Isomorphic edge disjoint subgraphs of hypergraphs

Paul Horn[*] Václav Koubek[†] Vojtěch Rödl[‡]

September 20, 2012

Abstract

We show that any k-uniform hypergraph with n edges contains two edge disjoint subgraphs of size $\tilde{\Omega}(n^{2/(k+1)})$ for $k = 4, 5$ and 6. This is best possible up to a logarithmic factor due to an upper bound construction of Erdős, Pach, and Pyber who show there exist k-uniform hypergraphs with n edges and with no two edge disjoint isomorphic subgraphs with size larger than $\tilde{O}(n^{2/(k+1)})$. Furthermore, our result extends results Erdős, Pach and Pyber who also established the lower bound for $k = 2$ (eg. for graphs), and of Gould and Rödl who established the result for $k = 3$.

Suppose G is an arbitrary k-uniform hypergraph with n edges. Let $\iota_k(G)$ denote the size of the largest pair of edge disjoint isomorphic subgraphs of G. Let

$$\iota_k(n) = \min_{\substack{G \\ |E(G)|=n}} \iota_k(G),$$

denote the largest size so that *every* k-uniform hypergraph with n edges contains two edge disjoint subgraphs of size at least $\iota_k(n)$. A natural question is to study the size of $\iota_k(n)$ for different values of k. This was first undertaken in [2] by Erdős, Pach and Pyber, who attribute the problem to Schönheim, and who proved that there exists constants C_k and C'_k depending only on k such that

$$C_k n^{2/(2k-1)} \le \iota_k(n) \le C'_k n^{2/(k+1)} \frac{\log n}{\log \log n}.$$

For $k = 2$, these are roughly of the same order and this result implies that every simple graph with n edges contains two edge disjoint isomorphic subgraphs of size $n^{2/3}$. The logarithmic gap was settled very recently in [5], who showed that $\iota_2(m) = \Theta((m \log m)^{2/3})$. Erdős, Pach, and Pyber also raised the question of narrowing the gap between the lower and upper bound for hypergraphs. In [4], Gould and Rödl proved that, up to a logarithmic factor, the upper bound is correct for $k = 3$ as well. That is, they showed $\iota_3(n) \ge C\sqrt{n}$.

The main result of this paper is to establish that the upper bound is correct, again up to logarithmic factors, for $k = 4, 5$, and 6. In particular we prove:

[*]Department of Mathematics, Harvard University and Department of Mathematics and Computer Science, Emory University, phorn@math.harvard.edu

[†]Department of Theoretical Computer Science and Mathematical Logic, Faculty of Mathematics and Physics, Charles University, koubek@ktiml.mff.cuni.cz

[‡]Department of Mathematics and Computer Science, Emory University, rodl@mathcs.emory.edu. This research partially supported by an Emory University research grant and by NSF grant DMS 080070

Theorem 1. *There exist constants c_4, c_5 and c_6 so that*

$$\iota_4(n) \geq c_4 n^{2/5},$$
$$\iota_5(n) \geq c_5 \frac{n^{1/3}}{\log n}, \text{ and}$$
$$\iota_6(n) \geq c_6 \frac{n^{2/7}}{\log^{35} n}.$$

Unfortunately as k increases, the problem seems to become harder still and $k = 7$ has difficulties that are not present when $k = 6$. However it still seems likely that upper bound should be essentially correct. The lower bound argument of Erdős, Pach and Pyber is inductive and in combination with Theorem 1 implies that

Corollary 1. *For all values of $k \geq 7$*

$$\iota_k(n) = \tilde{\Omega}(n^{\frac{2}{2k-5}}).$$

However our method allows us to improve the general lower bound of Erdős, Pach and Pyber while sidestepping the technical issues that prevent us from matching the upper bound in general.

Theorem 2. *For all values of $k \geq 7$*

$$\iota_k(n) = \tilde{\Omega}\left(n^{\frac{2}{2k-\log_2 k}}\right).$$

Of course, the real challenge is to establish that the upper bound is correct for all values of k. It would be interesting to show

$$\limsup_{k,n\to\infty} k\frac{\log \iota_k(n)}{\log n} = 2.$$

As a start, it would be interesting to show that instead one may that this is larger than $1 + \epsilon$ for some $\epsilon > 0$.

The remainder of the paper is organized as follows. In Section 1, we establish some preliminary lemmas which hold for all values of k. In Section 2, we establish Theorem 1. In Section 3, we establish Theorem 2.

1 Notation and Preliminaries

Since any k-uniform hypergraph with n edges contains a k-partite subgraph of size at least $\frac{n}{k!}$ we work throughout with k-uniform, k-partite hypergraphs. G will denote a k-uniform, k-partite hypergraph on vertex sets $V_1, \ldots, V_k$ and edge set E. Several times in the paper we make additional assumptions about G, which we show that we can make if Theorem 1 were not to hold.

Suppose $S \subseteq [k] = \{1, 2, \ldots, k\}$. For $e \in E$, we let $e^S = e \cap (\bigcup_{i\in S} V_i)$ denote the restriction of e to S. For $S \subset [k]$, an S-tuple is an $|S|$-tuple of vertices $\{v_i : v_i \in V_i, \forall i \in S\}$. If T is an S-tuple of vertices of G, we define the neighborhood of T to be

$$E(T) = \{e \in E : T \subseteq e\}.$$

For a set $S' \subseteq [k]$, and S-tuple T we define the S'-neighborhood to be the multiset

$$E^{S'}(T) = \{e^{S'} : e \in E(T)\},$$

the restriction of the edges in $E(T)$ to S'. Again, we emphasize that we view $E^{S'}(T)$ as a multiset, so $|E^{S'}(T)| = |E(T)|$.

Most often we care about the neighborhood of a single vertex and as a slight abuse of notation we define $E(v)$ to be $E(\{v\})$. Likewise, we are most often interested in the case where $S' = [i]$ $(= \{1, 2, \ldots, i\})$ for some $i \leq k$. Again as a slight abuse of notation, we define $E^i(T) = E^{[i]}(T)$ and $E^i = E^{[i]}$ to be a neighborhood or edge set restricted to the first i levels.

In the work that follows, we frequently study the action of permutations on the vertex sets, which induces an action on the edge set. Throughout, $\pi^j = (\pi_1, \ldots, \pi_j)$ will represent a collection of j permutations so that π_ℓ acts on V_ℓ. Suppose E_1 and E_2 are the edge sets of two k-uniform hypergraphs on $V_1, \ldots, V_k$ and π^S acts on the V_i such that $i \in S$. Then we define

$$\mathcal{I}_\pi^S(E_1, E_2) = |\{(e, f) \in E_1 \times E_2 : e \neq f \text{ and } \pi(e^S) = f^S\}|.$$

Note that this differs from simply $|\pi^S(E_1^S) \cap E_2^S|$; for instance if $E_2 = E_1$ and π acts identically on $V_1, \ldots V_k$, then $\mathcal{I}_\pi^{[k]}(E_1, E_2) = 0$. When $S = [j]$ we use the notation $\mathcal{I}_\pi^j(E, E)$, and when $j = k$ we suppress j entirely from the notation, that is we simply use $\mathcal{I}_\pi(E, E)$.

The relationship between studying the action of permutations on the vertex sets of a hypergraph and the problem at hand is encapsulated in the following.

Lemma 1. *Suppose $G = (V, E)$ is a k-uniform, k-partite hypergraph on $V_1, \ldots, V_k$ and π^k is a permutation of $V_1, \ldots, V_k$. Then G contains two edge disjoint subgraphs of size at least*

$$\frac{\mathcal{I}_{\pi^k}(E, E)}{3}.$$

Proof. We construct a directed graph Γ_π on vertex set $E(G)$, so that for any two distinct hyperedges $e, f \in E(G)$ we place the directed edge $e \to f$ if $\pi^k(e) = f$. Note that $|E(\Gamma_\pi)| = \mathcal{I}_{\pi^k}(E, E)$. Since $E(\Gamma_\pi)$ is defined by the permutation π^k, the in and out degrees of any vertex in Γ_{π^k} is at most one; that is Γ_π decomposes into directed cycles and paths. We thus may choose a matching from Γ_{π^k} of size at least $\frac{\mathcal{I}_{\pi^k}(E,E)}{3}$. Let $\mathcal{M}$ denote this matching; let $E_1 \subseteq E$ denote the set of all start vertices of edges in $\mathcal{M}$ and $E_2 \subseteq E$ denote the set of all end vertices. Then the hypergraphs (V, E_1) and (V, E_2) are edge disjoint, because $\mathcal{M}$ is a matching, isomorphic since $\pi^k(E_1) = E_2$, and of the desired size. ☐

Our goal, then will be to show we can find a permutation which yields $\mathcal{I}_\pi^k(E, E)$ of the proper size. We begin with an easy case, which follows from a first moment argument.

Lemma 2. *Suppose $G = (V, E)$ is a k-uniform, k-partite hypergraph on $V_1, \ldots, V_k$ with $|V_1| \leq \cdots \leq |V_k| \leq n^{\frac{2}{k+1}}$. If $|E| = Cn$ for some $C > 0$, then G contains two edge disjoint isomorphic subgraphs of size at least*

$$\left(\frac{C^2}{3} - o(1)\right) n^{2/(k+1)}.$$

Proof. Choose $\pi_1, \ldots, \pi_k$ to be uniformly random permutations of $V_1, \ldots, V_k$ respectively. Let $\pi^k = (\pi_1, \ldots, \pi_k)$. For an edge $e \in E(G)$, let X_e denote the Bernoulli random variable that is 1 if $\pi(e) = e'$ for some $e \neq e' \in E(G)$. Then

$$\mathbb{E}[X_e] = \frac{Cn - 1}{n^{2k/(k+1)}}.$$

By linearity of expectation,

$$\mathbb{E}[\mathcal{I}_{\pi^k}(E, E)] = \frac{Cn(Cn - 1)}{n^{2k/(k+1)}} = (C^2 - o(1)) n^{2/(k+1)}.$$

There thus exists a π^k with $\mathcal{I}_{\pi^k}(E, E)$ at least so large and an application of Lemma 1 completes the proof. ☐

Of course for a general k-uniform hypergraph with n edges, it is too much to assume that all partite sets are smaller than $n^{2/(k+1)}$. Our next observation, however, is that we may assume *most* of partite sets are so small.

Lemma 3. *Suppose $G = (V, E)$ is a k-uniform, k-partite hypergraph with n edges. Then either:*

1. *G contains two edge disjoint, isomorphic subgraph of size $\frac{1}{2}n^{2/(k+1)}$; or*
2. *There exists an $\ell \geq \lceil \frac{k+1}{2} \rceil$ and a set $E' \subseteq E(G)$ of size at least $\frac{1}{(3k)^\ell}|E|$ so that, possibly after reordering the partite sets,*
   (a) *E' is supported on sets $V'_1, \ldots, V'_\ell, V_{\ell+1}, \ldots, V_k$ with $|V'_1|, \ldots, |V'_\ell| \leq n^{2/(k+1)}$; and*
   (b) *There exists a set $E'' \subseteq E'$ of size at least $\frac{|E'|}{3}$ so that every $[\ell]$-tuple T consisting of one vertex each from $V'_1, \ldots, V'_\ell$ has $|E''(T)| \leq 1$.*

Proof. We assume that G does not contain two edge disjoint, isomorphic subgraphs of size $\frac{1}{2}n^{2/(k+1)}$. Choose $\ell \geq 0$ as large as possible so that $2(a)$ holds. That is, there exist sets $V'_1, \ldots, V'_\ell, V_{\ell+1}, \ldots, V_k$ with $|V'_1|, \ldots, |V'_\ell| \leq n^{2/(k+1)}$ and a set $E' \subseteq E$ with $|E'| \geq \frac{1}{(3k)^\ell}|E|$ supported on these sets. Note that a priori we may even have $\ell = 0$; but we will show that we satisfy that $\ell \geq \lceil \frac{k+1}{2} \rceil$.

