Combinatorics, Probability and Computing (2008) 17, 111–136. © 2007 Cambridge University Press
doi:10.1017/S0963548307008589 First published online 27 June 2007 Printed in the United Kingdom

The k-Core and Branching Processes

OLIVER RIORDAN

Royal Society Research Fellow, Department of Pure Mathematics and Mathematical Statistics,
University of Cambridge, Cambridge CB3 0WB, UK
(e-mail: O.M.Riordan@dpmms.cam.ac.uk)

Received 3 November 2005; revised 5 February 2007

The k-*core* of a graph G is the maximal subgraph of G having minimum degree at least k. In 1996, Pittel, Spencer and Wormald found the threshold λ_c for the emergence of a non-trivial k-core in the random graph $G(n, \lambda/n)$, and the asymptotic size of the k-core above the threshold. We give a new proof of this result using a local coupling of the graph to a suitable branching process. This proof extends to a general model of inhomogeneous random graphs with independence between the edges. As an example, we study the k-core in a certain power-law or 'scale-free' graph with a parameter c controlling the overall density of edges. For each $k \geqslant 3$, we find the threshold value of c at which the k-core emerges, and the fraction of vertices in the k-core when c is ε above the threshold. In contrast to $G(n, \lambda/n)$, this fraction tends to 0 as $\varepsilon \to 0$.

1. Introduction

The k-*core* of a graph G is the maximal subgraph of G with minimum degree at least k; if G has no such (non-trivial) subgraph, then the k-core of G is empty. This concept was introduced by Bollobás [3], in the context of finding large k-connected subgraphs of random graphs. As edges are added one by one to a graph, the k-core grows; in particular, it is empty up to some point, and then non-empty (and often large). The question of when the k-core emerges in a random graph (or random graph process) also arose in the context of finding the chromatic number of sparse random graphs; more specifically, let $G(n, p)$ be the Erdős–Rényi random graph with n vertices, in which the possible edges are present independently, each with probability p. If a graph has no k-core, it is k-colourable; Chvátal [8] used this to show that $G(n, 2.88/n)$ is w.h.p. 3-colourable. Here, as usual, an event holds *with high probability*, or w.h.p., if it holds with probability $1 - o(1)$ as $n \to \infty$.

It is natural to ask: For $k \geqslant 3$ fixed, what is the critical value $\lambda_c = \lambda_c(k)$ of λ above which a (non-empty) k-core first appears w.h.p. in $G(n, \lambda/n)$?

There is a natural 'guess' as to the answer, given by a branching process: let X_λ be a Galton–Watson branching process that starts with a single particle x_0 in generation zero, where the number of children of each particle has a Poisson distribution with mean λ, and these numbers

https://doi.org/10.1017/S0963548307008589 Published online by Cambridge University Press

are independent for different particles. Of course, X_λ provides a good 'local' model of the neighbourhood of a vertex of $G(n,\lambda/n)$. Let $\mathscr{B}^+$ be the event that x_0 has at least k children, each of which has at least $k-1$ children, each of which has at least $k-1$ children, each of which ..., *i.e.*, that x_0 is in a k-regular tree contained in X_λ. (We reserve the notation $\mathscr{B}$ for an associated event whose role in the analysis is more fundamental.) Let $\beta^+(\lambda)$ be the probability that X_λ has the property $\mathscr{B}^+$. One might expect that, up to probability $o(1)$, a vertex is in the k-core if and only if its neighbourhoods up to a suitable distance have a property corresponding to $\mathscr{B}^+$, and thus that the fraction of vertices of $G(n,\lambda/n)$ in the k-core is $\beta^+(\lambda)+o(1)$; this turns out to be the case.

Pittel, Spencer and Wormald [21] showed that, except at the critical point, the number of vertices in the k-core of $G(n,\lambda/n)$ is indeed $\beta^+(\lambda)n+o_p(n)$. In particular, the threshold λ_c for the emergence of the k-core is $\lambda_c=\inf\{\lambda:\beta^+(\lambda)>0\}$. Recently, simpler proofs of this result have been given, as well as generalizations to various other contexts: see, for example, [9, 19, 13, 6, 15, 10]. As far as we are aware, although the branching process heuristic was described already in [21], none of these proofs works by directly coupling the neighbourhoods of a vertex of the graph with the branching process; there seems to be a problem relating the inescapably global property of lying in the k-core to a simple local property. Here we give a new proof that does proceed in this way, using a carefully chosen (and not very simple) local property.

As the proof will be a little involved, we give a very rough outline to motivate what follows: we shall define a certain event $\mathscr{A}$ depending on the first D generations of X_λ, where $D=o(\log n)$ but $D/\log\log n\to\infty$, such that $\mathbb{P}(\mathscr{A})$ is almost as large as $\mathbb{P}(\mathscr{B}^+)$, and $\mathbb{P}(\mathscr{B}^+\mid\mathscr{A})=1-o(n^{-1})$. As $\mathscr{A}$ is a 'local' event, the number of vertices in $G(n,\lambda/n)$ with (neighbourhoods having the equivalent of) property $\mathscr{A}$ will be $\mathbb{P}(\mathscr{A})n+o_p(n)$. Defining $\mathscr{A}$ in the right way, we can show that, given that a vertex has property $\mathscr{A}$, with probability $1-o(n^{-1})$ it is the root of a k-regular tree of height at most D all of whose leaves have property $\mathscr{A}$. But then w.h.p. *every* vertex with property $\mathscr{A}$ is the root of such a tree, and the union of these trees has minimum degree at least k. Most of the work will be in the analysis of the branching process: having found the right event, the translation to the graph will be relatively straightforward.

Unfortunately, the event $\mathscr{A}$ we have to work with is rather complicated. Roughly speaking, we look for a tree T of height D that branches a little faster than a k-regular tree. Having found such a tree, if we select leaves of T independently with an appropriate probability (around $\beta(\lambda)$), it is very likely that we can find a k-regular tree whose leaves are a subset of the selected leaves. When we explore the neighbourhood of a vertex v in the graph, finding a tree T as above, we then wish to explore from each leaf w of T to see whether the neighbourhoods of w have the property $\mathscr{A}$. In doing this, it turns out that we cannot afford to ignore the edge along which we reached w when exploring from v, and that we must re-use parts of the tree T in establishing that w has property $\mathscr{A}$. This, together with the need to achieve (approximate) independence of the explorations from different leaves w, places rather subtle constraints on the properties $\mathscr{A}$ we can use.

The next section is devoted to the study of the k-core in $G(n,\lambda/n)$. In Section 3, we turn to inhomogeneous random graphs: an advantage of the direct branching process approach is that it extends easily to other random graph models in which the neighbourhoods of a vertex can be modelled by a suitable branching process. This includes the general sparse inhomogeneous model of Bollobás, Janson and Riordan [5]. Indeed, the present work started at the conference *Random Structures and Algorithms 2005* in Poznań, when Alan Frieze asked whether we knew the size of the k-core in this model. As a special case, we study the graph on $[n]$ in which edges

https://doi.org/10.1017/S0963548307008589 Published online by Cambridge University Press

are present independently, and the probability of an edge between i and j is $c/\sqrt{ij}$ for a parameter $c > 0$. We find the threshold (in terms of c) at which the k-core appears in this graph, and the size of the k-core above this threshold; in contrast to $G(n, \lambda/n)$, just above the threshold the k-core is small.

2. The uniform case

We start with some standard basic preliminaries. Let us write $\mathbb{P}_\lambda$ for the probability measure associated to the branching process X_λ; when discussing X_λ, we shall use the terms 'event' and 'property of the branching process' interchangeably, to refer to a (measurable) subset of the set of all rooted, unlabelled trees. When we say that a particle x of the branching process has a certain property, we mean that the process consisting of x (as the new root) and its descendants has the property. Throughout we write x_0 for the initial particle of X_λ. When λ is clear from the context (or specified in the notation for our probability measure), we may write X for X_λ. Many of the events and quantities we shall define below depend on k. Almost always, we regard $k \geqslant 2$ as fixed and suppress this dependence.

Let $\Psi_{\geqslant t}(\lambda) = \mathbb{P}(\mathrm{Po}(\lambda) \geqslant t)$ denote the probability that a Poisson random variable with mean λ takes a value that is at least t. Let $\mathscr{B}_d$ be the event that X contains a $(k-1)$-ary tree of height d with x_0 as the root, and let $\mathscr{B} = \lim_{d\to\infty} \mathscr{B}_d$ be the event that X contains an infinite $(k-1)$-ary tree with root x_0. Then $\mathbb{P}_\lambda(\mathscr{B}_0) = 1$. Also, each particle in the first generation of X_λ has probability $\mathbb{P}_\lambda(\mathscr{B}_d)$ of having property $\mathscr{B}_d$. As these events are independent for different particles, the number of particles in the first generation with property $\mathscr{B}_d$ has a Poisson distribution with mean $\lambda\mathbb{P}_\lambda(\mathscr{B}_d)$. Thus, $\mathbb{P}_\lambda(\mathscr{B}_{d+1}) = \Psi_{\geqslant k-1}\big(\lambda\mathbb{P}_\lambda(\mathscr{B}_d)\big)$.

Since $\Psi_{\geqslant k-1}(\lambda x)$ is a continuous, increasing function of x, it follows that

$$\mathbb{P}_\lambda(\mathscr{B}) = \lim_{d\to\infty} \mathbb{P}_\lambda(\mathscr{B}_d)$$

is given by the maximum solution p to the equation $p = \Psi_{\geqslant k-1}(\lambda p)$. We denote this solution by $\beta(\lambda)$. As in the Introduction, let $\mathscr{B}^+$ be the event that X contains an infinite k-regular tree with root x_0, *i.e.*, that x_0 has at least k children with property $\mathscr{B}$; we shall write $\mathscr{B}_d^+$ for the corresponding event depending only on the first d generations, *i.e.*, that x_0 has at least k children with property $\mathscr{B}_{d-1}$. Note that

$$\beta^+(\lambda) := \mathbb{P}_\lambda(\mathscr{B}^+) = \Psi_{\geqslant k}\big(\lambda\mathbb{P}_\lambda(\mathscr{B})\big) = \Psi_{\geqslant k}\big(\lambda\beta(\lambda)\big).$$

Let $\lambda_c = \inf\{\lambda : \beta(\lambda) > 0\}$. It is easy to see that (in this uniform case) the functions $\beta(\lambda)$ and $\beta^+(\lambda)$ have a jump at $\lambda = \lambda_c$ (for $k \geqslant 3$), and are continuous and strictly increasing for $\lambda \geqslant \lambda_c$.

Pittel, Spencer and Wormald [21] proved the following result, with sharper error estimates. As usual, if A_n is a sequence of random variables, we write $A_n = o_p(f(n))$ if $A_n/f(n) \to 0$ in probability, *i.e.*, if $|A_n| \leqslant \varepsilon f(n)$ w.h.p. for any $\varepsilon > 0$.

Theorem 2.1 ([21]). *Let $k \geqslant 3$ be fixed, and define λ_c and $\beta^+(\lambda)$ as above. If $\lambda \neq \lambda_c$ is constant, then the number of vertices of $G(n, \lambda/n)$ in the k-core is $\beta^+(\lambda)n + o_p(n)$ as $n \to \infty$.*

Note that the 2-core of a graph G is just the union of the cycles in G, together with all paths joining two cycles. It is easy to see that a large 2-core first appears in $G(n, \lambda/n)$ at the same

https://doi.org/10.1017/S0963548307008589 Published online by Cambridge University Press

time as the giant component, *i.e.*, when $\lambda > \lambda_c(2) = 1$, although there may be a small 2-core when $\lambda < 1$. With the weak error bounds above, Theorem 2.1 holds for $k = 2$ as well. For us, there will be no difference between the cases $k = 2$ and $k \geqslant 3$, although the latter is the more interesting.

The proof of the upper bound is easy; we postpone this to the end of the section. Our proof of the lower bound will require considerable preparation. Although the functions $\beta(\lambda)$ and $\beta^+(\lambda)$ are continuous except at $\lambda = \lambda_c$, we shall avoid using this fact, with an eye to generalizations.

If $\mathscr{E}_1$, $\mathscr{E}_2$ are properties of the branching process, with $\mathscr{E}_1$ depending only on the first d generations, let $\mathscr{E}_1 \circ \mathscr{E}_2$ denote the event that $\mathscr{E}_1$ holds if we delete from X all particles in generation d that do not have property $\mathscr{E}_2$. For example, with $\mathscr{B}_1$ the property of having at least $k-1$ children, as above, $\mathscr{B}_1 \circ \mathscr{B}_1 = \mathscr{B}_2$, the property of having at least $k-1$ children with at least $k-1$ children (and perhaps other children with fewer than $k-1$ children).

Let $\mathscr{R}_d$ be the event that $\mathscr{B}_d$ holds in a *robust* manner, meaning that $\mathscr{B}_d$ holds even after any single particle in generation d is deleted. Then, as $\mathscr{B} = \mathscr{B}_d \circ \mathscr{B}$, the event $\mathscr{R}_d \circ \mathscr{B}$ is the event that $\mathscr{B}$ holds after any single particle in generation d is deleted. Of course, whenever a particle is deleted, so are its descendants.

Lemma 2.2. *If $\lambda \mapsto \beta(\lambda)$ is continuous at λ then*

$$\mathbb{P}_\lambda(\mathscr{R}_d \circ \mathscr{B}) \nearrow \beta(\lambda) := \mathbb{P}_\lambda(\mathscr{B})$$

as $d \to \infty$.

