ssdnm
środowiskowe
studia doktoranckie
z nauk matematycznych

Bartosz Zaleski

Uniwersytet A. Mickiewicza w Poznaniu

On k-cores in random hypergraphs

Praca semestralna nr 2

(semestr zimowy 2012/13)

Opiekun pracy: Tomasz Łuczak

KAPITAŁ LUDZKI
NARODOWA STRATEGIA SPÓJNOŚCI

UNIA EUROPEJSKA
EUROPEJSKI
FUNDUSZ SPOŁECZNY

ON k-CORES IN RANDOM HYPERGRAPHS

BARTOSZ ZALESKI

ABSTRACT. In [5] Riordan showed that for every $k \geq 2$ random graph $G(n, \lambda/n)$ contains large $k-$core if λ is greater than some threshold λ_0. The ingenuity of Riordan's approach lays in coupling the consecutive neighbourhoods of vertices with a naturally defined branching process, and then showing that certain local properties of this process enforce global property of existence of the k-core. In this paper we prove, using Riordan's method, that a random 3-uniform hypergraph $G_3(n, \lambda/N)$, for $N = \binom{n}{2}$ and large enough λ, contains a large subhypergraph in which every vertex belong to at least k edges. This is a special case of a result of Molloy [4] which states that for any positive integers k and r, where $k + r > 2$, there exist a constant $c_{r,k}$ such that for $c < c_{r,k}$ with high probability there is no k-core in random r-uniform hypergraph $G_r(n, c/n^{r-1})$ and for $c > c_{r,k}$, such structure exists with high probability and is large. However, our argument is completely different from Molloy's approach.

1. INTRODUCTION

Let $G_3(n, p)$ be a random 3-uniform hypergraph on n vertices, such that each triple of vertices forms an edge independently, with probability p. By a k-core in a hypergraph we mean a subhypergraph in which every vertex belongs to at least k edges.

In [5] Riordan used tools from branching processes theory to show that random graph $G(n, \lambda/n)$ comprises a nontrivial k-core when $\lambda > \lambda_k$, for some threshold value λ_k. He utilised the fact that, for every vertex in $G(n, \lambda/n)$ its successive neighbourhoods might be coupled with specific Poisson process. We show that this method may also be applied to search for k-cores in 3-uniform random hypergraphs. Many of minor facts we present in this paper have their analogues in [5] and are just rewritten to fit our case. Because of that, we omit proofs of statements that does not significantly differ from the original ones.

We start with some convention. As the branching process with a single vertex in the first generation may be seen as a rooted tree, we shall often not distinguish between those two structures when describing their properties.

By $\Psi_{\geq t}(\lambda)$ (respectively $\Psi_t(\lambda)$) we mean the probability that Poisson distributed variable with mean λ assumes value at least t (respectively exactly t). Let X_λ be a branching process with single vertex x_0 in the starting generation in which the number of children of each particle has a Poisson distribution multiplied by two (i.e.

the probability that a particle has $2t$ children is equal to $\Psi_t(\lambda)$). We refer to the two particles that were generated by this branching process at the same time by a pair. To distinguish probability spaces of different branching processes, we shall use $\mathbb{P}_\lambda(\mathcal{A})$ to denote the probability of an event $\mathcal{A}$ in branching process X_λ. By $\mathcal{B}$ (respectively $\mathcal{B}^+$) we denote the event that the root of the process has at least $k-1$ pairs (respectively k pairs) of children, each of which have at least $k-1$ pair of direct descendants and so on. Moreover, by $\beta(\lambda)$ (respectively $\beta^+(\lambda)$) we mean probabilities of those events in process X_λ. Let $\mathcal{B}_d$ be the event that the branching process contains a $(2k-2)$-ary tree (in the sense that every inner particle of this tree has $k-1$ pairs of children) of height d with the root in the first generation. Thus, we can view the event $\mathcal{B}$ as the limit of events $\mathcal{B}_d$. On the other hand, as the numbers of children of different elements in the process are independent random variables, we get the following recursive (in some sense) equation for the probability of event $\mathcal{B}_d$

$$\mathbb{P}(\mathcal{B}_{d+1}) = \Psi_{\geq k-1}(\lambda(\mathbb{P}(\mathcal{B}_d))^2).$$

Thus, $\beta(\lambda)$ can be also defined as the largest positive solution to the equation

$$x = \Psi_{\geq k-1}(\lambda x^2).$$

Moreover, by a similar reasoning, we obtain that

$$\beta^+(\lambda) = \Psi_{\geq k}(\lambda(\beta(\lambda))^2).$$

Let λ_0 be the infimum of those values of λ for which the $\beta(\lambda)$ is positive. It is easy to see that both $\beta(\lambda)$ and $\beta^+(\lambda)$ are continuous and increasing for $\lambda > \lambda_0$. We shall use this observation in the proof of some auxiliary facts in the next section.

