Stability for Turán's theorem.

In this note we prove a version of the classical result of Erdös and Simonovits that a graph with no K_t subgraph and a number of edges close to the maximum is close to the extreme example. In particular, such a graph is nearly $(t-1)$-colorable. Our methods can be used to obtain similar stability results in a wider variety of situations.

We will use $V(G)$ to denote the set of vertices of a hypergraph G. Following the convention used in class G will be identified with its set of edges. In particular, $|G|$ denotes the number of edges in a hypergraph G. Given an r-graph H, let $\mathcal{E}x(H)$ denote the family of all r-graphs not containing H. Let

$$ex(n, H) := \max_{G \in \mathcal{E}x(H), |V(G)|=n} |G|,$$

and let the *Turán density of H* be defined as

$$\pi(H) := \lim_{n \to \infty} \frac{ex(n, H)}{\binom{n}{r}}.$$

We have shown in class that this limit exists.

The next lemma will demonstrate that almost all the vertices in a graph in $\mathcal{E}x(K_t)$ with density close to $\pi(K_t) = \frac{t-2}{t-1}$ have degree close to the average.

Lemma 1. *For every r-graph H and every $\varepsilon > 0$ there exists $\delta > 0$ and $n_0 > 0$ such that every r-graph $G \in \mathcal{E}x(H)$ with $|V(G)| \geq n_0$ and $|G| \geq (1-\delta)\pi(H)|V(G)|^r/r!$*

- *either contains a sub-r-graph G' with $n' := |V(G')| > (1-\varepsilon)n$ such that every vertex of G' belongs to more than $(1-\varepsilon)\pi(H)(n')^{r-1}/(r-1)!$ edges, or*

- *contains a sub-r-graph G' with $n' := |V(G')| = \lfloor (1-\varepsilon)n \rfloor$ and $|G'| > \pi(H)(n')^r/r!$.*

Proof. Let δ be chosen so that $(1-\delta)^2 > 1-\varepsilon/2$ and $\delta < r\varepsilon^2/2$. Let n_0 be chosen so that $(n'')^r \geq (n''-1)^r + (1-\delta)r(n'')^{r-1}$ for all $n'' \geq (1-\varepsilon)n_0$.
Let $n := |V(G)|$. If every vertex of G belongs to more than $(1-\varepsilon)\pi(H)n^{r-1}/(r-1)!$ edges the lemma holds. Otherwise, delete a vertex of G which belongs to at most these many edges to obtain an r-graph G_1. Repeat this procedure on G_1, deleting a vertex belonging to at most $(1-\varepsilon)\pi(H)(n-1)^{r-1}/(r-1)!$ edges, if necessary, to obtain a graph G_2, etc. If the procedure stops in less than εn steps the lemma holds. Otherwise, we obtain a graph $G' := G_k$ with $k = \lceil \epsilon n \rceil$. We have $n' := |V(G')| = \lfloor (1-\varepsilon)n \rfloor$ and it remains to upper bound $|G|$.
We prove by induction on l that

$$|G_l| \geq \left(1 - \frac{k-l}{k}\delta\right)\pi(H)\frac{(n-l)^r}{r!},$$

for $l \leq k$. The lemma will follow. The base case for $G_0 := G$ is immediate. For the induction step, let $n'' = n-(l-1)$. We have

$$\begin{aligned}
\frac{|G_l|}{\pi(H)} &\geq \frac{|G_{l-1}|}{\pi(H)} - (1-\varepsilon)\frac{(n'')^{r-1}}{(r-1)!} \\
&\geq \left(1 - \frac{k-l+1}{k}\delta\right)\frac{(n'')^r}{r!} - (1-\varepsilon)\frac{(n'')^{r-1}}{(r-1)!} \\
&\geq \left(1 - \frac{k-l+1}{k}\delta\right)\left(\frac{(n''-1)^r}{r!} + (1-\delta)\frac{(n'')^{r-1}}{(r-1)!}\right) - (1-\varepsilon)\frac{(n'')^{r-1}}{(r-1)!} \\
&\geq \left(1 - \frac{k-l}{k}\delta\right)\frac{(n''-1)^r}{r!} - \frac{\delta}{\varepsilon n}\frac{(n''-1)^r}{r!} + \frac{\varepsilon}{2}\frac{(n'')^{r-1}}{(r-1)!} \\
&\geq \left(1 - \frac{k-l}{k}\delta\right)\frac{(n''-1)^r}{r!} + \left(\frac{\varepsilon}{2} - \frac{\delta}{\varepsilon r}\right)(n'')^{r-1}(r-1)! \\
&\geq \left(1 - \frac{k-l}{k}\delta\right)\frac{(n''-1)^r}{r!},
\end{aligned}$$

as desired. In the chain of inequalities above, the induction hypothesis is used in the second line, the choice of n_0 in the third line, and the choice of δ in the fourth and fifth line. □

