

# $K_4^-$ -free triple systems without large stars in the complement

Dhruv Mubayi\*

Nicholas Spanier†

April 7, 2025

## Abstract

The  $n$ -star  $S_n$  is the  $n$ -vertex triple system with  $\binom{n-1}{2}$  edges all of which contain a fixed vertex, and  $K_4^-$  is the unique triple system with four vertices and three edges. We prove that the Ramsey number  $r(K_4^-, S_n)$  has order of magnitude  $n^2/\log n$ .

This confirms a conjecture of Conlon, Fox, He, Suk, Verstraëte and the first author. It also generalizes the well-known bound of Kim for the graph Ramsey number  $r(3, n)$ , as the link of any vertex in a  $K_4^-$ -free triple system is a triangle-free graph. Our method builds on the approach of Guo and Warnke who adapted Kim's lower bound for  $r(3, n)$  to the pseudorandom setting.

## 1 Introduction

A  $k$ -uniform hypergraph  $H$  (henceforth  $k$ -graph) is a pair  $(V(H), E(H))$  where  $E(H)$  is a collection of  $k$ -element subsets of  $V(H)$ . We often associate  $H$  with its edge set  $E(H)$ , when  $V(H)$  is obvious from context. Write  $K_n^k$  for the complete  $k$ -graph on  $n$  vertices. Given  $k$ -graphs  $G$  and  $H$ , the Ramsey number  $r(G, H)$  is the minimum  $N$  such that every red/blue coloring of  $E(K_N^k)$  has a monochromatic red copy of  $G$  or a monochromatic blue copy of  $H$ ; when  $H = K_n^k$ , we simply write  $r(G, n)$ . When  $G$  is fixed and  $n \rightarrow \infty$ , we often refer to this class of parameters as off-diagonal Ramsey numbers.

The off-diagonal graph Ramsey numbers  $r(s, n) = r(K_s, K_n)$  have been intensively studied for many decades. Erdős conjectured that for fixed  $s \geq 3$ , we have  $r(s, n) = n^{s-1+o(1)}$  and this was reiterated recently by the first author and Verstraëte [12]. Only the first two cases,  $s = 3, 4$  have been proved. The order of magnitude of  $r(3, n)$  is  $n^2/\log n$ , where the upper bound was obtained by Ajtai-Komlós-Szemerédi [1] and the lower bound was proved by Kim [7] and later by Bohman [3] using a different method. More recently, Mattheus and Verstraëte proved that  $r(4, n) = n^{3-o(1)}$  with the order of magnitude still undetermined. In this paper we consider an analogous parameter to  $r(3, n)$  in the hypergraph setting.

---

\*Department of Mathematics, Statistics and Computer Science, University of Illinois, Chicago, IL 60607. Email: mubayi@uic.edu. Research partially supported by NSF Awards DMS-1952786 and DMS-2153576.

†Department of Mathematics Statistics and Computer Science, University of Illinois, Chicago, IL 60607. Email: spanier2@uic.edu. Research partially supported by NSF Awards DMS-1952786 and DMS-2153576.

It is well-known that for each  $s \geq 4$ , the hypergraph Ramsey number  $r(K_4^3, K_n^3) > 2^{cn}$  for some positive constant  $c$ . On the other hand, there is a well-known large class  $\mathcal{C}$  of 3-graphs  $G$  for which it is known that  $r(G, n)$  is polynomial in  $n$ . Recently, it was conjectured [4] that  $r(G, n)$  is at most polynomial in  $n$  if and only if  $G \in \mathcal{C}$ . On the other hand, it is not clear whether  $G \notin \mathcal{C}$  implies that  $r(G, n)$  is exponential in  $n$ . In fact, the following result from [5] showed an intermediate growth rate for the following related problem. The link graph of a vertex  $v$  in a 3-graph  $H$  is the graph whose vertex set is  $V(H) \setminus \{v\}$  and edge set is  $\{yz : vyz \in E(H)\}$ . Write  $S_n$  for the  $n$ -vertex star: this is the 3-graph with vertex set  $[n] := \{1, \dots, n\}$  and edge set  $\{e \in \binom{[n]}{3} : 1 \in e\}$ . Alternatively,  $S_n$  is the 3-graph for which there is a vertex whose link graph is isomorphic to  $K_{n-1}$ . The 3-graph  $S_n$  is a natural way to generalize  $K_n$  to hypergraphs. It is proved in [5] that there are positive constants  $c, c'$  such that

$$2^{c \log^2 n} < r(K_4^3, S_n) < 2^{c' n^{2/3} \log n}. \quad (1)$$

This shows that  $r(K_4^3, S_n)$  is already superpolynomial in  $n$ . Therefore, one needs to consider sparser 3-graphs  $G$  to detect polynomial growth rates for  $r(G, S_n)$ . The smallest such nontrivial 3-graph  $G$  is  $K_4^-$ , the unique 3-graph with four vertices and three edges (if  $G$  is the 3-graph on four vertices and two edges, then it is trivial to see that  $r(G, S_n) = \Theta(n)$ ).

The link graph  $L$  of a vertex in an  $N$ -vertex  $K_4^-$ -free 3-graph  $H$  is a triangle-free graph on  $N - 1$  vertices, and hence by the result  $r(3, n) = O(n^2 / \log n)$  of [1], the independence number of  $L$  is at least  $\Omega(\sqrt{N \log N})$ . Consequently, the complement of every  $K_4^-$ -free 3-graph on  $N$  vertices contains an  $n$ -vertex star with  $n = \Omega(\sqrt{N \log N})$ . This shows

$$r(K_4^-, S_n) = O\left(\frac{n^2}{\log n}\right).$$

In the other direction, it was proved in [5] using the local lemma that  $r(K_4^-, S_n) = \Omega(n^2 / \log^2 n)$ . This led the authors of [5] to conjecture that

$$r(K_4^-, S_n) = \Theta\left(\frac{n^2}{\log n}\right). \quad (2)$$

A similar state of affairs existed for the graph case  $r(3, n)$  prior to Kim's improved lower bound. Note, however, that (2) is substantially stronger than Kim's result  $r(3, n) = \Omega(n^2 / \log n)$  as it posits the existence of a triple system such that *every* link graph is an optimal (in order of magnitude)  $r(3, n)$  graph. Also, as indicated in (1), hypergraph problems sometimes display novel phenomena compared to what we typically witness for graphs, so it was by no means a forgone conclusion that (2) should hold similar to the graph result  $r(3, n) = \Theta(n^2 / \log n)$ . In this paper, we prove (2).

**Theorem 1.** *There is an absolute constant  $c$  such that for all sufficiently large  $N$ , there is a  $K_4^-$ -free 3-graph on  $N$  vertices whose complement contains no copy of  $S_n$ , where  $n = c\sqrt{N \log N}$ . In other words,  $r(K_4^-, S_n) = \Omega(n^2 / \log n)$ .*

Our theorem follows from the following stronger result, which guarantees not only that for every vertex  $x$  and every  $n$ -set  $A$  omitting  $x$ , there is an edge of the form  $xyz$ , where  $yz \in \binom{A}{2}$ , but in fact that there are many such edges. It is more convenient to consider the bipartite setting as follows. Given a 3-graph  $H$ , disjoint sets  $A, B \subseteq V(H)$ , and  $x \in V(H) \setminus (A \cup B)$ , let  $e_{H,x}(A, B)$  be the

number of edges in  $H$  of the form  $xab$  where  $a \in A$  and  $b \in B$ . The following theorem is our main result.

**Theorem 2 (Main result).** *There exists  $\beta_0, D_0$  such that for all  $\delta \in (0, 1]$ ,  $\beta \in (0, \beta_0)$ ,  $C > \frac{D_0}{\delta^2 \sqrt{\beta}}$  and  $N > N_0(\delta, \beta, C)$ , there exists a  $K_4^-$  free 3-graph  $H$  on  $N$  vertices such that for every two disjoint  $n$ -vertex subsets  $A, B$  with  $n := C\sqrt{N \log N}$  and for every vertex  $x \notin A \cup B$ ,*

$$e_{H,x}(A, B) = (1 \pm \delta)\rho n^2$$

where  $\rho := \sqrt{\frac{\beta \log N}{3N}}$ .

There are two general approaches to obtain graphs or hypergraphs that have good Ramsey properties of the type sought in Theorems 1 and 2: the nibble method pioneered by Ajtai-Koml-Szemerédi [1, 2] and Rödl [13], and the differential equations method developed by Wormald [16] for combinatorial settings. We prove Theorem 2 using the nibble method. This was the approach used by Kim [7] in his proof of  $r(3, n) = \Omega(n^2 / \log n)$ . More recently, Guo and Warnke [6] established a stronger result than  $r(3, n) = \Omega(n^2 / \log n)$  in two ways. First, they not only found one triangle-free graph on  $N$  vertices with independence number  $O(\sqrt{N \log N})$ , but in fact decomposed most of the edges of  $K_N$  into such graphs. Second, they were able to obtain such a decomposition as long as their initial graph was sufficiently pseudorandom. Along the way, they introduced several simplifications of Kim's argument.

Our proof of Theorem 2 follows the broad outline of Kim's approach together with the additional technical ingredients provided by Guo and Warnke. However, creating  $K_4^-$ -free 3-graphs without large stars in their complement, rather than just triangle-free graphs with small independent sets, is more challenging. For example, certain random variables that we must track that appear in our semirandom algorithm do not even make sense to define in the graph case. New ideas are required to implement the additional technical steps needed to control their deviation probability from their mean (see, for example, the event  $\mathcal{N}_i^+$ , or the Proof of Lemma 12 in Section 7.7).

We end by remarking that we were not able to prove Theorem 2 using the random greedy  $K_4^-$ -free process, and this remains open.

## 2 Nibble Algorithm

Throughout the paper we use the notation  $xy$  and  $xyz$  for  $\{x, y\}$  and  $\{x, y, z\}$ , respectively.

We will form a  $K_4^-$ -free 3-graph on an  $N$ -element set  $V$  by iteratively selecting a set of edges and removing a subset of those edges to destroy all copies of  $K_4^-$ . At each step in the process, we will maintain a triple  $(E_i, H_i, O_i)$  where

- $E_i$  is a set of *picked* or *chosen* edges which may contain copies of  $K_4^-$ ,
- $H_i \subseteq E_i$  where  $H_i$  is  $K_4^-$ -free. We will iterate the process  $I$  times and our final 3-graph will be  $H_I$ .

- $O_i \subseteq \binom{V}{3} \setminus E_i$  is a set of *open* edges. No edge in  $O_i$  forms a  $K_4^-$  together with any two edges from  $E_i$ .

We will start by initializing  $E_0 = H_0 = \emptyset$  and  $O_0 = \binom{V}{3}$ . Assume that at step  $i$  we have sets  $E_i, H_i, O_i$  such that  $H_i \subseteq E_i$ ,  $E_i \cap O_i = \emptyset$ ,  $H_i$  is  $K_4^-$  free, and for any  $e \in O_i$  there do not exist  $f, g \in E_i$  such that  $efg$  forms a  $K_4^-$ . Set

$$\sigma := \frac{1}{\log^2 N} \quad \text{and} \quad p := \frac{\sigma}{\sqrt{N}} = \frac{1}{\sqrt{N} \log^2 N}. \quad (3)$$

Form a set of edges  $\Gamma_{i+1} \subseteq O_i$  by including each edge  $e \in O_i$  in  $\Gamma_{i+1}$  independently with probability  $p$ . Let  $E_{i+1} = E_i \cup \Gamma_{i+1}$ . Because  $\Gamma_{i+1} \subseteq O_i$ , the set  $E_{i+1}$  contains no  $K_4^-$  with exactly one edge in  $\Gamma_{i+1}$ .

There may be three edges in  $\Gamma_{i+1}$  which form a  $K_4^-$ , or two edges in  $\Gamma_{i+1}$  and one edge in  $E_i$  which together form a  $K_4^-$ . We gather these bad edges and remove a cleverly chosen subset of them so that there is no  $K_4^-$  of these two types after we have removed the edges. Then we add to  $H_i$  the edges that remain and form  $H_{i+1}$ . This process will ensure that  $H_{i+1}$  is  $K_4^-$  free. We do this as follows. First, let

$$\begin{aligned} B_{i+1}^2 &= \{ef \subseteq \Gamma_{i+1} : \exists g \in H_i, efg \text{ is a } K_4^-\} \\ B_{i+1}^3 &= \{efg \subseteq \Gamma_{i+1} : efg \text{ is a } K_4^-\} \\ B_{i+1} &= B_{i+1}^2 \cup B_{i+1}^3. \end{aligned}$$

Next, let  $D_{i+1}$  be a maximal collection of pairwise disjoint subsets of  $B_{i+1}$ . More precisely  $D_{i+1} \subseteq B_{i+1}$  with the property that for any  $ef \in B_{i+1}^2$ , at least one of  $e, f$  lies in some set in  $D_{i+1}$ , and for any  $e'f'g' \in B_{i+1}^3$ , at least one of  $e', f', g'$  lies in some set in  $D_{i+1}$ . Let

$$H_{i+1} = H_i \cup (\Gamma_{i+1} \setminus \bigcup_{S \in D_{i+1}} \bigcup_{e \in S} \{e\}).$$

In other words, we have added to  $H_i$  all sets that do not appear in an element of  $D_{i+1}$ .

Let us now prove that  $H_{i+1}$  is  $K_4^-$ -free. We may assume (by induction) that  $H_i$  is  $K_4^-$ -free. There is no  $K_4^-$  in  $H_{i+1}$  with exactly two edges in  $H_i$  by the definition of  $O_i$ . There is no  $K_4^-$  in  $H_{i+1}$  with exactly one edge  $g$  in  $H_i$  since the other two edges  $e, f$  must satisfy  $ef \in B_{i+1}^2$  so either  $e$  or  $f$  must lie in some pair from  $D_{i+1}$  which means it is not in  $H_{i+1}$ . Lastly, there is no  $K_4^-$  in  $H_{i+1}$  with all three edges  $e, f, g$  in  $H_{i+1} \setminus H_i$  since  $e, f, g$  must satisfy  $efg \in B_{i+1}^3$  so one of them must lie in some triple from  $D_{i+1}$  which means it is not in  $H_{i+1}$ .

We now explain how the set  $O_i$  is updated to  $O_{i+1}$ . The updated set  $O_{i+1}$  must have the property that each edge in  $O_{i+1}$  does not form a copy of  $K_4^-$  with already chosen edges, so that it will be possible to choose it in step  $i+1$ ; this copy of  $K_4^-$  can have both its remaining edges chosen in step  $i$ , or have one edge chosen in step  $i$  and one edge chosen in some previous step. Formally, for each  $e \in O_i$ , let

$$\hat{S}_i(e) = \{f \in O_i : \exists g \in E_i, efg \text{ is a } K_4^-\}.$$

Next, define

$$C_{i+1}^1 = \{e \in O_i : \hat{S}_i(e) \cap \Gamma_{i+1} \neq \emptyset\}$$

$$C_{i+1}^2 = \{e \in O_i : \exists f, g \in \Gamma_{i+1}, efg \text{ is a } K_4^-\}.$$

Our rough plan is to update  $O_i$  by removing all triples from  $\Gamma_{i+1} \cup C_{i+1}^1 \cup C_{i+1}^2$ . However, for technical reasons, it is convenient to remove some more edges from  $O_i$  as follows. Let  $Y_{i+1} \subseteq O_i$  be the (random) subset obtained by including each  $e \in O_i$  in  $Y_{i+1}$  independently with probability

$$\hat{p}_{e,i} := 1 - (1 - p)^{\max\{6\sqrt{N}q_i(\pi_i + \sqrt{\sigma}) - |\hat{S}_i(e)|, 0\}}. \quad (4)$$

We form  $O_{i+1}$  by removing from  $O_i$  the edges which were picked in  $\Gamma_{i+1}$ , the edges in  $C_{i+1}^1$  and  $C_{i+1}^2$ , and the edges in  $Y_{i+1}$ :

$$O_{i+1} = O_i \setminus (\Gamma_{i+1} \cup C_{i+1}^1 \cup C_{i+1}^2 \cup Y_{i+1}). \quad (5)$$

We will continue iterating this process until we have formed the set  $H_I$  where  $I := \lceil N^\beta \rceil$ . Now let  $H := H_I$ . Our goal is to prove that whp (with high probability)  $H$  satisfies the conditions in Theorem 2.

### 3 Expected Trajectories via Differential Equations

We use the following heuristics to predict the behavior of the size of our sets  $O_i$  and  $E_i$ . Assume that for all  $e \in \binom{V}{3}$ , we have

$$\mathbb{P}(e \in O_i) \approx q_i \quad \text{and} \quad \mathbb{P}(e \in E_i) \approx \frac{\pi_i}{\sqrt{N}},$$

where  $\pi_i = O(\sqrt{\log N})$  and these events are approximately independent for all  $e$ . Now notice that

$$\mathbb{P}(e \in E_{i+1}) - \mathbb{P}(e \in E_i) = \mathbb{P}(e \in \Gamma_{i+1} | e \in O_i) \cdot \mathbb{P}(e \in O_i) \approx pq_i = \frac{\sigma q_i}{\sqrt{N}}. \quad (6)$$

Then multiplying (6) by  $\sqrt{N}$  gives  $\pi_{i+1} - \pi_i \approx \sigma q_i$ . Also since  $E_0 = \emptyset$ , we have  $\pi_0 \approx 0$ .