Let $\mathcal{E}_1, \mathcal{E}_2$ be properties of the branching process, where $\mathcal{E}_1$ depends only on first d generations. Then, by $\mathcal{E}_1 \circ \mathcal{E}_2$ we denote the event that the process has property $\mathcal{E}_1$ even after removing all elements in d-th generation that does not have property $\mathcal{E}_2$. For example, if $\mathcal{O}$ is the property that a particle has no children, then $\mathcal{B}_1 \circ \mathcal{O}$ is the property that the root has at least $k-1$ pairs of children that have no children at all. With this notation we can write that $\mathcal{B}_1 \circ \mathcal{B}_1 = \mathcal{B}_2$.

Let $\mathcal{R}_d$ be the event that $\mathcal{B}_d$ holds even after a single pair of particles is deleted in the generation d. For example $\mathcal{R}_d \circ \mathcal{B}$ is the event that $\mathcal{B}$ holds even if we delete a pair of particles in generation d.

We will use similar notation, namely $\mathcal{E} \circ M$, to denote the property that $\mathcal{E}$ (dependant on d first generations) holds after deleting from the d-th generation all particles that were not marked (by some process of choosing them). By $\mathbb{P}^p_\lambda(\mathcal{E} \circ M)$ we denote the probability that, for X_λ, $\mathcal{E} \circ M$ holds if the marking of elements in generation d is done independently with probability p.

2. Main result

The main goal of this paper is to apply Riordan's method from [5] to show the following special case of more general Molloy's result [4].

Theorem 2.1. *Let $\lambda > \lambda_0$, $N = \binom{n}{2}$ and let $\beta^+(\lambda)$ be defined as in the previous section. Then, with high probability, hypergraph $G_3(n, \lambda/N)$ contains the k-core of size $\beta^+(\lambda)n + o(n)$.*

Let us briefly discuss the connection between this statement and branching processes. It is easy to observe that, for a fixed vertex v, the number of edges of hypergraph $G_3(n, \lambda/N)$ which contains v has asymptotically the Poisson distribution with mean λ. Moreover, for the values of edge probability we are interested in, the number of short cycles in the random hypergraph is very small. So we explore this hypergraph from the point of view of a single vertex by embedding an appropriate branching process rooted in this vertex. As the branching process which does not extinct is, by its nature, an infinite object we will show that there is a certain property, dependant on some small number of generations (chosen to allow embedding of part of the process into a hypergraph), close (in some sense) to the event $\mathcal{B}$. We will define this property to guarantee that if the particle has this property then, with high probability, there is a $(2k-2)$-ary tree rooted in this particle, of some height, such that each leaf of this tree also has this property. Then, after passing to the hypergraph, we will say that a vertex has the same property as the branching process if such tree can be embedded in its consecutive neighbourhoods. Taking the union of these trees we shall find the k-core we are looking for.

We begin with several facts concerning branching process X_λ. As proofs of the following lemmata are almost identical to the ones given in [5], we omit them.

Lemma 2.2. *For $\lambda > \lambda_0$ and $d \to \infty$,*

$$\mathbb{P}_\lambda(\mathcal{R}_d \circ \mathcal{B}) \nearrow \beta(\lambda).$$

□

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

Lemma 2.3. *Let $\lambda_1 < \lambda_2$ be fixed with $\lambda_1 > \lambda_0$. Then there exists d such that*

$$r(\lambda_2, d, \beta(\lambda_2)) \geq \beta(\lambda_1).$$

□

It will often be useful to look at our process from a slightly different point of view. As for each particle its children are generated in pairs, we can treat the obtained structure not as a tree, but (in some sense) as its triangulation. In Figure 1 we illustrate this process. We will often interchangeably use the notion of a tree and its triangular structure, which is, in some sense, equivalent to a hypergraph tree. For tree T by $\bar{T}$ we denote the described triangulation. We say that this triangular structure has property $\mathcal{M}_{\leq d}$ if all its leaves (vertices belonging to exactly $k-1$ triangles) are at distance at most d from the root. The following fact is an

