Saturation in Random Hypergraphs

Sahar Diskin * Ilay Hoshen * Dániel Korándi † Benny Sudakov ‡
Maksim Zhukovskii §

Abstract

Let K_n^r be the complete r-uniform hypergraph on n vertices, that is, the hypergraph whose vertex set is $[n] \coloneqq \{1, 2, \dots, n\}$ and whose edge set is $\binom{[n]}{r}$. We form $G^r(n,p)$ by retaining each edge of K_n^r independently with probability p.

An r-uniform hypergraph $H \subseteq G$ is *F-saturated* if H does not contain any copy of F, but any missing edge of H in G creates a copy of F. Furthermore, we say that H is *weakly F-saturated* in G if H does not contain any copy of F, but the missing edges of H in G can be added back one-by-one, in some order, such that every edge creates a new copy of F. The smallest number of edges in an F-saturated hypergraph in G is denoted by $\mathit{sat}(G,F)$, and in a weakly F-saturated hypergraph in G by $\mathit{w\text{-}sat}(G,F)$.

In 2017, Korándi and Sudakov initiated the study of saturation in random graphs, showing that for constant p, with high probability $\mathit{sat}(G(n,p),K_s) = (1+o(1))n\log_{\frac{1}{1-p}} n$, and $\mathit{w\text{-}sat}(G(n,p),K_s) = \mathit{w\text{-}sat}(K_n,K_s)$. Generalising their results, in this paper, we solve the suturation problem for random hypergraphs for every $2 \le r < s$ and constant p.

1. Introduction and main results

Denote by K_n^r the complete r-uniform hypergraph on n vertices, that is, the hypergraph whose vertex set is $[n] \coloneqq \{1, 2, \dots, n\}$ and whose edge set is $\binom{[n]}{r}$. For fixed r-uniform hypergraphs F and G, we say that a hypergraph $H \subseteq G$ is (strongly) *F-saturated* in G, if H does not contain any copy of F, but adding any edge $e \in E(G) \setminus E(H)$ to H creates a copy of F. We let $\mathit{sat}(G,F)$ denote the minimum number of edges in an F-saturated hypergraph in G. The problem of determining $\mathit{sat}(K_n^2,K_s^2)$ was raised by Zykov [19] in 1949, and independently by Erdős, Hajnal, and Moon [8] in 1964. They showed that $\mathit{sat}(K_n^2,K_s^2) = \binom{n}{2} - \binom{n-s+2}{2}$. Their result was later generalised by Bollobás [3] who showed that $\mathit{sat}(K_n^r,K_s^r) = \binom{n}{r} - \binom{n-s+r}{r}$.

We say that an r-uniform hypergraph $H \subseteq G$ is *weakly* F-saturated in G if H does not contain any copy of F, but the edges of $E(G) \setminus E(H)$ admit an ordering $e_1, \dots, e_k$ such that for each $i \in [k]$, the hypergraph $H_i = H \cup \{e_1, \dots, e_i\}$ contains a copy of F containing the edge e_i. We define $\mathit{w\text{-}sat}(G,F)$ to be the minimum number of edges in a weakly F-saturated hypergraph in G. Note that any H which is F-saturated in G is also weakly F-saturated in G, and hence $\mathit{w\text{-}sat}(G,F) \le \mathit{sat}(G,F)$. Thus, weak saturation can be viewed as a natural extension of saturation. The problem of determining $\mathit{w\text{-}sat}(K_n^r,K_s^r)$ was first raised by Bollobás [4], who conjectured that $\mathit{w\text{-}sat}(K_n^r,K_s^r) = \mathit{sat}(K_n^r,K_s^r)$. Using ingenious algebraic methods, this conjecture was verified by Lovász [17], Frankl [9], and Kalai [11, 12]:

*School of Mathematical Sciences, Tel Aviv University, Tel Aviv 6997801, Israel.
†Zürich, Switzerland.
‡Department of Mathematics, ETH, Zürich, Switzerland. Research supported in part by SNSF grant 200021_196965.
§Department of Computer Science, The University of Sheffield, Sheffield S1 4DP, United Kingdom.

Theorem 1.1 ([9, 11, 12, 17]).

$$w\text{-}sat(K_n^r, K_s^r) = sat(K_n^r, K_s^r) = \binom{n}{r} - \binom{n-s+r}{r}.$$

The binomial random graph $G(n,p)$ is obtained by retaining every edge of K_n^r independently, with probability p. Similarly, the binomial random r-uniform hypergraph $G^r(n,p)$ is obtained by retaining every edge of K_n^r independently with probability p. In 2017, Korándi and Sudakov [15] initiated the study of saturation and weak saturation in random graphs.

Theorem 1.2 ([15]). *Let $p \in (0,1)$ and let s be a constant. Then,* ***whp****,*

(a) $sat(G(n,p), K_s^2) = (1+o(1))n \log_{\frac{1}{1-p}} n.$

(b) $w\text{-}sat(G(n,p), K_s^2) = w\text{-}sat(K_n^2, K_s^2).$

Note that there is quite a stark difference between the (strong) saturation number and the weak saturation number. For weak saturation, **whp** the answer is the same as in the complete graph, whereas for saturation **whp** there is an additional $\log n$ factor compared with the complete graph.

Since the work of Korándi and Sudakov, there have been several papers devoted to the study of $sat(G(n,p), F)$ and $w\text{-}sat(G(n,p), F)$ for graphs F which are not cliques, in particular when p is constant. Mohammadian and Tayfeh-Rezaie [18] and Demyanov and Zhukovskii [6] proved tight asymptotic for stars, $F = K_{1,s}$. Considering cycles, Demidovich, Skorkin, and Zhukovskii [5] showed that **whp** $sat(G(n,p), C_m) = n + \Theta\left(\frac{n}{\log n}\right)$ for $m \geq 5$, and **whp** $sat(G(n,p), C_4) = \Theta(n)$. Considering a more general setting, Diskin, Hoshen, and Zhukovskii [7] showed that for every graph F, **whp** $sat(G(n,p), F) = O(n \log n)$, and gave sufficient conditions for graphs F for which **whp** $sat(G(n,p), F) = \Theta(n)$. As for weak saturation, Kalinichenko and Zhukovskii [14] gave sufficient conditions on F for which **whp** $w\text{-}sat(G(n,p), F) = w\text{-}sat(K_n, F)$, and Kalinichenko, Miralaei, Mohammadian, and Tayfeh-Rezaie [13] showed that, for any graph F, **whp** $w\text{-}sat(G(n,p), F) = (1+o(1))\, w\text{-}sat(K_n, F)$.

Another natural and challenging direction is to extend the results of Korándi and Sudakov to hypergraphs. Indeed, in their paper from 2017, they asked whether their results could be extended to r-uniform hypergraphs. In this paper, we answer that question in the affirmative:

Theorem 1. *Let $p \in (0,1)$, and let $2 \leq r < s$ be constants. Then,* ***whp****,*

(a) $w\text{-}sat(G^r(n,p), K_s^r) = w\text{-}sat(K_n^r, K_s^r).$

(b) $sat(G^r(n,p), K_s^r) = (1+o(1))\binom{n}{r-1} \log_{\frac{1}{1-p^{r-1}}} n.$

Similarly to the case of $r = 2$, for weak saturation **whp** the answer is the same as in the complete r-uniform hypergraph, whereas for saturation there is **whp** an additional $\log n$ factor compared with the complete r-uniform hypergraph. Furthermore, the proof of Theorem 1(a) can be extended to $p \geq n^{-\alpha}$, for some appropriately chosen constant $\alpha > 0$. In fact, we believe there exists a threshold for the property of weak saturation *stability*, that is, for when $w\text{-}sat(G^r(n,p), K_s^r) = w\text{-}sat(K_n^r, K_s^r)$ — for graphs, the respective result was obtained by Bidgoli, Mohammadian, Tayfeh-Rezaie, and Zhukovskii [2] — though the problem of finding the threshold itself looks very challenging.

The proof of Theorem 1(b) follows the ideas appearing in [15], with some adaptations and more careful treatment. Since the key ideas generalise from the graph setting to the hypergraph setting, we present only an outline of the proof in Section 4. However, for the sake of completeness and the reader's convenience, we present the full proof in Appendix A.

The weak saturation problem for random hypergraphs turns out to be much harder and requires several new ideas. Indeed, for the upper bound of Theorem 1(a), a novel and delicate construction is needed, see Section 2 for an outline of the proof.

The paper is structured as follows. In Section 2 we demonstrate the key ideas of the proof of Theorem 1(a). In Section 3 we prove Theorem 1(a). In Section 4, we provide a detailed sketch of the proof for Theorem 1(b), with the complete proof appearing in Appendix A. Throughout the paper, we will use the shorthand $G := G^r(n,p)$ and abbreviate r-uniform hypergraph to r-graph.

2. Outline of the proof of Theorem 1(a)

The lower bound will follow quite simply from Theorem 1.1, and that **whp** G is weakly saturated in K_n^r. The proof of the upper bound follows from a delicate construction. For the sake of clarity of presentation, and since it already illustrates all the main issues one needs to overcome, we consider the case when $r = 3$.

Defining cores. The natural first step, similar in spirit to the approach taken in the case when the host graph is a complete hypergraph, is to choose a *core* $C_0 \subseteq [n]$ of size $s-3$, such that $G[C_0] \cong K^3_{s-3}$. Indeed, in the case of the complete hypergraph, we could then set H to be the graph whose edges are of the form $e \subseteq [n]$ with $G[e \cup C_0] \cong K^3_s$ and $e \cap C_0 \neq \varnothing$. We could then activate all the remaining edges $e \subseteq [n] \setminus C_0$, since in K^3_n we have that $G[C_0 \cup e] \cong K^3_s$, and all the edges induced by $C_0 \cup e$, except e, are in H.

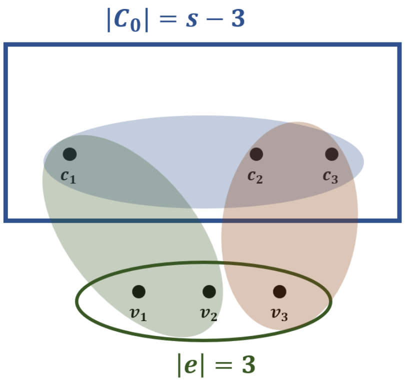

Figure 1: The edge e closes a copy with C_0 and can thus be activated. The three types of edges that are in H appear in shaded colours. The edges induced by C_0 appear in blue, the edges that contain one vertex from e and two vertices from C_0 appear in green, and the edges that contain two vertices from e and one vertex from C_0 appear in red.

However, unlike in the complete hypergraph, not all edges in the random hypergraph G (outside of C_0) close a copy of K^3_s with C_0. Thus, in what follows, we will choose additional *cores*, that is, sets of order $s-3$ which induce a copy of K^3_{s-3} in G. We further assume that cores can intersect only trivially, that is, if C_1, C_2 are two cores then either $C_1 = C_2$ or $C_1 \cap C_2 = \varnothing$.

For every $v \in [n] \setminus C_0$, if $G[C_0 \cup \{v\}] \cong K^3_{s-2}$, we choose the core $C_v := C_0$. Otheriwse, **whp** for every $v \in [n] \setminus C_0$ (for which $G[C_0 \cup \{v\}] \ncong K^3_{s-2}$), we can choose a core C_v such that every edge $f \subseteq C_v \cup \{v\} \cup C_0$, with $f \cap C_v \neq \varnothing$ is in G.

- Note that $G[C_0 \cup C_v] \cong K^3_{|C_0 \cup C_v|}$, and $|C_0 \cup C_v| = s-3$ when $C_v = C_0$, and $2(s-3)$ otherwise. Indeed, every edge e with $e \cap C_v \neq \varnothing$ is in G by the above assumption, and every $e \subseteq C_0$ is in G since $G[C_0] \cong K^3_{s-3}$.
- Thus, if $u \in C_v$ for some v, we have that $G[C_0 \cup \{u\}] \cong K^3_{s-2}$, and therefore $C_u = C_0$.

The core C_v will be useful for activating edges of the form $\{v, x, y\}$ when $x, y \in C_v$. However, as will be evident during the construction and activation of edges, defining cores for vertices will not suffice as there will be edges in the random hypergraph G, outside of these cores, which will not close a copy of K_s^3 with any of these cores. We thus now consider sets of pairs of vertices, $S = \{v_1, v_2\}$. If $v_1 \in C_{v_2}$ or $v_2 \in C_{v_1}$, we do not define $C_{\{v_1,v_2\}}$. If $v_1 \notin C_{v_2}$ and $v_2 \notin C_{v_1}$, we choose $C_S = C_{\{v_1,v_2\}}$ in the following manner. If for some $i \in \{1, 2\}$, every edge $f \subseteq C_{v_i} \cup S \cup C_0$ with $f \cap C_{v_i} \neq \varnothing$ is in G, we set $C_S = C_{v_i}$ (where if it holds for both v_1 and v_2, then it can be shown that $C_{v_1} = C_{v_2}$). As before, if $G[C_0 \cup \{v_1, v_2\}] \cong K_{s-1}^3$, we set $C_{\{v_1,v_2\}} = C_0$. Otherwise, **whp** there exists a core C_S such that for every $i \in \{1, 2\}$, every edge $f \subseteq C_S \cup S \cup C_{v_i} \cup C_0$ with $f \cap C_S \neq \varnothing$ is in G.

- Since every edge $e \subseteq C_S \cup S \cup C_{v_i} \cup C_0$ with $f \cap C_S \neq \varnothing$ is in G, we have that for every $c \in C_S$, $C_{\{v_i,c\}} = C_{v_i}$.
- Observe that $G[C_0 \cup C_{v_i} \cup C_S] \cong K^3_{|C_0 \cup C_{v_i} \cup C_S|}$. Indeed, we have shown before that $G[C_0 \cup C_{v_i}] \cong K^3_{2(s-3)}$, and by the choice of the cores, every edge $e \subseteq C_0 \cup C_{v_i} \cup C_S$ with $e \cap C_S \neq \varnothing$ is in G.
- Thus, if $X \subseteq C_{v_i} \cup C_S$ with $|X| = 2$, we have that $C_X = C_0$.