Next, we approximate  $O_i \setminus O_{i+1} = \Gamma_{i+1} \cup C_{i+1}^1 \cup C_{i+1}^2 \cup Y_{i+1} \approx C_{i+1}^1 \cup Y_{i+1}$ . Notice

$$\mathbb{E}(|\hat{S}_i(e)|) \approx 3(N-3)q_i \left(1 - \left(1 - \frac{\pi_i}{\sqrt{N}}\right)^2\right) \approx 3Nq_i \left(\frac{2\pi_i}{\sqrt{N}}\right) \approx 6\sqrt{N}q_i\pi_i. \quad (7)$$

Thus by the estimate for  $O_i \setminus O_{i+1}$  and from (4) and (7), we get that for all  $e \in O_i$

$$\mathbb{P}(e \in O_{i+1} | e \in O_i) \approx (1 - \hat{p}_{e,i})(1 - p)^{1+|\hat{S}_i(e)|} \approx (1 - p)^{6\sqrt{N}q_i(\pi_i + \sqrt{\sigma})} \approx 1 - 6q_i\pi_i\sigma.$$

This heuristic suggests that  $\mathbb{P}(e \in O_{i+1} | e \in O_i)$  depends only on the step  $i$  of our iteration, and this is the reason we introduced the stabilization probability  $\hat{p}_{e,i}$ . Therefore,

$$q_{i+1} - q_i \approx \mathbb{P}(e \in O_{i+1}) - \mathbb{P}(e \in O_i) = \mathbb{P}(e \in O_{i+1} | e \in O_i)\mathbb{P}(e \in O_i) - \mathbb{P}(e \in O_i) \approx -6q_i^2\pi_i\sigma.$$

Next suppose  $\pi_i \approx \Psi(i\sigma)$  where  $\Psi$  is a smooth function. Then

$$\Psi'(i\sigma) \approx \frac{\Psi((i+1)\sigma) - \Psi(i\sigma)}{\sigma} \approx \frac{\pi_{i+1} - \pi_i}{\sigma} \approx q_i$$

and for  $x = i\sigma$ ,

$$\Psi''(x) \approx \Psi''(i\sigma) = \frac{\Psi'((i+1)\sigma) - \Psi'(i\sigma)}{\sigma} \approx \frac{q_{i+1} - q_i}{\sigma} \approx -6q_i^2 \pi_i \approx -6(\Psi'(x))^2 \Psi(x).$$

Solving this differential equation yields

$$\Psi'(x) = e^{-3\Psi^2(x)}. \quad (8)$$

Moreover, writing  $f = \Psi^{-1}$  so that  $(f\Psi)(x) = x$  for  $x \geq 0$ , and taking derivatives we obtain

$$1 = (f\Psi)'(x) = f'(\Psi(x))\Psi'(x) = f'(\Psi(x))e^{-3\Psi^2(x)}.$$

Writing  $t = \Psi(x)$ , we get  $f'(t) = e^{3t^2}$ . Further, since  $\pi_0 \approx 0$  and  $q_0 \approx 1$ , we set  $\Psi(0) = 0$  and  $\Psi'(0) = 1$ . Consequently,

$$x = f(\Psi(x)) - f(\Psi(0)) = \int_{\Psi(0)}^{\Psi(x)} f'(t) dt = \int_0^{\Psi(x)} e^{3t^2} dt. \quad (9)$$

Now we will formally define  $\Psi(x)$  as the solution to the differential equation (8) with  $\Psi(0) = 0$ . Then (9) also holds and we define

$$q_i = \Psi'(i\sigma) \quad (10)$$

$$\pi_i = \sigma + \sum_{j=0}^{i-1} \sigma q_j. \quad (11)$$

## 4 Technical Estimates

In this section will collect some bounds on  $q_i$  and  $\pi_i$  which we will use throughout the proof of the main theorem. The proofs of these bounds will be given in the Appendix. Recall that

$$\sigma = \frac{1}{\log^2 N}, \quad \beta \in (0, \beta_0), \quad I = \lceil N^\beta \rceil.$$

**Lemma 3.** *Let  $\Psi(x)$ ,  $q_i$ , and  $\pi_i$  be defined as in (9), (10), and (11).*

$$\sqrt{\frac{\log(\sqrt{3}x)}{3}} - \frac{1}{\sqrt{3}} \leq \Psi(x) \leq \sqrt{\frac{\log(\sqrt{3}x)}{3}} + \frac{1}{\sqrt{3}} \text{ for } x \geq e \quad (12)$$

$$0 \leq q_i \leq q_0 = 1 \text{ for } i \geq 0 \quad (13)$$

$$\pi_{i+1} - \pi_i = \sigma q_i \text{ and } \pi_i - \Psi(i\sigma) \in [\sigma, 2\sigma] \text{ for all } i \geq 0 \quad (14)$$

$$\sqrt{\sigma} \pi_i \leq 1 \text{ for } 0 \leq i \leq I \quad (15)$$

$$q_i \pi_i^k \leq 1 \text{ for } 0 \leq i \leq I \text{ and } k \in \{1, 2\} \quad (16)$$

$$|(q_i - q_{i+1}) - 6\sigma q_i^2 \pi_i| \leq 16\sigma^2 q_i^2 \text{ for all } i \geq 0 \quad (17)$$

$$q_i \geq \frac{1}{10} N^{-\beta+o(1)} \text{ for all } 0 \leq i \leq I \quad (18)$$

$$\hat{p}_{e,i} \leq q_i \text{ for all } e \in O_i \quad (19)$$

$$0 \leq q_i - q_{i+1} \leq 12\sigma \min\{q_i, q_{i+1}, q_i \pi_i\} \quad (20)$$

In particular, if we let  $\beta_0 = \frac{1}{100}$ , then by (18) and  $\beta \leq \beta_0$ , there exists  $c > 0$  such that the following holds for all  $0 \leq i \leq I$  and  $j < 10$ :

$$q_i^j \sqrt{N} \geq N^{-j\beta+\frac{1}{4}} \geq N^c. \quad (21)$$

## 5 Events

Let  $X_i = (E_i, H_i, O_i, \Gamma_i, Y_i)$  and  $X = \{X_i\}_{0 \leq i \leq I}$ . We let  $\mathcal{F}_i$  denote the  $\sigma$ -algebra generated by  $\{X_j\}_{0 \leq j \leq i}$  and  $(\mathcal{F}_i)_{0 \leq i \leq I}$  be the natural filtration with  $\{X_i\}_{0 \leq i \leq I}$ . Recall that we use the notation  $abc$  for  $\{a, b, c\}$ .

Next we give definitions which apply to an edge  $xuv$  with a distinguished vertex  $x$ .

$$R_i(x, uv) := \{w \in [N] \setminus \{u, v, x\} : xuw, xvw \in O_i\}$$

$$S_i(x, uv) = \{w \in [N] \setminus \{u, v, w\} : |\{xuw, xvw\} \cap O_i| = |\{xuw, xvw\} \cap E_i| = 1\}$$

$$T_i(x, uv) := \{w \in [N] \setminus \{u, v, x\} : xuw, xvw \in E(i)\}$$

$$U_i(x, u, v, w) := \{z \in [N] \setminus \{x, u, v, w\} : xuv, xuw, xuz \in O_i \text{ and } xvw, uwz \in E_i\}$$

$$\hat{S}_i(x, uv) = \{xuw \in O_i : w \in S_i(x, uv)\}$$

The following figures depict the above definitions. Note for  $S_i(x, uv)$  that there are two types of edges, one where we include  $xuw \in O_i$  and  $xvw \in E_i$  and the other where we have  $xuw \in E_i$  and  $xvw \in O_i$ .

Note that  $S_i(x, uv)$  denotes a set of vertices which have one open edge and one closed edge with the pairs  $xu$  and  $xv$ , and  $\hat{S}_i(x, uv)$  describes the corresponding set of open edges. Thus  $|S_i(x, uv)| = |\hat{S}_i(x, uv)|$ .

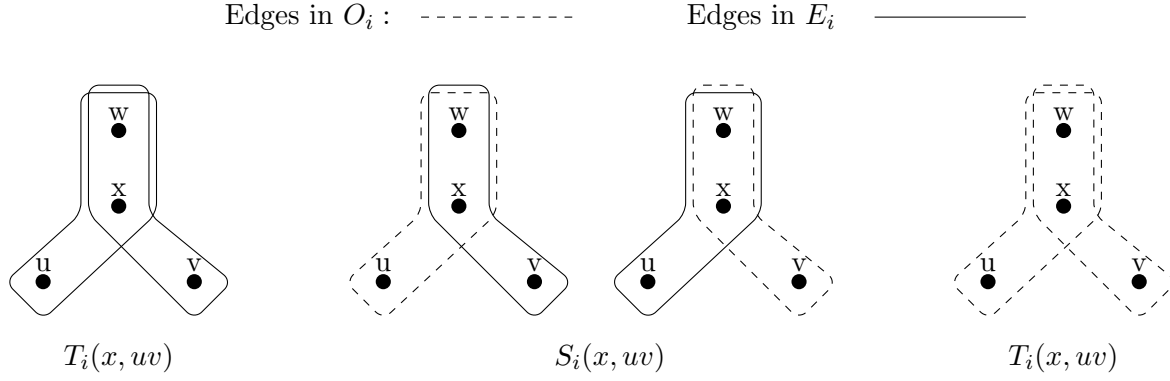


Figure 1:  $T_i(x, uv)$ ,  $S_i(x, uv)$ , and  $R_i(x, uv)$

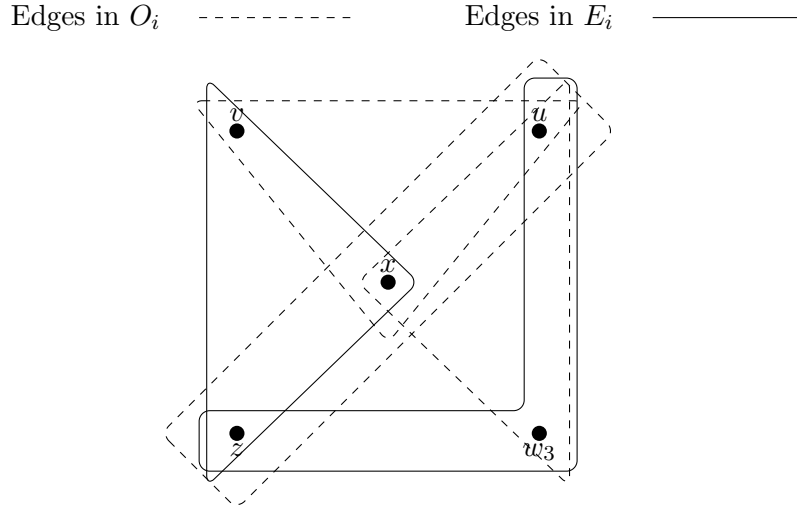


Figure 2:  $U_i(x, u, v, w)$

Given a set  $F$  of edges, and vertices  $u, v$ , let  $N_F(uv) = \{w : uvw \in F\}$  be the set of vertices that form an edge of  $F$  with  $u$  and  $v$ . Also given a vertex  $x$ , two vertex subsets  $A, B \subseteq [N] \setminus \{x\}$  and a set  $F$  of edges, let  $F(A, B, x) := \{xuv \in F : u \in A \text{ and } v \in B\}$ . Define the events

$$\mathcal{N}_i = \{|N_{O_i}(vx)| \leq q_i N \text{ and } |N_{\Gamma_i}(vx)| \leq 2q_{i-1}\sigma\sqrt{N} \text{ for all } x, v \in [N]\} \quad (22)$$

$$\mathcal{P}_i = \{|R_i(x, uv)| \leq q_i^2 N, |S_i(x, uv)| \leq 2q_i\pi_i\sqrt{N}, |T_i(x, uv)| \leq i(\log N)^9 \text{ for all } x, u, v \in [N]\} \quad (23)$$

$$\mathcal{P}_i^+ = \{U_i(x, u, v, w) \leq i(\log N)^9 \text{ for all } x, u, v, w \in [N]\} \quad (24)$$

By (12),  $\Psi^2(x) \sim (1/3)\log x$  as  $x \rightarrow \infty$ . Hence

$$q_I = \Psi'(I\sigma) = e^{-3\Psi^2(I\sigma)} = \Theta((I\sigma)^{-1}) = \Theta(N^{-\beta} \log^2 n).$$

Define constants

$$s := n\sigma^4 q_I^2 = \Theta\left(\frac{nN^{-2\beta}}{\log^4 N}\right) \quad \text{and} \quad \tau_i := 1 - \frac{\delta\pi_i}{2\pi_I}. \quad (25)$$



In particular note that since  $\pi_I \geq \pi_i$  then  $\tau_i \geq 1 - \frac{\delta}{2}$ .

Define the following events:

$$\mathcal{N}_i^+ := \left\{ |N_{\Gamma_i}(vx) \cap A| \leq p|A|(1 + N^{\frac{1}{4}+\beta}) \text{ for all } A \subseteq [N], v \in [N], x \in [N] \right\} \quad (26)$$

$$\mathcal{Q}_i^+ = \left\{ |O_i(A, B, x)| \leq q_i|A||B| \text{ for all disjoint } A, B \in \binom{[N]}{\geq s}, x \in [N] \setminus A \cup B \right\} \quad (27)$$

$$\mathcal{Q}_i := \left\{ \tau_i q_i |A||B| \leq |O_i(A, B, x)| \leq q_i |A||B| \text{ for all disjoint } A, B \in \binom{[N]}{n}, x \in [N] \setminus A \cup B \right\}. \quad (28)$$

We will be interested in the intersections of these events and hence we define

$$\mathcal{G}_i = \mathcal{N}_i \cap \mathcal{P}_i \cap \mathcal{P}^+ \cap \mathcal{Q}_i^+ \cap \mathcal{Q}_i \cap \mathcal{N}_i^+$$

$$\mathcal{G}_{\leq i} = \bigcap_{j=0}^i \mathcal{G}_j.$$

Recall that  $H = H_I$  and  $\rho = \sqrt{\beta(\log N)/3N}$ . Define the event

$$\mathcal{T} := \{e_{H,x}(A, B) = (1 \pm \delta)\rho n^2 \text{ for all disjoint } n\text{-sets } A, B \subseteq V(H), x \in V(H) \setminus (A \cup B)\} \quad (29)$$

We note that heuristically the number of edges we pick for  $e_{H,x}(A, B)$  at each stage is  $pq_i|A||B| = pq_i n^2$  and

$$\sum_{i=1}^I pq_i = \sum_{i=1}^I \frac{q_i \sigma}{\sqrt{N}} \approx \frac{\pi_I}{\sqrt{N}} \approx \sqrt{\frac{\beta \log N}{3N}} = \rho$$

where we used the definition of  $\pi_i$  from (11) for the third relation and the fact that  $\pi_i \approx \Psi(i\sigma) \approx \sqrt{(1/3) \log(\sqrt{3}i\sigma)}$  from Lemma 3 and  $I = \lceil N^\beta \rceil$  for the fourth relation. Thus, ignoring the error parameters, we expect that if  $\mathcal{Q}_i$  holds for all  $i \leq I$ , then  $\mathcal{T}$  holds as well.

The events  $\mathcal{G}_i$  and  $\mathcal{G}_{\leq i}$  are good events for the algorithm at step  $i$  and for all events up to and including step  $i$  respectively. Our main result is that the events  $\mathcal{T} \cap \mathcal{G}_{\leq I}$  hold with high probability, which then means the final  $K_4^-$ -free hypergraph  $H_I$  at the end of the process will have all of the desired properties to prove Theorem 2.

**Lemma 4.** *For  $\delta, \beta, C$  as in the statement of Theorem 2, we have that  $\mathbb{P}(\mathcal{T} \cap \mathcal{G}_{\leq I}) \geq 1 - N^{-\omega(1)}$ .*

In order to prove Lemma 4, we will prove that the probability we are not in the event  $\mathcal{G}_{i+1}$  conditioned on being in the event  $\mathcal{G}_i$  is exponentially small, and the probability of being in the event  $\neg \mathcal{T} \cap \mathcal{G}_{\leq I}$  is also exponentially small, such that when we multiply over the number of steps  $I = N^\beta$  we still get a small probability. More formally, we will prove the following Lemma.

**Lemma 5.** *Under the setup for Theorem 2, for all  $i \in [I - 1]$ ,*

$$\mathbb{P}(\neg \mathcal{G}_{\leq i+1} | \mathcal{G}_{\leq i}) \leq N^{-\omega(1)} \quad \text{and} \quad \mathbb{P}(\neg \mathcal{T} \cap \mathcal{G}_{\leq I}) \leq N^{-\omega(1)}.$$

## 6 Concentration bounds

Throughout the paper we will use the following two standard bounds. The following theorem is a well known version of the standard Chernoff bound.

**Theorem 6** ([11, 14]). *Let  $(X_\alpha)_{\alpha \in J}$  be a finite set of independent  $(0, 1)$  random variables and let  $X = \sum_{\alpha \in J} X_\alpha$ . Then if  $\lambda = \mathbb{E}(X)$*

$$\begin{aligned}\mathbb{P}(X \geq \lambda + t) &\leq \exp\left(-\frac{t^2}{2\lambda}\right) \\ \mathbb{P}(X \leq \lambda - t) &\leq \exp\left(-\frac{t^2}{2(\lambda + t)}\right)\end{aligned}$$

The next theorem of Mcdiarmid [10, 11], and Warnke [14] bounds a function of 0, 1 random variables when the differences between two inputs of the function are bounded when the inputs only differ in one place.

**Theorem 7** ([10, 11, 14]). *Let  $(\xi_\alpha)_{\alpha \in J}$  be a finite set of independent 0, 1 variables. Let  $f : \{0, 1\}^J \rightarrow \mathbb{R}$  be a decreasing function such that there exists  $(c_\alpha)_{\alpha \in J}$  so for all  $z, z' \in \{0, 1\}^J$  with  $z_\beta = z'_\beta$  for all  $\beta \neq \alpha$  we have  $|f(z) - f(z')| \leq c_\alpha$ . Set  $\lambda := \sum_{\alpha \in J} c_\alpha^2 \mathbb{P}(\xi_\alpha = 1)$  and  $X := f((\xi_\alpha)_{\alpha \in J})$ . For all  $t \geq 0$ ,*

$$\mathbb{P}(X \geq \mathbb{E}(X) + t) \leq \exp\left(-\frac{t^2}{2\lambda}\right). \quad (30)$$

Further, if  $C := \max_{\alpha \in J} c_\alpha$  and if we drop the decreasing assumption on  $f$ , then

$$\mathbb{P}(X \leq \mathbb{E}(X) - t) \leq \exp\left(-\frac{t^2}{2(\lambda + Ct)}\right) \quad (31)$$

The following Theorem which follows from Theorem 9 of Warnke [15], bounds the number of events which occur from a family where each has overlap with a limited number of events in the family.