Note that condition $2(b)$ cannot be satisfied unless $\ell \geq \lceil \frac{k+1}{2} \rceil$. Indeed, if $\ell < \frac{k+1}{2}$, then there are $o(n)$ $[\ell]$-tuples with one vertex from each of $V_1, \ldots, V_\ell$. But then any set E'' of $\frac{n}{3^{\ell+1}k^\ell}$ edges from E with the hypothesized support includes some $[\ell]$-tuple T with $E''(T) > 1$.

It remains to show that if $2(b)$ is not satisfied the maximality of ℓ is contradicted. Suppose $2(b)$ does not hold. Let $S = [k] \setminus [\ell]$. We greedily select pairs of distinct edges $\{e_1, f_1\}, \{e_2, f_2\}, \ldots$ so that $e_i^\ell = f_i^\ell$, but and e_i^S and f_i^S do not intersect $\bigcup_{j=1}^{i-1}(e_j \cup f_j)$. Note that we do not require that $e_i^S \cap f_i^S = \emptyset$. Observe that this process of selecting edges must halt with some $\{e_t, f_t\}$ where $t \leq \frac{1}{2}n^{2/(k+1)}$, as $E_1 = \{e_j : j \leq t\}$ and $E_2 = \{f_j : j \leq t\}$ induce edge disjoint isomorphic subgraphs of size t.

For $s = \ell + 1, \ldots, k$, we set $W_s = \bigcup_{i=1}^{t}(e_i \cup f_i) \cap V_s$. That is, W_s is the set of vertices in V_s that lie in some e_i^j or f_i^j. Note that $|W_s| \leq 2t \leq n^{2/(k+1)}$. Observe that for every $[\ell]$-tuple T of vertices from $V'_1, \ldots, V'_\ell$ at most one edge from $E'(T)$ does not intersect some W_s, as if two edges avoided all W_s the maximality of t is contradicted. Since we assume $2(b)$ is not satisfied, at most $\frac{|E'|}{3}$ edges lie in tuples T with $|E'(T)| \leq 1$. Therefore, at least $\frac{2|E'|}{3} \geq \frac{2n}{3(3k)^\ell}$ edges lie in $E'(T)$ for a tuple T with $|E'(T)| \geq 2$. For any of these tuples, at most one avoids intersecting some $\bigcup_{s=\ell+1}^{k} W_s$. Therefore $\frac{n}{3(3k)^\ell}$ edges intersect $\bigcup_{s=\ell+1}^{k} W_s$. Since there are only $k - \ell$ such sets, at least

$$\frac{n}{3(3k)^\ell} \cdot \frac{1}{k - \ell} \geq \frac{n}{3(3k)^{\ell+1}}$$

edges in E' all intersect the *same* W_s, which we denote W. Since $|W| \leq n^{2/(k+1)}$, taking $V'_{\ell+1} = W$ and E'' to be the set of all edges in E which intersect all of $V'_1, \ldots, V'_{\ell+1}$ (possibly reordering partite sets) contradicts the maximality of ℓ, completing the proof of the lemma. ☐

The essential idea of the proof of Theorem 1 is to show that, by permuting randomly on the levels with $|V_i| \leq n^{2/(k+1)}$, we can construct permutations on the larger levels so that, in total, enough edges are mapped to other edges, and then apply Lemma 1. In order to perform this task we require not only many small levels, as guaranteed by Lemma 3, but we must understand some structural properties of our edge set. The following lemma allows us to sacrifice a proportion of the edges to guarantee nice structural properties.

Lemma 4. *Suppose that $G = (V, E)$ is a k-uniform, k-partite hypergraph on $V_1, \ldots, V_k$ with Cn edges. Fix $S \subseteq [k]$ and define $\alpha = \alpha_S$ so that $\prod_{i \in [k] \setminus S} |V_i| = \alpha n^{(k-1)/(k+1)}$. Then either*

1. *There exists $E' \subseteq E$ with $|E'| \geq \frac{Cn}{2}$ so that every S-tuple T has $|E(T)| \leq \max\{\alpha, 1\}$; or*

2. *G contains two edge disjoint isomorphic subgraphs of size $\frac{C}{6}n^{2/(k+1)}$.*

Proof. Suppose that conclusion (1) of the Lemma does not hold. Then there exists a set E' of at least $\frac{Cn}{2}$ edges which share their S-tuple e^S with more than $\max\{\alpha, 1\}$ other edges. Consider the permutation π^k which fixes the levels V_i where $i \in S$, and acts uniformly on the levels V_j where $j \in [k] \setminus S$. For any edge $e \in E'$, let X_e denote the Bernoulli random variable that takes the value 1 if $\pi^k(e) = e'$ for some $e \neq e' \in E'$. We compute that

$$\mathbb{E}[X_e] \geq \frac{\max\{\alpha, 1\}}{\prod_{j \in [k] \setminus S} |V_j|} \geq \frac{\alpha}{\alpha n^{(k-1)/(k+1)}} = n^{-(k-1)/(k+1)}.$$

But then

$$\mathbb{E}[\mathcal{I}_{\pi^k}(E, E)] \geq \mathbb{E}[\mathcal{I}_{\pi^k}(E', E')] = \sum \mathbb{E}[X_e] \geq \frac{C}{2} n^{2/k+1}.$$

Choosing a π^k that obtains this bound, and applying Lemma 1 completes the proof. □

An almost immediate corollary of Lemmas 2, 3, and 4 is the following.

Corollary 2. *Suppose that $G = (V, E)$ a k-uniform hypergraph on n edges (which is not necessarily k-partite). Then either*

1. *G contains a k-partite subhypergraph on vertex set $V_1, \ldots, V_k$, with $|V_1| \leq \cdots \leq |V_k|$ and edge set E' with $|E'| \geq \epsilon n$, where $\epsilon = \frac{1}{(3k)^k 2^{2^k} k! 2^k}$, so that*

   (a) *There exists an ℓ with $\lceil \frac{k+1}{2} \rceil \leq \ell < k$ with $|V_1| \leq \cdots \leq |V_\ell| \leq n^{2/(k+1)}$, such that any $[\ell]$-tuple T has $|E(T)| \leq 1$.*

   (b) *For every $S \subseteq [k]$, if we define $\alpha = \left(\prod_{i \in [k] \setminus S} |V_i|\right)/n^{(k-1)/(k+1)}$, then every S-tuple T has $|E(T)| \leq \max\{\alpha, 1\}$.*

   (c) *Every vertex v in $V_{\ell+1}, V_{\ell+2}, \ldots, V_k$ has $|E(v)| \leq \frac{1}{\epsilon} n^{(k-1)/(k+1)}$; or*

2. *G contains two edge disjoint subhypergraphs of size $Kn^{(k-1)/(k+1)}$ for some constant $K > \frac{1}{3}\epsilon^2$.*

Proof. Making G k-partite, we lose at most a factor of $\frac{1}{k!}$ of the edges. Then we apply Lemma 3 and lose at most a factor of $\frac{1}{(3k)^k}$ of the edges and guarantee that we satisfy condition (1). Note that Lemma 2 guarantees that $\ell < k$. Iterating Lemma 4 over all subsets of $[k]$ loses us a factor of $\frac{1}{2}$ each time, and at most a factor of $\frac{1}{2^{2^k}}$ of the edges. After these applications, we have either found two disjoint isomorphic subgraphs satisfying (2), or we have remaining $\tilde{E} \subseteq E$ with $|\tilde{E}| \geq 2^k \epsilon |E|$ satisfying parts (a) and (b) of (1). To verify part that there is a set E' also satisfying (c), observe that if half of the edges of E' sit in vertices v in some V_j where $j \geq \ell$ with $|E(v)| \geq n^{(k-1)/(k+1)}$ then sacrificing a factor of two of the edges, we find a set that satisfies (a) and (b) with a larger value of ℓ or it satisfies (2). Iteration of this potentially loses a factor of 2^k, leaving us with a set E' of ϵn edges and completes the proof of the corollary. □

For the remainder of the paper, we will frequently choose random permutations π and then be interested in concentration of random variables of the form $\mathcal{I}_\pi(E, E)$ or $\mathcal{I}_\pi(E(T), E)$. These can be thought of as sums of dependent random variables, where dependency comes both from the structure of G and from the fact that π is a random *permutation*. This dependency will cause us some trouble. In order to avoid this trouble we give, in an Appendix, two concentration inequalities that we make use of in our proofs. In particular, we often use a variant of Talagrand's inequality for permutations by McDiarmid [6], which we state as Proposition 1 in the Appendix. We further use a concentration inequality of Chatterjee [1], stated as Proposition 2 in the Appendix, which gives a stronger concentration bound when it applies. Most often, we use the following inequality which follows directly from Talagrand's inequality.

To this end, let E be the edge set of a k-uniform k-partite hypergraph and $S \subseteq [k]$. As above let E^S denote the multiset obtained by restricting the edges of E to the vertex sets indexed by S. We call a set $\mathcal{M} \subset E$ an S-matching if no two elements of $\mathcal{M}$ intersect in the vertex sets indexed by S. Under this framework,

Lemma 5. *Suppose that $G = (V, E)$ is a k-uniform, k-partite hypergraph on $V_1, \ldots, V_k$. Suppose $\mathcal{M}$ is an S-matching for some set S, and let $\pi^S = (\pi_i : i \in S)$ denote a uniform permutation on the vertex sets indexed by S. Let $X = \mathcal{I}^S_{\pi^S}(\mathcal{M}, E)$ and ρ denote the maximum multiplicity of any edge in E^S. Then*

$$\mathbb{P}\left(X \geq \frac{|\mathcal{M}| \cdot |E|}{\prod_{i \in S} |V_i|} + \lambda\right) \leq \exp\left(-\frac{\lambda^2}{64\rho^2 |S| \left(2\frac{|\mathcal{M}| \cdot |E|}{\prod_{i \in S} |V_i|} + \lambda\right)}\right). \tag{i}$$

Proof. This is a direct result of Talagrand's inequality, Proposition 1 from the Appendix, all that is required are finding parameters c and r for Proposition 1 that are satisfied in this situation. Indeed, if two elements of a permutation are swapped, then X changes by at most 2 as $\mathcal{M}$ is a matching, so we take $c = 2\rho$. In order to certify that $X \geq s$ we need merely to exhibit s edges in $\mathcal{M}$ that map to s edges in E. Where these edges map is determined by a set of $|S|s$ coordinates of π. Thus we need to specify at most $|S|s$ coordinates of π to verify that $X \geq |S|$, so we take $r = |S|$. It is easy to verify that $\mathbb{E}[X] = \frac{|\mathcal{M}|(|E|-1)}{N}$, and the median is at most twice this by Markov's inequality. The result then follows immediately from Proposition 1. □

The following Lemma is a simple but useful variant of Markov's inequality:

Lemma 6. *Suppose X is a random variable so that $0 \leq X \leq n$. Then if $\mathbb{E}[X] > 2$,*

$$\mathbb{P}\left(X > \frac{\mathbb{E}[X]}{2}\right) \geq \frac{1}{n}.$$

Proof. If not,

$$\begin{aligned} \mathbb{E}[X] &\leq \frac{\mathbb{E}[X]}{2}\mathbb{P}\left(X \leq \frac{\mathbb{E}[X]}{2}\right) + n\mathbb{P}\left(X > \frac{\mathbb{E}[X]}{2}\right) \\ &\leq \frac{\mathbb{E}[X]}{2}\left(1 - \frac{1}{n}\right) + n \cdot \frac{1}{n} < \frac{\mathbb{E}[X]}{2} + 1, \end{aligned}$$

a contradiction. □

Another tool used extensively in our concentration arguments is the Hajnal-Szemerédi theorem,

Lemma 7 (The Hajnal-Szemerédi Theorem, [3])**.** *Suppose G is a graph on n vertices with maximum degree Δ. Then G can be partitioned into $\Delta + 1$ independent sets each of size $\lfloor n/(\Delta+1) \rfloor$ or $\lceil n/(\Delta+1) \rceil$.*

One final required lemma ensures the existence of a large matching in an integrally weighted digraph Γ. Here, $\omega(e)$ denotes the weight of an edge, and $\omega(S) = \sum_{e \in S} \omega(e)$ denotes the total weight of edges in S.

