ADVANCES IN COMBINATORICS, 2020:4, 52 pp.
www.advancesincombinatorics.com

A Completion of the Proof of the Edge-statistics Conjecture

Jacob Fox* Lisa Sauermann

Received 13 March 2019; Published 28 February 2020

Abstract: For given integers k and ℓ with $0 < \ell < \binom{k}{2}$, Alon, Hefetz, Krivelevich and Tyomkyn formulated the following conjecture: When sampling a k-vertex subset uniformly at random from a very large graph G, then the probability to have exactly ℓ edges within the sampled k-vertex subset is at most $e^{-1} + o_k(1)$. This conjecture was proved in the case $\Omega(k) \leq \ell \leq \binom{k}{2} - \Omega(k)$ by Kwan, Sudakov and Tran. In this paper, we complete the proof of the conjecture by resolving the remaining cases. We furthermore give nearly tight upper bounds for the probability described above in the case $\omega(1) \leq \ell \leq o(k)$. We also extend some of our results to hypergraphs with bounded edge size.

Key words and phrases: edge-statistics, inducibility

1 Introduction

Given integers $k \geq 1$ and $0 \leq \ell \leq \binom{k}{2}$, Alon, Hefetz, Krivelevich and Tyomkyn [1] asked the following question: If we choose a k-vertex subset uniformly at random from a very large graph, what is the maximum possible probability to obtain exactly ℓ edges within the random k-vertex subset? More precisely, for $n \geq k$ they defined $I(n,k,\ell)$ to be the maximum possible probability of the event to have exactly ℓ edges within a uniformly random k-vertex subset sampled from an n-vertex graph G (where the maximum is taken over all n-vertex graphs G). Observing that $I(n,k,\ell)$ is a monotone decreasing function of n, they defined the *edge-inducibility* $\text{ind}(k,\ell) = \lim_{n\to\infty} I(n,k,\ell)$.

Clearly, $0 \leq \text{ind}(k,\ell) \leq 1$. If $\ell = 0$ or $\ell = \binom{k}{2}$, it is easy to see that $\text{ind}(k,\ell) = 1$ by taking G to be an empty or a complete graph, respectively. However, Alon et al. [1] showed that for $0 < \ell < \binom{k}{2}$ the edge-inducibility $\text{ind}(k,\ell)$ is bounded away from 1. Furthermore, they formulated the following stronger conjecture, which they called the Edge-statistics Conjecture:

*Research supported by a Packard Fellowship and by NSF Career Award DMS-1352121.

 DOI: 10.19086/aic.12047

Conjecture 1.1 (The Edge-statistics Conjecture, [1])**.** *For all integers k and ℓ with $0 < \ell < \binom{k}{2}$, we have*

$$\mathrm{ind}(k,\ell) \le \frac{1}{e} + o_k(1).$$

The following observations of Alon et al. [1] motivate this conjecture and show that the constant $1/e$ in the conjecture is the smallest possible: By considering a random graph $G(n,p)$ with $p = 1/\binom{k}{2}$, one can see that $\mathrm{ind}(k,1) \ge e^{-1} + o_k(1)$. Furthermore, by considering a complete bipartite graph with parts of sizes n/k and $(k-1)n/k$, one obtains $\mathrm{ind}(k,k-1) \ge e^{-1} + o_k(1)$.

It is easy to see by considering graph complements that $\mathrm{ind}(k,\ell) = \mathrm{ind}(k,\binom{k}{2} - \ell)$. Hence, when studying the edge-inducibility $\mathrm{ind}(k,\ell)$, one may assume that $\ell \le \binom{k}{2}/2$.

As mentioned above, Alon et al. [1] proved that $\mathrm{ind}(k,\ell) < 1-\varepsilon$ for all $0 < \ell < \binom{k}{2}$ with some absolute constant ε. They furthermore established Conjecture 1.1 for $\Omega(k^2) \le \ell \le \binom{k}{2}/2$ by proving that $\mathrm{ind}(k,\ell) \le O(k^{-0.1})$ in this range. They made further partial progress towards their conjecture by showing that $\mathrm{ind}(k,\ell) \le \frac{1}{2} + o_k(1)$ if $\omega(k) \le \ell \le \binom{k}{2}/2$ as well as $\mathrm{ind}(k,\ell) \le \frac{3}{4} + o_k(1)$ for fixed $\ell > 1$ and $\mathrm{ind}(k,1) \le \frac{1}{2} + o_k(1)$.

Kwan, Sudakov and Tran [6] subsequently proved that $\mathrm{ind}(k,\ell) \le \sqrt{(k/\ell)} \cdot \log^{O(1)}(3\ell/k)$ for $k \le \ell \le \binom{k}{2}/2$. From considering a complete bipartite graph with appropriately sized parts, one can see that their bound is best-possible up to the logarithmic factors. Clearly, their result implies Conjecture 1.1 for $C \cdot k \le \ell \le \binom{k}{2}/2$, where C is a sufficiently large absolute constant.

Our main contribution in this paper is to resolve Conjecture 1.1 in the remaining cases, namely for $1 \le \ell \le C \cdot k$, where C is an appropriately chosen constant such that the result of Kwan et al. [6] yields Conjecture 1.1 for $C \cdot k \le \ell \le \binom{k}{2}/2$. More precisely, we prove the following:

Theorem 1.2. *For any constant C, the following is true: For all integers k and ℓ with $1 \le \ell \le C \cdot k$, we have*

$$\mathrm{ind}(k,\ell) \le \frac{1}{e} + o_k(1).$$

Here, the $o_k(1)$-term converges to zero as $k \to \infty$ while C is fixed (the term may depend on C, but not on ℓ).

As mentioned above, together with [6, Theorem 1.1], this theorem implies Conjecture 1.1. It furthermore improves [1, Theorem 1.6], which states the above-mentioned bounds for $\mathrm{ind}(k,\ell)$ while ℓ is fixed.

Theorem 1.2 implies that $\mathrm{ind}(k,1) = e^{-1} + o_k(1)$ and $\mathrm{ind}(k,k-1) = e^{-1} + o_k(1)$ (the correponding lower bounds were discussed above). It is an interesting question to study $\mathrm{ind}(k,\ell)$ in the ranges away from $\ell = 1$ and $\ell = k-1$. Alon et al. [1, Conjecture 6.1] conjectured that $\mathrm{ind}(k,\ell) = o_k(1)$ if $\omega(k) \le \ell \le \binom{k}{2}/2$. Kwan et al. [6] proved this conjecture as a consequence of their result that $\mathrm{ind}(k,\ell) \le \sqrt{(k/\ell)} \cdot \log^{O(1)}(3\ell/k)$ for $k \le \ell \le \binom{k}{2}/2$. Our next theorem shows that $\mathrm{ind}(k,\ell) = o_k(1)$ if $\omega(1) \le \ell \le o(k)$, which can be seen as an analogue of [1, Conjecture 6.1] in the sublinear range.

Theorem 1.3. *Let us fix a real number $\varepsilon > 0$. Then for all sufficiently large integers k, the following holds: For all integers ℓ with $1 \le \ell \le k/\log^4 k$, we have*

$$\mathrm{ind}(k,\ell) \le 90 \cdot \ell^{-1/4}$$

and for all integers ℓ with $k/\log^4 k \le \ell \le (1-\varepsilon)k/2$ we have

$$\mathrm{ind}(k,\ell) \le 90 \cdot k^{-1/4} \cdot \log k.$$

It is not hard to see that the exponent $-1/4$ in Theorem 1.3 is best-possible. Indeed, if $1 \le \ell \le k$ and $\ell = \binom{m}{2}$ for some positive integer m, then considering an n-vertex graph G consisting of a clique on $(m/k) \cdot n$ vertices (and no additional edges) shows that $\mathrm{ind}(k,\ell) \ge \Omega(m^{-1/2}) = \Omega(\ell^{-1/4})$.

Kwan et al. [6] started studying the edge-inducibility for r-uniform hypergraphs, a direction suggested by Alon et al. [1]. For a hypergraph G with at least k vertices, let $I(G,k,\ell)$ be the probability that a uniformly random k-vertex subset of G contains exactly ℓ edges. Now, for $r \ge 1$, $0 \le \ell \le \binom{k}{r}$ and $n \ge k$, let $I_r(n,k,\ell)$ be the maximum of $I(G,k,\ell)$ over all r-uniform hypergraphs G on n vertices. Again, $I_r(n,k,\ell)$ is a monotone decreasing function of n, and one can define $\mathrm{ind}_r(k,\ell) = \lim_{n\to\infty} I_r(n,k,\ell)$. With this notation, we have $\mathrm{ind}(k,\ell) = \mathrm{ind}_2(k,\ell)$. A natural analogue of Conjecture 1.1, suggested by Alon et al. [1], is that for fixed r one has $\mathrm{ind}_r(k,\ell) \le e^{-1} + o_k(1)$ for all $0 < \ell < \binom{k}{r}$. Note that $\mathrm{ind}_r(k,\ell) \ge e^{-1} + o_k(1)$ whenever $\ell = \binom{k-s}{r-s}$ for some $1 \le s \le r$. This can be seen by taking H to be an auxiliary random s-uniform hypergraph $G_s(n,p)$ with $p = 1/\binom{k}{s}$, and defining the edges of G to be all r-sets that contain one of the edges of H. For $s = r-1$, we in particular see that $\mathrm{ind}_r(k,k-r+1) \ge e^{-1} + o_k(1)$.

Our results for graphs for the sublinear range $1 \le \ell \le o(k)$ also extend to r-uniform hypergraphs. In fact, they also extend to not necessarily uniform hypergraphs all of whose edges have size bounded by r. Here, we insist that all edges of a hypergraph are non-empty. For $r \ge 1$, $0 \le \ell \le \binom{k}{r}$ and $n \ge k$, let $I_{\le r}(n,k,\ell)$ be the maximum of $I(G,k,\ell)$ over all hypergraphs G on n vertices all of whose edges have size at most r. Again, $I_{\le r}(n,k,\ell)$ is a monotone decreasing function of n, and one can define $\mathrm{ind}_{\le r}(k,\ell) = \lim_{n\to\infty} I_{\le r}(n,k,\ell)$. We clearly have $\mathrm{ind}_r(k,\ell) \le \mathrm{ind}_{\le r}(k,\ell)$. However, note that one cannot hope for $\mathrm{ind}_{\le r}(k,\ell) \le e^{-1} + o_k(1)$ for all $0 < \ell < \binom{k}{r}$, as $\mathrm{ind}_{\le r}(k,\ell) = 1$ whenever $\ell = \binom{k}{t}$ for some $1 \le t \le r-1$ (this can be seen by considering a complete t-uniform hypergraph).

Note that the case $r = 1$ is not interesting, as $I_1(n,k,\ell) = I_{\le 1}(n,k,\ell)$ is just given by maximizing a certain probability in a hypergeometric distribution. In particular, it is not hard to check that $\mathrm{ind}_1(k,\ell) = \mathrm{ind}_{\le 1}(k,\ell) \le e^{-1} + o_k(1)$ for all $1 \le \ell \le k-1$.

Our first result for hypergraphs is that $\mathrm{ind}_r(k,\ell) \le \mathrm{ind}_{\le r}(k,\ell) \le e^{-1} + o_k(1)$ if $1 \le \ell \le o(k)$. This is an immediate consequence of the following theorem.

Theorem 1.4. *For all positive integers r, k and ℓ with $1 \le \ell < k/r$, we have*

$$\mathrm{ind}_r(k,\ell) \le \mathrm{ind}_{\le r}(k,\ell) \le \frac{k}{k-r\ell} \cdot \frac{1}{e}.$$

The following theorem extends Theorem 1.3 to the hypergraph setting.

Theorem 1.5. *Let us fix a positive integer $r \ge 3$ and a real number $\varepsilon > 0$. Then for all sufficiently large integers k, we have*

$$\mathrm{ind}_r(k,\ell) \le \mathrm{ind}_{\le r}(k,\ell) \le 100 \cdot \ell^{-1/(2r)}$$

for all $1 \le \ell \le (1-\varepsilon)k/r$.

Again, the exponent $-1/(2r)$ is best-possible. Indeed, if $1 \le \ell \le k$ and $\ell = \binom{m}{r}$ for some positive integer m, then considering an n-vertex hypergraph G consisting of an r-uniform clique on $(m/k) \cdot n$ vertices (and no additional edges) shows that $\mathrm{ind}(k,\ell) \ge \Omega(m^{-1/2}) = \Omega(\ell^{-1/(2r)})$.

Another possible direction to extend the study of the edge-inducibility $\mathrm{ind}(k,\ell)$ for graphs, is to consider more restricted families of induced subgraphs on k-vertices with ℓ edges. One initial motivation for Alon et al. [1] to introduce and study edge-inducibilities was the close relationship to graph inducibilities, as introduced by Pippenger and Golumbic [8]. The inducibility of a k-vertex graph H is the maximum possible probability of the event to have an induced subgraph isomorphic to H when sampling a k-vertex subset uniformly at random from a large graph G (where again, we consider the limit as the number of vertices of G tends to infinity). The graph inducibility problem has attracted a lot of attention recently, see for example [3, 4, 5, 10]. If H has k vertices and ℓ edges, then the graph inducibility of H is clearly bounded by $\mathrm{ind}(k,\ell)$. We think that it is an interesting question to study notions of inducibility that lie between graph inducibility and edge-inducibility.

For example, one can ask the following question: What is the maximum possible probability of the event to have an induced forest with precisely ℓ edges when sampling a k-vertex subset uniformly at random from a large graph G? The following theorem implies an upper bound for this probability if $1 \le \ell \le \sqrt{k}/4$.

Theorem 1.6. *Let k and ℓ be positive integers with $1 \le \ell \le \sqrt{k}/4$ and let G be a graph on $n \ge k$ vertices. Then the number of k-vertex subsets $A \subseteq V(G)$ that induce a forest with exactly ℓ edges is at most*

$$50 \cdot \ell^{-1/2} \cdot \frac{n^k}{k!}.$$

Note that for sufficiently large n, Theorem 1.6 is tight up to a constant factor. Indeed, consider a graph G consisting of a complete bipartite graph with parts of sizes n/k and $\ell \cdot n/k$ and with $(k-\ell-1) \cdot n/k$ additional isolated vertices. When sampling a k-vertex subset uniformly at random from this graph, with probability $\Omega(\ell^{-1/2})$ we obtain an induced star with ℓ edges together with $k-\ell-1$ additional isolated vertices. Thus, Theorem 1.6 is tight up to a constant factor for all $1 \le \ell \le \sqrt{k}/4$. If $\ell = o(\sqrt{k})$ one can also consider a random graph $G(n,p)$ with $p = \ell/\binom{k}{2}$ and note that when sampling a k-vertex subset, one obtains an induced matching of exactly ℓ edges with probability $\Omega(\ell^{-1/2})$.

We remark that we did not attempt to optimize the constants in Theorems 1.3, 1.5 and 1.6.

Organization. We will first prove Theorem 1.4 in Section 2. Note that applying this theorem to $r = 2$ yields Theorem 1.2 in the case where $1 \le \ell \le o(k)$. In Section 3, we will prove Theorems 1.3, 1.5 and 1.6. The proofs will rely on three lemmas, which we will prove in Sections 4 and 5. A corollary of these lemmas will also be an important ingredient in the proof of Theorem 1.2. This proof can be found in Section 6, apart from the proofs of several other lemmas which we will postpone to Sections 7 to 9.

Notation. All logarithms are to base 2. For integers $m \ge 1$ and $0 \le t \le m$, let $(m)_t$ denote the falling factorial $\binom{m}{t} \cdot t! = m \cdot (m-1) \cdots (m-t+1) = \prod_{i=0}^{t-1}(m-i)$. Note that $(m)_0 = 1$.

All edges of all hypergraphs are assumed to be non-empty.

For a hypergraph G, let $V(G)$ denote its vertex set. For a vertex $v \in V(G)$, let $\deg_G(v)$ be its degree (which is the number of edges of G that contain v). Similarly, for a subset $W \subseteq V(G)$ and a vertex $v \in W$, let $\deg_W(v)$ denote the degree of v in the sub-hypergraph induced by W.

Furthermore, for a hypergraph G and a subset $W \subseteq V(G)$, let $e(W)$ be the number of edges of G that are subsets of W. Let us call a vertex $v \in W$ *non-isolated in* W if there is an edge $e \subseteq W$ with $v \in e$, and let us call v *isolated in* W otherwise. Finally, let $m(W)$ denote the number of vertices of W that are non-isolated in W. If all edges of G have size at most r, then we clearly have $m(W) \le r \cdot e(W)$.

For a graph G and a vertex $v \in V(G)$, let $N(v) \subseteq V(G) \setminus \{v\}$ denote the neighborhood of v in G.

All the $o_k(1)$-terms are supposed to only depend on k and C (and on none of the other variables or objects), and to converge to zero as $k \to \infty$ for every fixed C.

2 Proof of Theorem 1.4

In this section, we will prove Theorem 1.4. So let us fix positive integers r, k and ℓ satisfying $1 \le \ell < k/r$. In particular, $k \ge r+1 \ge 2$.

Furthermore, let G be a hypergraph on n vertices all of whose edges have size at most r, and let us assume that n is large with respect to k. We will show that the number of k-vertex subsets $A \subseteq V(G)$ such that the set A contains exactly ℓ edges is at most

$$\frac{k}{k-r\ell} \cdot \frac{1}{e} \cdot \frac{n^k}{k!} = (1+o_n(1)) \cdot \frac{k}{k-r\ell} \cdot \frac{1}{e} \cdot \binom{n}{k},$$

thus proving

$$\mathrm{ind}_r(k,\ell) \le \mathrm{ind}_{\le r}(k,\ell) \le \frac{k}{k-r\ell} \cdot \frac{1}{e}$$

as desired.

We may assume that there is at least one k-vertex subset $A \subseteq V(G)$ that contains exactly ℓ edges.

Let us introduce some more notation. For a subset $W \subseteq V(G)$, let

$$\mathcal{P}_{<r}(W) = \{X \subseteq W \mid |X| < r\}$$

be the family of all subsets of W of size smaller than r. Furthermore, for a subfamily $\mathcal{X} \subseteq \mathcal{P}_{<r}(W)$, let

$$\mathrm{U}(\mathcal{X}) = \bigcup_{X \in \mathcal{X}} X \subseteq W$$

be the union of all members of $\mathcal{X}$.

For a vertex $v \in V(G)$ and a subset $W \subseteq V(G) \setminus \{v\}$, let

$$\mathcal{N}(v,W) = \{X \subseteq W\} \mid X \cup \{v\} \in E(G)\}.$$

This means $\mathcal{N}(v,W)$ is the family of all subsets $X \subseteq W$ such that $X \cup \{v\}$ is an edge of the hypergraph G. In other words, $\mathcal{N}(v,W)$ consists of the sets $e \setminus \{v\}$ for all edges $e \in E(G)$ with $e \subseteq W \cup \{v\}$ and $v \in e$. As every edge of G has size at most r, we have $\mathcal{N}(v,W) \subseteq \mathcal{P}_{<r}(W)$. Also note that $|\mathcal{N}(v,W)|$ is the number of edges inside $W \cup \{v\}$ that contain v, hence

$$|\mathcal{N}(v,W)| = \deg_{W \cup \{v\}}(v).$$

The following definition will be crucial for our proof of Theorem 1.4.

Definition 2.1. A sequence $v_1,\dots,v_k$ of distinct vertices of G is called *good* if $e(\{v_1,\dots,v_k\}) = \ell$ and v_k is non-isolated in $\{v_1,\dots,v_k\}$. For $0 \le j \le k-1$ a sequence $v_1,\dots,v_j$ of distinct vertices of G is called *good* if it can be extended to a good sequence $v_1,\dots,v_k$.

Recall that the condition that v_k is non-isolated in $\{v_1,\dots,v_k\}$ means that there is at least one edge $e \subseteq \{v_1,\dots,v_k\}$ with $v_k \in e$. If $r = 2$ and G is a graph, this simply means that v_k has an edge to at least one of the vertices $v_1,\dots,v_{k-1}$.

Note that any good sequence $v_1,\dots,v_k$ must satisfy $e(\{v_1,\dots,v_{k-1}\}) < e(\{v_1,\dots,v_k\}) = \ell$ and consequently $e(\{v_1,\dots,v_j\}) < \ell$ for all $0 \le j \le k-1$. Furthermore note that for any good sequence $v_1,\dots,v_k$, we have $m(\{v_1,\dots,v_k\}) \ge 1$ as $e(\{v_1,\dots,v_k\}) = \ell \ge 1$.

For $0 \le j \le k-1$ and any good sequence $v_1,\dots,v_j$, set $\lambda(v_1,\dots,v_j) = 1$. On the other hand, for a good sequence $v_1,\dots,v_k$, set

$$\lambda(v_1,\dots,v_k) = \frac{1}{m(\{v_1,\dots,v_k\})}.$$

Now, for $0 \le j \le k-1$ and a good sequence $v_1,\dots,v_j$, set

$$\Lambda(v_1,\dots,v_j) = \sum_{\substack{v_{j+1} \text{ s.t.} \\ v_1,\dots,v_{j+1} \text{ good}}} \lambda(v_1,\dots,v_{j+1}).$$

Note that if $0 \le j \le k-2$, then $\Lambda(v_1,\dots,v_j)$ simply counts the number of choices for $v_{j+1} \in V(G)$ such that $v_1,\dots,v_{j+1}$ is a good sequence (as $\lambda(v_1,\dots,v_{j+1}) = 1$ for all such v_{j+1}).

Furthermore note that $\Lambda(v_1,\dots,v_j) > 0$ for any $0 \le j \le k-1$ and any good sequence $v_1,\dots,v_j$ (because there is at least one way to extend $v_1,\dots,v_j$ to a good sequence $v_1,\dots,v_k$).

For $1 \le j \le k$ and a good sequence $v_1,\dots,v_j$, set

$$\rho(v_1,\dots,v_j) = \prod_{i=1}^{j} \frac{\lambda(v_1,\dots,v_i)}{\Lambda(v_1,\dots,v_{i-1})} = \frac{\lambda(v_1)\cdot\lambda(v_1,v_2)\cdot\lambda(v_1,v_2,v_3)\cdots\lambda(v_1,\dots,v_j)}{\Lambda(\emptyset)\cdot\Lambda(v_1)\cdot\Lambda(v_1,v_2)\cdots\Lambda(v_1,\dots,v_{j-1})}.$$

Here, by slight abuse of notation, $\emptyset$ stands for the empty sequence.

Lemma 2.2. *For every $1 \le j \le k$, we have*

$$\sum_{v_1,\dots,v_j \text{ good}} \rho(v_1,\dots,v_j) = 1.$$

Proof. Let us prove the lemma by induction on j. Note that for $j = 1$, by definition, $\Lambda(\emptyset)$ is precisely the number of $v_1 \in V(G)$ such that v_1 is a good sequence. Thus,

$$\sum_{v_1 \text{ good}} \rho(v_1) = \sum_{v_1 \text{ good}} \frac{\lambda(v_1)}{\Lambda(\emptyset)} = \sum_{v_1 \text{ good}} \frac{1}{\Lambda(\emptyset)} = 1.$$

Now assume $2 \le j \le k$ and that we have already proved the lemma for $j-1$. For every good sequence

$v_1,\dots,v_j$, the sequence $v_1,\dots,v_{j-1}$ is also good. Hence

$$\begin{aligned}
\sum_{v_1,\dots,v_j \text{ good}} \rho(v_1,\dots,v_j) &= \sum_{v_1,\dots,v_j \text{ good}} \rho(v_1,\dots,v_{j-1}) \cdot \frac{\lambda(v_1,\dots,v_j)}{\Lambda(v_1,\dots,v_{j-1})} \\
&= \sum_{v_1,\dots,v_{j-1} \text{ good}} \ \sum_{\substack{v_j \text{ s.t.} \\ v_1,\dots,v_j \text{ good}}} \rho(v_1,\dots,v_{j-1}) \cdot \frac{\lambda(v_1,\dots,v_j)}{\Lambda(v_1,\dots,v_{j-1})} \\
&= \sum_{v_1,\dots,v_{j-1} \text{ good}} \left(\frac{\rho(v_1,\dots,v_{j-1})}{\Lambda(v_1,\dots,v_{j-1})} \cdot \sum_{\substack{v_j \text{ s.t.} \\ v_1,\dots,v_j \text{ good}}} \lambda(v_1,\dots,v_j) \right) \\
&= \sum_{v_1,\dots,v_{j-1} \text{ good}} \frac{\rho(v_1,\dots,v_{j-1})}{\Lambda(v_1,\dots,v_{j-1})} \cdot \Lambda(v_1,\dots,v_{j-1}) = \sum_{v_1,\dots,v_{j-1} \text{ good}} \rho(v_1,\dots,v_{j-1}) = 1,
\end{aligned}$$

using the induction hypothesis in the last step. □

Recall that we want to show that the number of k-vertex subsets $A \subseteq V(G)$ with $e(A) = \ell$ is at most

$$\frac{k}{k-r\ell} \cdot \frac{1}{e} \cdot \frac{n^k}{k!}.$$

Note that for every good sequence $v_1,\dots,v_k$, the set $A = \{v_1,\dots,v_k\}$ satisfies $A \subseteq V(G)$ and $e(A) = \ell$. Conversely, for every k-vertex subset $A \subseteq V(G)$ with $e(A) = \ell$ we can find certain labelings of the elements of A as $v_1,\dots,v_k$ such that $v_1,\dots,v_k$ is a good sequence (we just need to ensure that v_k is one of the $m(A)$ non-isolated vertices in A).

