}}(F, G_n) \geq (\tilde{t}(F, \underset{\sim}{\kappa}) - \varepsilon)n$ holds whp. By the first statement of Theorem 33, $\mathbb{E}n_{\mathrm{r}}(F, G_n) \leq \tilde{t}(F, \underset{\sim}{\kappa})n$. Combining these two bounds gives the result. □

Note that we do not directly control the variance of $n_{\mathrm{r}}(F, G_n)$; as we shall see in Section 7, there are natural examples where $n_{\mathrm{r}}(F, G_n)/n$ is concentrated about its finite mean even though its variance tends to infinity.

The very simplest case of Theorem 34 concerns edges; we stated this as a separate result in the introduction.

Proof of Theorem 3. Since all copies of K_2 in $G_n = G(n, \underset{\sim}{\kappa})$ are regular (and direct), $e(G_n) = n(K_2, G_n) = n_{\mathrm{r}}(K_2, G_n)$, and taking $F = K_2$ in Theorem 34 and using (55) yields $e(G_n)/n \xrightarrow{\mathrm{p}} \xi(\underset{\sim}{\kappa})$, which is the first claim of Theorem 3. It remains to show that $\mathbb{E}e(G_n) = \mathbb{E}n_{\mathrm{r}}(K_2, G_n) \to \xi(\underset{\sim}{\kappa})$. The lower bound follows from the first part, since convergence in probability implies $\liminf_{n\to\infty} \mathbb{E}e(G_n)/n \geq \xi(\underset{\sim}{\kappa})$, while Theorem 33 gives $\mathbb{E}e(G_n)/n \leq \tilde{t}(K_2, \underset{\sim}{\kappa}) = \xi(\underset{\sim}{\kappa})$, completing the proof. □

It is also easy to prove Theorem 3 directly, using truncations as in this section but avoiding many complications present in the general case.

By a *moment* of a kernel family $\underset{\sim}{\kappa}$ we shall mean any integral of the form

$$\int_{\mathcal{S}^d} \kappa_{F_1}\kappa_{F_2}\cdots\kappa_{F_r}\, d\mu(x_1)\cdots d\mu(x_d),$$

where $F_1, \ldots, F_r$ are not necessarily distinct, and each term κ_{F_i} is evaluated at some $|F_i|$-tuple of distinct x_j. The proof of Theorem 33 shows that for any connected F, $\mathbb{E}n_{\mathrm{x}}(F, G(n, \underset{\sim}{\kappa}))$ is bounded by a sum of moments of $\underset{\sim}{\kappa}$. This gives a very strong condition under which we can control $n_{\mathrm{x}}(F, G(n, \underset{\sim}{\kappa}))$.

Theorem 35. *Let $\underset{\sim}{\kappa}$ be a kernel family in which only finitely many kernels κ_F are non-zero. Suppose also that all moments of $\underset{\sim}{\kappa}$ are finite. Then for any connected F, $\mathbb{E}n_{\mathrm{x}}(F, G(n, \underset{\sim}{\kappa})) = O(1)$, and the conclusions of Theorem 34 apply with $n_{\mathrm{r}}(F, G_n)$ replaced by $n(F, G_n)$.*

Proof. This is essentially trivial from the comments above and Theorem 34. We omit the details. □

Example 30 shows that some conditions are necessary to control $n_{\mathrm{x}}(F, G(n, \underset{\sim}{\kappa}))$; we refer the reader to (49) for the description of the kernel family in this case. Plugging (49) into (56), in this case we have $\sigma_{K_2}(x, y) = 6d(x, y)^{\varepsilon-1} + a$ for some constant a (in fact, $a = 24\varepsilon^{-1}2^{-\varepsilon}$), and it easily follows that $\tilde{t}(P_2, \underset{\sim}{\kappa}) < \infty$. However, as shown in the discussion of that example, whp there is a vertex with degree at least $n^{1-3\varepsilon}$, and hence at least $n^{2-6\varepsilon}$ copies of P_2, which is much larger than n if $\varepsilon < 1/6$. In this case the problem is exceptional P_2s ijk arising from atoms $ij\ell$ and $jk\ell$: the corresponding moment

$$\int_{\mathcal{S}^4} \kappa_3(x_1, x_2, x_4)\kappa_3(x_2, x_3, x_4)$$

is infinite, due to the contribution from $d(x_2, x_4)^{2\varepsilon-2}$.

Of course, not all moments contribute to $\mathbb{E}n_{\mathrm{x}}(F, G_n)$; as we shall see in the next section, it is easy to obtain results similar to Theorem 35 under weaker assumptions in special cases. Also, in general it may happen that $n_{\mathrm{x}}(F, G_n)$ has infinite expectation (in the multigraph form), but is nonetheless often small, i.e., that the large expectation comes from the small probability of having a vertex in very many copies of F. Much more generally, it turns out that when $\underset{\sim}{\kappa}$ is integrable, whp all exceptional copies of F sit on a rather small set of vertices.

Theorem 36. *Let $\underset{\sim}{\kappa}$ be an integrable kernel family and F a connected graph, with $\tilde{t}(F, \underset{\sim}{\kappa})$ finite. Let $G_n = G(n, \underset{\sim}{\kappa})$.*

For any $\varepsilon > 0$, there is a $\delta > 0$ such that whp every *graph G'_n formed from G_n by deleting at most δn vertices has $n(F, G'_n) \geq (\tilde{t}(F, \underset{\sim}{\kappa}) - \varepsilon)n$.*

For any $\varepsilon > 0$ and any $\delta > 0$, whp there is some *graph G'_n formed from G_n by deleting at most δn vertices such that $n(F, G'_n) \leq (\tilde{t}(F, \underset{\sim}{\kappa}) + \varepsilon)n$.*

Together the statements above may be taken as saying that G_n contains essentially $(\tilde{t}(F, \underset{\sim}{\kappa}) + o_{\mathrm{p}}(1))n$ copies of F, where 'essentially' means that we may ignore $o(n)$ vertices. In other words, the 'bulk' of G_n contains this many copies of F, though a few exceptional vertices may meet many more copies.

Proof. We start with the second statement, since it is more or less immediate. Indeed, writing $\int \underset{\sim}{\kappa}$ for $\sum_{F\in\mathcal{F}} |F| \int_{\mathcal{S}^{|F|}} \kappa_F$, and considering truncations $\underset{\sim}{\kappa}^M$ as usual, from monotone convergence we have $\int \underset{\sim}{\kappa}^M \nearrow \int \underset{\sim}{\kappa}$ as $M \to \infty$. Let $\varepsilon > 0$, $\delta > 0$ and $\eta > 0$ be given. Since $\underset{\sim}{\kappa}$ is integrable, i.e., $\int \underset{\sim}{\kappa} < \infty$, there is some M such that $\int \underset{\sim}{\kappa}^M \geq \int \underset{\sim}{\kappa} - \delta\eta/2$. Coupling $G_n^M = G(n, \underset{\sim}{\kappa}^M)$ and $G_n = G(n, \underset{\sim}{\kappa})$ in the usual way, let us call a vertex *bad* if it meets an atom present in G_n but not G_n^M. The expected number of bad vertices is at most the expected sum of the orders of the extra atoms, which is at most $n(\int \underset{\sim}{\kappa} - \int \underset{\sim}{\kappa}^M) \leq \delta\eta n/2$. Hence the probability that there are more than δn bad vertices is at most $\eta/2$.

Deleting all bad vertices from G_n leaves a graph G'_n with at most $n(F, G_n^M)$ copies of F. Applying Theorem 33 this number is at most $(\tilde{t}(F, \underset{\sim}{\kappa}^M) + \varepsilon)n \leq (\tilde{t}(F, \underset{\sim}{\kappa}) + \varepsilon)n$ whp, so we see that if n is large enough, then with probability at least $1 - \eta$ we may delete at most δn vertices to leave G'_n with at most $(\tilde{t}(F, \underset{\sim}{\kappa}) + \varepsilon)n$ copies of F, as required.

Turning to the first statement, we may assume without loss of generality that $\underset{\sim}{\kappa}$ is bounded. Indeed, there is some truncation $\underset{\sim}{\kappa}^M$ with $\tilde{t}(F, \underset{\sim}{\kappa}^M) \geq \tilde{t}(F, \underset{\sim}{\kappa}) - \varepsilon/2$, and taking $G(n, \underset{\sim}{\kappa}^M) \subset G(n, \underset{\sim}{\kappa})$ as usual, it suffices to prove the same statement for $G(n, \underset{\sim}{\kappa}^M)$ with ε replaced by $\varepsilon/2$. Assuming $\underset{\sim}{\kappa}$ is bounded, then by Theorem 33 we have $n(F, \underset{\sim}{\kappa}) \geq (\tilde{t}(F, \underset{\sim}{\kappa}) - \varepsilon/2)n$ whp, so it suffices to prove that if $\underset{\sim}{\kappa}$ is bounded and $\varepsilon > 0$, then there is some $\delta > 0$ such that whp any δn vertices of $G_n = G(n, \underset{\sim}{\kappa})$ meet at most εn copies of F.

