ON THE RANDOM GREEDY F-FREE HYPERGRAPH PROCESS

DANIELA KÜHN, DERYK OSTHUS AND AMELIA TAYLOR

Abstract. Let F be a strictly k-balanced k-uniform hypergraph with $e(F) \geq |F| - k + 1$ and maximum co-degree at least two. The random greedy F-free process constructs a maximal F-free hypergraph as follows. Consider a random ordering of the hyperedges of the complete k-uniform hypergraph K_n^k on n vertices. Start with the empty hypergraph on n vertices. Successively consider the hyperedges e of K_n^k in the given ordering and add e to the existing hypergraph provided that e does not create a copy of F. We show that asymptotically almost surely this process terminates at a hypergraph with $\tilde{O}(n^{k-(|F|-k)/(e(F)-1)})$ hyperedges. This is best possible up to logarithmic factors.

1. Introduction

1.1. **Results.** Fix a k-uniform hypergraph F. In this paper, we study the following random greedy process, which constructs a maximal F-free k-uniform hypergraph. Assign a birthtime which is uniformly distributed in $[0, 1]$ to each hyperedge of the complete k-uniform hypergraph K_n^k on n vertices. Start with the empty hypergraph on n vertices at time $p = 0$. Increase p and each time that a new hyperedge is born, add it to the hypergraph provided that it does not create a copy of F (edges with equal birthtime are added in any order). Denote the resulting hypergraph at time p by $R_{n,p}$.

The random greedy graph process (i.e. the case when $k = 2$) has been studied for many graphs. The initial motivation (see for example [8]) was to study the Ramsey number $R(3, t)$. Indeed, the best current lower bounds on $R(3, t)$ were obtained via the study of the triangle-free process ([5], [10]). Osthus and Taraz [11] gave an upper bound on the number of edges in the graph $R_{n,1}$ when F is strictly 2-balanced, showing that a.a.s. $R_{n,1}$ has maximum degree $O(n^{1-(|F|-2)/(e(F)-1)}(\log n)^{1/(\Delta(F)-1)})$. (Here a.a.s. stands for 'asymptotically almost surely', i.e. for the property that an event occurs with probability tending to one as n tends to infinity.) Results for the cases when $F = C_4$ and $F = K_4$ were obtained independently by Bollobás and Riordan [7]. Bohman and Keevash [4] showed that a.a.s. $R_{n,1}$ has minimum degree $\Omega(n^{1-(|F|-2)/(e(F)-1)}(\log n)^{1/(e(F)-1)})$ whenever F is strictly 2-balanced and conjectured that this gives the correct order of magnitude. Improved upper bounds have been obtained for some graphs. For instance, the number of edges has been determined asymptotically when F is a cycle ([3], [5], [10], [12], [14]) and when $F = K_4$ ([15], [16]). Picollelli [13] determined asymptotically the number of edges when F is a diamond, i.e. the graph obtained by removing one edge from K_4. Note that this graph is not strictly 2-balanced.

Much less is known about the process when F is a k-uniform hypergraph and $k \geq 3$. The only known upper bound is due to Bohman, Mubayi and Picollelli [6], who studied the F-free process when F is a k-uniform generalisation of a graph triangle (with an application to certain

Date: November 27, 2015.

The research leading to these results was partially supported by the European Research Council under the European Union's Seventh Framework Programme (FP/2007–2013) / ERC Grant Agreements no. 258345 (D. Kühn) and 306349 (D. Osthus).

Ramsey numbers). In this paper, we obtain a generalisation of the upper bound in [11] to strictly k-balanced hypergraphs. Here we say that a k-uniform hypergraph F is *strictly k-balanced* if $|F| \geq k+1$ and for all proper subgraphs $F' \subsetneq F$ with $|F'| \geq k+1$ we have

$$\frac{e(F)-1}{|F|-k} > \frac{e(F')-1}{|F'|-k}.$$

We also need the following definition. Given a hypergraph H and $i \in \mathbb{N}$, we define the *maximum i-degree* of H by

$$\Delta_i(H) := \max\{d_H(U) : U \subseteq V(H), |U| = i\},$$

where $d_H(U)$ is the number of hyperedges in H containing U. For any k-uniform hypergraph, the *maximum co-degree* refers to the maximum $(k-1)$-degree.