By Lemma 2.2 applied to $j = k$, we have

$$\sum_{\substack{A \subseteq V(G) \\ |A|=k,\, e(A)=\ell}} \ \sum_{\substack{\text{labelings } v_1,\dots,v_k \text{ of } A \\ \text{s.t. } v_1,\dots,v_k \text{ is good}}} \rho(v_1,\dots,v_k) = 1.$$

In order to prove the desired bound on the number of k-vertex subsets A with $e(A) = \ell$, it therefore suffices to show that

$$\sum_{\substack{\text{labelings } v_1,\dots,v_k \text{ of } A \\ \text{s.t. } v_1,\dots,v_k \text{ is good}}} \rho(v_1,\dots,v_k) \geq \frac{k-r\ell}{k} \cdot e \cdot \frac{k!}{n^k}$$

for every k-vertex subset $A \subseteq V(G)$ with $e(A) = \ell$.

So from now on, let us fix a k-vertex subset $A \subseteq V(G)$ with $e(A) = \ell$.

Note that there are exactly $m(A) \cdot (k-1)!$ ways to label the elements of A by $v_1,\dots,v_k$ such that $v_1,\dots,v_k$ is a good sequence. This is because there are exactly $m(A)$ choices for $v_k \in A$ (as v_k needs to be non-isolated in A), and then any labeling $v_1,\dots,v_{k-1}$ of the remaining $k-1$ vertices yields a good sequence $v_1,\dots,v_k$.

It suffices to prove that for each of the $m(A)$ choices for $v_k \in A$, we have

$$\sum_{\substack{\text{labelings } v_1,\dots,v_{k-1} \text{ of } A\setminus\{v_k\} \\ \text{s.t. } v_1,\dots,v_k \text{ is good}}} \rho(v_1,\dots,v_k) \geq \frac{1}{m(A)} \cdot \frac{k-r\ell}{k} \cdot e \cdot \frac{k!}{n^k}.$$

So fix $v_k \in A$ such that v_k is non-isolated in A. Set $A' = A \setminus \{v_k\}$ and note that for any labeling $v_1,\dots,v_{k-1}$ of A', the sequence $v_1,\dots,v_k$ is good. So it suffices to prove

$$\sum_{\text{labelings } v_1,\dots,v_{k-1} \text{ of } A'} \rho(v_1,\dots,v_k) \geq \frac{1}{m(A)} \cdot \frac{k-r\ell}{k} \cdot e \cdot \frac{k!}{n^k}. \tag{2.1}$$

Let $d = \deg_A(v_k)$. Note that $1 \leq d \leq e(A) = \ell$ and

$$e(A') = e(A \setminus \{v_k\}) = e(A) - \deg_A(v_k) = \ell - d < \ell.$$

Furthermore, let $B \subseteq A'$ consist of those vertices in A' that are non-isolated in A'. Then we have $|B| = m(A') \leq r \cdot e(A') = r \cdot (\ell - d)$ and $e(B) = e(A') = \ell - d$.

Note that for any labeling $v_1,\dots,v_{k-1}$ of A' we have

$$\begin{aligned}\rho(v_1,\dots,v_k) &= \frac{\lambda(v_1)\cdot\lambda(v_1,v_2)\cdot\lambda(v_1,v_2,v_3)\cdots\lambda(v_1,\dots,v_k)}{\Lambda(\emptyset)\cdot\Lambda(v_1)\cdot\Lambda(v_1,v_2)\cdots\Lambda(v_1,\dots,v_{k-1})} \\ &= \frac{1\cdots 1\cdot(1/m(\{v_1,\dots,v_k\}))}{\Lambda(\emptyset)\cdot\Lambda(v_1)\cdot\Lambda(v_1,v_2)\cdots\Lambda(v_1,\dots,v_{k-1})} \\ &= \frac{1}{m(A)} \cdot \frac{1}{\Lambda(\emptyset)\cdot\Lambda(v_1)\cdot\Lambda(v_1,v_2)\cdots\Lambda(v_1,\dots,v_{k-1})}. \end{aligned} \tag{2.2}$$

Let us first analyze the quantity $\Lambda(v_1,\dots,v_{k-1})$. In particular, we will see that this quantity is independent of the labeling $v_1,\dots,v_{k-1}$ of A'.

For every labeling $v_1,\dots,v_{k-1}$ of A', the possible extensions of $v_1,\dots,v_{k-1}$ to a good sequence $v_1,\dots,v_{k-1},v'_k$ are given by those vertices $v'_k \in V(G)\setminus A'$ that satisfy $e(A'\cup\{v'_k\}) = \ell$ (note that then v'_k will automatically be non-isolated in $\{v_1,\dots,v_{k-1},v'_k\} = A'\cup\{v'_k\}$, because $e(A') = \ell - d < \ell$). As we always have

$$|\mathcal{N}(v'_k,A')| = \deg_{A'\cup\{v'_k\}}(v'_k) = e(A'\cup\{v'_k\}) - e(A') = e(A'\cup\{v'_k\}) - (\ell - d),$$

the condition $e(A'\cup\{v'_k\}) = \ell$ is equivalent to $|\mathcal{N}(v'_k,A')| = d$. Hence the extensions of $v_1,\dots,v_{k-1}$ to a good sequence $v_1,\dots,v_{k-1},v'_k$ are given by those $v'_k \in V(G)\setminus A'$ that satisfy $|\mathcal{N}(v'_k,A')| = d$. Thus,

$$\begin{aligned}\Lambda(v_1,\dots,v_{k-1}) &= \sum_{\substack{v'_k \text{ s.t.} \\ v_1,\dots,v_{k-1},v'_k \text{ good}}} \lambda(v_1,\dots,v_{k-1},v'_k) \\ &= \sum_{\substack{v'_k\in V(G)\setminus A' \\ |\mathcal{N}(v'_k,A')|=d}} \lambda(v_1,\dots,v_{k-1},v'_k) = \sum_{\substack{\mathcal{X}\subseteq\mathcal{P}_{<r}(A') \\ |\mathcal{X}|=d}} \sum_{\substack{v'_k\in V(G)\setminus A' \\ \mathcal{N}(v'_k,A')=\mathcal{X}}} \lambda(v_1,\dots,v_{k-1},v'_k) \end{aligned} \tag{2.3}$$

for every labeling $v_1,\dots,v_{k-1}$ of A'.

For every subfamily $\mathcal{X} \subseteq \mathcal{P}_{<r}(A')$ with $|\mathcal{X}| = d$, let $0 \le c(\mathcal{X}) \le 1$ be such that the number of vertices $v \in V(G)\setminus A'$ with $\mathcal{N}(v,A') = \mathcal{X}$ is precisely $c(\mathcal{X})\cdot n$. Then for every such $\mathcal{X}$ and every labeling $v_1,\dots,v_{k-1}$ of A', the second sum on the right-hand side of (2.3) has exactly $c(\mathcal{X})\cdot n$ summands. Note that $\mathcal{N}(v,A') = \mathcal{X}$ means that $X\cup\{v\} \in E(G)$ for all $X \in \mathcal{X}$, but for no other subsets $X \subseteq A'$.

Now, for every labeling $v_1,\dots,v_{k-1}$ of A', for every subfamily $\mathcal{X} \subseteq \mathcal{P}_{<r}(A')$ with $|\mathcal{X}| = d$ and for every vertex $v'_k \in V(G)\setminus A'$ with $\mathcal{N}(v'_k,A') = \mathcal{X}$, we claim that we have $m(\{v_1,\dots,v_{k-1},v'_k\}) = |B\cup \mathrm{U}(\mathcal{X})|+1$. Indeed, $\{v_1,\dots,v_{k-1},v'_k\} = A'\cup\{v'_k\}$. The edges inside $A'\cup\{v'_k\}$ are precisely the edges inside A' and the sets of the form $X\cup\{v'_k\}$ for $X \in \mathcal{N}(v'_k,A') = \mathcal{X}$. Hence the non-isolated vertices in $A'\cup\{v'_k\}$ are precisely the vertices in the set

$$\bigcup_{\substack{e\in E(G)\\ e\subseteq A'}} e \;\cup\; \bigcup_{X\in\mathcal{X}} (X\cup\{v'_k\}) = B\cup \mathrm{U}(\mathcal{X})\cup\{v'_k\}.$$

Here we used that $|\mathcal{X}| = d \ge 1$ and that B was defined as the set of non-isolated vertices in A'. Thus, we indeed have $m(\{v_1,\dots,v_{k-1},v'_k\}) = m(A'\cup\{v'_k\}) = |B\cup \mathrm{U}(\mathcal{X})|+1$. Hence

$$\lambda(v_1,\dots,v_{k-1},v'_k) = \frac{1}{m(\{v_1,\dots,v_{k-1},v'_k\})} = \frac{1}{|B\cup \mathrm{U}(\mathcal{X})|+1}.$$

Plugging this into (2.3), for every labeling $v_1,\dots,v_{k-1}$ of A' we obtain

$$\begin{aligned}\Lambda(v_1,\dots,v_{k-1}) &= \sum_{\substack{\mathcal{X}\subseteq\mathcal{P}_{<r}(A')\\ |\mathcal{X}|=d}} \; \sum_{\substack{v'_k\in V(G)\setminus A'\\ \mathcal{N}(v'_k,A')=\mathcal{X}}} \frac{1}{|B\cup \mathrm{U}(\mathcal{X})|+1}\\ &= \sum_{\substack{\mathcal{X}\subseteq\mathcal{P}_{<r}(A')\\ |\mathcal{X}|=d}} c(\mathcal{X})\cdot n\cdot \frac{1}{|B\cup \mathrm{U}(\mathcal{X})|+1} = n\cdot \sum_{\substack{\mathcal{X}\subseteq\mathcal{P}_{<r}(A')\\ |\mathcal{X}|=d}} \frac{c(\mathcal{X})}{|B\cup \mathrm{U}(\mathcal{X})|+1} = C\cdot n. \qquad (2.4)\end{aligned}$$

Here, we set

$$C = \sum_{\substack{\mathcal{X}\subseteq\mathcal{P}_{<r}(A')\\ |\mathcal{X}|=d}} \frac{c(\mathcal{X})}{|B\cup \mathrm{U}(\mathcal{X})|+1}.$$

In particular, we see that $\Lambda(v_1,\dots,v_{k-1})$ is independent of the labeling $v_1,\dots,v_{k-1}$ of A'. Note that $C > 0$ as $\Lambda(v_1,\dots,v_{k-1}) > 0$.

Now, let us find an upper bound for the terms $\Lambda(v_1,\dots,v_j)$ for $0 \le j \le k-2$ in the denominator on the right-hand side of (2.2). Recall that for any labeling $v_1,\dots,v_{k-1}$ of A' and any $0 \le j \le k-2$, the quantity $\Lambda(v_1,\dots,v_j)$ counts the number of choices for $v'_{j+1} \in V(G)$ such that $v_1,\dots,v_j,v'_{j+1}$ is a good sequence. So we clearly have $\Lambda(v_1,\dots,v_j) \le n$.

Suppose that the labeling $v_1,\dots,v_{k-1}$ of A' and $0 \le j \le k-2$ are such that $B \subseteq \{v_1,\dots,v_j\}$. Then $e(\{v_1,\dots,v_j\}) = e(B) = e(A') = \ell - d$. For any choice of $v'_{j+1} \in V(G)$ such that $v_1,\dots,v_j,v'_{j+1}$ is a good sequence, we must have $e(\{v_1,\dots,v_j,v'_{j+1}\}) < \ell$ and therefore

$$|\mathcal{N}(v'_{j+1},\{v_1,\dots,v_j\})| = \deg_{\{v_1,\dots,v_j,v'_{j+1}\}}(v'_{j+1}) = e(\{v_1,\dots,v_j,v'_{j+1}\}) - e(\{v_1,\dots,v_j\}) < d.$$

Recall that for any subfamily $\mathcal{X} \subseteq \mathcal{P}_{<r}(\{v_1,\dots,v_j\})$ with $|\mathcal{X}| = d$, there are precisely $c(\mathcal{X})\cdot n$ vertices $v \in V(G)\setminus A'$ with $\mathcal{N}(v,A') = \mathcal{X} \subseteq \mathcal{P}_{<r}(\{v_1,\dots,v_j\})$. For all these vertices v we have $\mathcal{N}(v,\{v_1,\dots,v_j\}) = \mathcal{N}(v,A') = \mathcal{X}$ and consequently $|\mathcal{N}(v,\{v_1,\dots,v_j\})| = |\mathcal{X}| = d$. Thus, none of these $c(\mathcal{X})\cdot n$ vertices are eligible choices for $v'_{j+1} \in V(G)$ such that $v_1,\dots,v_j,v'_{j+1}$ is a good sequence. So for any subfamily $\mathcal{X} \subseteq \mathcal{P}_{<r}(\{v_1,\dots,v_j\})$ with $|\mathcal{X}| = d$, we find $c(\mathcal{X})\cdot n$ vertices in $v \in V(G)\setminus A'$ that are ineligible to be chosen as $v'_{j+1} \in V(G)$. These $c(\mathcal{X})\cdot n$ vertices are disjoint for different $\mathcal{X} \subseteq \mathcal{P}_{<r}(\{v_1,\dots,v_j\})$, because they were given by the condition $\mathcal{N}(v,A') = \mathcal{X}$. Hence all in all we found at least

$$\sum_{\substack{\mathcal{X}\subseteq\mathcal{P}_{<r}(\{v_1,\dots,v_j\})\\ |\mathcal{X}|=d}} c(\mathcal{X})\cdot n$$

vertices that cannot be chosen as $v'_{j+1} \in V(G)$ such that $v_1,\dots,v_j,v'_{j+1}$ is a good sequence. Hence we can conclude that the number $\Lambda(v_1,\dots,v_j)$ of choices for $v'_{j+1} \in V(G)$ satisfies

$$\begin{aligned}\Lambda(v_1,\dots,v_j) \le n - \sum_{\substack{\mathcal{X}\subseteq\mathcal{P}_{<r}(\{v_1,\dots,v_j\})\\ |\mathcal{X}|=d}} c(\mathcal{X})\cdot n = n\cdot\left(1-\sum_{\substack{\mathcal{X}\subseteq\mathcal{P}_{<r}(\{v_1,\dots,v_j\})\\ |\mathcal{X}|=d}} c(\mathcal{X})\right)\\ \le n\cdot\exp\left(-\sum_{\substack{\mathcal{X}\subseteq\mathcal{P}_{<r}(\{v_1,\dots,v_j\})\\ |\mathcal{X}|=d}} c(\mathcal{X})\right)\end{aligned}$$

if $0 \le j \le k-2$ and $B \subseteq \{v_1,\dots,v_j\}$.

Thus, for every labeling $v_1,\dots,v_{k-1}$ of A', we obtain

$$\begin{aligned}&\Lambda(\emptyset)\cdot\Lambda(v_1)\cdot\Lambda(v_1,v_2)\cdots\Lambda(v_1,\dots,v_{k-2})\\ &\le \prod_{\substack{0\le j\le k-2 \text{ s.t.}\\ B\not\subseteq\{v_1,\dots,v_j\}}} n\cdot \prod_{\substack{0\le j\le k-2 \text{ s.t.}\\ B\subseteq\{v_1,\dots,v_j\}}} n\cdot\exp\left(-\sum_{\substack{\mathcal{X}\subseteq\mathcal{P}_{<r}(\{v_1,\dots,v_j\})\\ |\mathcal{X}|=d}} c(\mathcal{X})\right)\\ &= n^{k-1}\cdot\exp\left(-\sum_{\substack{0\le j\le k-2 \text{ s.t.}\\ B\subseteq\{v_1,\dots,v_j\}}}\ \sum_{\substack{\mathcal{X}\subseteq\mathcal{P}_{<r}(\{v_1,\dots,v_j\})\\ |\mathcal{X}|=d}} c(\mathcal{X})\right)\\ &= n^{k-1}\cdot\exp\left(-\sum_{\substack{\mathcal{X}\subseteq\mathcal{P}_{<r}(A')\\ |\mathcal{X}|=d}}\ \sum_{\substack{0\le j\le k-2 \text{ s.t.}\\ B\cup \mathrm{U}(\mathcal{X})\subseteq\{v_1,\dots,v_j\}}} c(\mathcal{X})\right).\end{aligned}$$

Note that for any given labeling $v_1,\dots,v_{k-1}$ of A' and any subfamily $\mathcal{X} \subseteq \mathcal{P}_{<r}(A')$, each index $0 \le j \le k-2$ satisfies $B\cup\mathrm{U}(\mathcal{X}) \subseteq \{v_1,\dots,v_j\}$ if and only if $v_{j+1} \in A'\setminus(B\cup\mathrm{U}(\mathcal{X}))$ comes after all vertices in $B\cup\mathrm{U}(\mathcal{X})$

in the labeling $v_1,\ldots,v_{k-1}$. Hence the number of indices $0 \le j \le k-2$ with $B \cup \mathrm{U}(\mathcal{X}) \subseteq \{v_1,\ldots,v_j\}$ is precisely the number of vertices $w \in A' \setminus (B \cup \mathrm{U}(\mathcal{X}))$ that come after all vertices in $B \cup \mathrm{U}(\mathcal{X})$ in the labeling $v_1,\ldots,v_{k-1}$. Thus, we can rewrite the previous inequality as

$$\Lambda(\emptyset)\cdot\Lambda(v_1)\cdot\Lambda(v_1,v_2)\cdots\Lambda(v_1,\ldots,v_{k-2}) \le n^{k-1}\cdot\exp\left(-\sum_{\substack{\mathcal{X}\subseteq\mathcal{P}_{<r}(A')\\ |\mathcal{X}|=d}}\ \sum_{\substack{w\in A'\setminus(B\cup\mathrm{U}(\mathcal{X}))\text{ s.t. }w\\ \text{after }B\cup\mathrm{U}(\mathcal{X})\text{ in }v_1,\ldots,v_{k-1}}} c(\mathcal{X})\right)$$

for every labeling $v_1,\ldots,v_{k-1}$ of A'.

Plugging this, together with (2.4), into (2.2) yields

$$\begin{aligned}\rho(v_1,\ldots,v_k) &= \frac{1}{m(A)}\cdot\frac{1}{\Lambda(\emptyset)\cdot\Lambda(v_1)\cdot\Lambda(v_1,v_2)\cdots\Lambda(v_1,\ldots,v_{k-2})}\cdot\frac{1}{\Lambda(v_1,\ldots,v_{k-2})}\\ &\ge \frac{1}{m(A)}\cdot\frac{1}{n^{k-1}}\cdot\exp\left(\sum_{\substack{\mathcal{X}\subseteq\mathcal{P}_{<r}(A')\\ |\mathcal{X}|=d}}\ \sum_{\substack{w\in A'\setminus(B\cup\mathrm{U}(\mathcal{X}))\text{ s.t. }w\\ \text{after }B\cup\mathrm{U}(\mathcal{X})\text{ in }v_1,\ldots,v_{k-1}}} c(\mathcal{X})\right)\cdot\frac{1}{C\cdot n}\\ &= \frac{1}{m(A)\cdot n^k}\cdot\frac{1}{C}\cdot\exp\left(\sum_{\substack{\mathcal{X}\subseteq\mathcal{P}_{<r}(A')\\ |\mathcal{X}|=d}}\ \sum_{\substack{w\in A'\setminus(B\cup\mathrm{U}(\mathcal{X}))\text{ s.t. }w\\ \text{after }B\cup\mathrm{U}(\mathcal{X})\text{ in }v_1,\ldots,v_{k-1}}} c(\mathcal{X})\right)\end{aligned}$$

for every labeling $v_1,\ldots,v_{k-1}$ of A'.

Summing this up for all labelings $v_1,\ldots,v_{k-1}$ of A', we obtain

$$\sum_{\text{labelings } v_1,\ldots,v_{k-1}\text{ of }A'} \rho(v_1,\ldots,v_k) \ge \frac{1}{m(A)\cdot n^k}\cdot\frac{1}{C}\cdot T, \tag{2.5}$$

where we define

$$T := \sum_{\substack{\text{labelings}\\ v_1,\ldots,v_{k-1}\text{ of }A'}} \exp\left(\sum_{\substack{\mathcal{X}\subseteq\mathcal{P}_{<r}(A')\\ |\mathcal{X}|=d}}\ \sum_{\substack{w\in A'\setminus(B\cup\mathrm{U}(\mathcal{X}))\text{ s.t. }w\\ \text{after }B\cup\mathrm{U}(\mathcal{X})\text{ in }v_1,\ldots,v_{k-1}}} c(\mathcal{X})\right).$$

On the other hand, the inequality between arithmetic and geometric mean yields

$$\begin{aligned}T &\ge (k-1)!\cdot\exp\left(\frac{1}{(k-1)!}\sum_{\substack{\text{labelings}\\ v_1,\ldots,v_{k-1}\text{ of }A'}}\ \sum_{\substack{\mathcal{X}\subseteq\mathcal{P}_{<r}(A')\\ |\mathcal{X}|=d}}\ \sum_{\substack{w\in A'\setminus(B\cup\mathrm{U}(\mathcal{X}))\text{ s.t. }w\\ \text{after }B\cup\mathrm{U}(\mathcal{X})\text{ in }v_1,\ldots,v_{k-1}}} c(\mathcal{X})\right)\\ &= (k-1)!\cdot\exp\left(\frac{1}{(k-1)!}\sum_{\substack{\mathcal{X}\subseteq\mathcal{P}_{<r}(A')\\ |\mathcal{X}|=d}}\ \sum_{w\in A'\setminus(B\cup\mathrm{U}(\mathcal{X}))}\ \sum_{\substack{\text{labelings } v_1,\ldots,v_{k-1}\text{ of }A'\\ \text{s.t. }w\text{ after }B\cup\mathrm{U}(\mathcal{X})}} c(\mathcal{X})\right).\end{aligned}$$

We claim that for every subfamily $\mathcal{X} \subseteq \mathcal{P}_{<r}(A')$ and every $w \in A' \setminus (B \cup \mathrm{U}(\mathcal{X}))$ the number of labelings $v_1, \dots, v_{k-1}$ of A' such that w comes after $B \cup \mathrm{U}(\mathcal{X})$ is precisely $(k-1)!/(|B \cup \mathrm{U}(\mathcal{X})|+1)$. Indeed, when picking a labeling $v_1, \dots, v_{k-1}$ of A' uniformly at random, the probability that w is the last element of the set $B \cup \mathrm{U}(\mathcal{X}) \cup \{w\}$ in that labeling is precisely

$$\frac{1}{|B \cup \mathrm{U}(\mathcal{X}) \cup \{w\}|} = \frac{1}{|B \cup \mathrm{U}(\mathcal{X})|+1}.$$

Consequently, the number of labelings $v_1, \dots, v_{k-1}$ of A' such that w comes after $B \cup \mathrm{U}(\mathcal{X})$ is indeed $(k-1)!/(|B \cup \mathrm{U}(\mathcal{X})|+1)$. Thus, we obtain

$$\begin{aligned} T &\geq (k-1)! \cdot \exp\left(\frac{1}{(k-1)!} \sum_{\substack{\mathcal{X} \subseteq \mathcal{P}_{<r}(A') \\ |\mathcal{X}|=d}} \sum_{w \in A' \setminus (B \cup \mathrm{U}(\mathcal{X}))} \frac{(k-1)!}{|B \cup \mathrm{U}(\mathcal{X})|+1} \cdot c(\mathcal{X}) \right) \\ &= (k-1)! \cdot \exp\left(\sum_{\substack{\mathcal{X} \subseteq \mathcal{P}_{<r}(A') \\ |\mathcal{X}|=d}} \sum_{w \in A' \setminus (B \cup \mathrm{U}(\mathcal{X}))} \frac{c(\mathcal{X})}{|B \cup \mathrm{U}(\mathcal{X})|+1} \right) \\ &= (k-1)! \cdot \exp\left(\sum_{\substack{\mathcal{X} \subseteq \mathcal{P}_{<r}(A') \\ |\mathcal{X}|=d}} (k-1-|B \cup \mathrm{U}(\mathcal{X})|) \cdot \frac{c(\mathcal{X})}{|B \cup \mathrm{U}(\mathcal{X})|+1} \right). \end{aligned}$$

Recall that $|B| \leq r \cdot (\ell - d)$. Furthermore observe that for every $\mathcal{X} \subseteq \mathcal{P}_{<r}(A')$ with $|\mathcal{X}| = d$, we have $|\mathrm{U}(\mathcal{X})| \leq (r-1) \cdot |\mathcal{X}| \leq (r-1) \cdot d$. Thus, for every $\mathcal{X} \subseteq \mathcal{P}_{<r}(A')$ with $|\mathcal{X}| = d$, we have $|B \cup \mathrm{U}(\mathcal{X})| \leq r \cdot \ell - d \leq r \cdot \ell - 1$ and consequently $k - 1 - |B \cup X| \geq k - r\ell$. Hence

$$\begin{aligned} T &\geq (k-1)! \cdot \exp\left(\sum_{\substack{\mathcal{X} \subseteq \mathcal{P}_{<r}(A') \\ |\mathcal{X}|=d}} (k - r\ell) \cdot \frac{c(\mathcal{X})}{|B \cup \mathrm{U}(\mathcal{X})|+1} \right) \\ &= (k-1)! \cdot \exp\left((k - r\ell) \cdot \sum_{\substack{\mathcal{X} \subseteq \mathcal{P}_{<r}(A') \\ |\mathcal{X}|=d}} \frac{c(\mathcal{X})}{|B \cup \mathrm{U}(\mathcal{X})|+1} \right) = (k-1)! \cdot \exp\left((k - r\ell) \cdot C \right), \end{aligned}$$

where the last step just used the definition of C. Plugging this into (2.5), we obtain

$$\begin{aligned} \sum_{\text{labelings } v_1, \dots, v_{k-1} \text{ of } A'} \rho(v_1, \dots, v_k) &\geq \frac{1}{m(A) \cdot n^k} \cdot \frac{1}{C} \cdot (k-1)! \cdot \exp\left((k - r\ell) \cdot C \right) \\ &= \frac{(k-1)!}{m(A) \cdot n^k} \cdot \frac{\exp\left((k - r\ell) \cdot C \right)}{C} \end{aligned}$$

Now we use the following lemma, which is a straightforward calculation.