Note that the second outcome of Lemma 1 can not occur for many choices of H. In particular, when $H = K_t$, it is impossible by Turán's theorem. We are now ready for our main result.

Theorem 2. *For every positive integer $t \geq 3$ and every $\varepsilon' > 0$ there exists $\delta > 0$ and $n_0 > 0$ so that every $G \in \mathcal{E}x(K_t)$ with $|V(G)| \geq n_0$ and*

$$|G| \geq (1-\delta)\frac{t-2}{t-1}\frac{|V(G)|^2}{2}$$

contains disjoint subsets of vertices $A_1, A_2, \ldots, A_{t-1}$ so that

$$|A_1 \cup A_2 \cup \ldots \cup A_{t-1}| \geq (1-\varepsilon')|V(G)|$$

and no edge of G joins to vertices in some A_i to each other.

Proof. Let $\varepsilon = \min\{\varepsilon'/2(t-2), 1/t^2\}$ and let n_0 and δ be chosen as in Lemma 1 for this value of ε, $r = 2$ and $H = K_t$. Let G' satisfy the first outcome of the lemma. (As we noted above the second outcome can not hold.) Let $n := |V(G)|$. As $|G'| \geq (1-\varepsilon)\frac{t-2}{t-1}\frac{n^2}{2}$, by the choice of ε and Turán's theorem G' contains a complete subgraph on $t-1$ vertices. Let $V = \{v_1, v_2, \ldots, v_{t-1}\}$ be set of vertices of this subgraph. For $i = 1, 2, \ldots, t-1$, let A_i consist of the set of vertices of G which are connected to all vertices of V except for v_i. No edge of G joins to vertices in some A_i to each other, as otherwise the corresponding vertices together with $V - \{v_i\}$ induce a complete subgraph of G on t vertices.
By the choice of G' and ε we have

$$\deg(v_i) \geq (1-\varepsilon)\frac{t-2}{t-1}(1-\varepsilon)n \geq \left(1 - \frac{\varepsilon'}{t-2}\right)\frac{t-2}{t-1}n.$$

Let $A := A_1 \cup A_2 \cup \ldots \cup A_{t-1}$. Then every vertex in $V(G) - A$ is joined to at most $t-3$ vertices in V. It follows that

$$\begin{aligned}\left(1 - \frac{\varepsilon'}{t-2}\right)(t-2)n \leq \sum_{i=1}^{t-1}\deg(v_i) &\leq (t-2)|A| + (t-3)(|V(G) - A|) \\ &= (t-3)n - |A|\end{aligned}$$

It follows that $|A| \geq (1-\varepsilon')n$, as desired. □

The above theorem can be routinely strengthened in several ways:

- By modifying the choice of δ one can remove the requirement on the minimum size of $|V(G)|$. Indeed, for an appropriate choice of δ and $|V(G)| < n_0$ one would have $|G| = \frac{t-2}{t-1}\frac{|V(G)|^2}{2}$. It follows from the first proof of Turán's theorem presented in class that in such a case the set A constructed in the proof of the Theorem 2 is equal to $V(G)$.

- The bound on the number of edges of G implies that by once again modifying the choice of δ one can guarantee that $|A_i| \geq \frac{1-\varepsilon}{t-1}n$ for every $i = 1, 2, \ldots, t-1$ and for $i \neq j$ there are at least

$$(1-\varepsilon)\left(\frac{n}{t-1}\right)^2$$

  edges joining A_i to A_j. Thus the graph G "differs" from a complete $(t-1)$-partite graph with sizes of parts as equal as possible (*the Turán graph*) by at most $O(\varepsilon n^2)$ edges.