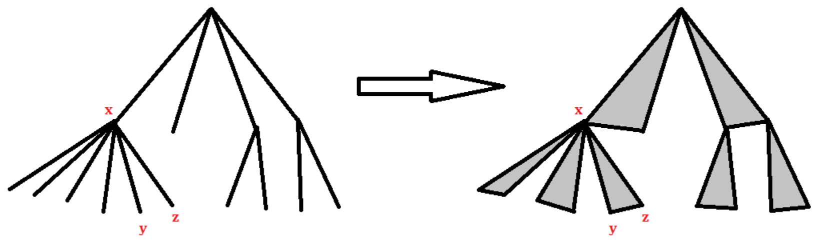

FIGURE 1. To get the desired triangulation we connect children y and z of an element x with an edge and treat triple xyz as a triangle.

analogous of a similar observation from [5].

Lemma 2.4. *Let $\bar{T}$ be a rooted triangular structure obtained from some rooted tree T minimal with respect to having property $\mathcal{R}_d \circ \mathcal{B}_1$. Let w be a vertex of $\bar{T}$ at distance d from the root and let y be the parent of w. If the structure $\bar{T}$ is regarded as rooted in w, then there is a substructure $\bar{W}$ with the property $\mathcal{M}_{\leq 2d+1}$, such that all leaves of $\bar{W}$ are leaves of $\bar{T}$.* □

The last problem we have to discuss before passing to the proof of Theorem 2.1 concerns short cycles (the length of a cycle is the number of triangles in it) in our hypergraph. Existence of such cycles impose differences between trees rooted in different vertices which complicates the argument. Firstly, let us specify what we mean as a cycle in hypergraph. We assume the most general definition of a cycle in a hypergraph, i.e. we consider as a cycle any sequence of unique edges, such that two consecutive edges (we suppose that the first and the last one are consecutive) share at least one vertex. A hypergraph without cycles will be called a tree. To avoid mentioned complications we shall study a hypergraph $\tilde{G} = \tilde{G}(n, l, p)$ which is hypergraph $G_3(n, p)$ conditioned on absence of cycles of length at most l. As it turns out, putting this constraint does not change the model much. We shall show that the probability of an edge in this model is roughly λ/N even, if we additionally condition on the existence of some small set of edges and the absence of some other set.

Lemma 2.5. *Let $\lambda > 0$, $l = l(n) = o(\log n)$ and $N = \binom{n}{2}$. For large enough n, as long as E_0, E_1, e are disjoint sets of possible edges of $\tilde{G}$ with $|E_1| \leq n^{1/5}$ and supposing that $E_1 \cup \{e\}$ does not contain a cycle of length at most l following* *inequality holds*

$$(1-n^{-1/6})\lambda/N \leq \mathbb{P}(e \in E(\tilde{G})|E_1 \subset E(\tilde{G}) \subset E_0^c).$$

Proof. We call $P \subset (E_0 \cup E_1 \cup \{e\})^c$ a precycle if it is minimal set of edges such that $P \cup E_1 \cup \{e\}$ contains a short cycle including e. Let G' a random hypergraph $\tilde{G}$ conditioned on the existence of every edge from E_1 and the absence of all the edges from E_0. Let $E \subset (E_0 \cup E_1)^c$ be the set of edges included in G'. Its obvious that possible configurations of edges in G' have the same relative probabilities as in $G_3(n, \lambda/N)$. Thus, the probability of e in G' conditioned on knowing the rest of G' is 0 if $G' \setminus \{e\}$ contains a precycle, and λ/N otherwise. So,

$$\begin{aligned}\mathbb{P}(e \in G') &= \lambda/N\mathbb{P}(G'\setminus\{e\} \text{ contains no precycle})\\ &= \lambda/N\mathbb{P}(E \text{ contains no precycle}|E \cup E_1 \text{ contains no cycle of length } l)\\ &\geq \lambda/N\mathbb{P}(E \text{ contains no precycle}),\end{aligned}$$

where the last inequality is a straightforward consequence of a special case of the FKG inequality, so called Harris lemma (cf. [1], [2]). W shall show that $\mathbb{P}(E \text{ contains a precycle}) \leq n^{-1/6}$. Notice that a precycle must contain a sequence of edges of length $1 \leq t \leq l$, such that each two consecutive edges share one or two vertices and the both first and the last edge contain at least one vertex from $V-$ the set of vertices included in $E_1 \cup \{e\}$. We shall call this sequence a path. Moreover, due to the minimality of a precycle, it is enough to consider those paths, in which no three consecutive triangles share two vertices. So, for a path of length t there are at most F_{t+1} (the $(t+1)$-th Fibonacci's number) ways to connect each two consecutive triangles together. Thus, the expected number of such paths is at most

$$C\sum_{t=1}^{l}\binom{|V|}{4}F_{t+1}n^{2t-1}(\lambda/N)^t \leq C'n^{-1/5}\sum_{t=1}^{l}F_{t+1}\lambda^t \leq C'n^{-1/5}O(1)^{o(\log n)} \leq n^{-1/6},$$

for some constants C, C' and large enough n. Hence, the probability that such path appears in E is at most $n^{-1/6}$ which bounds also the probability of a precycle. □