Constructing H. Let $H \subseteq G$ be a subhypergraph on $V(G) = [n]$, consisting of the following edges:

- Every edge $e \subseteq C_0$ is in H.
- For every $v \in [n] \setminus C_0$, we add to H all edges of the form $\{v\} \cup C'$ for every $C' \subseteq C_v$ of size two.
- For every $v \in [n] \setminus C_0$, we add to H all edges of the form $\{v\} \cup c_0 \cup c$ for every $c_0 \in C_0$ and $c \in C_v$. Indeed, every edge $f \subseteq C_v \cup \{v\} \cup C_0$ with $f \cap C_v \neq \varnothing$ is in G.
- For every $S = \{v_1, v_2\} \subseteq [n]$ for which C_S is defined, we add to H all edges of the form $S \cup \{c\}$ for every $c \in C_S$.

Recall that for every $S \subseteq V(G), |S| \leq 2$ and for every $X \subseteq C_S, |X| \leq 2$ we have that $C_X = C_0$. Thus, by this and by the above, we have added to H every edge $e \subseteq C_0 \cup C_S$ with $e \cap C_0 \neq \varnothing$.

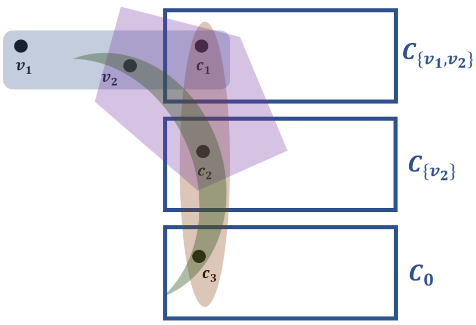

Figure 2: Illustration of a *chain* of cores, C_0, C_{v_2}, and $C_{\{v_1,v_2\}}$, together with edges that were added to H. The edge $\{v_1, v_2, c_1\}$ is added to H during the choice of $C_{\{v_1,v_2\}}$. The edge $\{v_2, c_1, c_2\}$ is added to H during the choice of $C_{\{v_2,c_1\}}$, where we note that by the properties of $C_{\{v_1,v_2\}}$, we have that $C_{\{v_2,c_1\}} = C_{v_2}$. The edge $\{c_1, c_2, c_3\}$ is added to H during the choice of $C_{\{c_1,c_2\}} = C_0$. Finally, the edge $\{v_2, c_2, c_3\}$ is added to H when we consider the pair $\{v_2, c_2\}$, where we observe that since $c_2 \in C_{v_2}$, we do not define $C_{\{v_2,c_2\}}$. Similar edges are added with the chain of cores C_0, C_{v_1}, and $C_{\{v_1,v_2\}}$.

Number of edges in H. The number of edges we have added to H is

$$\binom{s-3}{3}+\binom{n-s+3}{1}\binom{s-3}{2}+\binom{n-s+3}{2}\binom{s-3}{1}=\binom{n}{3}-\binom{n-s+3}{3}.$$

Indeed, the first term comes from the edges induced by C_0 which we have added to H.

For every $v \in [n] \setminus C_0$, we have chosen C_v (possibly $C_v = C_0$), $|C_v| = s-3$, and added all edges of the form $\{v\} \cup C'$ where $C' \subseteq C_v$ with $|C'| = 2$, contributing the second term in the sum.

Finally, for every $\{u, v\} \in [n] \setminus C_0$, if $u \notin C_v$ and $v \notin C_u$ (where we note that it may be that $C_v = C_u$), then we have chosen $C_{\{u,v\}}$, $|C_{\{u,v\}}| = s-3$, and added all edges of the form $\{u, v, c\}$ where $c \in C_{\{u,v\}}$. Otherwise, without loss of generality $v \in C_u$, and we have then added all edges of the form $\{u, v, c_0\}$ where $c_0 \in C_0$ – altogether contributing the third and final term in the sum.

Activating the edges in $E(G) \setminus E(H)$. First, note that we can activate all edges $e \subseteq [n] \setminus C_0$, such that $G[C_0 \cup e] \cong K_s^3$. Indeed, when defining the cores we ensured that for every $S \subsetneq e$ we have $C_S = C_0$, and thus all edges induced by $C_0 \cup \{e\}$, except e, are in H. We can thus activate e, which closes a copy of K_s^3 with C_0.

We can now activate the remaining edges $e = \{v_1, v_2, v_3\} \subseteq [n] \setminus C_0$. To that end, **whp** we can choose an *auxiliary clique* $\tilde{C} := \tilde{C}(e)$, such that $G[\tilde{C} \cup e] \cong K_s^3$ and for every $i \neq j \in \{1, 2, 3\}$, every edge $f \subseteq \tilde{C} \cup \{v_i, v_j\} \cup C_{\{v_i,v_j\}} \cup C_{v_i} \cup C_0$ with $f \cap \tilde{C} \neq \varnothing$ is in G. Crucially, observe that this choice of $\tilde{C}$ guarantees that for every $c \in \tilde{C}$, $C_{\{v_i,c\}} = C_{v_i}$.

Our goal is to show that all edges induced by $e \cup \tilde{C}$, except for e, are either in H or can be activated. We will then have that e closes a copy of K_s^3 with $\tilde{C}$, and can thus be activated as well.

Indeed, let us first consider the edges induced by $\tilde{C} \cup C_{v_i,v_j} \cup C_{v_i}$. By the choice of $\tilde{C}$ and the properties of chain of cores,

$$G\left[C_0 \cup \left(\tilde{C} \cup C_{\{v_i,v_j\}} \cup C_{v_i}\right)\right] \cong K^3_{|C_0 \cup \left(\tilde{C} \cup C_{\{v_i,v_j\}} \cup C_{v_i}\right)|}.$$

Thus, if the edge intersects C_0, we have added it to H when choosing C_0, and otherwise we have activated this edge with C_0.

Now, let us consider edges of the form $f = \{v_i, c_1, c_2\}$ where $c_1, c_2 \in \tilde{C}$. We claim that f closes a copy of K_s^3 with C_{v_i}, and can thus be activated. Indeed, we have activated all edges induced by $C_{v_i} \cup \tilde{C}$. Furthermore, we have added to H all edges of the form $\{v_i\} \cup C'$ where $C' \subseteq C_{v_i}$ with $|C'| = 2$. Since $C_{\{v_i,c_1\}} = C_{\{v_i,c_2\}} = C_{v_i}$, we have added to H all edges of the form $\{v_i, c_j, x\}$ where $c_j \in \{c_1, c_2\}$ and $x \in C_{v_i}$. We can therefore activate f, which closes a copy of K_s^3 with C_{v_i}.

Moving onwards, let us consider edges of the form $f = \{v_i, v_j, c\}$ where $c \in \tilde{C}$. Setting $S = \{v_i, v_j\}$, we claim that f closes a copy of K_s^3 with C_S, and can thus be activated. We have already activated edges induced by $\tilde{C} \cup C_S$. Furthermore, since for every $x \in C_S \cup \tilde{C}$ we have that $C_{\{v_i,x\}} = C_{v_i}$, edges of the form $\{v_i\} \cup C'$, where $C' \subseteq C_S \cup \tilde{C}$, close an edge with C_{v_i} and can thus be activated. Edges of the form $\{v_i, v_j, x\}$ where $x \in C_S$ have been added to H. Therefore, f closes a copy of K_s^3 with C_S.

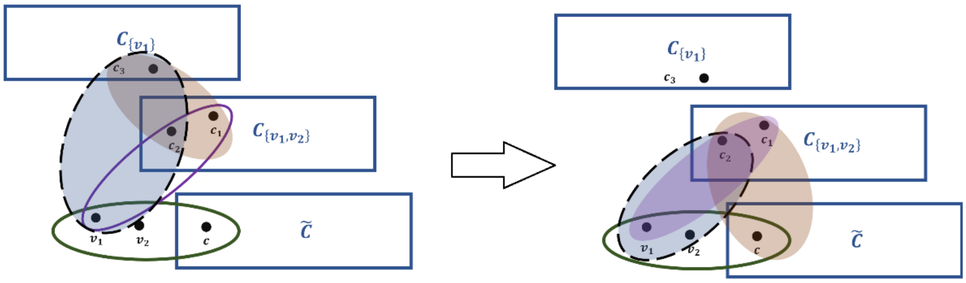

Figure 3: Illustration of a *chain* of activation, with the complexity of the construction evident already when $r = 3$. Towards activating an edge $\{v_1, v_2, v_3\}$ with $\tilde{C}$, we need to activate edges of the form $\{v_1, v_2, c\}$ where $c \in \tilde{C}$. To that end, we first activate all edges that form a clique with C_0, and in particular, all edges induced by a core. Then, as the left side illustrates, we can activate the edges of the form $\{v_1, c_1, c_2\}$ as they form a clique with C_{v_1}. Indeed, the edge $\{c_1, c_2, c_3\}$ forms a clique with C_0, and the edge $\{v_1, c_2, c_3\}$ is in H since $C_{\{v_1,c_2\}} = C_{v_1}$. We can then, as the right side illustrates, turn our attention to edges of the form $\{v_1, v_2, c\}$, which will close a clique with $C_{\{v_1,v_2\}}$. Here, the edge $\{c, c_1, c_2\}$ closes a clique with C_0 and thus has already been activated, and the edge $\{v_1, v_2, c_2\}$ is in H.

Finally, we are left with edges $\{v_1, v_2, v_3\}$ which intersect with C_0. To handle these types of edges, we once again choose an auxiliary clique $\tilde{C}$, with the same properties. The argument follows similarly to edges which do not intersect with C_0, where now we utilise the edges of the form $\{v_1, c_0, c'\}$, where $c_0 \in C_0$ and $c' \in C_{v_1}$, which we have added to H.

3. Weak saturation

Let $r \geq 3$ and $s \geq r + 1$ be integers. Let $p \in (0, 1]$ be a constant. Recall that $G := G^r(n, p)$. Set $\ell := s - r$.

Before delving into the construction of the upper-bound, let us note that the lower-bound follows quite simply from Theorem 1.1. Indeed, fix some edge $e \in K_n^r$. The probability that for a fixed set S of size ℓ we have that $e \cup G[S \cup e] \not\supseteq K_s^r$ is $1 - p^{\binom{s}{r}-1}$. This event is independent for every two disjoint sets of size ℓ, and thus the probability that e does not form a copy of K_s^r in $G \cup e$ is at most $\left(1 - p^{\binom{s}{r}-1}\right)^{\lfloor \frac{n}{\ell} \rfloor} = \exp(-\Theta(n))$. Union bound over $\binom{n}{r} \leq n^r$ edges $e \in K_n^r$ shows that **whp** every edge $e \in K_n^r$ belongs to a copy of K_s^r in $G \cup e$, and therefore **whp** G is weakly saturated in K_n^r. Now, if H is weakly saturated in G, then we can add to H edges one-by-one until we obtain G, and then keep adding edges until we reach K_n^r. But then, by Theorem 1.1, we have that H has at least $\binom{n}{r} - \binom{n-s+r}{r}$ edges.

In Section 3.1, we lay the groundwork for our proof: define the cores, construct the weakly saturated subhypergraph H, and count its edges. In Section 3.2, we show that there exists an ordering of the edges under which we can activate all edges of G that are not in H.

3.1. Laying the groundwork

Before we delve into the fine details, we note that what follows will be a natural extension of the construction in Section 2 to general r (together with their formalisation). We will once again define a core C_0, and inductively define cores for sets $S \subseteq V(G) \setminus C_0$ with size $1 \leq |S| \leq r - 1$. The construction of the cores will naturally have a *chain* property, that is, the core of S may (and will) depend on the cores of $S' \subsetneq S$. Once again, there will be several sets S for which we will not define a core, and instead draw relevant edges with vertices from C_0.

We begin by partitioning $V(G)$ into $\lfloor \frac{n}{\ell} \rfloor$ sets of size ℓ, denoted by $Q_1, \dots, Q_{\lfloor \frac{n}{\ell} \rfloor}$. Throughout the proof, we will use the following probabilistic lemma.

Lemma 3.1. *Let $k \geq 0$ be a constant. Then,* ***whp****, for every $S \subseteq V(G)$ with $|S| = k$, there exists Q_i, with $1 \leq i \leq \lfloor \frac{n}{\ell} \rfloor$, such that $S \cap Q_i = \varnothing$ and every edge $e \subseteq Q_i \cup S$ with $e \cap Q_i \neq \varnothing$ is in G.*

Proof. Fix S and fix i such that $Q_i \cap S = \varnothing$. The probability that every edge $e \subseteq Q_i \cup S$ with $e \cap Q_i \neq \varnothing$ is in G is at least $p^{\binom{\ell+k}{r}}$. There are at least $\lfloor \frac{n}{\ell} \rfloor - k$ different i such that $Q_i \cap S = \varnothing$. Therefore, the probability there doesn't exist such a Q_i is at most

$$\left(1 - p^{\binom{\ell+k}{r}}\right)^{\lfloor \frac{n}{\ell} \rfloor - k} = \exp\left(-\Theta(n)\right),$$

where we used our assumptions that $k, r, \ell,$ and p are constants. There are $\binom{n}{k}$ ways to choose S, and thus by the union bound, the probability of violating the statement of the lemma is at most

$$\binom{n}{k} \exp\left(-\Theta(n)\right) = o(1),$$

as required. $\square$

Defining the cores. Let i_0 be the first index in $\left[\lfloor \frac{n}{\ell} \rfloor\right]$ such that $G[Q_{i_0}] \cong K_\ell^r$, and let us set $C_0 = Q_{i_0}$ (note that by Lemma 3.1 **whp** such i_0 exists). For every $j \geq 0$, let i_j be the j-th index in $\left[\lfloor \frac{n}{\ell} \rfloor\right]$ such that $G[C_0 \cup Q_{i_j}] \cong K_{2\ell}^r$, and let us set $C_j = Q_{i_j}$ (once again, note that since $G[C_0] \cong K_\ell^r$, by Lemma 3.1 **whp** such C_j exist). We call these sets *cores*, and we enumerate them $C_0, C_1, \dots, C_m$ (we add to this sequence all sets that induce $K_{2\ell}^r$ with C_0). We continue assuming these cores have been defined deterministically.