Lemma 8. *Suppose Γ is a digraph with total weight $\omega(\Gamma)$, unweighted maximum out-degree $\Delta^+(\Gamma)$, and unweighted maximum in-degree $\Delta^-(\Gamma)$. Then there exists a weighted matching $\mathcal{M}$ in Γ with total weight at least*

$$\omega(\mathcal{M}) \geq \frac{\omega(\Gamma)}{4\Delta - 1},$$

where $\Delta = \max\{\Delta^+(\Gamma), \Delta^-(\Gamma)\}$.

Proof. Clearly the statement of the Lemma holds if $\omega(\Gamma) = 1$, and we proceed by induction on $\omega(\Gamma)$. Let $e = x \to y$ denote an edge so that $\omega(e)$ is maximal. Note that deletion of x and y destroys at most $2(\Delta^+(\Gamma) + \Delta^-(\Gamma)) - 1 \leq 4\Delta - 1$ edges (as the edge $x \to y$ is counted twice). Therefore by the maximality of $\omega(e)$ we have $\omega(\Gamma) > \omega(\Gamma \setminus \{x, y\}) \geq \omega(\Gamma) - (4\Delta - 1)\omega(e)$. Let $\mathcal{M}$ denote a maximum weight matching in Γ, and $\mathcal{M}'$ denote a maximum weight matching in $\Gamma \setminus \{x, y\}$. Further note that deletion of x and y only can decrease the maximum in- and out-degrees. By induction,

$$\begin{aligned}\omega(\mathcal{M}) &\geq \omega(e) + \omega(\mathcal{M}')\\ &\geq \omega(e) + \frac{\omega(\Gamma \setminus \{x, y\})}{4\Delta - 1}\\ &\geq \omega(e) + \frac{\omega(\Gamma) - (4\Delta - 1)\omega(e)}{4\Delta - 1}\\ &= \frac{\omega(\Gamma)}{4\Delta - 1}.\end{aligned}$$

as desired. □

2 The proof of Theorem 1

We are now ready to proceed with the proof of Theorem 1. The cases $k = 4$, and $k = 5$ are fairly simple with the machinery we built up in the previous section. We shall observe that the case $k = 6$ will require some additional work. The proof of the cases $k = 4$ and 5 will make the particular technical difficulties when $k = 6$ more clear. We address these difficulties after the proof in the cases of $k = 4$ and 5. Throughout the remainder of the paper $c_1, c_2, \ldots$ are taken to be positive constants, large enough so that the associated statements hold.

For the remainder of the paper, we will thus assume that G satisfies (1) from Corollary 2, namely:

Assumption: $G = (\bigcup_{i=1}^k V_i, E)$ is a k-partite subhypergraph on vertex set $V_1, \ldots, V_k$, with $|V_1| \leq \cdots \leq |V_k|$ and $|E| = \epsilon n$, so that

(a) There exists an ℓ with $\lceil \frac{k+1}{2} \rceil \leq \ell < k$ and with $|V_1| \leq \cdots \leq |V_\ell| \leq n^{2/(k+1)}$, and furthermore any $[\ell]$-tuple of vertices T has $|E(T)| \leq 1$.

(b) For every $S \subseteq [k]$, if we define $\alpha = \alpha_S = \left(\prod_{i \in [k] \setminus S} |V_i|\right)/n^{(k-1)/(k+1)}$, then every S-tuple T has $|E(T)| \leq \max\{\alpha, 1\}$.

(c) Every vertex v in $V_{\ell+1}, V_{\ell+2}, \ldots, V_k$ has $|E(v)| \leq \frac{1}{\epsilon} n^{(k-1)/(k+1)}$.

Remark: One may take $\epsilon \geq \frac{1}{(3k)^k 2^k 2^{2^k} k!}$.

Proof of Theorem 1, $k = 4, 5$. When $k = 4$, since $\ell \geq \lceil \frac{k+1}{2} \rceil$, the only possibility is that $\ell = 3$. Therefore our assumption is that we have a 4-uniform, 4-partite hypergraph with ϵn edges, and so that $|V_1| \leq |V_2| \leq |V_3| \leq n^{2/5}$, and $|V_4| \geq n^{2/5}$. Possibly by adding additional independent points to V_1, V_2 and V_3 we assume that $|V_1| = |V_2| = |V_3| = n^{2/5}$ for convenience. Furthermore, for $a = 1, 2, 3$ we have that any $\{a, 4\}$-tuple T of vertices has $E(T) \leq n^{1/5}$. This follows by Assumption (b), as in this case $\alpha = (n^{2/5})^2/n^{3/5} = n^{1/5}$. This implies that for any $v \in V_4$ one has $|E(v)| \leq 3n^{3/5}$ since for any $v' \in V_1 \cup V_2 \cup V_3$ we have that $|E(\{v, v'\})| \leq n^{1/5}$. Note that this is stronger than guaranteed by Assumption (c).

Let π^3 denote a uniformly random permutation on V_1, V_2, and V_3. Then if $X = \mathcal{I}_{\pi^3}(E, E)$, we have $\mathbb{E}[X] = \frac{\epsilon n(\epsilon n - 1)}{n^{6/5}} = (\epsilon^2 - o(1))n^{4/5}$ since for any of the ϵn edges in E, there are $\epsilon n - 1$ different target edges they may map to. On the other hand, if $X_v = \mathcal{I}^3_{\pi^3}(E(v), E)$ for $v \in V_4$, we have that $\mathbb{E}[X_v] \leq 3\epsilon n^{2/5}$ as each of the at most $3n^{3/5}$ edges in $E(v)$ has $\epsilon n - 1$ different target edges it might map to.

We desire to extend π^3 by choosing a permutation π_4 of V_4 to form π^4. For a fixed π^4, we have that

$$\mathcal{I}_{\pi^4}(E, E) = \sum_{u \in V_4} \mathcal{I}_{\pi^3}(E(u), E(\pi_4(u))). \tag{ii}$$

We wish to find π_4 such that $\mathcal{I}_{\pi^4}(E, E) \geq Cn^{2/5}$. Since $|V_4|$ is large, a random permutation does not work in expectation, so we must pick π_4 more carefully. To do so we define an auxiliary graph Γ_π as follows.

Γ_π is a weighted, directed graph on vertex set $V(\Gamma_\pi) = V_4$. The edge $v \to u$ occurs in Γ_π with weight $\mathcal{I}_{\pi^3}(E(v), E(u))$. That is, the edge $v \to u$ occurs with weight equal to the number of triplets in $E^3(v)$ which map to triplets in $E^3(u)$ under π^3. Suppose we find a matching $\mathcal{M}$ in Γ_π with total weight ω. Define π_4 so that $\pi(u) = v$ if $u \to v \in \mathcal{M}$, and let π_4 act arbitrarily on all other vertices. For such a π_4 and the extension π^4 of π^3 it yields, we have

$$\mathcal{I}_{\pi^4}(E, E) \geq \omega$$

by construction. Lemma 1 then guarantees two edge disjoint isomorphic subgraphs of size at least $\frac{\omega}{3}$. It suffices to find a matching with large weight, with is guaranteed by Lemma 8 so long as we bound the total weight of our Γ_π, and the maximum in- and out- degrees.

Note that the weight all edges of Γ_π is $X = \mathcal{I}^3_{\pi^3}(E, E)$. If $\Delta^+(\Gamma_\pi)$ and $\Delta^-(\Gamma_\pi)$ denote the maximum out- and in- degree of Γ_π, respectively, we simply need verify that with positive probability π^3 has $X \geq \frac{\epsilon^2}{2} n^{4/5}$, and $\Delta^+(\Gamma_\pi)$ and $\Delta^-(\Gamma_\pi)$ are $O(n^{2/5})$. If we can achieve such, then we are guaranteed a matching by Lemma 8. We define π_4 using our matching and apply Lemma 1 to finish.

First we bound $\Delta^+(\Gamma_\pi)$. For any vertex $v \in V_4$, Assumption (b) implies that any edge in $E^3(v)$ intersects fewer than $3n^{1/5}$ others in $E^3(v)$. Indeed, for an $e \in E^3(v)$; once a vertex $v' \in e^{[3]}$ is fixed, along with v we have $E(\{v, v'\}) < n^{1/5}$ as noted above. The Hajnal-Szemerédi theorem (Lemma 7) thus implies that $E(v)$ may be divided into at most $3n^{1/5}$ 3-matchings $\mathcal{M}_1, \ldots, \mathcal{M}_t$ of size at most $n^{2/5}$.

Let $Y_i = \mathcal{I}^3_{\pi^3}(\mathcal{M}_i, E)$. For any one of our matchings, $\frac{|\mathcal{M}_i||E|}{n^{6/5}} \leq n^{1/5}$. We apply Lemma 5, with parameters $\rho = 1$ (by Assumption (a)) and $S = [3]$, so that $|S| = 3$. Taking $\lambda = c_1\sqrt{n^{1/5}\log(n)}$ we have, for n sufficiently large,

$$\begin{aligned}
\mathbb{P}(Y_i \geq 3n^{1/5}) &\leq \mathbb{P}(Y_i \geq 2n^{1/5} + \lambda) \\
&\leq \exp\left(-\frac{(c_1\sqrt{n^{1/5}\log n})^2}{192(2n^{1/5} + 100\sqrt{n^{1/5}\log n})}\right) \\
&\leq \exp\left(-\frac{c_1^2 n^{1/5}\log(n)}{500n^{1/5}}\right) \leq n^{-c_2}.
\end{aligned}$$

We choose c_1 large enough to take $c_2 = 3$ ($c_1 = 100$ suffices). Applying the union bound, $Y_i < 3n^{1/5}$ for *every* matching with probability at least $1 - n^{-3}$ as there are only at most $3n^{1/5}$ matchings. As $X_v = \sum Y_i$, we have that $X_v < 9n^{2/5}$ with this probability. Note that X_v is the weighted out degree of v, which serves as an upper bound to the unweighted degree. Taking the union bound over all vertices in V_4 implies that $X_v \leq 9n^{2/5}$ for *every* $v \in V_4$ with probability at least $1 - n^{-2}$. On the other hand, observe that π^{-1} is also a uniformly chosen permutation and the out-degree in $\Gamma_{\pi^{-1}}$ is the in-degree in $\Gamma_{\pi^{-1}}$. Therefore, with probability at least $1 - n^{-2}$, the in-degree of every $v \in V_4$ is also bounded by $9n^{2/5}$. Combining, $\max\{\Delta^+(\Gamma_\pi), \Delta^-(\Gamma_\pi)\} < 9n^{2/5}$ with probability at least $1 - 2n^{-2}$.

On the other hand, we would like to say that the total weight, X, is large with reasonably high probability. Note that $X \leq \epsilon n$ deterministically. Since $\mathbb{E}[X] = (\epsilon^2 - o(1))n^{4/5}$, Lemma 6 implies that $X > \frac{\epsilon^2}{2} n^{4/5}$ with probability at least $\frac{1}{n}$. Since $\frac{1}{n} > 1 - (1 - 2n^{-2})$, there exists a permutation π^3 so that $X > \frac{\epsilon^2}{2} n^{4/5}$ and

$\max\{\Delta^+(\Gamma_\pi), \Delta^-(\Gamma_\pi)\} < 9n^{2/5}$. We then apply Lemma 8 to find a matching in Γ_π of total weight $\Omega(n^{2/5})$. We build from this permutation a matching, and apply Lemma 1 to find our two desired subgraphs.