Theorem 1.1. *Let $k \in \mathbb{N}$ be such that $k \geq 2$. Let F be a strictly k-balanced k-uniform hypergraph which has v vertices and $h \geq v-k+1$ hyperedges. Suppose $\Delta_{k-1}(F) \geq 2$. Then there exists a constant c such that a.a.s.*

$$\Delta_{k-1}(R_{n,1}) < t \qquad \text{where} \qquad t := cn^{1-\frac{v-k}{h-1}}(\log n)^{\frac{3}{\Delta_{k-1}(F)-1}-\frac{1}{h-1}}. \tag{1}$$

In particular, a.a.s. $R_{n,1}$ has at most tn^{k-1} hyperedges.

Note that Theorem 1.1 applies, for example, to all k-uniform cliques K_v^k on $v \geq k+1$ vertices and more generally to all balanced complete ℓ-partite k-uniform hypergraphs with $\ell \geq k$ and more than k vertices.

Bennett and Bohman [2] studied a random greedy independent set algorithm in certain quasi-random hypergraphs. This algorithm finds a maximal independent set by choosing vertices uniformly at random and adding them to the existing set provided they do not create a hyperedge. Note that we can define an $e(F)$-regular hypergraph H whose set of vertices is $E(K_n^k)$ and whose hyperedges correspond to all copies of F in K_n^k. In this case, the random greedy independent set process on H is exactly the F-free process. Their result can be applied in the context of the F-free process to show that if F is a strictly k-balanced k-uniform hypergraph and every vertex of F lies in at least two hyperedges, then a.a.s. $R_{n,1}$ has $\Omega(n^{k-(|F|-k)/(e(F)-1)}(\log n)^{1/(e(F)-1)})$ hyperedges. Up to logarithmic factors, this matches the upper bound given in Theorem 1.1.

1.2. **Open questions.** There are many natural open questions related to the random greedy F-free process. First, we discuss bounds on the number of edges in $R_{n,1}$ when F is an ℓ-cycle. Theorem 1.1 applies in the case when F is a k-uniform tight cycle. However, there are other natural notions of a hypergraph cycle: Given $\ell \in \mathbb{N}$ with $\ell < k$, we say that a k-uniform hypergraph $C_{\ell,h}$ is an *ℓ-cycle* of length h if there is a cyclic ordering of its vertices $x_1, \dots, x_{h(k-\ell)}$ and a corresponding ordering on its hyperedges $e_0, \dots, e_{h-1}$ such that $e_i = \{x_{i(k-\ell)+1}, \dots, x_{i(k-\ell)+k}\}$. So consecutive hyperedges on the cycle intersect in exactly ℓ vertices. The case when $\ell = k-1$ corresponds to $C_{\ell,h}$ being a tight cycle of length h. It is easy to check that all ℓ-cycles are strictly k-balanced, but only tight cycles satisfy the co-degree condition in Theorem 1.1. In the case when $\ell \geq k/2$, ℓ-cycles meet the conditions in [2]. We conjecture that the bound on the number of hyperedges in [2] is of the correct magnitude for any ℓ.

Conjecture 1.2. *Let $\ell, k \in \mathbb{N}$ be such that $k \geq 2$ and $k > \ell$ and let $F := C_{\ell,h}$ be the ℓ-cycle of length h. Then a.a.s. $R_{n,1}$ has $\Theta(n^{\frac{h\ell}{h-1}}(\log n)^{\frac{1}{h-1}})$ hyperedges.*

One motivation for Conjecture 1.2 is that $p = n^{h\ell/(h-1)-k}(\log n)^{1/(h-1)}$ is the threshold for the property that every hyperedge in $H_{n,p}$ lies in an ℓ-cycle of length h.

Another open problem would be to generalise Theorem 1.1 by finding an upper bound on the number of steps in the random greedy independent set process studied in [2].