Assigning cores to sets. Set $C_\varnothing = C_0$. For every vertex $v \in V \setminus C_0$, let $i(v)$ be the first index such that the following holds.

1. Every edge $e \subseteq \{v\} \cup C_0 \cup C_{i(v)}$ with $e \cap C_{i(v)} \neq \varnothing$ is in G.
2. $\{v\} \cap C_{i(v)} = \varnothing$.

We then set $C_{\{v\}} = C_{i(v)}$. Note that by the properties of the cores and by Lemma 3.1, **whp** such a $C_{\{v\}}$ exists for every $v \in V(G) \setminus C_0$. Furthermore, observe that the properties of $C_{\{v\}}$ imply that $G[\{v\} \cup C_{\{v\}}] = K_{\ell+1}^r$.

Now, we define cores for suitable subsets outside of C_0 by induction on their size. For every $j \in [2, r-1]$ and for every $S \subseteq V(G) \setminus C_0$ of size j, if

$$\text{there is no } S' \subsetneq S, \text{ such that } C_{S'} \text{ is defined and } S \setminus S' \subseteq C_{S'}, \qquad (S \text{ is core-definable})$$

we define C_S in the following way. Let $t \in [j-1]$. Let $i(S)$ be the first index such that the following holds for every sequence $\varnothing = S_0 \subsetneq \dots \subsetneq S_t \subsetneq S$ for which $C_{S_0}, \dots, C_{S_t}$ are defined.

(P1) Every edge $e \subseteq S \cup C_{i(S)} \cup C_{S_0} \cup \dots \cup C_{S_t}$ with $e \cap C_{i(S)} \neq \varnothing$ is in G.

(P2) $G\left[S_0 \cup C_{i(S)} \cup C_{S_0} \cup \dots \cup C_{S_t}\right] \cong K^r_{|S_0 \cup C_{i(S)} \cup C_{S_0} \cup \dots \cup C_{S_t}|}$.

(P3) $S \cap C_{i(S)} = \varnothing$.

We then set $C_S = C_{i(S)}$. Once again, note that by properties of the cores and by Lemma 3.1, **whp** for every $S \subseteq V(G) \setminus C_0$ of size j such that there is no $S' \subsetneq S$ with $C_{S'}$ defined and $S \setminus S' \subseteq C_{S'}$, we can find such a C_S. In what follows, we assume this holds deterministically.

Note that we have defined a core C_S for every set $S \subseteq V(G) \setminus C_0$ of size at most $r-1$ which is core-definable. Furthermore, if C_S is not defined, then there exists $S' \subseteq S$ for which $C_{S'}$ is defined and $S \setminus S' \subseteq C_{S'}$. Finally, note that if $S' \subseteq S$, then $i(S') \leq i(S)$ (given that both are defined) – indeed, the Properties (P1) through (P3) are closed under inclusion, and thus any index satisfying the properties for S must already satisfy the properties for S'.

Constructing H. Let $H \subseteq G$ be a subhypergraph consisting of the following edges:

(E1) Every edge $e \subseteq C_0$ is in H.

(E2) For every $\varnothing \neq S \subseteq V(G) \setminus C_0$ for which C_S is defined, we add to H all edges of the form $S \cup C'$ for every $C' \subseteq C_S$ of size $r - |S|$.

(E3) For every $\varnothing \neq S \subseteq V(G) \setminus C_0$ for which C_S is defined, we add to H all edges of the form $S \cup C' \cup C'_0$ for every $\varnothing \neq C' \subseteq C_S$ and $\varnothing \neq C'_0 \subseteq C_0$ satisfying $|C' \cup C'_0| = r - |S|$.

Note that edges of type (E1) are in G by the definition of the cores, and edges of type (E2) and type (E3) are in G by Property (P1) (and the first property when defining cores for singletons).

Number of edges in H. Note that, when counting the edges of H, we may assume that in (E3) we had that $C_S \neq C_0$, as otherwise these edges were added at step (E2).

For every $\varnothing \neq S \subseteq V(G) \setminus C_0$ for which C_S is defined, since $|C_S| = \ell$, we added $\binom{\ell}{r-|S|}$ edges of type (E2) to H. Since for every $S \subseteq V(G) \setminus C_0$ for which C_S is defined, and for every $C' \subseteq C_S$, we have that $C_{S \cup C'}$ is not defined, we can associate every edge of type (E3) with a set $X \subseteq V(G) \setminus C_0$ for which C_X is not defined. Then, for every $X \subseteq V(G) \setminus C_0$ for which C_X is not defined and $1 \leq |X| \leq r-1$, since $|C_0| = \ell$, we add $\binom{\ell}{r-|X|}$ edges of type (E3) to H. Therefore, the number of edges in H is

$$\binom{\ell}{r} + \sum_{i=1}^{r-1} \left(\sum_{\substack{S \subseteq V(G) \setminus C_0, \quad |S| = i \\ C_S \text{ is defined}}} \binom{\ell}{r-i} + \sum_{\substack{X \subseteq V(G) \setminus C_0, \quad |X| = i \\ C_X \text{ is not defined}}} \binom{\ell}{r-i} \right)$$
$$= \binom{\ell}{r} + \sum_{i=1}^{r-1} \binom{n-\ell}{i} \binom{\ell}{r-i} = \binom{n}{r} - \binom{n-\ell}{r}.$$

3.2. Activating the remaining edges

The argument for activating the remaining edges will be a natural extension of the argument given for the case $r = 3$ in Section 2, together with its formalisation.

Assigning an auxiliary core to every edge. Let us fix an edge $e \in E(G)$. By Lemma 3.1 applied with

$$S_{3.1} = e \cup \bigcup_{\substack{A \subsetneq e \\ A \text{ is core-definable}}} C_A,$$

we can choose an auxiliary core $\tilde{C} := \tilde{C}(e)$ from $Q_1, \dots Q_{\lfloor \frac{n}{\ell} \rfloor}$, such that the following holds. For every $S \subsetneq e$ and every $S_0 \subsetneq \cdots \subsetneq S_t \subsetneq S$ such that C_S is defined and C_{S_i} is defined for every $i \in [0, t]$:

($\overline{P}1$) Every edge $f \subseteq S \cup C_S \cup C_{S_0} \cup \cdots \cup C_{S_t} \cup \tilde{C}$ with $f \cap \tilde{C} \neq \varnothing$ is in G.

($\overline{P}2$) $G\left[S_0 \cup C_S \cup \tilde{C} \cup C_{S_0} \cup \cdots \cup C_{S_t}\right] \cong K^r_{|S_0 \cup C_S \cup \tilde{C} \cup C_{S_0} \cup \cdots \cup C_{S_t}|}$.

($\overline{P}3$) $(S \cup C_S \cup C_{S_0} \cup \cdots \cup C_{S_t}) \cap \tilde{C} = \varnothing$.

Core extension property. Recall that given $S \subseteq S \cup U \subseteq V(G)$, if $i(S)$ and $i(S \cup U)$ are defined, then $i(S) \leq i(S \cup U)$. When a set U is 'good enough', then we have $i(S \cup U) = i(S)$ – that is, $C_{S \cup U} = C_S$. In the following claim, we show that for a given S for which C_S is defined, we have that U is indeed 'good enough' whenever $U \subseteq C_{S_0} \cup \cdots \cup C_{S_t}$, for some $S \subsetneq S_0 \subsetneq \cdots \subsetneq S_t$. This 'extension' property of the cores will be important for us throughout the activation process, and in fact, we required Properties (P1) and (P2) when choosing our cores so that this extension property will hold.

Claim 3.2. *Let e_0 be an edge in G. Let $S \subseteq V(G) \setminus C_0$ be such that C_S is defined. For every set U such that $|S \cup U| \leq r-1$ and $U \subseteq C_{S_0} \cup \cdots \cup C_{S_t} \cup \tilde{C}(e_0)$ for some $S \subsetneq S_0 \subsetneq \cdots \subsetneq S_t \subsetneq e$ for which $C_{S_0}, \ldots, C_{S_t}$ are defined and different than C_S, we have that*

$$C_{S \cup U} = C_S.$$

Proof. We prove by induction on $|S \cup U|$, where the case when $|S \cup U| = 0$ follows trivially.

Let $S' \cup U' \subsetneq S \cup U$ with $S' \subseteq S$ and $U' \subseteq U$ such that $C_{S' \cup U'}$ is defined. Suppose towards contradiction that S' is not core-definable. Then, there exists $S'' \subsetneq S'$ such that $C_{S''}$ is defined and $S' \setminus S'' \subseteq C_{S''}$. But then, by the induction hypothesis, $C_{S'' \cup U'} = C_{S''}$, and by our assumption $(S' \cup U') \setminus (S'' \cup U') \subseteq C_{S''} = C_{S'' \cup U'}$. Therefore, $C_{S' \cup U'}$ is not defined — contradiction. Therefore, $C_{S'}$ is defined, and by the induction hypothesis, we have that $C_{S' \cup U'} = C_{S'}$.

Let us verify that $C_{S \cup U}$ is defined. Suppose towards contradiction that $S \cup U$ is not core-definable. Then, there exists $S' \cup U' \subsetneq S \cup U$ such that $S' \subseteq S$, $U' \subseteq U$, $C_{S' \cup U'}$ is defined, and $(S \cup U) \setminus (S' \cup U') \subseteq C_{S' \cup U'}$. We then have by the above that $C_{S' \cup U'} = C_{S'}$. Suppose first that $U' = U$. Then $(S \cup U) \setminus (S' \cup U') = S \setminus S' \subseteq C_{S'}$, contradicting the fact that C_S is defined. Otherwise, $U' \subsetneq U$. Then, by the induction hypothesis, $C_{S \cup U'} = C_S$. By our assumption, $(S \cup U') \setminus (S' \cup U') \subseteq C_{S' \cup U'}$, contradicting the fact that $C_{S \cup U'}$ is defined. Therefore, we conclude that $C_{S \cup U}$ is defined.

Therefore, it suffices to verify that $S \cup U$ satisfies Properties (P1) through (P3) with respect to C_S. Let $A_1 \subsetneq A_2 \subsetneq \cdots \subsetneq A_k \subsetneq S \cup U$ such that C_{A_i} is defined for every $i \in [k]$. By the above, we then have that $C_{A_i} = C_{S \cap A_i}$, and therefore when we consider C_{A_i}, we may assume that $A_i \subseteq S$.

1. First, let us show that every edge $e \subseteq (S \cup U) \cup C_S \cup C_{A_1} \cup \cdots \cup C_{A_k}$ with $e \cap C_S \neq \varnothing$ is in G, that is, let us verify Property (P1). Note that, by the above, there is some sequence of $S'_1 \subsetneq \cdots \subsetneq S'_k \subsetneq S$ such that $C_{A_j} = C_{S'_j}$ for every $j \in [k]$. Since $S \subsetneq S_0$, we have that $B_1 \subsetneq \cdots \subsetneq B_m$ where $B_1 = S'_1, \ldots, B_k = S'_k, B_{k+1} = S, B_{k+1} = S_0, \ldots, B_m = S_t$. Thus, if $e \cap \tilde{C}$, then by Property ($\overline{P}1$) with B_m we have that e is in G. Otherwise, let $\tau \in [t]$ be the maximal index such that $e \cap C_{S_\tau} \neq \varnothing$. We may assume that $\tau \geq 1$, otherwise $e \subseteq S \cup C_S \cup C_{A_1} \cup \ldots \cup C_{A_k}$ and then the above follows from Property (P1) with respect to S and C_S. Then, choosing $B_1 = A_1, \ldots, B_k = A_k, B_{k+1} = S, B_{k+2} = S_1, \ldots, B_m = S_\tau$, since $e \cap C_{S_\tau} \neq \varnothing$, we have that e is in G by Property (P1). Therefore, $S \cup U$ satisfies property (P1) with respect to C_S.

2. We now turn to show that $G[A_1 \cup C_S \cup C_{A_1} \cup \cdots \cup C_{A_k}] \cong K^r_{|A_1 \cup C_S \cup C_{A_1} \cup \cdots \cup C_{A_k}|}$ (note that if $A_1 \subseteq S$, then this follows from Property (P2) with respect to S and C_S). It suffices to show that $G[(A_1 \cup U) \cup C_S \cup C_{A_1} \cup \cdots \cup C_{A_k}] \cong K^r_{|(A_1 \cup U) \cup C_S \cup C_{A_1} \cup \cdots \cup C_{A_k}|}$, where we may assume as above that $A_k \subseteq S$.

To that end, note that by Property (P2) (or by Property ($\overline{P}2$) if U intersects with $\tilde{C}$) with respect to S_t and C_{S_t}, we have that for every $B_1 \subsetneq B_2 \subsetneq \cdots \subsetneq B_m \subseteq S_t$ such that C_{B_i} is defined for every $i \in [m]$, $G[B_1 \cup C_{S_t} \cup C_{B_1} \cup \cdots \cup C_{B_m} \cup \tilde{C}] \cong K^r_{|B_1 \cup C_{S_t} \cup C_{B_1} \cup \cdots \cup C_{B_m} \cup \tilde{C}|}$. We may thus choose $B_1 = A_1, \ldots, B_k = A_k, B_{k+1} = S, B_{k+2} = S_0, \ldots, B_m = S_t$, and since $U \subseteq C_{S_0} \cup \cdots \cup C_{S_t} \cup \tilde{C}$, we conclude that $S \cup U$ satisfies Property (P2).

3. Finally, let us show that $(S \cup U) \cap C_S = \varnothing$. Indeed, since C_S is defined, we have that $S \cap C_S = \varnothing$, and by our assumption, $C_{S_0}, \ldots, C_{S_t}$ are different from C_S (and in particular disjoint, due to the definition of the sequence Q_i, $i \in [\lfloor \frac{n}{\ell} \rfloor]$), and by property ($\overline{P}3$), $\tilde{C}$ is disjoint from C_S. Therefore, $U \cap C_S = \varnothing$ and $(S \cup U) \cap C_S = \varnothing$.