For $k = 5$, the largest change is we have two possibilities for ℓ; either $\ell = 3$ or $\ell = 4$.

The proof in the case where $\ell = 4$ is almost identical to the argument above. Again, we take a random permutation π^4 on $V_1, \ldots, V_4$ and create a graph Γ_π on V_5 where directed edges $v \to u$ are created with weight equal to $\mathcal{I}^4_{\pi^4}(E(v), E(u))$. In order to apply Lemma 8 and find a large matching in Γ_π we wish to bound the maximum in- and out-degree and the total weight.

To bound the maximum out-degree consider we bound $X_v = \mathcal{I}^4_{\pi^4}(E(v), E)$, the out degree of v. Similarly to before we have $\mathbb{E}[X_v] < n^{2/6} = n^{1/3}$. Also we have as an application of Assumption (b) that any two points v', v'' along with v define a unique edge. This implies that $|E(v)| < \binom{4}{2} n^{2/3} = 6n^{2/3}$. Since any edge $e \in E^4(v)$ intersects at most $6n^{1/3}$ others, by the Hajnal-Szemerédi theorem $E(v)$ decomposes into $6n^{1/3}$ 4-matchings $\mathcal{M}_1, \ldots, \mathcal{M}_t$ of size at most $n^{1/3}$. Here is primary difference between the arguments for $k = 4$ and this case of $k = 5$, which leads to an additional log factor: If we let $Y_i = \mathcal{I}^4_{\pi^4}(\mathcal{M}_i, E)$ we have that $\mathbb{E}[Y_i] \le \frac{\epsilon n^{4/3}}{n^{4/3}} = \epsilon$. Since $\mathbb{E}[Y_i]$ is of constant order, the most we can hope to hold with high probability (w.h.p.[1]) is that $Y_i < \log(n)$, and so the best w.h.p. bound achievable for X_v by this method is $O(n^{1/3}\log(n))$. We apply Lemma 5, noting $|S| = 4$ and $\rho = 1$. Taking $\lambda = c_3 \log(n)$, Lemma 5 implies that $Y_i < (c_3+1)\log(n)$ with probability strictly less than n^{-c_4}. We take c_3 large enough so that $c_4 = 3$. We take the union bound over all matchings and vertices for both in- and out- degrees and, with probability strictly less than $1 - 2n^{-2}$, we have that $\max\{\Delta^+(\Gamma_\pi), \Delta^-(\Gamma_\pi)\} < (c_3+1)n^{1/3}\log(n)$. On the other hand, as before, the total weight in Γ_π exceeds $\frac{\epsilon^2}{2} n^{2/3}$ with probability larger than $\frac{1}{n}$. Thus with positive probability, our desired matching exists Γ_π and an application of Lemma 1 yields isomorphic subgraph.

If $\ell = 3$, there is a slightly different scenario: we take a random permutation π^3 on V_1, V_2 and V_3 and then try to extend onto two large levels. Note that $\mathbb{E}[\mathcal{I}^3_{\pi^3}(E, E)] = \frac{\epsilon n(\epsilon n - 1)}{n} = (\epsilon^2 - o(1))n$. Fix π^3 arbitrarily so that $\mathcal{I}^3_{\pi^3}(E, E) > \frac{\epsilon^2}{2} n$.

As before we define an auxiliary graph Γ_π, but now we take the vertex set of Γ_π to be the pairs in V_4, V_5. If $x, x' \in V_4$ and $y, y' \in V_5$ we have $(x, y) \to (x', y')$ with multiplicity $\mathcal{I}^3_{\pi^3}(E(\{x, y\}), E(\{x', y'\})) \ge 1$. A matching in this graph is not quite what we want; we want $\approx n^{1/3}$ edges in Γ_π so that their source and destination vertices are all distinct. We proceed greedily. Greedily select an edge $(x, y) \to (x', y')$ in Γ_π with highest multiplicity, and destroy not only all incident edges, but all edges of Γ_π incident to some pair with any of the vertices x, x', y, y'. By Assumption (c), we have destroyed at most $\frac{4}{\epsilon} n^{2/3}$ total edges of Γ_π as at most $\frac{1}{\epsilon} n^{2/3}$ edges are in each of $E(x), E(x'), E(y)$ and $E(y')$. We can repeat this at least

$$\frac{\mathcal{I}^3_{\pi^3}(E, E)}{4n^{2/3}/\epsilon} \ge \frac{\epsilon^3}{8} n^{1/3} = \Omega(n^{1/3})$$

times. At the end, by constructing π_4 and π_5 from these mappings and choosing the rest arbitrarily we have ensured that π^5 has $\mathcal{I}_{\pi_5}(E, E) = \Omega(n^{1/3})$, and thus Lemma 1 completes the proof. □

Remark: Using the Hajnal-Szemerédi theorem to provide matchings in the proof may appear to overkill. The advantage to using Hajnal-Szemerédi, however, is that it guarantees matchings of the same size. Therefore all matchings 'act similarly' when we analyze how they act under our permutation.

Why, then, is the case $k = 6$ more difficult than the cases when $k = 4, 5$? There are two primary issues: There is first the case where $\ell = 5$ and hence there are 5 small partite sets, $V_1, \ldots, V_5$. Consider permuting $V_1, \ldots, V_5$ uniformly and defining a weighted directed graph Γ_π on V_6 as before. To build a permutation on

[1] When we colloquially say that a bound holds w.h.p., we mean a that the bound holds with probability $1 - O(n^{-K})$ for some large constant K. We use the term w.h.p. upper bound to refer to an upper bound that holds w.h.p. In all proofs, we then proceed to make this notion precise in the context of the particular bound in question.

V_6 we look for a large weighted matching in Γ_π. In the cases $k = 4, 5$, we could break the neighborhood $E(v)$ of a vertex $v \in V_k$ into matchings of the same size using Hajnal-Szemerédi and then consider the matchings independently. As already seen in the $k = 5$ proof, the number of matchings will continue to get larger. If the expected size of $|\pi(\mathcal{M}) \cap E| \ll 1$, Lemma 5 only gives a w.h.p. lower bound of $\mathcal{I}_\pi(\mathcal{M}, E) < C \log n$. If $k = 6$ and $\ell = 5$, $E(v)$ decomposes into $O(n^{3/7})$ matchings of size at most $n^{2/7}$, using Assumption (c) and the Hajnal-Szemerédi. Mapping matchings individually, we cannot guarantee a maximum degree in Γ_π smaller than $O(n^{3/7} \log(n))$. However to apply Lemma 8 and find a matching of weight $n^{2/7}$, we would like the maximum degree close to $n^{2/7}$ as well, because the total weight will be close to $n^{4/7}$.

Second, when $\ell = 4$, one may try to directly copy the proof from the case $k = 5$, defining a graph Γ_π between pairs in $V_5 \times V_6$ and defining π_5 and π_6 by choosing an appropriate 'matching' of edges from this graph. The immediate problem with this approach is that when $k = 5$ a deterministic bound on $|E(v)|$ (and hence on the number of edges of Γ_π destroyed by any choice) is sufficient to ensure that a greedy strategy succeeds. This deterministic bound is no longer sufficient. We will establish a w.h.p. probabilistic bound on $\mathcal{I}^4_\pi(E(v), E)$ so that we know that our strategy succeeds w.h.p.. But now, the structural knowledge about $E(v)$ by Assumption (b) is no longer sufficient to complete analysis in the manner we are accustomed to.

How do we handle these issues? Before diving into statement and proofs, which are somewhat more technical than the proofs so far, let us give a high-level description of what is needed for the case $k = 6$. The main new tool used in both $\ell = 4, 5$ is stronger intersection properties of E than those given by Lemma 4. In order to prove this we also define and find a matching in an auxiliary graph Γ_π. Previously when we bounded the degrees in Γ_π, we actually bounded the maximum weighted degree even though Lemma 8 only requires us to bound the unweighted degree. However in the proof of Lemma 9 below, it is critical that we only need bound the *unweighted* maximum degree of a vertex in Γ_π to find a weighted matching. The second idea needed to complete the proof when $\ell = 5$ is that in order to define our initial permutation π^5 which we extend, it is helpful to construct it in a semi-random manner. We first choose π^4 randomly, but fix it prior to choosing π_5 randomly to find π^5. Once π^5 is constructed in such a way, we extend again to find π^6.

For the remainder of the paper, we assume $k = 6$, and hence either $\ell = 4$ or $\ell = 5$, so we have either two or one larger partite sets, respectively.

Our next aim is to find a 'large' subhypergraph E' of an 6-uniform 6-partite hypergraph E satisfying the property that every quadruple T of points extends to at most one edge of E'.

For some quadruples $S \subseteq [6]$ this already follows from our Assumptions (a), (b), (c).

Indeed if $S \in \binom{[6]}{4}$ contains the indices of two (when $\ell = 4$) or the only one (when $\ell = 5$) larger partite sets then

$$\alpha_S = \Big(\prod_{i \in [6] \setminus S} |V_i| \Big) / n^{5/7} \leq \frac{n^{4/7}}{n^{5/7}} = n^{-1/7},$$

and Assumption (b) implies that any S-tuple T satisfies $|E(T)| \leq \max\{\alpha_S, 1\} \leq 1$.

Further if $\ell = 4$ and $S = [4]$ then $|E(T)| \leq 1$ by Assumption (a), for every S-tuple T.

Observed that the remaining $S \in \binom{[6]}{4}$ covered by neither of the above cases are those S so that $[6] \setminus S$ contains the index of precisely one larger set. In other words $[6] \setminus S = \{q, s\}$ where $|V_q| > n^{2/7}$ and $|V_s| \leq n^{2/7}$. This case is covered by the next lemma.

Lemma 9. *Let $S \subseteq [6]$ be a four-element set such that $[6] \setminus S = \{s, q\}$ where $|V_s| \leq n^{2/7}$ and $|V_q| > n^{2/7}$. Then for every 6-uniform 6-partite hypergraph $G = (V, E)$ with $|E| = cn$ for some $c > 0$ satisfying Assumptions (a), (b), (c) we have that*

1. *there exists $E' \subseteq E$ with $|E'| \geq \frac{cn}{4 \log^2 n}$ such that every S-tuple T has $|E'(T)| \leq 1$; or*
2. *G contains two edge disjoint isomorphic subgraphs of size $\Omega\left(\frac{cn^{2/7}}{\log^3 n}\right)$.*

Proof. We assume that $[6] \setminus S = \{q, s\}$ where $|V_q| > n^{2/7}$ and $|V_s| \leq n^{2/7}$.

If more than half of the edges on E are incident to vertices $x \in V_q$ with $|E(x)| > n^{4/7}$ we take V_q' be the set of these vertices. Note that $|V_q'| < cn^{3/7}$. Consider a permutation π that uniformly permutes V_q' and V_s while fixing the other levels. Then

$$\mathbb{E}[\mathcal{I}_\pi^{\{s,q\}}(E, E)] \geq \frac{|E|}{4n^{5/7}}.$$

There exists such a permutation with $\mathcal{I}_\pi^{s,q}(E, E)$ matching its expectation, and applying Lemma 1 to this permutation gives us condition (2) of the lemma. Thus, by discarding at most half the edges we assume that $|E(v)| < n^{4/7}$ for any $v \in V_q$.

For $x \in V_q$ and $y \in V_s$, let $d_{x,y} = |E(\{x, y\})|$. There exists a natural number p and a collection $\mathcal{P}$ of pairs $\{x, y\}$ where $x \in V_q$ and $y \in V_s$ such that

(α) $p \leq d_{x,y} \leq 2p$ for all $\{x, y\} \in \mathcal{P}$;

(β) $\sum_{\{x,y\} \in \mathcal{P}} d_{x,y} \geq \frac{|E|}{2 \log n}$.