**Theorem 8** (see [15]). *Let  $\{\xi_i\}_{i \in O}$  be a family of independent variables with values in  $\{0, 1\}$ . Let  $(Y_\alpha)_{\alpha \in L}$  be a finite family of variables  $Y_\alpha := \mathbb{1}_{\xi_i=1 \forall i \in \alpha}$  with  $\sum_{\alpha \in L} \mathbb{E}(Y_\alpha) \leq \mu$ . Define the function  $f : 2^L \rightarrow \mathbb{R}$  by  $f(J) = \max_{\beta \in J} |\alpha \in J : \alpha \cap \beta \neq \emptyset|$  for all  $J \subseteq L$ . Let  $Z_C := \max \sum_{\alpha \in J} Y_\alpha$  with the maximum taken over  $J \subseteq L$  with  $f(J) \leq C$ . Then for all  $C, t \geq 0$ ,*

$$\mathbb{P}(Z_C \geq \mu + t) \leq \exp\left(-\frac{t^2}{2C(\mu + t)}\right)$$

Since our algorithm forms our final graph  $H_I$  by removing a maximal set  $D_{i+1}$  of bad sets of edges, we will use the following Theorem of Krivelevich [8] to bound the number of edges removed.

**Theorem 9** ([8]). *Let  $Q$  be a finite set and let  $(\xi_i)_{i \in Q}$  be a set of indicator random variables with  $\mathbb{P}(\xi_i = 1) = p_i$  for all  $i \in Q$ . Let  $(Q(\alpha))_{\alpha \in J}$  be a family of subsets of  $Q$  where  $J$  is a finite set. Let  $X_\alpha = \prod_{i \in \alpha} \xi_i$  and let  $X = \sum_{\alpha \in J} X_\alpha$ . Let*

$$X_0 := \max\{m : \exists \alpha_1, \dots, \alpha_m \in J \text{ with } X_\alpha = 1 \text{ and } Q(\alpha_i) \cap Q(\alpha_j) = \emptyset, i \neq j\}.$$

Then

$$\mathbb{P}(X_0 \geq k) \leq \frac{\mathbb{E}(X)^k}{k!}.$$

## 7 Proof of Lemma 5

We will prove the statement about  $\mathbb{P}(\neg \mathcal{G}_{\leq i+1} | \mathcal{G}_{\leq i})$  by proving each of the statements

$$\begin{aligned}\mathbb{P}(\neg \mathcal{N}_{i+1} | \mathcal{G}_{\leq i}) &\leq N^{-\omega(1)} \\ \mathbb{P}(\neg \mathcal{P}_{i+1} | \mathcal{G}_{\leq i}) &\leq N^{-\omega(1)} \\ \mathbb{P}(\neg \mathcal{P}_{i+1}^+ | \mathcal{G}_{\leq i}) &\leq N^{-\omega(1)} \\ \mathbb{P}(\neg \mathcal{N}_{i+1}^+ | \mathcal{G}_{\leq i}) &\leq N^{-\omega(1)} \\ \mathbb{P}(\neg \mathcal{Q}_{i+1}^+ | \mathcal{G}_{\leq i}) &\leq N^{-\omega(1)} \\ \mathbb{P}(\neg \mathcal{Q}_{i+1} | \mathcal{G}_{\leq i}) &\leq N^{-\omega(1)}\end{aligned}$$

Before proceeding to the proofs, we prove two lemmas that will be used throughout the paper. In the lemmas below, we assume that  $\mathcal{G}_{\leq i}$  holds.

**Lemma 10.** *For all  $e \in O_i$ , we have that  $\mathbb{P}(e \notin C_{i+1}^1 \cup Y_{i+1}) - \frac{q_{i+1}}{q_i} \in [-7\sigma^{3/2}q_i, -5\sigma^{3/2}q_i]$ .*

*Proof.* Let  $e = xuv$  and first notice that since we are in the event  $\mathcal{G}_{\leq i}$  then  $\mathcal{P}_i$  holds and  $|S_i(x, uv)| \leq 2q_i\pi_i\sqrt{N}$ . Therefore,

$$|\hat{S}_i(e)| \leq |S_i(x, uv)| + |S_i(u, xv)| + |S_i(v, xu)| \leq 6q_i\pi_i\sqrt{N},$$

so  $6\sqrt{N}q_i(\pi_i + \sqrt{\sigma}) - |\hat{S}_i(e)| \geq 0$  and  $\hat{p}_{e,i} = 1 - (1-p)^{6\sqrt{N}q_i(\pi_i + \sqrt{\sigma}) - |\hat{S}_i(e)|}$ . Now we get

$$\mathbb{P}(e \notin C_{i+1}^1 \cup Y_{i+1}) = (1-p)^{|\hat{S}_i(e)|} (1 - \hat{p}_{e,i}) = (1-p)^{6\sqrt{N}q_i(\pi_i + \sqrt{\sigma})}.$$

Then using (3) and  $1 - ab \leq (1-a)^b \leq 1 - ab + a^2b^2$  when  $a \in [0, 1]$  and  $b \geq 2$  we get

$$1 - 6\sigma q_i(\pi_i + \sqrt{\sigma}) \leq \mathbb{P}(e \notin C_{i+1}^1 \cup Y_{i+1}) \leq 1 - 6\sigma q_i(\pi_i + \sqrt{\sigma}) + 36\sigma^2 q_i^2 (\pi_i + \sqrt{\sigma})^2. \quad (32)$$

Now since  $|(q_{i+1} - q_i) + 6\sigma q_i^2 \pi_i| \leq 16\sigma^2 q_i^2$  by Lemma 3 (17), we can rearrange to get

$$\left| \frac{q_{i+1}}{q_i} - (1 - 6\sigma q_i \pi_i) \right| \leq 16\sigma^2 q_i.$$

Then (17) together with (32) gives

$$\frac{q_{i+1}}{q_i} - 6\sigma^{3/2}q_i - 16\sigma^2 q_i \leq \mathbb{P}(e \notin C_{i+1}^1 \cup Y_{i+1}) \leq \frac{q_{i+1}}{q_i} - 6\sigma^{3/2}q_i + 16\sigma^2 q_i + 36\sigma^2 q_i^2 (\pi_i + \sqrt{\sigma})^2.$$

Applying Lemma 3 (16) and  $\sigma = \log^{-2} N$ , this completes the proof.  $\blacksquare$

The next lemma is notable and particularly useful as it applies to every subset  $J$  of triples.

**Lemma 11.** *Given  $\mathcal{G}_{\leq i}$ , then for any  $J \subseteq \binom{[N]}{3}$  we have that  $\sum_{e \in O_i} |\hat{S}_i(e) \cap J| \leq 6q_i\pi_i\sqrt{N}|J|$ .*

*Proof.* Note that for all  $e \in O_i$ , if  $f \in \hat{S}_i(e)$  then there is a  $K_4^-$  which contains  $e$  and  $f$  which are both in  $O_i$  and some edge  $g \in E_i$ , which means this same  $K_4^-$  satisfies the requirements for  $e \in \hat{S}_i(f)$ . Thus

$$\sum_{e \in O_i} |\hat{S}_i(e) \cap J| = \sum_{f \in J} \sum_{e \in O_i} \mathbb{1}_{f \in \hat{S}_i(e)} \leq \sum_{f \in J} \sum_{e \in O_i} \mathbb{1}_{e \in \hat{S}_i(f)} = \sum_{f \in J} |\hat{S}_i(f)| \leq |J|(2q_i)\pi_i\sqrt{N},$$

where the last inequality comes from the event  $\mathcal{P}_i$ .  $\blacksquare$

## 7.1 Bound on $\neg \mathcal{N}$

Throughout this section all expectations and probability are conditioned on  $\mathcal{F}_i$  and on the event  $\mathcal{G}_{\leq i}$ , and to improve readability, we will omit that in our notation in this subsection. For each concentration bound, we fix  $v$  and  $x$  till the very end when we take a union bound over all pairs  $v, x$ .

Clearly  $\mathbb{E}(|N_{\Gamma_{i+1}}(vx)|) = p|N_{O_i}(vx)| \leq pq_i N =: \mu$  since we are in the event  $\mathcal{G}_{\leq i}$ , so  $\mathcal{N}_i$  holds. Now by Theorem 6

$$\mathbb{P}(|N_{\Gamma_{i+1}}(vx)| \geq 2\mu) \leq \exp\left(\frac{-\mu^2}{2\mu}\right) \leq \exp\left(-\frac{\sigma\sqrt{N}q_i}{2}\right). \quad (33)$$

Then by Lemma 3 (18) and (33) we obtain  $\mathbb{P}(|N_{\Gamma_{i+1}}(vx)| \geq 2\mu) \leq e^{-\frac{N^c}{3}} < N^{-\omega(1)}$  where the  $c > 0$  comes from (21). Since the number of pairs  $v, x$  is  $\binom{N}{2}$ , a union bound completes the proof.

Next we consider  $|N_{O_i}(vx)|$ . First we will give an upper bound on  $\mathbb{E}(|N_{O_i}(vx)|)$ . Recall from (5) that

$$O_i \setminus O_{i+1} = \Gamma_{i+1} \cup C_{i+1}^1 \cup C_{i+1}^2 \cup Y_{i+1}.$$

Hence  $O_{i+1} \subseteq O_i \setminus (C_{i+1}^1 \cup Y_{i+1})$ , so

$$\mathbb{E}(|N_{O_{i+1}}(vx)|) \leq \sum_{u \in N_{O_i}(vx)} \mathbb{P}(xvu \notin C_{i+1}^1 \cup Y_{i+1}) =: X. \quad (34)$$

Recall that  $X = \sum_{u \in N_{O_i}(vx)} \mathbb{P}(xvu \notin C_{i+1}^1 \cup Y_{i+1})$ . Since we are in  $\mathcal{G}_{\leq i}$ ,  $|N_{O_i}(vx)| \leq q_i N$  by  $\mathcal{N}_i$ , and therefore by Lemma 10

$$\mathbb{E}(X) \leq |N_{O_i}(vx)| \cdot \mathbb{P}(xvu \notin C_{i+1}^1 \cup Y_{i+1}) \leq q_i N \left( \frac{q_{i+1}}{q_i} - \sigma^{3/2} q_i \right) = Nq_{i+1} - \sigma^{3/2} q_i^2 N. \quad (35)$$

Next, we will prove concentration around the expected value by using Theorem 7. The index set  $J$  from Theorem 7 will be made of two parts,  $J = J_\Gamma \cup J_Y$  where both  $J_\Gamma$  and  $J_Y$  are copies of  $O_i$ . Then  $\xi_\alpha$  from Theorem 7 will be defined as follows:

$$\xi_\alpha = \begin{cases} \mathbb{1}_{\alpha \in \Gamma_{i+1}} & \text{if } \alpha \in J_\Gamma, \\ \mathbb{1}_{\alpha \in Y_{i+1}} & \text{if } \alpha \in J_Y. \end{cases} \quad (36)$$

Then  $f$  in Theorem 7 will be defined so that  $X$  from Theorem 7 is the same as  $X$  from (34), and notice that  $f$  is decreasing since the presence of edges in  $\Gamma_{i+1}$  or  $Y_{i+1}$  only adds edges to  $C_{i+1}^1 \cup Y_{i+1}$ . We need to give bounds on  $\Delta_e$  which is the absolute value of the change in the size of  $(C_{i+1}^1 \cup Y_{i+1}) \cap (N_{O_i}(vx) \times \{v, x\})$  when we change whether or not an edge  $e$  is in  $\Gamma_{i+1}$ . Assume for now that  $\Delta_e \leq c_e$ . Similarly, let  $\Upsilon_e$  be the absolute value of the change in  $|(C_{i+1}^1 \cup Y_{i+1}) \cap (N_{O_i}(vx) \times \{v, x\})|$  when we change whether or not an edge  $e$  is in  $Y_{i+1}$  and assume that  $\Upsilon_e \leq c'_e$ .

Observe that changing whether or not  $e$  is in  $\Gamma_{i+1}$  only affects which edges are in  $C_{i+1}^1$ , and similarly changing whether an edge is in  $Y_{i+1}$  will only affect which edges are in  $Y_{i+1}$ . Consequently, when considering  $c_e$  we only need to consider the effect on  $C_{i+1}^1$ , and when considering  $c'_e$  we only need to consider the effect on  $Y_{i+1}$ .

First, to give a bound on  $\Delta_e$ , notice that  $\Delta_e$  by definition is the number of changes to

$$C_{i+1}^1 \cap (N_{O_i}(vx) \times \{v, x\}) = \{xvw \in O_i : \hat{S}_i(xvw) \cap \Gamma_{i+1} \neq \emptyset\}.$$

Since we only changed whether or not  $e$  was in  $\Gamma_{i+1}$ ,  $e$  can only affect elements of  $C_{i+1}^1 \cap (N_{O_i}(vx) \times \{v, x\})$  when  $e \in \hat{S}_i(vxw)$ . Further, if  $e \in \hat{S}_i(vxw)$  then  $vwx \in \hat{S}_i(e)$ . Indeed,  $e \in \hat{S}_i(vxw)$  implies  $e \in O_i$  and that there exists  $e' \in E_i$  with  $\{vwx\}ee'$  forming a  $K_4^-$ . Then  $e \in O_i$  and  $e' \in E_i$  so  $e\{vwx\}e'$  is a  $K_4^-$  which is the definition for  $vwx \in \hat{S}_i(e)$ . So the number of  $w$  such that  $vwx$  is in  $C_{i+1}^1 \cap (N_{O_i}(vx) \times \{v, x\})$  is at most  $|\hat{S}_i(e) \cap (\{vx\} \times N_{O_i}(vx))|$ . Thus for  $e = u_1u_2u_3$ ,

$$\Delta_e \leq |\hat{S}_i(e) \cap (N_{O_i}(vx) \times \{v, x\})| \leq |\hat{S}_i(e)| \leq 6q_i\pi_i\sqrt{N} \quad (37)$$

due to the bound  $\max\{|\hat{S}_i(u_1, u_2u_3)|, |S_i(u_2, u_1u_3)|, |S_i(u_3, u_1u_2)|\} \leq 2q_i\pi_i\sqrt{N}$  as we are assuming the event  $\mathcal{P}_i$ . Further,  $\sum_{e \in O_i} |\hat{S}_i(e) \cap (N_{O_i}(vx) \times \{v, x\})| \leq 6q_i\pi_i\sqrt{N}|N_{O_i}(vx)|$  by Lemma 11. Moreover, by (37) and Lemma 11,

$$\sum_{e \in O_i} \Delta_e^2 \leq 6q_i\pi_i\sqrt{N} \sum_{e \in O_i} \Delta_e \leq 6q_i\pi_i\sqrt{N} \sum_{e \in O_i} |\hat{S}_i(e) \cap (N_{O_i}(vx) \times \{v, x\})| \leq 36q_i^2\pi_i^2N(q_iN) \quad (38)$$

where the bound on  $|N_{O_i}(vx) \times \{v, x\}|$  in the last inequality comes from the event  $\mathcal{N}_i$ .

Next, we want to bound  $\Upsilon_e$ , but notice that since all edges are in  $Y_{i+1}$  independently, changing whether or not  $e$  is in  $Y_{i+1}$  can only affect whether the edge  $e$  itself (and no other edge) is in  $(C_{i+1}^1 \cup Y_{i+1}) \cap (N_{O_i}(vx) \times \{v, x\})$ . Therefore  $\Upsilon_e \leq 1$  and if  $\Upsilon_e = 1$ , then  $e \in N_{O_i}(vx) \times \{v, x\}$ . Thus by  $\mathcal{N}_i$ ,

$$\sum_{e \in O_i} \Upsilon_e^2 \leq |N_{O_i}(vx) \times \{v, x\}| = |N_{O_i}(vx)| \leq q_iN. \quad (39)$$

Now we can bound the value of  $\lambda$  from Theorem 7 using the fact that  $\hat{p}_{e,i} \leq q_i$  from Lemma 3 (19) and the bounds (38) and (39) by

$$\lambda := \sum_{e \in O_i} p\Delta_e^2 + \sum_{e \in O_i} \hat{p}_{e,i}\Upsilon_e^2 \leq p(36q_i^3\pi_i^2N^2) + q_i^2N \leq 37\sigma q_i^2N^{3/2}$$

where the final inequality followed from Lemma 3 (16). Finally, we apply Theorem 7 with  $\lambda$  and (35) to obtain

$$\mathbb{P}(|N_{O_{i+1}}(vx)| \geq q_{i+1}N) \leq \exp\left(-\frac{\sigma^3 q_i^4 N^2}{74\sigma q_i^2 N^{3/2}}\right) = \exp\left(-\left(\frac{1}{74}\right)\sigma^2 q_i^2 \sqrt{N}\right) \leq N^{-\omega(1)}$$

where the final inequality holds from Lemma 3 (18). Now taking a union bound over the  $\binom{N}{2}$  choices for  $v$  and  $x$  completes the proof.

## 7.2 Bound on $\neg\mathcal{P}$

Throughout this section we will again omit the conditioning on  $\mathcal{F}_i$  and the event  $\mathcal{G}_{\leq i}$  in our notation.

We will start with the proof for  $T_i(x, uv)$ . Clearly

$$|T_{i+1}(x, uv)| - |T_i(x, uv)| = \sum_{w \in R_i(x, uv)} \mathbb{1}_{\{xuw, xvw \in \Gamma_{i+1}\}} + \sum_{w \in S_i(x, uv)} \mathbb{1}_{\{xuw \in \Gamma_{i+1} \text{ or } xvw \in \Gamma_{i+1}\}} \quad (40)$$

Notice that in the second term of (40) for every  $w$  in  $S_i(x, uv)$  one of  $xuw$  or  $xvw$  is already in  $E_i$  so only one can potentially be in  $\Gamma_{i+1}$ . Now since for edges in  $O_i$  the event of being in  $\Gamma_{i+1}$  are independent Bernoulli random variables with probability  $p$ , we can bound the expected change in  $T_i(x, uv)$  using the bounds on  $|R_i(x, uv)|$  and  $|S_i(x, uv)|$  from (23) as follows.

$$\mathbb{E}(|T_{i+1}(x, uv)| - |T_i(x, uv)|) \leq (p^2)q_i^2 N + (p)2q_i\pi_i\sqrt{N} = \sigma^2 q_i^2 + 2q_i\pi_i\sigma \leq \sigma^2 + 2\sqrt{\sigma} \ll 1$$

where in the last two inequalities we used  $\sqrt{\sigma}\pi_i \leq 1$  and  $q_i \leq 1$  and  $\sigma \ll 1$  from Lemma 3 (13) and (15).