The random greedy independent set process can also be applied to study arithmetic progression free sets. Suppose $k, n \in \mathbb{N}$. The kAP-free process generates a subset I of $\mathbb{Z}_n$ which does not contain an arithmetic progression of length k as follows. The elements of $\mathbb{Z}_n$ are ordered uniformly at random. Each is then, in turn, added to the set I if it does not create a k term arithmetic progression. So this is another instance of the random greedy independent set algorithm, this time on the hypergraph with vertex set $\mathbb{Z}_n$ whose hyperedges are all arithmetic progressions of length k. When n is prime, Bennett and Bohman [2] showed that a.a.s. the kAP-free process generates a kAP-free set I of size $\Omega(n^{(k-2)/(k-1)}(\log n)^{1/(k-1)})$. It would be interesting to obtain a corresponding upper bound on I. (Note that an upper bound on the number of steps in the random greedy independent set process would imply an upper bound for the kAP-free process.)

1.3. **Sketch of the argument.** Rather than studying the random greedy process itself, we are able to prove Theorem 1.1 by obtaining precise information about the random binomial hypergraph $H_{n,p}$. (This idea was first used in [11].) More precisely, write $H_{n,p}$ for the random binomial k-uniform hypergraph on n vertices with hyperedge probability p, i.e., each hyperedge is included in $H_{n,p}$ with probability p, independently of all other hyperedges. We write $H^-_{n,p}$ for the hypergraph formed by removing all copies of F from $H_{n,p}$. Note that $H_{n,p}$ can also be viewed as the random hypergraph consisting of all hyperedges with birthtime at most p. Thus, for all $p \in [0, 1]$ we have

$$H^-_{n,p} \subseteq R_{n,p} \subseteq R_{n,1}.$$

We will always assume that K^k_n, $H_{n,p}$, $H^-_{n,p}$ and $R_{n,p}$ use the vertex set $[n]$.

In Section 2, we collect some large deviation inequalities. The proof of Theorem 1.1 is given in Section 3, the strategy is as follows. We first identify the largest point p where we can still use $H_{n,p}$ to approximate the behaviour of $H^-_{n,p}$ (i.e. for this p, only a small proportion of edges of $H_{n,p}$ lie in a copy of F). Now let U be a set of $k-1$ vertices in F such that $d_F(U) = \Delta_{k-1}(F)$. Let $\hat{F}$ be the subgraph of F obtained by deleting all those hyperedges which contain U. Let t be as in (1). Suppose for a contradiction that there exists a $(k-1)$-set V of degree t in $R_{n,1}$ and let T be the neighbourhood of V in $R_{n,1}$. We will show that in this case we would almost certainly find a copy α of $\hat{F}$ in $H^-_{n,p}[T \cup V]$ which maps U to V. Since $H^-_{n,p} \subseteq R_{n,1}$, α would also be a copy of $\hat{F}$ in $R_{n,1}[T \cup V]$ which maps U to V. But this actually yields a copy of F in $R_{n,1}$, a contradiction. So a.a.s. $\Delta_{k-1}(R_{n,1}) < t$. It is perhaps surprising that for our analysis the order of hyperedges added after this critical point p is irrelevant.

2. Tools

Let $\mathcal{S}$ be a collection of subsets of $E(K^k_n)$. For each $\alpha \in \mathcal{S}$, let I_α denote the indicator variable which equals one if all hyperedges in α lie in $H_{n,p}$ and zero otherwise. Set

$$X := \sum_{\alpha \in \mathcal{S}} I_\alpha \qquad \text{and} \qquad \mu := \mathbb{E}[X].$$

Let Y be the size of a largest hyperedge-disjoint collection of elements of $\mathcal{S}$ in $H_{n,p}$ (i.e. the maximum size of a set $\mathcal{S}' \subseteq \mathcal{S}$ such that $I_\alpha = 1$ for all $\alpha \in \mathcal{S}'$ and $\alpha \cap \alpha' = \emptyset$ for all distinct $\alpha, \alpha' \in \mathcal{S}'$). Erdős and Tetali [9] proved the following upper tail bound on Y.