Here the factor of two in (β) comes from the facts that half of the edges in E may already have been discarded. Define

$$E_1 = \bigcup_{\{x,y\} \in \mathcal{P}} E(\{x, y\}),$$

so that for every pair $x \in V_q$ and $y \in V_s$ either $p \leq |E_1(\{x, y\})| \leq 2p$ or $|E_1(\{x, y\})| = 0$.

We further define $d_y = |E_1(y)|$ for every $y \in V_s$. There exists a natural number a so that if we define

$$V_s' = \{y : a \leq d_y \leq 2a\},$$

we have that

$$\sum_{y \in V_s'} d_y \geq \frac{|E|}{2 \log^2 n}.$$

We take $E_2 = \bigcup_{y \in V_s'} E_1(y)$. For every S-tuple T such that $|E_2(T)|$ is odd we select (arbitrarily) one edge $e_T \in E_2(T)$. Let

$$E_3 = \{e_T : |E(T)| \text{ is odd.}\}.$$

If $|E_3| \geq \frac{|E|}{4 \log^2 n}$, then E_3 satisfies condition (1) of the Lemma.

Otherwise we define

$$E^* = E_2 \setminus E_3.$$

Then $|E^*| \geq \frac{|E|}{4 \log^2 n}$ and the hypergraph induced by E^* enjoys the following properties:

(i) for all pairs $x \in V_q$ and $y \in V_s'$ we have that $|E^*(\{x, y\})| \leq 2p$;

(ii) for each $y \in V_s'$ there are at most $\frac{2a}{p}$ vertices $x \in V_q$ with $|E^*(\{x, y\})| > 0$;

(iii) for each $x \in V_q$ there are at most $\frac{n^{4/7}}{p}$ vertices $y \in V_s'$ with $|E^*(\{x, y\})| > 0$.

Indeed the first condition follows from the definition of E_1 because $|E^*(\{x, y\})| \leq d_{x,y} \leq 2p$. The second one follows from the definition of E_2 and the fact that $p \leq d_{x,y}$ and the third one analogously follows because $|E(x)| \leq n^{4/7}$ for every $x \in V_q$.

Since $|E^*(T)|$ is even for every S-tuple T, we pair off the edges in $E^*(T)$. That is we partition $E^*(T)$ into pairs (e, e') which we call *partners*. Choosing these partners is slightly artificial, but gives a specified 'target' to every edge e. This will be useful for us for the rest of the argument.

Let π be a uniformly random permutation on V_s'. Define a weighted, directed graph Γ_π on V_q so that the weight of an edge $x \mapsto x'$ in Γ_π for $x, x' \in V_q$ is

$$|\{(e, e') \in E^*(x) : (e, e') \text{ are partners, and } e' = (e \setminus \{x, y\}) \cup \{x', \pi(y)\}\}|.$$

Since each edge has exactly one partner, the expected total weight the edges of Γ_π is (letting $\omega(\Gamma_\pi)$ denote this weight)

$$\mathbb{E}[\omega(\Gamma_\pi)] = \frac{|E^*|}{|V_s'|}.$$

Note that we gave each edge a partner partly to make sure that this weight was easy to compute, and partly to make understanding the degrees in Γ_π clearer later.

Our goal now is to show that there is a matching in Γ_π with total weight $\frac{cn^{2/7}}{320 \log^2 n}$ with positive probability. We then extend the matching to a permutation of V_q (similarly as in cases $k = 4, 5$), apply π to V_s' and fix level V_i with $i \in S$. This yields a permutation π^6, and an application of Lemma 1 yields condition 2 of the lemma.

In order to find such a matching we apply Lemma 8. Thus we must give an upper bound on the maximum (unweighted) in- and out-degree and a lower bound on the total weight. We give two different bounds on the total weight depending on p, our bound on $|E^*(\{x, y\})|$. If p is large, we have a deterministic bound on the unweighted maximum degree, and if p is small we have a probabilistic bound. We begin with the case where p is large.

Since $\mathbb{E}[\omega(\Gamma_\pi)] = \frac{|E^*|}{|V_s'|}$ we can select a permutation π with $\omega(\Gamma_\pi) \geq \frac{|E^*|}{|V_s'|}$. By (iii), for any vertex $x \in V_q$ there are at most $\frac{n^{4/7}}{p}$ vertices $y \in V_s'$ such that $|E^*(\{x, y\})| > 0$. If $x \to x'$ is an edge of Γ_π then there exist $y \in V_s'$ and $e \in E^*(\{x, y\})$ such that $(e \setminus \{x, y\} \cup \{x', \pi(y)\}) \in E^*$. By (ii), regardless of $\pi(y)$, this can happen for at most $\frac{2a}{p} \leq \frac{2n}{|V_s'|p}$ vertices x'. In total we have a deterministic bound on the out-degree (and likewise in-degree) of

$$\Delta^+(\Gamma_\pi) \leq \left(\frac{n^{4/7}}{p}\right)\left(\frac{2n}{|V_s'|p}\right) \leq \frac{2n^{11/7}}{|V_s'|p^2}.$$

If

$$p^2 \geq \frac{cn^{13/7}}{4 \log^2 n |E^*|}, \tag{iii}$$

then Lemma 8 guarantees a matching of size

$$\frac{|E^*|}{|V_s'|(4\Delta(\Gamma_\pi) - 1)} \geq \left(\frac{|E^*|}{|V_s'|}\right)\left(\frac{|V_s'|p^2}{8n^{11/7}}\right) \geq \frac{|E^*|cn^{13/7}}{32 \log^2 n |E^*| n^{11/7}} = \frac{cn^{2/7}}{32 \log^2 n}.$$

Applying Lemma 1, we then obtain condition 2 of the lemma.

Note that

$$\frac{cn^{13/7}}{4 \log^2 n |E^*|} \leq \frac{cn^{13/7}}{|E|} = n^{6/7},$$

so (iii) is satisfied so long as $p \geq n^{3/7}$. Therefore we can assume that $p \leq n^{3/7}$.

Under the condition that $p \leq n^{3/7}$ we would like to establish a w.h.p. upper bound on $\Delta^+(\Gamma_\pi)$ and $\Delta^-(\Gamma_\pi)$. Fix a vertex $x \in V_q$, and define the random variable Z_x to be the (unweighted) out-degree of x in Γ_π.

For each $x \in V_q$ and $y, y' \in V_s'$ we set

$$a^x_{y,y'} = |\{x' \in V_q : \exists e \in E^*(\{x,y\}) \text{ with } (e, e \setminus \{x,y\} \cup \{x',y'\}) \text{ partners}\}|.$$

Then let $Z^*_x = \sum_y a^x_{y,\pi(y)}$. We claim that Z^*_x is an upper bound for Z_x and a lower bound for the weighted degree of x in Γ_π. Indeed, Z^*_x undercounts the weighted degree as $a^x_{y,y'}$ counts only the number of x' such that for some $e \in E^*(\{x,y\})$ we have that $(e, e \setminus \{x,y\} \cup \{x',y'\})$ are partners. The weighted degree, on the other hand, counts the multiplicity of such edges. On the other hand, if $(e, e \setminus \{x,y\} \cup \{x',\pi(y)\})$ and $(e', e' \setminus \{x,y'\} \cup \{x',\pi(y')\})$ are both partners for different y, y', then Z_x counts both occurrences. Thus, Z^*_x overcounts the unweighted degree.

The expected weighted degree of x is at most $\frac{n^{4/7}}{|V_s'|}$, and hence

$$\mathbb{E}[Z_x] \leq \mathbb{E}[Z^*_x] \leq \frac{n^{4/7}}{|V_s'|}.$$

Since each of the (at most) $2p$ edges in $E^*(x,y)$ have a unique partner, $a^x_{y,y'} \leq 2p \leq 2n^{3/7}$. Recalling $|V_s'| \leq n^{2/7}$, we apply Chatterjee's inequality (Proposition 2 in the Appendix) with $\lambda = c_5 \frac{n^{5/7}\log n}{|V_s'|} > \mathbb{E}[Z^*_x]$ and $\rho = 2n^{3/7}$ and to Z^*_x to see that

$$\begin{aligned}
\mathbb{P}(Z_x \geq (c_5+1)\frac{n^{5/7}\log n}{|V_s'|}) &\leq \mathbb{P}(Z^*_x \geq \mathbb{E}[Z^*_x] + \lambda) \\
&\leq \exp\big(-\frac{\lambda^2}{4\rho\mathbb{E}[Z^*_x] + 2\rho\lambda}\big) \\
&\leq \exp\big(-\frac{\lambda}{6\rho}\big) \\
&\leq \exp\big(-\frac{c_5 n^{5/7}\log n}{12n^{3/7}|V_s'|}\big) \\
&\leq \exp\big(-c_6\log n\big) = n^{-c_6}.
\end{aligned}$$

We choose c_5 large enough so that $c_6 = 3$. Applying the union bound, we have that $Z_x < (c_5+1)\frac{n^{5/7}}{|V_s'|}\log n$ simultaneously for every $x \in V_q$ with probability at least $1 - n^{-2}$. Likewise, we have an identical bound on the (unweighted) in-degree of every $x \in V_q$. On the other hand, applying Lemma 6, if $\omega(\Gamma_\pi)$ denotes the total weight in Γ_π we have $\mathbb{P}(\omega(\Gamma_\pi) > \frac{|E^*|}{2|V_s'|}) > \frac{1}{n}$. Therefore a permutation π exists so that simultaneously $\omega(\Gamma_\pi) > \frac{|E^*|}{2|V_s'|}$ and

$$\max\{\Delta^-(\Gamma_\pi), \Delta^+(\Gamma_\pi)\} < (c_5+1)\frac{n^{5/7}}{|V_s'|}\log n.$$

Recalling $|E^*| > \frac{|E|}{4\log^2 n}$ and applying Lemma 8 yields a matching of weight $\Omega\big(\frac{|E|n^{2/7}}{\log^3 n}\big)$.

As before we define a new permutation π^6 by extending our matching to π_q on V_q, applying π on V_s' and fixing V_i where $i \neq s, q$. Applying Lemma 1 then yields two edge disjoint isomorphic subgraphs satisfying condition (2) of Lemma. Thus assuming no E' satisfying condition (1) exists, then condition (2) holds. ☐

Note that, in the case $\ell = 4$ there are 8 sets S that are not already covered by Assumption (a) and (b), and if $\ell = 5$ there are 5 such sets. After applying Lemma 9 repeatedly, we add the following assumption to Assumptions (a-c) above:

Assumption: G is a 6-partite, 6-uniform graphs with $|E| = \epsilon' n$, satisfying Assumptions (a-c) above and so that additionally

(d) For every set $S \in \binom{[6]}{4}$, every S-tuple T has $|E(T)| \leq 1$.

Remark: One can take $\epsilon' = \frac{\epsilon}{48 \log^{16} n}$.

We are now ready to proceed with the proof of Theorem 1 in the case where $k = 6$.

Proof of Theorem 1, k=6. We begin with the slightly easier case when $\ell = 4$. In this case there are two large partite sets, V_5 and V_6. Let π^4 denote a uniformly random permutation of $V_1, \ldots, V_4$. We wish to extend to permutations on V_5 and V_6. We would like to act as in the $k = 5, \ell = 3$ case, by repeatedly selecting pairs $\{u, v\}$ and $\{u', v'\}$ so that $\mathcal{I}^4_{\pi^4}(E(\{u, v\}), E^4(\{u', v'\})) > 0$, then removing all edges incident to each of u, u', v, and v'. Before we were able to proceed greedily, using our deterministic bounds on $E(u), E(u'), E(v)$ and $E(v')$. However in this case,

$$\mathbb{E}[\mathcal{I}^4_{\pi^4}(E, E)] = \frac{|E|(|E| - 1)}{n^{8/7}} = (\epsilon'^2 - o(1))n^{2-8/7} = (\epsilon'^2 - o(1))n^{6/7},$$

while, Assumption (c) guarantees only that we have $|E(v)| \leq \frac{1}{\epsilon} n^{5/7}$. Thus such a greedy approach will only provide two subgraphs of size $\tilde{O}(n^{1/7})$. In order to do better we replace the deterministic upper bound on these neighborhoods with a probabilistic one that holds with high probability. To accomplish this, we break the neighborhood of a vertex $v \in V_5 \cup V_6$ into matchings using Assumption (d).