After presenting those auxiliary facts we are ready to prove Theorem 2.1.

Proof of Theorem 2.1. Because of the continuity of both $\beta(\lambda)$ and $\beta^+(\lambda)$ for $\lambda > \lambda_0$, it is enough to show that, with high probability, for $\lambda_0 < \lambda < \lambda'$, hypergraph $G_3(n, \lambda'/N)$ contains the k-core of size at least $\beta^+(\lambda)n$. Fix λ and λ' satisfying this inequality and then choose any λ_2, where $\lambda < \lambda_2 < \lambda'$. Thus, for $\lambda_1 \nearrow \lambda_2$,

$$\Psi_{\geq k-1}(\lambda'\beta(\lambda_1)^2) \nearrow \Psi_{\geq k-1}(\lambda'\beta(\lambda_2)^2) > \Psi_{\geq k-1}(\lambda_2\beta(\lambda_2)^2) = \beta(\lambda_2)$$

and we can choose $\lambda < \lambda_1 < \lambda_2$ such that

$$\Psi_{\geq k-1}(\lambda'\beta(\lambda_1)^2) > \beta(\lambda_2). \tag{1}$$

From Lemma 2.3 we conclude, that there exists constant d such that

$$r(\lambda_2, d, \beta(\lambda_2)) \geq \beta(\lambda_1).$$

Moreover, from Lemma 2.2 we get that there exists constant d_1 such that $\mathbb{P}_{\lambda_2}(\mathcal{R}_{d_1} \circ \mathcal{B}) > \beta(\lambda_1)$. Let

$$\eta = \min\{\mathbb{P}_{\lambda_2}(\mathcal{R}_{d_1} \circ \mathcal{B}) - \beta(\lambda_1), (2k)^{-3d}\} > 0.$$

Because $\mathcal{R}_{d_1} \circ \mathcal{B}$ is measurable, there is an integer $L > d_1$ and event $\mathcal{L}_1$ depending only on firs L generations of branching process such that $\mathbb{P}_{\lambda_2}((\mathcal{R}_{d_1} \circ \mathcal{B}) \triangle \mathcal{L}_1) \leq \eta^2/2$. Let $\mathbb{E}_\lambda(\cdot)$ be the expectation corresponding to $\mathbb{P}_\lambda$ and let $X[L]$ denote the branching process X restricted to the first L generations. We have that

$$\eta^2/2 \geq \mathbb{P}_{\lambda_2}((\mathcal{R}_{d_1} \circ \mathcal{B})^c \cap \mathcal{L}_1) = \mathbb{E}_{\lambda_2}(\mathbb{1}_{\mathcal{L}_1} \mathbb{P}_{\lambda_2}((\mathcal{R}_{d_1} \circ \mathcal{B})^c | X[L])).$$

Next, let

$$\mathcal{L} = \mathcal{L}_1 \cap \{\mathbb{P}_{\lambda_2}(\mathcal{R}_{d_1} \circ \mathcal{B} | X[L]) \geq 1 - \eta\}.$$

As $\mathbb{E}_{\lambda_2}(\mathbb{1}_{\mathcal{L}_1} \mathbb{P}_{\lambda_2}((\mathcal{R}_{d_1} \circ \mathcal{B})^c | X[L])) \geq \eta \mathbb{P}_{\lambda_2}(\mathcal{L}_1 \backslash \mathcal{L})$, we have $\mathbb{P}_{\lambda_2}(\mathcal{L}_1 \backslash \mathcal{L}) \leq \eta/2$, hence

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

Moreover, if $\mathcal{L}$ holds, then

$$\mathbb{P}_{\lambda_2}(\mathcal{R}_L \circ \mathcal{B} | X[L]) \geq \mathbb{P}_{\lambda_2}(\mathcal{R}_{d_1} \circ \mathcal{B} | X[L]) \geq 1 - \eta \geq 1 - (2k)^{-3d}. \tag{3}$$