Lemma 2.3. *For all $x \in [0,\infty)$ we have*

$$x \cdot \exp(-(k-r\ell)\cdot x) \leq \frac{1}{k-r\ell} \cdot \frac{1}{e}.$$

Proof. The function $\varphi(x) = x \cdot \exp(-(k-r\ell)\cdot x)$ for $x \in [0,\infty)$ is non-negative, satisfies $\varphi(0) = 0$ and tends to zero for $x \to \infty$ (as $k - r\ell > 0$). Note that φ has derivative

$$\varphi'(x) = \exp(-(k-r\ell)\cdot x) - (k-r\ell)\cdot x \cdot \exp(-(k-r\ell)\cdot x),$$

hence it has its unique maximum at $x = 1/(k-r\ell)$. So we indeed have

$$x \cdot \exp(-(k-r\ell)\cdot x) = \varphi(x) \leq \varphi\left(\frac{1}{k-r\ell}\right) = \frac{1}{k-r\ell} \cdot \frac{1}{e}$$

for all $x \in [0,\infty)$. □

Applying Lemma 2.3 to $x = C$, we obtain

$$\frac{\exp((k-r\ell)\cdot C)}{C} = \frac{1}{C \cdot \exp(-(k-r\ell)\cdot C)} \geq (k-r\ell)\cdot e.$$

Thus,

$$\sum_{\text{labelings } v_1,\dots,v_{k-1} \text{ of } A'} \rho(v_1,\dots,v_k) \geq \frac{(k-1)!}{m(A)\cdot n^k} \cdot \frac{\exp((k-r\ell)\cdot C)}{C}$$
$$\geq \frac{(k-1)!}{m(A)\cdot n^k} \cdot (k-r\ell)\cdot e = \frac{1}{m(A)} \cdot \frac{k-r\ell}{k} \cdot e \cdot \frac{k!}{n^k},$$

which proves (2.1). This finishes the proof of Theorem 1.4.

3 Proof of Theorems 1.3, 1.5 and 1.6

The proofs of Theorem 1.3, 1.5 and 1.6 are based on the following three lemmas.

Lemma 3.1. *Let r, k and ℓ be positive integers and let G be a hypergraph on $n \geq k$ vertices all of whose edges have size at most r. Then for each real number c with $0 < c < \sqrt{k}/2$, the following holds: The number of k-vertex subsets $A \subseteq V(G)$ with $e(A) = \ell$ and $c \leq m(A) \leq \sqrt{k}/2$ is at most*

$$32 \cdot \sqrt{r} \cdot \frac{1}{\sqrt{c}} \cdot \frac{n^k}{k!}.$$

Lemma 3.2. *Let r, k and ℓ be positive integers and let G be a hypergraph on $n \geq k$ vertices all of whose edges have size at most r. Then for each real number c satisfying $\sqrt{k}/2 \leq c \leq k/(32r)$, the following holds: The number of k-vertex subsets $A \subseteq V(G)$ with $e(A) = \ell$ and $c \leq m(A) \leq 2c$ is at most*

$$44 \cdot \sqrt{r} \cdot k^{-1/4} \cdot \frac{n^k}{k!}.$$

Lemma 3.3. *Let r, k and ℓ be positive integers and let G be a hypergraph on $n \geq k$ vertices all of whose edges have size at most r. Then for each real number ε with $0 < \varepsilon < 1/2$, the following holds: The number of k-vertex subsets $A \subseteq V(G)$ with $e(A) = \ell$ and $\varepsilon k \leq m(A) \leq (1-\varepsilon)k$ is at most*

$$8 \cdot r^{1/4} \cdot \varepsilon^{-1/2} \cdot k^{-1/4} \cdot \frac{n^k}{k!}.$$

We will prove these three lemmas in the following two sections.

Lemmas 3.1, 3.2 and 3.3 imply the following corollary, from which we will deduce Theorems 1.3 and 1.5. Corollary 3.4 will also be used again in the proof of Theorem 1.2 in Section 6.

Corollary 3.4. *Let us fix a positive integer r and a real number $\varepsilon > 0$. Assume that k is an integer which is sufficiently large with respect to ε. Let ℓ be a positive integer and G a hypergraph on $n \geq k$ vertices all of whose edges have size at most r. Then for each real number $c' > 0$, there are at most*

$$\left(32 \cdot \sqrt{r} \cdot \frac{1}{\sqrt{c'}} + 23 \cdot \sqrt{r} \cdot k^{-1/4} \cdot \log k\right) \cdot \frac{n^k}{k!}$$

k-vertex subsets $A \subseteq V(G)$ with $e(A) = \ell$ and $c' \leq m(A) \leq (1-\varepsilon)k$.

Proof. We may assume without loss of generality that $\varepsilon < 1/(32r) < 1/2$, because otherwise we can just make ε smaller.

Let us divide the k-vertex subsets $A \subseteq V(G)$ with $e(A) = \ell$ and $c' \leq m(A) \leq (1-\varepsilon)k$ into three groups. The first group consists of those subsets A with $c' \leq m(A) < \sqrt{k}/2$, the second group of those with $\sqrt{k}/2 \leq m(A) < \varepsilon k$ and finally the third group of those with $\varepsilon k \leq m(A) \leq (1-\varepsilon)k$. We will bound the number of k-vertex subsets in each of these three groups separately, in each case using one of the three lemmas above.

For the first group, we claim that the number of k-vertex subsets $A \subseteq V(G)$ with $e(A) = \ell$ and $c' \leq m(A) < \sqrt{k}/2$ is at most

$$32 \cdot \sqrt{r} \cdot \frac{1}{\sqrt{c'}} \cdot \frac{n^k}{k!}.$$

If $c' \geq \sqrt{k}/2$, then there cannot be any such k-vertex subset $A \subseteq V(G)$, so this claim is trivially true. If $c' < \sqrt{k}/2$, then the claim follows directly from Lemma 3.1.

For the second group, we claim that the number of k-vertex subsets $A \subseteq V(G)$ with $e(A) = \ell$ and $\sqrt{k}/2 \leq m(A) < \varepsilon k$ is at most

$$22 \cdot \sqrt{r} \cdot k^{-1/4} \cdot \log k \cdot \frac{n^k}{k!}.$$

As k is sufficiently large with respect to ε, we may assume that $\varepsilon\sqrt{k} \geq 2$. Let us apply Lemma 3.2 several times with different values of c. More precisely, let us take

$$c_j = \frac{\sqrt{k}}{2} \cdot 2^j,$$

for $j = 0, 1, \dots, \lceil \log(\varepsilon\sqrt{k}) \rceil$. For each of these j, we have $c_j \geq \sqrt{k}/2$ and also

$$c_j \leq \frac{\sqrt{k}}{2} \cdot 2^{\lceil \log(\varepsilon\sqrt{k}) \rceil} \leq \frac{\sqrt{k}}{2} \cdot 2^{\log(\varepsilon\sqrt{k})+1} = \sqrt{k} \cdot (\varepsilon\sqrt{k}) = \varepsilon k < \frac{k}{32r},$$

where we used the assumption $\varepsilon < 1/(32r)$. Thus, all these values c_j satisfy the assumptions of Lemma 3.2. Thus, by Lemma 3.2, for each $j = 0, 1, \dots, \lceil \log(\varepsilon\sqrt{k}) \rceil$ the number of k-vertex subsets $A \subseteq V(G)$ with $e(A) = \ell$ and $(\sqrt{k}/2) \cdot 2^j = c_j \le m(A) \le 2c_j = (\sqrt{k}/2) \cdot 2^{j+1}$ is at most

$$44 \cdot \sqrt{r} \cdot k^{-1/4} \cdot \frac{n^k}{k!}.$$

Adding this up for all $j = 0, 1, \dots, \lceil \log(\varepsilon\sqrt{k}) \rceil$, we see that the total number of k-vertex subsets $A \subseteq V(G)$ with $e(A) = \ell$ and $\sqrt{k}/2 \le m(A) < (\sqrt{k}/2) \cdot 2^{\log(\varepsilon\sqrt{k})+1} = \varepsilon k$ is at most

$$\Big(\lceil \log(\varepsilon\sqrt{k}) \rceil + 1\Big) \cdot 44 \cdot \sqrt{r} \cdot k^{-1/4} \cdot \frac{n^k}{k!} < \frac{1}{2}\log k \cdot 44 \cdot \sqrt{r} \cdot k^{-1/4} \cdot \frac{n^k}{k!} = 22 \cdot \sqrt{r} \cdot k^{-1/4} \cdot \log k \cdot \frac{n^k}{k!},$$

where we used that $\lceil \log(\varepsilon\sqrt{k}) \rceil + 1 < \log(\varepsilon\sqrt{k}) + 2 \le \log(\sqrt{k}/32) + 2 < \log\sqrt{k} = \frac{1}{2}\log k$. This indeed proves our claim about the second group.

Finally, for the third group, Lemma 3.3 implies that the number of k-vertex subsets $A \subseteq V(G)$ with $e(A) = \ell$ and $\varepsilon k \le m(A) \le (1-\varepsilon)k$ is at most

$$8 \cdot r^{1/4} \cdot \varepsilon^{-1/2} \cdot k^{-1/4} \cdot \frac{n^k}{k!} \le \sqrt{r} \cdot k^{-1/4} \cdot \log k \cdot \frac{n^k}{k!}.$$

Here we used that k is sufficiently large with respect to ε, so that $\log k \ge 8 \cdot \varepsilon^{-1/2}$.

All in all, adding up our bounds for each of the three groups, we obtain that the total number of k-vertex subsets $A \subseteq V(G)$ with $e(A) = \ell$ and $c' \le m(A) \le (1-\varepsilon)k$ is at most

$$\left(32 \cdot \sqrt{r} \cdot \frac{1}{\sqrt{c'}} + 23 \cdot \sqrt{r} \cdot k^{-1/4} \cdot \log k\right) \cdot \frac{n^k}{k!},$$

as desired. □

In order to deduce Theorem 1.3 from Corollary 3.4, we will use the following claim.

Claim 3.5. *For every graph G and every subset $A \subseteq V(G)$ of its vertex set, we have*

$$\sqrt{2e(A)} \le m(A) \le 2e(A).$$

Proof. The inequality $m(A) \le 2e(A)$ follows directly from the definition of $m(A)$, because every vertex that is non-isolated in A needs to be incident with at least one of the $e(A)$ edges within A.

For the other inequality, note that the $e(A)$ edges inside A are all between the $m(A)$ non-isolated vertices. Therefore we must have

$$e(A) \le \binom{m(A)}{2} \le \frac{m(A)^2}{2},$$

which implies $m(A) \ge \sqrt{2e(A)}$. □

Now we are ready to prove Theorem 1.3

Proof of Theorem 1.3. Let $\varepsilon > 0$ be a fixed real number and let k be sufficiently large (with respect to ε). Furthermore, let ℓ be an integer with $1 \le \ell \le (1-\varepsilon)k/2$.

Now, consider any graph G on $n \ge k$ vertices. By Claim 3.5, for any subset $A \subseteq V(G)$ with $e(A) = \ell$ we have $\sqrt{2\ell} \le m(A) \le 2\ell \le (1-\varepsilon)k$. Thus, by Corollary 3.4 applied to $r = 2$ and $c' = \sqrt{2l}$ the number of subsets $A \subseteq V(G)$ with $e(A) = \ell$ is at most

$$\left(32 \cdot \sqrt{2} \cdot (2\ell)^{-1/4} + 23 \cdot \sqrt{2} \cdot k^{-1/4} \cdot \log k\right) \cdot \frac{n^k}{k!} \le \left(48 \cdot \ell^{-1/4} + 36 \cdot k^{-1/4} \cdot \log k\right) \cdot (1 + o_n(1)) \cdot \binom{n}{k}.$$

Here we used $\sqrt{2} \le 3/2$. Thus, we can conclude that

$$\mathrm{ind}(k,\ell) \le 48 \cdot \ell^{-1/4} + 36 \cdot k^{-1/4} \cdot \log k.$$

If $1 \le \ell \le k/\log^4 k$, then $\ell^{-1/4} \ge k^{-1/4} \cdot \log k$ and so we can conclude

$$\mathrm{ind}(k,\ell) \le 48 \cdot \ell^{-1/4} + 36 \cdot k^{-1/4} \cdot \log k \le 90 \cdot \ell^{-1/4}.$$

On the other hand, if $k/\log^4 k \le \ell \le (1-\varepsilon)k/2$, then $\ell^{-1/4} \le k^{-1/4} \cdot \log k$ and we obtain

$$\mathrm{ind}(k,\ell) \le 48 \cdot \ell^{-1/4} + 36 \cdot k^{-1/4} \cdot \log k \le 90 \cdot k^{-1/4} \cdot \log k.$$

This finishes the proof of Theorem 1.3. □

Now, let us turn to proving Theorem 1.5. Here, we will use the following claim.

Claim 3.6. *Let G be a hypergraph all of whose edges have size at most r. Then for every subset $A \subseteq V(G)$ of its vertex set with $e(A) \ge 2^{2r}$, we have*

$$\frac{1}{6} \cdot r \cdot e(A)^{1/r} \le m(A) \le r \cdot e(A).$$

Proof. Again, the inequality $m(A) \le r \cdot e(A)$ follows directly from the definition of $m(A)$, because every vertex that is non-isolated in A needs to be part of at least one of the $e(A)$ edges within A.

For the other inequality, note that the $e(A)$ edges inside A are subsets of the set consisting of the $m(A)$ non-isolated vertices in A. Since all edges have size at least 1 and at most r, we obtain

$$e(A) \le \binom{m(A)}{r} + \binom{m(A)}{r-1} + \dots + \binom{m(A)}{1}.$$

If $m(A) < 2r$, this yields $e(A) \le 2^{m(A)} < 2^{2r}$, a contradiction. Hence we must have $m(A) \ge 2r$. Then $\binom{m(A)}{r}$ is the largest of the binomial coefficients in the inequality above and we can conclude

$$e(A) \le r \cdot \binom{m(A)}{r} \le 2^r \cdot \binom{m(A)}{r} \le 2^r \cdot \left(\frac{e \cdot m(A)}{r}\right)^r = \left(\frac{2e \cdot m(A)}{r}\right)^r \le \left(\frac{6 \cdot m(A)}{r}\right)^r.$$

Thus, $m(A) \ge r \cdot e(A)^{1/r}/6$, as desired. □

Now it is easy to deduce Theorem 1.5 from Corollary 3.4.

Proof of Theorem 1.5. Let $r \geq 3$ and $\varepsilon > 0$ be fixed and let k be sufficiently large (with respect to r and ε). Furthermore, let ℓ be an integer with $1 \leq \ell \leq (1-\varepsilon)k/r$.

Note that we certainly have $\mathrm{ind}_{\leq r}(k,\ell) \leq 1$, so the bound on $\mathrm{ind}_{\leq r}(k,\ell) \leq 1$ claimed in Theorem 1.5 is trivial if $\ell \leq 100^{2r}$. So let us from now on assume that $\ell > 100^{2r}$

Consider any hypergraph G on $n \geq k$ vertices all of whose edges have size at most r. By Claim 3.6, for any subset $A \subseteq V(G)$ with $e(A) = \ell$ we have $\frac{1}{6} \cdot r \cdot \ell^{1/r} \leq m(A) \leq r \cdot \ell \leq (1-\varepsilon)k$. Thus, by Corollary 3.4 applied to $c' = \frac{1}{6} \cdot r \cdot \ell^{1/r}$ the number of subsets $A \subseteq V(G)$ with $e(A) = \ell$ is at most

$$\left(32 \cdot \sqrt{r} \cdot \left(\frac{1}{6} \cdot r \cdot \ell^{1/r}\right)^{-1/2} + 23 \cdot \sqrt{r} \cdot k^{-1/4} \cdot \log k\right) \cdot \frac{n^k}{k!}$$
$$\leq \left(96 \cdot \ell^{-1/(2r)} + 23 \cdot \sqrt{r} \cdot k^{-1/4} \cdot \log k\right) \cdot \frac{n^k}{k!} \leq 97 \cdot \ell^{-1/(2r)} \cdot (1 + o_n(1)) \cdot \binom{n}{k}.$$

Here we used $\sqrt{6} \leq 3$ and $23 \cdot \sqrt{r} \cdot k^{-1/4} \cdot \log k \leq k^{-1/(2r)} \leq \ell^{-1/(2r)}$ for k sufficiently large (recall that $r \geq 3$). Thus, we can conclude that

$$\mathrm{ind}_r(k,\ell) \leq \mathrm{ind}_{\leq r}(k,\ell) \leq 97 \cdot \ell^{-1/(2r)} \leq 100 \cdot \ell^{-1/(2r)},$$

as desired. □

Finally, Theorem 1.6 is an easy consequence of Lemma 3.1.

Proof of Theorem 1.6. As in the theorem statement, let k and ℓ be positive integers with $\ell \leq \sqrt{k}/4$ and let G be a graph on $n \geq k$ vertices. Note that every k-vertex subset $A \subseteq V(G)$ that induces a forest with exactly ℓ edges satisfies $\ell \leq \ell + 1 \leq m(A) \leq 2\ell \leq \sqrt{k}/2$. Hence, by Lemma 3.1 applied to $r = 2$ and $c = \ell$, the number of k-vertex subsets $A \subseteq V(G)$ that induce a forest with exactly ℓ edges is at most

$$32 \cdot \sqrt{2} \cdot \ell^{-1/2} \cdot \frac{n^k}{k!} \leq 50 \cdot \ell^{-1/2} \cdot \frac{n^k}{k!},$$

where we used that $\sqrt{2} \leq 3/2$. □

The next two sections will be devoted to proving Lemmas 3.1, 3.2 and 3.3. All three of these proofs follow the same general strategy. We will first prove Lemma 3.3 in Section 4, because the proof is slightly easier than the proofs of Lemmas 3.1 and 3.2. Afterwards, we will proof Lemmas 3.1 and 3.2 in Section 5. Let us now introduce some notation that will be relevant for the proofs of all three of the lemmas:

For a hypergraph G and a subset $B \subseteq V(G)$, call a vertex $v \in V(G) \setminus B$ *connected to* B if there is an edge $e \in E(G)$ with $v \in e$ and $e \setminus \{v\} \subseteq B$. Note that in the case where the singleton set $\{v\}$ is an edge of G, the vertex v is connected to every subset $B \subseteq V(G) \setminus \{v\}$. Also note that in the case where G is a graph, v being connected to B simply means that v has an edge to a vertex in B.

For a pair of subsets $B \subseteq A \subseteq V(G)$, let $h(A,B)$ denote the number of vertices $v \in A \setminus B$ that are connected to B. Furthermore, let $m(A,B)$ denote the number of vertices $v \in A \setminus B$ that are non-isolated in A. Clearly, $m(A,B) \leq m(A)$. Also note that every vertex $v \in A \setminus B$ that is connected to B is automatically non-isolated in A. Hence $h(A,B) \leq m(A,B)$. Finally, let $f(A,B) = m(A,B) - h(A,B)$ be the number of vertices $v \in A \setminus B$ that are non-isolated in A but not connected to B.

4 Proof of Lemma 3.3

In this section, we will prove Lemma 3.3. So let us fix positive integers r, k and ℓ and let G be a hypergraph on $n \geq k$ vertices all of whose edges have size at most r.

The following definition introduces the key concept for our proof. Recall that we defined the quantities $h(A,B)$, $m(A,B)$ and $f(A,B)$ at the end of Section 3.

Definition 4.1. For a real number $0 < \varepsilon < 1/2$, let a pair (A,B) of subsets $B \subseteq A \subseteq V(G)$ be called ε-*pleasant* if all of the following conditions are satisfied:

(i) $|A| = k$ and $e(A) = \ell$.

(ii) $f(A,B) = 0$, meaning that every vertex in $A \setminus B$ is either connected to B or isolated in A.

(iii) $h(A,B) \geq \frac{1}{4} r^{-1/2} \cdot \varepsilon \cdot \sqrt{k}$.

(iv) $|A \setminus B| - h(A,B) \geq \frac{1}{4} r^{-1/2} \cdot \varepsilon \cdot \sqrt{k}$.

The crucial insight for the proof of Lemma 3.3 is that for each fixed subset $B \subseteq V(G)$ we can bound the number of sets $A \subseteq V(G)$ such that (A,B) is ε-pleasant.

Lemma 4.2. *Let $0 < \varepsilon < 1/2$ be a real number and let $B \subseteq V(G)$ be a subset of size $|B| \leq k$. Let $a = k - |B|$. Then there are at most*

$$2 \cdot r^{1/4} \cdot \varepsilon^{-1/2} \cdot k^{-1/4} \cdot \frac{n^a}{a!}$$

different sets $A \subseteq V(G)$ such that (A,B) is an ε-pleasant pair.

We will prove Lemma 4.2 in the second subsection. First, we will prove Lemma 3.3 assuming Lemma 4.2 in the first subsection.

The idea of the proof of Lemma 3.3 is to show that for every k-vertex subset $A \subseteq V(G)$ with the properties specified in the lemma, we have a sufficiently high probability of generating an ε-pleasant pair (A,B) when choosing a random subset $B \subseteq A$ in an appropriate way. Combining this with Lemma 4.2 will give an upper bond for the number of such sets A.

4.1 Main proof

In this subsection we will prove Lemma 3.3 assuming Lemma 4.2. As in the statement of Lemma 3.3, let ε be a real number with $0 < \varepsilon < 1/2$.

Note that if $\varepsilon^2 \cdot k \leq 2^{12} \cdot r$, then

$$8 \cdot r^{1/4} \cdot \varepsilon^{-1/2} \cdot k^{-1/4} \cdot \frac{n^k}{k!} \geq \frac{n^k}{k!} \geq \binom{n}{k}.$$

So in this case Lemma 3.3 is trivially true as the total number of k-vertex subsets of $V(G)$ is only $\binom{n}{k}$. Thus, we can assume that $\varepsilon^2 \cdot k \geq 2^{12} \cdot r$.

As indicated at the end of the previous subsection, we start by proving that for each fixed k-vertex subset $A \subseteq V(G)$ with the properties in Lemma 3.3, we have a probability of at least $\frac{1}{2}$ to obtain an ε-pleasant pair (A,B) when choosing a random subset $B \subseteq A$ by picking each element of A independently with some appropriately chosen probability p.

Lemma 4.3. *Let $p = 1 - \frac{1}{2}r^{-1/2}k^{-1/2}$. Then for every k-vertex subset $A \subseteq V(G)$ with $e(A) = \ell$ and $\varepsilon k \leq m(A) \leq (1-\varepsilon)k$, the following holds: If we choose a subset $B \subseteq A$ randomly by taking each element of A into B with probability p (independently for all elements of A), then with probability at least $\frac{1}{2}$ the resulting pair (A,B) is ε-pleasant.*

Proof. Fix a subset $A \subseteq V(G)$ with $|A| = k$ as well as $e(A) = \ell$ and $\varepsilon k \leq m(A) \leq (1-\varepsilon)k$. As in the lemma, let the subset $B \subseteq A$ be chosen by randomly picking elements of A independently with probability p. Let $q = 1 - p = \frac{1}{2}r^{-1/2}k^{-1/2}$, then each element of A will be in $A \setminus B$ with probability q. We need to prove that the pair (A,B) satisfies the conditions in Definition 4.1 with probability at least $\frac{1}{2}$.

First, note that condition (i) in Definition 4.1 is automatically satisfied by the assumptions on A.

Next, we claim that condition (ii) holds with probability at least $\frac{3}{4}$. For condition (ii) to fail, there needs to be a vertex $v \in A \setminus B$ such that v is non-isolated in A but not connected to B. Each vertex $v \in A$ that is non-isolated in A is, by definition, contained in at least one edge $e \subseteq A$. For each fixed such v, the probability for $v \in A \setminus B$ is precisely q. Let us pick one edge $e \subseteq A$ with $v \in e$. For v to cause condition (ii) to fail, furthermore at least one vertex in $e \setminus \{v\}$ needs to be in $A \setminus B$ as well (otherwise v is connected to B). As $|e \setminus \{v\}| \leq r-1$, the probability of this happening is at most $(r-1)\cdot q$ (and this is independent of whether $v \in A \setminus B$). Thus, the total probability of v causing condition (ii) to fail is at most $q \cdot (r-1) \cdot q = (r-1)\cdot q^2$. Summing this over all the at most k choices of a non-isolated vertex $v \in A$, we see that the total probability of condition (ii) failing is at most

$$k \cdot (r-1) \cdot q^2 \leq k \cdot r \cdot q^2 = k \cdot r \cdot \frac{1}{4} \cdot \frac{1}{r \cdot k} = \frac{1}{4}.$$

Thus, condition (ii) indeed holds with probability at least $\frac{3}{4}$.