Theorem 2.1. [9]. *For every $a \in \mathbb{N}$, we have $\mathbb{P}[Y \geq a] \leq (e\mu/a)^a$.*

We also require a lower tail bound on Y. For all $\alpha, \alpha' \in \mathcal{S}$ with $\alpha \neq \alpha'$, we write $\alpha \sim \alpha'$ if $\alpha \cap \alpha' \neq \emptyset$. Define

$$\Delta := \sum_{\alpha' \sim \alpha} \mathbb{E}[I_\alpha I_{\alpha'}],$$

where the sum is over all ordered pairs $\alpha' \sim \alpha$ in $\mathcal{S}$. Also, let

$$\eta := \max_{\alpha \in \mathcal{S}} \mathbb{E}[I_\alpha] \qquad \text{and} \qquad \nu := \max_{\alpha \in \mathcal{S}} \sum_{\alpha' \in \mathcal{S}: \alpha' \sim \alpha} \mathbb{E}[I_{\alpha'}].$$

The following bound follows from Lemma 4.2 in Chapter 8 and Theorem A.15 in [1], see [11].

Theorem 2.2. *Let $\varepsilon > 0$. Then* $\mathbb{P}[Y \leq (1-\varepsilon)\mu] \leq e^{(1-\varepsilon)\mu\nu + \frac{\Delta}{2(1-\eta)} - \frac{\varepsilon^2 \mu}{2}}$.

3. Proof of Theorem 1.1

3.1. **Basic parameters.** Let F be a strictly k-balanced k-uniform hypergraph which has v vertices, h hyperedges and $d := \Delta_{k-1}(F) \geq 2$. Choose positive constants c_1, c_2 satisfying

$$1/c_1 \ll 1/c_2 \ll 1/v, 1/h.$$

(Here the notation $a \ll b$ means that we can find an increasing function f for which all of the conditions in the proof are satisfied whenever $a \leq f(b)$.) Given functions f and g, we will write $f = \tilde{O}(g)$ if there exists a constant c such that $f(n) \leq (\log n)^c g(n)$ for all sufficiently large n.

Set

$$p := \frac{1}{c_2 (n^{v-k} \log n)^{1/(h-1)}} \qquad \text{and} \qquad t := c_1 np (\log n)^{3/(d-1)}.$$

Here p is chosen to be as large as possible subject to the constraint that a.a.s. only a small proportion of the hyperedges of $H_{n,p}$ lie in a copy of F. For each $k+1 \leq i \leq v$, we define

$$h_i := \max\{e(F') : F' \subsetneq F, |F'| = i\}.$$

Since F is strictly k-balanced, we have

$$\frac{h-1}{v-k} > \frac{h_i - 1}{i-k}.$$

So for each $k+1 \leq i \leq v$ we can define a positive constant

$$\delta_i := i - k - (h_i - 1)\frac{v-k}{h-1} > 0. \tag{2}$$

Let

$$\delta := \min\{\delta_i : k+1 \leq i \leq v\}.$$

We will often use that for $k+1 \leq i \leq v$

$$n^{v-i} p^{h-h_i} \leq n^{v-i-\frac{v-k}{h-1}(h-h_i)} \stackrel{(2)}{=} n^{v-i-\frac{v-k}{h-1}(h-1-\frac{i-k-\delta_i}{v-k}(h-1))} = n^{-\delta_i} \leq n^{-\delta}. \tag{3}$$

Note that this bounds the expected number of extensions of a fixed subgraph of F on i vertices into copies of F in $H_{n,p}$.

3.2. **Many copies of F containing a fixed hyperedge.** For a given hyperedge $f \in E(K_n^k)$, an (r, f)-*cluster* is a collection $F_1, F_2, \ldots, F_r$ of r copies of F such that each F_i contains f and for each $1 < i \leq r$, there exists $f_i \in E(F_i)$ such that $f_i \notin E(F_j)$ for any $j < i$. Define $\mathcal{A}$ to be the event that $H_{n,p}$ has no $(\log n, f)$-cluster for any hyperedge f. We will bound the probability of $\mathcal{A}^c$, i.e., the probability that $H_{n,p}$ has a $(\log n, f)$-cluster for some hyperedge f.