Let us construct the following events recursively. Let $\mathcal{A}_0 = \mathcal{L}$ and for $t \geq 1$, $\mathcal{A}_t = \mathcal{R}_d \circ \mathcal{A}_{t-1}$. It is easy to see that

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

By induction on t, because of (2) and Lemma 2.3, we have that for every t

$$\mathbb{P}_{\lambda_2}(\mathcal{A}_t) \geq \beta(\lambda_1). \tag{4}$$

As $\lambda_2 < \lambda'$ and from (1)

$$\begin{aligned} \mathbb{P}_{\lambda'}(\mathcal{B}_1 \circ \mathcal{A}_t) &= \Psi_{\geq k-1}(\lambda' \mathbb{P}_{\lambda'}(\mathcal{A}_t)^2) \\ &\geq \Psi_{\geq k-1}(\lambda' \mathbb{P}_{\lambda_2}(\mathcal{A}_t)^2) \\ &\geq \Psi_{\geq k-1}(\lambda' \beta(\lambda_1)^2) \geq \beta(\lambda_2). \end{aligned} \tag{5}$$

Following the definition of $\mathbb{P}^p_\lambda(\mathcal{E} \circ M)$, for event $\mathcal{E}$ depending on first d generations, by $\mathbb{P}^p(\mathcal{E} \circ M | X[d])$ we denote the probability that, given the first d generations of branching process X, $\mathcal{E} \circ M$ holds if the particles in generation d are marked independently at random with probability p. Observe, that this probability does not depend on the value of λ.

From (3), if $\mathcal{L}$ holds, then

$$\mathbb{P}^{\beta(\lambda_2)}(\mathcal{R}_L \circ M | X[L]) = \mathbb{P}_{\lambda_2}(\mathcal{R}_L \circ \mathcal{B} | X[L]) \geq 1 - (2k)^{-3d}.$$

Let $\mathcal{E}_t = \mathcal{B}_{dt} \circ \mathcal{R}_L$. In other words, $\mathcal{E}_0 = \mathcal{R}_L$ and for $t > 0$, $\mathcal{E}_t = \mathcal{B}_d \circ \mathcal{E}_{t-1}$.

Claim 1. *Provided $\mathcal{A}_t$ holds,*

$$\mathbb{P}^{\beta(\lambda_2)}(\mathcal{E} \circ M | X[dt+L]) \geq 1 - (2k)^{-(2^t+2)d}.$$

We show this fact by induction on t. It is obvious for $t = 0$. Let $t > 0$ and assume that this inequality holds for $t-1$. Suppose that $\mathcal{A}_t = \mathcal{R}_d \circ \mathcal{A}_{t-1}$ holds and condition on first $dt+L$ generations of X. There exists a set Y of particles in generation d such that $\mathcal{A}_{t-1}$ is satisfied for every $y \in Y$, and $\mathcal{R}_d \circ M$ holds if we mark only elements in Y. Every tree witnessing $\mathcal{R}_d$ contains a subtree also witnessing $\mathcal{R}_d$ in which every particle has at most $2k$ children (k pairs of children). Thus we may assume that $|Y| \leq (2k)^d$. Next, we mark the particles in generation $dt + L$ independently with probability $\beta(\lambda_2)$. Let Y' be the set of particles from Y for which $\mathcal{E}_{t-1}$ holds. Each $y \in Y$ is included in Y' independently, so by the induction hypothesis the probability of such event is at least $1 - (2k)^{-(2^{t-1}+2)d}$. Hence

$$\mathbb{P}(|Y \backslash Y'| \geq 2) \leq \binom{|Y|}{2}((2k)^{-(2^{t-1}+2)d})^2 \leq (2k)^{2d}(2k)^{-(2^t+4)d} = (2k)^{-(2^t+2)d}.$$

If $|Y \backslash Y'| \leq 1$ then $\mathcal{B}_d$ holds if, in generation d, we keep only elements from Y'. Consequently $\mathcal{B}_d \circ \mathcal{E}_{t-1} \circ M = \mathcal{E}_t \circ M$ holds, proving that

$$\mathbb{P}^{\beta(\lambda_2)}(\mathcal{E} \circ M | X[dt+L]) \geq 1 - (2k)^{-(2^t+2)d}.$$

Let $\mathcal{E}$ be an event. By $\mathcal{B}^+_{\leq 2L+1} \circ \mathcal{E}$ we denote the event that there is a set S of *targets* in generations between 1 and $2L+1$ such that $\mathcal{E}$ holds for each target and in the tree consisting of targets and their ancestors every non-leaf has at least $2k$ (k pairs) children, i.e. this tree has property $\mathcal{M}_{\leq 2L+1}$. It follows from this definition that if a particle x is included in this tree, then also the particle that was generated with x must be included. Because we do not require that all elements in generations between 1 and $2L+1$ that have property $\mathcal{E}$ are taken as targets, the event $\mathcal{B}^+_{\leq 2L+1} \circ \mathcal{E}$ is increasing if $\mathcal{E}$ is increasing.