Note that if condition (ii) holds, then we have $h(A,B) = m(A,B) - f(A,B) = m(A,B)$. Therefore, in order to check conditions (iii) and (iv) we can consider $m(A,B)$ instead of $h(A,B)$.

We claim that with probability at least $\frac{7}{8}$ we have

$$m(A,B) \geq \frac{1}{2} \cdot q \cdot m(A). \tag{4.1}$$

Recall that $m(A,B)$ counts the number of vertices in $A \setminus B$ that are non-isolated in A. There are $m(A)$ vertices of A that are non-isolated in A and each of them ends up in $A \setminus B$ with probability q (and all independently). Thus, $m(A,B)$ is binomially distributed with parameters $m(A)$ and q. So using the standard Chernoff bound for the lower tail of binomially distributed variables (see, for example, [2, Theorem A.1.13]), we have

$$\mathbb{P}\left[m(A,B) < \frac{1}{2} \cdot q \cdot m(A)\right] < \exp\left(-\frac{\left(\frac{1}{2} \cdot q \cdot m(A)\right)^2}{2 \cdot q \cdot m(A)}\right) = \exp\left(-\frac{q}{8} \cdot m(A)\right) \leq \exp\left(-\frac{q}{8} \cdot \varepsilon \cdot k\right)$$
$$= \exp\left(-\frac{1}{16} \cdot \frac{1}{\sqrt{r} \cdot \sqrt{k}} \cdot \varepsilon \cdot k\right) = \exp\left(-\frac{1}{16} \cdot \frac{\varepsilon \cdot \sqrt{k}}{\sqrt{r}}\right) \leq \exp\left(-\frac{1}{16} \cdot 2^6\right) = e^{-4} \leq \frac{1}{8}$$

where we used $m(A) \geq \varepsilon k$ and $\varepsilon^2 \cdot k \geq 2^{12} \cdot r$. So (4.1) indeed holds with probability at least $\frac{7}{8}$.

Similarly, we claim that with probability at least $\frac{7}{8}$ we have

$$|A \setminus B| - m(A,B) \geq \frac{1}{2} \cdot q \cdot (k - m(A)). \tag{4.2}$$

Note that $|A \setminus B| - m(A,B)$ counts the number of vertices in $A \setminus B$ that are isolated in A. There are $k - m(A)$ vertices of A that are isolated in A and each of them ends up in $A \setminus B$ with probability q (and all independently). Thus, $|A \setminus B| - m(A,B)$ is binomially distributed with parameters $k - m(A)$ and q. So using the same computation as for inequality (4.1), we see that

$$\mathbb{P}\left[|A \setminus B| - m(A,B) < \frac{1}{2} \cdot q \cdot (k - m(A))\right] < \exp\left(-\frac{\left(\frac{1}{2} \cdot q \cdot (k - m(A))\right)^2}{2 \cdot q \cdot (k - m(A))}\right)$$
$$= \exp\left(-\frac{q}{8} \cdot (k - m(A))\right) \leq \exp\left(-\frac{q}{8} \cdot \varepsilon \cdot k\right) \leq \frac{1}{8}$$

where this time we used $m(A) \leq (1 - \varepsilon)k$. So (4.2) indeed holds with probability at least $\frac{7}{8}$.

All in all, condition (ii) fails with probability at most $\frac{1}{4}$, and each of (4.1) and (4.2) fails with probability at most $\frac{1}{8}$. Thus, with probability at least $\frac{1}{2}$ condition (ii) and both of the inequalities (4.1) and (4.2) hold. We claim that in this case the pair (A,B) is ε-pleasant. Indeed, we already saw that condition (i) is always satisfied by the assumptions on A. Furthermore, condition (ii) implies $h(A,B) = m(A,B) - f(A,B) = m(A,B)$. Thus, (4.1) gives (using $m(A) \geq \varepsilon k$)

$$h(A,B) = m(A,B) \geq \frac{1}{2} \cdot q \cdot m(A) \geq \frac{1}{2} \cdot q \cdot \varepsilon \cdot k = \frac{1}{4} \cdot \frac{1}{\sqrt{r} \cdot \sqrt{k}} \cdot \varepsilon \cdot k = \frac{1}{4} r^{-1/2} \cdot \varepsilon \cdot \sqrt{k}.$$

So condition (iii) holds. Finally, condition (iv) is satisfied, because (4.2) yields (using $m(A) \leq (1 - \varepsilon)k$ this time)

$$|A \setminus B| - h(A,B) = |A \setminus B| - m(A,B) \geq \frac{1}{2} \cdot q \cdot (k - m(A)) \geq \frac{1}{2} \cdot q \cdot \varepsilon \cdot k = \frac{1}{4} r^{-1/2} \cdot \varepsilon \cdot \sqrt{k}.$$

Thus, the pair (A,B) is indeed ε-pleasant if condition (ii) as well as the inequalities (4.1) and (4.2) hold. Hence we have shown that (A,B) is an ε-pleasant pair with probability at least $\frac{1}{2}$. $\square$

Let us now prove Lemma 3.3. So suppose for contradiction that there are more than

$$8 \cdot r^{1/4} \cdot \varepsilon^{-1/2} \cdot k^{-1/4} \cdot \frac{n^k}{k!}.$$

k-vertex subsets $A \subseteq V(G)$ with $e(A) = \ell$ and $\varepsilon k \leq m(A) \leq (1 - \varepsilon)k$.

Define $p = 1 - \frac{1}{2} r^{-1/2} k^{-1/2}$ as in Lemma 4.3. Now, let us choose a pair (A,B) of subsets $B \subseteq A \subseteq V(G)$ with $|A| = k$ randomly as follows: First, choose $A \subseteq V(G)$ uniformly at random among all subsets of size $|A| = k$. Then, choose a subset $B \subseteq A$ by taking each element of A into B with probability p (independently for all elements of A).

By our assumption above, the probability that A satisfies $e(A)=\ell$ and $\varepsilon k \le m(A) \le (1-\varepsilon)k$ is at least

$$8\cdot r^{1/4}\cdot \varepsilon^{-1/2}\cdot k^{-1/4}\cdot \frac{n^k}{k!}\cdot \binom{n}{k}^{-1} = 8\cdot r^{1/4}\cdot \varepsilon^{-1/2}\cdot k^{-1/4}\cdot \frac{n^k}{(n)_k}.$$

If this is the case, then by Lemma 4.3 the pair (A,B) is ε-pleasant with probability at least $\frac{1}{2}$. Thus, for the total probability that (A,B) is an ε-pleasant pair, we obtain

$$\mathbb{P}[(A,B) \text{ is } \varepsilon\text{-pleasant}] \ge 4\cdot r^{1/4}\cdot \varepsilon^{-1/2}\cdot k^{-1/4}\cdot \frac{n^k}{(n)_k}. \tag{4.3}$$

On the other hand, we can find an upper bound for the probability that (A,B) is ε-pleasant by conditioning on B. Note that B will always be a subset $B\subseteq V(G)$ with $|B|\le k$. So let us fix any subset $B\subseteq V(G)$ with $|B|\le k$ and let $a=k-|B|$. If we condition on having this particular set B as the result of our random draw of the pair (A,B), then there are

$$\binom{n-|B|}{k-|B|} = \binom{n-|B|}{a} = \frac{(n-|B|)_a}{a!}$$

possibilities for the set A and they are all equally likely (recall that we choose A uniformly at random among all subsets $A\subseteq V(G)$ with $|A|=k$). But by Lemma 4.2 at most

$$2\cdot r^{1/4}\cdot \varepsilon^{-1/2}\cdot k^{-1/4}\cdot \frac{n^a}{a!}$$

of these possible sets A have the property that (A,B) is ε-pleasant. Thus, when we condition on having this set B, the probability for (A,B) to be ε-pleasant is at most

$$2\cdot r^{1/4}\cdot \varepsilon^{-1/2}\cdot k^{-1/4}\cdot \frac{n^a}{(n-|B|)_a} = 2\cdot r^{1/4}\cdot \varepsilon^{-1/2}\cdot k^{-1/4}\cdot \frac{n^{k-|B|}}{(n-|B|)_{k-|B|}} \le 2\cdot r^{1/4}\cdot \varepsilon^{-1/2}\cdot k^{-1/4}\cdot \frac{n^k}{(n)_k}.$$

Since the set B was arbitrary, this shows that the total probability for (A,B) to be an ε-pleasant pair satisfies

$$\mathbb{P}[(A,B) \text{ is } \varepsilon\text{-pleasant}] \le 2\cdot r^{1/4}\cdot \varepsilon^{-1/2}\cdot k^{-1/4}\cdot \frac{n^k}{(n)_k}.$$

But now we have a contradiction to (4.3). This finishes the proof of Lemma 3.3.

4.2 Proof of Lemma 4.2

Now, let us finally prove Lemma 4.2. As in the statement of the lemma, fix $0<\varepsilon<1/2$ as well as a subset $B\subseteq V(G)$ of size $|B|\le k$, and let $a=k-|B|$.

Note that for every ε-pleasant pair (A,B) we have $|A|=k$ by condition (i) in Definition 4.1 and therefore $|A\setminus B| = k-|B| = a$.

The following definition introduces a central notion for our proof of Lemma 4.2.

Definition 4.4. Let us call a sequence $v_1,\dots,v_a$ of distinct vertices in $V(G)\setminus B$ *tidy* if $e(B\cup\{v_1,\dots,v_a\}) = \ell$ and if there exists an index $h\in\{1,\dots,a-1\}$ with

$$\frac{\varepsilon\cdot\sqrt{k}}{4\sqrt{r}} \le h \le a - \frac{\varepsilon\cdot\sqrt{k}}{4\sqrt{r}}$$

and such that

- each of the vertices $v_1,\dots,v_h$ is connected to B, and
- each of the vertices $v_{h+1},\dots,v_a$ is isolated in $B\cup\{v_1,\dots,v_a\}$.

Note that for a sequence $v_1,\dots,v_a$ as in Definition 4.4, the vertices $v_{h+1},\dots,v_a$ cannot be connected to B (since otherwise they would be non-isolated in $B\cup\{v_1,\dots,v_a\}$). So the number $h(B\cup\{v_1,\dots,v_a\},B)$ of vertices in $\{v_1,\dots,v_a\}$ that are connected to B is precisely h. Thus, for every tidy sequence $v_1,\dots,v_a$, the index h in Definition 4.4 must be equal to $h(B\cup\{v_1,\dots,v_a\},B)$.

An important observation is that for every set $A\subseteq V(G)$ such that (A,B) is ε-pleasant, we have $|A\setminus B| = a$ and we can label the vertices of $A\setminus B$ as $v_1,\dots,v_a$ in such a way that $v_1,\dots,v_a$ is a tidy sequence. In fact, there are many ways to label the vertices of $A\setminus B$ as $v_1,\dots,v_a$ so that $v_1,\dots,v_a$ is a tidy sequence. The following claim makes this precise.

Claim 4.5. *Let $A\subseteq V(G)$ be such that (A,B) is an ε-pleasant pair. Then there are at least*

$$h(A,B)!\cdot(a-h(A,B))!$$

different ways to label the vertices of $A\setminus B$ as $v_1,\dots,v_a$ such that $v_1,\dots,v_a$ is a tidy sequence.

Proof. First, note that by condition (i) in Definition 4.1 we have $|A\setminus B| = k-|B| = a$.

Let $h = h(A,B)$ and note that by conditions (iii) and (iv) in Definition 4.1 we have

$$\frac{\varepsilon\cdot\sqrt{k}}{4\sqrt{r}} \le h \le |A\setminus B| - \frac{\varepsilon\cdot\sqrt{k}}{4\sqrt{r}} = a - \frac{\varepsilon\cdot\sqrt{k}}{4\sqrt{r}}.$$

In particular $0<h<a$. As $h=h(A,B)$ is an integer, this implies $1\le h\le a-1$.

There are exactly $h=h(A,B)$ vertices in $A\setminus B$ that are connected to B. There are $h!$ ways to label these vertices as $v_1,\dots,v_h$. By condition (ii) in Definition 4.1, the remaining $a-h$ vertices in $A\setminus B$ are all isolated in A. There are $(a-h)!$ ways to label these vertices as $v_{h+1},\dots,v_a$. It is easy to see that for each labeling of $A\setminus B$ as $v_1,\dots,v_a$ obtained in this way, the sequence $v_1,\dots,v_a$ together with the index h satisfies the conditions in Definition 4.1.

All in all, this gives $h!\cdot(a-h)! = h(A,B)!\cdot(a-h(A,B))!$ labelings of $A\setminus B$ as $v_1,\dots,v_a$ such that $v_1,\dots,v_a$ is a tidy sequence. □

In particular, for each subset $A\subseteq V(G)$ such that (A,B) is an ε-pleasant pair, there is at least one way to label the vertices of $A\setminus B$ as $v_1,\dots,v_a$ such that $v_1,\dots,v_a$ is a tidy sequence. Thus, we may assume that there exists at least one tidy sequence $v_1,\dots,v_a$, as otherwise there are no sets $A\subseteq V(G)$ such that (A,B) is an ε-pleasant pair (and then Lemma 4.2 is trivially true).

Note that for every tidy sequence $v_1,\dots,v_a$ (with the index h as in Definition 4.4), we must have $e(B\cup\{v_1,\dots,v_h\}) = e(B\cup\{v_1,\dots,v_a\}) = \ell$, because the vertices $v_{h+1},\dots,v_a$ are all isolated in $B\cup\{v_1,\dots,v_a\}$ (so they are not part of any edges $e\subseteq B\cup\{v_1,\dots,v_a\}$). On the other hand, we have $e(B\cup\{v_1,\dots,v_{h-1}\}) < e(B\cup\{v_1,\dots,v_h\}) = \ell$, because v_h is connected to B (so there exists at least one edge $e\subseteq B\cup\{v_h\}$ containing v_h).

Hence we see that $e(B\cup\{v_1,\dots,v_{i-1}\}) < \ell$ for all i with $1\le i\le h$ and $e(B\cup\{v_1,\dots,v_{i-1}\}) = \ell$ for all i with $h+1\le i\le a$. Therefore, for all indices $i=1,\dots,a$, the following holds: If $e(B\cup\{v_1,\dots,v_{i-1}\}) < \ell$, then v_i is connected B, but if $e(B\cup\{v_1,\dots,v_{i-1}\}) = \ell$, then v_i is not connected to B (as v_i must be isolated in $B\cup\{v_1,\dots,v_a\}$).

Let us now choose a random tidy sequence $v_1,\dots,v_a$ according to the following procedure: If for some $1\le i\le a$ the vertices $v_1,\dots,v_{i-1}$ have already been chosen, choose v_i uniformly at random among all vertices in $V(G)\setminus(B\cup\{v_1,\dots,v_{i-1}\})$ such that $v_1,\dots,v_i$ can be extended to some tidy sequence.

For example, v_1 will be chosen uniformly at random among all vertices that occur as the first vertex in some tidy sequence. Then, v_2 will be chosen uniformly at random among all vertices that appear as the second vertex in a tidy sequence that starts with vertex v_1. Continuing this random process, in the end one obtains a tidy sequence $v_1,\dots,v_a$ (and in each step there is indeed at least one choice for the next vertex).

For every tidy sequence $v_1,\dots,v_a$, the following claim gives a lower bound on the probability that this particular sequence got chosen in the random procedure.

Claim 4.6. *Let $v_1,\dots,v_a$ be a tidy sequence and let the index h be as in Definition 4.4. Then the probability that the sequence $v_1,\dots,v_a$ was chosen during the random procedure is at least*

$$\frac{a^a}{h^h\cdot(a-h)^{a-h}}\cdot\frac{1}{n^a}.$$

Proof. Let X be the number of vertices in $V(G)\setminus B$ that are connected to B. Note that $1\le h\le a-1$ implies that v_1 is connected to B, but v_a is not connected to B (since it is isolated in $B\cup\{v_1,\dots,v_a\}$). In particular, we see that $0<X<n$.

For every $1\le i\le a$, let us analyze the number of choices we have when choosing v_i in the random procedure described above (after having chosen $v_1,\dots,v_{i-1}$).

If $1\le i\le h$, then by the above observations we have $e(B\cup\{v_1,\dots,v_{i-1}\})<\ell$ and, when choosing v_i, we must choose a vertex that is connected to B. Thus, there are at most X choices, and the probability that the desired vertex v_i is chosen is at least $1/X$.

On the other hand, if $h+1\le i\le a$, then by the above observations we have $e(B\cup\{v_1,\dots,v_{i-1}\})=\ell$ and, when choosing v_i, we must choose a vertex that is not connected to B. Thus, there are at most $n-X$ choices, and the probability that the desired vertex v_i is chosen is at least $1/(n-X)$.

So all in all, the probability that we make the desired choice in each step and obtain precisely the sequence $v_1,\dots,v_a$ is at least

$$\left(\frac{1}{X}\right)^h\cdot\left(\frac{1}{n-X}\right)^{a-h}.$$

Now, let $0<x<1$ be such that $X=xn$, then

$$\left(\frac{1}{X}\right)^h\cdot\left(\frac{1}{n-X}\right)^{a-h} = \left(\frac{1}{xn}\right)^h\cdot\left(\frac{1}{(1-x)n}\right)^{a-h} = \frac{1}{x^h(1-x)^{a-h}}\cdot\frac{1}{n^a}.$$

Note that by the inequality of arithmetic and geometric mean we have

$$x^h(1-x)^{a-h} = h^h \cdot (a-h)^{a-h} \cdot \left(\frac{x}{h}\right)^h \cdot \left(\frac{1-x}{a-h}\right)^{a-h}$$
$$\leq h^h \cdot (a-h)^{a-h} \cdot \left(\frac{h \cdot \frac{x}{h} + (a-h) \cdot \frac{1-x}{a-h}}{h+a-h}\right)^{h+a-h} = h^h \cdot (a-h)^{a-h} \cdot \left(\frac{1}{a}\right)^a.$$

Hence the probability to obtain precisely the sequence $v_1, \dots, v_a$ during the random procedure is at least

$$\left(\frac{1}{X}\right)^h \cdot \left(\frac{1}{n-X}\right)^{a-h} = \frac{1}{x^h(1-x)^{a-h}} \cdot \frac{1}{n^a} \geq \frac{a^a}{h^h \cdot (a-h)^{a-h}} \cdot \frac{1}{n^a},$$

which finishes the proof of the claim. □

Corollary 4.7. *Let $A \subseteq V(G)$ be such that (A,B) is an ε-pleasant pair. Then, when choosing $v_1, \dots, v_a$ according to the random procedure described above, we have $B \cup \{v_1, \dots, v_a\} = A$ with probability at least*

$$\frac{\varepsilon^{1/2} \cdot k^{1/4}}{2 \cdot r^{1/4}} \cdot \frac{a!}{n^a}.$$

Proof. Recall that by Claim 4.5 there are at least $h(A,B)! \cdot (a - h(A,B))!$ labelings of $A \setminus B$ as $v_1, \dots, v_a$ such that $v_1, \dots, v_a$ is a tidy sequence. Clearly, for each of these labelings we have $B \cup \{v_1, \dots, v_a\} = A$. Hence for each of these labelings the index h in Definition 4.4 equals $h(B \cup \{v_1, \dots, v_a\}, B) = h(A,B)$.

Thus, abbreviating $h(A,B)$ just by h, by Claim 4.6 each of these $h! \cdot (a-h)!$ tidy sequences $v_1, \dots, v_a$ is chosen during the random procedure with probability at least

$$\frac{a^a}{h^h \cdot (a-h)^{a-h}} \cdot \frac{1}{n^a}.$$

So all in all the probability of having $B \cup \{v_1, \dots, v_a\} = A$ is at least

$$h! \cdot (a-h)! \cdot \frac{a^a}{h^h \cdot (a-h)^{a-h}} \cdot \frac{1}{n^a} = \frac{h!}{h^h} \cdot \frac{(a-h)!}{(a-h)^{a-h}} \cdot \frac{a^a}{a!} \cdot \frac{a!}{n^a}.$$

Note that by conditions (iii) and (iv) in Definition 4.1 we have both $h \geq \frac{1}{4} r^{-1/2} \cdot \varepsilon \cdot \sqrt{k}$ and $a - h \geq \frac{1}{4} r^{-1/2} \cdot \varepsilon \cdot \sqrt{k}$. Furthermore, at least one of the two numbers h and $a-h$ is at least $\frac{1}{2}a$. Thus,

$$h \cdot (a-h) \geq \frac{\varepsilon \cdot \sqrt{k}}{4\sqrt{r}} \cdot \frac{1}{2} a = \frac{\varepsilon \cdot \sqrt{k}}{8\sqrt{r}} \cdot a.$$

Now, by Stirling's formula (see for example [9]) we have

$$\frac{h!}{h^h} \cdot \frac{(a-h)!}{(a-h)^{a-h}} \cdot \frac{a^a}{a!} \geq \left(\sqrt{2\pi} \cdot \sqrt{h} \cdot e^{-h}\right) \cdot \left(\sqrt{2\pi} \cdot \sqrt{a-h} \cdot e^{-a+h}\right) \cdot \left(e \cdot \sqrt{a} \cdot e^{-a}\right)^{-1}$$
$$= \frac{2\pi}{e} \cdot \sqrt{\frac{h \cdot (a-h)}{a}} \geq \frac{2\pi}{e} \cdot \sqrt{\frac{\varepsilon \cdot \sqrt{k}}{8\sqrt{r}}} = \frac{2\pi}{e \cdot \sqrt{2}} \cdot \frac{1}{2} \cdot \frac{\varepsilon^{1/2} \cdot k^{1/4}}{r^{1/4}} \geq \frac{6}{3 \cdot 2} \cdot \frac{\varepsilon^{1/2} \cdot k^{1/4}}{2 \cdot r^{1/4}} = \frac{\varepsilon^{1/2} \cdot k^{1/4}}{2 \cdot r^{1/4}},$$

Thus, the probability of having $B \cup \{v_1, \dots, v_a\} = A$ is at least

$$\frac{h!}{h^h} \cdot \frac{(a-h)!}{(a-h)^{a-h}} \cdot \frac{a^a}{a!} \cdot \frac{a!}{n^a} \geq \frac{\varepsilon^{1/2} \cdot k^{1/4}}{2 \cdot r^{1/4}} \cdot \frac{a!}{n^a},$$

as desired. □

Recall that we fixed the set $B \subseteq V(G)$ and we chose a tidy sequence $v_1, \dots, v_a$ according to the random procedure described above. For different sets $A \subseteq V(G)$ such that (A,B) is an ε-pleasant pair, the events $B \cup \{v_1, \dots, v_a\} = A$ are clearly disjoint. Hence, by Corollary 4.7 the number of different sets $A \subseteq V(G)$ such that (A,B) is an ε-pleasant pair can be at most

$$\left(\frac{\varepsilon^{1/2} \cdot k^{1/4}}{2 \cdot r^{1/4}} \cdot \frac{a!}{n^a} \right)^{-1} = 2 \cdot r^{1/4} \cdot \varepsilon^{-1/2} \cdot k^{-1/4} \cdot \frac{n^a}{a!},$$

which finishes the proof of Lemma 4.2.

5 Proof of Lemmas 3.1 and 3.2

This section is devoted to proving Lemmas 3.1 and 3.2. The general proof strategy is very similar to the proof of Lemma 3.3 in the previous section. However, we need slightly different versions of the notions introduced in Definitions 4.1 and 4.4, see Definitions 5.1 and 5.5 below.

As in Lemmas 3.1 and 3.2, let us fix positive integers r, k and ℓ and let G be a hypergraph on $n \geq k$ vertices all of whose edges have size at most r.

Recall that we defined the quantities $h(A,B)$, $m(A,B)$ and $f(A,B)$ at the end of Section 3.

Definition 5.1. For a real number $z > 0$, let a pair (A,B) of subsets $B \subseteq A \subseteq V(G)$ be called *z-nice* if all of the following conditions are satisfied:

(i) $|A| = k$ and $e(A) = \ell$.

(ii) $z \leq h(A,B) \leq \sqrt{k}/2$.

(iii) $|A \setminus B| \geq 2\sqrt{k} \cdot f(A,B)$.

(iv) $|A \setminus B| \geq \sqrt{k}$.

Similarly as in the proof of Lemma 3.3, for a given value of z and a fixed subset $B \subseteq V(G)$ we can bound the number of sets $A \subseteq V(G)$ such that (A,B) is z-nice:

Lemma 5.2. *Let $z > 0$ be a real number and let $B \subseteq V(G)$ be a subset of size $|B| \leq k$. Let $a = k - |B|$. Then there are at most*

$$\frac{4}{3} \cdot \frac{1}{\sqrt{z}} \cdot \frac{n^a}{a!}$$

different sets $A \subseteq V(G)$ such that (A,B) is a z-nice pair.

We will prove Lemma 5.2 in Subsection 5.3 at the end of this section. First, we will prove Lemmas 3.1 and 3.2 in the next two subsections assuming Lemma 5.2. Most of the arguments will be quite similar to those we used in the previous section to prove Lemma 3.3. However, some of the details are different, which makes it necessary to prove Lemmas 3.1 and 3.2 separately from Lemma 3.3.

5.1 Proof of Lemma 3.1

In this subsection we will prove Lemma 3.1 assuming Lemma 5.2. As in the statement of Lemma 3.1, let c be a real number with $0 < c < \sqrt{k}/2$. Note that if $c \le 32^2 \cdot r$, then

$$32 \cdot \sqrt{r} \cdot \frac{1}{\sqrt{c}} \cdot \frac{n^k}{k!} \ge \frac{n^k}{k!} \ge \binom{n}{k}.$$

So in this case Lemma 3.1 is trivially true as the total number of k-vertex subsets of $V(G)$ is only $\binom{n}{k}$. Thus, we can assume that $c \ge 32^2 \cdot r$.