□

Activating the edges. With this at hand, we are ready to prove that we can activate all the edges of G outside of H. We consider separately edges that intersect with C_0 and edges that do not.

Claim 3.3. *One can activate all edges $e \in V(G) \setminus C_0$.*

Proof. Let $e \subseteq V(G) \setminus C_0$ be of size r. We will show that e closes a copy of K^r_s with $\tilde{C} = \tilde{C}(e)$.

Let us prove by induction on $|S|$ that for every set $S \subsetneq e$, one can activate all edges (that are not in H) of the form $S \cup U$, where $U \subseteq \tilde{C} \cup C_{S_0} \cup \cdots \cup C_{S_t}$ for $S \subsetneq S_0 \subsetneq S_1 \subsetneq \cdots \subsetneq S_t \subsetneq e$ for which C_{S_i} is defined for all $i \in [t]$. Note that we need the above only for $U \subseteq \tilde{C}$, however, for the induction argument it will be easier to prove the aforementioned stronger statement.

For the base case of the induction, we consider $S = \varnothing$. Let $f \subseteq \tilde{C} \cup C_{S_0} \cup \cdots \cup C_{S_t}$ be an edge for some $\varnothing \subsetneq S_0 \subsetneq S_1 \subsetneq \cdots \subsetneq S_t \subsetneq e$. Note that by Property (P2), $G[f \cup C_0] = K^r_{|f \cup C_0|}$, and thus if $f \cap C_0 \neq \varnothing$, we have that f is in H. Let us thus suppose that $f \cap C_0 = \varnothing$. By (E1), all the edges induced by C_0 are in H. Moreover, for every $S' \subsetneq f$, since $G[f \cup C_0] = K^r_{|f \cup C_0|}$, we have that $C_{S'} = C_0$. Thus, by (E2), every edge of the form $S' \cup C'_0$, for every $C'_0 \subseteq C_0$ of size $r - |S'|$, is in H. Hence, $f \cup C_0$ is a clique in G, and all its edges but f are in H. Thus, we can activate f.

For the induction step, let $i \in [r-1]$. Suppose that for every $S' \subsetneq e$, $|S'| < i$, and $S' \subsetneq A_1 \subsetneq A_2 \subsetneq \cdots \subsetneq A_k \subsetneq e$ such that C_{A_i} exists for all $i \in [k]$, we have activated (or added to H) all edges of the form $S' \cup U'$ where $U' \subseteq \tilde{C} \cup C_{A_1} \cup \ldots \cup C_{A_k}$. Let $S \subsetneq e$ of size i. Let f be an edge of the form $S \cup U$ where $U \subseteq \tilde{C} \cup C_{S_0} \cup \cdots \cup C_{S_t}$ for some $S \subsetneq S_0 \subsetneq S_1 \subsetneq \cdots \subsetneq S_t \subsetneq e$ for which $C_{S_0}, \ldots, C_{S_t}$ are defined. We will consider two separate cases, determined by whether C_S is defined or not.

Assume first that C_S is defined. We will show that we can activate f by closing a copy of K^r_s induced by $f \cup C_S$. Indeed, note if $f \cap C_S \neq \varnothing$ then f is in H (indeed, by Claim 3.2, $C_{f \setminus C_S} = C_S$). We may thus suppose $f \cap C_S = \varnothing$. By Claim 3.2, we have that $C_{S \cup U'} = C_S$ for every $U' \subseteq U$. Thus, we added to H all edges of the form $S \cup U' \cup C'$ for every $U' \subsetneq U$ and $C' \subseteq C_S$. By the induction hypothesis, we have already activated (or have in H) all edges of the form $S' \cup U'$ for every $S' \subsetneq S$, $U' \subseteq U \cup C_S$. Hence, we can activate f by closing a copy of K^r_s induced by $f \cup C_S$.

Assume now that C_S is not defined. Then, there exists $S' \subsetneq S$ such that $C_{S'}$ is defined and $S \setminus S' \subseteq C_{S'}$. By Claim 3.2, we have that $C_{S' \cup U'} = C_{S'}$ for every $U' \subseteq U$. Hence, we have already added to H all edges of the form $S' \cup U' \cup C'$ for every $U' \subseteq U$ and $\varnothing \neq C' \subseteq C_{S'}$. In particular, letting $C' = S \setminus S'$, we get that the edge $f = S \cup U$ is already in H.

Altogether, all edges induced by $e \cup \tilde{C}$, other than e, are either in H or have been activated, and thus e can be activated by closing a copy of K^r_s with $\tilde{C}$. □

Claim 3.4. *One can activate all the edges intersecting with C_0.*

Proof. Let e be an edge such that $e \cap C_0 \neq \varnothing$. Once again, we will show that e closes a copy of K_s^r with $\tilde{C} = \tilde{C}(e)$ in H.

Let us prove by induction on $|S|$ that for every set $S \subsetneq e \setminus C_0$, one can activate (if not already in H) all edges of the form $S \cup U$ where $U \subseteq C_0 \cup \tilde{C} \cup C_{S_0} \cup \cdots \cup C_{S_t}$ for some $S \subsetneq S_0 \subsetneq S_1 \subsetneq \cdots \subsetneq S_t \subsetneq e \setminus C_0$ for which $C_{S_0}, \ldots, C_{S_t}$ were defined. We can then activate e as it closes a copy of K_S^r with $\tilde{C}$.

For the base case of the induction, we consider $S = \varnothing$. Let $f \subseteq C_0 \cup \tilde{C} \cup C_{S_0} \cup \cdots \cup C_{S_t}$ be an edge for some $S \subsetneq S_0 \subsetneq S_1 \subsetneq \cdots \subsetneq S_t \subsetneq e \setminus C_0$ for which $C_{S_0}, \ldots, C_{S_t}$ were defined. By Claim 3.3, we may assume that $f \cap C_0 \neq \varnothing$. By Claim 3.2, we have that $C_{f \setminus C_0} = C_S = C_0$ since $S = \varnothing$. Therefore, by (E2), f is in H.

For the induction step, let $i \in [r-1]$ and take $S \subsetneq e \setminus C_0$ of size i. Let f be an edge of the form $S \cup U$, where $U \subseteq C_0 \cup \tilde{C} \cup C_{S_0} \cup \cdots \cup C_{S_t}$ for some $S \subsetneq S_0 \subsetneq S_1 \subsetneq \cdots \subsetneq S_t \subsetneq e \setminus C_0$ for which $C_{S_0}, \ldots, C_{S_t}$ were defined. We will consider two separate cases, determined by whether C_S is defined or not.

Assume first that C_S is defined. If $U \cap C_0 = \varnothing$, then we can activate $S \cup U$ by Claim 3.3. If $U \cap C_0 \neq \varnothing$, note that by Claim 3.2, $C_{S \cup (U \setminus C_0)} = C_S$. We thus have that f closes a copy of K_s^r with C_S. Indeed, let $h \subseteq f \cup C_S$ be an edge with $h \neq f$. If $S \cap h \neq S$, then $|S \cap h| \leq i - 1$. Writing $S' = S \cap h$ and $U' = h \setminus S$, we want to apply the induction hypothesis with $S' \cup U'$. Indeed, note that $U' \subseteq C_0 \cup \tilde{C} \cup C_S \cup C_{S_0} \cup \cdots \cup C_{S_t}$, and in particular where $S' \subsetneq S \subsetneq S_0 \subsetneq \cdots \subsetneq S_t \subsetneq e$. Thus, by the induction hypothesis, we have already activated (or added to H) the edge h. Otherwise, we have $S \subseteq h$. Noting that by Claim 3.2, $C_{h \setminus C_0} = C_S$, by (E3) we have already added the edge h to H.

Assume now that C_S is not defined. Then, there exists $S' \subseteq S$ such that $C_{S'}$ is defined and $S \setminus S' \subseteq C_{S'}$. By Claim 3.2, $C_{S' \cup (U \setminus C_0)} = C_{S'}$. Thus, by (E3), we added to H all edges of the form $S' \cup (U \setminus C_0) \cup C' \cup C_0'$ for every $\varnothing \neq C' \subseteq C_{S'}$ and $C_0' \subseteq C_0$. In particular, letting $C' = S \setminus S' \subseteq C_{S'}$ and $C_0' = U \cap C_0$, we get that the edge $f = S' \cup (U \setminus C_0) \cup (S \setminus S') \cup (U \cap C_0) = S \cup U$ was added to H. ☐

4. Strong Saturation

We begin with the proof of the lower bound of Theorem 1(b), which sheds some light as to why typically $sat(G^r(n,p), K_s^r) = (1+o(1))\binom{n}{r-1} \log_{\frac{1}{1-p^{r-1}}}(n)$.

Proof of the lower bound of Theorem 1(b). Let us show that if H is K_s^r-saturated in G, then $e(H) \geq (1+o(1))\binom{n}{r-1} \log_{\frac{1}{1-p^{r-1}}} n$. Note that if H is K_s^r-saturated in G, then adding any $e \in E(G) \setminus E(H)$ creates a new copy of K_s^r, and in particular, a new copy of K_{r+1}^r. Let $\alpha = \frac{1}{1-p^{r-1}}$.

Given an $(r-1)$-subset $S \subseteq V(G)$, let $N_H(S) = \{v \in V(G) \colon \{v\} \cup S \in E(H)\}$. Let

$$A := \{S \subseteq V(G) \colon |S| = r-1 \text{ and } |N_H(S)| \geq \log_\alpha^2 n\},$$

and set $B = \binom{V(G)}{r-1} \setminus A$. Note that if $|A| = \Omega\left(n^{r-1}\right)$, then $|E(H)| = \Omega\left(n^{r-1} \log_\alpha^2 n\right)$, and we are done. We may thus assume that $|A| = o\left(n^{r-1}\right)$, and thus $|B| = (1+o(1))\binom{n}{r-1}$.

Let $S \in B$. We claim that there are at least $(1+o(1)) \log_\alpha n$ edges $e \in E(H)$, such that S is the only $(r-1)$-subset of e which is in B. This will imply that the number of edges in H is at least $|B| \log_\alpha n = (1+o(1))\binom{n}{r-1} \log_\alpha n$, as required. Let $u \in V(G)$ be such that $e = \{u\} \cup S \in E(G) \setminus E(H)$. Then $\{u\} \cup S$ closes a copy of K_{r+1}^r together with H, in particular, there is some $w \in V(G)$ such that $e \cup H[S \cup \{u\} \cup \{w\}] \cong K_{r+1}^r$. Note that for all but at most $\log_\alpha^4 n$ choices of u, we have that the only $(r-1)$-subset of $S \cup \{w\}$ that is in B is S. Indeed, $S \in B$, so there are at most $\log_\alpha^2 n$ choices of w such that $\{w\} \cup S \in E(H)$. Then, if at least one other $(r-1)$-subset $S \neq S' \subseteq S \cup \{w\}$ is such that $S' \in B$, then we have at most $\log_\alpha^2 n$ choices for this u as well, since $S' \cup \{u\} \in E(H)$.

Thus, we have a set U of at least $|N_G(S)| - \log_\alpha^4 - O(1)$ vertices u such that there exists $w = w(u)$ forming an edge with S and with every $\{u\} \cup S'$, where $S' \subset S$, $|S'| = r-2$, and such that S is the only $(r-1)$-subset of $S \cup \{w\}$ that is in B. Let us prove that $|\{w(u)\colon u \in U\}| \geq (1+o(1))\log_\alpha n$. This will immediately imply the desired upper bound.

To that end, it suffices to show that **whp** for any $(r-1)$-set S, and W of size at most $\log_\alpha n - 5\log_\alpha \log_\alpha n$, there are at least $2\log_\alpha^4 n$ vertices $u \in V(G)\setminus(S \cup W)$ such that $S \cup \{u\} \in E(G)$, and for every $w \in W$ there is some $X \subset S \cup \{u\}$ with $|X| = r-1$, $X \neq S$, such that $X \cup \{w\} \notin E(G)$. Fix S and W. The probability that u satisfies the above is $p' = p(1-p^{r-1})^{|W|} \geq p \cdot \frac{\ln^5 n}{n}$. These events are independent for different u, so the number of vertices satisfying this property is distributed as $Bin(n-(r-1)-|W|, p')$. Thus, by a standard Chernoff's bound, the probability that there are less than $2\log_\alpha^4 n$ such vertices is at most $\exp\left(-\Omega(\ln^5 n)\right)$. There are $\binom{n}{r-1}$ ways to choose S and $\sum_{i=1}^{\log_\alpha n}\binom{n}{i} \leq n^{\log_\alpha n}$ ways to choose W, and thus by the union bound there are **whp** at least $2\log_\alpha^4 n$ such vertices. □

We now turn to the upper bound of Theorem 1(b). As we follow here the proof strategy from [15] with minor adjustments, in the next section we give an outline of the proof. We give the full proof in Appendix A for the sake of completeness.

4.1. Outline of the upper bound of Theorem 1(b)

We will utilise two lemmas. The first one generalises a result of Krivelevich [16] to the case of hypergraphs:

Lemma 4.1. *Let $r \geq 3$ and $t \geq r$ be integers. There exist positive constants c_0 and c_1 such that if*

$$\rho \geq c_1 n^{-\frac{t+1-r}{\binom{t+1}{r}-1}} \quad \text{and} \quad k \geq c_0 n^{\frac{\binom{t}{r}(t+1-r)}{\left(\binom{t+1}{r}-1\right)(t-1)}} \log^{\frac{1}{t-1}} n,$$

*then **whp** $G^r(n,\rho)$ contains a subhypergraph H on $[n]$ such that*

- *H is K_{t+1}^r-free,*
- *every induced k-subhypergraph of H contains a copy of K_t^r.*

The proof of Lemma 4.1 follows, in general, ideas present in [16], utilising Janson's inequality.

We also require the following lemma, which shows that for ρ chosen as in that in Lemma 4.1, typically every two vertices u and v are not both contained in 'many' copies of K_t^r in $G^r(n,\rho)$. This lemma is a new ingredient in the proof — this was not needed in the graph setting of [15], but is necessary in the hypergraph setting.