Let $T = T(n) = o(n)$ satisfy $\frac{T}{\log \log n} \to \infty$. Let $\mathcal{A} = \mathcal{B}^+_{\leq 2L+1} \circ \mathcal{A}_T$. As $\mathcal{B}^+_1 \circ \mathcal{A}_T \subset \mathcal{A}$, from (4)

$$\mathbb{P}_{\lambda_2}(\mathcal{A}) \geq \Psi_{\geq k}(\lambda_2 \mathbb{P}_{\lambda_2}(\mathcal{A}_T)^2) \geq \Psi_{\geq k}(\lambda_2 \beta(\lambda_1)^2) \geq \Psi_{\geq k}(\lambda_1 \beta(\lambda_1)^2) = \beta^+(\lambda_1).$$

Let $s = 2L + 1 + dT + L = o(\log n)$, hence $\mathcal{A}$ depends on the first s generations of branching process. It is easy to check that in our branching process $\mathbb{E}_\lambda(|X[t]|^m) = O((2\lambda)^{mt})$. Substituting $t = 2s$, $m = 20$ and using Markov's inequality we have

$$\mathcal{P}_\lambda(|X[2s]| \geq n^{1/10}) = o(n^{-1}). \tag{6}$$

We are now ready to focus on the hypergraph $G_3(n, \lambda'/N)$. Let $\tilde{G} = \tilde{G}(n, 4s, \lambda')$ be the hypergraph obtained from $G = G_3(n, \lambda'/N)$ by conditioning on the absence of cycles of length at most $4s$. Let $c_k(G)$ be the number of vertices in the k-core of hypergraph G. Surely the event $\{c_k(G) \geq x\}$ is increasing, while absence of short cycles is a decreasing event. Thus, by the FKG inequality

$$\mathbb{P}(c_k(\tilde{G}) \geq x) \leq \mathbb{P}(c_k(G_3(n, \lambda'/N) \geq x)),$$

for any x. Hence we shall prove lower bound on the size of the core in $\tilde{G}$ that holds with high probability.

Let v be a random vertex of $\tilde{G}$. We shall uncover hypergraph using BFS algorithm in a following way. We start with v active and every other vertex untested. At every point of algorithm we pick an active vertex w closest to v (we count number of edges between v and w) and test all possible edges consisting of w and untested vertices. After checking all edges we mark w as tested. If such edge would appear we mark the two untested vertices it included as active. By Lemma 2.5, as long as we activate fewer than $2n^{1/5}$ vertices and conditioned on everything that happens before, every edge we test appears with probability $(1 + O(n^{-1/6}))\lambda'/N$. Furthermore we never attempt to test an edge that would close the cycle and because the number of untested vertices is $n - O(n^{-1/5})$ we are able, with high probability, to couple the number of new edges, that include w, with a Poisson distribution with mean λ'.

Let $G_v[t]$ be the subgraph of $\tilde{G}$ formed by the vertices within distance t from v. Obviously, as long as $t \leq 2s$, this subgraph is a tree. As the number of elements in $X_{\lambda'}[2s]$ is at most $n^{1/10}$ and because of the coupling above, we may couple, with high probability, $G_v[2s]$ with $X_{\lambda'}[2s]$. We will say that v has the property $\mathcal{A}$ if $G_v[2s]$, viewed as a branching process rooted at v, has this property. Hence, v has property $\mathcal{A}$ with probability $\mathbb{P}_{\lambda'}(\mathcal{A}) + o(1)$.