Lemma 3.1. *We have* $\mathbb{P}[\mathcal{A}^c] \leq n^{-k}$.

Proof. Fix some $f \in E(K_n^k)$. Write $Z_{r,f}$ for the number of (r, f)-clusters in $H_{n,p}$, so $Z_{1,f}$ counts copies of F which contain the hyperedge f. There are h hyperedges in F which could be mapped to f, so

$$\mathbb{E}[Z_{1,f}] \leq hn^{v-k}p^h \leq e^{-2k}$$

with room to spare. Let $r < \log n$ and consider a fixed (r, f)-cluster C in $H_{n,p}$. Let Z_C be the number of $(1, f)$-clusters in $H_{n,p}$ which contain at least one hyperedge which does not lie in C, so each of these $(1, f)$-clusters together with C forms an $(r + 1, f)$-cluster. Suppose that α is a $(1, f)$-cluster sharing $k + 1 \leq i \leq v$ vertices with C. The set of hyperedges shared by α and C form a proper subgraph of F on i vertices. Since F is strictly k-balanced, α and C can have at most h_i common hyperedges. This allows us to estimate $\mathbb{E}[Z_C]$ as

$$\mathbb{E}[Z_C] \leq hn^{v-k}p^{h-1} + \sum_{i=k+1}^{v} v^i (rv)^{i-k} n^{v-i} p^{h-h_i} \overset{(3)}{\leq} e^{-3k} + \tilde{O}(n^{-\delta}) \leq e^{-2k}.$$

If we sum over all (r, f)-clusters in K_n^k, we find that

$$\mathbb{E}[Z_{r+1,f}] \leq \mathbb{E}[Z_{r,f}]e^{-2k} \leq e^{-2(r+1)k}$$

and hence $\mathbb{E}[Z_{\log n,f}] \leq n^{-2k}$. By summing over all $f \in E(K_n^k)$, we obtain

$$\mathbb{P}[\mathcal{A}^c] \leq \binom{n}{k} n^{-2k} \leq n^{-k},$$

as required. □

3.3. **Estimating the number of extensions of a fixed set.** Recall that $d = \Delta_{k-1}(F)$. Let $U = \{u_1, u_2, \ldots, u_{k-1}\} \subseteq V(F)$ be such that $d_F(U) = d$. Let $N_F(U)$ denote the neighbourhood of U in F, i.e. $N_F(U) := \{x \in V(F) : U \cup \{x\} \in E(F)\}$. Define $\hat{F} \subseteq F$ which has vertex set $V(F)$ and all hyperedges $f \in E(F)$ such that $|f \cap U| \leq k-2$. Fix $T \subseteq [n]$ of size t and an ordered sequence $V = (v_1, v_2, \ldots, v_{k-1})$ of distinct vertices, where $v_i \in [n] \setminus T$ for each $1 \leq i \leq k - 1$. Given a hypergraph $H \subseteq K_n^k$, let $\mathcal{S}(H) = \mathcal{S}(H, T, V)$ be the set of all copies of $\hat{F}$ in H such that the following hold:

- for each $1 \leq i \leq k - 1$, u_i is mapped to v_i;
- $N_F(U)$ is mapped into T and
- $V(F) \setminus N_F(U)$ is mapped into $[n] \setminus T$.

We let $X := |\mathcal{S}(H_{n,p})|$ and $X^- := |\mathcal{S}(H_{n,p}^-)|$. Note that $X^- \leq X$ since $H_{n,p}^- \subseteq H_{n,p}$.

Note that if $T \subseteq N_{R_{n,1}}(V)$, then $\mathcal{S}(R_{n,1}) = \emptyset$, as otherwise we could find a copy of F in $R_{n,1}$. Since $H_{n,p}^- \subseteq R_{n,1}$, it follows that $X^- = 0$. So, in order to prove Theorem 1.1, it will suffice to prove that a.a.s. we have $X^- > 0$ for any choice of T, V.