Proof. Fix $0 < \varepsilon < 1$. Consider the natural coupling of X_λ and $X_{(1-\varepsilon)\lambda}$, obtained by constructing X_λ, and then deleting each edge (of the rooted tree) independently with probability ε, and taking for $X_{(1-\varepsilon)\lambda}$ the set of particles still connected to the root. In this coupling, if $\mathscr{B}$ does not hold for X_λ, it certainly does not hold for $X_{(1-\varepsilon)\lambda}$. Furthermore, if $\mathscr{B} \setminus (\mathscr{R}_d \circ \mathscr{B})$ holds for X_λ, then there is some particle x in generation d such that if x is deleted, then $\mathscr{B}$ no longer holds. But the probability that x is not deleted when passing to $X_{(1-\varepsilon)\lambda}$ is exactly $(1-\varepsilon)^d$. It follows that

$$\mathbb{P}_{(1-\varepsilon)\lambda}(\mathscr{B}) \leqslant \mathbb{P}_\lambda(\mathscr{R}_d \circ \mathscr{B}) + (1-\varepsilon)^d.$$

Since $\mathscr{R}_d \subset \mathscr{R}_{d+1}$, the sequence $\mathbb{P}_\lambda(\mathscr{R}_d \circ \mathscr{B})$ is increasing. Taking the limit of the inequality above,

$$\beta\big((1-\varepsilon)\lambda\big) = \mathbb{P}_{(1-\varepsilon)\lambda}(\mathscr{B}) \leqslant \lim_{d\to\infty} \mathbb{P}_\lambda(\mathscr{R}_d \circ \mathscr{B}).$$

Letting $\varepsilon \to 0$, the lemma follows. □

It will often be convenient to *mark* some subset of the particles in generation d. If $\mathscr{E}$ is an event depending on the first d generations, then we write $\mathscr{E} \circ M$ for the event that $\mathscr{E}$ holds after deleting all unmarked particles in generation d. We write $\mathbb{P}_\lambda^p(\mathscr{E} \circ M)$ for the probability that $\mathscr{E} \circ M$ holds when, given X_λ, we mark the particles in generation d independently with probability p. We suppress d from the notation, since it will be clear from the event $\mathscr{E}$. Let

$$r(\lambda, d, p) = \mathbb{P}_\lambda^p(\mathscr{R}_d \circ M).$$

https://doi.org/10.1017/S0963548307008589 Published online by Cambridge University Press

Lemma 2.3. *Let $\lambda_1 < \lambda_2$ be fixed, with $\lambda \mapsto \beta(\lambda)$ continuous at λ_1. Then there is a d such that*

$$r\big(\lambda_2, d, \beta(\lambda_1)\big) \geqslant \beta(\lambda_1). \tag{2.1}$$

Proof. Let us construct a branching process X' as follows. Start with X_{λ_1}. Then, independent of X_{λ_1}, add a $\mathrm{Po}(\lambda_2 - \lambda_1)$ number of 'extra' children of the initial particle. Each extra child then has descendants as in X_{λ_1}. Clearly, we may consider X' as a subset of X_{λ_2}. Given $d \geqslant 1$, let us mark each particle in generation d if it has property $\mathscr{B}$. Now the descendants of a particle in generation d have the distribution of X_{λ_1}, with independence for different particles. Thus, given the first d generations of X', we mark each particle in generation d independently with probability $p = \beta(\lambda_1)$. Hence, $r(\lambda_2, d, p)$ is at least the probability that $\mathscr{R}_d \circ \mathscr{B}$ holds in X'.

Each extra child has property $\mathscr{B}$ independently with probability p. Hence, the event $\mathscr{E}_1$ that there are at least k extra children with property $\mathscr{B}$ has probability $\delta = \Psi_{\geqslant k}((\lambda_2 - \lambda_1)p) > 0$. Now $\mathscr{E}_1$ is independent of X_{λ_1}. Also, if $\mathscr{E}_1$ holds in X', then so does $\mathscr{R}_1 \circ \mathscr{B}$, and hence $\mathscr{R}_d \circ \mathscr{B}$ for any d. Thus, the probability that $\mathscr{R}_d \circ \mathscr{B}$ holds in X' is at least

$$r_d = 1 - (1 - \delta)\big(1 - \mathbb{P}_{\lambda_1}(\mathscr{R}_d \circ \mathscr{B})\big).$$

By Lemma 2.2, as $d \to \infty$ we have $\mathbb{P}_{\lambda_1}(\mathscr{R}_d \circ \mathscr{B}) \to \beta(\lambda_1) = p$, so

$$r_d \to 1 - (1 - \delta)(1 - p) > p,$$

and there is a d with $r_d \geqslant p$, completing the proof. □

We shall use robustly branching trees, *i.e.*, rooted trees with height d having property $\mathscr{R}_d$, in two ways. The first is the obvious way: we shall use the fact that at least two leaves must be deleted in order to destroy all branching subtrees, *i.e.*, all subtrees with property $\mathscr{B}_d$. The second is less direct, and is described in the next lemma, for which we first need a definition.

We say that a finite rooted tree has the property $\mathscr{M}_{\leqslant d}$ if all its leaves are at distances between 1 and d from the root, and every non-leaf has degree at least k.

Lemma 2.4. *Let $\mathscr{T}$ be a rooted tree of height $d+1$ that is minimal with respect to having property $\mathscr{R}_d \circ \mathscr{B}_1$. Let w be a vertex of $\mathscr{T}$ at distance d from the root, and let y be the parent of w. If the graph $\mathscr{T}$ is regarded as a rooted tree with root w, then it has a subtree $\mathscr{W}$ with the property $\mathscr{M}_{\leqslant 2d+1}$, such that all leaves of $\mathscr{W}$ are leaves of $\mathscr{T}$.*

Proof. The proof is illustrated in Figure 1. Although it is not clear that a written proof adds anything to the figure, to be formal we include one.

Let $\mathscr{T}'$ denote the first d generations of $\mathscr{T}$, a tree with property $\mathscr{R}_d$ in which (by minimality) every non-leaf has at least $k-1$ children. By definition of the property $\mathscr{R}_d$, deleting w from $\mathscr{T}'$ leaves a tree with property $\mathscr{B}_d$; let $\mathscr{T}''$ be a minimal subtree of $\mathscr{T}' - w$ with property $\mathscr{B}_d$, so $\mathscr{T}''$ is a $(k-1)$-ary tree. In $\mathscr{T}$, there is a unique path P from w to $\mathscr{T}''$, meeting $\mathscr{T}''$ at z, say. Consider the tree $\mathscr{W}$ obtained by taking P together with the descendants in $\mathscr{T}$ of every vertex of P. Regarding this tree as rooted at w, the children of z in $\mathscr{W}$ are exactly its children in $\mathscr{T}''$, of which there are $k-1$. Furthermore, any vertex a of P other than z and w has the same number of children in $\mathscr{W}$ as in $\mathscr{T}$ (we lose one neighbour on P and gain the other), which is at least

https://doi.org/10.1017/S0963548307008589 Published online by Cambridge University Press

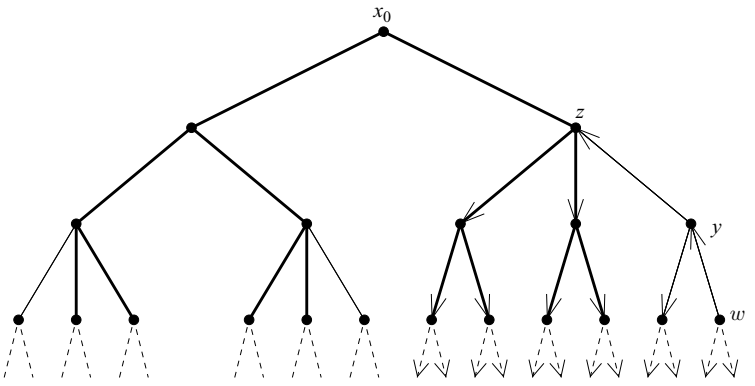


Figure 1. A tree $\mathscr{T}$ with property $\mathscr{R}_3 \circ \mathscr{B}_1$, where $k = 3$. The last generation is drawn with dotted lines; the solid tree $\mathscr{T}'$ has property $\mathscr{R}_3$. The thick lines are a minimal subtree $\mathscr{T}''$ of $\mathscr{T}' - w$ with property $\mathscr{B}_3$. The arrows consist of the shortest path wz from w to $\mathscr{T}''$ and all descendants of vertices in this path. Taking w as the root,they form a tree with property $\mathscr{M}_{\leqslant 5}$.

$k-1$. All other vertices of $\mathscr{W}$ have the same children in $\mathscr{W}$ as in $\mathscr{T}$, except for w, which has one additional child y in $\mathscr{W}$. Thus all non-leaves of $\mathscr{W}$ have at least $k-1$ children, the root w has at least k, and all leaves of $\mathscr{W}$ are leaves of $\mathscr{T}$. Observing finally that the height of $\mathscr{W}$ is at most $2d+1$, the result follows. □

Let us briefly recall Harris's lemma. Let X be a finite set (here consisting of some or all possible edges of $G(n,p)$), and let X_p be a random subset of X formed by selecting each element independently with probability p. An event $\mathscr{A} \subset \mathscr{P}(X)$ is *increasing* if $A \in \mathscr{A}$ and $A \subset B \subset X$ imply $B \in \mathscr{A}$, and *decreasing* if $A \in \mathscr{A}$ and $B \subset A$ imply $B \in \mathscr{A}$.

Lemma 2.5. *If $\mathscr{A}_1, \mathscr{A}_2 \subset \mathscr{P}(X)$ are increasing events and $0 < p < 1$, then*

$$\mathbb{P}(X_p \in \mathscr{A}_1 \cap \mathscr{A}_2) \geqslant \mathbb{P}(X_p \in \mathscr{A}_1)\mathbb{P}(X_p \in \mathscr{A}_2).$$

In other words, increasing events are positively correlated. Of course, it follows that decreasing events are also positively correlated. This was proved by Harris in 1960 [14], and rediscovered by Kleitman [18].

We shall need one simple lemma concerning $G(n,p)$; this is because we wish to prove a result about the neighbourhoods of *every* vertex, and cycles within these neighbourhoods will cause problems. To avoid this, we consider the random graph $\tilde{G} = \tilde{G}(n,\ell,\lambda)$ whose distribution is that of $G(n,\lambda/n)$ conditioned on the absence of any cycles of length at most ℓ. Although globally $G(n,\lambda/n)$ is very likely to contain cycles, locally, conditioning on their absence makes little difference: after this conditioning, given that a certain not too large set of edges is present, and that a certain other set of edges is absent, the probability that another edge is present is close to λ/n, as long as this edge would not complete a 'known' cycle.

Lemma 2.6. *Let a constant $\lambda > 0$ and a function $\ell = \ell(n) = o(\log n)$ be given. If n is large enough then, whenever E_0, E_1 and $\{e\}$ are disjoint sets of possible edges of $\tilde{G} = \tilde{G}(n,\ell,\lambda)$ with*

https://doi.org/10.1017/S0963548307008589 Published online by Cambridge University Press

$|E_1| \leqslant n^{1/3}$ such that $E_1 \cup \{e\}$ contains no cycle of length at most ℓ, we have

$$(1 - n^{-1/4})\lambda/n \leqslant \mathbb{P}\big(e \in E(\tilde{G}) \mid E_1 \subset E(\tilde{G}) \subset E_0^{\mathrm{c}}\big) \leqslant \lambda/n.$$

Proof. Let G' denote the (distribution of the) random graph $\tilde{G}$ conditioned on the presence of every edge in E_1 and the absence of every edge in E_0, so the probability we wish to bound is $\mathbb{P}(e \in G')$. Note that G' may be described as follows: start from $G(n, \lambda/n)$, and first condition on the presence of the edges in E_1 and the absence of the edges in E_0. At this point, each edge of $E_2 = (E_0 \cup E_1)^{\mathrm{c}}$ is present independently with probability λ/n. Writing $\mathscr{P}(S)$ for the power-set of a set S, we next condition on the event $\mathscr{D} \subset \mathscr{P}(E_2)$ that the set of edges E from E_2 present is such that the whole graph does not contain a short (length at most ℓ) cycle, *i.e.*, that $E \cup E_1$ does not contain a short cycle. In terms of the edges in E, the event $\mathscr{D}$ is a decreasing event. The upper bound (which we shall not in fact use) is thus essentially trivial: as the event $e \in E$ is increasing, by Harris's lemma

$$\mathbb{P}(e \in G') = \mathbb{P}(e \in E \mid \mathscr{D}) \leqslant \mathbb{P}(e \in E) = \lambda/n.$$

For the lower bound, let us call a set $P \subset (E_0 \cup E_1 \cup \{e\})^{\mathrm{c}}$ a *pre-cycle* if it is minimal subject to $P \cup E_1 \cup \{e\}$ containing a short cycle that includes the edge e. As legal configurations in the graph G' have the same relative probabilities in G' as in $G(n, \lambda/n)$, the conditional probability $\mathbb{P}(e \in G' \mid G' \setminus \{e\})$ is 0 if $G' \setminus \{e\}$ contains a pre-cycle, and λ/n otherwise. Thus,

$$\begin{aligned}
\mathbb{P}(e \in G') &= \mathbb{E}\left(\mathbb{P}(e \in G' \mid G' \setminus \{e\})\right) \\
&= (\lambda/n)\mathbb{P}(G' \setminus \{e\} \text{ contains no pre-cycle}) \\
&= (\lambda/n)\mathbb{P}(E \text{ contains no pre-cycle} \mid \mathscr{D}) \\
&\geqslant (\lambda/n)\mathbb{P}(E \text{ contains no pre-cycle}),
\end{aligned}$$

where the last step is again from Harris's lemma. Thus, it suffices to show that, with probability at least $1 - n^{-1/4}$, the set E contains no pre-cycle. But any pre-cycle must include a path of length $1 \leqslant t \leqslant \ell$ joining two vertices in V, the set of endpoints of edges in $E_1 \cup \{e\}$. Since $|V| \leqslant 2n^{1/3} + 2$, the expected number of such paths is at most

$$\sum_{t=1}^{\ell} \binom{|V|}{2} n^{t-1}(\lambda/n)^t \leqslant \frac{4n^{2/3}}{n} \sum_{t=1}^{\ell} \lambda^t \leqslant 4n^{-1/3} O(1)^{o(\log n)} \leqslant n^{-1/4}$$

if n is large enough. Hence, the probability that such a path is present is at most $n^{-1/4}$, and so is the probability that E contains a pre-cycle, completing the proof. □

We are now ready to prove Theorem 2.1.