Now using the bound on  $T_i(x, uv)$  from (23),

$$|T_{i+1}(x, uv)| = |\Delta T_{i+1}(x, uv)| + |T_i(x, uv)| \leq |\Delta T_{i+1}(x, uv)| + i \log^9 N$$

Thus Theorem 6 gives

$$\begin{aligned} \mathbb{P}(|T_{i+1}(x, uv)| \geq (i+1) \log^9 N) &\leq \mathbb{P}(|\Delta T_i(x, uv)| \geq \log^9 N) \\ &\leq \exp\left(-\frac{(\log^9 N - \sigma^2 q_i^2 - 2q_i\pi_i\sigma)^2}{2(\sigma^2 q_i^2 + 2q_i\pi_i\sigma)}\right) \\ &\leq \exp(-\log^{17} N) \\ &\leq N^{-\omega(1)} \end{aligned}$$

where the third inequality comes from (16). Thus taking a union bound over all  $3\binom{N}{3}$  choices for the vertex  $x$  and the vertices  $uv$  completes the proof.

Next, we will prove the bound on  $|R_i(x, uv)|$ . Since  $O_{i+1} \subseteq O_i \setminus (C_{i+1}^1 \cup Y_{i+1})$  we get

$$|R_{i+1}(x, uv)| \leq \sum_{w \in R_i(x, uv)} \mathbb{1}_{\{xuw, xvw \notin (C_{i+1}^1 \cup Y_{i+1})\}} =: X \quad (41)$$

Observe that

$$\begin{aligned} \mathbb{P}(xuw, xvw \notin C_{i+1}^1 \cup Y_{i+1}) &\leq (1 - \hat{p}_{xuw,i})(1 - \hat{p}_{xvw,i})(1 - p)^{|\hat{S}_i(x, uw)| + |\hat{S}_i(x, vw)| - |\hat{S}_i(x, uw) \cap \hat{S}_i(x, vw)|} \\ &= \mathbb{P}(xuw \notin C_{i+1}^1 \cup Y_{i+1}) \mathbb{P}(xvw \notin C_{i+1}^1 \cup Y_{i+1}) (1 - p)^{-|\hat{S}_i(x, uw) \cap \hat{S}_i(x, vw)|} \end{aligned}$$

where the first inequality comes from excluding some edges in  $\hat{S}_i(u, xw)$ ,  $\hat{S}_i(w, ux)$ ,  $\hat{S}_i(v, vx)$ , and  $\hat{S}_i(w, vx)$  which could also close edges.

Now if  $e \in |\hat{S}_i(x, uw) \cap \hat{S}_i(x, vw)|$  then since  $u \neq v$ ,  $e$  must be of the form  $xwz$  for some  $z \in [N] \setminus \{x, u, v, w\}$ , and  $xwz \in O_i$  and  $xuz, xvw \in E_i$ . This implies  $z \in T_i(x, uv)$ , so since  $\mathcal{P}_i$  holds the number of  $z$  satisfying these conditions is bounded above by  $i \log^9 N$  by (23). Thus

$$|\hat{S}_i(x, uw) \cap \hat{S}_i(x, vw)| \leq i \log^9 N. \quad (42)$$

Combining (41), (42), (23), and Lemma 10, we bound the expected value of  $X$  by

$$\begin{aligned} \mathbb{E}(X) &\leq (q_i^2 N) \left( \frac{q_{i+1}}{q_i} - 5\sigma^{3/2} q_i \right)^2 (1 - p)^{-i \log^9 N} \\ &\leq \left( q_{i+1}^2 N - 5q_{i+1}q_i^2 \sigma^{3/2} N + 25q_i^4 \sigma^3 N \right) (1 + pI \log^9 N) \\ &\leq q_{i+1}^2 N - 5q_i^3 (1 - 12\sigma) \sigma^{3/2} N + 25q_i^4 \sigma^3 N + q_i^2 N pI \log^9 N + 25q_i^4 \sigma^3 N pI \log^9 N \\ &\leq q_{i+1}^2 N - 4.9q_i^3 \sigma^{3/2} N + 25q_i^4 \sigma^3 N + \sqrt{N} (q_i^2 \sigma I \log^9 N + 25q_i^4 \sigma^4 I \log^9 N) \\ &\leq q_{i+1}^2 N - 4q_i^3 \sigma^{3/2} N \end{aligned}$$

where the third inequality uses  $q_{i+1} \geq q_i(1 - 12\sigma)$  from Lemma 3 (20) and  $q_{i+1} \leq q_i$ , and the last inequality uses  $q_i^3 \sigma^{3/2} \gg q_i^4 \sigma^3$  and  $q_i^3 \sigma^{3/2} N \gg \sqrt{N} q_i^2 \sigma I \log^9 N$  by Lemma 3 (18).

To prove concentration, we will apply Theorem 7. We define the index set  $J = J_\Gamma \cup J_Y$  and define  $\xi_\alpha$  as in (36). Then  $f$  will be defined so that  $X = f((\xi_\alpha)_{\alpha \in J})$ . We again let  $\Delta_e$  be the absolute value change of  $X$  when we change whether an edge  $e$  is in  $\Gamma_{i+1}$ , and let  $\Upsilon_e$  be the absolute value change of  $X$  when we change whether an edge  $e$  is in  $Y_{i+1}$ . Then notice that if changing whether  $e \in \Gamma_{i+1}$  changes whether  $xuw$  or  $xvw$  is in  $C_{i+1}^1$ , for some  $w \in R_i(x, uv)$ , there exists  $e' \in E_i$  so  $ee'\{xuw\}$  or  $ee'\{xvw\}$  is a  $K_4^-$  and then either  $\{xuw\}$  or  $\{xvw\}$  is in  $\hat{S}_i(e)$ . Thus  $\Delta_e \leq |\hat{S}_i(e) \cap (\{xu, xv\} \times R_i(x, uv))|$  and  $\Upsilon_e$  is only 1 when  $e = xuw$  or  $e = xv w$  for  $w \in R_i(x, uv)$  and 0 otherwise since changes edges being in  $Y_{i+1}$  doesn't impact other edges. Then we can bound  $\lambda$  in Theorem 7 by

$$\begin{aligned} \lambda &= \sum_{e \in O_i} p \Delta_e^2 + \sum_{e \in O_i} \hat{p}_{e,i} \Upsilon_e^2 \\ &\leq p(6q_i \pi_i \sqrt{N}) \sum_{e \in O_i} |\hat{S}_i(e) \cap (\{xu, xv\} \times R_i(x, uv))| + q_i(2q_i^2 N) \\ &\leq \frac{\sigma}{\sqrt{N}} (6q_i \pi_i \sqrt{N})^2 (2q_i^2 N) + 2q_i^3 N \\ &\leq 100\sigma q_i^3 N^{3/2} \end{aligned}$$

where the second inequality uses Lemma 11 together with the fact that  $\{xu, xv\} \times R_i(x, uv) \leq 2|R_i(x, uv)|$  and the last inequality uses Lemma 3 (16). Now Theorem 7 gives

$$\begin{aligned} \mathbb{P}(|R_{i+1}(x, uv)| \geq q_{i+1}^2 N) &\leq \mathbb{P}(X \geq \mathbb{E}(X) + 4q_i^3 \sigma^{3/2} N) \\ &\leq \exp\left(-\frac{16q_i^6 \sigma^3 N^2}{200\sigma q_i^3 N^{3/2}}\right) \\ &= \exp\left(-\frac{4}{25} q_i^3 \sigma^2 \sqrt{N}\right) \leq N^{-\omega(1)} \end{aligned}$$

where the last inequality follows from Lemma 3. Taking a union bound over the  $3\binom{N}{3}$  choices for  $x$  and  $uv$  completes the proof for all of the  $R_i(x, uv)$  variables.

Lastly we will prove the bounds for the variables  $S_{i+1}(x, uv)$ . Notice there are two types of edges which can be included in  $\hat{S}_{i+1}(x, uv)$ . The first type are edges which were in  $\hat{S}_i(x, uv)$  and did not leave  $\hat{S}_i(x, uv)$  during step  $i + 1$ . The second type are edges  $xuw$  where  $w \in R_i(x, uv)$  and  $xvw \in \Gamma_{i+1}$  or edges  $xvw$  where  $w \in R_i(x, uv)$  and  $xuw \in \Gamma_{i+1}$ .

We will prove bounds on the number of edges in  $\hat{S}_i(x, uv)$  of each of these types separately and then combine the bounds to prove that the required upper bound for  $S_i(x, uv)$  holds whp. To this end, define random variables

$$S_i^1(x, uv) := \sum_{e \in \hat{S}_i(x, uv)} \mathbb{1}_{e \notin C_{i+1}^1 \cup Y_{i+1}}$$

$$S_i^2(x, uv) := \sum_{w \in R_i(x, uv)} \mathbb{1}_{xuw \in \Gamma_{i+1} \text{ or } xv w \in \Gamma_{i+1}}$$

$$|S_{i+1}(x, uv)| \leq S_{i+1}^1(x, uv) + S_{i+1}^2(x, uv) \quad (43)$$

where inequality in (43) comes from the fact that edges can leave  $\hat{S}_i(x, uv)$  if they are in the sets  $\Gamma_{i+1}$  and  $C_{i+1}^2$ . Because we are in the event  $\mathcal{P}_i$ , (23) implies  $|S_i(x, uv)| \leq 2q_i\pi_i\sqrt{N}$  and  $|R_i(x, uv)| \leq q_i^2N$ . Together with Lemma 10 we obtain

$$\begin{aligned} \mathbb{E}(S_{i+1}^1(x, uv) + S_{i+1}^2(x, uv)) &\leq |S_i(x, uv)| \left( \frac{q_{i+1}}{q_i} - 5\sigma^{3/2}q_i \right) + 2p|R_i(x, uv)| \\ &\leq 2q_{i+1}\pi_i\sqrt{N} - 10q_i^2\pi_i\sigma^{3/2}\sqrt{N} + 2\sigma\sqrt{N}q_i^2 \\ &\leq 2q_{i+1}\pi_i\sqrt{N} - 10q_i^2\pi_i\sigma^{3/2}\sqrt{N} + 2\sigma\sqrt{N}q_i(q_{i+1} + 12\sigma q_i) \\ &= 2q_{i+1}\sqrt{N}(\pi_i + \sigma q_i) - 10q_i^2\pi_i\sigma^{3/2}\sqrt{N} + 24\sigma^2q_i^2\sqrt{N} \\ &\leq 2q_{i+1}\pi_{i+1}\sqrt{N} - 9q_i^2\pi_i\sigma^{3/2}\sqrt{N} \end{aligned}$$

where the third inequality follows from (20) and the last inequality follows from the fact that  $\pi_{i+1} = \pi_i + \sigma q_{i+1}$  by (11) and  $\sigma^2 \ll \pi_i\sigma^{3/2}$  by Lemma 3 (12) and (14). Now if we can prove that whp the values of  $S_{i+1}^1(x, uv)$  and  $S_{i+1}^2(x, uv)$  are both not more than  $q_i^2\pi_i\sigma^2\sqrt{N}$  larger than their respective expected values, we will have proven the bound on  $|S_i(x, uv)|$ .

Now to prove concentration for  $S_i^1(x, uv)$ , we will bound  $\lambda$  in Theorem 7. Let  $\Delta_e$  be the absolute value of the change in  $S_i^1(x, uv)$  when we change whether  $e$  is in  $\Gamma_{i+1}$  and let  $\Upsilon_e$  be the absolute value change in  $S_i^1(x, uv)$  when we change whether  $e$  is in  $Y_{i+1}$ . Similar to the proofs of concentration for  $|N_{O_i}(vx)|$  and  $|R_i(x, uv)|$ ,  $e$  only affects whether an edge  $f$  is in  $S_{i+1}(x, uv)$  or not if  $e$  and  $f$  are in a  $K_4^-$  together with one other edge in  $E_i$ , which would then mean  $f \in \hat{S}_i(e)$ . Thus if we let  $e = w_1w_2w_3$

$$\Delta_e \leq |\hat{S}_i(e) \cap \hat{S}_i(x, uv)| \leq |\hat{S}_i(e)| \leq |S_i(w_1, w_2w_3)| + |S_i(w_2, w_1w_3)| + |S_i(w_3, w_1w_2)| \leq 6q_i\pi_i\sqrt{N}.$$

However, we can obtain a much better bound on  $|\hat{S}_i(e) \cap \hat{S}_i(x, uv)|$  than  $|\hat{S}_i(e)|$  when  $e \neq xuv$ . If  $w_1w_2w_3 \cap xuv = \emptyset$  then  $|\hat{S}_i(e) \cap \hat{S}_i(x, uv)| = 0$  since any triple which is in a  $K_4^-$  with  $e$  must contain two of the vertices  $w_1, w_2, w_3$  but then this triple cannot contain both  $x$  and one of  $v$  or  $u$ . If  $|w_1w_2w_3 \cap xuv| = 1$ , then  $|\hat{S}_i(e) \cap \hat{S}_i(x, uv)| \leq 4$  since if a triple contains both two of  $w_1, w_2, w_3$  and two of  $x, u, v$ , it must contain the unique overlapping vertex and then one of the remaining vertices from each of  $w_1w_2w_3$  and  $xuv$ . Finally, if  $|w_1w_2w_3 \cap xuv| = 2$  then if  $f \in \hat{S}_i(e) \cap \hat{S}_i(x, uv)$  it either includes one of the shared vertices between  $w_1w_2w_3$  and  $xuv$  and both of the vertices in  $w_1w_2w_3$  and  $xuv$  which are not shared, or it contains both of the shared vertices and a third vertex. Thus there are only two choices for  $f$  which contain only one of the shared vertices. When  $f$  contains both of the shared vertices, let the last vertex of  $f$  be  $z$ . Notice that  $f$  must contain  $x$  by the definition of  $\hat{S}_i(x, uv)$  and similarly  $f$  must contain exactly one of  $u, v$ .

**Case 1.** Suppose that  $f = xuz$ . Without loss of generality let  $w_1 = x$  and  $w_2 = u$  and  $f = w_1w_2z$ . Recall that  $e = w_1w_2w_3$ . Now since  $f \in \hat{S}_i(e)$  there must be an edge in  $E_i$  out of edges comprised of the vertices  $x, u, w_3, z$  and  $xuz, xuw_3 \in O_i$ . Thus we have the following two subcases.

**Subcase 1a.**  $xw_3z \in E_i$ . Since  $xvz, xw_3z \in E_i$  we have  $z \in T_i(x, vw_3)$ .

**Subcase 1b.**  $uw_3z \in E_i$ . Since  $xuz, xuv, xuw_3 \in O_i$  and  $xvz, uw_3z \in E_i$  then  $z \in U_i(x, u, v, w_3)$  (See Figure 3 below).



**Case 2.** Suppose that  $f = x v z$ . As in Case 1, we obtain  $z \in T_i(x, u w_3)$  or  $z \in U_i(x, v, u, w_3)$ .

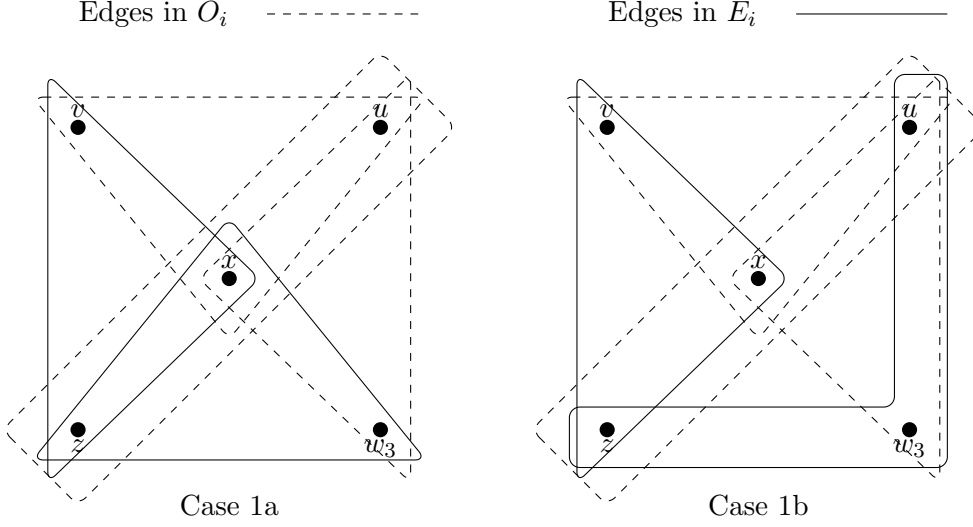


Figure 3:

This means the number of  $f \in \hat{S}_i(e) \cap \hat{S}_i(x, uv)$  is less than  $4i(\log N)^9$ . Putting this together we get that if  $e \neq xuv$  then  $\Delta_e \leq \max\{4i(\log N)^9, 4\} \leq I\sigma^{-5}$ . Hence applying Lemma 11 with  $J = \hat{S}_i(x, uv)$ ,

$$p \sum_{e \in O_i} \Delta_e^2 \leq p((6q_i\pi_i\sqrt{N})^2 + I\sigma^{-5}(6q_i\pi_i\sqrt{N})|S_i(x, uv)|) \leq 13I\sigma^{-4}\sqrt{N}q_i^2\pi_i^2$$

Further,  $\Upsilon_e \leq 1$  so  $\hat{p}_{e,i} \sum_{e \in O_i} \Upsilon_e^2 \leq q_i \sum_{e \in O_i} \Upsilon_e^2 \leq q_i |S_{x,uv}(i)| \leq 2q_i^2\pi_i\sqrt{N}$ . Then we can bound  $\lambda$  in Theorem 7 by

$$\lambda \leq p \sum_{e \in O_i} \Delta_e^2 + q_i \sum_{e \in O_i} \Upsilon_e^2 \leq 15I\sigma^{-4}\sqrt{N}q_i^2\pi_i^2.$$

Theorem 7 gives

$$\mathbb{P}(S_i^1(x, uv) \geq \mathbb{E}(S_i^1(x, uv) + q_i^2\pi_i\sigma^2\sqrt{N}) \leq \exp\left(\frac{-q_i^4\pi_i^2\sigma^4N}{30I\sigma^{-4}\sqrt{N}q_i^2\pi_i^2}\right) = \exp\left(\frac{-q_i^2\sigma^8\sqrt{N}}{30I}\right) \leq N^{-\omega(1)} \quad (44)$$

where the last inequality comes from (18).