Let v be a fixed vertex of F, and for $1 \leq i \leq n$ let a_i denote the number of homomorphisms from F to G_n mapping v to vertex i. Let F' be the graph formed from two copies of F meeting only at v. Then there are exactly a_i^2 homomorphisms from F' to G_n mapping v to i, so in total there are $\sum a_i^2$ homomorphisms from F' to G_n. Now the image of any homomorphism from F' to G_n is a connected subgraph F'' of G_n, and each such subgraph is the image of $O(1)$ homomorphisms. Applying Theorem 33 to each of the $O(1)$ possible isomorphism types of F'', it follows that there is some constant C such that, whp,

$$\sum_i a_i^2 = \hom(F', G_n) \leq Cn.$$

When the upper bound holds, given any set $S \subset [n]$ with $|S| \leq \delta n$, by the Cauchy–Schwartz inequality we have

$$\sum_{i \in S} a_i \leq \sqrt{|S|} \sqrt{\sum_i a_i^2} \leq \sqrt{\delta n}\sqrt{Cn} = \sqrt{C\delta} n.$$

Repeating the argument above for each vertex v of F and summing, we see that there is some $C' < \infty$ (given by the sum of at most $|F|$ constants corresponding to C above) such that whp for any $\delta > 0$, and any set S of at most δn vertices of G_n, there are at most $\sqrt{C'\delta} n$ homomorphisms from F to G_n mapping any vertex of F into S. This condition implies that S meets at most $\sqrt{C'\delta} n$ copies of F, so choosing δ such that $\sqrt{C'\delta} < \varepsilon$, we see that whp any δn vertices meet at most εn copies of F. As noted above, the first statement of the theorem follows. $\square$

7 A power-law graph with clustering

Our aim in this paper has been to introduce a very general family of sparse random graph models, showing that despite the generality, the models are still susceptible to mathematical analysis. The question of which special cases of the model may be relevant in applications is a very broad one, and not our focus. Nevertheless, in the light of the motivation of the model, we shall investigate one special case. We should like to show that, with an appropriate choice of kernel family, our model gives rise to graphs with power-law degree distributions, with various ranges of the degree exponent, the clustering coefficient (see (66)), and the mixing coefficient (see (69)). We achieve this in the simplest possible way, considering a 'rank 1' version of the model in which we add only edges and triangles. We do not claim that this particular model is appropriate for any particular real-world example; nevertheless, it shows the potential of our model to produce graphs that are similar to real-world graphs, where similarity is measured by the values of these important and much studied parameters.

Throughout this section we fix three parameters, $\alpha > 1$, and $A, B \geq 0$ with $A + B > 0$. We consider one specific kernel family $\underset{\sim}{\kappa}$ on $\mathcal{S} = (0, 1]$ with μ Lebesgue measure. Our kernel family has only two non-zero kernels, κ_2, corresponding to edges, and κ_3 to triangles, with

$$\kappa_2(x, y) = Ax^{-1/\alpha}y^{-1/\alpha}$$

and

$$\kappa_3(x, y, z) = Bx^{-1/\alpha}y^{-1/\alpha}z^{-1/\alpha}.$$

We could of course consider many other possible functions, but these seem the simplest and most natural for our purposes. It would be straightforward to carry out computations such as those that follow with each of the αs above replaced by a different constant, for example, although we should symmetrize the kernels in this case. However, one of these exponents would determine the power law, and it seems most natural to take them all equal.

For convenience, we define

$$\beta_k = \int_0^1 x^{-k/\alpha}\,dx = \begin{cases} \frac{\alpha}{\alpha-k}, & \alpha > k, \\ \infty, & \alpha \leq k. \end{cases} \tag{60}$$

In particular, $\beta_1 = \alpha/(\alpha - 1)$. We then have

$$\int_{\mathcal{S}^2} \kappa_2 = A\beta_1^2 \quad \text{and} \quad \int_{\mathcal{S}^3} \kappa_3 = B\beta_1^3,$$

so $\underset{\sim}{\kappa}$ is integrable. Also, for the asymptotic edge density in Theorem 3,

$$\xi(\underset{\sim}{\kappa}) = \int_{\mathcal{S}^2} \kappa_2 + 3\int_{\mathcal{S}^3} \kappa_3 = A\beta_1^2 + 3B\beta_1^3. \tag{61}$$

In the following subsections we apply our general results to determine various characteristics of this particular random graph $G_n = G(n, \underset{\sim}{\kappa})$.

7.1 Degree distribution

From (33) and symmetry of κ_2 and κ_3 we see that

$$\lambda_{K_2,1}(x) = \lambda_{K_2,2}(x) = \int_{\mathcal{S}} \kappa_2(x, y)\,d\mu(y) = A\beta_1 x^{-1/\alpha},$$

while for $j = 1, 2, 3$,

$$\lambda_{K_3,j}(x) = \int_{\mathcal{S}^2} \kappa_3(x, y, z)\,d\mu(y)\,d\mu(z) = B\beta_1^2 x^{-1/\alpha}.$$

Since an edge contributes 1 to the degree of each endvertex, while a triangle contributes 2 to the degrees of its vertices, for each x, the measure λ_x defined by (34) is given by

$$\lambda_x = 2A\beta_1 x^{-1/\alpha}\delta_1 + 3B\beta_1^2 x^{-1/\alpha}\delta_2.$$

Theorem 29 then tells us that the degree distribution of $G_n = G(n, \underset{\sim}{\kappa})$ converges to the mixed compound Poisson distribution $\mathrm{MCPo}(\Lambda)$, where Λ is the random measure corresponding to λ_x with x chosen uniformly from $(0, 1]$.

Note that if $B = 0$, then the limiting degree distribution is mixed Poisson, while if $A = 0$, almost all degrees are even and the degrees divided by 2 have a mixed Poisson distribution.

For the power law, note that the mean $\lambda(x)$ of λ_x is simply

$$\lambda(x) = (2A\beta_1 + 6B\beta_1^2)x^{-1/\alpha} = cx^{-1/\alpha},$$

where $0 < c = 2A\beta_1 + 6B\beta_1^2 = 2\xi(\underset{\sim}{\kappa})/\beta_1 < \infty$ is a constant depending on A, B and α. Choosing x randomly from $(0, 1]$, for any $k > c$ we have

$$\mathbb{P}(\lambda(x) > k) = \mathbb{P}(x < (k/c)^{-\alpha}) = (k/c)^{-\alpha},$$

so the distribution of $\lambda(x)$ has a power-law tail. Using the concentration properties of Poisson distributions with large means, arguing as in the proof of Corollary 13.1 of [10], it follows easily that

$$\mathbb{P}(\mathrm{MCPo}(\Lambda) > k) \sim (k/c)^{-\alpha}$$

as $k \to \infty$, so the asymptotic degree distribution does indeed have a power-law tail with (cumulative) exponent α.

Let $d_k = \mathbb{P}(\mathrm{MCPo}(\Lambda) = k)$, so by Theorem 29, the asymptotic fraction of vertices with degree k is simply d_k. If $A > 0$ then it is not hard to check that in fact

$$d_k \sim c' k^{-\alpha-1} \tag{62}$$

as $k \to \infty$, where $0 < c' = \alpha c^\alpha < \infty$, so the degree distribution is power-law in this stronger sense. If $A = 0$, then $d_k = 0$ if k is odd, but (62) still holds for even k, for a different (doubled) constant c'.

7.2 The phase transition and the giant component

From (3), we have $\kappa_{\mathrm{e}}(x, y) = (2A + 6B\beta_1)x^{-1/\alpha}y^{-1/\alpha}$, which we may rewrite as $\kappa_{\mathrm{e}}(x, y) = \psi(x)\psi(y)$, where

$$\psi(x) = (2A + 6B\beta_1)^{1/2}x^{-1/\alpha}.$$

By Theorems 5 and 6, the largest component of G_n is of size $\rho(\underset{\sim}{\kappa})n + o_{\mathrm{p}}(n)$, and there is a giant component, i.e., $\rho(\underset{\sim}{\kappa}) > 0$, if and only if $\|T_{\kappa_{\mathrm{e}}}\| > 1$. In this case κ_{e} is 'rank 1' in the terminology of [10], and we have

$$\|T_{\kappa_{\mathrm{e}}}\| = \|\psi\|_2^2 = (2A + 6B\beta_1)\beta_2.$$

Hence, fixing $\alpha > 2$ and thus β_1 and β_2, there is a giant component if and only if $2A + 6B\alpha/(\alpha - 1) > (\alpha - 2)/\alpha$.