Proof of Theorem 2.1. We start with the more difficult bound, the lower bound on the size of the k-core. Note that $\beta(\lambda)$ and $\beta^+(\lambda)$ are continuous except at $\lambda = \lambda_{\mathrm{c}}$. (We shall not use this fact elsewhere in this proof.) Hence it suffices to prove that if $\lambda < \lambda'$, then w.h.p. the k-core of $G(n, \lambda'/n)$ contains at least $\beta^+(\lambda)n$ vertices. Let us fix $\lambda < \lambda'$. We may assume that $\lambda \geqslant \lambda_{\mathrm{c}}$, as otherwise there is nothing to prove. Fix any $\lambda < \lambda_2 < \lambda'$ with $\beta(\cdot)$ continuous at λ_2. (An increasing function has at most countably many discontinuities.) Letting $\lambda_1 \nearrow \lambda_2$, we have

https://doi.org/10.1017/S0963548307008589 Published online by Cambridge University Press

$\beta(\lambda_1) \nearrow \beta(\lambda_2)$. Hence,

$$\Psi_{\geqslant k-1}\big(\lambda'\beta(\lambda_1)\big) \nearrow \Psi_{\geqslant k-1}\big(\lambda'\beta(\lambda_2)\big) > \Psi_{\geqslant k-1}\big(\lambda_2\beta(\lambda_2)\big) = \beta(\lambda_2).$$

Thus we may choose $\lambda < \lambda_1 < \lambda_2$ so that

$$\Psi_{\geqslant k-1}\big(\lambda'\beta(\lambda_1)\big) > \beta(\lambda_2). \tag{2.2}$$

We may also assume that $\beta(\lambda)$ is continuous at λ_1.

By Lemma 2.3 there is a constant d such that (2.1) holds; fix such a d. Once $\beta(\lambda)$ is positive, it is strictly increasing, so $\beta(\lambda_2) > \beta(\lambda_1)$. Hence, by Lemma 2.2, there is a d_1 such that $\mathbb{P}_{\lambda_2}(\mathscr{R}_{d_1} \circ \mathscr{B}) > \beta(\lambda_1)$.

Set

$$\eta = \min\{\mathbb{P}_{\lambda_2}(\mathscr{R}_{d_1} \circ \mathscr{B}) - \beta(\lambda_1), k^{-3d}\} > 0.$$

As $\mathscr{R}_{d_1} \circ \mathscr{B}$ is measurable, there is an integer L and an event $\mathscr{L}_1$ depending only on the first L generations of the branching process such that $\mathbb{P}_{\lambda_2}\big((\mathscr{R}_{d_1} \circ \mathscr{B}) \triangle \mathscr{L}_1\big) \leqslant \eta^2/2$, where $\triangle$ denotes symmetric difference. We may and shall assume that $L > d_1$. Writing $1_{\mathscr{E}}$ for the indicator function of an event $\mathscr{E}$, and $X[L]$ for the first L generations of the branching process, we have

$$\eta^2/2 \geqslant \mathbb{P}_{\lambda_2}\big((\mathscr{R}_{d_1} \circ \mathscr{B})^{\mathrm{c}} \cap \mathscr{L}_1\big) = \mathbb{E}_{\lambda_2}\big(1_{\mathscr{L}_1}\mathbb{P}_{\lambda_2}\big((\mathscr{R}_{d_1} \circ \mathscr{B})^{\mathrm{c}} \mid X[L]\big)\big), \tag{2.3}$$

where $\mathbb{E}_{\lambda_2}$ is the expectation corresponding to $\mathbb{P}_{\lambda_2}$. Set

$$\mathscr{L} = \mathscr{L}_1 \cap \Big\{\mathbb{P}_{\lambda_2}\big(\mathscr{R}_{d_1} \circ \mathscr{B} \mid X[L]\big) \geqslant 1 - \eta\Big\},$$

noting that the event $\mathscr{L}$ depends only on the first L generations of X. Since the expectation appearing in (2.3) is at least $\eta\mathbb{P}_{\lambda_2}\big(\mathscr{L}_1 \setminus \mathscr{L}\big)$, we have $\mathbb{P}_{\lambda_2}\big(\mathscr{L}_1 \setminus \mathscr{L}\big) \leqslant \eta/2$, so

$$\mathbb{P}_{\lambda_2}(\mathscr{L}) \geqslant \mathbb{P}_{\lambda_2}(\mathscr{L}_1) - \eta/2 \geqslant \mathbb{P}_{\lambda_2}(\mathscr{R}_{d_1} \circ \mathscr{B}) - \eta^2/2 - \eta/2 \geqslant \beta(\lambda_1). \tag{2.4}$$

Also, recalling that $L > d_1$, so $\mathscr{R}_L \circ \mathscr{B} \supset \mathscr{R}_{d_1} \circ \mathscr{B}$, whenever $\mathscr{L}$ holds we have

$$\mathbb{P}_{\lambda_2}\big(\mathscr{R}_L \circ \mathscr{B} \mid X[L]\big) \geqslant \mathbb{P}_{\lambda_2}\big(\mathscr{R}_{d_1} \circ \mathscr{B} \mid X[L]\big) \geqslant 1 - \eta \geqslant 1 - k^{-3d}. \tag{2.5}$$

Let $\mathscr{A}_0 = \mathscr{L}$, and for $t \geqslant 1$ set $\mathscr{A}_t = \mathscr{R}_d \circ \mathscr{A}_{t-1}$. Thus, $\mathscr{A}_t$ is a 'recursively robust' version of the event $\mathscr{B}_{dt} \circ \mathscr{L}$. From the independence of the descendants of different particles in generation d of X_{λ_2} we have

$$\mathbb{P}_{\lambda_2}(\mathscr{A}_t) = r\big(\lambda_2, d, \mathbb{P}_{\lambda_2}(\mathscr{A}_{t-1})\big).$$

Thus, by (2.4), (2.1), and induction on t, we have

$$\mathbb{P}_{\lambda_2}(\mathscr{A}_t) \geqslant \beta(\lambda_1) \tag{2.6}$$

for every t. Recalling that $\lambda_2 < \lambda'$, let us note for later that

$$\mathbb{P}_{\lambda'}(\mathscr{B}_1 \circ \mathscr{A}_t) = \Psi_{\geqslant k-1}\big(\lambda'\mathbb{P}_{\lambda'}(\mathscr{A}_t)\big) \geqslant \Psi_{\geqslant k-1}\big(\lambda'\mathbb{P}_{\lambda_2}(\mathscr{A}_t)\big) \geqslant \Psi_{\geqslant k-1}\big(\lambda'\beta(\lambda_1)\big) > \beta(\lambda_2), \tag{2.7}$$

where the final inequality is from (2.2).

If $\mathscr{E}$ depends on the first d generations of X, let $\mathbb{P}^p(\mathscr{E} \circ M \mid X[d])$ denote the conditional probability given the first d generations that $\mathscr{E} \circ M$ holds if we mark particles in generation d independently with probability p. Note that this conditional probability does not depend on λ.

https://doi.org/10.1017/S0963548307008589 Published online by Cambridge University Press

From (2.5), whenever $\mathscr{A}_0 = \mathscr{L}$ holds we have

$$\mathbb{P}^{\beta(\lambda_2)}(\mathscr{R}_L \circ M \mid X[L]) = \mathbb{P}_{\lambda_2}(\mathscr{R}_L \circ \mathscr{B} \mid X[L]) \geqslant 1 - k^{-3d}. \tag{2.8}$$

Let $\mathscr{E}_t = \mathscr{B}_{dt} \circ \mathscr{R}_L$, so $\mathscr{E}_0 = \mathscr{R}_L$ and $\mathscr{E}_t = \mathscr{B}_d \circ \mathscr{E}_{t-1}$ for $t > 0$. We claim that for any $t \geqslant 0$, whenever $\mathscr{A}_t$ holds, then

$$\mathbb{P}^{\beta(\lambda_2)}(\mathscr{E}_t \circ M \mid X[dt + L]) \geqslant 1 - k^{-(2^t+2)d}. \tag{2.9}$$

For $t = 0$, this is just (2.8). We prove the claim by induction on t; let us assume that $t > 0$, and that (2.9) holds with t replaced by $t - 1$. We condition on the first $dt + L$ generations of X, and assume that $\mathscr{A}_t$ holds. Since $\mathscr{A}_t = \mathscr{R}_d \circ \mathscr{A}_{t-1}$, there is a set Y of particles in generation d such that $\mathscr{A}_{t-1}$ holds for each $y \in Y$, and $\mathscr{R}_d \circ M$ holds if we mark only the particles in Y. As any tree witnessing $\mathscr{R}_d$ contains a subtree witnessing $\mathscr{R}_d$ in which each particle has at most k children, we may assume that $|Y| \leqslant k^d$. Now let us mark each particle in generation $dt + L$ independently with probability $\beta(\lambda_2)$, and let Y' be the set of particles y in Y such that $\mathscr{E}_{t-1} \circ M$ holds for y. Each $y \in Y$ is included independently in Y', and, by the induction hypothesis, the probability that a particular $y \in Y$ is included is at least $1 - k^{-(2^{t-1}+2)d}$. Thus

$$\mathbb{P}(|Y \setminus Y'| \geqslant 2) \leqslant \binom{|Y|}{2} \left(k^{-(2^{t-1}+2)d}\right)^2 \leqslant k^{2d} k^{-(2^t+4)d} = k^{-(2^t+2)d}.$$

From the definition of $\mathscr{R}_d$, whenever $|Y \setminus Y'| \leqslant 1$ then $\mathscr{B}_d$ holds if we keep in generation d only the particles in Y'. But then $\mathscr{B}_d \circ \mathscr{E}_{t-1} \circ M = \mathscr{E}_t \circ M$ holds, proving (2.9). This proves the claim by induction.

Given an event $\mathscr{E}$, let $\mathscr{B}^+_{\leqslant 2L+1} \circ \mathscr{E}$ be the event that there is a set S of *targets* in generations between 1 and $2L + 1$ such that $\mathscr{E}$ holds for each target, and in the tree consisting of the targets and their ancestors every non-leaf (*i.e.*, every non-target) has degree at least k; in the notation of Lemma 2.4, we require the targets and their ancestors to form a tree with property $\mathscr{M}_{\leqslant 2L+1}$. Note that we need not (and in general must not) take all particles in generations between 1 and $2L + 1$ that have property $\mathscr{E}$ to be targets. Thus, if $\mathscr{E}$ is increasing, so is $\mathscr{B}^+_{\leqslant 2L+1} \circ \mathscr{E}$.

Let $T = T(n)$ be any function with $T = o(\log n)$ and $T / \log\log n \to \infty$. Note that, until now, the branching process events we have defined do not depend on n. Set $\mathscr{A} = \mathscr{B}^+_{\leqslant 2L+1} \circ \mathscr{A}_T$. Since $\mathscr{B}^+_1 \circ \mathscr{A}_T \subset \mathscr{A}$, from (2.6) we have

$$\mathbb{P}_{\lambda_2}(\mathscr{A}) \geqslant \Psi_{\geqslant k}\big(\lambda_2 \mathbb{P}_{\lambda_2}(\mathscr{A}_T)\big) \geqslant \Psi_{\geqslant k}\big(\lambda_2 \beta(\lambda_1)\big) > \Psi_{\geqslant k}\big(\lambda_1 \beta(\lambda_1)\big) = \beta^+(\lambda_1).$$

Let $s = 2L + 1 + dT + L = o(\log n)$, so $\mathscr{A}$ depends on the first s generations of the branching process. By standard properties of branching processes, for any constant λ we have

$$\mathbb{P}_\lambda\big(|X[2s]| \geqslant n^{1/10}\big) = o(n^{-1}), \tag{2.10}$$

where $|X[2s]|$ is the total size of the first $2s$ generations of the branching process. Indeed, a simple calculation shows that for any m we have $\mathbb{E}_\lambda(|X[t]|^m) = O(\lambda^{mt})$ as $t \to \infty$. Thus $\mathbb{E}_\lambda\big(|X[2s]|^{20}\big) = O\big(\lambda^{40s}\big) = o(n)$, and (2.10) follows by Markov's inequality.

With this branching-process preparation behind us, we are now ready to turn to the graph $G(n, \lambda'/n)$.

Let $\tilde{G} = \tilde{G}(n, 4s, \lambda')$ be the graph obtained from $G(n, \lambda'/n)$ by conditioning on the absence of any cycles of length at most $4s$. Let $c_k(G)$ denote the number of vertices in the k-core of a graph

https://doi.org/10.1017/S0963548307008589 Published online by Cambridge University Press

G. As $\{c_k(G) \geqslant x\}$ is an increasing event, while the absence of short cycles is a decreasing event, we have

$$\mathbb{P}(c_k(\tilde{G}) \geqslant x) \leqslant \mathbb{P}\big(c_k(G(n, \lambda'/n)) \geqslant x\big)$$

for any x. Hence it suffices to prove a lower bound for $c_k(\tilde{G})$ that holds w.h.p. instead of the corresponding bound for $c_k(G(n, \lambda'/n))$. We shall work entirely with $\tilde{G}$.

Let v be a random vertex of $\tilde{G}$, and explore its successive neighbourhoods $\Gamma_t(v)$, $t \leqslant 2s$, in the usual way, where $\Gamma_t(v)$ is the set of vertices of $\tilde{G}$ at graph distance t from v. We start with v 'active' and all other vertices 'untested'. At each step we pick an 'active' vertex w closest to v, and test edges from w to untested vertices one by one, marking any neighbour of w found in this way as 'active'. After testing all possible edges from w we mark w as 'tested'. By Lemma 2.6, as long as we have reached at most $n^{1/3}$ vertices, conditional on everything so far each test succeeds with probability $(1 + O(n^{-1/4}))\lambda'/n$; we never attempt to test an edge that might complete a cycle. As the number of untested vertices is $n - O(n^{1/3})$, we may couple the number of new neighbours of w found with a Poisson distribution with mean λ' so that the two numbers agree with probability $1 - O(n^{-1/4})$.