We start by proving that for each fixed k-vertex subset $A \subseteq V(G)$ with the properties in Lemma 3.1, we have a probability of at least $\frac{1}{4}$ to obtain a z-nice pair (A,B) when choosing a random subset $B \subseteq A$ by picking each element of A independently with probability p (for appropriate choices of the parameters z and p).

Lemma 5.3. *Let $z = \frac{1}{32r}c$ and $p = 1 - \frac{1}{8r}$. Then for every k-vertex subset $A \subseteq V(G)$ with $e(A) = \ell$ and $c \le m(A) \le \sqrt{k}/2$, the following holds: If we choose a subset $B \subseteq A$ randomly by taking each element of A into B with probability p (independently for all elements of A), then with probability at least $\frac{1}{4}$ the resulting pair (A,B) is z-nice.*

Proof. Fix a subset $A \subseteq V(G)$ with $|A| = k$ as well as $e(A) = \ell$ and $c \le m(A) \le \sqrt{k}/2$. As in the lemma, let the subset $B \subseteq A$ be chosen by randomly picking elements of A independently with probability p. Let $q = 1 - p = \frac{1}{8r}$, then each element of A will be in $A \setminus B$ with probability q.

We claim that with probability at least $\frac{1}{2}$ we have

$$f(A,B) \le 2r \cdot q^2 \cdot m(A) = \frac{1}{32r} \cdot m(A). \tag{5.1}$$

Recall that $f(A,B)$ counts the number of vertices in $A \setminus B$ that are non-isolated in A but not connected to B. In total there are $m(A)$ non-isolated vertices v in A. For each such v in order to count towards $f(A,B)$, it must satisfy $v \in A \setminus B$ and for each edge $e \subseteq A$ with $v \in e$ there must be at least one vertex $v' \in e \setminus \{v\}$ with $v' \in A \setminus B$. The probability for $v \in A \setminus B$ is exactly q. Let us just pick one edge $e \subseteq A$ with $v \in e$ (such an edge exists as v is non-isolated in A). The probability for there to exist $v' \in e \setminus \{v\}$ with $v' \in A \setminus B$ is at most $(r-1) \cdot q$ (because there are at most $|e \setminus \{v\}| \le r-1$ possibilities for v' and each of them is in $A \setminus B$ with probability q), and this is independent of whether $v \in A \setminus B$. Thus, the total probability for v to count towards $f(A,B)$ is at most $(r-1) \cdot q^2$. Summing this over all $m(A)$ choices for v, we obtain

$$\mathbb{E}[f(A,B)] \le (r-1) \cdot q^2 \cdot m(A) \le r \cdot q^2 \cdot m(A).$$

Clearly, $f(A,B)$ is a non-negative random variable, so by Markov's inequality we have

$$\mathbb{P}[f(A,B) \ge 2r \cdot q^2 \cdot m(A)] \le \frac{\mathbb{E}[f(A,B)]}{2r \cdot q^2 \cdot m(A)} \le \frac{r \cdot q^2 \cdot m(A)}{2r \cdot q^2 \cdot m(A)} = \frac{1}{2}.$$

This shows that (5.1) indeed holds with probability at least $\frac{1}{2}$.

Now, we claim that with probability at least $\frac{7}{8}$ we have

$$m(A,B) \geq \frac{1}{2} \cdot q \cdot m(A) = \frac{1}{16r} \cdot m(A). \tag{5.2}$$

Recall that $m(A,B)$ counts the number of vertices in $A \setminus B$ that are non-isolated in A. As there are $m(A)$ vertices of A that are non-isolated in A, we can see that $m(A,B)$ is binomially distributed with parameters $m(A)$ and q. By [2, Theorem A.1.13], we therefore have

$$\mathbb{P}\left[m(A,B) < \frac{1}{2} \cdot q \cdot m(A)\right] < \exp\left(-\frac{\left(\frac{1}{2} \cdot q \cdot m(A)\right)^2}{2 \cdot q \cdot m(A)}\right) = \exp\left(-\frac{q}{8} \cdot m(A)\right)$$
$$= \exp\left(-\frac{1}{64r} \cdot m(A)\right) \leq e^{-3} \leq \frac{1}{8}$$

where we used $m(A) \geq c \geq 32^2 \cdot r \geq 3 \cdot 64r$. So (5.2) indeed holds with probability at least $\frac{7}{8}$.

Furthermore, by the assumptions on A we always have

$$h(A,B) \leq m(A,B) \leq m(A) \leq \sqrt{k}/2. \tag{5.3}$$

Finally, we claim that with probability at least $\frac{7}{8}$ we have

$$|A \setminus B| \geq \frac{1}{2} \cdot q \cdot k = \frac{1}{16r} \cdot k. \tag{5.4}$$

Indeed, $|A \setminus B|$ is binomially distributed with parameters $|A| = k$ and q. So again using [2, Theorem A.1.13], we have

$$\mathbb{P}\left[m(A,B) < \frac{1}{2} \cdot q \cdot k\right] < \exp\left(-\frac{\left(\frac{1}{2} \cdot q \cdot k\right)^2}{2 \cdot q \cdot k}\right) = \exp\left(-\frac{q}{8} \cdot k\right) = \exp\left(-\frac{1}{64r} \cdot k\right) \leq e^{-3} \leq \frac{1}{8}$$

where we used $k \geq \sqrt{k}/2 \geq c \geq 32^2 \cdot r \geq 3 \cdot 64r$. So (5.4) indeed holds with probability at least $\frac{7}{8}$.

All in all, (5.1) fails with probability at most $\frac{1}{2}$, each of (5.2) and (5.4) fails with probability at most $\frac{1}{8}$, and (5.3) always holds. Thus, with probability at least $\frac{1}{4}$ all the inequalities (5.1) to (5.4) hold. We claim that in this case the pair (A,B) is z-nice.

First, note that condition (i) in Definition 5.1 is automatically satisfied by the assumptions on A. The upper bound $h(A,B) \leq \sqrt{k}/2$ in condition (ii) follows from (5.3). Furthermore, note that (5.1) and (5.2) imply

$$h(A,B) = m(A,B) - f(A,B) \geq \frac{1}{16r} \cdot m(A) - \frac{1}{32r} \cdot m(A) = \frac{1}{32r} \cdot m(A) \geq \frac{1}{32r} \cdot c = z,$$

which gives the lower bound in condition (ii). For condition (iii), note that the inequalities (5.4) and (5.1) imply (using $m(A) \leq \sqrt{k}/2 \leq \sqrt{k}$)

$$|A \setminus B| \geq \frac{1}{16r} \cdot k = 2\sqrt{k} \cdot \frac{1}{32r} \cdot \sqrt{k} \geq 2\sqrt{k} \cdot \frac{1}{32r} \cdot m(A) \geq 2\sqrt{k} \cdot f(A,B).$$

For condition (iv), note that $16r \le 32^2 \cdot r \le c \le \sqrt{k}/2 \le \sqrt{k}$, so (5.4) gives

$$|A \setminus B| \ge \frac{1}{16r} \cdot k \ge \frac{1}{\sqrt{k}} \cdot k = \sqrt{k}.$$

Thus, the pair (A,B) is indeed z-nice if all the inequalities (5.1) to (5.4) hold. Hence we have shown that (A,B) is a z-nice pair with probability at least $\frac{1}{4}$. □

We will now prove Lemma 3.1. Let us suppose for contradiction that there are more than

$$32 \cdot \sqrt{r} \cdot \frac{1}{\sqrt{c}} \cdot \frac{n^k}{k!}.$$

k-vertex subsets $A \subseteq V(G)$ with $e(A) = \ell$ and $c \le m(A) \le \sqrt{k}/2$.

Define $z = \frac{1}{32r}c$ and $p = 1 - \frac{1}{8r}$ as in Lemma 5.3. Now, let us choose a pair (A,B) of subsets $B \subseteq A \subseteq V(G)$ with $|A| = k$ randomly as follows: First, choose $A \subseteq V(G)$ uniformly at random among all subsets of size $|A| = k$. Then, choose a subset $B \subseteq A$ by taking each element of A into B with probability p (independently for all elements of A).

The probability that A satisfies $e(A) = \ell$ and $c \le m(A) \le \sqrt{k}/2$ is at least

$$32 \cdot \sqrt{r} \cdot \frac{1}{\sqrt{c}} \cdot \frac{n^k}{k!} \cdot \binom{n}{k}^{-1} = 32 \cdot \sqrt{r} \cdot \frac{1}{\sqrt{c}} \cdot \frac{n^k}{(n)_k}.$$

If this is the case, then by Lemma 5.3 the pair (A,B) is z-nice with probability at least $\frac{1}{4}$. So we obtain

$$\mathbb{P}[(A,B) \text{ is } z\text{-nice}] \ge 8 \cdot \sqrt{r} \cdot \frac{1}{\sqrt{c}} \cdot \frac{n^k}{(n)_k}. \tag{5.5}$$

However, we can find an upper bound on the probability that (A,B) is z-nice by conditioning on B. Note that B will always be a subset $B \subseteq V(G)$ with $|B| \le k$. So let us fix any subset $B \subseteq V(G)$ with $|B| \le k$ and let $a = k - |B|$. Conditioning on having this particular set B as the result of our random draw of the pair (A,B), there are

$$\binom{n-|B|}{k-|B|} = \binom{n-|B|}{a} = \frac{(n-|B|)_a}{a!}$$

possibilities for the set A and they are all equally likely. But by Lemma 5.2 at most

$$\frac{4}{3} \cdot \frac{1}{\sqrt{z}} \cdot \frac{n^a}{a!} = \frac{4}{3} \cdot \frac{1}{\sqrt{\frac{1}{32r}c}} \cdot \frac{n^a}{a!} < \frac{4}{3} \cdot 6 \cdot \sqrt{r} \cdot \frac{1}{\sqrt{c}} \cdot \frac{n^a}{a!} = 8 \cdot \sqrt{r} \cdot \frac{1}{\sqrt{c}} \cdot \frac{n^a}{a!}$$

of these possible sets A have the property that (A,B) is z-nice. Thus, when we condition on having this set B, the probability for (A,B) to be z-nice is less than

$$8 \cdot \sqrt{r} \cdot \frac{1}{\sqrt{c}} \cdot \frac{n^a}{(n-|B|)_a} = 8 \cdot \sqrt{r} \cdot \frac{1}{\sqrt{c}} \cdot \frac{n^{k-|B|}}{(n-|B|)_{k-|B|}} \le 8 \cdot \sqrt{r} \cdot \frac{1}{\sqrt{c}} \cdot \frac{n^k}{(n)_k}.$$

Since the set B was arbitrary, this shows

$$\mathbb{P}[(A,B) \text{ is } z\text{-nice}] < 8 \cdot \sqrt{r} \cdot \frac{1}{\sqrt{c}} \cdot \frac{n^k}{(n)_k},$$

a contradiction to (5.5). This finishes the proof of Lemma 3.1.

5.2 Proof of Lemma 3.2

In this subsection we will prove Lemma 3.2 assuming Lemma 5.2. The proof is very similar to the proof of Lemma 3.1 in the previous subsection. This time, let c be a real number with $\sqrt{k}/2 \le c \le k/(32r)$. Note that if $k^{1/4} \le 44 \cdot \sqrt{r}$, then

$$44 \cdot \sqrt{r} \cdot k^{-1/4} \cdot \frac{n^k}{k!} \ge \frac{n^k}{k!} \ge \binom{n}{k}.$$

and Lemma 3.2 is trivially true. Thus, we can assume that $k^{1/4} \ge 44 \cdot \sqrt{r}$, which means $k \ge 44^4 \cdot r^2$.

Lemma 5.4. *Let $z = \frac{1}{64r}\sqrt{k}$ and $p = 1 - \frac{1}{16r}\sqrt{k} \cdot c^{-1}$. Then for every k-vertex subset $A \subseteq V(G)$ with $e(A) = \ell$ and $c \le m(A) \le 2c$, the following holds: If we choose a subset $B \subseteq A$ randomly by taking each element of A into B with probability p (independently for all elements of A), then with probability at least $\frac{1}{4}$ the resulting pair (A,B) is z-nice.*

Proof. Fix a subset $A \subseteq V(G)$ with $|A| = k$ as well as $e(A) = \ell$ and $c \le m(A) \le 2c$. As in the lemma, let the subset $B \subseteq A$ be chosen by randomly picking elements of A independently with probability p. Let $q = 1 - p = \frac{1}{16r}\sqrt{k} \cdot c^{-1}$, so each element of A will be in $A \setminus B$ with probability q.

We claim that with probability at least $\frac{1}{2}$ we have

$$f(A,B) \le 2r \cdot q^2 \cdot m(A). \tag{5.6}$$

With the same argument as in the proof of Lemma 5.3, we can see that $\mathbb{E}[f(A,B)] \le r \cdot q^2 \cdot m(A)$ and again by Markov's inequality we have

$$\mathbb{P}[f(A,B) \ge 2r \cdot q^2 \cdot m(A)] \le \frac{\mathbb{E}[f(A,B)]}{2r \cdot q^2 \cdot m(A)} \le \frac{r \cdot q^2 \cdot m(A)}{2r \cdot q^2 \cdot m(A)} = \frac{1}{2}.$$

This shows that (5.6) indeed holds with probability at least $\frac{1}{2}$.

Next, we claim that with probability at least $\frac{15}{16}$ we have

$$m(A,B) \ge \frac{1}{2} \cdot q \cdot m(A). \tag{5.7}$$

As in the proof of Lemma 5.3 we see that $m(A,B)$ is binomially distributed with parameters $m(A)$ and q. So again using [2, Theorem A.1.13], we have

$$\mathbb{P}\left[m(A,B) < \frac{1}{2} \cdot q \cdot m(A)\right] < \exp\left(-\frac{\left(\frac{1}{2} \cdot q \cdot m(A)\right)^2}{2 \cdot q \cdot m(A)}\right) = \exp\left(-\frac{1}{8} \cdot q \cdot m(A)\right)$$
$$\le \exp\left(-\frac{1}{8} \cdot q \cdot c\right) = \exp\left(-\frac{1}{8} \cdot \frac{1}{16r} \cdot \sqrt{k}\right) = \exp\left(-\frac{1}{128r} \cdot \sqrt{k}\right) \le e^{-4} \le \frac{1}{16}$$

where this time we used $m(A) \ge c$ and $\sqrt{k} \ge 44^2 \cdot r \ge 4 \cdot 128r$. So (5.7) indeed holds with probability at least $\frac{15}{16}$.

Furthermore, we claim that with probability at least $\frac{15}{16}$ we have

$$m(A,B) \le \sqrt{k}/2. \tag{5.8}$$

Note that

$$q \cdot m(A) \le q \cdot 2c = 2 \cdot \frac{1}{16r}\sqrt{k} = \frac{1}{8r}\sqrt{k} \le \frac{\sqrt{k}}{8}.$$

Therefore, again using that $m(A,B)$ is binomially distributed with parameters $m(A)$ and q, the standard Chernoff bound for the upper tail of binomially distributed variables (see, for example, [2, Theorem A.1.4]) gives

$$\mathbb{P}\left[m(A,B) > \frac{\sqrt{k}}{2}\right] < \exp\left(-2 \cdot \left(\frac{\sqrt{k}}{2} - q \cdot m(A)\right)^2 \cdot \frac{1}{m(A)}\right) \le \exp\left(-2 \cdot \left(\frac{3}{8}\sqrt{k}\right)^2 \cdot \frac{1}{m(A)}\right)$$
$$= \exp\left(-\frac{9}{32} \cdot k \cdot \frac{1}{m(A)}\right) \le \exp\left(-\frac{9}{32} \cdot k \cdot \frac{1}{k/16}\right) = \exp\left(-\frac{9}{2}\right) \le e^{-4} \le \frac{1}{16}.$$

Here we used that $m(A) \le 2c \le k/(16r) \le k/16$. So (5.8) holds with probability at least $\frac{15}{16}$.

Finally, we claim that with probability at least $\frac{15}{16}$ we have

$$|A \setminus B| \ge \frac{1}{2} \cdot q \cdot k. \tag{5.9}$$

Using that $|A \setminus B|$ is binomially distributed with parameters $|A| = k$ and q and applying [2, Theorem A.1.13], we obtain

$$\mathbb{P}\left[m(A,B) < \frac{1}{2} \cdot q \cdot k\right] < \exp\left(-\frac{\left(\frac{1}{2} \cdot q \cdot k\right)^2}{2 \cdot q \cdot k}\right) = \exp\left(-\frac{q}{8} \cdot k\right) = \exp\left(-\frac{1}{128r} \cdot \frac{\sqrt{k}}{c} \cdot k\right)$$
$$\le \exp\left(-\frac{1}{128r} \cdot \frac{\sqrt{k}}{k/(32r)} \cdot k\right) = \exp\left(-\frac{\sqrt{k}}{4}\right) \le e^{-4} \le \frac{1}{16},$$

where we used $c \le k/(32r)$ and $k \ge 44^4 \cdot r^2 \ge 16^2$. So (5.9) holds with probability at least $\frac{15}{16}$.

All in all, (5.6) fails with probability at most $\frac{1}{2}$, and each of (5.7), (5.8) and (5.9) fails with probability at most $\frac{1}{16}$. Thus, with probability at least $\frac{1}{4}$ all the inequalities (5.6) to (5.9) hold. We claim that in this case the pair (A,B) is z-nice.

First, note that condition (i) in Definition 5.1 is automatically satisfied by the assumptions on A. For the upper bound in condition (ii), note that (5.8) implies $h(A,B) \le m(A,B) \le \sqrt{k}/2$. For the lower bound in condition (ii), first note that by the assumption $c \ge \sqrt{k}/2$ we have

$$q = \frac{1}{16r} \cdot \frac{\sqrt{k}}{c} \le \frac{1}{16r} \cdot \frac{\sqrt{k}}{\sqrt{k}/2} = \frac{1}{8r}.$$

Hence (5.6) and (5.7) imply

$$h(A,B) = m(A,B) - f(A,B) \geq \frac{1}{2} \cdot q \cdot m(A) - 2r \cdot q^2 \cdot m(A) = q \cdot m(A) \cdot \left(\frac{1}{2} - 2r \cdot q\right)$$
$$\geq q \cdot m(A) \cdot \left(\frac{1}{2} - 2r \cdot \frac{1}{8r}\right) = \frac{1}{4} \cdot q \cdot m(A) \geq \frac{1}{4} \cdot q \cdot c = \frac{1}{4} \cdot \frac{1}{16r} \cdot \sqrt{k} = \frac{1}{64r} \cdot \sqrt{k} = z,$$

which gives the lower bound in condition (ii). For condition (iii), note that the inequalities (5.6) and (5.9) imply (using $m(A) \leq 2c$)

$$2\sqrt{k} \cdot f(A,B) \leq 2\sqrt{k} \cdot 2r \cdot q^2 \cdot m(A) \leq 4\sqrt{k} \cdot r \cdot q^2 \cdot 2c = 8\sqrt{k} \cdot r \cdot q \cdot \frac{1}{16r}\sqrt{k} = \frac{1}{2} \cdot q \cdot k \leq |A \setminus B|.$$

For condition (iv), note that using $c \leq k/(32r)$, inequality (5.9) gives

$$|A \setminus B| \geq \frac{1}{2} \cdot q \cdot k = \frac{1}{2} \cdot \frac{1}{16r} \cdot \frac{\sqrt{k}}{c} \cdot k \geq \frac{1}{32r} \cdot \frac{\sqrt{k}}{k/(32r)} \cdot k = \sqrt{k}.$$

Thus, the pair (A,B) is indeed z-nice if the inequalities (5.6) to (5.9) hold. Hence we have shown that (A,B) is a z-nice pair with probability at least $\frac{1}{4}$. □

Let us now prove Lemma 3.2. So suppose for contradiction that there are more than

$$44 \cdot \sqrt{r} \cdot k^{-1/4} \cdot \frac{n^k}{k!}$$

k-vertex subsets $A \subseteq V(G)$ with $e(A) = \ell$ and $c \leq m(A) \leq 2c$.

Define $z = \frac{1}{64r}\sqrt{k}$ and $p = 1 - \frac{1}{16r}\sqrt{k} \cdot c^{-1}$ as in Lemma 5.4. Again, let us choose a pair (A,B) of subsets $B \subseteq A \subseteq V(G)$ with $|A| = k$ randomly as follows: First, choose $A \subseteq V(G)$ uniformly at random among all subsets of size $|A| = k$. Then, choose a subset $B \subseteq A$ by taking each element of A into B with probability p (independently for all elements of A).

The probability that A satisfies $e(A) = \ell$ and $c \leq m(A) \leq 2c$ is at least

$$44 \cdot \sqrt{r} \cdot k^{-1/4} \cdot \frac{n^k}{k!} \cdot \binom{n}{k}^{-1} = 44 \cdot \sqrt{r} \cdot k^{-1/4} \cdot \frac{n^k}{(n)_k},$$

and in this case, by Lemma 5.3 the pair (A,B) is z-nice with probability at least $\frac{1}{4}$. Thus,

$$\mathbb{P}[(A,B) \text{ is } z\text{-nice}] \geq 11 \cdot \sqrt{r} \cdot k^{-1/4} \cdot \frac{n^k}{(n)_k}. \tag{5.10}$$

On the other hand, we can again find an upper bound on the probability that (A,B) is z-nice by conditioning on B. Let us fix any subset $B \subseteq V(G)$ with $|B| \leq k$ and let $a = k - |B|$. Conditioning on having this particular set B as the result of our random draw of the pair (A,B), there are

$$\binom{n-|B|}{k-|B|} = \binom{n-|B|}{a} = \frac{(n-|B|)_a}{a!}$$

possibilities for the set A and they are all equally likely. By Lemma 5.2 at most

$$\frac{4}{3}\cdot\frac{1}{\sqrt{z}}\cdot\frac{n^a}{a!}=\frac{4}{3}\cdot\frac{1}{\sqrt{\frac{1}{64r}\sqrt{k}}}\cdot\frac{n^a}{a!}=\frac{4}{3}\cdot 8\cdot\sqrt{r}\cdot k^{-1/4}\cdot\frac{n^a}{a!}<11\cdot\sqrt{r}\cdot k^{-1/4}\cdot\frac{n^a}{a!}$$

of these possible sets A have the property that (A,B) is z-nice. Thus, conditioning on having this set B, the probability for (A,B) to be z-nice is less than

$$11\cdot\sqrt{r}\cdot k^{-1/4}\cdot\frac{n^a}{(n-|B|)_a}=11\cdot\sqrt{r}\cdot k^{-1/4}\cdot\frac{n^{k-|B|}}{(n-|B|)_{k-|B|}}\le 11\cdot\sqrt{r}\cdot k^{-1/4}\cdot\frac{n^k}{(n)_k}.$$

Since the set B was arbitrary, we obtain

$$\mathbb{P}[(A,B)\text{ is }z\text{-nice}]<11\cdot\sqrt{r}\cdot k^{-1/4}\cdot\frac{n^k}{(n)_k},$$

which contradicts (5.10). This finishes the proof of Lemma 3.2.

5.3 Proof of Lemma 5.2

In this subsection, we will prove Lemma 5.2. The proof is overall quite similar to the proof of Lemma 4.2, but some of the details differ significantly.

As in the statement of the lemma, fix $z>0$ as well as a subset $B\subseteq V(G)$ of size $|B|\le k$, and let $a=k-|B|$. Furthermore, let $s=\lfloor\sqrt{k}/2\rfloor$.

Note that for every z-nice pair (A,B) we have $|A|=k$ by condition (i) in Definition 5.1 and therefore $|A\setminus B|=k-|B|=a$. Thus, by condition (iv) in Definition 5.1 we must have $a\ge\sqrt{k}$ if there exists at least one z-nice pair (A,B). This means that we may assume that $a\ge\sqrt{k}$, because otherwise Lemma 5.2 is trivially true. This in particular implies $a\ge 2\cdot\lfloor\sqrt{k}/2\rfloor=2s$.

Furthermore, if there exists some z-nice pair (A,B), we must have $0<z\le h(A,B)\le\sqrt{k}/2$ by condition (ii) in Definition 5.1. As $h(A,B)$ is an integer, this implies $h(A,B)\ge 1$ and hence $\sqrt{k}/2\ge 1$. Thus, we must have $s=\lfloor\sqrt{k}/2\rfloor\ge 1$ if there exists at least one z-nice pair (A,B). So we may assume $s\ge 1$ from now on.

Definition 5.5. We call a sequence $v_1,\dots,v_a$ of distinct vertices in $V(G)\setminus B$ *tame* if $e(B\cup\{v_1,\dots,v_a\})=\ell$ and if there exists a positive integer h with $z\le h\le s$ such that

- none of the vertices $v_1,\dots,v_{a-s}$ is connected to B,
- each of the vertices $v_{a-s+1},\dots,v_{a-s+h}$ is connected to B, and
- each of the vertices $v_{a-s+h+1},\dots,v_a$ is isolated in $B\cup\{v_1,\dots,v_a\}$.

Note that for a sequence $v_1,\dots,v_a$ as in Definition 5.5, the vertices $v_{a-s+h+1},\dots,v_a$ cannot be connected to B (since otherwise they would be non-isolated in $B\cup\{v_1,\dots,v_a\}$). Thus, the number $h(B\cup\{v_1,\dots,v_a\},B)$ of vertices in $\{v_1,\dots,v_a\}$ that are connected to B is precisely h. Hence for every tame sequence $v_1,\dots,v_a$, the integer h in Definition 5.5 must equal $h(B\cup\{v_1,\dots,v_a\},B)$.