Lemma 4.2. *Let $r \geq 3$ and $t \geq 4$ satisfying $t \geq r$ be integers. Let $c > 0$ be a constant and let $p = cn^{-(t+1-r)/(\binom{t+1}{r}-1)}$. Then, **whp**, for every two vertices u and v, the number of copies of K_t^r in $G^r(n,\rho)$, which contains u and v, is at most $2\binom{n}{t-2}\rho^{\binom{t}{r}}$.*

The proof of Lemma 4.2 follows from an application of Markov's inequality to a high-moment random variable, together with the FKG inequality.

With these two lemmas at hand, we can now describe our construction. We say that an edge $e \in E(G) \setminus E(H)$ can be *completed* if it closes a copy of K_s^r in $H \cup \{e\}$. Recall that we aim to construct a K_s^r-free subhypergraph $H \subseteq G$ with $(1+o(1))\binom{n}{r-1}\log_{\frac{1}{1-p^{r-1}}} n$ edges, such that every edge $e \in E(G) \setminus E(H)$ can be completed. Throughout the proof, we make use of the following observation, already appearing in [15].

Observation 4.3. *It suffices to construct a K_s^r-free graph that completes all but $o(n^{r-1}\log n)$ edges. Indeed, by adding each of the uncompleted edges to H, if necessary, we obtain a K_s^r-saturated graph with an asymptotically equal number of edges.*

We set $\alpha = (1-p^{r-1})^{-1}$ and $\beta = (1-p^{\binom{s}{r}-\binom{s-r}{r}-1})^{-1}$. Throughout the rest of this section, unless explicitly stated otherwise, the base of the logarithms is α. We then set aside three disjoint subsets of $[n]$, A_1, A_2, and A_3, of sizes $a_1 := \frac{1}{p}\log n\left(1+\frac{3}{\log\log n}\right)$, $a_2 := sr\log_\beta n$, and $a_3 = \frac{a_2}{\log^{1/r} a_2}$, respectively. We let $B = [n]\setminus(A_1\cup A_2\cup A_3)$, shorthand $t = s-r$, and recall that $G \sim G^r(n,p)$.

We now define H to be a subhypergraph of G with the following edges:

- all edges of G intersecting both A_1 and B with at most t vertices from A_1; and,
- if $t \geq r$, we take $H[A_1] \subsetneq G[A_1]$ to be K_{t+1}^r-free, such that there exists a copy of K_t^r in every induced subhypergraph of $H[A_1]$ of size at least $c_0 a_1^{\frac{\binom{t}{r}(t+1-r)}{\left(\binom{t+1}{r}-1\right)(t-1)}}\log^{1/(t-1)} a_1$. Moreover, every two vertices in A_1 are contained in at most $2\binom{a_1}{t-2}(c_1a_1)^{-\frac{t+1-r}{\binom{t+1}{r}-1}}$ copies of K_t^r in $H[A_1]$. Otherwise, if $t<r$, we take $H[A_1]$ to be the empty hypergraph.

Note that, **whp**, by Lemmas 4.1 and 4.2 the subhypergraph described in the second bullet exists. Furthermore, the number of edges we have in H now is indeed $(1+o(1))\binom{n}{r-1}\log_{\frac{1}{1-p^{r-1}}} n$. Moreover, since $H[A_1]$ is K_{t+1}^r-free, if $t+1\geq r$ we see that H is K_s^r-free. If $t+1\leq r-1$, we note that we do not have edges intersecting both A_1 and B with at least $t+1$ vertices in A_1, and thus H is K_s^r-free.

Given an $(r-1)$-set of vertices $S\subseteq[n]$, the neighbourhood of S in $X\subseteq[n]$ is given by $N_X(S) := \{v\in X\colon \{v\}\cup S\in E(G)\}$. Given an $(r-1)$-set of vertices $S\subseteq B$, we say that S is *good* if $|N_{A_1}(S)|\geq pa_1 - \frac{\log n}{\log\log n}$, and otherwise we say that S is *bad*. Finally, we say that an edge $e\in G[B]$ is good if there exists at least one good $(r-1)$-subset $S\subseteq e$, and otherwise say that the edge e is *bad*.

We first utilise the subhypergraph in $H[A_1]$ in order to activate almost all the good edges $e\in E(G[B])$. Observe that good edges e have some $S\subseteq e$, with $|S|=r-1$ and such that S has a large neighbourhood in A_1. Fix an r-subset $e\subseteq B$. Note that given $S\subseteq e$ with $|S|=r-1$ and its neighbourhood in A_1, we have that

$$N_{S,e} := \left|\left\{v\in N_{A_1}(S)\colon \{v\}\cup S'\in E(G) \text{ for all } S'\subseteq e, |S'|=r-1\right\}\right| \sim Bin(|N_{A_1}(S)|, p^{r-1}).$$

By the choice of $H[A_1]$, we can then expose the edges of the form $S'\cup\{v\}$ in G, where $S'\subseteq e$ and $v\in N_{A_1}(S)$, and then using Chernoff's bound find in $N_{S,e}$ many copies of K_t^r. Now, note that given two copies of K_t^r in $H[N_{S,e}]$ the events that each of them closes a copy of K_s^r with e are dependent only if these copies intersect in two vertices. To bound these dependencies, we utilise the fact that in $H[A_1]$ every two vertices do not have many copies of K_t^r sharing them. This, together with Janson's inequality, allows us to show that **whp**, for all but $o(n^{r-1}\log n)$ good edges in $e\in E(G[B])$, there is at least one copy of K_t^r that closes a copy of K_s^r with e. Note that by Observation 4.3, this in fact completes our treatment for all the good edges in $G[B]$.

Before turning to bad edges in $G[B]$, let us discuss the difference in the treatment of good edges above from the setting of graphs in [15]. Indeed, in the case of graphs, a key part of the argument for the upper bound is to consider vertices v with a large neighbourhood in A_1. Then, fixing v and considering edges $uv\in E(G[B]))$, it suffices to find a copy of K_{s-2}^2 in $N_v\cap N_u$ – whose size is distributed according to $Bin(|N_v|,p)$. On the other hand, when considering hypergraphs, we consider instead an $(r-1)$-set $S\subseteq e$ with a substantial neighbourhood in A_1, $N_{S,e}$. Then, when considering edges $\{u\}\cup S\in E(G[B])$, it no longer suffices to find a copy of K^r_{s-r} in $\tilde{N}_{S,u} = \bigcap_{S' \subset S \cup \{u\}, |S'|=r-1} N_{S',e}$. Indeed, one needs to further consider edges with $1 \le \ell < r-1$ vertices from $S \cup \{u\}$, and $r-\ell$ vertices from $\tilde{N}_{S,u}$. As noted above, this further creates possible dependencies which do not appear in the case of graphs, and, in particular, this is why Lemma 4.2 is important to us in the hypergraph setting.

We now turn to bad edges in $G[B]$. As there is a non-negligible portion of them, we require the set A_2 to complete them. By Lemma 4.1 there exists a subhypergraph $H_2 \subseteq G[A_2]$ which is K^r_{t+1}-free and every large enough subset in it induces a copy of K^r_t, when $t \ge r$. Moreover, there are at least $r \log_\beta n$ vertex-disjoint copies of K^r_t in $H_2[A_2]$. We can thus add to H the edges of $H_2[A_2]$, edges intersecting both A_2 and B in G with at least two and at most t vertices from A_2, and edges of the form $S \cup \{v\}$ where $S \subseteq B$ is bad and $v \in A_2$. The first two are of order $o(n^{r-1}\log n)$, and using Chernoff's bound, one can show that there are at most $\frac{n^{r-1}}{\log n}$ bad $(r-1)$-subsets $S \subseteq B$, and thus the last set of edges is also of order $o(n^{r-1}\log n)$. Similarly to before, H is still K^r_s-free. Using Markov's inequality, **whp** we can now close all bad edges in $G[B]$ (recall that a bad edge has that all its $(r-1)$-subsets are bad).

Note that we have not added any edge to H intersecting both B and A_2, whose intersection with B is a good $(r-1)$-subset. As there are $\Omega(n^{r-1})$ good $(r-1)$-subsets in B, and $\Omega(\log n)$ vertices in A_2, we cannot ignore these edges, and this is the reason for setting aside the set A_3. Once again, we find in $G[A_3]$ a subhypergraph H_3 which is K^r_{t+1}-free, and every large enough subset in it induces a copy of K^r_t, when $t \ge r$.

The choice of edges which we add to H now is slightly more delicate. We add to H the edges of H_3, all edges intersecting both B and A_3 with at least two and at most t vertices from A_3, edges of the form $S \cup \{v\}$ where $S \subseteq B$ is good and $v \in A_3$, and edges intersecting both $B \cup A_2$ and A_3 with one vertex from A_2 and at least one and at most t vertices from A_3. This careful choice of edges allows us to show that the hypergraph H is still K^r_s-free. As the argument here is slightly more delicate, let us state it explicitly. Since there are no edges intersecting both A_1 and $A_2 \cup A_3$, it suffices to consider $X \subseteq B \cup A_2 \cup A_3$ of size s, and suppose towards contradiction that X induces a clique in H. Since $H[B]$ is an empty hypergraph, we must have that $|X \cap B| \le r-1$. Since we only added edges intersecting A_3, we may assume that X contains vertices from A_3. If $|X \cap A_3| \ge t+1$, we are done by similar arguments to before. Furthermore, if there are at least two vertices in $A_2 \cap X$, note that there are no edges in H containing two vertices from A_2 and at least one vertex from A_3, leading to a contradiction. Thus, we are left with the case where $|X \cap A_2| = 1, |X \cap A_3| = t, |X \cap B| = r-1$. Now, if $X \cap B$ is bad, there are no edges of the form $(X \cap B) \cup \{v\}$, $\{v\} \in A_3$, in H, and if it is good, then there are no edges of the form $(X \cap B) \cup \{v\}$, $\{v\} \in A_2$, in H — either way, we get that X cannot induce a clique. Let us note that this last part of the argument is also quite different from the argument in the setting of graphs. Indeed, in [15], to ensure that H remains K^2_s-free, one needs to consider a partition of A_2 into independent sets, and partition A_3 accordingly. Here, adding edges intersecting both $B \cup A_2$ and A_3 which intersect A_2 in at most one vertex suffices to show that H remains K^r_s-free.

Furthermore, due to the size of A_3, the number of edges we added at this stage is of order $o(n^{r-1}\log n)$. Using similar arguments to before, we can show that we can activate all but $o(n^{r-1}\log n)$ of the edges of the form $S \cup \{v\}$ where $S \subseteq B$ is good and $v \in A_2$. Again, by Observation 4.3, this completes our treatment for all edges of the form $S \cup \{v\}$ where $S \subseteq B$ is good and $v \in A_2$.

Finally, note that we have not completed all the edges: we did not complete edges of the form $S \cup \{v\}$ where $S \subseteq B$ is bad and $v \in A_3$ as well as some of the edges induced by $A_1 \cup A_2 \cup A_3$. Furthermore, if $t < r$, we have not completed edges instersecting both B and $A_1 \cup A_2 \cup A_3$ with more than t vertices in either A_1, A_2, or A_3. However, there are at most $o(n^{r-1}\log n)$ edges of these types, and we are thus done by Observation 4.3.

References

[1] N. Alon and J. H. Spencer. *The probabilistic method.* Hoboken, NJ: John Wiley & Sons, fourth edition, 2016.

[2] M. Bidgoli, A. Mohammadian, B. Tayfeh-Rezaie, and M. Zhukovskii. Threshold for weak saturation stability. *J. Graph Theory*, accepted for publication.

[3] B. Bollobás. On generalized graphs. *Acta Math. Acad. Sci. Hungar.*, 16:447–452, 1965.

[4] B. Bollobás. Weakly k-saturated graphs. In *Beiträge zur Graphentheorie (Kolloquium, Manebach, 1967)*, pages 25–31. B. G. Teubner Verlagsgesellschaft, Leipzig, 1968.

[5] Y. Demidovich, A. Skorkin, and M. Zhukovskii. Cycle saturation in random graphs. *SIAM J. Discrete Math.*, 37(3):1359–1385, 2023.

[6] S. Demyanov and M. Zhukovskii. Tight concentration of star saturation number in random graphs. *Discrete Math.*, 346(10):113572, 2023.

[7] S. Diskin, I. Hoshen, and M. Zhukovskii. A jump of the saturation number in random graphs? *arXiv preprint arXiv:2303.12046*, 2023.

[8] P. Erdős, A. Hajnal, and J. W. Moon. A problem in graph theory. *Amer. Math. Monthly*, 71:1107–1110, 1964.

[9] P. Frankl. An extremal problem for two families of sets. *European J. Combin.*, 3(2):125–127, 1982.

[10] S. Janson, T. Luczak, and A. Rucinski. An exponential bound for the probability of nonexistence of a specified subgraph in a random graph. In *Random graphs*, volume 87, pages 73–87, 1990.

[11] G. Kalai. Weakly saturated graphs are rigid. In *Convexity and graph theory (Jerusalem, 1981)*, volume 87 of *North-Holland Math. Stud.*, pages 189–190. North-Holland, Amsterdam, 1984.

[12] G. Kalai. Hyperconnectivity of graphs. *Graphs Combin.*, 1(1):65–79, 1985.

[13] O. Kalinichenko, M. Miralaei, A. Mohammadian, and B. Tayfeh-Rezaie. Weak saturation numbers in random graphs. *arXiv preprint arXiv:2306.10375*, 2023.

[14] O. Kalinichenko and M. Zhukovskii. Weak saturation stability. *European Journal of Combinatorics*, 114:103777, 2023.

[15] D. Korándi and B. Sudakov. Saturation in random graphs. *Random Structures Algorithms*, 51(1):169–181, 2017.

[16] M. Krivelevich. Bounding Ramsey numbers through large deviation inequalities. *Random Structures Algorithms*, 7(2):145–155, 1995.

[17] L. Lovász. Flats in matroids and geometric graphs. In *Combinatorial surveys (Proc. Sixth British Combinatorial Conf., Royal Holloway Coll., Egham, 1977)*, pages 45–86. Academic Press, London-New York, 1977.