Let $G_v[t]$ be the subgraph of $\tilde{G}$ formed by the vertices within distance t of v, noting that for $t \leqslant 2s$ this graph is by definition a tree. Since, w.h.p., $X_{\lambda'}[2s]$ contains at most $n^{1/10}$ particles, w.h.p. $G_v[2s]$ contains at most $n^{1/10}$ vertices. It follows from the coupling above that we may couple $G_v[2s]$ and $X_{\lambda'}[2s]$ so as to agree in the natural sense w.h.p. Let us say that v has property $\mathscr{A}$ if $G_v[2s]$ has property $\mathscr{A}$ when viewed as a branching process, an event that depends only on $G_v[s]$. We have shown that v has property $\mathscr{A}$ with probability $\mathbb{P}_{\lambda'}(\mathscr{A}) + o(1)$.

To establish concentration of the number of vertices v with property $\mathscr{A}$, we use a simple trick also used in [5]. Let v and w be independently chosen random vertices of $\tilde{G}$. Since the probability that v and w are within distance $2s$ is $o(1)$, we may couple $G_v[s]$ and $G_w[s]$ with independent copies of $X_{\lambda'}[s]$ so as to agree w.h.p. Hence, the probability that both v and w have property $\mathscr{A}$ is $\mathbb{P}_{\lambda'}(\mathscr{A})^2 + o(1)$. Writing N for the number of vertices with property $\mathscr{A}$, we thus have

$$\mathbb{E}(N) = \mathbb{P}_{\lambda'}(\mathscr{A})n + o(n), \quad \mathbb{E}(N^2) = \mathbb{P}_{\lambda'}(\mathscr{A})^2 n^2 + o(n^2).$$

Hence (by Chebyshev's inequality), N/n converges in probability to

$$\mathbb{P}_{\lambda'}(\mathscr{A}) \geqslant \mathbb{P}_{\lambda_2}(\mathscr{A}) \geqslant \beta^+(\lambda_1) > \beta^+(\lambda).$$

Thus,

$$\mathbb{P}(N \leqslant \beta^+(\lambda)n) = o(1).$$

To complete the proof of the lower bound on the size of the k-core, it thus suffices to show that $c_k(\tilde{G}) \geqslant N$ w.h.p. We do this by showing that w.h.p. *every* vertex with property $\mathscr{A}$ is in the k-core.

Let v be a random vertex of $\tilde{G}$. Let us test whether v has property $\mathscr{A}$ by exploring its neighbourhoods as follows. Throughout we regard the graph $G_v[2s]$ as a branching process with v as initial particle; this is valid since the graph contains no cycles. Thus we shall speak of the children and descendants of a vertex. If w is a vertex in generation t (*i.e.*, at graph distance t from v), we shall write $D(w)$ for the set of descendants of w in generation $t + (dT + L)$. Let S_1 and S_2 be initially empty sets of vertices. At each stage, every vertex in S_1 will have property $\mathscr{A}_T$ in the

https://doi.org/10.1017/S0963548307008589 Published online by Cambridge University Press

tree/branching process rooted at v. In other words, w and its descendants will form a tree with property $\mathscr{A}_T$. Recall that $\mathscr{A} = \mathscr{B}^+_{\leqslant 2L+1} \circ \mathscr{A}_T$.

First, let us reveal the descendants of v unto generation $dT + L + 1$. For each child w of v, examine its descendants to relative generation $dT + L$, to see whether w has property $\mathscr{A}_T$. If so, put w into S_1 and the vertices in $D(w)$ into S_2. If S_1 is large enough to guarantee that v has $\mathscr{A}$, stop. (At this stage, this happens if and only if $|S_1| \geqslant k$.) Otherwise, examine each w in generation 2 that is not a descendant of a vertex in S_1 in turn. For each w, reveal its descendants up to $dT + L$ generations later, and test whether w has $\mathscr{A}_T$. If so, include w into S_1 and the vertices of $D(w)$ into S_2. Continue in this way until either (a) the set S_1 shows that $\mathscr{A}$ holds (*i.e.*, S_1 contains a set of targets witnessing $\mathscr{B}^+_{\leqslant 2L+1} \circ \mathscr{A}_T$), or (b) we have tested all vertices in generations up to $2L + 1$ not descended from vertices in S_1 for property $\mathscr{A}_T$ without finding such a set. In case (b), $\mathscr{A} = \mathscr{B}^+_{\leqslant 2L+1} \circ \mathscr{A}_T$ does not hold. Hence, v has $\mathscr{A}$ if and only if (a) holds.

Let us condition on v having property $\mathscr{A}$. The key points are that this event is guaranteed by the vertices in S_2 and their ancestors, and that we have not examined the children of any vertex in S_2. Let us suppose that at most $n^{1/10}$ vertices have been examined, an event of probability $1 - o(n^{-1})$ by (2.10). We now examine the vertices $w \in S_2$ one by one. For each w we explore its descendants for the next $1 + dT + L$ generations to test whether, in the branching process rooted at v, the particle w has the property $\mathscr{B}_1 \circ \mathscr{A}_T$. We abandon the exploration associated to a given w if we reach more than $n^{1/10}$ vertices in this exploration. If the test is successful, we mark w. Note that, from (2.10), the probability that any exploration is abandoned is $o(n^{-1})$. Provided no previous exploration has been abandoned, the argument above shows that we may couple the descendants of w to agree with a branching process $X_{\lambda'}$ with probability $1 - o(1)$. Hence, if n is large enough, the probability that we mark w is at least

$$\mathbb{P}_{\lambda'}(\mathscr{B}_1 \circ \mathscr{A}_T) - o(1) \geqslant \beta(\lambda_2),$$

using (2.7). In summary, ignoring an error probability of $1 - o(n^{-1})$, we can view each $w \in S_2$ as marked independently with probability (at least) $\beta(\lambda_2)$.

Let $\mathscr{T}_v$ be the tree rooted at v with leaves the marked vertices in S_2. Since $T/\log\log n \to \infty$, we have $k^{-(2^T+2)d} \leqslant n^{-10}$ if n is large enough. Hence, from (2.9), with probability $1 - o(n^{-1})$ every vertex of S_1 has property $\mathscr{E}_T$ guaranteed by its descendants in $\mathscr{T}_v$, so with probability $1 - o(n^{-1})$ the tree $\mathscr{T}_v$ has the property $\mathscr{A}' = \mathscr{B}^+_{\leqslant 2L+1} \circ \mathscr{E}_T$ when viewed as a branching process. Let us suppose that $\mathscr{T}_v$ has property $\mathscr{A}'$, and let $\mathscr{T}'_v$ be a minimal subtree with this property. Recalling that $\mathscr{E}_T = \mathscr{B}_{Td} \circ \mathscr{R}_L$, note that every non-leaf of $\mathscr{T}'_v$ has degree at least k in $\mathscr{T}'_v$.

We claim that each leaf w of $\mathscr{T}'_v$ has property $\mathscr{A}$ in the graph $\tilde{G}$. To see this, let x be the ancestor of w in $\mathscr{T}'_v$ that is L generations above w, so the subtree $\mathscr{T}'$ of $\mathscr{T}'_v$ rooted at x is a tree of height L that is minimal with respect to having property $\mathscr{R}_L$. All leaves of $\mathscr{T}'$ are leaves of $\mathscr{T}'_v$ and hence marked, so they have the property $\mathscr{B}_1 \circ \mathscr{A}_T$ in the tree $G_v[2s]$ rooted at v. Hence we may take $k - 1$ children of each leaf and add them to $\mathscr{T}'$ to form a tree $\mathscr{T}$ with property $\mathscr{R}_L \circ \mathscr{B}_1$ in which every leaf has property $\mathscr{A}_T$. By Lemma 2.4, taking w as the root, this tree $\mathscr{T}$ has a subtree $\mathscr{W}$ with property $\mathscr{M}_{\leqslant 2L+1}$ all of whose leaves are leaves of $\mathscr{T}$ and hence have property $\mathscr{A}_T$. Thus w has property $\mathscr{B}^+_{\leqslant 2L+1} \circ \mathscr{A}_T = \mathscr{A}$ in the graph, as claimed.

We have shown that with probability $1 - no(n^{-1}) = 1 - o(1)$ *every* vertex v with property $\mathscr{A}$ is the root of a tree $\mathscr{T}'_v$ in which each non-leaf has degree at least k and each leaf has property $\mathscr{A}$ (in the graph $\tilde{G}$). When this happens, the union of the trees $\mathscr{T}'_v$ is a subgraph of $\tilde{G}$ with minimum

https://doi.org/10.1017/S0963548307008589 Published online by Cambridge University Press

degree at least k containing all N vertices with property $\mathscr{A}$, so $c_k(\tilde{G}) \geqslant N$. This completes our proof of the lower bound.

The proof of the upper bound on $c_k(G(n,\lambda/n))$ is much simpler: the events $\mathscr{B}_d^+$ decrease to $\mathscr{B}^+$ as $d \to \infty$. Hence, given any $\varepsilon > 0$, there is a d with $\mathbb{P}_\lambda(\mathscr{B}_d^+) \leqslant \beta^+(\lambda) + \varepsilon$. Exploring the first d neighbourhoods of a random vertex v of $G(n,\lambda/n)$ as above, the probability that we encounter a cycle is $o(1)$. Let us say that v has property $\mathscr{B}_d^+$ if its d-neighbourhood is a tree with property $\mathscr{B}_d^+$. Considering the local exploration described above for $\tilde{G}$, in the simpler context of $G(n,\lambda/n)$ it is easy to check that $\mathbb{P}_\lambda(\mathscr{B}_d^+)n + o_p(n)$ vertices v have property $\mathscr{B}_d^+$, while $o_p(n)$ vertices have a cycle in their d-neighbourhood. But any vertex in the k-core must have one of these properties, so

$$c_k(G(n,\lambda/n)) \leqslant \mathbb{P}_\lambda(\mathscr{B}_d^+)n + o_p(n) \leqslant \beta^+(\lambda)n + \varepsilon n + o_p(n).$$

As $\varepsilon > 0$ was arbitrary, this completes the proof. □

Theorem 2.1 implies that the natural coupling between the neighbourhoods of a vertex v in $G(n,\lambda/n)$ and the branching process X_λ can be adapted to the k-core. With $k \geqslant 2$ fixed as usual, let $C_k(G)$ denote the k-core of G, and let X'_λ denote the (possibly empty) branching process consisting of all particles in X_λ that are in a k-regular tree in X_λ containing the initial particle.

Corollary 2.7. *Let $k \geqslant 2$, $\lambda > 0$ and L be fixed. If $\beta(\lambda)$ is continuous at λ, then the first L neighbourhoods in $C_k(G(n,\lambda/n))$ of a random vertex of $G(n,\lambda/n)$ may be coupled with the first L generations of X'_λ so as to agree w.h.p.*

Proof. We use only the result of Theorem 2.1, not anything from the proof. We may assume that $\beta(\lambda) > 0$, as otherwise $C_k = C_k(G(n,\lambda/n))$ contains $o_p(n)$ vertices, while X'_λ is empty with probability 1. Since $\mathscr{B}_d^+ \searrow \mathscr{B}^+$, if $d(n) \to \infty$ then $\mathbb{P}_\lambda(\mathscr{B}_{d(n)}^+) = \mathbb{P}_\lambda(\mathscr{B}^+) + o(1)$. Note that $\mathscr{B}_{d(n)}^+ \supset \mathscr{B}^+$.

If $d(n) \to \infty$ sufficiently slowly then, by standard arguments, we may couple the branching process X_λ with the neighbourhood process $Y = Y(v, G(n,\lambda/n))$ so that w.h.p. they agree for the first $d(n)$ generations. (This follows from the d constant case of the coupling.) If v is in the k-core C_k, then either $\mathscr{B}_{d(n)}^+$ holds for Y, or the first $d(n)$ neighbourhoods of v contain a cycle. Now

$$\mathbb{P}(Y \in \mathscr{B}_{d(n)}^+) = \mathbb{P}(X_\lambda \in \mathscr{B}_{d(n)}^+) + o(1) = \mathbb{P}(X_\lambda \in \mathscr{B}^+) + o(1) = \mathbb{P}(v \in C_k) + o(1),$$

where the first step is from the coupling and the last from Theorem 2.1. Provided $d(n) = o(\log n)$, which we may assume, $o_p(n)$ vertices v have a cycle in their $d(n)$-neighbourhoods. Thus, w.h.p. either all three of $\{v \in C_k\}$, $\{Y \in \mathscr{B}_{d(n)}^+\}$ and $\{X_\lambda \in \mathscr{B}^+\}$ hold, or none.

The argument above holds with $d(n) - L$ in place of $d(n)$. Let us call a vertex w *exceptional* if its neighbourhoods to distance $d(n) - L$ contain a cycle, or these neighbourhoods have the property corresponding to $\mathscr{B}_{d(n)-L}^+$ but w is not in the k-core. All other vertices are *normal*. Then the probability that a given vertex w is exceptional is $o(1)$, so there are $o_p(n)$ exceptional w. Since, for any fixed t, the t-neighbourhood of any set of $o_p(n)$ vertices in $G(n,\lambda/n)$ has size $o_p(n)$, returning to the random vertex v, with probability $1 - o(1)$ all vertices within distance L of v are normal.

https://doi.org/10.1017/S0963548307008589 Published online by Cambridge University Press

Suppose that the coupling above succeeds for the neighbourhoods of a random vertex v to distance $d(n)$, and let S be the set of vertices near (*i.e.*, within distance L of) v that are in the k-core. As w.h.p. every vertex near v is normal, w.h.p. S is the set of vertices corresponding to particles y in the first L generations of X_λ with the property $\mathscr{B}^+_{d(n)-L}$ in the tree X_λ with y as root. As L is fixed and $d(n) \to \infty$, this latter set agrees w.h.p. with the set Z of particles z in the first L generations of X_λ having property $\mathscr{B}^+$ in X_λ with z as root, *i.e.*, contained in some infinite k-regular tree. But the set of particles in Z reachable from the root x_0 by a path in Z is just the first L generations of X'_λ, so w.h.p. the first L generations of X'_λ correspond to the first L neighbourhoods of v in the k-core, as required. □

The argument above shows that, once we have the asymptotic number of vertices in the k-core C_k, we can count up to $o_p(n)$ all 'local' structures in C_k, including, for example, the number of vertices of a given degree. Of course, we can also estimate the sum of the degrees, say, and hence the number of edges. More generally, we can estimate the sum of any function of the L-neighbourhood with a well-behaved tail, where the contribution from the $o(1)$-probability case when the coupling breaks down is small.