Next, to prove concentration for  $S_i^2(x, uv)$ , notice that for all  $w \in R_i(x, uv)$  the inclusion of the edges  $xuw$  or  $xvw$  in  $\Gamma_{i+1}$  are independent identically distributed Bernoulli random variables. Also,  $\mathbb{E}(S_i^2(x, uv)) \leq 2pq_i^2N = 2\sigma q_i^2\sqrt{N}$  and  $\sigma\pi_i \leq 1$  by Lemma 3 (15). We use Theorem 6 to obtain

$$\mathbb{P}(S_i^2(x, uv) \leq \mathbb{E}(S_i^2(x, uv)) + \sigma^2 q_i^2 \pi_i \sqrt{N}) \leq \exp\left(\frac{-(\sigma^2 q_i^2 \pi_i \sqrt{N})^2}{2(2\sigma q_i^2 \sqrt{N})}\right) = \exp\left(\frac{-\sigma^3 q_i^2 \pi_i^2 \sqrt{N}}{4}\right) \leq N^{-\omega(1)} \quad (45)$$

where the last inequality holds by Lemma 3. Now taking a union bound over all choices for  $x$  and  $uv$  and combining (44) and (45) completes the proof that  $\mathbb{P}(\neg \mathcal{P}_{i+1} | \mathcal{G}_{\leq i}) \leq n^{-\omega(1)}$ .

### 7.3 Bound on $\neg\mathcal{P}^+$

We will omit the conditioning on  $\mathcal{F}_i$  and the event  $\mathcal{G}_{\leq i}$  in the notation in this section. For  $z$  to be in  $U_{i+1}(x, u, v, w)$  but not in  $U_i(x, u, v, w)$ , at least one of the edges  $xvz, uzw$  must have been added to  $\Gamma_{i+1}$  in this round (see Figure 2). In other words there are three cases for how  $xvz$  and  $uwz$  are in  $E_{i+1}$

1.  $xvz \in \Gamma_{i+1}$  and  $uwz \in E_i$
2.  $xvz \in E_i$  and  $uwz \in \Gamma_{i+1}$
3.  $xvz \in \Gamma_{i+1}$  and  $uwz \in \Gamma_{i+1}$ .

Further, for  $z$  to be in  $U_{i+1}$  all three of  $xuv, xuw$ , and  $xuz$  must be in  $O_{i+1}$ . In particular, in the first and third cases where  $xvz \in \Gamma_{i+1}$ ,  $z \in R_i(x, uv)$ , and in the second and third cases where  $uwz \in \Gamma_{i+1}$ ,  $z \in R_i(u, xw)$ . Further, in the first case where  $uwz \in E_i$ ,  $z \in S_i(u, xw)$  and similarly in the second case where  $xvz \in E_i$   $z \in S_i(x, uv)$ .

Then defining

$$\begin{aligned} U_1 &:= \sum_{z \in R_i(x, uv) \cap S_i(u, xw)} \mathbb{1}_{xvz \in \Gamma_{i+1}} \leq \sum_{z \in S_i(u, xw)} \mathbb{1}_{xvz \in \Gamma_{i+1}} =: U_1^+ \\ U_2 &:= \sum_{z \in S_i(x, uv) \cap R_i(u, xw)} \mathbb{1}_{uwz \in \Gamma_{i+1}} \leq \sum_{z \in S_i(x, uv)} \mathbb{1}_{uwz \in \Gamma_{i+1}} =: U_2^+ \\ U_3 &:= \sum_{z \in R_i(x, uv) \cap R_i(u, xw)} \mathbb{1}_{xvz, uwz \in \Gamma_{i+1}} \leq \sum_{z \in R_i(x, uv)} \mathbb{1}_{xvz, uwz \in \Gamma_{i+1}} =: U_3^+ \end{aligned}$$

we obtain

$$|U_{i+1}(x, u, v, w) \setminus U_i(x, u, v, w)| \leq U_1 + U_2 + U_3 \leq U_1^+ + U_2^+ + U_3^+ =: X.$$

Using (24) yields

$$|U_{i+1}(x, u, v, w)| \leq i(\log N)^9 + X. \quad (46)$$

We first upper bound the expected value of  $X$  using (23) and (46) as

$$\mathbb{E}(X) \leq 4pq_i\pi_i\sqrt{N} + p^2q_i^2N. \quad (47)$$

Since  $X$  is a sum of independent  $(0, 1)$  random variables, by Theorem 6 and (46) and (47)

$$\begin{aligned} \mathbb{P}(|U_{i+1}(x, u, v, w)| \geq (i+1)\log^9 N) &\leq \mathbb{P}(X \geq \log^9 N) \\ &\leq \exp\left(-\frac{(\log^9 N - 4pq_i\pi_i\sqrt{N} - p^2q_i^2N)^2}{2(4pq_i\pi_i\sqrt{N} + p^2q_i^2N)}\right) \\ &\leq \exp\left(-\frac{(\log^9 N - 4\sigma q_i\pi_i - \sigma^2q_i^2)^2}{8\sigma q_i\pi_i + \sigma^2q_i^2}\right) \\ &\leq \exp(-\log^{17} N) \\ &\leq N^{-\omega(1)}. \end{aligned}$$

## 7.4 Proof of $\mathcal{N}^+$

The random variable  $|N_{\Gamma_{i+1}}(vx) \cap A|$  is a sum of  $|N_{O_i}(vx) \cap A|$  Bernoulli random variables each with probability  $p$ . Thus,

$$\mathbb{E}(|N_{\Gamma_{i+1}}(vx) \cap A|) = p|N_{O_i}(vx) \cap A| \leq p|A|$$

Applying Theorem 6 gives

$$\begin{aligned} \mathbb{P}(|N_{\Gamma_{i+1}}(vx) \cap A| \geq p|A|(1 + N^{\frac{1}{4}+\beta})) &\leq \exp\left(-\frac{(p|A|N^{\frac{1}{4}+\beta})^2}{2p|A|}\right) \\ &= \exp\left(-\frac{1}{2}p|A|N^{\frac{1}{2}+2\beta}\right) \\ &= \exp\left(-\frac{1}{2}\sigma|A|N^{2\beta}\right). \end{aligned}$$

By a union bound over all choices of  $v, x$  and  $A$ , we obtain

$$\begin{aligned} \mathbb{P}(\exists v, x, A : |N_{\Gamma_{i+1}}(vx) \cap A| \geq p|A|(1 + N^{\frac{1}{4}+\beta})) &< N^{2+A} \exp\left(-\frac{1}{2}\sigma N^{2\beta}|A|\right) \\ &= \exp\left((2 + |A|) \log n - \frac{1}{2}\sigma N^{2\beta}|A|\right) \\ &< N^{-\omega(1)}. \end{aligned}$$

## 7.5 Proof of $\mathcal{Q}^+$

Recalling that  $s = n\sigma^4 q_I^2 = C\sqrt{N \log N} \sigma^4 q_I^2$ , we begin with the case where  $|A| = |B| \geq s$ . Since  $O_{i+1} \subseteq O_i \setminus (C_{i+1}^1 \cup Y_{i+1})$

$$|O_{i+1}(A, B, x)| \leq |O_i(A, B, x) \setminus (C_{i+1}^1 \cup Y_{i+1})| = \sum_{f \in O_i(A, B, x)} \mathbb{1}_{f \notin C_{i+1}^1 \cup Y_{i+1}} =: X \quad (48)$$

We will first bound  $\mathbb{E}(X)$  and then we will use Theorem 7 to prove concentration. Using  $\mathcal{Q}_i^+$  and Lemma 10

$$\begin{aligned} \mathbb{E}(X) &= |O_i(A, B, x)| \mathbb{P}(f \notin C_{i+1}^1 \cup Y_{i+1}) \\ &\leq q_i |A| |B| \left(\frac{q_{i+1}}{q_i} - 5\sigma^{3/2} q_i\right) \\ &\leq q_{i+1} |A| |B| - 5q_i^2 \sigma^{3/2} |A| |B|. \end{aligned}$$

To bound the probability that  $X \geq \mathbb{E}(X) + 5q_i^2 \sigma^{3/2} |A| |B|$ , we will use Lemma 7. For every  $e \in O_i$ , let  $\Delta_e$  be the absolute change in  $X$  when we change whether  $e$  is in  $\Gamma_{i+1}$ . Similarly let  $\Upsilon_e$  be the absolute change in  $X$  when we change whether  $e$  is in  $Y_{i+1}$ .

First, to bound  $\Delta_e$ , let  $e = w_1 w_2 w_3$  and in this section assume  $a \in A$  and  $b \in B$ . If  $|\{w_1, w_2, w_3\} \cap (A \cup B \cup \{x\})| \leq 1$ , then any  $K_4^-$  which contains  $e$  does not contain an edge of the form  $xab$ , so

$\Delta_e = 0$ . Next, if  $x \notin \{w_1, w_2, w_3\}$  and  $|\{w_1, w_2, w_3\} \cap (A \cup B)| \geq 2$ , then if an edge of the form  $xab$  is in a  $K_4^-$  together with  $e$ , the vertices of this  $K_4^-$  must be  $\{w_1, w_2, w_3, x\}$  so  $\Delta_e \leq 3$ .

Finally, without loss of generality let  $w_1 = x$  and assume  $|\{w_2, w_3\} \cap (A \cup B)| \geq 1$ . If  $xab$  is an edge which is in  $C_{i+1}^1$  only when  $e \in \Gamma_{i+1}$ , and without loss of generality  $w_2 = b$  then either the edge  $xw_3a$  or  $w_2w_3a$  must be in  $E_i$ . The number of these edges can be bounded by  $\sum_{j=1}^i (|N_{\Gamma_j}(xw_3) \cap A| + |N_{\Gamma_j}(w_2w_3) \cap A|)$ . Because  $\mathcal{N}_i^+$  holds, this is no more than  $20p|A|N^{\frac{1}{4}+\beta}I$ . Thus

$$\Delta_e \leq 40p|A|N^{\frac{1}{4}+\beta}I \quad (49)$$

where the extra factor of 2 comes from the possibility that both  $w_2$  and  $w_3$  are in  $A \cup B$ . Further, since changing whether  $e \in \Gamma_{i+1}$  only changes if  $xab$  is open if  $xab \in \hat{S}_i(e)$ ,  $\Delta_e \leq |\hat{S}_i(e) \cap O_i(A, B, x)|$ . Then using (49) together with Lemma 11, we obtain

$$p \sum_{e \in O_i} c_e^2 \leq 40p^2|A|N^{\frac{1}{4}+\beta}I(6q_i\pi_i\sqrt{N}(q_i|A|^2)) \leq 240\sigma^2N^{2\beta-\frac{1}{4}}q_i^2\pi_i|A|^3.$$

Further  $\Upsilon_e \leq \mathbb{1}_{e \in O_i(A, B, x)}$  so

$$q_i \sum_{e \in O_i} \hat{c}_e^2 \leq q_i^2|A|^2 \leq q_i^2|A|^3 \left( \frac{1}{C\sigma^4q_i^2} N^{-\frac{1}{2}} \log^{-\frac{1}{2}} N \right) \leq \sigma^4N^{2\beta-\frac{1}{4}}q_i^2\pi_i|A|^3$$

where the second inequality comes from  $|A| \geq s$ . The last inequality holds, with room to spare, due to the power of  $N$  which is  $-1/2 + 2\beta + o(1)$  on the left and  $-1/4 + 2\beta$  on the right. Hence

$$\lambda \leq p \sum_{e \in O_i} c_e^2 + q_i \sum_{e \in O_i} \hat{c}_e^2 \leq 250\sigma^2N^{2\beta-\frac{1}{4}}q_i^2\pi_i|A|^3 \quad (50)$$

Applying Theorem 7 gives

$$\begin{aligned} \mathbb{P}(|O_{i+1}(A, B, x)| \geq q_{i+1}|A|^2) &\leq \mathbb{P}(X \geq \mathbb{E}(X) + q_i^2\sigma^{3/2}|A|^2) \\ &\leq \exp\left(-\frac{q_i^4\sigma^3|A|^4}{500\sigma^2N^{2\beta-\frac{1}{4}}q_i^2\pi_i|A|^3}\right) \\ &\leq \exp\left(-\frac{1}{500}q_i^3\sigma|A|N^{\frac{1}{4}-2\beta}\right) \\ &\leq \exp\left(-\frac{1}{500}|A|N^{\frac{1}{4}-6\beta}\right) \end{aligned}$$

where the third inequality uses (16) and the fourth inequality uses (18). By the union bound over all choices of  $x, A$  we obtain

$$\begin{aligned} \mathbb{P}(\exists x, A, B, |A| = |B| : |O_i(A, B, x)| \geq q_{i+1}|A||B|) &\leq N^{2|A|+1} \exp\left(-\frac{1}{500}|A|N^{\frac{1}{4}-6\beta}\right) \\ &\leq \exp\left((2|A|+1)\log N - \frac{1}{500}|A|N^{\frac{1}{4}-6\beta}\right) \\ &\leq N^{-\omega(1)} \end{aligned}$$

which proves the result when  $|A| = |B|$ .

Now assume  $|A| > |B| \geq s$ . Then

$$|O_i(A, B, x)| = \frac{\sum_{A' \subset A, |A'|=|B|} |O_i(A', B, x)|}{\binom{|A|-1}{|B|-1}}.$$

Now applying the fact that  $\mathcal{Q}_i^+$  holds for  $|A| = |B|$  with probability at least  $1 - N^{-\omega(1)}$  we obtain

$$|O_i(A, B, x)| \leq \frac{\binom{|A|}{|B|} q_i |B|^2}{\binom{|A|-1}{|B|-1}} = q_i |A| |B|.$$

with probability at least  $1 - N^{-\omega(1)}$ .

## 7.6 Proof of $\mathcal{Q}$

Recall we are to show that if  $A, B \in \binom{[N]}{n}$  are disjoint and  $x \notin A \cup B$ , then  $\tau_{i+1} q_{i+1} n^2 \leq |O_i(A, B, x)| \leq q_{i+1} |A| |B|$ . We begin with the following Lemma.

**Lemma 12.** *Let  $\mathcal{Q}_{A,B,x}$  be the event that the following hold where  $x \in [N]$  and  $A, B \in \binom{[N]}{n}$ :*

$$\begin{aligned} X_1 &:= |O_i(A, B, x) \setminus (C_{i+1}^1 \cup Y_{i+1})| \in \left[ |O_i(A, B, x)| \left( \frac{q_{i+1}}{q_i} - 8\sigma^{3/2} q_i \right), |O_i(A, B, x)| \frac{q_{i+1}}{q_i} \right] \\ X_2 &:= |O_i(A, B, x) \cap C_{i+1}^2| \leq |O_i(A, B, x)| (6\sigma^2 q_i) + 100\sigma q_i^2 \sqrt{N} n \\ X_3 &:= |O_i(A, B, x) \cap \Gamma_{i+1}| \leq |O_i(A, B, x)| (2\sigma^2 q_i). \end{aligned}$$

Then  $\mathbb{P}(\neg \mathcal{Q}_{A,B,x} \cap \mathcal{N}_{i+1} \cap \mathcal{P}_{i+1}) \leq N^{-\omega(n)}$  for all choices of  $A, B, x$ .

Before proving Lemma 12, we will use Lemma 12 to prove  $\mathbb{P}(\neg \mathcal{Q}_{i+1} \cap \mathcal{N}_{i+1} \cap \mathcal{P}_{i+1}) \leq N^{-\omega(1)}$ . Assume for all  $x, A, B$ , the event  $\mathcal{Q}_{A,B,x}$  holds. Let us show that  $\mathcal{Q}_{i+1}$  holds.

Let  $A, B \in \binom{[N]}{n}$  and  $x \in [N] \setminus (A \cup B)$ . Recall that an edge in  $O_i(A, B, x)$  remains in  $O_{i+1}(A, B, x)$  if it is not in  $C_{i+1}^1 \cup Y_{i+1} \cup C_{i+1}^2 \cup \Gamma_{i+1}$ . Therefore

$$X_1 - X_2 - X_3 \leq |O_{i+1}(A, B, x)| \leq X_1.$$

For the upper bound, notice that since  $\mathcal{G}_{\leq i}$  holds,  $|O_i(A, B, x)| \leq q_i |A| |B|$  by (28), so

$$|O_{i+1}(A, B, x)| \leq X_1 \leq |O_i(A, B, x)| \frac{q_{i+1}}{q_i} \leq q_{i+1} |A| |B|.$$

For the lower bound, recall that  $|A| = |B| = n = C\sqrt{N \log N}$  and  $\sigma = 1/\log^2 N$  and  $\tau_i \leq 1$ . Then

$$\begin{aligned}
X_1 - X_2 - X_3 &\geq |O_i(A, B, x)| \left( \frac{q_{i+1}}{q_i} - 8\sigma^{3/2}q_i - 8\sigma^2q_i \right) - 100\sigma q_i^2 \sqrt{N}n \\
&\geq (\tau_i q_i n^2) \left( \frac{q_{i+1}}{q_i} - 16\sigma^{3/2}q_i \right) - 100\sigma q_i^2 \sqrt{N}n \\
&\geq n^2 \left( \tau_i q_{i+1} - 16\sigma^{3/2}q_i^2 - \frac{100\sigma q_i^2}{C\sqrt{\log N}} \right) \\
&= n^2 \left( \tau_i q_{i+1} - 16\sigma^{3/2}q_i^2 - \frac{100\sigma^{5/4}q_i^2}{C} \right) \\
&\geq q_{i+1}n^2 \left( \tau_i - \left( \frac{200\sigma^{5/4}q_i}{C} \right) \left( \frac{q_i}{q_{i+1}} \right) \right)
\end{aligned}$$

where in the second inequality we used (28) and in the equality we used  $\sigma^{1/4} = \frac{1}{\sqrt{\log N}}$ . Now recalling that  $\tau_i = 1 - \frac{\delta\pi_i}{2\pi_I}$  and therefore  $\tau_i - \tau_{i+1} = \frac{\delta\sigma q_i}{2\pi_I}$  together with  $\sqrt{\frac{\log(\sqrt{3}x)}{3}} - \frac{1}{\sqrt{3}} \leq \Psi(x) \leq \sqrt{\frac{\log(\sqrt{3}x)}{3}} + \frac{1}{\sqrt{3}}$ ,  $|\pi_i - \Psi(i\sigma)| \leq 2\sigma$ , and  $I = \lceil N^\beta \rceil$  from Lemma 3 (12) and (14) gives

$$\begin{aligned}
\tau_{i+1} &= \tau_i - \frac{\delta\sigma q_i}{2\pi_I} \\
&\leq \tau_i - \frac{\delta\sigma q_i}{2 \left( \sqrt{\frac{\log(\sqrt{3}(I\sigma))}{3}} + \left( \frac{1}{\sqrt{3}} + 2\sigma \right) \right)} \\
&\leq \tau_i - \frac{\delta\sigma q_i}{2.1 \sqrt{\frac{0.5 \log 3 + \beta \log N + \log \sigma}{3}}} \\
&\leq \tau_i - \frac{\sqrt{3}\delta\sigma q_i}{(2.2)\sqrt{\beta \log N}} \\
&= \tau_i - \frac{\sqrt{3}\delta q_i \sigma^{5/4}}{(2.2)\sqrt{\beta}}.
\end{aligned}$$

Now using  $q_i/(q_{i+1}) \leq 12\sigma + 1 \leq 2$  from Lemma 3 (20) and choosing constant  $C$  so that  $400/C \leq \frac{\sqrt{3}\delta}{(2.2)\sqrt{\beta}}$  gives  $X_1 - X_2 - X_3 \geq q_{i+1}n^2\tau_{i+1}$  as desired.