Claim 2. *With high probability the number of vertices with property $\mathcal{A}$ in $\tilde{G}$ is at least $\beta^+(\lambda)n$.*

We show the claim following the argument from [5]. Let A be this number and let v and w be randomly chosen vertices of $\tilde{G}$. The probability that v and w are within distance $2s$ is $o(1)$. Thus we may couple $G_v[2s]$ and $G_w[2s]$ with independent copies of $X_{\lambda'}[2s]$. Hence, the probability that both v and w have property $\mathcal{A}$ is $\mathbb{P}_{\lambda'}(\mathcal{A})^2 + o(1)$. Thus

$$\begin{aligned}\mathbb{E}(A) &= \mathbb{P}_{\lambda'}(\mathcal{A})n + o(n),\\ \mathbb{E}(A^2) &= \mathbb{P}_{\lambda'}(\mathcal{A})^2n^2 + o(n^2).\end{aligned}$$

By Chebyshev's inequality, A/n converges in probability to

$$\mathbb{P}_{\lambda'}(\mathcal{A}) \geq \mathbb{P}_{\lambda_2}(\mathcal{A}) \geq \beta^+(\lambda_1) \geq \beta^+(\lambda).$$

To finish the proof of the lower bound on the size of the k-core, we shall show the following fact.

Claim 3. *With high probability every vertex with property $\mathcal{A}$ belongs to a k-core.*

Let v be a random vertex of $\tilde{G}$. Because of the coupling above, from now on, we shall regard the hypergraph $G_v[2s]$ as the corresponding branching process. This allows us to speak about the children and descendants of a particular vertex. For a vertex w in generation t we denote by $D(w)$ the set of its descendants in generation $t+(dT+L)$. Let us check whether v has property $\mathcal{A}$ using following algorithm. Let S_1 and S_2 be initially empty sets of vertices. At each step of the algorithm every vertex in S_1 will have property $\mathcal{A}_T$.

First, let us observe descendants of v up to $dT+L+1$ generation. Then, for every child w of v let us examine its descendants up to relative generation $dT+L$. If such w has property $\mathcal{A}_T$ put it into S_1 and vertices from $D(w)$ into S_2. If at any point in S_1 there are at least k pairs of vertices (w_1, w_2), (w_3, w_4), $\dots$ and (w_{k-1}, w_k), such that $(w, w_1, w_2), (w, w_3, w_4), \dots,$ and (w, w_{k-1}, w_k) are edges in $\tilde{G}$, we stop (as if so, vertex v has property $\mathcal{A}$). Otherwise, we check each w in generation 2, that is not a descendant of a vertex from S_1. For each such vertex reveal its descendants up to relative generation $dT+L$ and test whether it has property $\mathcal{A}_T$. If so, we add it to S_1 and $D(w)$ to S_2. We continue this process, taking testing descendants of v from further generation until one of the following occurs

(1) the set S_1 shows that v has property $\mathcal{A}$,
(2) we have checked all vertices in generations up to $2L+1$, that were not descendants of vertices from S_1, for having the property $\mathcal{A}_T$ and have not found the set S_1 that would show that v has property $\mathcal{A}$.

If the second case occurs, vertex v does not have property $\mathcal{A}$. Hence, v has $\mathcal{A}$ if and only if the first case holds.

Let us condition on v having the property $\mathcal{A}$. Let us suppose that we examined at most $n^{1/10}$ vertices (which happens with probability $1-o(n^{-1})$ and let us check the vertices from S_2 (observe, that their children were checked in the previous stage). For each $w \in S_2$ we reveal its descendants up to $1+dT+L$ generation to test whether it has the property $\mathcal{B}_1 \circ \mathcal{A}_T$. If w has this property, we mark it. Observe, that we reach more than $n^{1/10}$ vertices in this process with probability $o(n^{-1})$ (vertices in $1+dT+L$ generation from w lay in at most $dT+L+2L+1+dT+L+1 \le 2s$ generation from v). Then, if we have not reached $n^{1/10}$ vertices yet, with probability $1-o(1)$, we can couple the descendants of w with a branching process $X_{\lambda'}$. Hence, for large enough n, the probability that we mark w is at least

$$\mathbb{P}_{\lambda'} \circ \mathcal{A}_T - o(1) \ge \beta(\lambda_2),$$

where the last inequality follows from (5). Putting these together and omitting the error term of $1-o(n^{-1})$, we can consider each $w \in S_2$ marked independently with probability at least $\beta(\lambda_2)$.

Let $\mathcal{T}_v$ be the tree rooted in v with leaves the marked vertices in S_2. For large enough n, $(2k)^{-2^T+2}d \leq n^{-1}$ and, by Claim 1, with probability $1 - o(n^{-1})$ every vertex in S_1 has property $\mathcal{E}_T$ even if we restrict its descendants to $\mathcal{T}_v$. Thus, with probability $1 - o(n^{-1})$ the tree $\mathcal{T}_v$ has property $\mathcal{A}' = \mathcal{B}^+_{\leq 2L+1} \circ \mathcal{E}_T$. Let $\mathcal{T}'_v$ be the minimal subtree with this property. Observe that each non-leaf in $\mathcal{T}'_v$ has at least $2k$ (k pairs) of children.