Similarly as in the proof of Lemma 4.2, we can show that for every set $A\subseteq V(G)$ such that (A,B) is z-nice, there are many ways to label the vertices of $A\setminus B$ as $v_1,\dots,v_a$ so that $v_1,\dots,v_a$ is a tame sequence:

Claim 5.6. *Let $A \subseteq V(G)$ be such that (A,B) is a z-nice pair. Then there are at least*

$$\frac{3}{4} \cdot h(A,B)! \cdot (a - h(A,B))!$$

different ways to label the vertices of $A \setminus B$ as $v_1, \dots, v_a$ such that $v_1, \dots, v_a$ is a tame sequence.

Proof. First, note that by condition (i) in Definition 5.1 we have $|A \setminus B| = k - |B| = a$.

Let $h = h(A,B)$ and note that by condition (ii) in Definition 5.1 we have $z \le h \le \sqrt{k}/2$, and therefore $h \le \lfloor \sqrt{k}/2 \rfloor = s$ as $h = h(A,B)$ is an integer. Thus, $z \le h \le s$. From $z > 0$, we also obtain $h \ge 1$.

There are exactly $h = h(A,B)$ vertices in $A \setminus B$ that are connected to B. Label these h vertices as $v_{a-s+1}, \dots, v_{a-s+h}$; there are $h!$ ways how to do that. Now, pick a labeling of the remaining $a-h$ vertices in $A \setminus B$ as $v_1, \dots, v_{a-s}$ and $v_{a-s+h+1}, \dots, v_a$ uniformly at random from the $(a-h)!$ possible labelings.

Then $v_1, \dots, v_a$ are always distinct vertices in $V(G) \setminus B$ and furthermore $e(B \cup \{v_1, \dots, v_a\}) = e(A) = \ell$ (by condition (i) in Definition 5.1). We already saw that $z \le h \le s$. Note that by the choice of the labeling, the first and second condition in Definition 5.5 are also automatically satisfied. So we just need to make sure that with a sufficiently high probability the third condition is satisfied as well.

In order for the third condition in Definition 5.5 to fail, there needs to be a vertex $v \in A \setminus B$ that is non-isolated in $B \cup \{v_1, \dots, v_a\} = A$ and obtains a label v_i with $a - s + h + 1 \le i \le a$. In particular this vertex is not one of the $h = h(A,B)$ vertices in $A \setminus B$ that are connected to B, because otherwise v would have already obtained one of the labels $v_{a-s+1}, \dots, v_{a-s+h}$ in the first step. This means that v is one of the $f(A,B) = m(A,B) - h(A,B)$ vertices in $A \setminus B$ that are non-isolated in A but not connected to B. So there are at most $f(A,B)$ possibilities for the vertex v and for each of them the probability that v is labeled v_i for some $a - s + h + 1 \le i \le a$ is at most

$$\frac{s-h}{a-h} \le \frac{s}{a}.$$

So by a simple union bound, the third condition in Definition 5.5 holds with probability at least

$$1 - f(A,B) \cdot \frac{s}{a} = 1 - \frac{s \cdot f(A,B)}{a} \ge 1 - \frac{1}{4} = \frac{3}{4},$$

where we used that

$$a = |A \setminus B| \ge 2\sqrt{k} \cdot f(A,B) \ge 4 \cdot \lfloor \sqrt{k}/2 \rfloor \cdot f(A,B) = 4 \cdot s \cdot f(A,B)$$

by condition (iii) in Definition 5.1.

Hence for at least $\frac{3}{4}(a-h)!$ labelings of the $a-h$ vertices in the second step as $v_1, \dots, v_{a-s}$ and $v_{a-s+h+1}, \dots, v_a$, the resulting sequence $v_1, \dots, v_a$ is tame. Recall that we also had $h!$ choices in the first step (to distribute the labels $v_{a-s+1}, \dots, v_{a-s+h}$). Thus, the total number of ways to label the vertices of $A \setminus B$ as $v_1, \dots, v_a$ such that $v_1, \dots, v_a$ is a tame sequence is at least

$$h! \cdot \frac{3}{4}(a-h)! = \frac{3}{4} \cdot h! \cdot (a-h)! = \frac{3}{4} \cdot h(A,B)! \cdot (a - h(A,B))!,$$

as desired. □

In particular, for each subset $A \subseteq V(G)$ such that (A,B) is a z-nice pair, there is a labeling of $A \setminus B$ as $v_1,\dots,v_a$ such that $v_1,\dots,v_a$ is a tame sequence. Thus, we may assume that there exists at least one tame sequence $v_1,\dots,v_a$, as otherwise Lemma 5.2 is trivially true.

Note that for every tame sequence $v_1,\dots,v_a$ (with the integer h as in Definition 5.5), we must have $e(B\cup\{v_1,\dots,v_{a-s+h}\}) = e(B\cup\{v_1,\dots,v_a\}) = \ell$, because the vertices $v_{a-s+h+1},\dots,v_a$ are all isolated in $B\cup\{v_1,\dots,v_a\}$. On the other hand, we have $e(B\cup\{v_1,\dots,v_{a-s+h-1}\}) < e(B\cup\{v_1,\dots,v_{a-s+h}\}) = \ell$, because v_{a-s+h} is connected to B.

Therefore we obtain that $e(B\cup\{v_1,\dots,v_{i-1}\}) < \ell$ for all i with $a-s+1 \le i \le a-s+h$ and that $e(B\cup\{v_1,\dots,v_{i-1}\}) = \ell$ for all i with $a-s+h+1 \le i \le a$. Therefore, for all i with $a-s+1 \le i \le a$, the following holds: If $e(B\cup\{v_1,\dots,v_{i-1}\}) < \ell$, then v_i is connected B, but if $e(B\cup\{v_1,\dots,v_{i-1}\}) = \ell$, then v_i is not connected to B (as v_i must be isolated in $B\cup\{v_1,\dots,v_a\}$). Furthermore, for all i with $1 \le i \le a-s$, the vertex v_i is also not connected to B.

As in the proof of Lemma 4.2, let us choose a random tame sequence $v_1,\dots,v_a$ according to the following procedure: If for some $1 \le i \le a$ the vertices $v_1,\dots,v_{i-1}$ have already been chosen, choose v_i uniformly at random among all vertices in $V(G)\setminus(B\cup\{v_1,\dots,v_{i-1}\})$ such that $v_1,\dots,v_i$ can be extended to some tame sequence.

Claim 5.7. *Let $v_1,\dots,v_a$ be a tame sequence and let the integer h be as in Definition 5.5. Then the probability that the sequence $v_1,\dots,v_a$ was chosen during the random procedure is at least*

$$\frac{a^a}{h^h\cdot(a-h)^{a-h}}\cdot\frac{1}{n^a}.$$

Proof. Let X be the number of vertices in $V(G)\setminus B$ that are connected to B. Note that $a-s \ge 2s-s \ge 1$ implies that v_1 is not connected to B, whereas v_{a-s+1} is connected to B. In particular, we see that $0 < X < n$.

For every $1 \le i \le a$, let us analyze the number of choices we have when choosing v_i in the random procedure described above (after having chosen $v_1,\dots,v_{i-1}$).

When choosing v_i for $1 \le i \le a-s$, we must choose a vertex that is not connected to B. Thus, there are at most $n-X$ choices, and the probability that the desired vertex v_i is chosen is at least $1/(n-X)$.

If $a-s+1 \le i \le a-s+h$, then by the above observations we have $e(B\cup\{v_1,\dots,v_{i-1}\}) < \ell$ and, when choosing v_i, we must choose a vertex that is connected to B. Thus, there are at most X choices, and the probability that the desired vertex v_i is chosen is at least $1/X$.

If $a-s+h+1 \le i \le a$, then by the above observations we have $e(B\cup\{v_1,\dots,v_{i-1}\}) = \ell$ and, when choosing v_i, we must choose a vertex that is not connected to B. Thus, there are at most $n-X$ choices, and the probability that the desired vertex v_i is chosen is at least $1/(n-X)$.

So all in all, the probability that we make the desired choice in each step and obtain precisely the sequence $v_1,\dots,v_a$ is at least

$$\left(\frac{1}{n-X}\right)^{a-s}\cdot\left(\frac{1}{X}\right)^{h}\cdot\left(\frac{1}{n-X}\right)^{s-h} = \left(\frac{1}{X}\right)^{h}\cdot\left(\frac{1}{n-X}\right)^{a-h}.$$

In exactly the same way as in the proof of Claim 4.6, we can show that

$$\left(\frac{1}{X}\right)^{h}\cdot\left(\frac{1}{n-X}\right)^{a-h} \ge \frac{a^a}{h^h\cdot(a-h)^{a-h}}\cdot\frac{1}{n^a}.$$

This finishes the proof of Claim 5.7. □

Corollary 5.8. *Let $A \subseteq V(G)$ be such that (A,B) is a z-nice pair. Then, when choosing $v_1,\dots,v_a$ according to the random procedure described above, we have $B \cup \{v_1,\dots,v_a\} = A$ with probability at least*

$$\frac{3}{4}\cdot\sqrt{z}\cdot\frac{a!}{n^a}.$$

Proof. Recall that by Claim 5.6 there are at least $\frac{3}{4}\cdot h(A,B)!\cdot(a-h(A,B))!$ labelings of $A\setminus B$ as $v_1,\dots,v_a$ such that $v_1,\dots,v_a$ is a tame sequence. For each of these labelings we clearly have $B\cup\{v_1,\dots,v_a\}=A$ and the integer h in Definition 5.5 equals $h(B\cup\{v_1,\dots,v_a\},B)=h(A,B)$.

Abbreviating $h(A,B)$ just by h, by Claim 5.7 each of these $\frac{3}{4}\cdot h!\cdot(a-h)!$ tame sequences $v_1,\dots,v_a$ is chosen during the random procedure with probability at least

$$\frac{a^a}{h^h\cdot(a-h)^{a-h}}\cdot\frac{1}{n^a}.$$

Therefore the probability of having $B\cup\{v_1,\dots,v_a\}=A$ is at least

$$\frac{3}{4}\cdot h!\cdot(a-h)!\cdot\frac{a^a}{h^h\cdot(a-h)^{a-h}}\cdot\frac{1}{n^a}=\frac{3}{4}\cdot\frac{h!}{h^h}\cdot\frac{(a-h)!}{(a-h)^{a-h}}\cdot\frac{a^a}{a!}\cdot\frac{a!}{n^a}.$$

By Stirling's formula (see for example [9]) we have

$$\begin{aligned}\frac{h!}{h^h}\cdot\frac{(a-h)!}{(a-h)^{a-h}}\cdot\frac{a^a}{a!} &\geq \left(\sqrt{2\pi}\cdot\sqrt{h}\cdot e^{-h}\right)\cdot\left(\sqrt{2\pi}\cdot\sqrt{a-h}\cdot e^{-a+h}\right)\cdot\left(e\cdot\sqrt{a}\cdot e^{-a}\right)^{-1}\\ &=\frac{2\pi}{e}\cdot\sqrt{h\cdot\frac{a-h}{a}}\geq\frac{2\pi}{e}\cdot\sqrt{z\cdot\frac{1}{2}}=\frac{2\pi}{e\cdot\sqrt{2}}\cdot\sqrt{z}\geq\frac{6}{3\cdot 2}\cdot\sqrt{z}=\sqrt{z},\end{aligned}$$

where in the second inequality we used that $z\leq h\leq s\leq a/2$. Consequently, the probability of having $B\cup\{v_1,\dots,v_a\}=A$ is at least

$$\frac{3}{4}\cdot\frac{h!}{h^h}\cdot\frac{(a-h)!}{(a-h)^{a-h}}\cdot\frac{a^a}{a!}\cdot\frac{a!}{n^a}\geq\frac{3}{4}\cdot\sqrt{z}\cdot\frac{a!}{n^a},$$

as desired. □

Recall that we fixed the set $B\subseteq V(G)$ and we chose a tame sequence $v_1,\dots,v_a$ according to the random procedure described above. As the events $B\cup\{v_1,\dots,v_a\}=A$ are clearly disjoint for different sets A, by Corollary 5.8 the number of sets $A\subseteq V(G)$ such that (A,B) is a z-nice pair can be at most

$$\left(\frac{3}{4}\cdot\sqrt{z}\cdot\frac{a!}{n^a}\right)^{-1}=\frac{4}{3}\cdot\frac{1}{\sqrt{z}}\cdot\frac{n^a}{a!}.$$

This finishes the proof of Lemma 5.2.

6 Proof of Theorem 1.2

Now, we will finally prove Theorem 1.2. This section contains the main part of the proof of the theorem, but the proofs of several lemmas will be postponed to the following three sections.

We may assume without loss of generality that the constant C in the statement of Theorem 1.2 satisfies $C \geq 3$ (otherwise we can just make C larger). Furthermore, we may assume that k is sufficiently large and in particular that $k \geq 10^3 C$ and $k \geq 4 \cdot \log^{10} k$.

If $1 \leq \ell \leq \sqrt{k}$, then Theorem 1.4 applied to $r = 2$ implies

$$\operatorname{ind}(k,\ell) = \operatorname{ind}_2(k,\ell) \leq \operatorname{ind}_{\leq 2}(k,\ell) \leq \frac{k}{k-2\ell} \cdot \frac{1}{e} \leq \frac{k}{k-2\sqrt{k}} \cdot \frac{1}{e} = \frac{1}{e} + o_k(1).$$

This proves Theorem 1.2 in the case $1 \leq \ell \leq \sqrt{k}$. So from now on let us assume that $\sqrt{k} \leq \ell \leq C \cdot k$.

Let n be large with respect to k and consider a graph G on n vertices. We will show that for sufficiently large n, the graph G has at most

$$\left(\frac{1}{e} + o_k(1)\right) \cdot \frac{n^k}{k!}$$

different k-vertex subsets $A \subseteq V(G)$ with $e(A) = \ell$. Proving this for all sufficiently large n and all n-vertex graphs G gives

$$\operatorname{ind}(k,\ell) \leq \frac{1}{e} + o_k(1),$$

and finishes the proof of Theorem 1.2.

We can assume that the graph G has at least $e^{-1} \cdot n^k/k!$ different k-vertex subsets $A \subseteq V(G)$ with $e(A) = \ell$, because otherwise we are already done.

Our proof is based on partitioning the vertex set $V(G)$ into vertices of high, medium and low degree. These notions are defined in the following definition.

Definition 6.1. Let us say that a vertex $v \in V(G)$ has *low degree* if $\deg_G(v) \leq 10C \cdot n/k$. Furthermore, let us say that v has *high degree* if $\deg_G(v) \geq 10C \cdot n/(\log^2 k)$. Finally, if $10C \cdot n/k < \deg_G(v) < 10C \cdot n/(\log^2 k)$, let us say that v has *medium degree*.

Let V_{high} be the set of high degree vertices, V_{med} be the set of medium degree vertices, and V_{low} the set of low degree vertices. Then V_{high}, V_{med} and V_{low} form a partition of $V(G)$ (since $k \geq 4 \cdot \log^{10} k > \log^2 k$).

Lemma 6.2. $|V_{\text{low}}| \geq n/12$.

We postpone the proof of Lemma 6.2 to Section 7.

Let $\gamma = e^{-120C}/48$. Note that $0 < \gamma < 1/64 < 1/2$ and γ only depends on the constant C.

In order to prove the desired bound on the number of k-vertex subsets $A \subseteq V(G)$ with $e(A) = \ell$, we will distinguish between those subsets A with $A \cap V_{\text{high}} = \emptyset$ and those with $A \cap V_{\text{high}} \neq \emptyset$. Let us first investigate the k-vertex subsets $A \subseteq V(G)$ with $e(A) = \ell$ and $A \cap V_{\text{high}} = \emptyset$.

Lemma 6.3. *If we choose a k-vertex subset $A \subseteq V_{\text{low}} \cup V_{\text{med}}$ uniformly at random, then we have $m(A) > (1-\gamma)k$ with probability at most $o_k(1)$.*

We will prove Lemma 6.3 in Section 8. The proof will be based on a second moment computation. Using this lemma as well as Corollary 3.4, we can now bound the number of k-vertex subsets $A \subseteq V(G)$ with $e(A) = \ell$ and $A \cap V_{\text{high}} = \emptyset$.

Corollary 6.4. *The number of k-vertex subsets $A \subseteq V(G)$ with $e(A) = \ell$ and $A \cap V_{\text{high}} = \emptyset$ is at most*

$$o_k(1) \cdot \frac{n^k}{k!}.$$

Proof. Lemma 6.3 implies that the number of k-vertex subsets $A \subseteq V(G)$ with $A \cap V_{\text{high}} = \emptyset$ and $m(A) > (1-\gamma)k$ is at most

$$o_k(1) \cdot \binom{|V_{\text{low}} \cup V_{\text{med}}|}{k} \leq o_k(1) \cdot \binom{n}{k} \leq o_k(1) \cdot \frac{n^k}{k!}.$$

In particular, the number of k-vertex subsets $A \subseteq V(G)$ with $e(A) = \ell$, $A \cap V_{\text{high}} = \emptyset$ and $m(A) > (1-\gamma)k$ is at most $o_k(1) \cdot n^k/k!$.

On the other hand, recall that by Claim 3.5 each k-vertex subset $A \subseteq V(G)$ with $e(A) = \ell$ satisfies $m(A) \geq \sqrt{2\ell}$. So by Corollary 3.4 applied to $r = 2$, $\varepsilon = \gamma$ and $c' = \sqrt{2\ell}$, the number of k-vertex subsets $A \subseteq V(G)$ with $e(A) = \ell$ and $m(A) \leq (1-\gamma)k$ is at most

$$\left(32 \cdot \sqrt{2} \cdot \frac{1}{(2\ell)^{1/4}} + 23 \cdot \sqrt{2} \cdot k^{-1/4} \cdot \log k\right) \cdot \frac{n^k}{k!} \leq \left(48 \cdot k^{-1/8} + 36 \cdot k^{-1/4} \cdot \log k\right) \cdot \frac{n^k}{k!} = o_k(1) \cdot \frac{n^k}{k!}.$$

Here we used our assumption that $\ell \geq \sqrt{k}$. Hence, the number of k-vertex subsets $A \subseteq V(G)$ with $e(A) = \ell$, $A \cap V_{\text{high}} = \emptyset$ and $m(A) \leq (1-\gamma)k$ is also at most $o_k(1) \cdot n^k/k!$.

Combining the bounds for these two cases, we see that total number of k-vertex subsets $A \subseteq V(G)$ with $e(A) = \ell$ and $A \cap V_{\text{high}} = \emptyset$ is at most $o_k(1) \cdot n^k/k!$, as desired. □

At this point, it remains to consider the k-vertex subsets $A \subseteq V(G)$ with $e(A) = \ell$ and $A \cap V_{\text{high}} \neq \emptyset$. The following lemma further restricts the subsets to consider.

Lemma 6.5. *There are at most $o_k(1) \cdot n^k/k!$ different k-vertex subsets $A \subseteq V(G)$ with the property that there exists a vertex $v \in A$ with*

$$\left|\deg_A(v) - \frac{k-1}{n} \cdot \deg_G(v)\right| > \sqrt{k \cdot \log k}.$$

We will prove Lemma 6.5 in Section 7. The proof is a relatively standard Chernoff bound computation.

The following definition captures the properties of the k-vertex subsets $A \subseteq V(G)$ that still need to be considered to finish the proof of Theorem 1.2.

Definition 6.6. Let us call a k-vertex subset $A \subseteq V(G)$ *interesting* if $e(A) = \ell$ and $A \cap V_{\text{high}} \neq \emptyset$ and if every vertex $v \in A$ satisfies

$$\left|\deg_A(v) - \frac{k-1}{n} \cdot \deg_G(v)\right| \leq \sqrt{k \cdot \log k}.$$

The next lemma implies an upper bound on the number of interesting k-vertex subsets $A \subseteq V(G)$.

Lemma 6.7. *If we choose a k-vertex subset $A \subseteq V(G)$ uniformly at random, then A is interesting with probability at most*

$$\frac{1}{e} + o_k(1).$$

We will prove Lemma 6.7 in Section 9. Now, let us finish the proof of Theorem 1.2.

Corollary 6.8. *The number of k-vertex subsets $A \subseteq V(G)$ with $e(A) = \ell$ and $A \cap V_{\text{high}} \neq \emptyset$ is at most*

$$\left(\frac{1}{e} + o_k(1)\right) \cdot \frac{n^k}{k!}.$$

Proof. Each k-vertex subset $A \subseteq V(G)$ with $e(A) = \ell$ and $A \cap V_{\text{high}} \neq \emptyset$ is either interesting or it contains a vertex $v \in A$ with $|\deg_A(v) - \deg_G(v) \cdot (k-1)/n| > \sqrt{k \cdot \log k}$. By Lemma 6.5, the number of k-vertex subsets $A \subseteq V(G)$ with this second property is at most $o_k(1) \cdot n^k/k!$.

On the other hand, by Lemma 6.7, the number of interesting k-vertex subsets $A \subseteq V(G)$ is at most

$$\left(\frac{1}{e} + o_k(1)\right) \cdot \binom{n}{k} \leq \left(\frac{1}{e} + o_k(1)\right) \cdot \frac{n^k}{k!}.$$

Thus, the total number of k-vertex subsets $A \subseteq V(G)$ with $e(A) = \ell$ and $A \cap V_{\text{high}} \neq \emptyset$ is also at most $(e^{-1} + o_k(1)) \cdot n^k/k!$, as desired. □

Now, combining Corollaries 6.4 and 6.8 yields that the total number of k-vertex subsets $A \subseteq V(G)$ with $e(A) = \ell$ is at most

$$\left(\frac{1}{e} + o_k(1)\right) \cdot \frac{n^k}{k!}.$$

This finishes the proof of Theorem 1.2.

7 Proof of Lemmas 6.2 and 6.5

In this section, we will prove Lemma 6.2 and 6.5. As a preparation, we start with the following lemma, which is an easy consequence of the Chernoff bound for binomial random variables.

Lemma 7.1. *Suppose we choose a random sequence $v_1, \dots, v_k$ of vertices in $V(G)$ by independently picking each $v_i \in V(G)$ uniformly at random. Let Z be the random variable counting the number of indices $j \in \{2, \dots, k\}$ such that v_1 and v_j are adjacent (and in particular $v_1 \neq v_j$). Then we have*

$$\mathbb{P}\left[|Z - \deg_G(v_1) \cdot (k-1)/n| > \sqrt{k \cdot \log k}\right] < \frac{2}{k^2}.$$

Furthermore, for the conditional probability of $Z \leq \frac{1}{2} \cdot \deg_G(v_1) \cdot (k-1)/n$ conditioned on $v_1 \notin V_{\text{low}}$, we have

$$\mathbb{P}\left[Z \leq \frac{1}{2} \cdot \deg_G(v_1) \cdot \frac{k-1}{n} \,\middle|\, v_1 \notin V_{\text{low}}\right] < e^{-3}.$$

Proof. Let us condition on the choice of v_1, so fix any $v_1 \in V(G)$. Then for each $j \in \{2,\dots,k\}$ the vertex v_j will adjacent to v_1 with probability $\deg_G(v_1)/n$, and this is independent for all $j \in \{2,\dots,k\}$. Hence, when fixing v_1, we see that Z is a binomially distributed random variable with parameters $k-1$ and $\deg_G(v_1)/n$. In particular, the expectation of Z when v_1 is fixed equals $\deg_G(v_1)\cdot(k-1)/n$. So by a Chernoff bound, for example in the form of [2, Corollary A.1.7], we obtain

$$\mathbb{P}\left[|Z-\deg_G(v_1)\cdot(k-1)/n| > \sqrt{k\cdot\log k}\right] \le \mathbb{P}\left[|Z-\deg_G(v_1)\cdot(k-1)/n| > \sqrt{k\cdot\ln k}\right]$$
$$< 2\exp\left(-2\left(\sqrt{k\cdot\ln k}\right)^2\cdot\frac{1}{k-1}\right) = 2\exp\left(-2\cdot\ln k\cdot\frac{k}{k-1}\right) < 2\exp(-2\cdot\ln k) = \frac{2}{k^2},$$

where in the first step we used that $\log k > \ln k$. Since this holds for any choice of $v_1 \in V(G)$, we also obtain the desired bound for the unconditional probability of $|Z-\deg_G(v_1)\cdot(k-1)/n| > \sqrt{k\cdot\log k}$. This proves the first inequality.

For the second inequality, let us again condition on the choice of v_1, but this time we assume $v_1 \notin V_{\text{low}}$. As before, when fixing $v_1 \notin V_{\text{low}}$, we see that Z is a binomially distributed random variable with parameters $k-1$ and $\deg_G(v_1)/n$. By the Chernoff bound for lower tails of binomial random variables (see for example [2, Corollary A.1.13]), we have

$$\mathbb{P}\left[Z \le \frac{1}{2}\cdot\deg_G(v_1)\cdot\frac{k-1}{n}\right] = \mathbb{P}\left[Z \le \frac{1}{2}\cdot(k-1)\cdot\frac{\deg_G(v_1)}{n}\right]$$
$$< \exp\left(-\frac{\left(\frac{1}{2}\cdot(k-1)\cdot\deg_G(v_1)/n\right)^2}{2\cdot(k-1)\cdot\deg_G(v_1)/n}\right) = \exp\left(-\frac{1}{8}\cdot(k-1)\cdot\frac{\deg_G(v_1)}{n}\right)$$
$$\le \exp\left(-\frac{1}{8}\cdot(k-1)\cdot\frac{10C}{k}\right) \le \exp(-C) \le e^{-3},$$

where we used that $\deg_G(v_1) \ge 10C\cdot n/k$ and also $k \ge 5$ and $C \ge 3$. Since this holds for every $v_1 \notin V_{\text{low}}$, we obtain the second inequality in Lemma 7.1. □

Let us now prove Lemma 6.2, which states that $|V_{\text{low}}| \ge n/12$.