Lemma 3.2. *Given $T \subseteq [n]$ of size t and an ordered sequence $V = (v_1, v_2, \dots, v_{k-1})$ of distinct vertices, where $v_i \in [n] \setminus T$ for each $1 \le i \le k-1$, define X^- as above. Then*

$$\mathbb{P}[(X^- = 0) \cap \mathcal{A}] \le 2n^{-2t}.$$

Proof. Write $\mathcal{S} := \mathcal{S}(K_n^k)$. Note that

$$\begin{aligned} \mu_1 := \mathbb{E}[X] &\ge \binom{t}{d}\binom{n-t-k+1}{v-d-k+1}p^{h-d} \ge \frac{tt^{d-1}n^{v-d-k+1}p^{h-d}}{d^d v^v} \\ &= \frac{tc_1^{d-1}n^{v-k}p^{h-1}(\log n)^3}{d^d v^v} = \frac{c_1^{d-1}}{d^d v^v c_2^{h-1}} t(\log n)^2 \ge 24h^2 t(\log n)^2. \end{aligned} \tag{4}$$

Let $\mathcal{S}'(H_{n,p})$ be a hyperedge-disjoint collection of elements of $\mathcal{S}(H_{n,p})$ of maximum size and let $Y_1 := |\mathcal{S}'(H_{n,p})|$. In order to apply Theorem 2.2, we will estimate ν, Δ and η.

First we estimate ν. Define

$$\nu^* := \max_{\alpha \in \mathcal{S}} \sum_{\alpha' \in \mathcal{S} : \alpha' \sim \alpha} \mathbb{E}[I_{\alpha'} \mid I_\alpha = 1]$$

and note that $\nu \le \nu^*$. We count the expected number of elements $\alpha' \in \mathcal{S}(H_{n,p}) \setminus \{\alpha\}$ sharing at least one hyperedge with some fixed element $\alpha \in \mathcal{S}$. Note that α and α' must share at least two vertices outside V by the definition of $\hat{F}$. We let $k+1 \le i+j \le v$ denote the number of shared vertices, where i is the number of vertices shared in T. Consider any $\alpha' \in \mathcal{S} \setminus \{\alpha\}$ sharing $i+j$ vertices with α. Let K be the hypergraph on $i+j$ vertices formed by the set of hyperedges shared by α and α'. Let K' be the hypergraph on $i+j$ vertices obtained from K by adding all hyperedges of the form $V \cup x$ for each of the i vertices $x \in T$ shared by α and α'. Since $K' \subsetneq F$, $e(K') \le h_{i+j}$ and so α and α' can have at most $h_{i+j} - i$ common hyperedges. Then

$$\begin{aligned} \nu \le \nu^* &\le \sum_{i+j=k+1}^{v} v^{i+j} t^{d-i} n^{v-d-j} p^{h-d-(h_{i+j}-i)} \\ &= \sum_{i+j=k+1}^{v} v^{i+j} (c_1(\log n)^{\frac{3}{d-1}})^{d-i} n^{v-(i+j)} p^{h-h_{i+j}} \overset{(3)}{=} \tilde{O}(n^{-\delta}) = o(1). \end{aligned}$$

Since Δ counts the expected number of ordered pairs of elements in $\mathcal{S}(H_{n,p})$ which share at least one hyperedge, we have

$$\Delta \le \mu_1 \nu^* = o(\mu_1).$$

Finally, the probability of a fixed element in $\mathcal{S}$ being present in $H_{n,p}$ is given by

$$\eta = p^{h-d} = o(1).$$

So we can apply Theorem 2.2 to see that

$$\mathbb{P}[Y_1 \le \mu_1/2] \le e^{-\mu_1/10} \overset{(4)}{\le} n^{-2t}. \tag{5}$$

We define a *cluster* (α, F') to be the union of an element $\alpha \in \mathcal{S}'(H_{n,p})$ and a copy F' of F in $H_{n,p}$ which share at least one hyperedge. Note that deleting F' from $H_{n,p}$ to form $H_{n,p}^-$ will destroy α.