Turning to the normalized size $\rho(\underset{\sim}{\kappa})$ of the giant component, Theorem 10 allows us to calculate this in terms of the solution to a functional equation. Usually this is intractable, but for the special $\underset{\sim}{\kappa}$ we are considering this simplifies greatly, as in the rank 1 case of the edge-only model; see Section 16.4 of [10], or Section 6.2 of [34]. Indeed, writing $\rho(x)$ for the survival probability of $\mathfrak{X}_{\underset{\sim}{\kappa}}(x)$, from (6) we have

$$S_{\underset{\sim}{\kappa}}(\rho)(x) = \int_0^1 2Ax^{-1/\alpha}y^{-1/\alpha}\rho(y)\,dy + \int_0^1\int_0^1 3Bx^{-1/\alpha}y^{-1/\alpha}z^{-1/\alpha}(\rho(y) + \rho(z) - \rho(y)\rho(z))\,dy\,dz,$$

which simplifies to

$$S_{\underset{\sim}{\kappa}}(\rho)(x) = x^{-1/\alpha}(2AC + 6B\beta_1 C - 3BC^2),$$

where

$$C = \int_0^1 x^{-1/\alpha}\rho(x)\,dx. \tag{63}$$

By Lemma 7, we have $\rho(x) = 1 - \exp(-S_{\underset{\sim}{\kappa}}(\rho)(x))$, so

$$\rho(x) = 1 - \exp\bigl(-(2AC + 6B\beta_1 C - 3BC^2)x^{-1/\alpha}\bigr). \tag{64}$$

Although we defined C in terms of ρ, we can view C as an unknown constant, define ρ by (64), and substitute back into (63). The function ρ then solves (7) if and only if C solves

$$C = \int_0^1 x^{-1/\alpha}\left(1 - \exp\bigl(-\bigl((2A + 6B\beta_1)C - 3BC^2\bigr)x^{-1/\alpha}\bigr)\right), \tag{65}$$

and every solution to (7) arises in this way. In particular, by Theorem 10 there is a unique positive solution above the critical point. Transforming the integral using the substitution $y = x^{-1/\alpha}$, one can rewrite the right hand side of (65) in terms of an incomplete gamma function, although it is not clear this is informative. The point is that the form of $\rho(x)$ is given by (64), and the constant can in principle be found as the solution to an equation, and can easily be found numerically for given values of A, B and α.

7.3 Subgraph densities

In the following subsections we shall need expressions for $\tilde{t}(F, \underset{\sim}{\kappa})$ for various small graphs F, where $\tilde{t}(F, \underset{\sim}{\kappa})$, defined by (53) and (54), may be thought of as the asymptotic density of copies of F in the kernel family $\underset{\sim}{\kappa}$.

We start with direct copies of F. Since all atoms are edges or triangles, the only graphs F that can be produced directly are edges, triangles, and P_2s, i.e., paths with 2 edges.

Putting the specific kernels κ_2 and κ_3 into the formulae (56) and (57) from the previous section, we have

$$\sigma_{K_2}(x,y) = (2A + 6B\beta_1)x^{-1/\alpha}y^{-1/\alpha},$$

and

$$\sigma_{P_2}(x,y,z) = 6Bx^{-1/\alpha}y^{-1/\alpha}z^{-1/\alpha},$$

while

$$\sigma_{K_3}(x,y,z) = 6\kappa_{K_3}(x,y,z) = 6Bx^{-1/\alpha}y^{-1/\alpha}z^{-1/\alpha}.$$

Edges may be formed only directly, so either from (53) and (54) or from (55), we have

$$\tilde{t}(K_2, \underset{\sim}{\kappa}) = \frac{1}{2}\int \sigma_{K_2}(x,y)\,d\mu(x)\,d\mu(y) = A\beta_1^2 + 3B\beta_1^3,$$

which agrees, as it should, with (61). Since a triangle is 2-connected, it has no non-trivial tree decomposition, and (53) and (54) give

$$\tilde{t}(K_3, \underset{\sim}{\kappa}) = \frac{1}{6}\int_{\mathcal{S}^3} 6\kappa_3(x,y,z)\,d\mu(x)\,d\mu(y)\,d\mu(z) = B\beta_1^3,$$

which may also be seen by noting that the only regular copies of a triangle are those directly corresponding to κ_3.

A copy of P_2 may be formed by a single triangular atom (a direct copy), but may also be formed by two edges from different atoms Hence, as in Example 32,

$$\begin{aligned}\tilde{t}(P_2, \underset{\sim}{\kappa}) &= \frac{1}{2}\int (\sigma_{P_2}(x,y,z) + \sigma_{K_2}(x,y)\sigma_{K_2}(y,z))\,d\mu(x)\,d\mu(y)\,d\mu(z)\\ &= \frac{1}{2}\left(6B\beta_1^3 + \int (2A + 6B\beta_1)^2 x^{-1/\alpha}y^{-2/\alpha}z^{-1/\alpha}\,d\mu(x)\,d\mu(y)\,d\mu(z)\right)\\ &= 3B\beta_1^3 + \frac{(2A+6B\beta_1)^2\beta_1^2}{2}\beta_2.\end{aligned}$$

In particular, if $\alpha \le 2$ then $\tilde{t}(P_2, \underset{\sim}{\kappa})$ is infinite.

For $S_3 = K_{1,3}$, the star with three edges, there are two types of tree-decompositions: three edges or one edge and one copy of P_2, the latter occurring in 3 different ways. (There are no direct copies.) Hence,

$$\begin{aligned}t(S_3, \underset{\sim}{\kappa}) &= \int \sigma_{K_2}(x_1,x_2)\sigma_{K_2}(x_1,x_3)\sigma_{K_2}(x_1,x_4) + 3\int \sigma_{P_2}(x_2,x_1,x_3)\sigma_{K_2}(x_1,x_4)\\ &= (2A+6B\beta_1)^3\beta_1^3\beta_3 + 18B(2A+6B\beta_1)\beta_1^3\beta_2\end{aligned}$$

and thus

$$\tilde{t}(S_3, \underset{\sim}{\kappa}) = \frac{1}{6}(2A+6B\beta_1)^3\beta_1^3\beta_3 + 3B(2A+6B\beta_1)\beta_1^3\beta_2.$$

Finally, for P_3, there are again two types of tree-decompositions: three edges or one edge and one copy of P_2, the latter now occurring in 2 different ways. Hence,

$$\tilde{t}(P_3,\underset{\sim}{\kappa}) = \frac{1}{2}\int \sigma_{K_2}(x_1,x_2)\sigma_{K_2}(x_2,x_3)\sigma_{K_2}(x_3,x_4) + \frac{2}{2}\int \sigma_{P_2}(x_1,x_2,x_3)\sigma_{K_2}(x_3,x_4)$$
$$= \frac{1}{2}(2A+6B\beta_1)^3\beta_1^2\beta_2^2 + 6B(2A+6B\beta_1)\beta_1^3\beta_2.$$

As we shall see, the counts above are enough to calculate two more interesting parameters of the graph $G_n = G(n,\underset{\sim}{\kappa})$.

7.4 The clustering coefficient

The *clustering coefficient* $C(G)$ of a graph G was introduced by Watts and Strogatz [35] as a measure of the extent to which neighbours of a random vertex in G tend to be joined directly to each other. After the degree distribution, it is one of the most studied parameters of real-world networks. As discussed in [13], for example, there are several different definitions of such clustering coefficients. One of these turns out to be most convenient for mathematical analysis, and is also very natural; following [13], we call this coefficient $C_2(G)$. (Hopefully there will be no confusion with our earlier use of $C_2(G)$ for the number of vertices in the 2nd largest component.) The coefficient $C_2(G)$ may be defined as a certain weighted average of the 'local clustering coefficients' at individual vertices, but is also simply given by

$$C_2(G) = \frac{3n(K_3,G)}{n(P_2,G)}, \tag{66}$$

a ratio that is easily seen to lie between 0 and 1.

Now from above we have $\tilde{t}(K_3,\underset{\sim}{\kappa}) = B\beta_1^3 < \infty$. Hence, by Theorem 34,

$$n_{\mathrm{r}}(K_3,G_n) = B\beta_1^3 n + o_{\mathrm{p}}(n),$$

where, as usual, $G_n = G(n,\underset{\sim}{\kappa})$. We shall return to exceptional copies of K_3 shortly.