Proof of Lemma 6.2. For a sequence $v_1,\dots,v_k$ of (not necessarily distinct) vertices in $V(G)$, define $Z(v_1,\dots,v_k)$ to be the number of indices $j \in \{2,\dots,k\}$ such that v_1 and v_j are adjacent. We first claim that there are at least $n^k/(2e)$ sequences $v_1,\dots,v_k$ with $Z(v_1,\dots,v_k) \le 4C$.

Recall that we assumed at the beginning of Section 6 that there are at least $e^{-1}\cdot n^k/k!$ different k-vertex subsets $A \subseteq V(G)$ with $e(A) = \ell$. Each of these subsets A satisfies

$$\sum_{v\in A}\deg_A(v) = 2e(A) = 2\ell \le 2C\cdot k.$$

Hence each of these subsets A can contain at most $k/2$ vertices v with $\deg_A(v) \ge 4C$. So for each k-vertex subset $A \subseteq V(G)$ with $e(A) = \ell$, there exist at least $k/2$ vertices $v \in A$ with $\deg_A(v) \le 4C$.

Furthermore note that for each k-vertex subset $A \subseteq V(G)$, for any labeling of the vertices of A as $v_1,\dots,v_k$, we have $Z(v_1,\dots,v_k) = \deg_A(v_1)$, since both of these quantities just count how many of the vertices $v_2,\dots,v_k$ are adjacent to v_1. So from each k-vertex subset $A\subseteq V(G)$ with $e(A)=\ell$, we can generate at least $(k/2)\cdot(k-1)!$ labelings of the vertices of A as $v_1,\dots,v_k$ such that $Z(v_1,\dots,v_k)\le 4C$. Indeed, we can just choose $v_1\in A$ such that $\deg_A(v_1)\le 4C$ (and we have at least $k/2$ choices to do so), and then choose any labeling of the $k-1$ remaining vertices as $v_2,\dots,v_k$. Since there are at least $e^{-1}\cdot n^k/k!$ different k-vertex subsets $A\subseteq V(G)$ with $e(A)=\ell$, in total that gives at least

$$\frac{1}{e}\cdot\frac{n^k}{k!}\cdot\frac{k}{2}\cdot(k-1)!=\frac{1}{2e}\cdot n^k$$

sequences $v_1,\dots,v_k$ with $Z(v_1,\dots,v_k)\le 4C$, as claimed above.

Now suppose that we choose a random sequence $v_1,\dots,v_k$ of vertices in $V(G)$ by picking each element uniformly at random in $V(G)$, as in Lemma 7.1. Let $Z=Z(v_1,\dots,v_k)$ be the random variable counting the number of indices $j\in\{2,\dots,k\}$ such that v_1 and v_j are adjacent. Since there are at least $n^k/(2e)$ sequences $v_1,\dots,v_k$ with $Z(v_1,\dots,v_k)\le 4C$, we have

$$\mathbb{P}[Z\le 4C]\ge\frac{1}{2e}. \tag{7.1}$$

On the other hand, by the second part of Lemma 7.1, we have

$$\mathbb{P}\left[Z\le\frac{1}{2}\cdot\deg_G(v_1)\cdot\frac{k-1}{n}\,\middle|\,v_1\notin V_{\text{low}}\right]<e^{-3}.$$

Recall that for every $v_1\notin V_{\text{low}}$, we have $\deg_G(v_1)\ge 10C\cdot n/k$ and therefore (as $k\ge 10^3C\ge 5$)

$$\frac{1}{2}\cdot\deg_G(v_1)\cdot\frac{k-1}{n}\ge\frac{1}{2}\cdot 10C\cdot\frac{n}{k}\cdot\frac{k-1}{n}=5C\cdot\frac{k-1}{k}\ge 4C.$$

Thus,

$$\mathbb{P}\left[Z\le 4C\mid v_1\notin V_{\text{low}}\right]\le\mathbb{P}\left[Z\le\frac{1}{2}\cdot\deg_G(v_1)\cdot\frac{k-1}{n}\,\middle|\,v_1\notin V_{\text{low}}\right]<e^{-3}<\frac{1}{4e}. \tag{7.2}$$

We clearly have

$$\mathbb{P}[Z\le 4C]=\mathbb{P}[v_1\in V_{\text{low}}]\cdot\mathbb{P}\left[Z\le 4C\mid v_1\in V_{\text{low}}\right]+\mathbb{P}[v_1\notin V_{\text{low}}]\cdot\mathbb{P}\left[Z\le 4C\mid v_1\notin V_{\text{low}}\right].$$

Together with (7.1) and (7.2), this implies

$$\frac{1}{2e}\le\mathbb{P}[Z\le 4C]\le\mathbb{P}[v_1\in V_{\text{low}}]\cdot 1+1\cdot\mathbb{P}\left[Z\le 4C\mid v_1\notin V_{\text{low}}\right]\le\mathbb{P}[v_1\in V_{\text{low}}]+\frac{1}{4e}.$$

So we obtain

$$\mathbb{P}[v_1\in V_{\text{low}}]\ge\frac{1}{2e}-\frac{1}{4e}=\frac{1}{4e}.$$

Since v_1 was chosen uniformly at random inside $V(G)$, this means that

$$|V_{\text{low}}|\ge\frac{1}{4e}\cdot n\ge\frac{n}{12}.$$

This finishes the proof of Lemma 6.2. □

Now, let us prove Lemma 6.5.

Proof of Lemma 6.5. As before, for a sequence $v_1,\dots,v_k$ of (not necessarily distinct) vertices in $V(G)$, let $Z(v_1,\dots,v_k)$ be the number of indices $j\in\{2,\dots,k\}$ such that v_1 and v_j are adjacent. We will prove Lemma 6.5 by considering the number of sequences $v_1,\dots,v_k$ such that

$$|Z(v_1,\dots,v_k)-\deg_G(v_1)\cdot(k-1)/n|>\sqrt{k\cdot\log k}. \tag{7.3}$$

The first part of Lemma 7.1 states that when choosing a sequence $v_1,\dots,v_k$ randomly by independently picking each element uniformly from $V(G)$, the probability for (7.3) to hold is less than $2/k^2$. Hence there are at most

$$\frac{2}{k^2}\cdot n^k$$

sequences $v_1,\dots,v_k$ satisfying (7.3).

On the other hand, for each k-vertex subset $A\subseteq V(G)$ with the property that there exists a vertex $v\in A$ with $\left|\deg_A(v)-\frac{k-1}{n}\cdot\deg_G(v)\right|>\sqrt{k\cdot\log k}$, we can generate at least $(k-1)!$ sequences $v_1,\dots,v_k$ with $\{v_1,\dots,v_k\}=A$ that satisfy (7.3). Indeed, first choose $v_1\in A$ to be a vertex with $\left|\deg_A(v_1)-\frac{k-1}{n}\cdot\deg_G(v_1)\right|>\sqrt{k\cdot\log k}$ (there is at least one such choice) and then choose any labeling of the remaining $k-1$ vertices in A as $v_2,\dots,v_k$ (there are $(k-1)!$ choices). Thus, we obtain at least $(k-1)!$ labelings of the vertices of A as $v_1,\dots,v_k$ such that $\left|\deg_A(v_1)-\frac{k-1}{n}\cdot\deg_G(v_1)\right|>\sqrt{k\cdot\log k}$. Now, noting that for each such labeling we have $Z(v_1,\dots,v_k)=\deg_A(v_1)$, we see that we indeed generated $(k-1)!$ sequences $v_1,\dots,v_k$ with $\{v_1,\dots,v_k\}=A$ that satisfy (7.3). Because of the condition $\{v_1,\dots,v_k\}=A$, these $(k-1)!$ sequences are disjoint for different choices of the set A.

Thus, the total number of possible k-vertex subsets $A\subseteq V(G)$ with the property that there exists a vertex $v\in A$ with $\left|\deg_A(v)-\frac{k-1}{n}\cdot\deg_G(v)\right|>\sqrt{k\cdot\log k}$ can be at most

$$\frac{(2/k^2)\cdot n^k}{(k-1)!}=\frac{2}{k}\cdot\frac{n^k}{k!}=o_k(1)\cdot\cdot\frac{n^k}{k!}.$$

This proves Lemma 6.5. □

8 Proof of Lemma 6.3

In this section, we will prove Lemma 6.3. So let us assume that we choose a k-vertex subset $A\subseteq V_{\text{low}}\cup V_{\text{med}}$ uniformly at random. We need to prove that we have $m(A)>(1-\gamma)k$ with probability at most $o_k(1)$.

First, let $n'=|V_{\text{low}}\cup V_{\text{med}}|$. Then each vertex $v\in V_{\text{low}}\cup V_{\text{med}}$ is contained in A with probability k/n'. We clearly have $n'\le n$. On the other hand, Lemma 6.2 implies $n'\ge |V_{\text{low}}|\ge n/12$.

Let Y be the random variable counting the number of vertices $v\in V_{\text{low}}$ with $v\in A$ and such that v is isolated in A. Since $m(A)$ counts the number of non-isolated vertices in A, we always have $m(A)+Y\le |A|=k$. In particular, this gives

$$\mathbb{P}[m(A)>(1-\gamma)k]\le\mathbb{P}[Y<\gamma k].$$

Hence it suffices to prove that the probability for $Y<\gamma k$ is at most $o_k(1)$.

For each vertex $v \in V_{\text{low}}$, let Y_v be the indicator random variable for the event that $v \in A$ and v is isolated in A. By definition we have

$$Y = \sum_{v \in V_{\text{low}}} Y_v.$$

For each $v \in V_{\text{low}}$, let $0 \le \delta(v) \le 1$ be such that $\delta(v)n' = |N(v) \cap (V_{\text{low}} \cup V_{\text{med}})|$. Then the number of vertices in $V_{\text{low}} \cup V_{\text{med}} \setminus \{v\}$ that are not adjacent to v is precisely

$$|V_{\text{low}} \cup V_{\text{med}}| - 1 - |N(v) \cap (V_{\text{low}} \cup V_{\text{med}})| = n' - 1 - \delta(v)n' = (1-\delta(v))n' - 1.$$

Since $v \in V_{\text{low}}$, we have $\delta(v)n' \le |N(v)| = \deg_G(v) \le 10C \cdot n/k$. As $n' \ge n/12$, this gives $\delta(v) \le 120C/k \le 1/4$ (recall that we assumed $k \ge 10^3 C$ at the beginning of Section 6).

First, for each vertex $v \in V_{\text{low}}$, let us analyze the expectation $\mathbb{E}[Y_v]$. This is the probability that $v \in A$ and v is isolated in A. Note that we have $v \in A$ with probability k/n'. Conditioning on this, in order for v to be isolated in A, the remaining $k-1$ vertices have to be chosen in such a way that none of them is adjacent to v. Since there are precisely $(1-\delta(v))n' - 1$ vertices in $V_{\text{low}} \cup V_{\text{med}} \setminus \{v\}$ that are not adjacent to v, the probability for this to happen is

$$\frac{(1-\delta(v))n'-1}{n'-1} \cdot \frac{(1-\delta(v))n'-2}{n'-2} \cdots \frac{(1-\delta(v))n'-k+1}{n'-k+1} = \frac{((1-\delta(v))n'-1)_{k-1}}{(n'-1)_{k-1}}.$$

Hence we obtain

$$\mathbb{E}[Y_v] = \frac{k}{n'} \cdot \frac{((1-\delta(v))n'-1)_{k-1}}{(n'-1)_{k-1}}.$$

Recall that we are operating under the assumption that n is sufficiently large with respect to k. Hence, $n' \ge n/12$ and $(1-\delta(v))n' \ge n'/2 \ge n/24$ are also large with respect to k. Hence, the falling factorials in the numerator and denominator of the previous expression can be approximated by powers and we obtain

$$\mathbb{E}[Y_v] = (1+o_k(1)) \cdot \frac{k}{n'} \cdot \left(\frac{(1-\delta(v))n'}{n'}\right)^{k-1} = (1+o_k(1)) \cdot \frac{k}{n'} \cdot (1-\delta(v))^{k-1} \tag{8.1}$$

for each $v \in V_{\text{low}}$.

Recall that we have $n' \le n$ and $\delta(v) \le 120C/k$ for each $v \in V_{\text{low}}$. Hence we obtain

$$\mathbb{E}[Y_v] = (1+o_k(1)) \cdot \frac{k}{n'} \cdot (1-\delta(v))^{k-1} \ge (1+o_k(1)) \cdot \frac{k}{n} \cdot \left(1 - \frac{120C}{k}\right)^{k-1} = (1+o_k(1)) \cdot \frac{k}{n} \cdot e^{-120C},$$

where we used that $\left(1-\frac{120C}{k}\right)^{k-1}$ converges to e^{-120C} as $k \to \infty$.

Letting $\mu = \mathbb{E}[Y]$, we can use this to obtain a lower bound on μ. Indeed, as $|V_{\text{low}}| \ge n/12$ by Lemma 6.2, we have

$$\mu = \mathbb{E}[Y] = \sum_{v \in V_{\text{low}}} \mathbb{E}[Y_v] \ge |V_{\text{low}}| \cdot (1+o_k(1)) \cdot \frac{k}{n} \cdot e^{-120C} \ge (1+o_k(1)) \cdot \frac{e^{-120C}}{12} \cdot k.$$

So as long as k is sufficiently large, we have

$$\mu = \mathbb{E}[Y] \ge \frac{e^{-120C}}{24} \cdot k = 2\gamma \cdot k. \tag{8.2}$$

We would like to find an upper bound on the variance $\mathbb{E}[(Y-\mu)^2] = \mathbb{E}[Y^2] - \mu^2$. Note that

$$\mathbb{E}[Y^2] = \sum_{(v,w)\in V_{\text{low}}^2} \mathbb{E}[Y_v Y_w].$$

For every pair $(v,w) \in V_{\text{low}}^2$, let $0 \le \delta(v,w) \le 1$ be such that

$$\delta(v,w)n' = |N(v) \cap N(w) \cap (V_{\text{low}} \cup V_{\text{med}})|,$$

which is the number of vertices in $V_{\text{low}} \cup V_{\text{med}}$ that are adjacent to both v and w. Note that we clearly have $\delta(v,w) \le \delta(v)$ and $\delta(v,w) \le \delta(w)$.

We can now analyze the expectation $\mathbb{E}[Y_v Y_w]$ for each pair $(v,w) \in V_{\text{low}}^2$. Note that $Y_v Y_w$ is an indicator random variable for the event that $v,w \in A$ and both of v and w are isolated in A. This is impossible if v and w are connected by an edge, so in this case we have $\mathbb{E}[Y_v Y_w] = 0$. Otherwise, the probability for $v,w \in A$ is

$$\frac{k}{n'} \cdot \frac{k-1}{n'-1}$$

and in order for v and w to be isolated in A, the remaining $k-2$ vertices have to be chosen in such a way that none of them is adjacent to v or to w. The number of vertices in $V_{\text{low}} \cup V_{\text{med}} \setminus \{v,w\}$ that are not adjacent to v or to w is

$$\begin{aligned}|V_{\text{low}} \cup V_{\text{med}}| - 2 - |N(v) \cap (V_{\text{low}} \cup V_{\text{med}})| - |N(w) \cap (V_{\text{low}} \cup V_{\text{med}})| + |N(v) \cap N(w) \cap (V_{\text{low}} \cup V_{\text{med}})| \\ = n' - 2 - \delta(v)n' - \delta(w)n' + \delta(v,w)n' = (1 - \delta(v) - \delta(w) + \delta(v,w))n' - 2.\end{aligned}$$

Thus, if v and w are not adjacent, we obtain

$$\begin{aligned}\mathbb{E}[Y_v Y_w] &= \frac{k}{n'} \cdot \frac{k-1}{n'-1} \cdot \frac{((1-\delta(v)-\delta(w)+\delta(v,w))n'-2)_{k-2}}{(n'-2)_{k-2}} \\ &= (1+o_k(1)) \frac{k^2}{n'^2} \frac{((1-\delta(v)-\delta(w)+\delta(v,w))n')^{k-2}}{n'^{k-2}} = (1+o_k(1)) \frac{k^2}{n'^2} (1-\delta(v)-\delta(w)+\delta(v,w))^{k-2},\end{aligned}$$

again using that $n' \ge n/12$ is large compared to k as well as $\delta(v) \le 1/4$ and $\delta(w) \le 1/4$. So for every pair $(v,w) \in V_{\text{low}}^2$ (independently of whether or not v and w are adjacent) we have

$$\mathbb{E}[Y_v Y_w] \le (1+o_k(1)) \frac{k^2}{n'^2} (1-\delta(v)-\delta(w)+\delta(v,w))^{k-2} = (1+o_k(1)) \frac{k^2}{n'^2} (1-\delta(v)-\delta(w)+\delta(v,w))^{k-1},$$

where in the last step we used that $1 \ge 1 - \delta(v) - \delta(w) + \delta(v,w) \ge 1 - (240C/k)$ as $\delta(v) \le 120C/k$, $\delta(w) \le 120C/k$ and $\delta(v,w) \le \delta(w)$.

Note that using $\delta(v,w) \le \delta(w)$ again, this in particular gives

$$\mathbb{E}[Y_v Y_w] \le (1+o_k(1)) \cdot \frac{k^2}{n'^2} \cdot (1-\delta(v))^{k-1} = (1+o_k(1)) \cdot \frac{k}{n'} \cdot \mathbb{E}[Y_v] \tag{8.3}$$

for all pairs $(v,w) \in V_{\text{low}}^2$, where in the last step we plugged in (8.1).

On the other hand, for all pairs $(v,w) \in V_{\text{low}}^2$ we also obtain

$$\begin{aligned}
\mathbb{E}[Y_v Y_w] \le (1+o_k(1)) \cdot \frac{k^2}{n'^2} \cdot (1-\delta(v)-\delta(w)+\delta(v,w))^{k-1} \\
\le (1+o_k(1)) \cdot \frac{k^2}{n'^2} \cdot (1-\delta(v)-\delta(w)+\delta(v)\delta(w)+\delta(v,w))^{k-1} \\
= (1+o_k(1)) \cdot \frac{k^2}{n'^2} \cdot (1-\delta(v))^{k-1} \cdot (1-\delta(w))^{k-1} \cdot \left(1+\frac{\delta(v,w)}{(1-\delta(v))\cdot(1-\delta(w))}\right)^{k-1} \\
\le (1+o_k(1)) \cdot \mathbb{E}[Y_v] \cdot \mathbb{E}[Y_w] \cdot (1+2\delta(v,w))^{k-1} \le (1+o_k(1)) \cdot \mathbb{E}[Y_v] \cdot \mathbb{E}[Y_w] \cdot e^{2k\cdot\delta(v,w)}, \quad (8.4)
\end{aligned}$$

where in the second-last step we used (8.1) as well as $(1-\delta(v))\cdot(1-\delta(w)) \ge 9/16 \ge 1/2$ (since $\delta(v) \le 1/4$ and $\delta(w) \le 1/4$).

Let us call a pair $(v,w) \in V_{\text{low}}^2$ *special* if $\delta(v,w) \ge 1/(k\cdot\log k)$ and *non-special* otherwise.

For each non-special pair $(v,w) \in V_{\text{low}}^2$ we have $2k\cdot\delta(v,w) < 2/\log k$ and so (8.4) gives

$$\mathbb{E}[Y_v Y_w] \le (1+o_k(1)) \cdot \mathbb{E}[Y_v] \cdot \mathbb{E}[Y_w] \cdot e^{2/\log k} = (1+o_k(1)) \cdot \mathbb{E}[Y_v] \cdot \mathbb{E}[Y_w].$$

Hence

$$\begin{aligned}
\sum_{\substack{(v,w)\in V_{\text{low}}^2\\ \text{non-special}}} \mathbb{E}[Y_v Y_w] \le (1+o_k(1)) \cdot \sum_{\substack{(v,w)\in V_{\text{low}}^2\\ \text{non-special}}} \mathbb{E}[Y_v]\cdot\mathbb{E}[Y_w] \le (1+o_k(1)) \cdot \sum_{(v,w)\in V_{\text{low}}^2} \mathbb{E}[Y_v]\cdot\mathbb{E}[Y_w] \\
= (1+o_k(1)) \cdot \mathbb{E}[Y]^2 = (1+o_k(1)) \cdot \mu^2 \quad (8.5)
\end{aligned}$$

Recall that each vertex $v \in V_{\text{low}}$ satisfies $\deg_G(v) \le 10C\cdot n/k$. In particular, v has at most $10C\cdot n/k$ neighbors $u \in V_{\text{low}} \cup V_{\text{med}}$. Each of these neighbors u has degree $\deg_G(u) \le 10C\cdot n/(\log^2 k)$, so u has at most $10C\cdot n/(\log^2 k)$ neighbors $w \in V_{\text{low}}$. Thus, for each $v \in V_{\text{low}}$, the number of triples $(v,w,u) \in V_{\text{low}} \times V_{\text{low}} \times (V_{\text{low}} \cup V_{\text{med}})$ such that u is a common neighbor of v and w is at most

$$10C\cdot\frac{n}{k}\cdot 10C\cdot\frac{n}{\log^2 k} = 100C^2\cdot\frac{n^2}{k\cdot\log^2 k}.$$

On the other hand the number of such triples $(v,w,u) \in V_{\text{low}} \times V_{\text{low}} \times (V_{\text{low}} \cup V_{\text{med}})$ is precisely

$$\sum_{w\in V_{\text{low}}} |N(v)\cap N(w)\cap(V_{\text{low}}\cup V_{\text{med}})| = \sum_{w\in V_{\text{low}}} \delta(v,w)n'.$$

So we obtain (using $n' \ge n/12$)

$$\sum_{w\in V_{\text{low}}} \delta(v,w) \le \frac{1}{n'}\cdot 100C^2\cdot\frac{n^2}{k\cdot\log^2 k} \le \frac{1}{n'}\cdot 120^2C^2\cdot\frac{n'^2}{k\cdot\log^2 k} = 120^2C^2\cdot\frac{n'}{k\cdot\log^2 k}.$$

For each $w \in V_{\text{low}}$ such that (v,w) is special we have $\delta(v,w) \ge 1/(k\cdot\log k)$. Thus, for each $v \in V_{\text{low}}$, the number of $w \in V_{\text{low}}$ such that (v,w) is special can be at most

$$\frac{120^2C^2\cdot n'/(k\cdot\log^2 k)}{1/(k\cdot\log k)} = 120^2C^2\cdot\frac{n'}{\log k}.$$

Using (8.3), we now obtain

$$\sum_{\substack{(v,w)\in V_{\text{low}}^2\\ \text{special}}} \mathbb{E}[Y_vY_w] \le (1+o_k(1))\cdot\frac{k}{n'}\cdot \sum_{\substack{(v,w)\in V_{\text{low}}^2\\ \text{special}}} \mathbb{E}[Y_v]$$
$$\le (1+o_k(1))\cdot\frac{k}{n'}\cdot\sum_{v\in V_{\text{low}}}\left(120^2C^2\cdot\frac{n'}{\log k}\cdot\mathbb{E}[Y_v]\right)$$
$$= (1+o_k(1))\cdot 120^2C^2\cdot\frac{k}{\log k}\cdot\sum_{v\in V_{\text{low}}}\mathbb{E}[Y_v] = (1+o_k(1))\cdot 120^2C^2\cdot\frac{k}{\log k}\cdot\mu \le o_k(1)\cdot\mu^2, \quad (8.6)$$

where in the last step we used that $k/\log k \le o_k(1)\cdot\mu$ as $\mu \ge (e^{-120C}/24)\cdot k$ by (8.2).

Combining (8.5) and (8.6), we now obtain

$$\mathbb{E}[Y^2] = \sum_{(v,w)\in V_{\text{low}}^2}\mathbb{E}[Y_vY_w] = \sum_{\substack{(v,w)\in V_{\text{low}}^2\\ \text{non-special}}}\mathbb{E}[Y_vY_w] + \sum_{\substack{(v,w)\in V_{\text{low}}^2\\ \text{special}}}\mathbb{E}[Y_vY_w] \le (1+o_k(1))\cdot\mu^2.$$

Thus,

$$\mathbb{E}[(Y-\mu)^2] = \mathbb{E}[Y^2]-\mu^2 = o_k(1)\cdot\mu^2.$$

Now, Markov's inequality gives

$$\mathbb{P}[Y\le\mu/2] \le \mathbb{P}[(Y-\mu)^2\ge\mu^2/4] \le \frac{\mathbb{E}[(Y-\mu)^2]}{\mu^2/4} \le \frac{o_k(1)\cdot\mu^2}{\mu^2/4} = o_k(1).$$

Using that $\mu/2 \ge \gamma k$ by (8.2), this implies

$$\mathbb{P}[Y\le\gamma k] \le \mathbb{P}[Y\le\mu/2] \le o_k(1).$$

This finishes the proof of Lemma 6.3.