Fix $v \in V_5 \cup V_6$. Let $X_v = \mathcal{I}^4_{\pi^4}(E(v), E)$ denote the number of edges in $E^4(v)$ which are mapped by π^4 to an edge in E^4. If $\Delta = \max_{v \in V_5 \cup V_6} X_v$, our greedy strategy really loses at most Δ edges each time as opposed to $n^{5/7}$. The key point is that many of the edges in $E^4(v)$ do not map to edges in E^4, so we do not lose by deleting them. We have

$$\mathbb{E}[X_v] \leq \frac{|E(v)||E|}{n^{8/7}} \leq \frac{\epsilon'}{\epsilon} n^{4/7},$$

and so we need to show X_v is not too much larger than its expectation. Assumption (d) implies that any edge $e \in E(v)$ intersects less than $12n^{4/7}$ other edges in $E(v)$. This follows, as after fixing v and a vertex $x \in e$, choosing any two additional vertices uniquely defines an edge. Our concentration is at this point standard: using the Hajnal-Szemerédi theorem $E(v)$ can be partitioned into at most $12n^{4/7}$ matchings $\mathcal{M}_1, \ldots, \mathcal{M}_t$ each of size at most $\frac{1}{12\epsilon} n^{1/7}$. As in the case where $k = 5$, and $\ell = 4$, we now apply Lemma 5 with $\lambda = c_7 \log(n)$, and apply the union bound to say that $X_v < c_8 n^{4/7} \log(n)$ for *every* $v \in V_5 \cup V_6$. Choosing c_7 large enough, this occurs with probability greater than $1 - n^{-2}$. Applying Lemma 6 we have

$$\mathcal{I}^4_{\pi^4}(E, E) > \frac{1}{2}(\epsilon'^2 - o(1))n^{6/7}$$

with probability greater than $\frac{1}{n}$. With positive probability, there is a π^4 satisfying both. Now our strategy above works: Fix such a π^4 and greedily choose pairs $\{x, y\}$ and $\{x', y'\}$ so that π^4 maps an edge in $E^4(\{x, y\})$ to an edge in $E^4(\{x', y'\})$ and delete all edges containing both. At each step at most $c_8 n^{4/7} \log(n)$ edges out of the $\frac{1}{2}(\epsilon'^2 - o(1))n^{6/7}$ are deleted. This can be repeated $\frac{1}{2c_8 \log(n)}(\epsilon'^2 - o(1))n^{2/7}$ times. As in the proof in the $k = 5$ case, these pairs define π_5 and π_6 and hence π^6 with $\mathcal{I}_{\pi^6}(E, E) > \frac{1}{2c_8 \log(n)}(\epsilon'^2 - o(1))n^{2/7}$. Applying Lemma 1 exhibits two edge disjoint isomorphic subgraphs. Note, since $\epsilon' = \Omega(\log^{-16}(n))$, these subgraphs are of size $\Omega(\log^{-33}(n)n^{2/7})$.

The final challenge is the case where $\ell = 5$. We choose a permutation π^5 on $V_1, \ldots, V_5$ and define an auxiliary graph Γ_π on V_6 as before. If we proceed completely at random, Γ_π will have expected weight $O(n^{4/7})$. To guarantee a large matching, the maximum (unweighted) degree in Γ_π must be of the order $n^{2/7}$. As noted in the discussion after the proof of Theorem 1 in the $k = 4, 5$ case, breaking $E(v)$ for $v \in V_6$ into matchings as before requires $O(n^{3/7})$ matchings. This is not enough to prove a w.h.p. maximum degree of $n^{2/7}$.

Instead, we proceed in two steps. First we take a random permutation π^4 on $V_1, \dots, V_4$. Then we will extend it to a permutation π^5 on $V_1, \dots, V_5$. For technical reasons we require that $|E(x)|$ is not too large for all vertices $x \in V_5$. Note, there exists $V_5' \subseteq V_5$ with

$$\forall x \in V_5' : \quad |E(x)| \leq \frac{2n}{|V_5'|} \tag{iv}$$

and $\sum_{x \in V_5'} |E(x)| \geq \frac{|E|}{\log n}$. Replace V_5 with such a set. This may destroy the fact that V_5 is the second largest set, but this shall not bother us. Therefore from here on, we assume that vertices in $x \in V_5$ satisfy $|E(x)| \leq \frac{2n}{|V_5'|}$, and there are $\epsilon'' n$ edges, where $\epsilon'' = \Omega(\log^{-17}(n))$.

To make the argument clear, we define π^4 to be a uniformly random permutation on $V_1, \dots, V_4$ and π_5 to be a uniformly random permutation on V_5 and take $\pi^5 = (\pi^4, \pi_5)$. This, of course, is just another way of saying π^5 is uniform on $V_1, \dots, V_5$. We will condition on π^4, however, and want this to be clear. Let Ω denote the set of all permutations on $V_1, \dots, V_5$ and $\mathcal{F} = 2^\Omega$. Then the probability space from whence π^5 comes is the uniform distribution on Ω, and $\mathcal{F}$ is the associated σ-field. Let $\sigma(\pi^4) \subset \mathcal{F}$ be the σ-field generated by π^4. For $v \in V_6$ and $u, u' \in V_5$, we define the random variables

$$\begin{aligned}
X_v &= \mathcal{I}^5_{\pi^5}(E(v), E) \\
Y_v &= \mathbb{E}[X_v | \sigma(\pi^4)] \\
X^v_{u,u'} &= \mathcal{I}^4_{\pi^4}(E^4(\{v, u\}), E^4(u')) \\
X &= \mathcal{I}^5_{\pi^5}(E, E) = \sum_v X_v \\
Y &= \mathbb{E}[X | \sigma(\pi^4)]
\end{aligned}$$

Claim: There exists a π^4 and positive constants c_9, c_{10} which satisfy the following conditions for all $v \in V_6$ and $u, u' \in V_5$:

(i) $Y \geq \frac{1}{2}\mathbb{E}[X]$

(ii) $X^v_{u,u'} \leq c_9 n^{2/7} \log(n)$.

(iii) $Y_v \leq c_{10} \frac{n^{4/7} \log n}{|V_5|}$

Assuming the claim, we now proceed to complete the proof of Theorem 1. Then we will return to the proof of the claim.

For the remainder of the proof we fix π^4 satisfying the conclusions of the claim. Consider a uniform permutation π_5 on V_5. Let $\pi = (\pi^4, \pi_5)$ with our particular choice for π^4 fixed. Thus π denotes a permutation where the only randomness is in π_5. We denote π^5 a truly uniform permutation on $V_1, \dots, V_5$. We define Γ_π to be an auxiliary graph on V_6, where edges $v \to v'$ are present with weight $\mathcal{I}^5_\pi(E(v), E(v'))$. By our choice of π^4, the expected weight in Γ_π is the random variable Y defined above (which depended only on π^4). The random variables Y_v defined above are the expected weighted degree of a vertex in Γ^π.

We desire a large weighted matching in Γ_π. In order to apply Lemma 8 to find a matching with large weight, we must upper bound the maximum degree in Γ_π, and lower bound the total weight.

Fix a vertex $v \in V_6$. Let $Z_v = \mathcal{I}^5_\pi(E(v), E)$ denote the weighted out-degree of $v \in \Gamma_\pi$, which serves as a bound on the unweighted degree. Note that by the definition of Y_v above,

$$\mathbb{E}[Z_v] = Y_v \leq c_9 \frac{n^{4/7} \log n}{|V_5|}.$$

Further note that we may write $Z_v = \sum X^v_{u,\pi_5(u)}$, so Z_v is in a form where Chatterjee's inequality applies.

For vertices u, u' in V_5, we set $a_{u,u'} = X^v_{u,u'}$. Letting $\rho = X^v_{u,u'} < c_{10} n^{2/7} \log(n)$ by Claim (iii), we apply Chatterjee's inequality. Set $\lambda = c_{11} \frac{n^{4/7} \log n}{|V_5|}$ and ρ to be our bound from Claim (iii). If $c_{11} > c_9$, we obtain

$$\mathbb{P}\left(Z_v \geq 2c_{11} \frac{n^{4/7} \log n}{|V_5|}\right) \leq \exp\left(-\frac{(c_{11} \frac{n^{4/7} \log n}{|V_5|})^2}{6 \cdot (c_{10} n^{2/7} \log(n)) \cdot (c_{11} \frac{n^{4/7} \log n}{|V_5|} \cdot)}\right)$$
$$\leq \exp\left(-c_{12} \frac{n^{2/7} \log(n)}{|V_5|}\right) \leq n^{-c_{12}}.$$

Choosing c_{11} to be sufficiently large with respect to c_{10} and c_9, we may take $c_{12} = 3$. With such a choice, we apply the union bound to guarantee that all vertices satisfy $Z_v \leq 2 \cdot c_{11} \frac{n^{4/7} \log n}{|V_5|}$, with probability at least $1 - n^{-2}$. Likewise they enjoy an identical upper bound on the in-degree, so with probability at least $1 - 2n^{-2}$

$$\max\{\Delta^+(\Gamma_{\pi^5}), \Delta^-(\Gamma_{\pi^5})\} < 2 \cdot c_{11} \frac{n^{4/7} \log n}{|V_5|}.$$

The expected weight of Γ_π is Y, and $Y \geq \frac{1}{2}\mathbb{E}[X]$ by Claim (i). By Lemma 6, with probability at least $\frac{1}{n}$, the total weight is at least half of its expected weight. Recalling that $|E| = \Omega\big(\frac{n}{\log^{17}(n)}\big)$ (as we lost a factor of $\log(n)$ after assumption (d) by normalizing degrees in V_5) this is at least,

$$\frac{1}{2} Y \geq \frac{1}{4}\mathbb{E}[X] = \mathbb{E}[\mathcal{I}^5_{\pi^5}(E, E)] = \frac{1}{4} \frac{|E|(|E| - 1)}{n^{8/7}|V_5|} = \Omega\left(\log^{-34}(n) \frac{n^{6/7}}{|V_5|}\right).$$

(Again, for clarity, π^5 denotes a *uniformly randomly chosen* permutation of $V_1, \dots, V_5$ while π is the permutation where π^4 is a fixed permutation and π_5 is chosen at random.)

Hence there exists a π_5 which extends π^4 to give the proper maximum degree and total weight to Γ_π and we proceed as before, applying Lemma 8 to find a matching of size $\Omega(\log^{-35}(n) n^{2/7})$ in Γ_π and applying Lemma 1 to complete the result. This completes the proof of the theorem. $\square$

It remains to prove the Claim. For convenience, we recall that π^5 is a uniformly random permutation of $V_1, \dots, V_5$, where π^4 denotes the permutation on the first 4 levels. We defined the following set of random variables. Also recall the assumption that for each vertex $u \in V_5$, we have $|E(u)| < \frac{2n}{|V_5|}$.

$$X_v = \mathcal{I}^5_{\pi^5}(E(v), E)$$
$$Y_v = \mathbb{E}[X_v | \sigma(\pi^4)]$$
$$X^v_{u,u'} = \mathcal{I}_{\pi^4}(E^4(\{v, u\}), E^4(u')).$$
$$X = \mathcal{I}^5_{\pi^5}(E, E) = \sum_v X_v$$
$$Y = \mathbb{E}[X | \sigma(\pi^4)]$$

Recall we wish to show there exists a π^4 and constants c_9 and c_{10} which satisfy the following conditions for all $v \in V_6$ and $u, u' \in V_5$:

$$Y \geq \frac{1}{2}\mathbb{E}[X] \tag{v}$$
$$X^v_{u,u'} \leq c_9 n^{2/7} \log(n) \tag{vi}$$
$$Y_v \leq c_{10} \frac{n^{4/7} \log n}{|V_5|} \tag{vii}$$

Proof of Claim. We begin with the $X^v_{u,u'}$. Fix $v \in V_6$ and $u, u' \in V_5$.