If $\alpha \le 2$ then $\tilde{t}(P_2,\underset{\sim}{\kappa})$ is infinite, and G_n will whp contain more than $O(n)$ copies of P_2. Note that this is to be expected given the exponent of the asymptotic degree distribution, since in this case the expected square degree is infinite.

From now on we suppose that $\alpha > 2$, so $\tilde{t}(P_2,\underset{\sim}{\kappa})$ is finite. Suppose for the moment that exceptional copies of P_2 and K_3 are negligible, i.e., that

$$n_{\mathrm{x}}(P_2,G_n),\ n_{\mathrm{x}}(K_3,G_n) = o_{\mathrm{p}}(n). \tag{67}$$

By Theorem 34, we have $n_{\mathrm{r}}(P_2,G_n)/n = \tilde{t}(P_2,\underset{\sim}{\kappa}) + o_{\mathrm{p}}(1)$ and $n_{\mathrm{r}}(K_3,G_n)/n = \tilde{t}(K_3,\underset{\sim}{\kappa}) + o_{\mathrm{p}}(1)$, so it follows that

$$C_2(G_n) = \frac{3\tilde{t}(K_3,\underset{\sim}{\kappa})}{\tilde{t}(P_2,\underset{\sim}{\kappa})} + o_{\mathrm{p}}(1) = c_2(A,B,\alpha) + o_{\mathrm{p}}(1)$$

where, from the formulae in Subsection 7.3,

$$c_2(A,B,\alpha) = \frac{3B\beta_1^3}{3B\beta_1^3 + 2(A+3B\beta_1)^2\beta_1^2\beta_2}, \tag{68}$$

with β_1, β_2 given by (60). It follows that with the degree exponent $\alpha > 2$ fixed, this special case of our model can achieve any possible value of the clustering coefficient, with the trivial exception of 1 (achieved only by graphs that are vertex disjoint unions of cliques). Indeed, $c_2(A, 0, \alpha) = 0$ for any A, while taking $A = 0$ we have

$$c_2(0,B,\alpha) = \frac{1}{1+6B\beta_1\beta_2},$$

which is decreasing as a function of B, and tends to 1 as $B \to 0$ and to 0 as $B \to \infty$.

Let us note in passing that by Theorem 34, if $2 < \alpha \le 4$ then $n_{\mathrm{r}}(P_2, G_n)/n$ is concentrated around its finite mean even though its variance, which involves the expected 4th power of the degree of a random vertex, tends to infinity.

So far we considered only regular copies of P_2 and K_3; we now turn our attention to exceptional copies. Unfortunately, for any α, some moment of our kernel is infinite, so Theorem 35 does not apply. However, it is easy to describe the set of moments relevant to the calculation of $\mathbb{E}n_{\mathrm{x}}(F, G_n)$ for the graphs F we consider.

Suppose that F is an exceptional triangle (or P_2; the argument is then almost identical) in $G_n = G(n, \underset{\sim}{\kappa})$. Since F has (at most) three edges, there are at most 3 atoms F_i contributing edges to F. Let H be the union of these atoms, considered as a multigraph. For example, if F is the triangle abc, then H might consist of the union of the three triangles abd, bcd, and cad. In some sense this will turn out to be the 'worst' case.

Let us fix the isomorphism type of H, defined in the obvious way. Let h be the total number of vertices in H, and write $r = \sum_i(|F_i| - 1) - (h-1)$ for the 'redundancy' of H. Since F is exceptional, $r \ge 1$. The expected number of exceptional F arising in this way is exactly $n_{(h)}n^{-\sum_i(|F_i|-1)}$ times a certain integral of products of κ_2 and κ_3. From the form of κ_2 and κ_3, we may write this as

$$\frac{n_{(h)}}{n^{r+h-1}} \int_{\mathcal{S}^h} x_1^{-n_1/\alpha} \cdots x_h^{-n_h/\alpha}\, d\mu(x_1)\cdots d\mu(x_h),$$

where n_i is the number of the atoms F_j that contain the ith vertex of H. The initial factor is at most $n^{1-r} \le 1$, while the integral is finite unless $n_i \ge \alpha$ for some i. Since H is made up of at most 3 atoms F_j, we always have $n_i \le 3$, so if $\alpha > 3$ then the relevant integrals (i.e., the relevant moments) are finite, and we have $\mathbb{E}n_{\mathrm{x}}(K_3, G_n), \mathbb{E}n_{\mathrm{x}}(P_2, G_n) = O(1)$, which certainly implies (67).

In fact, we do not need to assume that $\alpha > 3$. Suppose that $2 < \alpha \le 3$. Then in the multigraph version of the model, $\mathbb{E}n_{\mathrm{x}}(K_3, G_n) = \infty$. (Consider, for example, three triangles sitting on 4 vertices as above.) On the other hand, this does not mean that $n_{\mathrm{x}}(K_3, G_n)$ is often large. Indeed, when we choose our vertex types uniformly from $(0, 1]$, whp there is no vertex whose type x is at most $\delta = 1/(n \log n)$, say. Conditioning on this very likely event $\mathcal{A}$, we may consider the restrictions of κ_2 and κ_3 to $(\delta, 1]^2$ and $(\delta, 1]^3$, respectively. Now the expected number of copies of some pattern H is at most a constant times

$$n^{1-r} \left(\int_\delta^1 x^{-3/\alpha} \, dx \right)^s ,$$

where s is the number of vertices i of H with $n_i = 3$. Since the graph K_3 (or P_2) we are trying to form has maximum degree 2, every vertex of H with $n_i = 3$ corresponds to a redundancy, so we always have $r \geq s$. Up to constants and a power of $\log n$ the integral is $n^{(3-\alpha)/\alpha} \leq \sqrt{n}$, and it follows that

$$\mathbb{E}(n_{\mathrm{x}}(K_3, G_n) \mid \mathcal{A}) = O(\sqrt{n}) = o(n).$$

Since $\mathbb{P}(\mathcal{A}) = 1 - o(1)$, it follows that $n_{\mathrm{x}}(K_3, G_n) = o_{\mathrm{p}}(n)$, even though its expectation would not suggest this. The same holds for $n_{\mathrm{x}}(P_2, G_n)$, so we see that (67) does indeed hold for any $\alpha > 2$, and the clustering coefficient is indeed concentrated about $c_2(A, B, \alpha)$.

7.5 The mixing coefficient

Another interesting parameter of real networks is the extent to which the degrees of the two ends of a randomly chosen edge tend to correlate; positive correlation is known as *assortative mixing*, and negative correlation as *disassortative mixing*. To define this precisely, let G be any graph, and let vw be an edge of G chosen uniformly at random. More precisely, let (v, w) be chosen uniformly at random from all $2e(G)$ ordered pairs corresponding to edges of G. Let D_v and D_w denote the degrees of v and w; we view these as random variables. Since the events $\{v = v_1, w = v_2\}$ and $\{v = v_2, w = v_1\}$ have the same probability, the random vertices v and w have the same distribution, so D_v and D_w have the same distribution.

Let

$$a(G) = \frac{\mathrm{Cov}(D_v, D_w)}{\sqrt{\mathrm{Var}(D_v)\,\mathrm{Var}(D_w)}} = \frac{\mathrm{Cov}(D_v, D_w)}{\mathrm{Var}(D_v)}. \tag{69}$$

Here G is fixed, and all expectations are with respect to the random choice of (v, w). Thus $a(G)$ is simply the correlation coefficient between the degrees of the two ends of a randomly chosen edge, so $-1 \leq a(G) \leq 1$, and $a(G) > 0$ corresponds to assortative mixing and $a(G) < 0$ to disassortative mixing. This mixing coefficient was introduced by Callaway, Hopcroft, Kleinberg, Newman and Strogatz [19], building on work of Krapivsky and Redner [30], and has been studied by many people, for example Newman [33]. In [19], $a(G)$ is denoted $\rho(G)$; we avoid this notation as it clashes with our notation for the survival probability of a branching process.

