

# Extending Mubayi and Turán's Algorithm to 3-graphs

Ferran Espuña

Let  $G$  be a 3-graph with  $n$  vertices and  $m = \epsilon n^3$  edges. A polynomial time algorithm is given to find a  $K(q, q, q)$  in  $G$  for

$$q = c_\epsilon^{(3)} \sqrt{\log n}$$

As long as (insert condition here).

Note that this result is tight up to the constant  $c_\epsilon^{(3)}$ , as proved in [1]. This result is a generalization of the result in 2-graphs by [3], and algorithm will be analogous to the one given there. The procedure is as follows:

1. Choose parameters  $q < r < n$  depending on  $n$  and  $\epsilon$ .
2. Let  $R$  be the set of  $r$  vertices with the highest degree in  $G$ .
3. find a subset  $Q \subset R$  with  $q$  vertices such that there is a large  $S \subset \binom{[n] \setminus Q}{2}$  satisfying  $xyz \in E(G) \forall \{x, y\} \in S, z \in Q$ . Say, of size  $s$ .
4. Apply the algorithm of [3] to find a  $K(q, q)$  in the 2-graph induced by  $S$ . Say, we find partition  $S \supset U \cup V$ .

If successful, a  $K(q, q, q)$  has been found in  $G$  with parts  $U, V, Q$ . The problem is now to find parameters  $q, r$  such that the above procedure is successful and the algorithm runs in polynomial time.

**Lemma 1.** *As long as  $r < \epsilon n$ , there are at least  $\epsilon r n^2$  edges in  $G$  with exactly one vertex in  $R$ .*

*Proof.* The sum of the degrees in  $G$  is  $3m$ . Therefore, by the pigeonhole principle,

$$\sum_{v \in R} d(v) \geq r \cdot \frac{3m}{n} = 3\epsilon r n^2$$

However, here we are overcounting:

- The edges with only one vertex in  $R$  are counted exactly once.
- The edges with two vertices in  $R$  are counted twice. The contribution of these is at most  $r(r-1)(n-r) < r^2 n$

- The edges with all vertices in  $R$  are counted three times. The contribution of these is at most  $r(r-1)(r-2) < r^3 < r^2n$

Therefore, the condition will hold as long as  $r^2n < \epsilon rn^2 \iff r < \epsilon n$ .  $\square$

Next, a counting argument in the style of [2] is used to guarantee the existence of  $Q$  and  $S$ . The size  $q$  of  $Q$  will be left as a parameter to be determined later, and the size  $s$  of  $S$  will be determined by the following lemma:

**Lemma 2.** *Under the same assumptions as in Lemma 1, and assuming  $r \leq n/2$ , there is a subset  $Q \subset R$  of size  $q$  and a subset  $S \subset \binom{[n] \setminus Q}{2}$  of size*

$$s := \frac{n^2}{8} \left( \frac{\epsilon}{e} \right)^q$$

such that  $xyz \in E(G) \forall \{x, y\} \in S, z \in Q$ .

*Proof.* Let  $E$  be the set of edges with exactly one vertex in  $R$ . for every  $\{x, y\} \in \binom{[n] \setminus R}{2}$ , let  $E_{xy}$  be the set of edges in  $E$  containing  $x$  and  $y$ . Finally, let

$$T = \left\{ P \subset E_{xy} : x, y \in \binom{[n] \setminus R}{2}, |P| = q \right\}$$

On the one hand, the number of elements in  $T$  is

$$\sum_{\{x, y\} \in \binom{[n] \setminus R}{2}} \binom{|E_{xy}|}{q} \geq \binom{n-r}{2} \binom{\epsilon rn^2 / \binom{n-r}{2}}{q} > \binom{n-r}{2} \binom{\epsilon r}{q} \geq \frac{n^2}{8} \binom{\epsilon r}{q}$$

Where the first inequality follows from Jensen's inequality and the third from the fact that  $r \leq n/2$ .

On the other hand, there are only  $\binom{r}{q}$  possible  $q$ -subsets of  $R$ . By the pigeonhole principle, one of these (say  $Q$ ) must be the set of vertices in  $R$  associated with  $P_j$  for  $k$  different  $P_j \in T$ , where  $k$  is

$$\frac{|T|}{\binom{r}{q}} > \frac{n^2 \binom{\epsilon r}{q}}{8 \binom{r}{q}} \geq \frac{n^2 \left( \frac{\epsilon r}{q} \right)^q}{8 \left( \frac{er}{q} \right)^q} = \frac{n^2}{8} \left( \frac{\epsilon}{e} \right)^q = s$$

$\square$

Now, the algorithm of [3] is applied to the graph  $G'$  with vertex set  $[n]$  and edge set  $S$ .

This yields a  $K(q', q') \subset G'$  with

$$q' = \left\lfloor \frac{\ln(n/2)}{\ln(2en^2/s)} \right\rfloor = \left\lfloor \frac{\ln(n/2)}{\ln(16e^{q+1}/\epsilon^q)} \right\rfloor$$

For the found subgraph to be a  $K(q, q, q)$ , it is necessary that  $q' \geq q$ . A sufficient condition is that

$$q \leq \frac{\ln(n/2)}{\ln(16e^{q+1}/\epsilon^q)} - 1 = \frac{\ln(n/2)}{\ln(16e) - q \ln(e/\epsilon)} - 1$$

This is true for

$$0 \leq q \leq \frac{\ln(16e) + \sqrt{(\ln(16e))^2 + 4 \ln(n/(32e)) \ln(e/\epsilon)}}{2 \ln(e/\epsilon)}$$

so a valid value for  $q$  is

$$q = \frac{\sqrt{4 \ln(n/(32e)) \ln(e/\epsilon)}}{2 \ln(e/\epsilon)} = \frac{\sqrt{\ln(n/(32e))}}{\sqrt{\ln(e/\epsilon)}}$$

## References

- [1] P. Erdős. On extremal problems of graphs and generalized graphs. *Israel Journal of Mathematics*, 2(3):183–190, September 1964.
- [2] Sós V. Turán P. Kóvari, T. On a problem of k. zarankiewicz. *Colloquium Mathematicae*, 3(1):50–57, 1954.
- [3] Dhruv Mubayi and György Turán. Finding bipartite subgraphs efficiently. *Information Processing Letters*, 110(5):174–177, 2010.