We have shown that  $\neg \mathcal{Q}_{i+1} \cap \mathcal{N}_{i+1} \cap \mathcal{P}_{i+1} \subset \cup_{x,A,B} \neg \mathcal{Q}_{A,B,x} \cap \mathcal{N}_{i+1} \cap \mathcal{P}_{i+1}$ . Finally, applying Lemma 12 and union bounding over all choices of  $x, A, B$  gives

$$\begin{aligned}
\mathbb{P}(\neg \mathcal{Q}_{i+1} \cap \mathcal{N}_{i+1} \cap \mathcal{P}_{i+1}) &\leq \mathbb{P}(\exists x, A, B : \neg \mathcal{Q}_{A,B,x} \cap \mathcal{N}_{i+1} \cap \mathcal{P}_{i+1}) \\
&\leq N^{1+2n-\omega(n)} \\
&\leq N^{-\omega(n)}.
\end{aligned}$$

Now all that is needed to prove  $\mathcal{Q}_{i+1}$  holds with high probability is to prove Lemma 12.

## 7.7 Proof of Lemma 12

### 7.7.1 $X_1$

First we prove the bound on  $X_1$ . The upper bound follows from the proof for the event  $\mathcal{Q}_i^+$  since  $X_1$  is exactly the  $X$  from (48) and  $n > s$ . For the lower bound, we apply Lemma 10 to show

$$\mathbb{E}(X_1) = \sum_{e \in O_i A, B, x} \mathbb{P}(e \notin C_{i+1, A, B, x}^1 \cup Y_{i+1}) \geq |O_i(A, B, x)| \left( \frac{q_{i+1}}{q_i} - 7\sigma^{3/2} q_i \right)$$

Now using the same value for  $\lambda$  from (50) and  $C = 40p|A|N^{\frac{1}{4}+\beta}I$  which follows from (49) in the proof of  $\mathcal{Q}^+$ , we can apply the second half of Theorem 7 to show

$$\mathbb{P}\left(X_1 \leq |O_i(A, B, x)| \left( \frac{q_{i+1}}{q_i} - 8\sigma^{3/2} q_i \right)\right) \leq \mathbb{P}(X_1 \leq \mathbb{E}(X_1) - \sigma^{3/2} q_i |O_i(A, B, x)|) \leq \exp(-D)$$

where

$$D = \frac{(\sigma^{3/2} q_i |O_i(A, B, x)|)^2}{2((250\sigma^2 N^{2\beta-\frac{1}{4}} q_i^2 \pi_i |A|^3) + (40p|A|N^{\frac{1}{4}+\beta} I)(\sigma^{3/2} q_i |O_i(A, B, x)|))}.$$

Note that

$$\begin{aligned} D &\geq \frac{\sigma^3 \tau_i^2 q_i^4 n^4}{500\sigma^2 N^{2\beta-\frac{1}{4}} q_i^2 \pi_i n^3 + 40\sigma^{5/2} N^{-\frac{1}{4}+2\beta} n^3 q_i^2} \\ &\geq n N^{\frac{1}{4}-10\beta} \end{aligned}$$

where the last inequality uses (16) and (18) from Lemma 3. Since  $e^{-D} = N^{-\omega(n)}$  we are done.

### 7.7.2 $X_2$

We now prove the bound on  $X_2$ . While one might expect that this would be relatively easy to show as it seems unlikely that two edges are picked in step  $i$  which then form a  $K_4^-$  with an edge of  $O_i(A, b, x)$ , this is the most difficult part of the proof.

We will break up  $X_2$  into four components and handle each one separately. First we introduce some definitions and notation. The following auxiliary variables depend on the step of the process  $i$ , a distinguished vertex  $x$ , and two vertex subsets  $A, B \subseteq [N] \setminus \{x\}$  with  $|A| = |B|$ .

$$z_{A,B}(i) := \sigma^4 q_i^2 |A|$$

$$W(A, B, x, i) := \{w \in [N] : |N_{\Gamma_{i+1}}(xw) \cap (A \cup B)| \geq z_{A,B}(i)\}$$

$$\hat{C}_{i+1, A, B, x}^2 := \{xuv : \exists w \notin W(A, B, x, i) : xuw, xvw \in \Gamma_{i+1}\}$$

Whenever the context for  $A$ ,  $B$ ,  $x$ , and  $i$  are clear, we will write  $z$  for  $z_{A,B}(i)$  and  $W$  for  $W(A, B, x, i)$ . We write  $a$  (resp.  $b$ ) for a generic vertex in  $A$  (resp.  $B$ ).

$$\begin{aligned}\hat{X}_2 &:= |\{xab \in O_i(A, B, x) \cap C_{i+1}^2 : \exists w \notin W(A, B, x) \text{ and } |\{xwa, xwb, wab\} \cap \Gamma_{i+1}| \geq 2\}| \\ X_2^{AB} &:= |\{xab \in O_i(A, B, x) \cap C_{i+1}^2 : \exists w \in W(A, B, x) \text{ and } xwa, xwb \in \Gamma_{i+1}\}| \\ X_2^A &:= |\{xab \in O_i(A, B, x) \cap C_{i+1}^2 : \exists w \in W(A, B, x) \text{ and } xwa, wab \in \Gamma_{i+1}\}| \\ X_2^B &:= |\{xab \in O_i(A, B, x) \cap C_{i+1}^2 : \exists w \in W(A, B, x) \text{ and } xwb, wab \in \Gamma_{i+1}\}|.\end{aligned}$$

Since for  $abx \in O_i(A, B, x) \cap C_{i+1}^2$  there must be some  $w$  with  $|\{xwa, xwb, wab\} \cap \Gamma_{i+1}| \geq 2$ , we deduce that  $X_2 \leq \hat{X}_2 + X_2^{AB} + X_2^A + X_2^B$ .

With the aim of applying Theorem 8, we can upper bound  $\hat{X}_2$  with

$$\hat{X}_2 = \sum_{e \in O_i(A, B, x)} \mathbb{1}_{e \in \hat{C}_{i+1, A, B, x}^2} \leq \sum_{xab \in O_i(A, B, x)} \sum_{w \in V \setminus W(A, B, x, i)} \mathbb{1}_{|\{axw, bxw, wab\} \cap \Gamma_{i+1}| \geq 2} =: \hat{X}_2^+.$$

Now let

$$\begin{aligned}L_{AB} &:= \{\{xwa, xwb\} \subseteq O_i : xab \in O_i, w \notin \{x, a, b\}, a \in A, b \in B\} \\ L_A &:= \{\{xwa, wab\} \subseteq O_i : xab \in O_i, w \notin \{x, a, b\}, a \in A, b \in B\} \\ L_B &:= \{\{xwb, wab\} \subseteq O_i : xab \in O_i, w \notin \{x, a, b\}, a \in A, b \in B\} \\ L &:= L_{AB} \cup L_A \cup L_B.\end{aligned}$$

Now with Theorem 8 in mind, for each  $e \in O_i := O$ , let  $\xi_e$  be the indicator random variable for  $e \in \Gamma_{i+1}$ . For  $\alpha \in L$ , let  $Y_\alpha = \mathbb{1}_{\xi_e=1 \forall e \in \alpha} = \mathbb{1}_{\alpha \subseteq \Gamma_{i+1}}$ . In order to bound  $\sum_{\alpha \in L} \mathbb{E}(Y_\alpha)$ , note that if  $\{xwa, xwb\} \in L_{AB}$  then  $w \in R_i(x, ab)$ . This shows that

$$|L_{AB}| \leq |O_i(A, B, x)| |R_i(x, ab)|,$$

since  $xab \in O_i(A, B, x)$  and  $w \in R_i(x, ab)$  means that  $xwa$  and  $xwb$  are open.

Similarly, if  $\{xwa, wab\} \in L_A$ , then  $w \in R_i(a, xb)$  and if  $\{xwb, wab\} \in L_B$  then  $w \in R_i(b, xa)$ . Thus

$$|L| \leq |O_i(A, B, x)| (|R_i(x, ab)| + |R_i(a, xb)| + |R_i(b, xa)|).$$

Since  $|R_i(u_1, u_2 u_3)| \leq q_i^2 N$  by (23) due to  $\mathcal{P}_i$ , and  $q_i \leq 1$ , we can give the upper bound

$$\sum_{\alpha \in L} \mathbb{E}(Y_\alpha) \leq p^2 |O_i(A, B, x)| (3 \max\{|R_i(x, ab)|\}) \leq 3\sigma^2 q_i |O_i(A, B, x)| =: \mu.$$

Now we define  $K \subseteq L$  as follows:

$$\begin{aligned}K_{AB} &:= \{\{xwa, xwb\} \in L_{AB} : w \notin W(A, B, x, i) \text{ and } xwa, xwb \in \Gamma_{i+1}\} \\ K_A &:= \{\{xwa, wab\} \in L_A : w \notin W(A, B, x, i) \text{ and } xwa, wab \in \Gamma_{i+1}\} \\ K_B &:= \{\{xwb, wab\} \in L_B : w \notin W(A, B, x, i) \text{ and } xwb, wab \in \Gamma_{i+1}\} \\ K &:= K_{AB} \cup K_A \cup K_B.\end{aligned}$$



Note that  $\hat{X}_2^+ = \sum_{\alpha \in K} Y_\alpha$ . Next, recalling from Theorem 8 that  $f(J) := \max_{\beta \in J} |\alpha \in J : \alpha \cap \beta \neq \emptyset|$ , we claim that  $f(K) \leq 8z$ . Then since in Theorem 8,  $Z_{8z}$  is defined as the maximum over  $J \subseteq L$  with  $f(J) \leq 8z$  of  $\sum_{\alpha \in J} Y_\alpha$ , if  $f(K) \leq 8z$  then  $\hat{X}_2^+ \leq Z_{8z}$ . Finally, we apply Theorem 8 to conclude using  $t = \mu$  that

$$\mathbb{P}(\hat{X}_2 \geq 6\sigma^2 q_i |O_i(A, B, x)|) \leq \mathbb{P}(Z_{8z} \geq 2\mu) \leq \exp(-\mu^2/(32z\mu)) = \exp(-\mu/(32z)).$$

We now provide the details of these assertions.

We now show that  $f(K) \leq 8z$ . Let  $\beta \in K$  and let  $e \in \beta$ . First, if  $x \notin e$  notice that there are at most 3 elements  $\alpha \in K$  such that  $e \in \alpha$ . Further, if  $x \in e$ , then the number of elements  $\alpha \in K$  such that  $e \in \alpha$  is bounded by  $|N_{\Gamma_{i+1}}(xw) \cap (A \cup B)|$ , and since for all  $\beta \in K$ ,  $w \notin W(A, B, x, i)$ , we obtain  $|N_{\Gamma_{i+1}}(xw) \cap (A \cup B)| < z$ . Thus for all  $\beta \in K$

$$|\alpha \in K : \alpha \cap \beta \neq \emptyset| \leq \sum_{e \in \beta} |\alpha \in K : e \in \alpha| \leq \sum_{e \in \beta} 3 + \sum_{v \in e \setminus W(A, B, x, i)} |N_{\Gamma_{i+1}}(xv) \cap (A \cup B)| \leq 8z.$$

Now using Theorem 8 since  $\hat{X}_2^+ \leq \sum_{\alpha \in K} Y_\alpha \leq Z_{8z}$ , set  $\mu = 3\sigma^2 q_i |O_i(A, B, x)|$ , and  $C = 8z$ . Using the fact that  $|O_i(A, B, x)| \geq \tau_i q_i n^2$  by (28) since we are in  $\mathcal{G}_{\leq i}$ , we obtain

$$\begin{aligned} \mathbb{P}(\hat{X}_2 \geq 6\sigma^2 q_i |O_i(A, B, x)|) &\leq \mathbb{P}(Z_{8z} \geq 2\mu) \\ &\leq \exp\left(-\frac{\mu^2}{32\mu z}\right) \\ &\leq \exp\left(-\frac{3\tau_i q_i^2 \sigma^2 n^2}{32\sigma^4 q_i^2 n}\right) \leq N^{-\omega(n)}. \end{aligned}$$

Next we prove the bound for  $X_2^{AB}$ . By definition

$$X_2^{AB} \leq \sum_{w \in W(A, B, x, i)} |O_i(N_{\Gamma_{i+1}}(wx) \cap A, N_{\Gamma_{i+1}}(wx) \cap B, x)|.$$

If

$$\min\{|N_{\Gamma_{i+1}}(wx) \cap A|, |N_{\Gamma_{i+1}}(wx) \cap B|\} \geq z = \sigma^4 q_i^2 |A| \geq n\sigma^2 q_i^2 = s,$$

then we use the fact that  $\mathcal{Q}_i^+$  holds and obtain

$$O_i(N_{\Gamma_{i+1}}(wx) \cap A, N_{\Gamma_{i+1}}(wx) \cap B, x) \leq q_i |N_{\Gamma_{i+1}}(wx) \cap A| |N_{\Gamma_{i+1}}(wx) \cap B|.$$

Otherwise one of these terms is less than  $z$ . Writing

$$M := \max\{|N_{\Gamma_{i+1}}(wx) \cap A|, |N_{\Gamma_{i+1}}(wx) \cap B|\},$$

this gives

$$|O_i(N_{\Gamma_{i+1}}(wx \cap A), N_{\Gamma_{i+1}}(wx) \cap B, x)| \leq zM.$$

Putting both bounds together we obtain

$$\begin{aligned} |O_i(N_{\Gamma_{i+1}}(wx \cap A), N_{\Gamma_{i+1}}(wx) \cap B, x)| &\leq q_i |N_{\Gamma_{i+1}}(wx) \cap A| |N_{\Gamma_{i+1}}(wx) \cap B| + zM \\ &\leq (q_i |N_{\Gamma_{i+1}}(wx)| + z) |N_{E_i \cup \Gamma_{i+1}}(wx) \cap (A \cup B)|. \end{aligned} \tag{51}$$

We will now use the following Lemma from [6] which bounds the sum of the sizes of a collection of sets whose intersection is bounded. We will provide a proof as well for completeness.

**Lemma 13.** *Let  $(U_i)_{i \in I}$  be a family of subsets  $U_i \subseteq U$ , and suppose  $|U_i| \geq z \geq \sqrt{4|U|y}$  and  $|U_i \cap U_j| \leq y$  for  $i \neq j$ . Then  $|I| \leq \frac{2|U|}{z}$  and  $\sum_{i \in I} |U_i| \leq 2|U|$ .*

*Proof.* Suppose for contradiction  $|I| > \frac{2|U|}{z}$  and let  $J \subseteq I$  with  $|J| = \lfloor \frac{2|U|}{z} \rfloor + 1$ . Note that  $y \leq \frac{z^2}{4|U|}$  and  $\frac{z}{2} \leq \frac{|U_i|}{2}$  for all  $i \in I$ . Then for any  $i \in J$

$$\sum_{j \in J, j \neq i} |U_i \cap U_j| \leq (|J| - 1)y \leq \frac{2|U|y}{z} \leq \frac{z}{2} \leq \frac{|U_i|}{2}.$$

Then

$$|U| \geq |\cup_{i \in J} U_i| \geq \sum_{i \in J} \left( |U_i| - \sum_{j \in J, j \neq i} |U_i \cap U_j| \right) \geq \sum_{i \in J} \frac{|U_i|}{2} \geq |J| \frac{z}{2} = \left( \lfloor \frac{2|U|}{z} \rfloor + 1 \right) \frac{z}{2} > |U|$$

which is a contradiction. Hence  $|I| \leq \frac{2|U|}{z}$ . Now for any  $i \in I$

$$\sum_{j \in I, j \neq i} |U_i \cap U_j| \leq (|I| - 1)y \leq \frac{2|U|y}{z} \leq \frac{z}{2} \leq \frac{|U_i|}{2}$$

as before and

$$|U| \geq |\cup_{i \in I} U_i| \geq \sum_{i \in I} \left( |U_i| - \sum_{j \in I, j \neq i} |U_i \cap U_j| \right) \geq \sum_{i \in I} \frac{|U_i|}{2}. \quad (52)$$

■

We will let  $W(A, B, x, i) = I$ ,  $A \cup B = U$ ,  $U_w = N_{E_i \cup \Gamma_{i+1}}(wx) \cap (A \cup B)$  in Lemma 13. As  $\mathcal{P}_{i+1}$  holds, for all  $w_1, w_2 \in W$ ,

$$\begin{aligned} |(N_{E_i \cup \Gamma_{i+1}}(w_1x) \cap (A \cup B)) \cap (N_{E_i \cup \Gamma_{i+1}}(w_2x) \cap (A \cup B))| &\leq |N_{E_{i+1}}(w_1x) \cap N_{E_{i+1}}(w_2x)| \\ &\leq T_{i+1}(x, w_1w_2) \\ &\leq I(\log N)^9. \end{aligned}$$