We define an auxiliary graph G as follows. For each element of $\mathcal{S}'(H_{n,p})$ which lies in a cluster, choose one. These clusters form the vertices of G. Draw an edge between two vertices in G if the corresponding clusters share a hyperedge. We will use that

$$|G| \le (\Delta(G) + 1)\alpha(G) \tag{6}$$

(which holds for all graphs) to bound the number of vertices in G. We will show that with sufficiently high probability $|G| < Y_1$. (This in turn implies that at least one element of $\mathcal{S}'(H_{n,p})$ will remain in $H^-_{n,p}$, i.e. $X^- > 0$.)

First, we bound $\alpha(G)$. Let X_2 be the number of clusters in $H_{n,p}$. We estimate $\mu_2 := \mathbb{E}[X_2]$, breaking the sum into parts depending on the number i of vertices shared by α and F' in each cluster (α, F'). For $k+1 \le i \le v$, we use that α and F intersect in a proper subgraph of F and thus can have at most h_i common hyperedges. The first term in our bound on μ_2 corresponds to those clusters (α, F') where α and F' share exactly one hyperedge:

$$\mu_2 = \mathbb{E}[X_2] \le \mu_1 h^2 n^{v-k} p^{h-1} + \sum_{i=k+1}^{v} \mu_1 v^i n^{v-i} p^{h-h_i}$$

$$\overset{(3)}{\le} \mu_1 h^2 n^{v-k} p^{h-1} + O(\mu_1 n^{-\delta}) \le \mu_1/(12e^2h^2\log n). \tag{7}$$

Let Y_2 be the size of a largest hyperedge-disjoint collection of clusters in $H_{n,p}$. We note that $\alpha(G) \le Y_2$ and use Theorem 2.1 to bound Y_2:

$$\mathbb{P}\left[\alpha(G) \ge \mu_1/(12h^2\log n)\right] \le \mathbb{P}\left[Y_2 \ge \mu_1/(12h^2\log n)\right] \le \left(\frac{e\mu_2 12h^2\log n}{\mu_1}\right)^{\mu_1/(12h^2\log n)}$$

$$\overset{(7)}{\le} e^{-\mu_1/(12h^2\log n)} \overset{(4)}{\le} n^{-2t}. \tag{8}$$

We now bound $\Delta(G)$. Assume that $\mathcal{A}$ holds, that is, $H_{n,p}$ does not contain a $(\log n, f)$-cluster for any hyperedge f. Fix some hyperedge $f \in E(H_{n,p})$. Let $\mathcal{F}$ be a collection of clusters (α_i, F_i) such that $f \in E((\alpha_i, F_i))$ for each i and $\alpha_i \ne \alpha_j$ if $i \ne j$. Suppose that $|\mathcal{F}| \ge h\log n+1$. For each cluster (α_i, F_i) in $\mathcal{F}$, let e_i be a hyperedge shared by α_i and F_i. The α_i are hyperedge-disjoint by the definition of $\mathcal{S}'(H_{n,p})$, so $f \in E(F_i)$ for all but at most one cluster $(\alpha_i, F_i) \in \mathcal{F}$ where $f \in E(\alpha_i)$. If $\mathcal{F}$ contains such a cluster, delete it. Then, starting with $i = 1$, if (α_i, F_i) has not already been deleted, delete from $\mathcal{F}$ any clusters (α_j, F_j) with $j > i$ such that e_j lies in (α_i, F_i). Do this for each i in turn. Since the α_i are hyperedge-disjoint, at each step we delete at most $h-1$ clusters from $\mathcal{F}$. So a collection $\mathcal{F}' \subseteq \mathcal{F}$ of at least $\log n$ clusters remains. But the set of all F_i such that $(\alpha_i, F_i) \in \mathcal{F}'$ contains a $(\log n, f)$-cluster in $H_{n,p}$. Therefore, $|\mathcal{F}| < h\log n + 1$. Since every cluster has less than $2h$ hyperedges, we must have

$$\Delta(G) < 2h^2\log n. \tag{9}$$

So, if $\mathcal{A}$ holds, if $\alpha(G) < \mu_1/(12h^2\log n)$ and if $|Y_1| \ge \mu_1/2$, then

$$|G| \overset{(6),(9)}{\le} (2h^2\log n + 1)\mu_1/(12h^2\log n) \le \mu_1/4 < |Y_1|.$$