We shall show that every leaf in $\mathcal{T}'_v$ has property $\mathcal{A}$ in the hypergraph $\tilde{G}$. Let w be the leaf and let x be its ancestor in $\mathcal{T}'_v$, that is L generations above w. Moreover let $\mathcal{T}'$ be the subtree of $\mathcal{T}'_v$ of height L, rooted at x and minimal with respect to having property $\mathcal{R}_L$. Observe, that w was added to S_2 as the descendant of v in generation $t + dT + L$ for some t. Thus all leaves of $\mathcal{T}'$ are also leaves $\mathcal{T}'_v$. So they have property $\mathbb{B}_1 \circ \mathcal{A}_T$ in the tree $G_v[2s]$ rooted in v. Let $\mathcal{T}$ be formed from $\mathcal{T}'$ by adding $k-1$ pairs of children to each leaf, such that $\mathcal{T}$ has property $\mathcal{R}_L \circ \mathcal{B}_1$ and all of its leaves have property $\mathcal{A}_T$. By Lemma 2.4, there is a subtree $\mathcal{W}$ of $\mathcal{T}$ rooted at w, with property $\mathcal{M}_{\leq 2L+1}$. All leaves of $\mathcal{W}$ are leaves of $\mathcal{T}$ so they have property $\mathcal{A}_T$. Thus w has property $\mathcal{A}$ in the hypergraph.

So far, we have shown that with probability $1 - n \cdot o(n^{-1}) = 1 - o(1)$ every vertex v with property $\mathcal{A}$ is the root of a tree $\mathcal{T}'_v$ in which every non-leaf belongs to at least k edges and each leaf has also property $\mathcal{A}$ in hypergraph $\tilde{G}$. Taking the union of those trees we obtain a subgraph of $\tilde{G}$ in which every vertex belong to at least k edges. Moreover, this subgraph contain all vertices with property $\mathcal{A}$, so $c_k(\tilde{G}) \geq A$, which completes the proof of lower bound on the size of a k-core of our original hypergraph.

It remains to show the upper bound on $c_k(G_3(n, \lambda/N))$.

Claim 4. *For any $\varepsilon > 0$,*

$$c_k(G_3(n, \lambda/N)) \leq \mathbb{P}_\lambda(\mathcal{B}^+_d)n + o(n) \leq \beta^+(\lambda)n + \epsilon n + o(n).$$

Let $\varepsilon > 0$ and v be a random vertex of $G_3(n, \lambda/N)$. Let us explore the first d neighbourhoods of v, where d is chosen such that $\mathbb{P}_\lambda(\mathcal{B}^+_d) \leq \beta^+(\lambda) + \varepsilon$. The probability that we encounter a cycle is $o(1)$. We say that v has property $\mathcal{B}^+_d$ if its first d neighbourhoods form a tree with property $\mathcal{B}^+_d$. Observe, that if v belongs to k-core then it must have one of the following properties

(1) there is a cycle in the first d neighbourhoods of v,
(2) v has the property $\mathcal{B}^+_d$.

It is easy to check that the number of vertices having the first property converges in probability to $o(n)$, where the number of those with the second one converges to $\mathbb{P}_\lambda(\mathcal{B}^+_d)n + o(n)$. Thus,

$$c_k(G_3(n, \lambda/N)) \leq \mathbb{P}_\lambda(\mathcal{B}^+_d)n + o(n) \leq \beta^+(\lambda)n + \epsilon n + o(n).$$

Since ε can be arbitrarily small positive constant, the assertion follows. □

References

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

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

[3] T. Łuczak, *Size and connectivity of the k-core of a random graph*, Discrete Mathematics **91** (2003), 61–68.

[4] M. Molloy, *Cores in random hypergraphs and boolean formulas*, Random Structures and Algorithms **27** (2005), 124–135.

[5] O. Riordan, *The $k-$core and branching processes*, Combinatorics, Probability & Computing (2007), 111–136.

[6] B. Pittel, J. Spencer, N. Wormald, *Sudden emergence of a giant k-core in a random graph*, J. Comb. Theory Ser. B **67** (2003), 111–151.

Faculty of Mathematics and Computer Science, Adam Mickiewicz University, ul. Umultowska 87, 61-614 Poznań, Poland