We begin by noting that by Assumption (b),

$$E^4(\{v,u\}) \le \frac{\prod_{i=1}^4 |V_i|}{n^{5/7}}.$$

Recall after normalizing V_5 in the proof, by equation (iv), we have $|E^4(u')| \le \frac{2n}{|V_5|}$. We proceed by breaking into matchings as before. Any edge $e \in E^4(\{v,u\})$ intersects fewer than $12n^{2/7}$ other edges. This also follows from assumption (b): we already are considering $E^4(\{v,u\})$, thus fixing two vertices, and selecting another vertex uniquely defines (at most) one edge. As we have done before, we apply the Hajnal-Szemerédi theorem to break up $E^4(\{v,x\})$ into at most $12n^{2/7}$ matchings each of size at most $\left(\prod_{i=1}^4 |V_i|\right)/n$. We now apply Lemma 5. Recall that we are interested in $\mathcal{I}^4_{\pi^4}(E^4(\{v,u\}), E^4(u'))$. By Assumption (a), the maximum multiplicity of an edge in $E^4(u')$ is one, which corresponds to ρ in our application. Fix a matching $\mathcal{M}_i$, let $X_i = \mathcal{I}^4_{\pi^4}(\mathcal{M}_i, E^4(u))$. Then

$$\mathbb{E}[X_i] = \frac{|\mathcal{M}_i||E^4(u')|}{\prod_{i=1}^4 |V_i|} \le \frac{(\prod_{i=1}^4 |V_i|/n)\cdot(2n/|V_5|)}{\prod_{i=1}^4 |V_i|} = \frac{2}{|V_5|} \le 1.$$

Setting parameters $\lambda = c_{13}\log(n)$ and $\rho = 1$, and finally noting $\mathbb{E}[X_i] \le \lambda$ we have,

$$\begin{aligned}\mathbb{P}\left(X_i \ge (c_{13}+1)\log n\right) &\le \mathbb{P}\left(X_i \ge \mathbb{E}[X_i] + \lambda\right)\\ &\le \exp\left(-\frac{\lambda^2}{256(2\mathbb{E}[X]+\lambda)}\right)\\ &\le \exp\left(-\frac{\lambda}{1000}\right) \le n^{-c_{14}}.\end{aligned}$$

Choose c_{13} large enough so that $c_{14} = 3$. For instance, $c_{13} = 10^4$ suffices. Taking a union bound over all matchings, we have that

$$X^v_{u,u'} \le \sum_i X_i \le 12c_{13}n^{2/7}\log(n),$$

with probability at least $1 - n^{-2}$. Taking $c_9 = 12c_{13}$ finishes the proof of (ii).

Now we show that there exists a c_{10} so that $Y_v < c_{10}\frac{n^{4/7}\log n}{|V_5|}$ with high probability. Note that Y_v can be explicitly written in terms of π^4. The key observation is that by Assumption (d), every [4]-tuple extends to at most one edge. Thus for $e \in E(v)$ and $e' \in E^4$, if $\pi^4(e^4) = e'$, in order for $\pi^5(e^5) \cap E^5$ to be non empty, there is a unique target for the fifth element of e. Thus,

$$Y_v = \frac{1}{|V_5|}\mathcal{I}^4_{\pi^4}(E(v), E).$$

Consequently, in order to show concentration of Y_v it suffices to show concentration of $Y_v^* = \mathcal{I}^4_{\pi^4}(E(v), E)$.

As before, any edge in $E^4(v)$ intersects fewer than $12n^{4/7}$ other edges. Applying Hajnal-Szemerédi, we decompose $E(v)$ into at most $12n^{4/7}$ matchings, $\mathcal{M}_1, \ldots, \mathcal{M}_t$. Assumption (c) asserts $|E^4(v)| \le \frac{1}{\epsilon}n^{5/7}$, so these matchings are of size at most $\frac{1}{12\epsilon}n^{1/7}$. Write $Y_v^* = \sum_i \mathcal{I}^4_{\pi^4}(\mathcal{M}_i, E)$. Then

$$\mathbb{E}[\mathcal{I}^4_{\pi^4}(\mathcal{M}_i, E)] = \frac{|\mathcal{M}_i||E|}{n^{8/7}} \le 1.$$

We apply Lemma 5 as above to each matching, with $\lambda = c_{14}\log(n)$, $\rho = 1$, and $|S| = 4$ to show that with probability at least $1 - n^{-c_{15}}$ all vertices $v \in V_6$ satisfy

$$Y_v^* \le 2c_{14}n^{4/7}\log n$$

and hence

$$Y_v \leq 24c_{14}\frac{n^{4/7}\log n}{|V_5|}.$$

Selecting c_{14} large enough ensures that we may take $c_{15} = 2$, and we take $c_{10} = 24c_{14}$.

Finally note $\mathbb{E}[Y] = \mathbb{E}[X]$ and as $Y \leq n$, Lemma 6 implies that $Y \geq \frac{\mathbb{E}[X]}{2}$ with probability at least $\frac{1}{n}$.

Since all $X_v^{u,u'} < c_9 n^{2/7}\log(n)$ with probability $1 - n^{-2}$, all $Y_v < c_{10}\frac{n^{4/7}\log n}{|V_5|}$ with probability $1 - n^{-2}$ and $Y \geq \frac{\mathbb{E}[X]}{2}$ with proability at least $\frac{1}{n}$, there is a π^4 which simultaneously satisfies (v), (vi) and (vii) completing the proof of the claim. □

3 The proof of Theorem 2

In this section we use our general technique to improve the lower bound of Erdős, Pach and Pyber for $k \geq 7$. In order to establish

$$\iota_k(n) \geq C_k n^{2/(2k-1)},$$

they proceeded inductively as follows. Suppose G is a k-uniform hypergraph, with maximum degree Δ and we have already established a lower bound on $\iota_{k-1}(n)$. Then by considering the $k-1$ uniform hypergraph on Δ edges which is the the neighborhood of a vertex of maximum degree, G contains two edge disjoint isomorphic subgraphs of size $\iota_{k-1}(\Delta)$. On the other hand, G contains a matching of size $\frac{n}{k\Delta}$, and hence two edge disjoint isomorphic subgraphs of size $\frac{n}{2k\Delta}$. Minimizing the maximum of these quantities over Δ and using their result for $\iota_2(n)$ as a base case gives their result. Starting with $\iota_6(n)$ gives $\iota_k(n) = \tilde{\Omega}(n^{2/(2k-5)})$, a very mild improvement.

We now give a slightly more sophisticated argument based on our ideas which gives a slightly better bound. To recall, we wish to prove

Theorem 2. *For all values of $k \geq 7$*

$$\iota_k(n) = \tilde{\Omega}\left(n^{\frac{2}{2k-\log_2 k}}\right).$$

Proof of Theorem 2. Assume that $\varepsilon_k > 0$, and we wish to find a sufficient condition so that $\iota_k(n) = \tilde{\Omega}(n^{\varepsilon_k})$. At the end we will take $\varepsilon_k = \frac{2}{2-log_2 k}$, but it is more convenient now to work with generic ε_k. A simple modification to the proof of Lemma 3 gives:

Lemma 10. *Suppose $G = (V, E)$ is a k-uniform, k-partite hypergraph with n edges. Then either:*

1. *G contains two edge disjoint, isomorphic subgraph of size $\frac{1}{2}n^{\varepsilon_k}$; or*
2. *There exists an $\ell \geq \lceil\frac{1}{\varepsilon_k}\rceil$ and a set $E' \subseteq E(G)$ of size at least $\frac{1}{(3k)^\ell}|E|$ so that*
   - *(a) E' is supported on sets $V_1', \ldots, V_\ell', V_{\ell+1}, \ldots, V_k$ with $|V_1'|, \ldots, |V_\ell'| \leq n^{\varepsilon_k}$; and*
   - *(b) There exists a set $E'' \subseteq E'$ of size at least $\frac{|E'|}{3}$ so that every $[\ell]$-tuple T consisting of one vertex each from $V_1', \ldots, V_\ell'$ has $|E''(T)| \leq 1$.*

We thus assume that G is a k-uniform hypergraph on vertex sets

$$|V_1| = |V_2| = \cdots = |V_\ell| = n^{\varepsilon_k} < |V_{\ell+1}| \leq \cdots \leq |V_k|,$$

for some $\ell \geq \lceil \frac{1}{\varepsilon_k} \rceil$. Further, we have $|E| = n$ and every $[\ell]$-tuple extends uniquely (i.e. for any $[\ell]$-tuple T, $|E(T)| \leq 1$). By losing a factor of at most $\log^{k-\ell}(n)$ edges, we may assume

$$|E(v)| \leq \frac{2n}{|V_i|} \text{ for } i = \ell+1, \ell+2, \dots, k \text{ and all } v \in V_i. \tag{viii}$$

This leaves $\tilde{\Omega}(n)$ edges.

For $j = \ell+1, \ell+2, \dots, k$, we define α_j so that $|V_j| = n^{\alpha_j}$. Note then that $\alpha_{\ell+1} \leq \alpha_{\ell+2} \leq \cdots \leq \alpha_k$.

Observe first that if $2 - \ell\varepsilon_k - \sum_{j=\ell+1}^{k} \alpha_j \geq \varepsilon_k$, then for a random permutation π^k that permutes all levels uniformly at random we have

$$\mathbb{E}[\mathcal{I}_{\pi^k}(E, E)] = \tilde{\Omega}(n^{\varepsilon_k}).$$

Selecting a permutation beating the expectation and applying Lemma 1, we have our desired subgraphs. Thus we may assume that

$$2 - \ell\varepsilon_k - \sum_{j=\ell+1}^{k} \alpha_j < \varepsilon_k. \tag{ix}$$

Secondly consider $j = \ell, \ell+1, \dots, k-1$. Then for a permutation π^j randomly permuting levels $V_1, \dots, V_j$ we have

$$\mathbb{E}[\mathcal{I}^j_{\pi^j}(E, E)] = \tilde{\Omega}\big(n^{2-\ell\varepsilon_k - \sum_{i=\ell+1}^{j} \alpha_i}\big). \tag{x}$$

Suppose

$$\alpha_{j+1} \geq (\ell+1)\varepsilon_k + \sum_{i=\ell+1}^{j} \alpha_i - 1, \tag{xi}$$

and fix a permutation π^j matching the expectation (x). In this case we prove that the permutation π^j can be extended to a permutation π so that $\mathcal{I}_\pi(E, E) = \tilde{\Omega}(n^{\varepsilon_k})$. Let $S = [k] \setminus [j]$. Select a pair of S-tuples T and T' so that $\mathcal{I}^j_{\pi^j}(E(T), E(T')) \geq 1$ (that is, there is an extension of T mapped by π^j to an extension of T'). Then delete all other edges incident to T and T' and choose π to map T to T'. We greedily repeat this process as long as possible. Deleting all edges at each step ensures that π remains well defined. For a fixed $m \in \{j+1, j+2, \dots, k\}$, deleting all edges intersecting T and T' at this level causes a deletion of at most

$$2\frac{2n}{|V_i|} = 4n^{1-\alpha_m} \leq 4n^{1-\alpha_{j+1}} \leq 4n^{2-(\ell+1)\varepsilon_k - \sum_{i=\ell+1}^{j} \alpha_i}$$

edges, by the maximum degree condition (viii) of vertices in V_{j+1}. Here the extra factor of two compared with (viii) comes from deleting edges incident to both T and T'. Each selection deletes $O(n^{2-(\ell+1)\varepsilon_k - \sum_{i=\ell+1}^{j} \alpha_i})$ edges, and as we chose π^j beating its expectation, $\mathcal{I}^j_{\pi^j}(E, E) = \tilde{\Omega}\big(n^{2-\ell\varepsilon_k - \sum_{i=\ell+1}^{j} \alpha_i}\big)$. Therefore the greedy process continues for

$$\frac{\tilde{\Omega}(n^{2-\ell\varepsilon_k - \sum_{i=\ell+1}^{j} \alpha_i})}{O(n^{2-(\ell+1)\varepsilon_k - \sum_{i=\ell+1}^{j} \alpha_i})} = \tilde{\Omega}(n^{\varepsilon_k})$$

steps. Each step increases $\mathcal{I}_\pi(E, E)$ by at least one, therefore this greedy process yields a π with $\mathcal{I}_\pi(E, E) = \tilde{\Omega}(n^{\varepsilon_k})$. Applying Lemma 1 yields desired isomorphic subgraphs.