[18] A. Mohammadian and B. Tayfeh-Rezaie. Star saturation number of random graphs. *Discrete Math.*, 341(4):1166–1170, 2018.

[19] A. A. Zykov. On some properties of linear complexes. *Mat. Sbornik N.S.*, 24(66):163–188, 1949.

A. Proof of the upper bound of Theorem 1(b)

A.1. Preliminaries

Let us first state (several versions of) Janson's inequality, which we will use later on. In some cases, we will need to bound the probability of the existence of many edge-disjoint copies of some graph in $G(n, p)$. Janson's inequality is a well-known tool, frequently utilised in such problems.

Let Ω be a finite universal set and let R be a random subset of Ω containing every element $r \in \Omega$ with probability p_r, independently. Let I be a finite set of indices, and let $\{A_i\}_{i \in I}$ be subsets of Ω. For every $i \in I$, let X_i be the indicator random variable of the event that $B_i \subset R$ and let $X = \sum_{i \in I} X_i$. For $i, j \in I$, we write $i \sim j$ if $A_i \cap A_j \neq \varnothing$. Define

$$\mu := \mathbb{E}[X] = \sum_{i \in I} \mathbb{P}(X_i = 1) \quad \text{and}$$
$$\Delta := \sum_{i \sim j} \mathbb{P}(X_i = X_j = 1),$$

where the sum in Δ goes over all ordered pairs $i, j \in I$ such that $i \sim j$.

Theorem A.1 (Janson's inequality [10])**.** $\mathbb{P}(X = 0) \leq e^{-\mu^2/2\Delta}$.

We also use versions of Janson's inequality (see, for example, [1]) which bound the probability of the event that we do not have many disjoint copies of A_i's in R. We call a set of indices $J \subseteq I$ a *disjoint family* if for every $i \neq j \in J$ we have $A_i \cap A_j = \varnothing$. Moreover, we call a set of indices $J \subseteq I$ a *maximal disjoint family* if J is a disjoint family and there is no $i \in I \setminus J$ such that $A_i \cap A_j = \varnothing$ for every $j \in J$. For every s, denote by D_s and M_s the events that there exists a disjoint family of size s and a maximal disjoint family of size s, respectively.

Lemma A.2. *For every integer s,*

$$\mathbb{P}(D_s) \leq \frac{\mu^s}{s!}.$$

For a probability bound on the size of a maximal disjoint family, we need another notation. Define

$$\nu = \max_{j \in I} \sum_{i \sim j} \mathbb{P}(X_i = 1).$$

Lemma A.3. *For every integer s,*

$$\mathbb{P}(M_s) \leq \frac{\mu^s}{s!} e^{-\mu + s\nu + \Delta/2}.$$

Let us derive from Lemma A.3 an upper bound to the probability that there exists a maximal disjoint family of size $s \leq (1-\epsilon)\mu$ for some $\epsilon \in (0, 1)$.

Corollary A.4. *If $\Delta = o(\mu)$ and $\nu = o(1)$, then, for every $\epsilon \in (0, 1)$,*

$$\mathbb{P}\left(\bigcup_{s \leq (1-\epsilon)\mu} M_s \right) \leq e^{-\epsilon^2 \mu/3}.$$

Proof. By the union bound,

$$\mathbb{P}\left(\bigcup_{s \leq (1-\epsilon)\mu} M_s \right) \leq \sum_{s=1}^{(1-\epsilon)\mu} \mathbb{P}(M_s) \leq \sum_{s=1}^{(1-\epsilon)\mu} \frac{\mu^s}{s!} e^{-\mu} e^{s\nu + \Delta/2}$$
$$\leq e^{\epsilon^3 \mu} \sum_{s=1}^{(1-\epsilon)\mu} \frac{\mu^s}{s!} e^{-\mu},$$

where the second inequality is true by Lemma A.3 and the last inequality is true since, by the assumptions of the corollary, $s\nu + \Delta/2 \le \epsilon^2\mu/6$. Note that the sum above equals to the probability that the value of a Poisson random variable with mean μ is at most $(1-\epsilon)\mu$. Thus, by a typical bound on the tails of the Poisson distribution (see, for example, [1, Theorem A.1.15]),

$$\mathbb{P}\left(\bigcup_{s\le(1-\epsilon)\mu} M_s\right) \le e^{\epsilon^2\mu/6}e^{-\epsilon^2\mu/2} = e^{-\epsilon^2\mu/3}.$$

□

We also make use of the following bounds (see, for example, [15, Claim 2.1]).

Claim A.5. *Let $0 < p < 1$ be a constant and $X \sim Bin(n,p)$ be a binomial variable. Then, for sufficiently large n and for any $a > 0$,*

1. $\mathbb{P}(X \ge np + a) \le e^{-\frac{a^2}{2(np+a/3)}}$,
2. $\mathbb{P}(X \le np - a) \le e^{-\frac{a^2}{2np}}$ *and*
3. $\mathbb{P}\left(X \le \frac{n}{\log^2 n}\right) \le (1-p)^{n-\frac{n}{\log n}}$.

A.2. Properties of random hypergraphs

In this section, we collect some properties of the random hypergraph that will be useful to us throughout the proof.

Lemma A.6. *Let $r \ge 3$ and $t \ge 4$ satisfying $t \ge r$ be integers. Let $c > 0$ be a constant and let $p = cn^{-(t+1-r)/(\binom{t+1}{r}-1)}$. Then,* ***whp****, for every two vertices u and v, the number of copies of K_t^r in $G^r(n,p)$, which contains u and v, is at most $2\binom{n}{t-2}p^{\binom{t}{r}}$.*

Proof. Fix two vertices u and v. Denote by X the number of copies of K_t^r in $G \sim G^r(n,p)$ which contain u and v. By the union bound over all the choices of u and v, it suffices to prove that

$$\mathbb{P}\left(X > 2\binom{n}{t-2}p^{\binom{t}{r}}\right) \ll n^{-2}.$$

Set $L = K\log n$ where K is a sufficiently large constant. For every subset $T \subset V(G)$ of size t, denote by Y_T the indicator random variable of the event $G[T] \cong K_t^r$. Below, we consider only those t-subsets of $V(G)$ that contain u and v (in particular, sum up over such sets in the computation of the L-th moment of X below). We have

$$\begin{aligned}
\mathbb{E}[X^L] &= \sum_{\substack{T_1,\dots,T_L\subset V(G)\\ |T_1|=\dots=|T_L|=t}} \mathbb{E}[Y_{T_1}\cdot\ldots\cdot Y_{T_L}]\\
&= \sum_{\substack{T_1,\dots,T_{L-1}\subset V(G)\\ |T_1|=\dots=|T_{L-1}|=t}} \mathbb{E}[Y_{T_1}\cdot\ldots\cdot Y_{T_{L-1}}] \sum_{T_L\subset V(G),|T_L|=t} \mathbb{E}[Y_{T_L} \mid Y_{T_1} = \ldots = Y_{T_{L-1}} = 1]\\
&= \sum_{\substack{T_1,\dots,T_{L-1}\subset V(G)\\ |T_1|=\dots=|T_{L-1}|=t}} \mathbb{E}[Y_{T_1}\cdot\ldots\cdot Y_{T_{L-1}}]\cdot\mathbb{E}[X \mid Y_{T_1} = \ldots = Y_{T_{L-1}} = 1]. \qquad (1)
\end{aligned}$$

Let H be a hypergraph with $e(H) \le L\cdot t$. By FKG inequality, $\mathbb{E}[X \mid H \subset G] \ge \mathbb{E}X$. On the other hand,

$$\mathbb{E}[X \mid H \subset G] \le \binom{n}{t-2}p^{\binom{t}{r}} + \sum_{j=r}^{t} e(H)^{\binom{j}{r}} n^{t-j} p^{\binom{t}{r}-\binom{j}{r}}.$$

The first summand in the right-hand side above is an upper bound to the expectation of the number of copies of K_t^r which does not share any edges with H while the second summand is an upper bound to the expectation of the number of copies of K_t^r which does share at least one edge with H. The index j indicates how many vertices in the copy of K_t^r belongs to H. So we have at most $\binom{j}{r}$ edges which are already in H. The term $e(H)^{\binom{j}{r}}$ is an upper bound to the number of choices of the common j vertices, n^{t-j} is an upper bound to the number of choices of the remaining $t-j$ vertices and $p^{\binom{t}{r}-\binom{j}{r}}$ is an upper bound to the probability that this copy appears in G conditioned on $H \subseteq G$. We will show that the second summand on the right-hand side of the inequality is asymptotically smaller than

$$\mathbb{E}[X] = (1+o(1))\binom{n}{t-2}p^{\binom{t}{r}}, \tag{2}$$

implying $\mathbb{E}[X \mid H \subset G] \leq (1+o(1))\mathbb{E}[X]$. Note that, for every $j = r, r+1, \ldots, t$, we have

$$\frac{n^{t-j}p^{\binom{t}{r}-\binom{j}{r}}}{n^{t-2}p^{\binom{t}{r}}} = n^{2-j}p^{-\binom{j}{r}} = \Theta\left(n^{2-j+\frac{\binom{j}{r}(t+1-r)}{\binom{t+1}{r}-1}}\right).$$

Thus, it is sufficient to show that the last expression tends to 0 as $n \to \infty$, which is true if

$$\frac{\binom{j}{r}(t+1-r)}{\binom{t+1}{r}-1} < j-2.$$

To show the latter, we will use the following claim.

Claim A.7. $\frac{\binom{t}{r}(t+1-r)}{(\binom{t+1}{r}-1)(t-1)} < 1 - \frac{r-1}{t}$.

Proof. We have

$$\begin{aligned}
\frac{\binom{t}{r}(t+1-r)}{\left(\binom{t+1}{r}-1\right)(t-1)} &< \frac{\binom{t}{r}}{\binom{t+1}{r}-1} = \frac{\binom{t}{r}}{\binom{t}{r}+\binom{t}{r-1}-1} \\
&= \frac{\binom{t}{r}}{\binom{t}{r}+\binom{t}{r}\cdot\frac{r}{t-r+1}-1} = \frac{1}{1+\frac{r}{t-r+1}-\frac{1}{\binom{t}{r}}} \\
&\leq 1-\frac{r}{t-r+1}+\frac{1}{\binom{t}{r}} \leq 1-\frac{r}{t-r+1}+\frac{1}{t-r+1} \\
&= 1-\frac{r-1}{t-r+1} < 1-\frac{r-1}{t}.
\end{aligned}$$

□

By Claim A.7,

$$\begin{aligned}
\frac{\binom{j}{r}(t+1-r)}{\binom{t+1}{r}-1} &< \frac{\binom{j}{r}}{\binom{t}{r}}\left(1-\frac{r-1}{t}\right)(t-1) \leq \frac{\binom{j}{r}}{\binom{t}{r}}(t-2) = \frac{j(j-1)\ldots(j-r+1)}{t(t-1)\ldots(t-r+1)}(t-2) \\
&\overset{r\geq 3}{\leq} j-2,
\end{aligned}$$

as needed.

Thus, we conclude that $\mathbb{E}[X \mid H \subset G] = (1+o(1))\mathbb{E}[X]$ uniformly over all H with $e(H) \leq L \cdot t$. By induction, from (1), we have

$$\mathbb{E}[X^L] = \Big((1+o(1))\mathbb{E}[X]\Big)^L.$$

From (2), by Markov's inequality, we get

$$\mathbb{P}\left(X > 2\binom{n}{t-2}p^{\binom{t}{r}}\right) \leq \frac{\mathbb{E}[X^L]}{((2+o(1))\mathbb{E}[X])^L} = \left(\frac{(1+o(1))\mathbb{E}[X]}{2\mathbb{E}[X]}\right)^L = (1/2+o(1))^L \ll n^{-2},$$

where the last asymptotical inequality is true if K is sufficiently large. $\square$

Lemma A.8. *Let $r \geq 3$ and $t \geq r$ be integers. There exist positive constants c_0 and c_1 such that if*

$$p \geq c_1 n^{-\frac{t+1-r}{\binom{t+1}{r}-1}} \quad \text{and} \quad k \geq c_0 n^{\frac{\binom{t}{r}(t+1-r)}{(\binom{t+1}{r}-1)(t-1)}} \log^{\frac{1}{t-1}} n,$$

*then **whp** $G^r(n,p)$ contains a subhypergraph H on $[n]$ such that*

- *H is K^r_{t+1}-free,*
- *every induced k-subhypergraph of H contains a copy of K^r_t.*

Proof. We may assume that k is the smallest possible integer satisfying the requirement in the statement of Lemma A.8, that is

$$k = \left\lceil c_0 n^{\frac{\binom{t}{r}(t+1-r)}{(\binom{t+1}{r}-1)(t-1)}} \log^{\frac{1}{t-1}} n \right\rceil.$$

Also, since the property that the lemma claims is increasing, without the loss of generality, we may set

$$p = c_1 n^{-\frac{t+1-r}{\binom{t+1}{r}-1}}.$$

Let $G \sim G^r(n,p)$. Let $\mathcal{K}$ be a set of all graphs obtained as a union of K^r_t and K^r_{t+1} sharing at least r vertices (we include a single representative of every isomorphism class into $\mathcal{K}$). For $S \subset [n]$,

- denote by X_S the maximum size of a set of pairwise edge-disjoint copies of K^r_t in $G[S]$;
- denote by Y_S the number of subhypergraphs $H \subset G$ isomorphic to a hypergraph from $\mathcal{K}$ and such that $H[V(H) \cap S]$ contains a copy of K^r_t;
- denote by Z_S the maximum size of a set of pairwise edge-disjoint hypergraphs $H \subset G$ isomorphic to a hypergraph from $\mathcal{K}$ and such that $H[V(H) \cap S]$ contains a copy of K^r_t.