Let $f(v, G)$ be an isomorphism-invariant function of graphs G rooted at a vertex v. As in [5], we call f an *L-neighbourhood function* if it depends only on the subgraph of G induced by the vertices within distance L of v. (Thus the degree of v in G is a 1-neighbourhood function.) We interpret $f(X_\lambda)$ in the natural way, viewing the branching process as a tree with root x_0.

Theorem 2.8. *Let $k \geqslant 2$, $\lambda \neq \lambda_c(k)$, $L \geqslant 1$ and an L-neighbourhood function f be given. If $f(v, G)$ is bounded by a polynomial of the number of vertices in the L-neighbourhood of v, then*

$$S_n := \frac{1}{n} \sum_{v \in C_k} f(v, C_k) \xrightarrow{\mathrm{p}} \mathbb{E}\, f(X'_\lambda),$$

where C_k denotes the k-core of $G(n, \lambda/n)$.

The simple proof follows exactly that of the corresponding result for the 2-core of a more general graph, Lemma 11.11 of [5]. As the result is very unsurprising in the light of Corollary 2.7, we omit the proof. The condition that f be polynomially bounded can be replaced by a fourth moment condition, as in [5].

3. Inhomogeneous random graphs

Many random graph models have been considered in which edges are independent but different possible edges have different probabilities, including numerous 'sparse' models generalizing $G(n, \lambda/n)$, where the expected number of edges is linear in the number of vertices; see [5] and the references therein. A very general model of this type was introduced by Bollobás, Janson and Riordan in [5]; let us recall the definitions.

A *ground space* is a pair $(\mathscr{S}, \mu)$, where $\mathscr{S}$ is a separable metric space and μ is a Borel probability measure on $\mathscr{S}$. Mostly, we shall consider the cases $\mathscr{S}$ finite, with μ any (strictly positive) measure on $\mathscr{S}$, and $\mathscr{S} = (0, 1]$ or $\mathscr{S} = [0, 1]$, with μ Lebesgue measure.

https://doi.org/10.1017/S0963548307008589 Published online by Cambridge University Press

A set $A \subset \mathscr{S}$ is a *μ-continuity set* if A is (Borel) measurable and $\mu(\partial A) = 0$, where ∂A is the boundary of A.

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

$$v_n(A) := \#\{i : x_i^{(n)} \in A\}/n \xrightarrow{\mathrm{p}} \mu(A) \tag{3.1}$$

as $n \to \infty$, for every μ-continuity set A.

The sequence $x_i^{(n)}$ will give the *type* of each vertex i in a graph on $[n]$ still to be defined. The convergence condition (3.1) says that distribution of the types of the vertices is essentially μ. More precisely, the empirical distribution v_n of the types converges in probability to μ. In the *finite-type* case, where $\mathscr{S} = [r] = \{1, 2, \dots, r\}$ and $\mu\{i\} > 0$ for $i = 1, 2, \dots, r$, the condition says that the number n_i of vertices of each type i satisfies $n_i/n \xrightarrow{\mathrm{p}} \mu\{i\}$. No assumption is made about the dependence between the types of different vertices.

A *kernel* κ on a ground space $(\mathscr{S}, \mu)$ is a symmetric non-negative (Borel) measurable function on $\mathscr{S} \times \mathscr{S}$. A kernel on a vertex space $(\mathscr{S}, \mu, (\mathbf{x}_n)_{n\geqslant 1})$ is simply a kernel on $(\mathscr{S}, \mu)$.

Finally, let κ be a kernel on the vertex space $\mathscr{V}$. Given the (random) sequence $(x_1^{(n)}, \dots, x_n^{(n)})$, let $G^{\mathscr{V}}(n, \kappa)$ be the random graph on $[n]$ in which edges are present independently, and the probability that the edge ij is present is

$$p_{ij} := \min\{\kappa(x_i^{(n)}, x_j^{(n)})/n, 1\}. \tag{3.2}$$

In other words, we assign vertices types from $\mathscr{S}$ with the types asymptotically distributed according to μ, and then join vertices of types x and y with probability $\kappa(x, y)/n$. If κ takes the constant value λ, then $G^{\mathscr{V}}(n, \kappa)$ is exactly $G(n, \lambda/n)$.

The model just described includes many models of inhomogeneous graphs previously defined. As the model is very general, it may help to bear in mind a few special cases. The most fundamental is the finite-type case described above. This is an extremely natural generalization of $G(n, \lambda/n)$, and has been considered earlier by several authors (*e.g.*, Söderberg [23]), perhaps with slightly different assumptions on the distribution of the types. (For example, the types of the vertices being independent, or n_i equal to $\lfloor \mu\{i\}n \rfloor$ or $\lceil \mu\{i\}n \rceil$.) Several other interesting special cases have $\mathscr{S} = (0, 1]$, μ Lebesgue measure, and $x_i^{(n)} = i/n$, so the probability that the edge ij is present is $p_{ij} = \kappa(i/n, j/n)/n$. Taking $\kappa(x, y) = 1/\max\{x, y\}$ gives $p_{ij} = 1/\max\{i, j\}$; minor variants of the corresponding random graph were introduced independently by Dubins in 1984 (see [16]) and by Callaway, Hopcroft, Kleinberg, Newman and Strogatz [7] in 2001; see [5].

In order to prove results about $G^{\mathscr{V}}(n, \kappa)$, some additional assumptions are needed to avoid pathologies. Following [5], we assume throughout that κ is continuous almost everywhere on $\mathscr{S} \times \mathscr{S}$, that

$$\int_{\mathscr{S}\times\mathscr{S}} \kappa(x, y)\, d\mu(x)\, d\mu(y) < \infty, \tag{3.3}$$

and that

$$\frac{1}{n}\,\mathbb{E}\, e\big(G^{\mathscr{V}}(n, \kappa)\big) \to \frac{1}{2}\iint_{\mathscr{S}\times\mathscr{S}} \kappa(x, y)\, d\mu(x)\, d\mu(y) \tag{3.4}$$

as $n \to \infty$. The last condition says that the number of edges of $G^{\mathscr{V}}(n, \kappa)$ is 'what it should be', at least in expectation. Without this condition, it is impossible to relate the behaviour of $G^{\mathscr{V}}(n, \kappa)$

https://doi.org/10.1017/S0963548307008589 Published online by Cambridge University Press

to that of κ (or rather, to that of the branching process X_κ defined below): changing κ on a set of measure zero should not affect the model, but $G^\mathscr{V}(n,\kappa)$ may depend *only* on the values of κ on such as set, for example, when $x_i^{(n)} = i/n$. Surprisingly, the very natural and rather weak condition (3.4) is enough to enable many results about $G^\mathscr{V}(n,\kappa)$ to be proved; see [5] for a discussion of this.

We shall need one further definition: a kernel κ on a ground space $(\mathscr{S},\mu)$ is *reducible* if

$$\exists A \subset \mathscr{S} \text{ with } 0 < \mu(A) < 1 \text{ such that } \kappa = 0 \text{ a.e. on } A \times (\mathscr{S} \setminus A),$$

and *irreducible* otherwise. Roughly speaking, reducible kernels correspond to disconnected graphs; much of the time, nothing is lost by considering only irreducible kernels.

The actual definitions in [5] are slightly more general in two ways; firstly, the kernel κ is allowed to depend somewhat on n. We shall not bother with this additional generality here. Secondly, the number of vertices of $G^\mathscr{V}(n,\kappa)$ need not be exactly n, but may be $n + o_p(n)$. This latter relaxation complicates only the notation, not the proofs.

The key to the analysis of $G^\mathscr{V}(n,\kappa)$ turns out to be the multi-type Galton–Watson branching process X_κ associated to $(\mathscr{S},\mu,\kappa)$. This starts with a single particle x_0 whose type is distributed according to μ. Each particle of type x has a set of children whose types are distributed as a Poisson process on $\mathscr{S}$ with intensity $\kappa(x,y)\,d\mu(y)$. In other words, the number of children with types in a subset $A \subset \mathscr{S}$ has a Poisson distribution with mean $\int_A \kappa(x,y)\,d\mu(y)$, and these numbers are independent for disjoint sets A and for different particles; see, for example, Kallenberg [17]. Sometimes it will be convenient to start the process with a particle of a fixed type x instead of a random type. We write $X_\kappa(x)$ for this branching process.

If κ takes the constant value λ, then X_κ is just X_λ as defined in Section 2. In general, the relationship of X_κ to $G^\mathscr{V}(n,\kappa)$ is the same as that of X_λ to $G(n,\lambda/n)$. For example, under suitable regularity conditions, which certainly hold in the finite-type case, the first few neighbourhoods of a random vertex v of $G^\mathscr{V}(n,\kappa)$ may be coupled with the branching process X_κ in the natural sense (so that the type of a vertex is the same as the type of the corresponding particle in the branching process); see [5]. One of the main results of [5] is that, under suitable mild assumptions, the size of the giant component in $G^\mathscr{V}(n,\kappa)$ is asymptotically n times the probability that the branching process X_κ never dies out. This generalizes the classical result giving the size of the giant component of $G(n,\lambda/n)$ in terms of X_λ.

We shall write $\mathbb{P}_\kappa$ for the probability measure associated to X_κ, and $\mathbb{P}_{\kappa,x}$ for that associated to $X_\kappa(x)$. Thus, if $\mathscr{A}$ is some property of rooted trees, we write $\mathbb{P}_\kappa(X_\kappa \in \mathscr{A})$, or simply $\mathbb{P}_\kappa(X \in \mathscr{A})$, for the probability that X_κ has this property when viewed as a tree.

Turning to the k-core, let us define $\beta(\kappa)$ and $\beta^+(\kappa)$ as before, but for X_κ. Thus $\beta(\kappa) = \mathbb{P}_\kappa(\mathscr{B})$, and $\beta^+(\kappa) = \mathbb{P}_\kappa(\mathscr{B}^+)$. We shall also write $\beta_x(\kappa)$ for $\mathbb{P}_{\kappa,x}(\mathscr{B})$, and $\beta_x^+(\kappa)$ for $\mathbb{P}_{\kappa,x}(\mathscr{B}^+)$. By analogy with the result for $G(n,\lambda/n)$, we expect the k-core of $G^\mathscr{V}(n,\kappa)$ to have size $\beta^+(\kappa)n + o_p(n)$, at least under suitable conditions.

3.1. The finite-type case

The strategy of starting with the finite-type case, and then using approximation and monotonicity arguments to attack the general case, is used throughout [5]. We use the same strategy here; we have written the proof of Theorem 2.1 so that it carries over almost immediately to the finite-type case. Recall that we write $c_k(G)$ for $|C_k(G)|$, where $C_k(G)$ is the k-core of a graph G.

https://doi.org/10.1017/S0963548307008589 Published online by Cambridge University Press

Theorem 3.1. *Let $k \geqslant 2$ be fixed. Let κ be a kernel on a vertex space $\mathscr{V} = (\mathscr{S}, \mu, (\mathbf{x}_n)_{n\geqslant 1})$ with $\mathscr{S} = [r]$, $r \geqslant 1$, and $\mu\{i\} > 0$ for each i. If the function $\lambda \mapsto \beta(\lambda\kappa)$ is continuous at $\lambda = 1$, then*

$$c_k(G^{\mathscr{V}}(n,\kappa)) = \beta^+(\kappa)n + o_p(n).$$

Proof. We may assume without loss of generality that κ is irreducible; indeed, in the finite-type case, if κ is reducible the graph $G^{\mathscr{V}}(n,\kappa)$ may be written as the disjoint union of two or more graphs given by instances of the same model with irreducible kernels. (As each n_i is random, these graphs have a random number of vertices, but this does not matter.)

As in the proof of Theorem 2.1, the upper bound is easy. Indeed, arguing as for $G(n,\lambda/n)$, it is easy to see that for fixed d one can couple the d-neighbourhood of a random vertex v of $G^{\mathscr{V}}(n,\kappa)$ with the first d generations of X_κ to agree with probability $1 - o(1)$. The upper bound on the k-core follows as in the uniform case.

For the lower bound, it again suffices to prove that if $\lambda < \lambda'$, then w.h.p. the k-core of $G^{\mathscr{V}}(n,\lambda'\kappa)$ contains at least $\beta^+(\lambda\kappa)n$ vertices. The proof is essentially the same as in the uniform case, *mutatis mutandis*; we indicate the changes briefly.

The proof of Lemma 2.2 extends unchanged. Since $\mu\{i\} > 0$ for each i, it then follows that

$$\mathbb{P}_{\lambda\kappa,i}(\mathscr{R}_d \circ \mathscr{B}) \nearrow \beta_i(\lambda\kappa)$$

as $d \to \infty$, for each $i \in \mathscr{S} = [r]$. For $\mathbf{p} = (p_1, \ldots, p_r)$, let $r_i(\lambda\kappa, d, \mathbf{p})$ be the probability that $\mathscr{R}_d \circ M$ holds if we mark the particles in generation d of $X_{\lambda\kappa}(i)$ independently, marking a particle of type j with probability p_j. In place of Lemma 2.3 we obtain

$$r_i(\lambda_2\kappa, d, \mathbf{p}) \geqslant \beta_i(\lambda_1\kappa)$$

for every i, where $p_j = \beta_j(\lambda_1\kappa)$; in the proof, we add $\mathrm{Po}\big((\lambda_2 - \lambda_1)\kappa(i,j)\big)$ 'extra' particles of some type j with $\kappa(i,j) > 0$. The irreducibility of κ guarantees that such a type j exists, and also that $\beta_j(\lambda_1\kappa) > 0$ for every j whenever $\beta(\lambda_1\kappa) > 0$.