9 Proof of Lemma 6.7

In this section we will prove Lemma 6.7, which states that a k-vertex subset $A\subseteq V(G)$ chosen uniformly at random is interesting with probability at most $e^{-1}+o_k(1)$.

First we start with an easy observation: In order for A to be interesting, we need $A\cap V_{\text{high}}\neq\emptyset$. Each of the $|V_{\text{high}}|$ vertices of V_{high} is contained in A with probability k/n. Hence the probability for $A\cap V_{\text{high}}\neq\emptyset$ is at most $|V_{\text{high}}|\cdot k/n$, and in particular the probability that A is interesting is at most $|V_{\text{high}}|\cdot k/n$. Thus, Lemma 6.7 is definitely true if $|V_{\text{high}}|\le e^{-1}\cdot n/k$. So from now, we may assume that $|V_{\text{high}}|\ge e^{-1}\cdot n/k$. Since n is large with respect to k, this in particular implies $|V_{\text{high}}|\ge k$.

Also recall that by Lemma 6.2 we have $|V_{\text{med}}\cup V_{\text{low}}|\ge|V_{\text{low}}|\ge n/12$, so we also have $|V_{\text{med}}\cup V_{\text{low}}|\ge k$ as long as n is large enough with respect to k.

For a uniformly random k-vertex subset $A\subseteq V(G)$, let us consider the quantity $|A\cap V_{\text{high}}|$. Clearly $0\le|A\cap V_{\text{high}}|\le k$. For each $j=0,\dots,k$, let p_j be the probability that $|A\cap V_{\text{high}}|=j$ when A is chosen uniformly at random among all k-vertex subsets of $V(G)$. Then $p_0+\dots+p_k=1$.

The following claim can be checked via a straightforward calculation. We provide a proof in the appendix.

Claim 9.1. *We have $p_j \le e^{-1} + o_k(1)$ for all $1 \le j \le k-1$.*

We could now model the uniform random choice of the k-vertex subset $A \subseteq V(G)$ in the following way: First, choose a random variable $a \in \{0,\dots,k\}$ by taking $a = j$ with probability p_j for each $j = 0,\dots,k$. Then choose an a-vertex subset of V_{high} uniformly at random and independently of that choose a $(k-a)$-vertex subset of $V(G) \setminus V_{\text{high}} = V_{\text{low}} \cup V_{\text{med}}$ uniformly at random. Then taking A to be the union of those two sets results in a uniformly random choice of A among all k-vertex subsets of $V(G)$.

However, we will need a slightly more complicated variant of this idea. So instead, let us model the random choice of A as follows, where each of the three steps is performed independently.

- First choose a sequence $v_1,\dots,v_k$ of distinct vertices in V_{high} uniformly at random among all such sequences of distinct vertices.
- Then choose a k-vertex subset $W \subseteq V_{\text{med}} \cup V_{\text{low}}$ uniformly at random. Now, choose a vertex $w_1 \in W$ uniformly at random. Afterwards, choose a vertex $w_2 \in W \setminus \{w_1\}$ uniformly at random. Continue like this, for each $i = 1,\dots,k$ choosing a vertex $w_i \in W \setminus \{w_1,\dots,w_{i-1}\}$ uniformly at random. This results in a sequence $w_1,\dots,w_k$ of distinct vertices of $V_{\text{med}} \cup V_{\text{low}}$ with $W = \{w_1,\dots,w_k\}$.
- Finally, choose a random variable $a \in \{0,\dots,k\}$ by taking $a = j$ with probability p_j for each $j = 0,\dots,k$.

After having made all these random choices, set $A = \{v_1,\dots,v_a\} \cup \{w_{a+1},\dots,w_k\}$.

Note that for every fixed a, the set $\{v_1,\dots,v_a\}$ will be uniformly random among all a-vertex subsets of V_{high} and the set $\{w_{a+1},\dots,w_k\}$ will be uniformly random among all $(k-a)$-vertex subsets of $V_{\text{med}} \cup V_{\text{low}}$. Therefore, this random procedure yields the same probability distribution for the set A as the slightly simpler procedure described above. In particular, we see that the resulting set A will have a uniform distribution among all k-vertex subsets of $V(G)$.

So let us from now on assume that the random set A is chosen in the described way. Our goal is to show that A is interesting with probability at most $e^{-1} + o_k(1)$.

If A is interesting, we must have $A \cap V_{\text{high}} \ne \emptyset$ and therefore $a = |A \cap V_{\text{high}}| \ge 1$.

Furthermore, if A is interesting, for each vertex v_j with $1 \le j \le a$ we must have

$$\left| \deg_A(v_j) - \frac{k-1}{n} \cdot \deg_G(v_j) \right| \le \sqrt{k \cdot \log k}. \tag{9.1}$$

Hence, for each vertex v_j with $1 \le j \le a$,

$$\begin{aligned} \deg_A(v_j) \ge \frac{k-1}{n} \cdot \deg_G(v_j) - \sqrt{k \cdot \log k} &\ge \frac{k-1}{n} \cdot 10C \cdot \frac{n}{\log^2 k} - \sqrt{k \cdot \log k} \\ &\ge 10C \cdot \frac{k/2}{\log^2 k} - \sqrt{k \cdot \log k} \ge 5C \cdot \frac{k}{\log^2 k} - \frac{k}{\log^2 k} \ge 4C \cdot \frac{k}{\log^2 k}, \end{aligned}$$

where we used that $v_j \in V_{\text{high}}$ as well as the assumption $k \ge 4 \cdot \log^{10} k$ made at the beginning of Section 6.

On the other hand, if A is interesting, we must also have $e(A) = \ell$ and therefore

$$\sum_{j=1}^{a} \deg_A(v_j) \le \sum_{v \in A} \deg_A(v) = 2e(A) = 2\ell \le 2C \cdot k.$$

Thus, if A is interesting, we obtain

$$a \cdot 4C \cdot \frac{k}{\log^2 k} \le \sum_{j=1}^{a} \deg_A(v_j) \le 2C \cdot k$$

and consequently $a \le (\log^2 k)/2 \le \log^2 k$.

So we have shown that if A is interesting, then we must have $1 \le a \le \log^2 k$.

Note that we always have

$$e(A) = e(\{v_1,\dots,v_a\} \cup \{w_{a+1},\dots,w_k\}) = \sum_{j=1}^{a} \deg_A(v_j) - e(\{v_1,\dots,v_a\}) + e(\{w_{a+1},\dots,w_k\}).$$

If A is interesting, then we have $e(A) = \ell$ and therefore

$$e(W \setminus \{w_1,\dots,w_a\}) = e(\{w_{a+1},\dots,w_k\}) = \ell + e(\{v_1,\dots,v_a\}) - \sum_{j=1}^{a} \deg_A(v_j).$$

Now, (9.1) and $a \le \log^2 k$ yield

$$\begin{aligned} e(W \setminus \{w_1,\dots,w_a\}) &\ge \ell - \sum_{j=1}^{a} \left(\frac{k-1}{n} \cdot \deg_G(v_j) + \sqrt{k \cdot \log k} \right) \\ &\ge \ell - \frac{k-1}{n} \cdot \left(\sum_{j=1}^{a} \deg_G(v_j) \right) - \log^2 k \cdot \sqrt{k \cdot \log k} \\ &\ge \ell - \frac{k-1}{n} \cdot \left(\sum_{j=1}^{a} \deg_G(v_j) \right) - \sqrt{k} \cdot \log^3 k \end{aligned}$$

as well as

$$\begin{aligned} e(W \setminus \{w_1,\dots,w_a\}) &\le \ell + \binom{a}{2} - \sum_{j=1}^{a} \left(\frac{k-1}{n} \cdot \deg_G(v_j) - \sqrt{k \cdot \log k} \right) \\ &\le \ell + (\log^2 k)^2 - \frac{k-1}{n} \cdot \left(\sum_{j=1}^{a} \deg_G(v_j) \right) + \log^2 k \cdot \sqrt{k \cdot \log k} \\ &\le \ell - \frac{k-1}{n} \cdot \left(\sum_{j=1}^{a} \deg_G(v_j) \right) + \sqrt{k} \cdot \log^3 k. \end{aligned}$$

All in all, if A is interesting, we must have $1 \le a \le \lfloor \log^2 k \rfloor$ and

$$\ell - \frac{k-1}{n} \left(\sum_{j=1}^{a} \deg_G(v_j) \right) - \sqrt{k} \cdot \log^3 k \le e(W \setminus \{w_1,\dots,w_a\}) \le \ell - \frac{k-1}{n} \left(\sum_{j=1}^{a} \deg_G(v_j) \right) + \sqrt{k} \cdot \log^3 k.$$

For $t = 1, \dots, \lfloor \log^2 k \rfloor$, let us define $\mathscr{E}_t$ to be the event

$$\ell - \frac{k-1}{n}\left(\sum_{j=1}^{t} \deg_G(v_j)\right) - \sqrt{k}\cdot\log^3 k \le e(W \setminus \{w_1, \dots, w_t\}) \le \ell - \frac{k-1}{n}\left(\sum_{j=1}^{t} \deg_G(v_j)\right) + \sqrt{k}\cdot\log^3 k.$$

Note that this event $\mathscr{E}_t$ only depends on $v_1, \dots, v_t$ as well as W and $w_1, \dots, w_t$. In particular, $\mathscr{E}_t$ is not influenced by the random choice of a and the random choices of $w_{t+1}, \dots, w_k$ inside W (recall that we chose $w_{t+1}, \dots, w_k$ after W and $w_1, \dots, w_t$).

We saw above that in order for A to be interesting, we must have $1 \le a \le \lfloor \log^2 k \rfloor$ and the event $\mathscr{E}_a$ needs to hold. Since each of the events $\mathscr{E}_t$ is independent of the choice of a, this implies

$$\mathbb{P}[A \text{ interesting}] \le \mathbb{P}\left[1 \le a \le \lfloor \log^2 k \rfloor \text{ and } \mathscr{E}_a \text{ holds}\right] = \sum_{t=1}^{\lfloor \log^2 k \rfloor} \mathbb{P}[a = t]\cdot \mathbb{P}[\mathscr{E}_t] = \sum_{t=1}^{\lfloor \log^2 k \rfloor} p_t \cdot \mathbb{P}[\mathscr{E}_t]$$

By Claim 9.1 we have $p_t \le e^{-1} + o_k(1)$ for $t = 1, \dots, \lfloor \log^2 k \rfloor$. So we obtain

$$\mathbb{P}[A \text{ interesting}] \le \left(\frac{1}{e} + o_k(1)\right)\cdot \sum_{t=1}^{\lfloor \log^2 k \rfloor} \mathbb{P}[\mathscr{E}_t]. \tag{9.2}$$

Our goal is now to prove that $\sum_{t=1}^{\lfloor \log^2 k \rfloor} \mathbb{P}[\mathscr{E}_t] \le 1 + o_k(1)$. First, note that by the inclusion-exclusion principle we have

$$0 \le \mathbb{P}\left[\text{no } \mathscr{E}_t \text{ for } t = 1, \dots, \lfloor \log^2 k \rfloor \text{ holds}\right] \le 1 - \sum_{t=1}^{\lfloor \log^2 k \rfloor} \mathbb{P}[\mathscr{E}_t] + \sum_{1 \le t < t' \le \lfloor \log^2 k \rfloor} \mathbb{P}[\mathscr{E}_t \wedge \mathscr{E}_{t'}].$$

Thus,

$$\sum_{t=1}^{\lfloor \log^2 k \rfloor} \mathbb{P}[\mathscr{E}_t] \le 1 + \sum_{1 \le t < t' \le \lfloor \log^2 k \rfloor} \mathbb{P}[\mathscr{E}_t \wedge \mathscr{E}_{t'}]. \tag{9.3}$$

We will now find an upper bound for each of the terms $\mathbb{P}[\mathscr{E}_t \wedge \mathscr{E}_{t'}]$ for $1 \le t < t' \le \lfloor \log^2 k \rfloor$.

So let us, for now, fix some indices t and t' with $1 \le t < t' \le \lfloor \log^2 k \rfloor$. Let $W' = W \setminus \{w_1, \dots, w_t\}$. If both $\mathscr{E}_t$ and $\mathscr{E}_{t'}$ hold, we must have

$$e(W') = e(W \setminus \{w_1, \dots, w_t\}) \ge \ell - \frac{k-1}{n}\cdot\left(\sum_{j=1}^{t} \deg_G(v_j)\right) - \sqrt{k}\cdot\log^3 k$$

and

$$e(W' \setminus \{w_{t+1}, \dots, w_{t'}\}) = e(W \setminus \{w_1, \dots, w_{t'}\}) \le \ell - \frac{k-1}{n}\cdot\left(\sum_{j=1}^{t'} \deg_G(v_j)\right) + \sqrt{k}\cdot\log^3 k.$$

Hence

$$\begin{aligned}
\sum_{j=t+1}^{t'} \deg_{W'}(w_j) &\geq e(W') - e(W' \setminus \{w_{t+1}, \dots, w_{t'}\}) \\
&\geq \left(\ell - \frac{k-1}{n} \cdot \left(\sum_{j=1}^{t} \deg_G(v_j)\right) - \sqrt{k} \cdot \log^3 k\right) - \left(\ell - \frac{k-1}{n} \cdot \left(\sum_{j=1}^{t'} \deg_G(v_j)\right) + \sqrt{k} \cdot \log^3 k\right) \\
&= \frac{k-1}{n} \cdot \left(\sum_{j=t+1}^{t'} \deg_G(v_j)\right) - 2 \cdot \sqrt{k} \cdot \log^3 k \geq \frac{k-1}{n} \cdot (t'-t) \cdot 10C \cdot \frac{n}{\log^2 k} - 2 \cdot \sqrt{k} \cdot \log^3 k \\
&\geq (t'-t) \cdot 9C \cdot \frac{k}{\log^2 k} - 2 \cdot \sqrt{k} \cdot \log^3 k \geq (t'-t) \cdot 8C \cdot \frac{k}{\log^2 k},
\end{aligned}$$

where we used $\deg_G(v_j) \geq 10C \cdot n/(\log^2 k)$ for $j = t+1, \dots, t'$ as $v_j \in V_{\text{high}}$, as well as $(k-1)/k \geq 9/10$ and $2\sqrt{k} \cdot \log^3 k \leq k/(\log^2 k)$ since we assumed $k \geq 4 \cdot \log^{10} k$ at the beginning of Section 6. Thus, at least one of the vertices w_j for $j = t+1, \dots, t'$ must satisfy $\deg_{W'}(w_j) \geq 8C \cdot k/(\log^2 k)$ if both of the events $\mathscr{E}_t$ and $\mathscr{E}_{t'}$ hold.

Let us now fix any choice of $v_1, \dots, v_k$, W and $w_1, \dots, w_t$ such that the event $\mathscr{E}_t$ holds. We want to prove that then the probability for $\mathscr{E}_{t'}$ to hold as well is small (the choice of $w_{t+1}, \dots, w_k$ in $W \setminus \{w_1, \dots, w_t\}$ is still random as described above). As before, let $W' = W \setminus \{w_1, \dots, w_t\}$. Since $\mathscr{E}_t$ holds we have

$$e(W') = e(W \setminus \{w_1, \dots, w_t\}) \leq \ell - \frac{k-1}{n} \cdot \left(\sum_{j=1}^{t} \deg_G(v_j)\right) + \sqrt{k} \cdot \log^3 k \leq C \cdot k + \sqrt{k} \cdot \log^3 k \leq 2C \cdot k,$$

where we used the assumption $k \geq 4 \cdot \log^{10} k$ again. Thus,

$$\sum_{w \in W'} \deg_{W'}(w) = 2e(W') \leq 4C \cdot k.$$

This in particular implies that the number of vertices $w \in W'$ with $\deg_{W'}(w) \geq 8C \cdot k/(\log^2 k)$ is at most $(\log^2 k)/2$. We saw above that in order for the event $\mathscr{E}_{t'}$ to hold in addition to $\mathscr{E}_t$, we must have $\deg_{W'}(w_j) \geq 8C \cdot k/(\log^2 k)$ for at least one index $j \in \{t+1, \dots, t'\}$. There are $t'-t$ choices for this index j and for each fixed $j \in \{t+1, \dots, t'\}$ the vertex w_j is a uniformly random element of $W' = W \setminus \{w_1, \dots, w_t\}$. Hence for each fixed j the probability of $\deg_{W'}(w_j) \geq 8C \cdot k/(\log^2 k)$ is at most

$$\frac{(\log^2 k)/2}{|W'|} = \frac{(\log^2 k)/2}{k-t} \leq \frac{(\log^2 k)/2}{k - \log^2 k} \leq \frac{(\log^2 k)/2}{k/2} = \frac{\log^2 k}{k}.$$

By the union bound, the probability for there to exist an index $j \in \{t+1, \dots, t'\}$ with $\deg_{W'}(w_j) \geq 8C \cdot k/(\log^2 k)$ is therefore at most

$$(t'-t) \cdot \frac{\log^2 k}{k} \leq (\log^2 k) \cdot \frac{\log^2 k}{k} = \frac{\log^4 k}{k},$$

where we used that $1 \leq t < t' \leq \lfloor \log^2 k \rfloor$. This proves that for any fixed choice of $v_1, \dots, v_k$, W and $w_1, \dots, w_t$ such that the event $\mathscr{E}_t$ holds, the probability for the event $\mathscr{E}_{t'}$ to hold as well is at most $(\log^4 k)/k$. Hence, the overall probability for both $\mathscr{E}_t$ and $\mathscr{E}_{t'}$ to hold is also at most $(\log^4 k)/k$.

So we have shown that for any fixed indices t and t' with $1 \le t < t' \le \lfloor \log^2 k \rfloor$, we have

$$\mathbb{P}[\mathscr{E}_t \wedge \mathscr{E}_{t'}] \le \frac{\log^4 k}{k}.$$

Thus, (9.3) yields

$$\sum_{t=1}^{\lfloor \log^2 k \rfloor} \mathbb{P}[\mathscr{E}_t] \le 1 + \sum_{1 \le t < t' \le \lfloor \log^2 k \rfloor} \mathbb{P}[\mathscr{E}_t \wedge \mathscr{E}_{t'}] \le 1 + \sum_{1 \le t < t' \le \lfloor \log^2 k \rfloor} \frac{\log^4 k}{k} \le 1 + \frac{\log^8 k}{k} = 1 + o_k(1).$$

Plugging this into (9.2), we obtain

$$\mathbb{P}\left[A \text{ interesting}\right] \le \left(\frac{1}{e} + o_k(1)\right) \cdot \sum_{t=1}^{\lfloor \log^2 k \rfloor} \mathbb{P}[\mathscr{E}_t] \le \left(\frac{1}{e} + o_k(1)\right) \cdot (1 + o_k(1)) = \frac{1}{e} + o_k(1),$$

as desired. This finishes the proof of Lemma 6.7.

Appendix

Here, we provide a proof of Claim 9.1. As mentioned above, the proof is a straightforward calculation.

Let $1 \le j \le k-1$. By definition, p_j is the probability that a uniformly random k-vertex subset $A \subseteq V(G)$ satisfies $|A \cap V_{\text{high}}| = j$. Hence

$$p_j = \binom{k}{j} \cdot \frac{(|V_{\text{high}}|)_j \cdot |(V_{\text{low}} \cup V_{\text{med}}|)_{k-j}}{(n)_k} \le \binom{k}{j} \cdot \frac{|V_{\text{high}}|^j \cdot |V_{\text{low}} \cup V_{\text{med}}|^{k-j}}{(n)_k}$$
$$= (1 + o_k(1)) \cdot \binom{k}{j} \cdot \frac{|V_{\text{high}}|^j \cdot |V_{\text{low}} \cup V_{\text{med}}|^{k-j}}{n^k},$$

where in the last step we used that n is large in terms of k.

Let x be such that $|V_{\text{high}}| = xn$, then $|V_{\text{low}} \cup V_{\text{med}}| = (1-x)n$. Recall that we assumed $|V_{\text{high}}| \ge k$ and $|V_{\text{low}} \cup V_{\text{med}}| \ge k$ at the beginning of Section 9, and consequently $0 < x < 1$. Now we have

$$p_j \le (1 + o_k(1)) \cdot \binom{k}{j} \cdot \frac{(xn)^j \cdot ((1-x)n)^{k-j}}{n^k} = (1 + o_k(1)) \cdot \frac{k!}{j! \cdot (k-j)!} \cdot x^j \cdot (1-x)^{k-j}.$$

Note that the inequality between arithmetic and geometric mean implies

$$x^j (1-x)^{k-j} = j^j \cdot (k-j)^{k-j} \cdot \left(\frac{x}{j}\right)^j \cdot \left(\frac{1-x}{k-j}\right)^{k-j}$$
$$\le j^j \cdot (k-j)^{k-j} \cdot \left(\frac{j \cdot \frac{x}{j} + (k-j) \cdot \frac{1-x}{k-j}}{j + k - j}\right)^{j+k-j} = j^j \cdot (k-j)^{k-j} \cdot \left(\frac{1}{k}\right)^k.$$

Thus,

$$p_j \le (1+o_k(1)) \cdot \frac{k!}{j! \cdot (k-j)!} \cdot x^j \cdot (1-x)^{k-j} \le (1+o_k(1)) \cdot \frac{j^j}{j!} \cdot \frac{(k-j)^{k-j}}{(k-j)!} \cdot \frac{k!}{k^k}.$$

Hence it suffices to prove

$$\frac{j^j}{j!} \cdot \frac{(k-j)^{k-j}}{(k-j)!} \cdot \frac{k!}{k^k} \le \frac{1}{e} + o_k(1)$$

for all $1 \le j \le k-1$. For this, we can assume without loss of generality that $1 \le j \le k/2$.

If $2 \le j \le k/2$, then $j \cdot (k-j) \ge 2 \cdot (k-2)$ and so Stirling's formula (see for example [9]) yields

$$\frac{j^j}{j!} \cdot \frac{(k-j)^{k-j}}{(k-j)!} \cdot \frac{k!}{k^k} \le \left(\sqrt{2\pi} \cdot \sqrt{j} \cdot e^{-j}\right)^{-1} \left(\sqrt{2\pi} \cdot \sqrt{k-j} \cdot e^{-k+j}\right)^{-1} \left((1+o_k(1))\sqrt{2\pi} \cdot \sqrt{k} \cdot e^{-k}\right)$$
$$= (1+o_k(1))\frac{1}{\sqrt{2\pi}} \cdot \sqrt{\frac{k}{j \cdot (k-j)}} \le (1+o_k(1))\frac{1}{\sqrt{2\pi}} \cdot \sqrt{\frac{k}{2 \cdot (k-2)}} = (1+o_k(1))\frac{1}{\sqrt{4\pi}} < \frac{1}{e} + o_k(1),$$

where in the last step we used $4\pi > 12 > 9 > e^2$. So it only remains to consider the case $j = 1$. Then we simply have

$$\frac{1^1}{1!} \cdot \frac{(k-1)^{k-1}}{(k-1)!} \cdot \frac{k!}{k^k} = \frac{(k-1)^{k-1}}{k^{k-1}} = \left(1 - \frac{1}{k}\right)^{k-1} \le e^{-(k-1)/k} = e^{-1} \cdot e^{1/k} = \frac{1}{e} + o_k(1).$$

This finishes the proof of Claim 9.1.

Acknowledgments

We would like to thank the anonymous referee for their helpful comments.

Remark. After completing this work, we learned that Martinsson, Mousset, Noever, and Trujić [7] also found a completion of the proof of the Edge-statistics Conjecture.

References

[1] N. Alon, D. Hefetz, M. Krivelevich, and M. Tyomkyn, *Edge-statistics on large graphs*, preprint, 2018, arXiv:1805.06848. 1, 2, 3, 4

[2] N. Alon and J. H. Spencer, *The Probabilistic Method*, 4th ed., Wiley, 2016. 19, 27, 29, 30, 39

[3] J. Balogh, P. Hu, B. Lidický, and F. Pfender, *Maximum density of induced 5-cycle is achieved by an iterated blow-up of 5-cycle*, European J. Combin. **52** (2016), 47–58. 4

[4] D. Hefetz and M. Tyomkyn, *On the inducibility of cycles*, J. Combin. Theory Ser. B **133** (2018), 243–258. 4

[5] D. Král', S. Norin, and J. Volec, *A bound on the inducibility of cycles*, J. Combin. Theory Ser. A **161** (2019), 359–363. 4

[6] M. Kwan, B. Sudakov, and T. Tran, *Anticoncentration for subgraph statistics*, J. Lond. Math. Soc. (2) **99** (2019), 757–777. 2, 3

[7] A. Martinsson, F. Mousset, A. Noever, and M. Trujić, *The edge-statistics conjecture for* $\ell \ll k^{6/5}$, Israel J. Math. **234** (2019), 677–690. 51

[8] N. Pippenger and M. C. Golumbic, *The inducibility of graphs*, J. Combin. Theory Ser. B **19** (1975), 189–203. 4

[9] H. Robbins, *A remark on Stirling's formula*, Amer. Math. Monthly **62** (1955), 26–29. 24, 35, 51

[10] R. Yuster, *On the exact maximum induced density of almost all graphs and their inducibility*, J. Combin. Theory Ser. B **136** (2019), 81–109. 4

AUTHORS

Jacob Fox
Department of Mathematics
Stanford University
Stanford, USA
jacobfox@stanford.edu
https://stanford.edu/~jacobfox

Lisa Sauermann
Department of Mathematics
Stanford University
Stanford, USA
lsauerma@stanford.edu
https://stanford.edu/~lsauerma