Let A_S be the event that $X_S > \lambda Z_S$ where

$$\lambda := \binom{t}{r}\left(\binom{t+1}{r} - 1\right) + 1.$$

Claim A.9. *If A_S holds for every subset $S \subset V(G)$ of size k, then G contains a subhypergraph H which is K^r_{t+1}-free and every induced k-subhypergraph of H contains a copy of K^r_t.*

Proof. Let $\mathcal{H}$ be an inclusion-maximal family of edge-disjoint copies of K^r_{t+1} in G. Denote by $G' \subset G$ the subhypergraph obtained by deleting from G all edges that appear in a clique from $\mathcal{H}$. Note that G' is K^r_{t+1}-free. We claim that every induced k-subhypergraph of G' contains K^r_t, i.e. G' is the desired hypergraph.

Assume by contradiction that there exists a k-set $S \subset [n]$ such that there is no copy of K^r_t in $G'[S]$. Let $\mathcal{K}_S$ be a family of edge-disjoint copies of K^r_t in $G[S]$ of size X_S. Consider an auxiliary graph G_S with the vertex set $\mathcal{K}_S$ and the edge set defined as follows: cliques K^1, K^2 from $\mathcal{K}_S$ are adjacent if there exists $K' \in \mathcal{H}$ sharing an edge with K^1 and sharing an edge with

K^2. Fix $K \in \mathcal{K}_S$. Since cliques in $\mathcal{H}$ are edge-disjoint, K has common edges with at most $\binom{t}{r}$ elements of $\mathcal{H}$. On the other hand, each clique from $\mathcal{H}$ that has a common edge with K shares edges with at most $\binom{t+1}{r} - 1$ other cliques from $\mathcal{K}_S$. Thus, the maximum degree of G_S is at most $\binom{t}{r}\left(\binom{t+1}{r} - 1\right) = \lambda - 1$. Hence, G_S contains an independent set of size at least X_S/λ. However, since $X_S/\lambda > Z_S$, this corresponds to a family of strictly more than Z_S edge-disjoint copies of hypergraphs from $\mathcal{K}$ having a copy of K_t^r in S. Hence, there exists a copy of K_t^r in $G[S]$ which does not share any edges with elements from $\mathcal{H}$ and thus this copy exists also in $G'[S]$, a contradiction. □

To finish the proof of Lemma A.8 it suffices to prove the following claim.

Claim A.10. ***Whp*** *A_S holds for every $S \subset V(G)$ of size k.*

Proof. Fix $S \subset V(G)$ of size k. We first bound a lower-tail of X_S using Corollary A.4. Next, we bound an upper-tail of Z_S using Lemma A.2. We use the notations of Janson's inequality described in Section A.1.

Let us start with a lower-tail estimate of X_S. Denote by D_S the number of copies of K_t^r in $G[S]$. We have

$$\mu_D := \mathbb{E}[D_S] = \binom{k}{t} p^{\binom{t}{r}},$$

$$\Delta_D := \sum_{\substack{T_1, T_2 \subset S \\ |T_1|=|T_2|=t \\ |T_1 \cap T_2| \geq r}} \mathbb{P}(G[T_1] \cong G[T_2] \cong K_t^r) = \mu_D \sum_{i=r}^{t-1} \binom{t}{i} \binom{k-t}{t-i} p^{\binom{t}{r} - \binom{i}{r}},$$

$$\nu_D = \max_{T \subset S, |T|=t} \sum_{i=r}^{t-1} \binom{t}{i} \binom{k-t}{t-i} p^{\binom{t}{r}}.$$

To make sure we can apply Corollary A.4, we show that $\mu_D = o(1)$ and $\Delta_D = o(\mu_D)$. We have

$$\nu_D = \Theta\left(k^{t-r} p^{\binom{t}{r}}\right) = \Theta\left(n^{\frac{(t-r)\binom{t}{r}(t+1-r)}{\left(\binom{t+1}{r}-1\right)(t-1)} - \frac{\binom{t}{r}(t+1-r)}{\binom{t+1}{r}-1}} \log^{\frac{t-r}{t-1}} n\right)$$
$$= \Theta\left(n^{\frac{\binom{t}{r}(t+1-r)}{\binom{t+1}{r}-1}\left(\frac{t-r}{t-1} - 1\right)} \log^{\frac{t-r}{t-1}} n\right) = o(1).$$

In order to show that $\Delta_D = o(\mu_D)$, we need to show that $k^{t-i} p^{\binom{t}{r} - \binom{i}{r}} = o(1)$ for every integer $r \leq i \leq t-1$. Fix such an integer i. Then,

$$k^{t-i} p^{\binom{t}{r} - \binom{i}{r}} = n^{\frac{(t-i)\binom{t}{r}(t+1-r)}{\left(\binom{t+1}{r}-1\right)(t-1)} - \left(\binom{t}{r} - \binom{i}{r}\right)\frac{t+1-r}{\binom{t+1}{r}-1}} \log^{\frac{t-i}{t-1}} n$$
$$= n^{\frac{t+1-r}{\binom{t+1}{r}-1}\left(\frac{(t-i)\binom{t}{r}}{t-1} - \binom{t}{r} + \binom{i}{r}\right)} \log^{\frac{t-i}{t-1}} n$$
$$= n^{\frac{t+1-r}{\binom{t+1}{r}-1}\left(-\frac{i-1}{t-1}\binom{t}{r} + \binom{i}{r}\right)} \log^{\frac{t-i}{t-1}} n = o(1),$$

where the last equality is true since $\binom{t}{r} > \binom{i}{r} \cdot \frac{t-1}{i-1}$. Indeed,

$$\binom{t}{r} > \binom{i}{r} \cdot \frac{t-1}{i-1} \iff \frac{t!}{(t-r)!} > \frac{i!}{(i-r)!} \cdot \frac{t-1}{i-1} \iff \frac{t(t-1) \cdot \ldots \cdot (t-r+1)}{i(i-1) \cdot \ldots \cdot (i-r+1)} > \frac{t-1}{i-1},$$

and the last inequality is true since we have $t > i$ and $r \geq 2$ and thus

$$\frac{t(t-1)\cdot\ldots\cdot(t-r+1)}{i(i-1)\cdot\ldots\cdot(i-r+1)} \geq \frac{t}{i}\cdot\frac{t-1}{i-1} > \frac{t-1}{i-1}.$$

Therefore, by Corollary A.4, for every $\epsilon \in (0,1)$,

$$\mathbb{P}(X_S < (1-\epsilon)\mu_D) \leq e^{-\epsilon^2\mu_D/3}. \tag{3}$$

Next, we analyse the upper-tail of Z_S. We will denote $\mathbb{E}[Z_S]$ and $\mathbb{E}[Y_S]$ by μ_Z and μ_Y, respectively. Clearly, $Z_S \leq Y_S$ and thus $\mu_Z \leq \mu_Y$. Thus, by Lemma A.2,

$$\mathbb{P}(Z_S \geq 5\mu_Y) \leq \frac{\mu_Z^{5\mu_Y}}{(5\mu_Y)!} \leq \left(\frac{e\mu_Z}{5\mu_Y}\right)^{5\mu_Y} \leq e^{-(\log 5-1)5\mu_Y}. \tag{4}$$

Calculating μ_Y,

$$\mu_Y = \Theta\left(k^t \sum_{\ell=r}^{t} n^{t-\ell+1} p^{\binom{t+1}{r}+\binom{t}{r}-\binom{\ell}{r}}\right).$$

For every $t \geq \ell > r$,

$$\begin{aligned}
&k^t n^{t-r+1} p^{\binom{t+1}{r}+\binom{t}{r}-\binom{r}{r}} \gg k^t n^{t-\ell+1} p^{\binom{t+1}{r}+\binom{t}{r}-\binom{\ell}{r}} \\
&\iff n^{\ell-r} \gg p^{1-\binom{\ell}{r}} \\
&\iff \ell - r > -\left(1-\binom{\ell}{r}\right)\frac{t+1-r}{\binom{t+1}{r}-1} \\
&\iff \frac{\ell-r}{\binom{\ell}{r}-1} > \frac{t+1-r}{\binom{t+1}{r}-1}.
\end{aligned}$$

The last inequality is true since the function $f(x) = \frac{x-r}{\binom{x}{r}-1}$ is decreasing when $x > r$. Thus, there exist positive constants c_2 and c_3, which do not depend on c_0 and c_1, such that

$$c_2\binom{k}{t} n^{t-r+1} p^{\binom{t}{r}+\binom{t+1}{r}-1} \leq \mu_Y \leq c_3\binom{k}{t} n^{t-r+1} p^{\binom{t}{r}+\binom{t+1}{r}-1}.$$

Let us choose c_1 such that $c_3 n^{t-r+1} p^{\binom{t+1}{r}-1} = \frac{1}{10\lambda}$. Recall that $\mu_D = \binom{k}{t}p^{\binom{t}{r}}$. We have

$$\frac{10\lambda c_3}{c_2} \geq \frac{\mu_D}{\mu_Y} \geq 10\lambda.$$

Notice that if $X_S > 0.5\mu_D$ and $Z_S < \mu_D/(2\lambda)$, then $X_S > \lambda Z_S$ and thus A_S holds. Therefore, the probability that A_S does not hold is at most

$$\begin{aligned}
\mathbb{P}(X_S \leq 0.5\mu_D) + \mathbb{P}(Z_S \geq \mu_D/(2\lambda)) &\leq \mathbb{P}(X_S \leq 0.5\mu_D) + \mathbb{P}(Z_S \geq 5\mu_Y) \\
&\leq e^{-\mu_D/12} + e^{-5\mu_D(\log 5-1)c_2/10\lambda c_3} \leq e^{-c_4\mu_D},
\end{aligned}$$

for some constant c_4 which does not depend on c_0 and c_1. The second inequality is true by (3) and (4). We have

$$\begin{aligned}
\mu_D &\geq \left(\frac{k}{t}\right)^t p^{\binom{t}{r}} \geq \left(\frac{c_0}{t}\right)^t n^{\frac{t}{t-1}\cdot\frac{\binom{t}{r}(t+1-r)}{\binom{t+1}{r}-1}} \log^{\frac{t}{t-1}} n \cdot c_1^t n^{-\frac{\binom{t}{r}(t+1-r)}{\binom{t+1}{r}-1}} \\
&= \left(\frac{c_0 c_1}{t}\right)^t n^{\frac{\binom{t}{r}(t+1-r)}{(\binom{t+1}{r}-1)(t-1)}} \log^{\frac{t}{t-1}} n.
\end{aligned}$$

Hence,

$$\mathbb{P}(A_S) \le \exp\left(-c_4\left(\frac{c_0c_1}{t}\right)^t n^{\frac{\binom{t}{r}(t+1-r)}{(\binom{t+1}{r}-1)(t-1)}} \log^{\frac{t}{t-1}} n\right).$$

By the union bound over all choices of S, the probability that there exists S such that A_S does not hold is at most

$$\begin{aligned}\mathbb{P}\left(\bigcup_S A_S\right) &\le \binom{n}{k}\exp\left(-c_4\left(\frac{c_0c_1}{t}\right)^t n^{\frac{\binom{t}{r}(t+1-r)}{(\binom{t+1}{r}-1)(t-1)}} \log^{\frac{t}{t-1}} n\right)\\ &\le \exp\left(\left\lceil c_0 n^{\frac{\binom{t}{r}(t+1-r)}{(\binom{t+1}{r}-1)(t-1)}} \log^{\frac{t}{t-1}} n\right\rceil - c_4\left(\frac{c_0c_1}{t}\right)^t n^{\frac{\binom{t}{r}(t+1-r)}{(\binom{t+1}{r}-1)(t-1)}} \log^{\frac{t}{t-1}} n\right)\\ &= o(1),\end{aligned}$$

where the last equality is true if c_0 and c_1 are sufficiently large. □

The proof of Lemma A.8 is complete. □

A.3. Proof of the upper bound of Theorem 1(b)

Set $G := G^r(n,p)$. Set $\alpha = \frac{1}{1-p^{r-1}}$ and $\beta = \frac{1}{1-p^{\binom{s}{r}-\binom{s-r}{r}-1}}$. Unless stated otherwise, then the base of all the logs in the proof is α. Set

$$a_1 = \frac{1}{p}\log n\left(1+\frac{3}{\log\log n}\right), \quad a_2 = s\log_\beta(n^r) \quad \text{and} \quad a_3 = \frac{a_2}{(\log a_2)^{1/r}}.$$

Let A_1, A_2, A_3 be disjoint subsets of $V(G)$ of sizes a_1, a_2, a_3, respectively. Let B denote the set of the remaining vertices of $V(G)$. Set $t = s - r$ and

$$q = c_1 a_1^{-\frac{t+1-r}{\binom{t+1}{r}-1}}.$$

If $t < r$, let H_1 be the empty hypergraph on A_1. Otherwise, if $t \ge r$, let $H_1 \subseteq G[A_1]$ be the following spanning subhypergraph, which by Lemmas A.6 and A.8 **whp** exists:

- H_1 is K^r_{t+1}-free.
- There exists a copy of K^r_t in every induced subhypergraph of H_1 of size at least

$$T = c_0 a_1^{\frac{\binom{t}{r}(t+1-r)}{(\binom{t+1}{r}-1)(t-1)}} \log^{\frac{1}{t-1}} a_1.$$

- Every two vertices in A_1 are contained in at most $2\binom{a_1}{t-2} q^{\binom{t}{r}}$ copies of K^r_t in H_1.

Let H be a subhypergraph of G with the following edges:

- All edges of H_1.
- All edges of G of the form $A' \cup B'$ where $A' \subseteq A_1$ with $1 \le |A'| \le t$ and $B' \subseteq B$ with $0 < |B'| = s - |A'|$.

We will say that an edge $e \in E(G) \setminus E(H)$ can be *completed* if it closes a copy of K_s^r together with $E(H)$. First, observe that there are no copies of K_s^r in H. Indeed, consider a set X of s vertices from $A_1 \cup B$, and suppose towards contradiction that X induces a clique in H. Since $H[B]$ is an empty hypergraph, we have that $|X \cap B| \leq r-1$, that is, $|X \cap A_1| \geq s-(r-1) = t+1$. Thus, if $t+1 \geq r$, since $H[A_1]$ is K_{t+1}^r-free we have that H is K_s^r-free. Otherwise, if $t+1 \leq r-1$, we note that there are no edges with $t+1$ vertices from A_1 and $r-(t+1)$ vertices from B.