As we have ruled out (xi), the remaining case is that for every $j = \ell, \ell+1, \dots, k-1$ we have that

$$\alpha_{j+1} < (\ell+1)\varepsilon_k + \sum_{i=\ell+1}^{j} \alpha_i - 1.$$

By induction, we prove that

$$\alpha_{j+1} < 2^{j-\ell}((\ell+1)\varepsilon_k - 1) \tag{xii}$$

for $j \leq k-1$. If $j = \ell$, then $\sum_{i=\ell+1}^{\ell} \alpha_j$ is empty so,

$$\alpha_{\ell+1} < (\ell+1)\varepsilon_k + \sum_{i=\ell+1}^{\ell} \alpha_j - 1 = 2^{j-\ell}((\ell+1)\varepsilon_k - 1),$$

so the (xii) holds for $j = \ell$. Assuming that the (xii) holds for all $\ell \leq i < j$, then

$$\begin{aligned} \alpha_{j+1} &< (\ell+1)\varepsilon_k + \sum_{i=\ell+1}^{j} \alpha_i - 1 \\ &< (\ell+1)\varepsilon_k - 1 + \sum_{i=\ell+1}^{j} 2^{i-\ell-1}((\ell+1)\varepsilon_k - 1) \\ &= ((\ell+1)\varepsilon_k - 1)\left(1 + \sum_{i=0}^{j-\ell-1} 2^i\right) \\ &= 2^{j-\ell}((\ell+1)\varepsilon_k - 1), \end{aligned}$$

completing the induction.

Note that

$$\sum_{i=\ell+1}^{k} \alpha_i \leq \sum_{i=\ell+1}^{k} 2^{i-\ell-1}((\ell+1)\varepsilon_k - 1) \leq 2^{k-\ell}((\ell+1)\varepsilon_k - 1),$$

so

$$2 - (\ell+1)\varepsilon_k < \sum_{i=\ell+1}^{k} \alpha_i < 2^{k-\ell}((\ell+1)\varepsilon_k - 1),$$

where the lower bound comes from (ix) and the upper bound from our induction. Subtracting the left hand side from the right hand side yields

$$(2^{(k-\ell)} + 1)(\ell+1)\varepsilon_k - 2^{k-\ell} - 2 > 0. \tag{xiii}$$

Recall that $\ell \geq \frac{1}{\varepsilon_k}$ by Lemma 10. It thus suffices to show that $\varepsilon_k = \frac{2}{2k - \log_2 k}$ yields a contradiction of (xiii) for all $k \geq \ell \geq \frac{1}{\varepsilon_k}$ if $k \geq 7$. This contradiction yields the proof of Theorem 2.

An easy calculus problem reveals that the LHS of (xiii) is an increasing function for $k \geq 7$, and for $\ell \geq \frac{1}{\varepsilon_k} = k - \frac{\log_2 k}{2}$. Thus it suffices to show that

$$(2^{\frac{\log_2 k}{2}} + 1)(\frac{2}{2k - \log_2 k} + 1) - 2^{\frac{\log_2 k}{2}} - 2 < 0.$$

But this is

$$(\sqrt{k} + 1)(\frac{2}{2k - \log_2 k} + 1) - \sqrt{k} - 2 = \frac{2\sqrt{k}}{2k - \log_2 k} - 1.$$

This is a decreasing function of k, and negative for $k = 7$. This completes the proof.

□

4 Appendix: Concentration Inequalities

In order to establish Theorem 1 we need, several times, to show that random variables are concentrated on their mean. The difficulty in this is that although the random variables we are interested in are sums of other random variables they are not independent. This arises as we deal with random variables that are functions of random permutations. Additionally, we have to show concentration of random variables $X = \sum X_i$ where $\mathbb{E}[X]$ is small compared to the number of terms in the summation. This makes applying martingale inequalities, such as Azuma-Hoeffding inequality difficult.

We have two primary tools. First is a version of Talagrand's inequality established by McDiarmid in [6] for random variables that are a function of independent random permutations. In the form we use it, it states the following:

Proposition 1. *Suppose $V_1, \ldots, V_k$ are sets and let $\pi = (\pi_1, \ldots, \pi_k)$ be a family of independent random permutations, so that π_i is chosen uniformly at random from the set $\mathrm{Sym}(V_i)$ of all permutations of V_i. Let c and r be positive constants, and suppose the nonnegative real-valued function $h : \prod_{i=1}^k \mathrm{Sym}(V_i) \to \mathbb{R}$ satisfies the following condition for each $\sigma = (\sigma_1, \ldots, \sigma_k) \in \Omega$:*

- *Swapping any two elements in any σ_i for $i = 1, 2, \ldots, k$ can change the value of $h(\sigma)$ by at most c*
- *If $h(\sigma) = s$, then in order to certify that $h(\sigma) \geq s$, we need to specify at most rs coordinates of σ. That is, if $h(\sigma) \geq s$, there exists a set of rs coordinates of σ so that any permutation σ' agreeing with σ on these rs coordinates also has $h(\sigma') \geq s$.*

Then, if $Z = h(\pi)$ and m denotes the median of Z we have for each $\lambda \geq 0$,

$$\mathbb{P}(Z \geq \mathrm{Med}(X) + \lambda) \leq 2 \exp\left(-\frac{\lambda^2}{16rc^2(\mathrm{Med}(X) + \lambda)}\right).$$

Essentially, this result implies that if we have a (not too large) deterministic bound on the change that occurs when we switch two entries of a permutation we get tight concentration. A slight annoyance is that this gives concentration around the median as opposed to the mean. In most cases where Talagrand is applied, c is small and the median and mean are close together. For our applications, the fact that the median is at most twice the expectation (by Markov's inequality) will suffice.

Sometimes, however, Talagrand's inequality does not suffice for us because c is too large. If there is only one random permutation involved, we are able to do somewhat better by applying a result of Chatterjee. For a particular example of where this is necessary, consider the proof of the final claim. The following is a direct consequence of the proof of Proposition 1.1 and of Theorem 1.5 in [1]

Proposition 2. *Let $\{a_{ij}\}_{1 \leq i,j \leq n}$ be a collection of numbers in $[0, \rho]$. Let $X = \sum_{i=1}^n a_{i\pi(i)}$, where π is a uniformly randomly chosen element of Sym(n). Then*

$$\mathbb{P}(X \geq \mathbb{E}[X] + \lambda) \leq \exp\left(-\frac{\lambda^2}{4\rho\mathbb{E}[X] + 2\rho\lambda}\right) \tag{xiv}$$

In our applications, we will have that $\lambda > \mathbb{E}[X]$ and hence we will often use Chatterjee's inequality in the following form:

Corollary 3. *Let $\{a_{ij}\}_{1 \leq i,j \leq n}$ be a collection of numbers in $[0, \rho]$. Let $X = \sum_{i=1}^n a_{i\pi(i)}$, where π is a uniformly randomly chosen element of S_n. Then for $\lambda > \mathbb{E}[X]$*

$$\mathbb{P}(X \geq \mathbb{E}[X] + \lambda) \leq \exp\left(-\frac{\lambda}{6\rho}\right) \tag{xv}$$

Remark 1: In [1], Proposition 1.1 gives a derivation of $\mathbb{P}(|X - \mathbb{E}[X]| \geq \lambda)$, which introduces a factor of 2. Since we only need the upper tail, we can avoid this factor. Chatterjee also considers the case where $a_{i\pi(i)} \in [0, 1]$ as opposed to $[0, c]$; dividing by c yields our result.

Remark 2: Suppose $a_{i,j}$ only took the values $0, 1, 2, \ldots, C$, and $X = \sum a_{i,\pi(i)}$. Then one can apply Talagrand's Inequality, Proposition 1, with $r = 1$ and $c = C$, to show that

$$\mathbb{P}(X \geq \mathrm{Med}(X) + \lambda) \leq 2 \exp\left(-\frac{\lambda^2}{16C^2(\mathrm{Med}(X) + \lambda)}\right).$$

On the other hand, Chatterjee's inequality gives a bound like

$$\mathbb{P}(X \geq \mathbb{E}[X] + \lambda) \leq \exp\left(-\frac{\lambda^2}{4C\mathbb{E}[X] + C\lambda}\right).$$

Note the dependence in the exponent of C^2 versus C. For C a small constant, this hardly matters but for C large (as in our applications) this can matter quite a bit. A similar phenomenon can be observed when $X = \sum a_i X_i$ where X_i are independent $0/1$ valued random variables and $a_i \in \{1, 2, 3, \ldots, C\}$. Then the standard version of Talagrand's inequality gives

$$\mathbb{P}(X \geq \mathrm{Med}(X) + \lambda) \leq 2 \exp\left(-\frac{\lambda^2}{4C^2(\mathrm{Med}(X) + \lambda)}\right),$$

while an appropriate version of the Chernoff bounds gives

$$\mathbb{P}(X \geq \mathbb{E}[X] + \lambda) \leq \exp\left(-\frac{\lambda^2}{2\mathbb{E}[X^2] + C\lambda/3}\right) \leq \exp\left(-\frac{\lambda^2}{2C\mathbb{E}[X] + C\lambda/3}\right).$$

Thus Chatterjee's inequality gives us something more in line with the Chernoff bounds when it applies. Unfortunately as we often have to concentrate functions of several permutations, the dependence structure becomes more complicated and hence Talagrand's inequality becomes more applicable.

References

[1] S. Chatterjee, Stein's method for concentration inequalities, *Probability theory and related fields*, **138**, no. 1-2, 2007, 305–321

[2] P. Erdős, J. Pach, and L. Pyber, Isomorphic subgraphs in a graph,*Combinatorics (Eger, 1987),* 553556, Colloq. Math. Soc. Jnos Bolyai, **52**, North-Holland, Amsterdam, 1988,

[3] A. Hajnal, and E. Szemerédi, Proof of a conjecture of Erdős, In *Combinatorial Theory and its Applications, Vol. 2*, (Ed. P. Erdős, A. Rényi, and V. T. Sós). Amsterdam, Netherlands: North-Holland, 1970, 601–623

[4] R. Gould, and V. Rödl, On isomorphic subgraphs.*Discrete Math.* **118** (1993), no. 1-3, 259262

[5] C. Lee, P. Loh, and B. Sudakov, Self-similarity of graphs, preprint.

[6] C. McDiarmid, Concentration for independent permutations, *Combinatorics, Probability and Computing*, **11**, no. 2, 2002, 163–178