Let  $y := I(\log N)^9$  and note that  $z = \sigma^4 q_i^2 n \geq \sqrt{4(2n)(I(\log N)^9)}$ . Also since for all  $w \in W(A, B, x, i)$  we have  $|N_{\Gamma_{i+1}}(wx) \cap (A \cup B)| \geq z$ , then  $|U_w| \geq z$ . Thus Lemma 13 gives

$$\sum_{w \in W(A, B, x, i)} |N_{E_i \cup \Gamma_{i+1}}(wx) \cap (A \cup B)| \leq 4n.$$

Since  $\mathcal{N}_{i+1}$  holds, for all  $w \in W(A, B, x, i)$ , we have  $|N_{\Gamma_{i+1}}(wx)| \leq 2q_i \sigma \sqrt{N}$ . Also,  $z = \sigma^4 q_i^2 n \ll 2q_i^2 \sigma \sqrt{N}$ , so

$$q_i |N_{\Gamma_{i+1}}(wx)| + z \leq 3q_i^2 \sigma \sqrt{N}.$$

Putting (51) and Lemma 13 together, we obtain

$$X^{AB} \leq \sum_{w \in W(A, B, x, i)} (q_i |N_{\Gamma_{i+1}}(wx)| + z) |N_{E_i \cup \Gamma_{i+1}}(wx) \cap (A \cup B)| \leq (3q_i^2 \sigma \sqrt{N})(4n).$$

The last bound we need to prove to bound  $X_2$  is a bound on  $X^A$  and  $X^B$ . We will show the bound on  $X^A$  and the argument for  $X^B$  is symmetric. Since  $\mathcal{N}^+$  holds,  $|N_{\Gamma_{i+1}}(aw) \cap B| \leq 10p|B|(N^{\frac{1}{4}+\beta})$  for all  $a$  and  $w$ . Thus we can bound  $X_2^A$  by

$$\begin{aligned} X_2^A &\leq \sum_{w \in W(A, B, x, i)} \sum_{a \in N_{\Gamma_{i+1}}(xw) \cap A} |N_{\Gamma_{i+1}}(aw) \cap B| \\ &\leq \left(10pnN^{\frac{1}{4}+\beta}\right) \sum_{w \in W(A, B, x, i)} |N_{\Gamma_{i+1}}(wx) \cap A| \\ &\leq 10C\sigma\sqrt{\log N}N^{\frac{1}{4}+\beta} \sum_{w \in W(A, B, x, i)} |N_{E_{i+1}}(wx) \cap (A \cup B)|. \end{aligned}$$

Now we apply Lemma 13 with  $I = W(A, B, x, i)$ ,  $U = A \cup B$ , and  $U_w = N_{E_{i+1}}(wx) \cap (A \cup B)$ . Again,  $|U_w| \geq z$  since  $w \in W(A, B, x, i)$  and  $\Gamma_{i+1} \subseteq E_{i+1}$ , and  $|U_{w_1} \cap U_{w_2}| \leq |T_{i+1}(x, w_1 w_2)| \leq I(\log N)^9 =: y$ . Also,  $z \geq \sqrt{4ny}$ , so Lemma 13 yields

$$\sum_{w \in W(A, B, x, i)} |N_{E_{i+1}}(wx) \cap (A \cup B)| \leq 4n.$$

Altogether,

$$X_2^A \leq 10C\sigma\sqrt{\log N}N^{\frac{1}{4}+\beta}(4n).$$

Therefore, whp,

$$\begin{aligned} X_2 &\leq \hat{X}_2 + X_2^{AB} + X_2^A + X_2^B \\ &\leq 6\sigma^2 q_i |O_i(A, B, x)| + 3q_i^2 \sigma \sqrt{N}(4n) + 2(10C\sigma\sqrt{\log N})N^{\frac{1}{4}+\beta}(4n) \\ &\leq 6\sigma^2 q_i |O_i(A, B, x)| + 100^2 \sqrt{N}n. \end{aligned}$$

### 7.7.3 $X_3$

Finally, we prove the bound on  $X_3$ . Since  $X_3$  is a sum of iid Bernoulli random variables each with probability  $p$ ,

$$\mathbb{E}(X_3) = p|O_i(A, B, x)| = \frac{\sigma}{\sqrt{N}}|O_i(A, B, x)| \ll \sigma^2 q_i |O_i(A, B, x)| =: t,$$

where the last inequality comes from Lemma 3 (18). Applying Theorem 6 together with the bound  $|O_i(A, B, x)| \geq \tau_i q_i |A||B|$  since  $\mathcal{G}_{\leq i}$  holds gives

$$\mathbb{P}(X_3 > |O_i(A, B, x)|(2\sigma^2 q_i)) \leq \mathbb{P}(X_3 \geq \mathbb{E}(X_3) + t) \leq \exp\left(-\frac{t^2}{4t}\right) \leq \exp\left(-\frac{\sigma^2 q_i^2 \tau_i |A||B|}{4}\right).$$

By Lemma 3 (18),  $|A| = |B| = C\sqrt{N \log N}$  and  $\tau_i \geq 1 - \frac{\delta}{2}$ , we have  $\mathbb{P}(X_3 > |O_i(A, B, x)|2\sigma^2 q_i) \leq N^{-\omega(n)}$ .

## 8 Bound on $\neg\mathcal{T}$

To complete the proof of Lemma 5, recall that  $B_{i+1}$  is a collection of bad subsets of  $\Gamma_{i+1}$  and  $H_{i+1}$  is comprised of the edges selected in  $\Gamma_{i+1}$  which remain after removing the edges from  $D_{i+1}$ , which is a maximal subset of  $B_{i+1}$ . Rather than conditioning on  $\mathcal{G}_{\leq i}$  as we did in previous sections, here we fix some  $x \in [N]$  and disjoint sets  $A$  and  $B$  with  $|A| = |B| = n$  and we consider the number of edges of the form  $xab$  added over the entire process where  $a \in A$  and  $b \in B$ . We break this quantity up into those added into  $E_i$  and those which remain in  $H_i$  by defining

$$X(A, B, x) := \sum_{i=0}^{I-1} |O_i(A, B, x) \cap \Gamma_{i+1}|$$

$$Y(A, B, x) := \sum_{i=0}^{I-1} |O_i(A, B, x) \cap (\cup_{S \in D_{i+1}} \cup_{e \in S} e)|.$$

We now define  $\mu^+$  and  $\mu^-$  in order to bound  $X(A, B, x)$  and  $Y(A, B, x)$  as

$$\mu^+ := p \sum_{i=0}^{I-1} q_i n^2$$

$$\mu^- := p \sum_{i=0}^{I-1} \tau_i q_i n^2.$$

Since  $H$  only includes edges picked in  $E_i$  which were not in  $D_i$  for some  $i$

$$X(A, B, x) - Y(A, B, x) \leq e_{H,x}(A, B) \leq X(A, B, x). \quad (53)$$

Thus if we prove

$$X(A, B, x) \in \left[ \left(1 - \frac{\delta}{2}\right) \mu^-, \left(1 + \frac{\delta}{2}\right) \mu^+ \right] \quad (54)$$

$$Y(A, B, x) \leq \frac{\delta^2 \mu^-}{9} \quad (55)$$

then (53), (54), and (55) and  $0 < \delta < 1$  will show

$$\left(1 - \frac{\delta}{2} - \frac{\delta^2}{9}\right) p \sum_{i=0}^{I-1} \tau_i q_i n^2 \leq e_{H,x}(A, B) \leq \left(1 + \frac{\delta}{2}\right) p \sum_{i=0}^{I-1} q_i n^2. \quad (56)$$

Observe that (11) implies

$$p \sum_{i=0}^{I-1} q_i n^2 = \frac{n^2}{\sqrt{N}} \sum_{i=0}^{|I|-1} \sigma q_i = \frac{n^2}{\sqrt{N}} (\pi_I - \sigma).$$

By (12) and (14)

$$\pi_I \in \left[ \sqrt{\frac{\log(\sqrt{3}) + \beta \log N}{3}} - \frac{1}{\sqrt{3}} + \sigma, \sqrt{\frac{\log(\sqrt{3}) + \beta \log N}{3}} + \frac{1}{\sqrt{3}} + 2\sigma \right]. \quad (57)$$

Since  $\rho = \sqrt{\frac{\beta \log N}{3N}}$  we obtain

$$e_{H,x}(A, B) \leq \left(1 + \frac{\delta}{2}\right) p \sum_{i=0}^{I-1} q_i n^2 \leq \left(1 + \frac{\delta}{2}\right) \frac{n^2}{\sqrt{N}} (\pi_I - \sigma) < (1 + \delta) \rho n^2.$$

Similarly since  $\tau_I = 1 - \frac{\delta}{2}$  and  $\mu^- \geq \tau_I \mu^+$  then (56) gives  $e_{H,x}(A, B) \geq (1 - \delta + \frac{5\delta}{36}) \mu^+$ . Then by a similar argument to the upper bound  $e_{H,x}(A, B) \geq (1 - \delta) \rho n^2$ . Hence, showing (54) and (55) hold for all  $x, A, B$  with probability  $1 - N^{-\omega(1)}$  will complete the proof of Lemma 5.

To bound  $X(A, B, x)$ , we define

$$X_{i+1}^+ := \mathbb{1}_{\mathcal{G}_i} \sum_{e \in O_i(A, B, x)} \mathbb{1}_{e \in \Gamma_{i+1}}$$

and

$$X^+ := \sum_{i=0}^{I-1} X_{i+1}^+.$$

Next define  $Z_{i+1}^+ = \text{Bin}(q_i n^2, p)$  to be independent random variables. As  $\mathbb{P}(-\mathcal{G}_i) = N^{-\omega(1)}$  and we are in the event  $\mathcal{G}_i$  we have  $|O_i(A, B, x)| \leq q_i n^2$ . Consequently,  $\mathbb{P}(X_{i+1}^+ \geq t) \leq \mathbb{P}(Z_{i+1}^+ \geq t)$  for all  $t$  and for all  $i$ . Then let  $Z^+ = \sum_{i=0}^{I-1} Z_{i+1}^+$  and notice that this is equal in distribution to  $\text{Bin}\left(\sum_{i=0}^{I-1} q_i n^2, p\right)$ . Also, for all  $t$

$$\mathbb{P}(X(A, B, x) \geq t \text{ and } \mathcal{G}_{\leq I}) \leq \mathbb{P}(X^+ \geq t) \leq \mathbb{P}(Z^+ \geq t).$$

Similarly, defining  $Z_{i+1}^- := \text{Bin}(\tau_i q_i n^2, p)$  and  $Z^- := \text{Bin}\left(\sum_{i=0}^{I-1} \tau_i q_i n^2, p\right)$  gives

$$\mathbb{P}(X(A, B, x) \leq t \text{ and } \mathcal{G}_{\leq I}) \leq \mathbb{P}(Z^- \leq t).$$

Let  $\mu^+ = \mathbb{E}(Z^+) = p \sum_{i=0}^{I-1} q_i n^2$  and  $\mu^- = \mathbb{E}(Z^-) = p \sum_{i=0}^{I-1} \tau_i q_i n^2$ . Now we can apply Theorem 6 to  $Z^+$  and  $Z^-$  to get

$$\mathbb{P}\left(X(A, B, x) \geq \left(1 + \frac{\delta}{2}\right) \mu^+ \text{ and } \mathcal{G}_{\leq I}\right) \leq \mathbb{P}\left(Z^+ \geq \left(1 + \frac{\delta}{2}\right) \mu^+\right) \leq \exp\left(-\frac{\delta^2 \mu^+}{8}\right)$$

$$\mathbb{P}\left(X(A, B, x) \leq \left(1 - \frac{\delta}{2}\right) \mu^- \text{ and } \mathcal{G}_{\leq I}\right) \leq \mathbb{P}\left(Z^- \leq \left(1 - \frac{\delta}{2}\right) \mu^-\right) \leq \exp\left(-\frac{\delta^2 \mu^-}{8(1 + \frac{\delta}{2})}\right).$$

By (11)

$$\mu^+ = \frac{n^2}{\sqrt{N}} \sum_{i=0}^{I-1} \sigma q_i = \frac{n^2}{\sqrt{N}} (\pi_I - \sigma).$$

Using the bound on  $\pi_I$  from (57) gives  $\mu^+ \geq \frac{C}{\sqrt{3}}\sqrt{\log N}n(\sqrt{\beta \log N} - 1)$ , and similarly since  $\tau_i = 1 - \frac{\delta\pi_i}{2\pi_I} \geq 1 - \frac{\delta}{2}$  and  $0 \leq \delta \leq 1$ , then  $\mu^- \geq \frac{C}{2\sqrt{3}}\sqrt{\log N}n(\sqrt{\beta \log N} - 1)$ . Thus

$$\mathbb{P}\left(X(A, B, x) \notin [(1 - \frac{\delta}{2})\mu^-, (1 + \frac{\delta}{2})\mu^+] \text{ and } \mathcal{G}_{\leq I}\right) \leq 2 \exp\left(-\frac{C\sqrt{B}\delta^2}{64\sqrt{3}}n \log N\right).$$

By the union bound

$$\begin{aligned} & \mathbb{P}\left(\exists x, A, B : X(A, B, x) \notin \left[\left(1 - \frac{\delta}{2}\right)\mu^-, \left(1 + \frac{\delta}{2}\right)\mu^+\right] \text{ and } \mathcal{G}_{\leq I}\right) \\ & \leq 2N^{2n+1} \exp\left(-\frac{C\sqrt{B}\delta^2}{64\sqrt{3}}n \log N\right) \\ & \leq 2 \exp\left((2n+1) \log N - \frac{C\sqrt{B}\delta^2}{64\sqrt{3}}n \log N\right) \\ & \leq N^{-\omega(1)} \end{aligned}$$

where the last inequality follows if we let  $C > \frac{144\sqrt{3}}{\delta^2\sqrt{\beta}}$ .

Now turning to  $Y(A, B, x)$  and recalling that  $B_{i+1}$  is the set of subsets of  $\Gamma_{i+1}$  which create a  $K_4^-$  together with  $H_i$  and  $D_{i+1}$  is a maximal subset of  $B_{i+1}$ , we introduce

$$\hat{Y}(A, B, x) := \sum_{i=0}^{I-1} |O_i(A, B, x) \cap (\cup_{S \in B_{i+1}} \cup_{e \in S} e)|.$$

Let

$$\begin{aligned} Y_{i+1}(A, B, x) &:= |O_i(A, B, x) \cap (\cup_{S \in D_{i+1}} \cup_{e \in S} e)|, \\ \hat{Y}_{i+1}(A, B, x) &:= |O_i(A, B, x) \cap (\cup_{S \in B_{i+1}} \cup_{e \in S} e)|. \end{aligned}$$

Let  $y := \frac{\delta^2\mu^-}{9}$ . Writing

$$Y_y^I = \{\mathbf{y} = (y_1, \dots, y_I) \in \mathbb{N}^I : \sum_{i=1}^I y_i = y\},$$

$$\begin{aligned} \mathbb{P}(Y(A, B, x) \geq y \text{ and } \mathcal{G}_{\leq I}) &\leq \sum_{\mathbf{y} \in \mathbb{Y}_y^I} \mathbb{P}(\cap_{i=0}^{I-1} Y_{i+1}(A, B, x) \geq y_{i+1} \text{ and } \mathcal{G}_{\leq i+1}) \\ &= \sum_{\mathbf{y} \in \mathbb{Y}_y^I} \prod_{i=0}^{I-1} \mathbb{P}(Y_{i+1} \geq y_{i+1} | \cap_{j=0}^{i-1} Y_{j+1}(A, B, x) \geq y_{j+1} \text{ and } \mathcal{G}_{\leq j+1}). \end{aligned} \tag{58}$$

Here we will use the concentration inequality Theorem 9 to bound  $Y_{i+1}(A, B, x)$ . To this end we

define

$$Y_{i+1}^+(A, B, x) := \sum_{e \in O_i(A, B, x)} \left( \sum_{\{e', e'' \in O_i: ee'e'' \text{ is a } K_4^-\}} \mathbb{1}_{ee'e'' \in D_{i+1}} + \sum_{\{e' \in O_i, e'' \in E_i: ee'e'' \text{ is a } K_4^-\}} \mathbb{1}_{ee' \in D_{i+1}} \right),$$

$$\hat{Y}_{i+1}^+(A, B, x) := \sum_{e \in O_i(A, B, x)} \left( \sum_{\{e', e'' \in O_i: ee'e'' \text{ is a } K_4^-\}} \mathbb{1}_{ee'e'' \in B_{i+1}} + \sum_{\{e' \in O_i, e'' \in E_i: ee'e'' \text{ is a } K_4^-\}} \mathbb{1}_{ee' \in B_{i+1}} \right),$$

and note that  $Y_{i+1}(A, B, x) \leq Y_{i+1}^+(A, B, x)$  and  $\hat{Y}_{i+1}(A, B, x) \leq \hat{Y}_{i+1}^+(A, B, x)$ .