Fortunately, we need no new theory to evaluate $a(G)$ for $G = G(n, \underset{\sim}{\kappa})$, since $a(G)$ can be expressed in terms of small subgraph counts. More precisely, for any graph G,

$$\mathbb{E}(D_v-1)=\frac{1}{2e(G)}\sum_i\sum_{j\sim i}(d_i-1)=\frac{1}{2e(G)}\sum_i d_i(d_i-1)=\frac{n(P_2,G)}{e(G)},$$

where i runs over all vertices of G, then j over all neighbours of i, and d_i is the degree of vertex i in G. Also,

$$\mathbb{E}\big((D_v-1)(D_w-1)\big)=\frac{1}{2e(G)}\sum_i\sum_{j\sim i}(d_i-1)(d_j-1)=\frac{2n(P_3,G)+6n(K_3,G)}{2e(G)},$$

so

$$\operatorname{Cov}(D_v,D_w)=\operatorname{Cov}(D_v-1,D_w-1)=\frac{\big(n(P_3,G)+3n(K_3,G)\big)e(G)-n(P_2,G)^2}{e(G)^2}.$$

Also,

$$2e(G)\mathbb{E}\big((D_v-1)(D_v-2)\big)=\sum_i\sum_{j\sim i}(d_i-1)(d_i-2)=\sum_i d_i(d_i-1)(d_i-2)=6n(S_3,G),$$

where $S_3=K_{1,3}$ is the star with 3 edges. Thus

$$\begin{aligned}\operatorname{Var}(D_v)=\operatorname{Var}(D_v-1)&=\mathbb{E}\big((D_v-1)(D_v-2)\big)+\mathbb{E}(D_v-1)-(\mathbb{E}(D_v-1))^2\\&=\frac{3n(S_3,G)e(G)+n(P_2,G)e(G)-n(P_2,G)^2}{e(G)^2}.\end{aligned}$$

Hence

$$a(G)=\frac{\big(n(P_3,G)+3n(K_3,G)\big)e(G)-n(P_2,G)^2}{3n(S_3,G)e(G)+n(P_2,G)e(G)-n(P_2,G)^2}. \tag{70}$$

In well-behaved cases, for example for bounded kernel families, it follows from our results here (Theorems 33–35) that if $G_n=G(n,\underset{\sim}{\kappa})$, then

$$a(G_n)=a(\underset{\sim}{\kappa})+o_{\mathrm{p}}(1), \tag{71}$$

where

$$a(\underset{\sim}{\kappa})=\frac{\tilde{t}(P_3,\underset{\sim}{\kappa})\xi(\underset{\sim}{\kappa})+3\tilde{t}(K_3,\underset{\sim}{\kappa})\xi(\underset{\sim}{\kappa})-\tilde{t}(P_2,\underset{\sim}{\kappa})^2}{3\tilde{t}(S_3,\underset{\sim}{\kappa})\xi(\underset{\sim}{\kappa})+\tilde{t}(P_2,\underset{\sim}{\kappa})\xi(\underset{\sim}{\kappa})-\tilde{t}(P_2,\underset{\sim}{\kappa})^2}, \tag{72}$$

with $\xi(\underset{\sim}{\kappa})=\tilde{t}(K_2,\underset{\sim}{\kappa})$; see (55).

Returning to our present specific example, substituting in the expressions for $\tilde{t}(\cdot,\underset{\sim}{\kappa})$ in Subsection 7.3, the ratio (72) turns out to be

$$a(\underset{\sim}{\kappa})=\frac{3AB\beta_1^5}{\big(4\tilde{\xi}^4(\beta_1\beta_3-\beta_2^2)+2(A+6B\beta_1)\tilde{\xi}^2\beta_2+3AB\beta_1\big)\beta_1^4}, \tag{73}$$

where $\tilde{\xi}=(A+3B\beta_1)=\xi(\underset{\sim}{\kappa})/\beta_1^2$. Let us make a few comments on these expressions.

Firstly, the coefficients $\tilde{t}(K_2, \underset{\sim}{\kappa})$, $\tilde{t}(P_2, \underset{\sim}{\kappa})$, $\tilde{t}(K_3, \underset{\sim}{\kappa})$ and $\tilde{t}(P_3, \underset{\sim}{\kappa})$ are finite for all $\alpha > 2$, while $\tilde{t}(S_3, \underset{\sim}{\kappa})$ is finite if and only if $\alpha > 3$. For the numerator, one can argue for P_3 as for P_2 and K_3 above to show that the number of exceptional copies of P_3 is $o_{\mathrm{p}}(n)$ and thus negligible for every $\alpha > 2$, and hence $n(P_3, G_n) = \tilde{t}(P_3, \underset{\sim}{\kappa})n + o_{\mathrm{p}}(n)$ by Theorem 34. Consequently, the numerator in (70) (with $G = G_n$) divided by n^2 converges in probability to the numerator in (72), and this limit is finite. For $\alpha > 3$, one can argue in the same way to show that the number of exceptional copies of S_3 is negligible, so $n(S_3, G_n) = \tilde{t}(S_3, \underset{\sim}{\kappa})n + o_{\mathrm{p}}(n)$ and (71) does indeed hold. For $\alpha \le 3$, when $\tilde{t}(S_3, \underset{\sim}{\kappa}) = \infty$, Theorem 34 implies that $n(S_3, \underset{\sim}{\kappa})/n \xrightarrow{\mathrm{p}} \infty$, so in this case $a(G_n) \xrightarrow{\mathrm{p}} 0 = a(\underset{\sim}{\kappa})$, for the not very interesting reason that $\mathrm{Var}(D_v)$ is unbounded while $\mathrm{Cov}(D_v, D_w)$ is not. In any case, we have shown that (71) holds in our example for every $\alpha > 2$.

Secondly, we see that $0 \le a(\underset{\sim}{\kappa}) < \infty$ for every $\alpha > 2$, with $a(\underset{\sim}{\kappa}) > 0$ whenever $\alpha > 3$ and we add both edges and triangles (i.e., if both A and B are non-zero).

Thirdly, if A and B are both positive and comparable but very small, then it is easy to see that $a(\underset{\sim}{\kappa})$ is close to 1, for the simple reason that the graph then consists of rather few (though still order n) edges and triangles, which are almost all vertex disjoint. In this case we almost always have either $D_v = D_w = 1$, if we pick an edge component, or $D_v = D_w = 2$ if we pick an edge of a triangle. This is also easily checked algebraically from (73): the denominator is of the form $3AB\beta_1^5 + O((A+B)^3)$, which is asymptotically equal to the numerator if $A, B \to 0$ with A/B bounded above and below. It follows that as A and B are varied, $a(\underset{\sim}{\kappa})$ can take any value between 0 and 1, with 1 excluded.

Finally, it is easy to check that the form of $a(\underset{\sim}{\kappa})$ as a function of A, B and α is very different from that of $c_2(A, B, \alpha)$ given in (68). It follows that with the degree exponent $\alpha > 3$ fixed, if we vary A and B we may vary the clustering coefficient and $a(\underset{\sim}{\kappa})$ independently, subject to certain inequalities.

It so happens that in the example considered here, $a(\underset{\sim}{\kappa})$ is always non-negative, but it is easy to give examples where $a(\underset{\sim}{\kappa}) < 0$. Indeed, this arises already in the edge-only case (of the kind we treated in [10]), with two elements of weights $\mu\{1\} = p$ and $\mu\{2\} = q = 1 - p$ with $0 < p < 1$, taking $\kappa_2(1, 1) = 0$, $\kappa_2(2, 2) = 0$ and $\kappa_2(1, 2) = A > 0$. In symbols,

$$\kappa_2(x, y) = A\mathbf{1}[x \neq y],$$

where $\mathbf{1}[E]$ is the indicator function of the event E.

For this kernel (family)

$$\begin{aligned} \kappa_{\mathrm{e}}(x, y) &= 2\kappa_2(x, y) = 2A\mathbf{1}[x \neq y], \\ \xi(\underset{\sim}{\kappa}) &= \int \kappa_2 = 2Apq, \\ \sigma_{K_2}(x, y) &= 2\kappa_2(x, y) = \kappa_{\mathrm{e}}(x, y) = 2A\mathbf{1}[x \neq y]. \end{aligned}$$

Expanding the integrals as sums, it follows that

$$\tilde{t}(K_2, \underset{\sim}{\kappa}) = \frac{1}{2} \int \sigma_{K_2}(x, y)\, d\mu(x)\, d\mu(y) = \xi(\underset{\sim}{\kappa}) = 2Apq;$$

$$
\begin{aligned}
\tilde{t}(P_2, \underset{\sim}{\kappa}) &= \frac{1}{2}\int \sigma_{K_2}(x,y)\sigma_{K_2}(y,z)\,d\mu(x)\,d\mu(y)\,d\mu(z) = 2A^2(pqp+qpq)\\
&= 2A^2pq;\\
\tilde{t}(K_3, \underset{\sim}{\kappa}) &= 0\\
\tilde{t}(S_3, \underset{\sim}{\kappa}) &= \frac{1}{6}\int \sigma_{K_2}(x_1,x_2)\sigma_{K_2}(x_1,x_3)\sigma_{K_2}(x_1,x_4)\,d\mu(x_1)\,d\mu(x_2)\,d\mu(x_3)\,d\mu(x_4)\\
&= \frac{4}{3}A^3(pq^3+qp^3) = \frac{4}{3}A^3pq(p^2+q^2);\\
\tilde{t}(P_3, \underset{\sim}{\kappa}) &= \frac{1}{2}\int \sigma_{K_2}(x_1,x_2)\sigma_{K_2}(x_2,x_3)\sigma_{K_2}(x_3,x_4)\,d\mu(x_1)\,d\mu(x_2)\,d\mu(x_3)\,d\mu(x_4)\\
&= 4A^3(pqpq+qpqp) = 8A^3p^2q^2.
\end{aligned}
$$

Substituting these expressions into (72) and simplifying, we find that

$$
a(\underset{\sim}{\kappa}) = -\frac{A(p-q)^2}{A(p-q)^2+1}.
$$

Hence $a(\underset{\sim}{\kappa}) \le 0$, and we have disassortative mixing as soon as $p \neq q$, i.e., when $p \in (0, \frac{1}{2}) \cup (\frac{1}{2}, 1)$. We see also that the coefficient $a(\underset{\sim}{\kappa})$ can be made to take any value in $(-1, 0]$ by choosing the parameters suitably.