Lemma 2.6 adapts immediately, as Harris's Lemma applies to random subsets obtained by selecting elements independently, even if the selection probabilities are different for different elements. Next, in place of (2.2) we have

$$\Psi_{\geqslant k-1}\left(\lambda' \sum_j \kappa(i,j)\beta_j(\lambda_1\kappa)\right) \geqslant \beta_i(\lambda_2\kappa)$$

for each i. Arguing as for (2.4), (2.5) but taking

$$\eta = \min\{k^{-3d}, \min_i\{\big(\mathbb{P}_{\lambda_2\kappa,i}(\mathscr{R}_{d_1} \circ \mathscr{B}) - \beta_i(\lambda_1)\big)\mu\{i\}\}\} > 0,$$

we find an event $\mathscr{L}$ depending on the first L generations with

$$\mathbb{P}_{\lambda_2\kappa}(\mathscr{R}_L \circ \mathscr{B} \mid X[L]) \geqslant 1 - k^{-3d}$$

whenever $\mathscr{L}$ holds, and $\mathbb{P}_{\lambda_2\kappa,i}(\mathscr{L}) \geqslant \beta_i(\lambda_1\kappa)$ for each i. It follows as before that $\mathbb{P}_{\lambda_2\kappa,i}(\mathscr{A}_t) \geqslant \beta_i(\lambda_1\kappa)$ for each i, and the rest of the proof is essentially unchanged. □

Remark. The continuity condition in Theorem 3.1 is non-trivial: unlike in the uniform case, $\lambda \mapsto \beta(\lambda\kappa)$ may have several discontinuities. One way this can happen is when κ is reducible, so

https://doi.org/10.1017/S0963548307008589 Published online by Cambridge University Press

the graph decomposes into two separate pieces, whose k-cores may emerge at different points. Perhaps surprisingly, this is not the only way. Let $k = 3$, and set $\mathscr{S} = \{1, 2\}$ with $\mu\{1\} = \mu\{2\} = 1/2$. Let $\kappa(1, 1) = 2000$, $\kappa(2, 2) = 2$ and $\kappa(1, 2) = \kappa(2, 1) = 1/100$, say. It is easy to check that a k-core first emerges near $\lambda = \lambda_c/1000$, where λ_c is the critical parameter for the emergence of a 3-core in $G(n, \lambda/n)$; at this point, the vertices of type 1 form a uniform random graph with large enough average degree to contain a k-core. When λ is close to λ_c, the probability $p_2 = \beta_2(\lambda\kappa)$ that a vertex of type 2 has $k - 1$ neighbours in the k-core is related to the largest solution to $p = \Psi_{\geqslant k-1}(\lambda p + \lambda p_1/200)$, where p_1 is the (unknown) probability that a vertex of type 1 has $k - 1$ neighbours in the k-core. Since $0 < p_1 < 1$, it is easy to check that the solution jumps near $\lambda = \lambda_c$, and in fact that p_2 jumps from around $\Psi_{\geqslant k}(\lambda_c/100)$ to around $\beta(\lambda_c)$.

As in the uniform case, Theorem 3.1 implies the equivalents of Corollary 2.7 and Theorem 2.8 for the finite-type case. The statements and proofs are direct translations of those in Section 2, so we omit them.

3.2. The general case

Theorem 3.1 extends easily to more general kernels under some mild assumptions. To state these, we need another definition: a kernel κ' is *regular finitary* if there is a partition of $\mathscr{S}$ into a finite number r of μ-continuity sets S_i, such that κ' is constant on each $S_i \times S_j$. Such kernels correspond to finite-type kernels in an obvious way.

When studying the k-core in $G(n, \lambda/n)$, we assumed that $\beta(\cdot)$ (or, equivalently, $\beta^+(\cdot)$) was continuous from below at λ, *i.e.*, that $\beta(\lambda_n) \nearrow \beta(\lambda)$ for every sequence $\lambda_n \nearrow \lambda$. Of course, it makes no difference if we consider only one sequence λ_n, as long as this increases strictly to λ. It is this latter, formally weaker, condition that we shall adapt to general kernels, with the restriction that our kernels tending up from below be of finite type.

When studying $G^{\mathscr{S}}(n, \kappa)$, we shall assume that the functional $\kappa \mapsto \beta^+(\kappa)$ is *continuous from below at* κ, in the weak sense that

$$\beta^+(c_m\kappa_m) \to \beta^+(\kappa) \tag{3.5}$$

for *some* sequence κ_m of regular finitary kernels with $\kappa_m(x, y) \leqslant \kappa(x, y)$ for all x and y, and some sequence c_m of real numbers with $c_m < 1$ for all m. (Equivalently, we require $\beta^+(\kappa_m) \to \beta^+(\kappa)$ for some finite-type kernels κ_m with $\sup_{x,y} \kappa_m(x, y)/\kappa(x, y) < 1$ for all m. When $|\mathscr{S}| = 1$, this is equivalent to continuity from below in the usual sense.) As we have seen, this is a non-trivial condition; however, we expect it to hold for almost all κ, in some imprecise sense. We return to this later. Also, it may well be the case that, if κ is irreducible, then this condition is equivalent to the condition that $\beta^+(\kappa_m) \to \beta^+(\kappa)$ for *all* sequences κ_m of (arbitrary) kernels increasing to κ (almost everywhere). Note that β^+ is always continuous from above: if κ_m is a sequence of kernels decreasing pointwise to κ, then one can couple the branching processes X_{κ_m} and X_κ so that $X_{\kappa_1} \supset X_{\kappa_2} \supset \cdots$, and $X_\kappa = \bigcap_m X_{\kappa_m}$. It follows that $\beta^+(\kappa_m) \searrow \beta^+(\kappa)$.

Our continuity assumption can be viewed as the assertion that $o(n)$ edges will not change the size of the k-core much. Due to the flexibility of the model, it would be unreasonable to expect a precise result without this condition; as shown in [5], the condition (3.4) used to exclude pathologies still permits the insertion of $o_p(n)$ edges into $G^{\mathscr{S}}(n, \kappa)$ in a more or less arbitrary way, while changing κ on a set of measure zero, which does not alter $\beta^+(\kappa)$.

https://doi.org/10.1017/S0963548307008589 Published online by Cambridge University Press

Our second assumption will state essentially that a small set of 'exceptional' vertices (or edges) cannot have many vertices in its L-neighbourhood, for any fixed L. Perhaps the most natural form of this assumption involves counting paths: let $P_\ell(G)$ denote the number of ℓ-edge paths in a graph G, let

$$\alpha_\ell(\kappa) := \frac{1}{2}\int_{\mathscr{S}^{\ell+1}} \kappa(x_0,x_1)\kappa(x_1,x_2)\cdots\kappa(x_{\ell-1},x_\ell)\,d\mu(x_0)\cdots d\mu(x_\ell),$$

and suppose that

$$\frac{1}{n}\,\mathbb{E}\left(P_\ell\left(G^\mathscr{V}(n,\kappa)\right)\right) \to \alpha_\ell(\kappa) < \infty \tag{3.6}$$

as $n\to\infty$, for each $\ell \geqslant 1$. The convergence condition says essentially that the expected number of paths in $G^\mathscr{V}(n,\kappa)$ is 'what it should be'. As shown in [5], it holds whenever κ is bounded, or whenever the types of the vertices are independent with distribution μ. Theorem 17.1 in [5] shows that convergence in expectation in (3.6) implies convergence in probability; we shall not directly use this fact.

By Lemma 7.3(ii) of [5], there is a sequence of regular finitary kernels κ_m with $\kappa_m \leqslant \kappa$ pointwise, such that $\kappa_m(x,y) \nearrow \kappa(x,y)$ for almost all $(x,y)\in\mathscr{S}\times\mathscr{S}$. From dominated convergence we have $\alpha_\ell(\kappa_m)\nearrow\alpha_\ell(\kappa)$ for each ℓ. It is easy to check that for a regular finitary kernel κ' we have

$$\frac{1}{n}\,\mathbb{E}\left(P_\ell\left(G^\mathscr{V}(n,\kappa')\right)\right) \to \alpha_\ell(\kappa') \text{ and } \frac{1}{n}P_\ell\left(G^\mathscr{V}(n,\kappa')\right) \xrightarrow{\mathrm{p}} \alpha_\ell(\kappa').$$

In particular, if (3.6) holds then as $m\to\infty$ we have

$$\lim_{n\to\infty}\frac{1}{n}\,\mathbb{E}\left(P_\ell\left(G^\mathscr{V}(n,\kappa_m)\right)\right) \to \lim_{n\to\infty}\frac{1}{n}\,\mathbb{E}\left(P_\ell\left(G^\mathscr{V}(n,\kappa)\right)\right) < \infty. \tag{3.7}$$

Since $\kappa_m\leqslant\kappa$, we may couple $G^\mathscr{V}(n,\kappa_m)$ and $G^\mathscr{V}(n,\kappa)$ so that $G^\mathscr{V}(n,\kappa_m)\subset G^\mathscr{V}(n,\kappa)$. Condition (3.7) says essentially that, if ℓ is fixed and $m(n)\to\infty$, then almost all paths of length ℓ in $G^\mathscr{V}(n,\kappa)$ are already present in the subgraph $G^\mathscr{V}(n,\kappa_{m(n)})$.

The actual assumption we shall need is a tiny bit weaker. We shall assume that there is an increasing sequence κ_m of regular finitary kernels with $\kappa_m\leqslant\kappa$ pointwise, and a coupling $G^\mathscr{V}(n,\kappa_m)\subset G^\mathscr{V}(n,\kappa)$, with the following property: for any $\ell\geqslant 1$ and any $\varepsilon>0$ there is an $m_0=m_0(\ell,\varepsilon)$ such that

$$\mathbb{P}\big(|V_{\ell,m}|\geqslant\varepsilon n\big)\leqslant\varepsilon \tag{3.8}$$

for all $m\geqslant m_0$ and all large enough n, where $V_{\ell,m}$ is the set of vertices that are endpoints of a path of length ℓ in $G^\mathscr{V}(n,\kappa)$ not present in $G^\mathscr{V}(n,\kappa_m)$. In other words, if $m(n)\to\infty$, then $|V_{\ell,m(n)}|=o_p(n)$. As a path has a bounded number of endpoints, (3.7) immediately implies this condition: all paths present in $G^\mathscr{V}(n,\kappa_m)$ are also present in $G^\mathscr{V}(n,\kappa)$, and the expected number of additional paths is small.

Assuming continuity and (3.8), it is very easy to extend Theorem 3.1.

Theorem 3.2. *Let $k\geqslant 2$ be fixed. Let κ be a kernel on a vertex space $\mathscr{V}$. Suppose that κ is continuous almost everywhere on $\mathscr{S}\times\mathscr{S}$, and that* (3.3) *and* (3.4) *hold. If, in addition, $\kappa\mapsto$*

https://doi.org/10.1017/S0963548307008589 Published online by Cambridge University Press

$\beta^+(\kappa)$ is continuous from below at κ in the sense of (3.5)*, and* (3.8) *holds, then*

$$\frac{1}{n}c_k\big(G^{\mathscr{V}}(n,\kappa)\big) \xrightarrow{\mathrm{p}} \beta^+(\kappa)$$

as $n \to \infty$.

Proof. Let κ_m be a sequence of regular finitary kernels with $\kappa_m \leqslant \kappa$ pointwise, and c_m a sequence of real numbers with $c_m < 1$, such that $\beta^+(c_m\kappa_m) \to \beta^+(\kappa)$. Such sequences exists by our continuity assumption (3.5). Let $\varepsilon > 0$ be arbitrary.

Since $\beta^+(c_m\kappa_m) \to \beta^+(\kappa)$, there is an m with $\beta^+(c_m\kappa_m) \geqslant \beta^+(\kappa) - \varepsilon$. Since $c \mapsto \beta^+(c\kappa_m)$ is an increasing function of c, there is a c with $c_m < c < 1$ at which this function is continuous. As $c\kappa_m(x,y) \leqslant \kappa_m(x,y) \leqslant \kappa(x,y)$ for all x and y, we may couple $G^{\mathscr{V}}(n,c\kappa_m)$ and $G^{\mathscr{V}}(n,\kappa)$ so that $G^{\mathscr{V}}(n,c\kappa_m) \subset G^{\mathscr{V}}(n,\kappa)$ for every n. Applying Theorem 3.1 to the finite-type kernel corresponding to $c\kappa_m$ in the natural way, we have

$$c_k\big(G^{\mathscr{V}}(n,\kappa)\big) \geqslant c_k\big(G^{\mathscr{V}}(n,c\kappa_m)\big) = \beta^+(c\kappa_m)n + o_p(n),$$

so w.h.p.

$$c_k\big(G^{\mathscr{V}}(n,\kappa)\big) \geqslant (\beta^+(c\kappa_m) - \varepsilon)n \geqslant (\beta^+(c_m\kappa_m) - \varepsilon)n \geqslant (\beta^+(\kappa) - 2\varepsilon)n.$$

To prove the upper bound, let $\kappa_m \leqslant \kappa$ be a sequence of regular finitary kernels satisfying (3.8). (We shall not assume that $\beta^+(\kappa_m) \to \beta^+(\kappa)$ for this sequence.) Since $\mathscr{B}_d^+ \searrow \mathscr{B}^+$, there is a d with $\mathbb{P}_\kappa(\mathscr{B}_d^+) \leqslant \beta^+(\kappa) + \varepsilon$. By (3.8), for any $\eta > 0$ there is an m such that

$$\mathbb{P}\big(|V_{\ell,m}| \geqslant \varepsilon n/d\big) \leqslant \eta/d \tag{3.9}$$

for all $1 \leqslant \ell \leqslant d$ and all large enough n.