Recall that  $B_{i+1}^2$  is the set of pairs of edges  $\{e, e'\}$  added to  $\Gamma_{i+1}$  where there is a third edge  $e'' \in H_i \subseteq E_i$  so  $e, e', e''$  comprise a  $K_4^-$ , and  $B_{i+1}^3$  is the set of triples of edges added to  $\Gamma_{i+1}$  which make a  $K_4^-$ . Further, recalling that  $B_{i+1} = B_{i+1}^2 \cup B_{i+1}^3$ ,  $\hat{Y}_{i+1}^+(A, B, x)$  includes edges from sets in  $B_{i+1}^2$  and edges from sets  $B_{i+1}^3$ . If  $xab$  is the edge in  $\hat{Y}_{i+1}^+(A, B, x)$  then the number of  $w \in [N]$  which can be the fourth vertex of the  $K_4^-$  in the case where one edge from the  $K_4^-$  is in  $E_i$  is  $|S_i(x, ab)| + |S_i(a, bx)| + |S_i(b, ax)| \leq 6q_i\pi_i\sqrt{N}$  assuming  $\mathcal{G}_{\leq i}$ . Similarly the number of  $w$  that can be the fourth vertex from a  $K_4^-$  where none of the edges were in  $E_i$  is  $|R_i(x, ab)| + |R_i(a, xb)| + |R_i(b, ax)| \leq 3q_i^2N$  assuming  $\mathcal{G}_{\leq i}$ . Hence assuming  $\mathcal{G}_{\leq i}$

$$\begin{aligned} \mathbb{E}(\hat{Y}_{i+1}^+(A, B, x)) &\leq |O_i(A, B, x)| \left( \mathbb{P}\left(xab \in (\cup_{S \in B_{i+1}^2} \cup_{e \in S} e)\right) + \mathbb{P}\left(xab \in (\cup_{S \in B_{i+1}^3} \cup_{e \in S} e)\right) \right) \\ &\leq q_i n^2 p^2 \left( 6q_i\pi_i\sqrt{N} \right) + q_i n^2 p^3 (3q_i^2 N). \end{aligned}$$

By Lemma 3 we have  $\max\{q_i^2, q_i\pi_i\} \leq 1$ . Recalling  $p = \sigma/\sqrt{N}$  we obtain

$$\mathbb{E}(\hat{Y}_{i+1}^+(A, B, x)) \leq 10\sigma p q_i n^2 =: \mu_{i+1}.$$

Our plan now is to apply Theorem 9. To this end, let  $Q = O_i$ ,  $\xi_i = \mathbb{1}_{e \in \Gamma_{i+1}}$ ,

$$\begin{aligned} J^2 &:= \{(e, e') : e \in O_i(A, B, x), e' \in O_i, \exists e'' \in E_i \text{ s.t. } ee'e'' \text{ is a } K_4^-\} \\ J^3 &:= \{(e, e', e'') : e \in O_i(A, B, x), e' \in O_i, e'' \in O_i, ee'e'' \text{ is a } K_4^-\} \\ J &:= J^2 \cup J^3. \end{aligned}$$

For all  $\alpha \in J^2$  let  $Q(\alpha) = \{e, e'\}$  and for  $\alpha \in J^3$  let  $Q(\alpha) = \{e, e', e''\}$ . Now  $X_0 = Y_{i+1}^+(A, B, x)$  and  $X = \hat{Y}_{i+1}^+(A, B, x)$  so by Theorem 9

$$\mathbb{P}(Y_{i+1}(A, B, x) \geq y_{i+1}) \leq \mathbb{P}(Y_{i+1}^+(A, B, x) \geq y_{i+1}) \leq \frac{\left(\mathbb{E}(\hat{Y}_{i+1}^+(A, B, x))\right)^{y_{i+1}}}{y_{i+1}!} \leq \left(\frac{\mu_{i+1}e}{y_{i+1}}\right)^{y_{i+1}}$$

where the last inequality comes from Stirling's formula.

Then for all  $y_{i+1}$  such that  $y_{i+1} \geq \frac{10\mu_{i+1}}{\sqrt{\sigma}}$  and assuming  $\mathcal{G}_{\leq I}$  we get

$$\mathbb{P}(Y_{i+1}(A, B, x) \geq y_{i+1}) \leq \sigma^{\frac{y_{i+1}}{2}}. \quad (59)$$

Turning to  $y_{i+1} < \frac{10\mu_{i+1}}{\sqrt{\sigma}}$ , trivially  $\mathbb{P}(Y_{i+1} \geq y_{i+1}) \leq 1$ . Then using  $\mu^- = p \sum_{i=0}^{I-1} \tau_i q_i n^2$  and  $\mu_{i+1} = 10\sigma p q_i n^2$  and  $\tau_i \geq \frac{1}{2}$ , we obtain

$$\sum_{0 < i < I: y_{i+1} < \frac{10\mu_{i+1}}{\sqrt{\sigma}}} y_{i+1} \leq \frac{10}{\sqrt{\sigma}} \sum_{i=0}^{I-1} \mu_{i+1} \leq \frac{10}{\sqrt{\sigma}} 20\sigma \mu^- \ll y. \quad (60)$$

Recalling the notation  $Y_y^I = \{\mathbf{y} = (y_1, \dots, y_I) \in \mathbb{N}^I : \sum y_i = y\}$ , and applying (58), (59), and (60),

$$\begin{aligned}
\mathbb{P}(Y(A, B, x) \geq y \text{ and } \mathcal{G}_{\leq I}) &\leq \sum_{\mathbf{y} \in Y_y^I} \prod_{i: y_{i+1} \geq \frac{10\mu_{i+1}}{\sqrt{\sigma}}} \mathbb{P}(Y_{i+1} \geq y_{i+1} | \cap_{j=0}^{i-1} Y_{j+1} \geq y_{j+1} \text{ and } \mathcal{G}_{\leq j+1}) \\
&\leq \sum_{\mathbf{y} \in Y_y^I} \sigma^{\frac{y}{2} - o(y)} \\
&\leq (y+2)^I \sigma^{\frac{y}{4}} \\
&\leq (2y)^{N^\beta} (\log N)^{-\frac{y}{2}} \leq (\log N)^{\frac{-\delta^2 \mu^-}{36}}
\end{aligned}$$

where the last inequality follows from  $(\log N)^y \gg y^I$  since  $y = \Theta(\mu^-)$  and  $\mu^- \geq \frac{C\sqrt{\beta}}{2\sqrt{3}}(\log N)n \gg I$ . By the union bound

$$\begin{aligned}
&\mathbb{P}(\exists x, A, B : Y(A, B, x) \geq y \text{ and } \mathcal{G}_{\leq I}) \\
&\leq N^{2n+1} (\log N)^{\frac{-\delta^2 \mu^-}{36}} \\
&\leq N^{2n+1} (\log N)^{-\frac{\delta^2 C \sqrt{\beta}}{72\sqrt{3}} (\log N)n} \\
&\leq \exp \left( (2n+1) \log N - \frac{\delta^2 C \sqrt{\beta}}{72\sqrt{3}} (\log N \log \log N)n \right) \\
&\leq N^{-\omega(1)}.
\end{aligned}$$

## References

- [1] M. Ajtai, J. Komlós and E. Szemerédi, A note on Ramsey numbers, J. Combin. Theory Ser. A **29** (1980), no. 3, 354–360.
- [2] M. Ajtai, J. Komlós and E. Szemerédi, A dense infinite Sidon sequence, European J. Combin. **2** (1981), no. 1, 1–11.
- [3] T. Bohman, The triangle-free process, Adv. Math. **221** (2009), no. 5, 1653–1677.
- [4] D. Conlon, J. Fox, B. Gunby, X. He, D. Mubayi, A. Suk, J. Verstraëte and H. Yu, When are off-diagonal hypergraph Ramsey numbers polynomial?, submitted
- [5] D. Conlon, J. Fox, X. He, D. Mubayi, A. Suk and J. Verstraëte., Hypergraph Ramsey numbers of cliques versus stars, Random Structures Algorithms **63** (2023), no. 3, 610–623.
- [6] H. Guo and L. Warnke, Packing nearly optimal Ramsey  $R(3, t)$  graphs, Combinatorica **40** (2020), no. 1, 63–103.
- [7] J. H. Kim, The Ramsey number  $R(3, t)$  has order of magnitude  $t^2 / \log t$ , Random Structures Algorithms **7** (1995), no. 3, 173–207.
- [8] M. Krivelevich, Bounding Ramsey numbers through large deviation inequalities, Random Structures Algorithms **7** (1995), no. 2, 145–155.



- [9] S. Mattheus and J. Verstraëte, The asymptotics of  $r(4, t)$ , *Ann. of Math. (2)* **199** (2024), no. 2, 919–941.
- [10] C. J. H. McDiarmid, On the method of bounded differences, in *Surveys in combinatorics, 1989 (Norwich, 1989)*, 148–188, London Math. Soc. Lecture Note Ser., 141, Cambridge Univ. Press, Cambridge.
- [11] C. J. H. McDiarmid, Concentration, in *Probabilistic methods for algorithmic discrete mathematics*, 195–248, Algorithms Combin., 16, Springer, Berlin.
- [12] D. Mubayi and J. Verstraëte, A note on pseudorandom Ramsey graphs, *J. Eur. Math. Soc. (JEMS)* **26** (2024), no. 1, 153–161.
- [13] V. Rödl, On a packing and covering problem, *European J. Combin.* **6** (1985), no. 1, 69–78.
- [14] L. Warnke, On the method of typical bounded differences, *Combin. Probab. Comput.* **25** (2016), no. 2, 269–299.
- [15] L. Warnke, Upper tails for arithmetic progressions in random subsets, *Israel J. Math.* **221** (2017), no. 1, 317–365.
- [16] N. C. Wormald, Differential equations for random processes and random graphs, *Ann. Appl. Probab.* **5** (1995), no. 4, 1217–1235.

## 9 Appendix

### 9.1 Proof of Lemma 3

1. To prove (12) first we will prove the following fact for  $x \geq e$

$$\int_0^{\sqrt{\log x}-1} e^{t^2} dt \leq x \leq \int_0^{\sqrt{\log x}+1} e^{t^2} dt \quad (61)$$

For the upper bound, notice that the rectangle in  $\mathbb{R}^2$  defined by

$$R := \{(t, y) \mid \sqrt{\log x} \leq t \leq \sqrt{\log x} + 1, 0 \leq y \leq x\}$$

clearly has an area of  $x$  and falls under the curve  $y = e^{t^2}$ . Then since for  $e^{t^2} \geq 0$  the area of  $R$  is no more than  $\int_0^{\sqrt{\log x}+1} e^{t^2} dt$ , proving the upper bound.

For the lower bound, for all integers  $0 \leq i \leq \sqrt{\log x} - 1$  define the regions

$$R_i := \{(t, y) : i - 1 \leq t \leq i, 0 \leq y \leq e^{i^2}\}$$

Now let

$$S = \cup_{i=0}^{\sqrt{\log x}-1} R_i$$

and notice that the area of  $S$  is at least  $\int_0^{\sqrt{\log x}-1} e^{t^2} dt$ . Also, the area of

$$\text{Area}(S) = \sum_{i=0}^{\sqrt{\log x}-1} e^{i^2} \leq (\sqrt{\log x} - 1)e^{(\sqrt{\log x}-1)^2} = x \frac{e(\sqrt{\log x} - 1)}{e^{\sqrt{\log x}}} \leq x$$

where the last inequality holds since  $x \geq e$ . This completes the lower bound and proves (61).

Recall that  $x = \int_0^{\Psi(x)} e^{3t^2} dt$  by the definition of  $\Psi(x)$ . Since for  $t \geq 0$  we have  $e^{3t^2} \geq 0$ , and for  $a, b \geq 0$  we have  $a \leq b$  iff  $\int_0^a e^{3t^2} dt \leq \int_0^b e^{3t^2} dt$ . Note that

$$\int_0^{\sqrt{\frac{\log(\sqrt{3}x)}{3}} \pm \frac{1}{\sqrt{3}}} e^{3t^2} dt = \frac{1}{\sqrt{3}} \int_0^{\sqrt{\log(\sqrt{3}x)} \pm 1} e^{u^2} du$$

Using (61), we obtain

$$\int_0^{\sqrt{\frac{\log(\sqrt{3}x)}{3}} - \frac{1}{\sqrt{3}}} e^{3t^2} dt \leq x \leq \int_0^{\sqrt{\frac{\log(\sqrt{3}x)}{3}} + \frac{1}{\sqrt{3}}} e^{3t^2} dt \quad (62)$$

The bounds on  $\Psi(x)$  follow from (62).

2. For (13) note  $\Psi(x)$  is strictly increasing for  $x \geq 0$  and  $\Psi'(x)$  is strictly decreasing for  $x \geq 0$  since  $\Psi'(x) = \exp(-3\Psi^2(x))$ . Together with  $\Psi'(0) = 1$  this gives that  $q_i \leq 1$  for all  $i$ , and notice that since  $q_i = \exp(-3\Psi^2(i\sigma))$  then we also get that  $q_i \geq 0$ .
3. For (14) the first part follows trivially from the definition of  $\pi_i$ .

For the second part of this statement, note that since  $\Psi(x)$  is strictly increasing and  $\Psi'(x)$  is strictly decreasing for  $x \geq 0$ . Therefore, for all  $j < i$ ,

$$\sigma\Psi'(j\sigma) \geq \Psi((j+1)\sigma) - \Psi(j\sigma) \geq \sigma\Psi'((j+1)\sigma)$$

Then notice that  $\Psi(i\sigma) = \sum_{j=0}^{i-1} (\Psi((j+1)\sigma) - \Psi(j\sigma))$  since  $\Psi(0) = 0$ , and also recall that  $\pi_i = \sigma + \sum_{j=0}^{i-1} \sigma q_j = \sigma + \sum_{j=0}^{i-1} \sigma\Psi'(j\sigma)$ . Thus we get

$$\begin{aligned} \pi_i - \Psi(i\sigma) &= \left( \sigma + \sum_{j=0}^{i-1} \sigma\Psi'(j\sigma) \right) - \Psi(i\sigma) \\ &\geq \sigma + \left( \sum_{j=0}^{i-1} \Psi((j+1)\sigma) - \Psi(j\sigma) \right) - \left( \sum_{j=0}^{i-1} \Psi((j+1)\sigma) - \Psi(j\sigma) \right) = \sigma. \end{aligned}$$

Similarly

$$\pi_i - \Psi(i\sigma) = \left( \sigma + \sigma\Psi'(0) - \sigma\Psi'(i\sigma) + \sum_{j=0}^{i-1} \sigma\Psi'((j+1)\sigma) \right) - \Psi(i\sigma) \leq 2\sigma$$

where the last inequality follows since  $\sigma\Psi'(0) = \sigma$  and  $\Psi'(i\sigma) \geq 0$ .

4. Turning to (15), by the definition of  $\pi_i$  we have that  $\pi_i \leq \pi_I$  for all  $0 \leq i \leq I$ . Then

$$\pi_i \leq \sqrt{\frac{\log(\sqrt{3}(I\sigma))}{3}} + 1 \leq \sqrt{\frac{\beta \log N + 1}{3}} \ll \log N = \sigma^{-1/2}.$$

5. Next for (16) we start by noting that for all  $x > 0$ ,  $e^{-3x^2}x^2 \leq \frac{1}{2}$  and  $e^{-3x^2}x \leq \frac{1}{2}$ . Indeed, notice that for  $x > 1$   $3x^2 \geq 2\ln x + \ln \frac{1}{2}$ , for  $\frac{1}{2} \leq x \leq 1$ ,  $-3x^2 \leq \ln \frac{1}{2}$  and  $\ln x \leq 0$ , and for  $0 < x < \frac{1}{2}$ ,  $\ln x \leq \ln \frac{1}{2}$  and  $-3x^2 \leq 0$ . Then applying this inequality with  $x = \Psi(i\sigma)$  together with (14) gives

$$q_i \pi_i^2 \leq e^{-3\Psi^2(i\sigma)} (\Psi(i\sigma) + 2\sigma)^2 = e^{-3\Psi^2(i\sigma)} (\Psi^2(i\sigma) + 4\Psi(i\sigma)\sigma + 4\sigma^2) \leq 1.$$

6. To prove (17) we will use the fact that for some  $s \in [i\sigma, (i+1)\sigma]$ ,

$$\Psi'((i+1)\sigma) = \Psi'(i\sigma) + \Psi''(i\sigma)(\sigma) + \frac{\Psi'''(s)}{2}(\sigma^2)$$

From the definition of  $\Psi(x)$ , it can easily be seen that  $\Psi''(i\sigma) = -6(\Psi'(i\sigma))^2\Psi(i\sigma)$  and  $\Psi'''(i\sigma) = -6(\Psi'(i\sigma))^3(1 + 2\Psi^2(i\sigma))$ . Recalling  $q_i = \Psi'(i\sigma)$  and  $\Psi(i\sigma) \in [\pi_i - 2\sigma, \pi_i + 2\sigma]$ , we obtain

$$q_{i+1} - q_i \in [-6\sigma q_i^2(\pi_i + 2\sigma) - 3\sigma^2 q_i^3(1 + 2(\pi_i + 2\sigma)^2), -6\sigma q_i^2(\pi_i - 2\sigma) - 3\sigma^2 q_i^3(1 + 2(\pi_i - 2\sigma)^2)]$$

Now since  $6\sigma q_i^2(2\sigma) + 3\sigma^2 q_i^3(1 + 2(\pi_i + 2\sigma)^2) \leq 16\sigma^2 q_i^2$ , we obtain  $|(q_{i+1} - q_i) + 6\sigma q_i^2 \pi_i| \leq 16\sigma^2 q_i^2$ .

7. To prove (18) note that since for  $x \geq e$  we have that  $\sqrt{\frac{\log(\sqrt{3}x)}{3}} - \frac{1}{\sqrt{3}} \leq \Psi(x) \leq \sqrt{\frac{\log(\sqrt{3}x)}{3}} + \frac{1}{\sqrt{3}}$  and  $\Psi'(x) = \exp(-3\Psi^2(x))$ ,  $\Psi'(x) = \frac{x^{-1+o(1)}}{e\sqrt{3}}$ . Then using the fact that  $I = \lceil N^\beta \rceil$  and since  $\Psi'(x)$  is strictly decreasing we have  $q_i \geq q_I$  for all  $0 \leq i \leq I$ , and  $q_i \geq \frac{N^{-\beta+o(1)}}{e\sqrt{3}}$ .

8. The inequality (19) follows from

$$\hat{p}_{e,i} \leq 1 - (1-p)^{6\sqrt{N}q_i(\pi_i + \sqrt{\sigma})} \leq 6\sigma q_i(\pi_i + \sqrt{\sigma}) \leq q_i.$$

9. Finally for (20), note that since  $\Psi$  is increasing and  $\Psi'$  is decreasing

$$\begin{aligned} |q_i - q_{i+1}| &\leq \sigma \max_{s \in [i\sigma, (i+1)\sigma]} |\Psi''(s)| \\ &\leq \max_{s \in [i\sigma, (i+1)\sigma]} 6\sigma (\Psi'(s))^2 \Psi(s) \\ &\leq 6\sigma (\Psi'(i\sigma))^2 \Psi((i+1)\sigma) \\ &\leq 6\sigma q_i^2(\pi_{i+1} - \sigma) \\ &\leq 6\sigma q_i^2 \pi_i. \end{aligned}$$

Then using  $q_i, q_i \pi_i \leq 1$  bounds the bounds  $q_i - q_{i+1} < 6\sigma q_i$  and  $q_i - q_{i+1} < 6\sigma q_i \pi_i$ , and for the last bound, notice these imply that  $\frac{q_i}{q_{i+1}} < \frac{1}{1-6\sigma} < 2$ , so  $q_i < 2q_{i+1}$ . Then  $q_i - q_{i+1} < 12\sigma q_{i+1}$ .