One can easily combine the simple example above with that considered in the bulk of this section to give graphs with power-law degree distributions with various values of the clustering coefficient and of $a(G_n)$, now with negative values of $a(G_n)$ possible. Perhaps the simplest way of giving such graphs is to divide the type space $(0,1]$ into two intervals $I_1 = (0, x_0]$ and $I_2 = (x_0, 1]$, take $\varphi(x) = x^{-1/\alpha}$ on I_1 and $\varphi(x) = (x-x_0)^{-1/\alpha}$ on I_2, to set $\kappa_2(x,y) = A_1\varphi(x)\varphi(y)$ if one of x is in I_1 and the other in I_2, and $\kappa_2(x,y) = A_2\varphi(x)\varphi(y)$ otherwise, and to define $\kappa_3(x,y,z)$ to be some constant times $\varphi(x)\varphi(y)\varphi(z)$, where the constant depends on how many of x, y and z line in I_1.

8 Limits of sparse random graphs

Although our main focus in this paper was the introduction of the model $G(n, \underset{\sim}{\kappa})$, and the study of the existence and size of the giant component in this graph, we shall close by briefly discussing some connections to earlier work that arise when considering the local structure of $G(n, \underset{\sim}{\kappa})$.

Let us start by considering subgraph counts. As before, let $\mathcal{G}$ consist of one representative of each isomorphism class of finite graphs, and let $\mathcal{F} \subset \mathcal{G}$ consist of the connected graphs in $\mathcal{G}$. Given two graphs F and G, let $\hom(F,G)$ be the number of homomorphisms from F to G, and $\mathrm{emb}(F,G)$ the number of embeddings, so $\mathrm{emb}(F,G) = n(F,G)\,\mathrm{aut}(F)$. Writing G_n for a graph with n vertices, in the dense case, where G_n has $\Theta(n^2)$ edges, one can combine the normalized subgraph or embedding counts

$$
s(F, G_n) = n(F,G_n)/n(F,K_n) = \mathrm{emb}(F,G_n)/\,\mathrm{emb}(F,K_n)
$$

to define a metric that turns out to have very nice properties. (Often one uses the equivalent homomorphism densities $t(F, G_n) = \hom(F, G_n)/n^{|F|}$, but when we come to sparse graphs embeddings are more natural than homomorphisms.) A sequence (G_n) converges in this *subgraph metric* if and only if there are constants $s(F)$, $F \in \mathcal{F}$, such that $s(F, G_n) \to s(F)$ for each $F \in \mathcal{F}$. Lovász and Szegedy [32] characterised the possible limits $(s(F))_{F\in\mathcal{F}}$, both in terms of kernels and algebraically.

Borgs, Chayes, Lovász, Sós and Vesztergombi [17, 18] introduced the cut metric $\delta_\square$ that we used in Section 4. They showed that this metric is equivalent to the subgraph metric, as well as to various other notions of convergence for sequences of dense graphs. One of the nicest features of these results is that for every point in the completion of the space of finite graphs (with respect to any of these metrics), there is a natural random graph model (called a W-random graph in [32]) that produces sequences of graphs tending to this point. (See also Diaconis and Janson [23], where connections to certain infinite random graphs are described.)

Turning to sparse graphs, as described in [16], the situation is much less simple. When G_n has $\Theta(n)$ edges, as here, the natural normalization is to consider, for each connected F,

$$\tilde{s}(F, G_n) = \mathrm{emb}(F, G_n)/n = \mathrm{aut}(F)n(F, G_n)/n.$$

Under suitable additional assumptions on the sequences G_n, one can again combine these counts to define a metric, and consider the possible limit points. Unfortunately, not much is known about these; see the discussion in [16].

Turning to our present model, Theorem 35 shows that if $\underset{\sim}{\kappa}$ is a kernel family with only finitely many non-zero kernels and all moments finite, then $\tilde{s}(F, G_n) \xrightarrow{\mathrm{p}} t(F, \underset{\sim}{\kappa})$ for all connected F, where $G_n = G(n, \underset{\sim}{\kappa})$ and $t(F, \underset{\sim}{\kappa})$ is given by (53). This suggests the following question.

Question 1. Is there a simple characterization of those vectors $(t_F)_{F\in\mathcal{F}}$ for which there is an integrable kernel family $\underset{\sim}{\kappa}$ such that $t_F = t(F, \underset{\sim}{\kappa})$ for all $F \in \mathcal{F}$?

As unbounded kernel families may cause technical difficulties, it may make sense to ask the same question with the restriction that $\underset{\sim}{\kappa}$ should be bounded.

Note that Question 1 is very different from the question answered by Lovász and Szegedy [32]: our definition of $t(F, \underset{\sim}{\kappa})$ is different from the corresponding notion studied there, since it is adapted to the setting of sparse graphs. In particular, if $\underset{\sim}{\kappa}$ consists only of a single kernel κ_2 (as in [32]), then we have $t(F, \underset{\sim}{\kappa}) = 0$ for any F that is not a tree.

As discussed in [16, Question 7.1], it is an interesting question to ask whether, for various natural metrics on sparse graphs, one can provide natural random graph models corresponding to points in the completion. For those vectors (t_F) where the answer to Question 1 is yes, the model $G(n, \underset{\sim}{\kappa})$ provides a positive answer (at least if $\underset{\sim}{\kappa}$ is bounded, say). But these points will presumably only be a very small subset of the possible limits, so there are many corresponding models still to be found.

As noted in [16, Section 6.6], rather than considering subgraph counts $\tilde{s}(F, G_n)$, for graphs with $\Theta(n)$ edges it is more natural to consider directly the probability that the t-neighbourhood of a random vertex v is a certain graph F; the subgraph counts may be viewed as moments of these probabilities.

More precisely, let $\mathcal{G}^{\mathrm{r}}$ be the set of isomorphism classes of connected, locally finite rooted graphs, and for $t \geq 0$, let $\mathcal{G}_t^{\mathrm{r}}$ be the set of isomorphism classes of finite connected rooted graphs with *radius* at most t, i.e., in which all vertices are within distance t of the root. A probability distribution π on $\mathcal{G}^{\mathrm{r}}$ naturally induces a probability distribution π_t on each $\mathcal{G}_t^{\mathrm{r}}$, obtained by taking a π-random element of $\mathcal{G}^{\mathrm{r}}$ and deleting any vertices at distance more than t from the root. Given $F \in \mathcal{G}_t^{\mathrm{r}}$ and a graph G_n with n vertices, let $p_t(F, G_n)$ be the probability that a random vertex v of G_n has the property that its neighbourhoods up to distance t form a graph isomorphic to F, with v as the root. A sequence (G_n) with $|G_n| \to \infty$ has *local limit* π if

$$p_t(F, G_n) \to \pi_t(F)$$

for every $F \in \mathcal{G}_t^{\mathrm{r}}$ and all $t \geq 0$. This notion has been introduced in several different contexts under different names: Aldous and Steele [4] used the term 'local weak limit', and Aldous and Lyons [3], 'random weak limit'. Also, Benjamini and Schramm [7] defined a corresponding 'distributional limit' of certain random graphs. Notationally it is convenient to map a graph G_n to the point $\phi(G_n) = (p_t(F, G_n)) \in X = \prod_t [0,1]^{\mathcal{G}_t^{\mathrm{r}}}$, and to define $\phi(\pi)$ similarly. Taking any metric d on X giving rise to the product topology, we obtain a metric d_{loc} on the set of graphs together with probability distributions on $\mathcal{G}^{\mathrm{r}}$, and (G_n) has local limit π if and only if $d_{\mathrm{loc}}(G_n, \pi) \to 0$.