We will now show that most of the edges in $G[B]$ can be completed in H. Given an $(r-1)$-set of vertices $S \subset V$, the *neighbourhood* of S in $X \subset V$ is

$$N_X(S) \coloneqq \{v \in X \colon \{v\} \cup S \in E(G)\}.$$

Furthermore, given an $(r-1)$-set of vertices $S \subset B$, we say that S is *good* if

$$|N_{A_1}(S)| \geq m \coloneqq \left(1 + \frac{2}{\log\log n}\right)\log n,$$

and otherwise we say that S is *bad*. Moreover, we say that an edge $e \in G[B]$ is *good* if there exists at least one good $(r-1)$-subset $S \subset e$ and otherwise we say that e is *bad*.

Claim A.11. ***Whp****, all good edges in $G[B]$ but $o(n^{r-1}\log n)$ can be completed.*

Proof. Fix an $(r-1)$-set $S \subset B$ and expose all edges of the form $\{v\} \cup S$, $v \in A_1$. Let us assume that these edges justify the validity of the event that S is good. Fix an r-subset $e \subset B$ containing S, noting that we then have that e is good. We will show that the probability that e cannot be completed is less than $\frac{2}{n}$. If so, then the expectation of the number of good edges in $G[B]$ which cannot be completed is $O(n^{r-1})$ and then, by Markov's inequality, there are **whp** $o(n^{r-1}\log n)$ such edges.

Since S is good, the size of the neighbourhood of S in A_1, denoted by $N_{A_1}(S)$, is at least m. We will give an upper bound to the probability that there is no copy of $K_t^r \subseteq H[N_{A_1}(S)]$ which closes a copy of K_s^r together with e. Notice that, for a given vertex $v \in N_{A_1}(S)$, the probability that $\{v\} \cup S' \in E(G)$ for all $S' \subset e$ of size $r-1$ is exactly p^{r-1}. Define

$$N_{S,e} \coloneqq \{v \in N_{A_1}(S) \colon \{v\} \cup S' \in E(G) \text{ for all } S' \subset e,\ |S'| = r-1\}.$$

Then, given S and $N_{A_1}(S)$, $|N_{S,e}| \sim \mathrm{Bin}(|N_{A_1}(S)|, p^{r-1})$. By Claim A.5,

$$\mathbb{P}\left(|N_{S,e}| \geq \frac{|N_{A_1}(S)|}{\log^2 |N_{A_1}(S)|}\right) \geq 1 - (1-p^{r-1})^{|N_{A_1}(S)| - \frac{|N_{A_1}(S)|}{\log |N_{A_1}(S)|}} \geq 1 - \frac{1}{n},$$

where the last inequality holds since

$$|N_{A_1}(S)| - \frac{|N_{A_1}(S)|}{\log |N_{A_1}(S)|} \geq m\left(1 - \frac{1}{\log\log n}\right) > \log n.$$

Let us expose all edges of the form $\{v\} \cup S'$, $S' \subset e$, $v \in N_{A_1}(S)$, and let us suppose from now that $|N_{S,e}| \geq \frac{|N_{A_1}(S)|}{\log^2 |N_{A_1}(S)|}$ and thus $|N_{S,e}| > a_1^{1-\epsilon}$ for some small constant $\epsilon > 0$. Denote the family of copies of K_t^r in $H[N_{S,e}]$ by $\mathcal{K}$. Since in every induced subhypergraph of $H[N_{S,e}]$ of size T there is a copy of K_t^r and every copy of K_t^r appears in at most $\binom{|N_{S,e}|-t}{T-t}$ subsets of size T, we have

$$|\mathcal{K}| \geq \frac{\binom{|N_{S,e}|}{T}}{\binom{|N_{S,e}|-t}{T-t}} > \left(\frac{|N_{S,e}|}{T}\right)^t = \Omega\left(a_1^{t - t\epsilon - \frac{t\binom{t}{r}(t+1-r)}{\left(\binom{t+1}{r}-1\right)(t-1)}} \log^{-\frac{t}{t-1}} a_1\right).$$

By Claim A.7,

$$\left(\frac{|N_{S,e}|}{T}\right)^t = \Omega\left(a_1^{r-1-t\epsilon}\log^{-\frac{t}{t-1}} a_1\right) \geq a_1^{1+\epsilon_1},$$

where the last inequality is true for some small enough $\epsilon_1 > 0$.

Denote by F the number of pairs $(K, K') \in \mathcal{K}$ satisfying $|V(K) \cap V(K')| \geq 2$. We have that $F = O\left(|\mathcal{K}|a^{t-2}q^{\binom{t}{r}}\right)$. Note that $\frac{1}{T^t} = \Omega\left(\frac{q^{\binom{t}{r}}\log a_1}{T}\right)$ and thus

$$|\mathcal{K}| = \Omega\left(a_1^{t-t\epsilon}q^{\binom{t}{r}}T^{-1}\log a_1\right) \geq a_1^{t-2t\epsilon}q^{\binom{t}{r}}T^{-1}.$$

Then,

$$\frac{|\mathcal{K}|^2}{F} = \Omega\left(\frac{|\mathcal{K}|}{a_1^{t-2}q^{\binom{t}{r}}}\right) = \Omega\left(\frac{a_1^{t-2t\epsilon}}{Ta_1^{t-2}}\right).$$

By Claim A.7, $T \leq a_1^{1-\frac{1}{t}}$. Hence,

$$\frac{|\mathcal{K}|^2}{F} = \Omega\left(a_1^{2-2t\epsilon-1+\frac{1}{t}}\right) \geq a_1^{1+\epsilon_2}, \tag{5}$$

for some $\epsilon_2 > 0$.

Recall that we want to give an upper bound for the probability that e cannot be completed. For every $K \in \mathcal{K}$, denote by B_K the event that $G[K \cup e] \cong K_s$. Note that, if B_K holds, then e can be completed. We will give an upper bound for $\mathbb{P}\left(\cap_{K\in\mathcal{K}}B_K^c\right)$ using Theorem A.1. For every $K \in \mathcal{K}$, we already have the edges induced by K and edges with one vertex from K and $r-1$ vertices from e. Therefore, two events B_K and $B_{K'}$ are dependent only if $|K \cap K'| \geq 2$. Moreover, the probability that B_K holds is at most $p^{\binom{s}{r}} = O(1)$. We have,

$$\mu := \sum_{K\in\mathcal{K}} \mathbb{P}(B_k) = \Theta(|\mathcal{K}|),$$

$$\Delta := \sum_{\substack{K,K'\in\mathcal{K}\\ |V(K)\cap V(K')|\geq 2}} \mathbb{P}(B_K \cap B_{K'}) < F.$$

Hence, by Theorem A.1,

$$\mathbb{P}\left(\cap_{K\in\mathcal{K}}B_K^c\right) \leq e^{-\mu^2/2\Delta} \leq e^{-\Theta(|\mathcal{K}|^2/F)} \leq e^{-a_1^{1+\epsilon_3}} < \frac{1}{n},$$

where the third inequality is true for some $0 < \epsilon_3 < \epsilon_2$ by (5). Therefore, the expected number of uncompleted good edges in $G[B]$ is $O(n^{r-1})$ and thus, by Markov's inequality, **whp** there are $o(n^{r-1}\log n)$ uncompleted good edges in $G[B]$. □

Before dealing with bad edges in $G[B]$, the following claim gives an upper bound to the number of bad $(r-1)$-sets in B.

Claim A.12. *There are at most $\frac{n^{r-1}}{\log n}$ bad $(r-1)$-subsets $S \subset B$.*

Proof. Fix an $(r-1)$-subset $S \subset B$. Recall that

$$N_{A_1}(S) = \{v \in A_1 \colon \{v\} \cup S \in E(G)\}.$$

Note that $|N_{A_1}(S)| \sim \text{Bin}(a_1, p)$. By Claim A.5,

$$\mathbb{P}(|N_{A_1}(S)| < m) = \mathbb{P}\left(|N_{A_1}(S)| < a_1p - \frac{\log n}{\log\log n}\right) \leq e^{-\log n/4\log\log n} = o(1/\log n).$$

Hence, the claim is true by Markov's inequality. □

We will use A_2 to complete bad edges in $G[B]$. By Lemma A.8, there exists a spanning subhypergraph $H_2 \subseteq G[A_2]$ which is K_{t+1}^r-free and, for some sufficiently small δ with respect to t, every subset of $a_2^{1-\delta}$ vertices induces a copy of K_t^r. Observe that there are at least

$$\frac{a_2}{t} - a_2^{1-\delta} \geq \frac{a_2}{s} = \log_\beta(n^r)$$

vertex-disjoint copies of K_t^r in $H_2[A_2]$.

Add to H the following edges:

- edges from H_2,
- edges intersecting both A_2 and B in the graph G with at least 2 and at most t vertices from A_2,
- edges of the form $S \cup \{v\}$ where $S \subseteq B$ is a bad $(r-1)$-subset and $v \in A_2$.

Before showing how to complete bad edges utilising these edges, let us first note that there are still no copies of K_s^r in H. As we only added edges induced by $A_2 \cup B$, and there are no edges intersecting both A_1 and A_2, we may consider a set $X \subseteq A_2 \cup B$ of s vertices, and suppose towards contradiction that X induces a clique in H. Since $H[B]$ is empty, we have that $|X \cap B| \leq r-1$ and $|X \cap A_2| \geq s-r+1 = t+1$. Similar to before, if $t+1 \geq r$, we have a contradiction since $H[A_2]$ is K_{t+1}^r-free, and, if $t+1 \leq r-1$, then note that we have not added to H edges with $s-r+1 = t+1$ vertices from A_2 and the rest from B, and we once again obtain a contradiction.

Let us now consider some bad edge e in $G[B]$. Given a copy of $K_t^r \subseteq H[A_2]$, the probability that this copy fails to complete e is exactly $1/\beta$. Since we have at least $\log_\beta(n^r)$ vertex-disjoint copies of K_t^r in $H[A_2]$, the probability that e cannot be completed is at most n^{-r}. By Claim A.12, the number of bad edges in B is at most $\frac{n^{r-1}}{\log n} \cdot n$. Therefore, the expectation of the number of bad edges that cannot be completed is at most $1/\log n$. By Markov's inequality, **whp** all the bad edges in $G[B]$ can be completed.

Next, we consider edges of the form $S \cup \{v\}$ where S is a good $(r-1)$-set in B and $v \in A_2$. We have $\Theta(n^{r-1}\log n)$ edges of this form, so we are not allowed to add them to H; thus, we will have to complete most of them. We will use A_3 for that.

Once again, **whp** there exists a spanning subhypergraph $H_3 \subset G[A_3]$ which is K_{t+1}^r-free and such that every set of $a_3^{1-\delta}$ vertices in H_3, for some small constant $\delta > 0$, contains a copy of K_t^r. Let us add to H the following edges:

- edges from H_3,
- edges intersecting both B and A_3 in the graph G with at least two and at most t vertices from A_3,
- edges of the form $S \cup \{v\}$ where $S \subseteq B$ is a good $(r-1)$-subset and $v \in A_3$,
- edges intersecting both $B \cup A_2$ and A_3 in the graph G with exactly one vertex from A_2 and at most t vertices from A_3.

Let us first show that H is still K_s^r-free. Since no edges are intersecting both A_1 and $A_2 \cup A_3$, it suffices to consider $X \subseteq B \cup A_2 \cup A_3$ of size s, and suppose towards contradiction that X induces a clique in H. Once again, we note that $|X \cap B| \leq r-1$. Now, if there are no vertices from A_3, we are done by the previous arguments, as we did not add new edges which do not intersect A_3. If there are at least $t+1$ vertices from A_3, we are done by similar arguments to before. Furthermore, if there are at least two vertices in A_2, note that there are no edges in H containing two vertices from A_2 and at least one vertex from A_3, leading to a contradiction.

Thus, we are left with the case where $|X \cap A_2| = 1, |X \cap A_3| = t, |X \cap B| = r-1$. Now, if $X \cap B$ is bad, there are no edges of the form $(X \cap B) \cup \{v\}$, $v \in A_3$, in H, and if it is good, then there are no edges of the form $(X \cap B) \cup \{v\}$, $v \in A_2$, in H — either way, we get that X cannot induce a clique.

Let us now consider an edge $e \in G$ of the form $S \cup \{v\}$ where S is a good $(r-1)$-set in B and $v \in A_2$. As before, there are at least a_3/s vertex-disjoint copies of K_t^r in $H[A_3]$. The probability that a given copy fails to complete e is exactly $1/\beta$. Thus, the probability that e cannot be completed is at most $\beta^{-a_3/s}$. Therefore, the expectation of such edges e that cannot be completed is at most

$$n^{r-1} a_2 \beta^{-a_3/s} \ll n^{r-1} \log n,$$

and thus, by Markov's inequality, **whp** all but at most $o(n^{r-1} \log n)$ edges of the form $S \cup \{v\}$, where S is a good $(r-1)$-set in B and $v \in A_2$, are completed.

Let us count the edges in H:

- Edges intersecting both B and A_1. There are $(1+o(1))\binom{n}{r-1} \log n$ such edges.
- Edges intersecting both B and A_2. There are $o(n^{r-1})$ such edges.
- Edges intersecting both $B \cup A_2$ and A_3. There are $o(n^{r-1} \log n)$ such edges.
- Some of the edges induced by $A_1 \cup A_2 \cup A_3$, of which there are at most $o(n)$.

Note that we have completed all edges except for the following: we did not complete edges of the form $S \cup \{v\}$ where $S \subseteq B$ is bad and $v \in A_3$ as well as some of the edges induced by $A_1 \cup A_2 \cup A_3$. Furthermore, if $t < r$, we have not completed edges intersecting both B and $A_1 \cup A_2 \cup A_3$ with more than t vertices in either A_1, A_2, or A_3. However, there are at most $o(n^{r-1} \log n)$ edges of these types, and we are thus done by Observation 4.3. Hence, H has $(1+o(1))\binom{n}{r-1} \log n$ edges and $o(n^{r-1} \log n)$ additional edges make it saturated.