Let v be a vertex of $G^{\mathscr{V}}(n,\kappa)$, so v is also a vertex of $G^{\mathscr{V}}(n,\kappa_m)$. If v is in the k-core of $G^{\mathscr{V}}(n,\kappa)$, then either the d-neighbourhood $\Gamma_d(v, G^{\mathscr{V}}(n,\kappa))$ of v in $G^{\mathscr{V}}(n,\kappa)$ contains a tree with the property $\mathscr{B}_d^+$, or it contains a cycle. Thus, one of the following three cases holds: (i) $\Gamma_d(v, G^{\mathscr{V}}(n,\kappa_m))$ contains a tree with property $\mathscr{B}_d^+$, (ii) $\Gamma_d(v, G^{\mathscr{V}}(n,\kappa_m))$ contains a cycle, or (iii) $\Gamma_d(v, G^{\mathscr{V}}(n,\kappa))$ contains an edge not present in $G^{\mathscr{V}}(n,\kappa_m)$. In case (iii), v is an endpoint of a path of length at most d contained in $G^{\mathscr{V}}(n,\kappa)$ but not in $G^{\mathscr{V}}(n,\kappa_m)$, *i.e.*, $v \in \bigcup_{\ell \leqslant d} V_{\ell,m}$. From (3.9), with probability $1-\eta$ there are at most εn such v. By the finite-type equivalent of Corollary 2.7, the number of vertices with property (i) is

$$\mathbb{P}_{\kappa_m}(\mathscr{B}_d^+)n + o_p(n) \leqslant \mathbb{P}_\kappa(\mathscr{B}_d^+)n + o_p(n) \leqslant \beta^+(\kappa)n + \varepsilon n + o_p(n).$$

Also, by the same result, or by directly counting short cycles (noting that κ_m is bounded), only $o_p(n)$ vertices v have property (ii). It follows that, for any η, if n is large enough we have $c_k(G^{\mathscr{V}}(n,\kappa)) \leqslant (\beta^+(\kappa) + 3\varepsilon)n$ with probability at least $1 - \eta - o(1)$.

As $\eta > 0$ was arbitrary, we have shown that w.h.p.

$$(\beta^+(\kappa) - 2\varepsilon)n \leqslant c_k(G^{\mathscr{V}}(n,\kappa)) \leqslant (\beta^+(\kappa) + 3\varepsilon)n.$$

As $\varepsilon > 0$ was arbitrary, this completes the proof. □

As noted above, the rather cumbersome assumption (3.8) holds whenever (3.6) holds. In the next section, we shall consider an interesting example which satisfies (3.8) but not (3.6).

https://doi.org/10.1017/S0963548307008589 Published online by Cambridge University Press

In some sense, under mild assumptions, Theorem 3.2 is the complete answer to the question: 'How large is the k-core in $G^{\mathscr{V}}(n,\kappa)$?' In another sense, it is just the beginning: it remains to determine $\beta^+(\kappa)$. Although there is no further combinatorics involved, depending on the form of the solution required, this is in general very difficult. However, it is very easy to determine $\beta^+(\kappa)$ in terms of the solution to a certain functional equation.

From the definition of $\beta_x(\kappa)$ and the definition of the branching process, we have

$$\beta_x(\kappa) = \Psi_{\geqslant k-1}\left(\int_{\mathscr{S}} \kappa(x,y)\beta_y(\kappa)\,d\mu(y)\right). \tag{3.10}$$

Indeed, starting with a particle of type x, the number of children having property $\mathscr{B}$ is Poisson with mean $\int \kappa(x,y)\beta_y(\kappa)\,d\mu(y)$. With κ fixed, we regard (3.10) as a functional equation in a function $x \mapsto \beta_x(\kappa)$. Arguing as in the uniform case, it is easy to see that $\beta_x(\kappa)$ is given by the maximum solution to this equation (*i.e.*, the supremum of all solutions, which is itself a solution). Arguing as for (3.10), we have

$$\beta_x^+(\kappa) = \Psi_{\geqslant k}\left(\int_{\mathscr{S}} \kappa(x,y)\beta_y(\kappa)\,d\mu(y)\right), \tag{3.11}$$

while, of course,

$$\beta^+(\kappa) = \int_{\mathscr{S}} \beta_x^+(\kappa)\,d\mu(x). \tag{3.12}$$

Turning to the continuity of $\kappa \mapsto \beta^+(\kappa)$ at κ, suppose that $\kappa_m \nearrow \kappa$, and define γ_x by

$$\gamma_x = \lim_{m\to\infty} \beta_x(\kappa_m). \tag{3.13}$$

This limit exists as the sequence is increasing and bounded by 1. By monotone convergence,

$$\int_{\mathscr{S}} \kappa(x,y)\gamma_y\,d\mu(y) = \lim_{m\to\infty} \int_{\mathscr{S}} \kappa_m(x,y)\beta_y(\kappa_m)\,d\mu(y)$$

for every x, so

$$\begin{aligned}\Psi_{\geqslant k-1}\left(\int_{\mathscr{S}} \kappa(x,y)\gamma_y\,d\mu(y)\right) &= \lim_{m\to\infty} \Psi_{\geqslant k-1}\left(\int_{\mathscr{S}} \kappa_m(x,y)\beta_y(\kappa_m)\,d\mu(y)\right)\\ &= \lim_{m\to\infty} \beta_x(\kappa_m) = \gamma_x.\end{aligned} \tag{3.14}$$

In other words, γ_x also satisfies the functional equation (3.10). Also, again by monotone convergence,

$$\lim_{m\to\infty} \beta^+(\kappa_m) = \Psi_{\geqslant k}\left(\int_{\mathscr{S}} \kappa(x,y)\gamma_y\,d\mu(y)\right). \tag{3.15}$$

In sufficiently nice cases, this can be used to establish the continuity of κ, by showing that the functional equation (3.10) has only one strictly positive solution, and using a simple special form of κ_1 to ensure that $\lim_{m\to\infty} \beta^+(\kappa_m) \geqslant \beta^+(\kappa_1) > 0$.

3.3. Case study: a power-law or 'scale-free' graph

One of the most studied examples of the general model $G^{\mathscr{V}}(n,\kappa)$ is the following: let $\mathscr{S} = (0,1]$, with μ Lebesgue measure, let $\kappa(x,y) = c/\sqrt{xy}$, and set $x_i^{(n)} = i/n$. Then the edges of $G^{\mathscr{V}}(n,\kappa)$

https://doi.org/10.1017/S0963548307008589 Published online by Cambridge University Press

are present independently, and the probability that the edge ij is present is $c/\sqrt{ij}$, or, if $c > \sqrt{2}$, the minimum of this quantity and 1. This random graph $G^{\curlyvee}(n,\kappa)$ is the 'mean-field' version of the 'scale-free' random graph introduced by Barabási and Albert in [2] as a model of the world-wide web, and studied in many papers, especially in the computer science and statistical physics literature.

For this kernel, $\alpha_\ell(\kappa) = \infty$ for all $\ell \geqslant 2$: the 'early' vertices (with i/n small) have large degree, and, while the average degree is bounded as $n \to \infty$, the average square degree is not. If ℓ is fixed, i/n is bounded away from zero, and $\varepsilon \to 0$ slowly, then the expected number of paths of length ℓ from vertex i to an 'early' vertex j with $j \leqslant \varepsilon n$ is large (order $(\log n)^{\ell-1}$), and one can check that (3.7) does not hold. On the other hand, it is easy to check that the probability that there is such a path tends to zero uniformly in n as $\varepsilon \to 0$. Indeed, this probability is bounded by the expected number of *late–early* paths of length at most ℓ starting at i, where a late–early path is a path all of whose vertices have indices at least εn, apart from the last vertex which does not; see [22, Section 4].

Let $\kappa^{(\varepsilon)}$ be the 'truncated' kernel that agrees with κ on $[\varepsilon,1]^2$ and is zero otherwise; in the graph, this corresponds to deleting all edges incident with early vertices. From the observation above, if $\varepsilon = \varepsilon(n) \to 0$, then $o_p(n)$ vertices are within graph distance ℓ of an early vertex, so only $o_p(n)$ vertices of $G^{\curlyvee}(n,\kappa)$ are incident with a path of length ℓ in $G^{\curlyvee}(n,\kappa)$ not present in $G^{\curlyvee}(n,\kappa^{(\varepsilon)})$. Letting $\varepsilon \to 0$ slowly enough, it is easy to approximate the bounded kernels $\kappa^{(\varepsilon)}$ by regular finitary kernels (for example, step functions), and so to deduce that (3.8) holds. Thus we shall be able to apply Theorem 3.2, if we can establish the required continuity of β^+.

Since we shall vary the parameter c, it will be convenient to write κ as $c\kappa_0$, where $\kappa_0(x,y) = 1/\sqrt{xy}$. The size of the giant component in $G^{\curlyvee}(n,c\kappa_0)$ was found in [22]; in Section 6.1 it was shown that the size is $\sigma(c)n + o_p(n)$, where $\sigma(c)$ is the survival probability of the branching process $X_{c\kappa_0}$. (This is a special case of the main result of [5].) In Section 6.2 of [22], this survival probability is calculated in terms of the exponential integral. It turns out that there is a giant component for *any* $c > 0$, although it is extremely small when c is small: $\sigma(c) \sim 2e^{1-\gamma}\exp(-1/(2c))$ as $c \to 0$, where γ is Euler's constant.

One might expect that for k fixed, there would be a k-core in $G^{\curlyvee}(n,c\kappa_0)$ for any $c > 0$, perhaps with size a constant fraction of that of the giant component when c is small. In fact, there is a positive threshold above which the k-core first appears. Just above this threshold, the k-core is small: this is in sharp contrast to $G(n,\lambda/n)$. In this result we write β_k^+ for β^+ when it is necessary to indicate the dependence on k.

Theorem 3.3. *Let* $\kappa_0(x,y) = 1/\sqrt{xy}$. *For* $c > 0$, *let* $G^{\curlyvee}(n,c\kappa_0)$ *be the graph on* $[n] = \{1,2,\dots,n\}$ *in which edges are present independently, and the probability that* i *and* j *are joined is* $\min\{c/\sqrt{ij},1\}$. *For each* $k \geqslant 2$ *we have*

$$c_k\big(G^{\curlyvee}(n,c\kappa_0)\big) = \beta_k^+(c\kappa_0)n + o_p(n).$$

If $k \geqslant 3$, *then* $\beta_k^+(c\kappa_0) = 0$ *for* $c \leqslant (k-2)/2$, *while*

$$\beta_k^+(c\kappa_0) \sim \frac{(k-1)!^{2/(k-2)}}{(k-1)(k-2)}\varepsilon^{2/(k-2)} \tag{3.16}$$

when $c = (1+\varepsilon)(k-2)/2$ *and* $\varepsilon \to 0$ *from above.*

https://doi.org/10.1017/S0963548307008589 Published online by Cambridge University Press

If $k = 2$, then $\beta_k^+(c\kappa_0) > 0$ for every $c > 0$, and

$$\beta_k^+(c\kappa_0) \sim \frac{1}{2c} e^{2-2\gamma} \exp(-1/c) \tag{3.17}$$

as $c \to 0$.

Thus, there is always a 'giant' 2-core; for small c its normalized size is essentially the square of that of the giant component (times a factor $\Theta(1/c)$, which is logarithmic in terms of the normalized size of the giant component). For $k \geqslant 3$, the k-core emerges at a positive threshold, $c = (k-2)/2$, and does so slowly. In the terminology of [5], for $k \geqslant 3$ the emergence of the k-core exhibits a phase transition of *exponent* $2/(k-2)$, where this is the exponent of ε appearing above. This contrasts with the exponent 0 transition in $G(n, \lambda/n)$. (The term 'order' is often used in this context, but not always in the same way.)

Proof. As noted above, the graph $G^{\mathscr{V}}(n, c\kappa_0)$ satisfies (3.8). Thus the first statement follows from Theorem 3.2 once we have established the required continuity of β^+ at $c\kappa_0$. We return to this later.

For the second and third statements, let us write β_x for $\beta_x(c\kappa_0)$, which depends on c and also on k. Then equation (3.10) becomes

$$\beta_x = \Psi_{\geqslant k-1}\left(\int_0^1 \frac{c\beta_y}{\sqrt{xy}}\, dy \right). \tag{3.18}$$

Define $A = A_k(c)$ by

$$A = A_k(c) = \int_0^1 \frac{c\beta_y}{\sqrt{y}}\, dy. \tag{3.19}$$

Then

$$\beta_x = \Psi_{\geqslant k-1}(A/\sqrt{x}), \tag{3.20}$$

so to determine β_x it remains to determine A. Substituting (3.20) into (3.19), we see that $A = cf(A)$, where

$$f(B) = f_k(B) := \int_0^1 \frac{\Psi_{\geqslant k-1}(B/\sqrt{y})}{\sqrt{y}}\, dy.$$

So far the argument is very similar to that in Section 6.2 of [22]; indeed, for $k = 2$ the event $\mathscr{B}$ is the event that the branching process survives, so β_x is the probability of survival starting with a particle of type x; this is denoted $S_\infty(x)$ in [22] (see p. 919), where it is given by (3.20) with

$$A_2(c) \sim e^{1-\gamma} \exp(-1/(2c)) \tag{3.21}$$

as $c \to 0$, where γ is Euler's constant. It turns out that the case $k \geqslant 3$ is much easier to handle.

Suppose that $k \geqslant 3$. Then, substituting $y = B^2 x^{-2}$, so $x = B/\sqrt{y}$, we have

$$f_k(B) = \int_B^\infty \frac{x}{B} \Psi_{\geqslant k-1}(x) 2B^2 x^{-3}\, dx = 2B g_k(B),$$

where

$$g_k(B) := \int_B^\infty \Psi_{\geqslant k-1}(x) x^{-2}\, dx. \tag{3.22}$$

https://doi.org/10.1017/S0963548307008589 Published online by Cambridge University Press

For any given $k \geqslant 3$, it is straightforward to calculate $f_k(B)$ explicitly. Indeed, $\Psi_{\geqslant k-1}(x) = \sum_{t \geqslant k-1} x^t \exp(-x)/t!$, and for $t \geqslant 2$ we have $\int_0^\infty x^{t-2} \exp(-x)\, dx = (t-2)!$. Thus, for $k \geqslant 3$,

$$g_k(0) = \sum_{t \geqslant k-1} \frac{(t-2)!}{t!} = \sum_{t \geqslant k-1} \frac{1}{t(t-1)} = \frac{1}{k-2}.$$

In particular, $g_3(0) = 1$. It is easy to verify by differentiating that

$$g_3(B) = \frac{1 - e^{-B}}{B},$$

so

$$f_3(B) = 2(1 - e^{-B}) = 2B - B^2 + O(B^3).$$

Also, for each $k \geqslant 4$,

$$g_k(B) - g_{k+1}(B) = \int_B^\infty \frac{x^{k-3}}{(k-1)!} \exp(-x)\, dx,$$

which is $\exp(-B)$ times a polynomial in B that may be easily evaluated. Rather than do this, let us note that, as $B \to 0$,

$$g_k(0) - g_k(B) = \int_0^B \frac{x^{k-1}}{(k-1)!} x^{-2} + O(x^{k-2})\, dx = \frac{B^{k-2}}{(k-2)(k-1)!} + O(B^{k-1}),$$

so

$$f_k(B) = \frac{2B}{k-2}\left(1 - \frac{B^{k-2}}{(k-1)!} + O(B^{k-1})\right). \tag{3.23}$$

Note that $f_k(B)/B = 2g_k(B)$ is decreasing (from the form of (3.22)). Thus the equation $B = cf_k(B)$ has a positive solution if and only if $cf_k'(0) > 1$, *i.e.*, if and only if $c > (k-2)/2$. Furthermore, when $c > (k-2)/2$, the solution $A_k(c)$ is unique.