As noted in [16], under suitable assumptions (which will hold here if $\underset{\sim}{\kappa}$ is bounded, for example), the two notions of convergence described above are equivalent, and one can pass from the limiting normalized subgraph counts $\tilde{s}(F)$ to the distribution π and *vice versa*. Also, if κ is a bounded kernel, then the random graphs $G(n, \kappa)$ defined in [10] have as local limit a certain distribution associated to π. This latter observation extends to the present model, and as we shall now see, no boundedness restriction is needed.

Given an integrable hyperkernel $\underset{\sim}{\kappa}$, let $G_{\underset{\sim}{\kappa}}$ be the random (potentially infinite) rooted graph associated to the branching process $\mathfrak{X}_{\underset{\sim}{\kappa}}$. This is defined in the natural way: we take the root of $\mathfrak{X}_{\underset{\sim}{\kappa}}$ as the root vertex, for each child clique of the root we take a complete graph in $G_{\underset{\sim}{\kappa}}$, with these cliques sharing only the root vertex. Each child w of the root then corresponds to a non-root vertex in one of these cliques, and we add further cliques meeting only in w to correspond to the child cliques of w, and so on.

More generally, given an integrable kernel family $\underset{\sim}{\kappa} = (\kappa_F)_{F \in \mathcal{F}}$, we may define a random rooted graph $G_{\underset{\sim}{\kappa}}$ in an analogous way; we omit the details. We write $\pi_{\underset{\sim}{\kappa}}$ for the probability distribution on $\mathcal{G}^{\mathrm{r}}$ associated to $G_{\underset{\sim}{\kappa}}$.

Theorem 37. *Let $\underset{\sim}{\kappa}$ be an integrable kernel family and let $G_n = G(n, \underset{\sim}{\kappa})$. Then $d_{\mathrm{loc}}(G_n, \pi_{\underset{\sim}{\kappa}}) \xrightarrow{\mathrm{p}} 0$.*

In fact, we conjecture that almost sure convergence holds for any coupling of the G_n for different n, and in particular if the different G_n are taken to be independent. (The case of independent G_n is the extreme case, which by standard arguments implies a.s. convergence for every other coupling too; a.s. convergence in this case is known as *complete convergence*.)

Writing $\pi_{\underset{\sim}{\kappa},t}$ for the probability distribution on $\mathcal{G}_t^{\mathrm{r}}$ induced by $\pi_{\underset{\sim}{\kappa}}$, by definition we have $d_{\mathrm{loc}}(G_n,\pi_{\underset{\sim}{\kappa}}) \xrightarrow{\mathrm{p}} 0$ if and only if

$$p_t(F,G_n) \xrightarrow{\mathrm{p}} \pi_{\underset{\sim}{\kappa},t}(F) \tag{74}$$

for each t and each $F \in \mathcal{G}_t^{\mathrm{r}}$. The special case where $\underset{\sim}{\kappa}$ is a bounded hyperkernel is essentially immediate: (74) is simply a formal statement of the local coupling established for bounded hyperkernels in Section 3. Exactly the same argument applies to a bounded kernel family. For the extension to general kernel families we need a couple of easy lemmas.

Lemma 38. *Let $\underset{\sim}{\kappa}$ be an edge integrable kernel family. For any $\varepsilon > 0$ there is a $\delta = \delta_1(\underset{\sim}{\kappa},\varepsilon) > 0$ such that whp any δn vertices of $G(n,\underset{\sim}{\kappa})$ meet at most εn edges.*

Proof. This is an extension of Proposition 8.11 of [10]; the proof carries over *mutatis mutandis*, using Theorem 33 with $F = P_2$ to bound the sum of the squares of the vertex degrees in the bounded case. The key step is to use edge integrability to find a bounded kernel family $\underset{\sim}{\kappa}'$ such that $G(n,\underset{\sim}{\kappa}')$ may be regarded as a subgraph of $G(n,\underset{\sim}{\kappa})$ containing all but at most $\varepsilon n/2 + o_{\mathrm{p}}(n)$ of the edges. □

It turns out that we can weaken edge integrability to integrability. The price we pay is that we cannot control the number of edges incident to a small set of vertices, but only the size of the neighbourhood. As usual, given a set A of vertices in a graph G, we write $N^t(A)$ for the set of vertices at graph distance at most t from A, so $A \subset N(A) = N^1(A) \subset N^2(A) \cdots$.

Lemma 39. *Let $\underset{\sim}{\kappa}$ be an integrable kernel family. For any $\varepsilon > 0$ there is a $\delta = \delta_2(\underset{\sim}{\kappa},\varepsilon) > 0$ such that whp every set A of at most δn vertices of $G(n,\underset{\sim}{\kappa})$ satisfies $|N(A)| \le \varepsilon n$.*

Proof. Replacing each atom by a clique, we may and shall assume that $\underset{\sim}{\kappa}$ is a hyperkernel. Let $\underset{\sim}{\kappa}'$ be the kernel family obtained from $\underset{\sim}{\kappa}$ by replacing each clique by a star. Since $\underset{\sim}{\kappa}$ is integrable, $\underset{\sim}{\kappa}'$ is edge integrable. Let $\delta_1(\varepsilon) = \delta_1(\underset{\sim}{\kappa}',\varepsilon)$ be the function given by Lemma 38, and set $\delta = \delta_1(\delta_1(\varepsilon)) > 0$. Then whp every set A of at most δn vertices of $G(n,\underset{\sim}{\kappa}')$ has $|N(A)| \le \delta_1(\varepsilon)n$ and hence $|N^2(A)| \le \varepsilon n$. Coupling $G(n,\underset{\sim}{\kappa})$ and $G(n,\underset{\sim}{\kappa}')$ in the obvious way, vertices adjacent in $G(n,\underset{\sim}{\kappa})$ are at distance at most 2 in $G(n,\underset{\sim}{\kappa}')$, and the result follows. □

Let $v(G(n,\underset{\sim}{\kappa}))$ be the sum of the orders of the atoms making up $G(n,\underset{\sim}{\kappa})$. Our final lemma relates this sum to $\int \underset{\sim}{\kappa} = \sum_F |F| \int_{\mathcal{S}^{|F|}} \kappa_F$.

Lemma 40. *Let $\underset{\sim}{\kappa} = (\kappa_F)_{F\in\mathcal{F}}$ be an integrable kernel family. Then $v(G(n,\underset{\sim}{\kappa})) = n\int\underset{\sim}{\kappa} + o_{\mathrm{p}}(n)$.*

Proof. Let X_r be the number of atoms with order r, and $X = v(G(n,\underset{\sim}{\kappa})) = \sum_{r\geq 2} rX_r$. Let $c = \int\underset{\sim}{\kappa}$, and let c_r be the contribution to $\int\underset{\sim}{\kappa}$ from kernels corresponding to graphs F with r vertices, so $\mathbb{E}(rX_r) = (n)_r c_r/n^{r-1} \sim nc_r$ and $\sum c_r = c < \infty$. Given $\varepsilon > 0$, there is an R such that $\sum_{r\leq R} c_r \geq c - \varepsilon$. Each X_r has a Poisson distribution and is thus concentrated about its mean, so whp

$$X \geq \sum_{r\leq R} rX_r \geq \sum_{r\leq R} c_r n - \varepsilon n \geq (c - 2\varepsilon)n.$$

Writing $(x)_+$ for $\max\{x, 0\}$, since ε was arbitrary we have shown that $(c - X/n)_+ \xrightarrow{\mathrm{p}} 0$. Since $c - X/n$ is bounded, it follows that $\mathbb{E}(c - X/n)_+ \to 0$. But $\mathbb{E}X \leq cn$, so $\mathbb{E}(X/n - c)_+ = \mathbb{E}(X/n - c) + \mathbb{E}(c - X/n)_+ \to 0$. Hence $(X/n - c)_+ \xrightarrow{\mathrm{p}} 0$, so $X/n \xrightarrow{\mathrm{p}} c$ as claimed. □

Combining the last two lemmas, we can now prove Theorem 37.

Proof of Theorem 37. As noted after the statement of the theorem, the case where $\underset{\sim}{\kappa}$ is bounded is straightforward.

Let $\underset{\sim}{\kappa}$ be an integrable kernel family, and let $G_n = G(n,\underset{\sim}{\kappa})$. Fix $t \geq 1$, $F \in \mathcal{G}_t^{\mathrm{r}}$, and $\varepsilon > 0$. It suffices to prove that

$$|p_t(F, G_n) - \pi_{\underset{\sim}{\kappa},t}(F)| \leq \varepsilon + o_{\mathrm{p}}(1). \tag{75}$$

Then letting $\varepsilon \to 0$ we have $p_t(F, G_n) \xrightarrow{\mathrm{p}} \pi_{\underset{\sim}{\kappa},t}(F)$, so (74) holds. Since t and F are arbitrary, this implies $d_{\mathrm{loc}}(G_n, \pi_{\underset{\sim}{\kappa}}) \xrightarrow{\mathrm{p}} 0$.