Thus,

$$\mathbb{P}[(X^- = 0) \cap \mathcal{A}] = \mathbb{P}[(|G| = Y_1) \cap \mathcal{A}] \le \mathbb{P}[Y_1 \le \mu_1/2] + \mathbb{P}[\alpha(G) \ge \mu_1/(12h^2\log n)] \overset{(5),(8)}{\le} 2n^{-2t},$$

as desired. □

3.4. **Combining the bounds.** We now use Lemmas 3.1 and 3.2 to prove Theorem 1.1.

Proof of Theorem 1.1. Define $\mathcal{B}$ to be the event that there exist $T \subseteq [n]$ of size t and an ordered sequence $V = (v_1, v_2, \ldots, v_{k-1})$ of distinct vertices such that $v_i \in [n] \setminus T$ for each

$1 \leq i \leq k-1$ and $X^- = 0$. As remarked before Lemma 3.2, $\Delta_{k-1}(R_{n,1}) \geq t$ implies $\mathcal{B}$. So we can apply Lemmas 3.1 and 3.2 to see that

$$\mathbb{P}[\Delta_{k-1}(R_{n,1}) \geq t] \leq \mathbb{P}[\mathcal{B}] \leq \mathbb{P}[\mathcal{A}^c] + \mathbb{P}[\mathcal{A} \cap \mathcal{B}] \leq n^{-k} + n^{t+k-1}(2n^{-2t}) = o(1).$$

This completes the proof of Theorem 1.1. □

References

[1] N. Alon and J.H. Spencer, *The Probabilistic Method*, Wiley-Interscience, New York (1992).

[2] P. Bennett and T. Bohman, A note on the random greedy independent set algorithm, *arXiv:1308.3732*, (2013).

[3] T. Bohman, The triangle-free process, *Advances in Math.* **221** (2009), 1653–1677.

[4] T. Bohman and P. Keevash, The early evolution of the H-free process, *Invent. Math.* **181** (2010), 291–336.

[5] T. Bohman and P. Keevash, Dynamic concentration of the triangle-free process, *arXiv:1302.5963*, (2013).

[6] T. Bohman, D. Mubayi and M. Picollelli, The independent neighborhoods process, *arXiv:1407.7192*, (2014).

[7] B. Bollobás and O. Riordan, Constrained graph processes, *Electronic J. Combin.* **7** (2000), R18.

[8] P. Erdős, S. Suen and P. Winkler, On the size of a random maximal graph, *Random Structures Algorithms* **6** (1995), 309–318.

[9] P. Erdős and P. Tetali, Representation of integers as the sum of k terms, *Random Structures Algorithms* **1** (1990), 245–261.

[10] G. Fiz Pontiveros, S. Griffiths and R. Morris, The triangle-free process and $R(3,k)$, *arXiv:1302.6279*, (2013).

[11] D. Osthus and A. Taraz, Random maximal H-free graphs, *Random Structures Algorithms* **18** (2001), 61–82.

[12] M. Picollelli, The final size of the C_ℓ-free process, *SIAM Journal on Discrete Math.* **28(3)** (2014), 1276–1305.

[13] M. Picollelli, The diamond-free process, *Random Structures Algorithms* **45** (2014), 513-551.

[14] L. Warnke, The C_ℓ-free process, *Random Structures Algorithms* **44** (2014), 490–526.

[15] L. Warnke, When does the K_4-free process stop?, *Random Structures Algorithms* **44** (2014), 355–397.

[16] G. Wolfovitz, The K_4-free process, *arXiv:1008.4044*, (2010).

Daniela Kühn, Deryk Osthus, Amelia Taylor
School of Mathematics
University of Birmingham
Edgbaston
Birmingham
B15 2TT
UK

E-mail addresses: {`d.kuhn, d.osthus`}`@bham.ac.uk, a.m.taylor@pgr.bham.ac.uk`