Let $c_0 = c_0(k) = (k-2)/2$, and let $c = (1+\varepsilon)c_0$ with $\varepsilon > 0$. From (3.23) and the fact that $f_k(B)/B$ is decreasing in B, it follows that $A = A_k(c) \to 0$ as $\varepsilon \to 0$. Furthermore, as $\varepsilon \to 0$,

$$\begin{aligned} A = cf_k(A) &= \frac{c}{c_0} A\left(1 - (1+o(1))\frac{A^{k-2}}{(k-1)!}\right) \\ &= A + \varepsilon A - (1+o(1))A\frac{A^{k-2}}{(k-1)!}, \end{aligned}$$

so

$$A = A_k(c) = A_k((1+\varepsilon)c_0) \sim \left((k-1)!\varepsilon\right)^{1/(k-2)}. \tag{3.24}$$

Recalling that any solution to the functional equation (3.18) has the form (3.20) with A satisfying $A = cf(A)$, and that the probability $\beta_x = \mathbb{P}_{c\kappa,x}(\mathscr{B})$ is given by the maximum solution to this equation, we have shown that if $c \leqslant c_0(k)$, then $\beta_x = 0$ for all x, while if $c > c_0(k)$, then

$$\beta_x = \Psi_{\geqslant k-1}(A_k(c)/\sqrt{x}).$$

Using (3.11) and (3.12), we have $\beta^+(c\kappa_0) = 0$ if $c \leqslant c_0(k)$. Otherwise,

$$\beta^+(c\kappa_0) = \int_{x=0}^1 \Psi_{\geqslant k}(A_k(c)/\sqrt{x})\, dx. \tag{3.25}$$

https://doi.org/10.1017/S0963548307008589 Published online by Cambridge University Press

Let

$$h(B) = \int_{x=0}^{1} \Psi_{\geqslant k}(B/\sqrt{x})\,dx = \int_{y=B}^{\infty} \Psi_{\geqslant k}(y) 2B^2 y^{-3}\,dy.$$

As $B \to 0$ we have

$$\frac{h(B)}{2B^2} \to \int_0^{\infty} \sum_{t \geqslant k} \frac{y^{t-3}}{t!} e^{-y}\,dy = \sum_{t \geqslant k} \frac{(t-3)!}{t!} = \frac{1}{2(k-1)(k-2)}.$$

Hence, from (3.24) and (3.25),

$$\beta^+\big((1+\varepsilon)c_0\kappa_0\big) \sim \frac{(k-1)!^{2/(k-2)}}{(k-1)(k-2)} \varepsilon^{2/(k-2)},$$

which is exactly (3.16).

Returning to the case $k = 2$, in this case, by reducing to the exponential integral as in [22], one can show that $h(B) \sim \log(1/B)B^2$ as $B \to 0$. Using (3.21), (3.17) follows.

To complete the proof of Theorem 3.3, it remains to show that there is a sequence κ_m of finite-type kernels with $\sup_{x,y} \kappa_m/(c\kappa_0) < 1$ for every m, such that $\beta^+(\kappa_m) \to \beta^+(c\kappa_0)$. We may assume that $c > (k-2)/2$, as otherwise $\beta^+(c\kappa_0) = 0$ and there is nothing to prove. Given a sequence κ_m tending up to κ, let γ_x be defined by (3.13). Then, from (3.14), the function $x \mapsto \gamma_x$ also satisfies the functional equation (3.18). We have shown above that this functional equation has exactly two solutions, the zero function, and β_x. Hence, writing γ^+ for $\lim_{m\to\infty} \beta^+(\kappa_m)$, from (3.15) we have $\gamma^+ = 0$ or $\gamma^+ = \beta^+(c\kappa_0)$. We shall rule out the former case by constructing the first element κ_1 of our approximating sequence suitably. As a first step, we consider a bounded kernel that approximates $c\kappa_0$.

Let $(k-2)/2 < c' < c$ be fixed. Given $\varepsilon > 0$, let $\kappa^{(\varepsilon)}$ be the kernel given by $\kappa^{(\varepsilon)}(x,y) = c'\kappa_0(x,y) = c'/\sqrt{xy}$ if $x,\ y \geqslant \varepsilon$, and $\kappa^{(\varepsilon)}(x,y) = 0$ otherwise. Arguing exactly as for $c\kappa_0$, we have

$$\beta_x(\kappa^{(\varepsilon)}) = \Psi_{\geqslant k-1}(A^{(\varepsilon)}/\sqrt{x})$$

for $\varepsilon \leqslant x \leqslant 1$, and $\beta_x(\kappa^{(\varepsilon)}) = 0$ otherwise, where $A^{(\varepsilon)}$ is the largest solution to $A^{(\varepsilon)} = c' f^{(\varepsilon)}(A^{(\varepsilon)})$, with

$$f^{(\varepsilon)}(B) = \int_{\varepsilon}^{1} \frac{\Psi_{\geqslant k-1}(B/\sqrt{y})}{\sqrt{y}}\,dy.$$

Since $c' > (k-2)/2$, there is a $B > 0$ with $c' f(B) > B$; fix such a B. As $\varepsilon \to 0$, we have $f^{(\varepsilon)}(B) \nearrow f(B)$, so there is an $\varepsilon > 0$ with $c' f^{(\varepsilon)}(B) > B$. It follows that $A^{(\varepsilon)} > 0$, and hence that $\beta^+(\kappa^{(\varepsilon)}) > 0$. To complete the proof, let κ_1 be a regular finitary kernel with $c'\kappa^{(\varepsilon)}(x,y) \leqslant \kappa_1(x,y) \leqslant \frac{1}{2}(c+c')\kappa(x,y)$ for all x, y. Such a kernel is easy to construct as κ is continuous and bounded away from 0 on the compact set $[\varepsilon,1]^2$, where it coincides with $\kappa^{(\varepsilon)}$. It is easy to construct a sequence of finite-type kernels κ_m with $\sup_{x,y} \kappa_m/(c\kappa_0) < 1$, with κ_m tending up to $c\kappa_0$ and starting with this particular κ_1. Then $\gamma^+ \geqslant \beta^+(\kappa_1) \geqslant \beta^+(c'\kappa^{(\varepsilon)}) > 0$. Since $\gamma^+ = 0$ or $\gamma^+ = \beta^+(c\kappa_0)$, we have $\gamma^+ = \beta^+(c\kappa_0)$ as required. □

There is a sense in which Theorem 3.3 does not illustrate the full power of Theorem 3.2. Indeed, the kernel $c/\sqrt{xy}$ has a special property: it may be written as $\kappa(x,y) = \phi(x)\phi(y)$ for some

https://doi.org/10.1017/S0963548307008589 Published online by Cambridge University Press

function ϕ on $\mathscr{S}$. Such kernels are called *rank 1* in [5]. The branching process corresponding to a rank 1 kernel is much simpler than the general case: roughly speaking, while the distribution of the *number* of children of a particle depends on its type, the distribution of their *types* does not.

There is a minor variant of the model $G^{\times}(n,\kappa)$ where the edge probabilities are taken as $p_{ij} = \kappa(x_i^{(n)}, x_j^{(n)})/(n + \kappa(x_i^{(n)}, x_j^{(n)}))$, so $p_{ij}/(1-p_{ij}) = \kappa(x_i^{(n)}, x_j^{(n)})/n$. As in [5], all our results here apply equally to this variant. As noted in Section 16.4 of [5], using this variant, if κ has rank 1 then, conditional on the degree sequence, $G^{\times}(n,\kappa)$ is equally likely to be any graph with the given degree sequence. Thus the structure of $G^{\times}(n,\kappa)$ is simpler than in the general case. Also, $G^{\times}(n,\kappa)$ is then closely related to random graphs defined by first fixing a degree sequence (perhaps exactly, or perhaps asymptotically), and then choosing a random graph with this degree sequence.

Graphs of this form have been studied by many authors, including Molloy and Reed [20] in the general case, and Aiello, Chung and Lu [1] in the power-law case. The k-core of such graphs has been studied by Janson and Luczak [15] and by Fernholz and Ramachandran [13]; it is probable that Theorem 3.3 could be proved by the methods of either of these papers, although the calculation of $\beta^+(c\kappa_0)$ must still be carried out.

In fact, Fernholz and Ramachandran studied the k-core in random graphs with a given power-law degree sequence with exponent α: they assume a limiting fraction $i^{-\alpha}/\zeta(\alpha)$ of vertices with degree i for each $i \geqslant 1$. When $\alpha = 3$, this model is very close to $G^{\times}(n, c\kappa_0)$, where the degree distribution satisfies $\mathbb{P}(\deg(v) = i) \sim ai^{-3}$ as $i \to \infty$, for a constant a depending on c. Note, however, that, having fixed the degree exponent, the model considered here is much more flexible, due to the presence of the parameter c, which allows control of the overall number of edges. For this reason, the result in [13], that w.h.p. there is a k-core if $\alpha < 3$, and w.h.p. there is no k-core when $\alpha \geqslant 3$, gives no insight into the transition studied in Theorem 3.3. Roughly speaking, the $\alpha = 3$ case of this result corresponds to showing that there is no 3-core in $G^{\times}(n, c\kappa_0)$ for a specific c. However, the correspondence is not direct: in the model of [13], the presence or absence of the k-core depends very much on the entire degree distribution, not just its asymptotics, and in $G^{\times}(n, c\kappa_0)$, the distribution of the small degrees does not follow exactly a power-law.

It would be interesting to use Theorem 3.2 to compute the size of the k-core in examples of $G^{\times}(n,\kappa)$ where κ does not have rank one. A particularly interesting case is the kernel $\kappa(x,y) = c/\max\{x,y\}$ on $(0,1]^2$, corresponding to the 'uniformly grown random graph' proposed by Dubins (as an infinite random graph) in 1984 (see [16]). A closely related model was introduced by Callaway, Hopcroft, Kleinberg, Newman and Strogatz [7] in 2001. The emergence of the giant component in this model shows particularly interesting behaviour: at $c = 1/4$ there is an 'infinite order' phase transition; see [11, 12, 4, 22]. The functional equation (3.10) is likely to be harder to handle in this case, but may well still be tractable. Indeed, although the exact solution could not be calculated, good bounds on the size of the giant component just above the transition were obtained in [22] by bounding the solution to a related functional equation for the giant component in a generalization of this model.

Acknowledgements

The author would like to thank Alan Frieze for suggesting studying the k-core of $G^{\times}(n,\kappa)$, and Béla Bollobás and Svante Janson for helpful discussions on this topic.

https://doi.org/10.1017/S0963548307008589 Published online by Cambridge University Press

References

[1] Aiello, W., Chung, F. and Lu, L. (2001) A random graph model for power law graphs. *Experiment. Math.* **10** 53–66.

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

[3] Bollobás, B. (1984) The evolution of sparse graphs. In *Graph Theory and Combinatorics: Cambridge 1983*, Academic Press, London, pp. 35–57.

[4] Bollobás, B., Janson, S. and Riordan, O. (2005) The phase transition in a uniformly grown random graph has infinite order. *Random Struct. Alg.* **26** 1–36.

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

[6] Cain, J. and Wormald, N. (2006) Encores on cores. *Electron. J. Combin.* **13** R81.

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

[8] Chvátal, V. (1991) Almost all graphs with 1.44n edges are 3-colorable. *Random Struct. Alg.* **2** 11–28.

[9] Cooper, C. (2004) The cores of random hypergraphs with a given degree sequence. *Random Struct. Alg.* **25** 353–375.

[10] Darling, R. W. R. and Norris, J. R. (2005) Differential equation approximations for Markov chains. Manuscript.

[11] Dorogovtsev, S. N., Mendes, J. F. F. and Samukhin, A. N. (2001) Anomalous percolation properties of growing networks. *Phys. Rev. E* **64** 066110.

[12] Durrett, R. (2003) Rigorous result for the CHKNS random graph model. In *Proc. Discrete Random Walks 2003* (C. Banderier and C. Krattenthaler, eds), *Discrete Mathematics and Theoretical Computer Science* **AC** 95–104. `http://dmtcs.loria.fr/proceedings/`

[13] Fernholz, D. and Ramachandran, V. (2003) The giant k-core of a random graph with a specified degree sequence. Manuscript.

[14] Harris, T. E. (1960) A lower bound for the critical probability in a certain percolation process. *Proc. Cam. Philos. Soc.* **56** 13–20.

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

[16] Kalikow, S. and Weiss, B. (1988) When are random graphs connected? *Israel J. Math.* **62** 257–268.

[17] Kallenberg, O. (2002) *Foundations of Modern Probability*, 2nd edn, Springer, New York.

[18] Kleitman, D. J. (1966) Families of non-disjoint subsets, *J. Combin. Theory* **1** 153–155.

[19] Molloy, M. (2005) Cores in random hypergraphs and Boolean formulas. *Random Struct. Alg.* **27** 124–135.

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

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

[22] Riordan, O. (2005) The small giant component in scale-free random graphs. *Combin. Probab. Comput.* **14** 897–938.

[23] Söderberg, B. (2002) General formalism for inhomogeneous random graphs. *Phys. Rev. E* **66** 066121.

https://doi.org/10.1017/S0963548307008589 Published online by Cambridge University Press