Applying Lemma 39 t times, there is a $\delta > 0$ such that whp any set A of at most δn vertices of G_n satisfies $|N^t(A)| \leq \varepsilon n/2$. Since $\underset{\sim}{\kappa}$ is integrable, there is a bounded kernel family $\underset{\sim}{\kappa}^M$ which satisfies $\underset{\sim}{\kappa}^M \leq \underset{\sim}{\kappa}$ pointwise and $\int\underset{\sim}{\kappa} - \int\underset{\sim}{\kappa}^M \leq \delta/2$. As $M \to \infty$, we have $\underset{\sim}{\kappa}^M \nearrow \underset{\sim}{\kappa}$ pointwise, and it follows that $\pi_{\underset{\sim}{\kappa}^M,t}(F) \to \pi_{\underset{\sim}{\kappa},t}(F)$; the argument is as for Theorem 19(i). Taking M large enough, we may thus assume that $\left|\pi_{\underset{\sim}{\kappa}^M,t}(F) - \pi_{\underset{\sim}{\kappa},t}(F)\right| \leq \varepsilon/2$. Let $G'_n = G(n,\underset{\sim}{\kappa}^M)$. Since $\underset{\sim}{\kappa}^M$ is bounded, we have $p_t(F, G'_n) \xrightarrow{\mathrm{p}} \pi_{\underset{\sim}{\kappa}^M,t}(F)$. Coupling G_n and G'_n as usual so that $G'_n \subset G_n$, let B be the set of vertices incident with an atom present in G_n but not G'_n. By Lemma 40 we have $|B| \leq \delta n$ whp, so whp no more than $\varepsilon n/2$ vertices are within distance t of vertices in B. But then $|p_t(F, G_n) - p_t(F, G'_n)| \leq \varepsilon/2$ whp, and (75) follows. □

The general question of which probability distributions on $\mathcal{G}^{\mathrm{r}}$ arise as local limits of sequences of finite graphs seems to be rather difficult. There is a natural necessary condition noted in different forms in all of [3, 4, 7]; see also [16, Section 6.6]. Aldous and Lyons [3] asked whether this condition is sufficient, emphasizing the importance of this open question. Let us finish with a related but perhaps much simpler question: given $\underset{\sim}{\kappa}$, we defined $\mathfrak{X}_{\underset{\sim}{\kappa}}$ as a branching process in which the particles have types. But in the corresponding random graph $G_{\underset{\sim}{\kappa}}$ these types are not recorded. This means that $\underset{\sim}{\kappa}$ cannot simply be read out of the distribution of $G_{\underset{\sim}{\kappa}}$, i.e., out of $\pi_{\underset{\sim}{\kappa}}$. This suggests the following question.

Question 2. Which probability distributions on $\mathcal{G}^{\mathrm{r}}$ are of the form $\pi_{\underset{\sim}{\kappa}}$ for some integrable kernel family $\underset{\sim}{\kappa}$?

Acknowledgement. Part of this research was done during visits of SJ to the University of Cambridge and Trinity College in 2007, and to the Isaac Newton Institute in Cambridge, funded by a Microsoft fellowship, in 2008.

References

[1] R. Albert and A.-L. Barabási, Statistical mechanics of complex networks, *Rev. Mod. Phys.* **74** (2002), 47–97.

[2] R. Albert, H. Jeong and A.-L. Barabási, Error and attack tolerance of complex networks, *Nature* **406** (2000), 378–382.

[3] D. Aldous and R. Lyons, Processes on unimodular random networks, *Electronic Journal of Probability* **12** (2007), 1454–1508.

[4] D. Aldous and J.M. Steele, The objective method: probabilistic combinatorial optimization and local weak convergence, in *Probability on discrete structures*, *Encyclopaedia Math. Sci.* **110**, Springer (2004), pp. 1–72.

[5] A.-L. Barabási and R. Albert, Emergence of scaling in random networks, *Science* **286** (1999), 509–512.

[6] A. D. Barbour, L. Holst and S. Janson, *Poisson Approximation.* Oxford University Press, Oxford, UK, 1992.

[7] I. Benjamini and O. Schramm, Recurrence of distributional limits of finite planar graphs, *Electron. J. Probab.* **6** (2001), no. 23, 13 pp. (electronic).

[8] B. Bollobás, *Modern Graph Theory.* Springer-Verlag, New York, 1998. xiv+394 pp.

[9] B. Bollobás, C. Borgs, J. Chayes and O. Riordan, Percolation on dense graph sequences, to appear in *Annals of Probability.* `arXiv:0701346v2.`

[10] B. Bollobás, S. Janson and O. Riordan, The phase transition in inhomogeneous random graphs, *Random Structures and Algorithms*, **31** (2007), 3–122.

[11] B. Bollobás, S. Janson and O. Riordan, Spread-out percolation in $\mathbf{R}^d$, *Random Structures and Algorithms* **31** (2007), 239–246.

[12] B. Bollobás, S. Janson and O. Riordan, The cut metric, branching processes and random graphs, in preparation.

[13] B. Bollobás and O. Riordan, Mathematical results on scale-free random graphs, in Handbook of Graphs and Networks, Stefan Bornholdt and Heinz Georg Schuster (eds.), Wiley-VCH, Weinheim (2002), 1–34.

[14] B. Bollobás and O. Riordan, Robustness and vulnerability of scale-free random graphs, *Internet Mathematics* **1** (2003), 1–35.

[15] B. Bollobás and O. Riordan, The diameter of a scale-free random graph, *Combinatorica* **24** (2004), 5–34.

[16] B. Bollobás and O. Riordan, Sparse graphs: metrics and random models, preprint (2007). `arXiv:0708.1919`.

[17] C. Borgs, J. T. Chayes, L. Lovász, V. T. Sós and K. Vesztergombi, Convergent sequences of dense graphs I: Subgraph frequencies, metric properties and testing, preprint (2006) (revised 2007). `arXiv:0702004`.

[18] C. Borgs, J.T. Chayes, L. Lovász, V.T. Sós and K. Vesztergombi, Convergent sequences of dense graphs II: Multiway cuts and statistical physics. Preprint, March 2007.

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

[20] D. S. Callaway, M. E. J. Newman, S. H. Strogatz and D. J. Watts, Network robustness and fragility: percolation on random graphs, *Phys. Rev. Lett.* **85** (2000), 5468–5471.

[21] R. Cohen, K. Erez, D. ben-Avraham and S. Havlin, Resilience of the Internet to random breakdowns, *Phys. Rev. Lett.* **85** (2000), 4626–4628.

[22] C. Cooper and A. Frieze, A general model of web graphs, *Random Structures and Algorithms* **22** (2003), 311–335.

[23] P. Diaconis and S. Janson, Graph limits and exchangeable random graphs, *Rendiconti di Matematica* **28** (2008), 33–61.

[24] S.N. Dorogovtsev and J.F.F. Mendes, Evolution of networks, *Adv. Phys.* **51** (2002), 1079.

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

[26] M. Faloutsos, P. Faloutsos and C. Faloutsos, On power-law relationships of the internet topology, SIGCOMM 1999, *Comput. Commun. Rev.* **29** (1999), 251.

[27] A. Frieze and R. Kannan, Quick approximation to matrices and applications, *Combinatorica* **19** (1999), 175–220.

[28] S. Janson, Standard representation of multivariate functions on a general probability space, preprint (2008). `arXiv:0801.0196.`

[29] S. Janson, Asymptotic equivalence and contiguity of some random graphs, preprint (2008). `arXiv:0802.1637.`

[30] P.L. Krapivsky and S. Redner, Organization of growing random networks, *Phys. Rev. E* **66** (2001), 066123.

[31] R. Kumar, P. Raghavan, S. Rajagopalan, D. Sivakumar, A. Tomkins and E. Upfal, Stochastic models for the web graph, FOCS 2000.

[32] L. Lovász and B. Szegedy, Limits of dense graph sequences, *J. Combin. Theory B* **96** (2006), 933–957.

[33] M.E.J. Newman, Assortative mixing in networks, *Phys. Rev. Lett.* **89** (2002), 208701(4).

[34] O. Riordan, The small giant component in scale-free random graphs, *Combinatorics, Probability and Computing* **14** (2005), 897–938.

[35] D. J. Watts and S. H. Strogatz, Collective dynamics of 'small-world' networks, *Nature* **393** (1998), 440–442.