Inverting the Turán problem

Christopher Cox — **Carnegie Mellon University**
Joseph Briggs — **Carnegie Mellon University**

Abstract

Classical questions in extremal graph theory concern the asymptotics of $\mathrm{ex}(G, \mathcal{H})$ where $\mathcal{H}$ is a fixed family of graphs and $G = G_n$ is taken from a "standard" increasing sequence of host graphs $(G_1, G_2, \dots)$, most often K_n or $K_{n,n}$. Inverting the question, we can instead ask how large $|E(G)|$ can be with respect to $\mathrm{ex}(G, \mathcal{H})$. We show that the standard sequences indeed maximize $|E(G)|$ for some choices of $\mathcal{H}$, but not for others. Many interesting questions and previous results arise very naturally in this context, which also, unusually, gives rise to sensible extremal questions concerning multigraphs and non-uniform hypergraphs.

For a graph G and a family of graphs $\mathcal{H}$, the extremal number of $\mathcal{H}$ in G is defined to be

$$\mathrm{ex}(G, \mathcal{H}) = \max\{|E(F)| : F \subseteq G \text{ and } H \not\subseteq F \text{ for any } H \in \mathcal{H}\}.$$

When the family consists only of a single graph, $\mathrm{ex}(G, H)$ is used in place of $\mathrm{ex}(G, \{H\})$.

A typical example of this is when $\mathcal{H} = \{C_3, C_4, C_5, \dots\}$ is the collection of all cycles, in which case the extremal number is simply the *graphic matroid rank* of G, an important graph parameter in its own right.

The Turán problem, one of the cornerstones of extremal graph theory, concerns the behavior of $\mathrm{ex}(K_n, H)$ for a fixed H when n is large. The first result along these lines is a theorem of Mantel which states that $\mathrm{ex}(K_n, K_3) = \lfloor n^2/4 \rfloor$. Turán obtained a version for K_t in place of K_3, in particular obtaining $\mathrm{ex}(K_n, K_t) = \left(1 - \frac{1}{t-1} + o(1)\right)\frac{n^2}{2}$ where $o(1) \to 0$ as $n \to \infty$. In a similar spirit, the Erdős–Stone Theorem states that if $\chi = \chi(H)$ is the chromatic number of H, then $\mathrm{ex}(K_n, H) = \left(1 - \frac{1}{\chi-1} + o(1)\right)\frac{n^2}{2}$. The Erdős–Stone Theorem asymptotically answers the Turán problem, except when H is bipartite, in which case the bound becomes $o(n^2)$. In this situation, known as the degenerate case, the asymptotic behavior of very few graphs is known and is an active area of research.

Most approaches in the case of a bipartite graph instead ask about $\mathrm{ex}(K_{n,n}, H)$, which is known as the Zarankiewicz problem. This is often seen as a more natural question and provides bounds on the Turán problem as $\frac{1}{2}\mathrm{ex}(K_n, H) \leq \mathrm{ex}(K_{n/2,n/2}, H) \leq \mathrm{ex}(K_n, H)$ for bipartite H. In the special case of $H = C_4$, the incidence graphs showing tightness for the Zarankiewicz problem were spotted a few years before polarity graphs showing tightness for the Turán problem.

With this in mind, we set out to explore a framework in which to ask: what is the most "natural" or "best" host graph for a fixed family of graphs? This suggests optimizing a particular monotone graph parameter over all host graphs G where $\mathrm{ex}(G, \mathcal{H})$ is bounded, the simplest of which is just the edge count. Thus we define the following extremal function for $\mathcal{H}$:

$$\mathcal{E}_k(\mathcal{H}) := \sup\{|E(G)| : \mathrm{ex}(G, \mathcal{H}) < k\}.$$

In other words, for a family $\mathcal{H}$, we would like to determine the host graph G with the most edges such that any k edges from G contain some copy of $H \in \mathcal{H}$. In other words, G is best at "forcing" a copy of some $H \in \mathcal{H}$. When the family consists only of a single graph, we write $\mathcal{E}_k(H)$ in place of $\mathcal{E}_k(\{H\})$. Note that it is necessary to consider the supremum here as $\mathcal{E}_k(\mathcal{H})$ may be infinite. In

particular, $\mathcal{E}_k(K_{1,t}) = \mathcal{E}_k(tK_2) = \infty$ for $k \geq t$ as for any $s \geq t$, $\mathrm{ex}(K_{1,s}, K_{1,t}) = t-1 = \mathrm{ex}(sK_2, tK_2)$, despite both host graphs having s edges. However, we will later show that stars and matchings classify all families having $\mathcal{E}_k(\mathcal{H}) = \infty$.

In a similar fashion to the original Turán problem, this paper considers two questions:

- What are the asymptotics of $\mathcal{E}_k(\mathcal{H})$?
- When $\mathcal{E}_k(\mathcal{H})$ can be determined precisely, which host graphs G attain $|E(G)| = \mathcal{E}_k(\mathcal{H})$?

On the one hand, we will show that for non-bipartite H, this parameter behaves more or less as one might expect. For example, the following theorem is close in spirit to the Erdős–Stone Theorem:

Theorem 1. *If $\mathcal{H}$ is a family of graphs with $\rho = \min\{\chi(H) : H \in \mathcal{H}\} \geq 3$, then*

$$\mathcal{E}_k(\mathcal{H}) = \left(1 + \frac{1}{\rho - 2} + o(1)\right) k.$$

Recalling our motivation from the Zarankiewicz problem, we show that complete bipartite graphs are optimal hosts for at least one natural family, namely the collection $\mathcal{C}_e := \{C_4, C_6, \dots\}$ of even cycles:

Theorem 2. *For $k \geq 4$, $\mathcal{E}_k(\mathcal{C}_e) = \left\lfloor \frac{k^2}{4} \right\rfloor$, with $K_{\lfloor k/2 \rfloor, \lceil k/2 \rceil}$ being the unique extremal graph for $k \geq 6$.*

On the other hand, this is already a challenge for the case $H = K_{2,2}$:

Question. What is $\mathcal{E}_k(C_4)$ and what is the optimal host graph?

One peculiar feature of our question is that it is sensible even for multigraphs or nonuniform hypergraphs. We let $\mathcal{E}_k^*(\mathcal{H})$ denote the maximum number of edges among host *multigraphs* G with $\mathrm{ex}(G, \mathcal{H}) < k$. The parameter $\mathcal{E}_k^*(\mathcal{H})$ will be important in proving bounds on $\mathcal{E}_k(\mathcal{H})$ when $\mathcal{H}$ is a family of simple graphs. However, we do not even know the following:

Conjecture. If $\mathcal{H}$ consists only of simple graphs, then $\mathcal{E}_k(\mathcal{H}) = \mathcal{E}_k^*(\mathcal{H})$.

Curiously, for *non-uniform* graphs H without parallel edges, the above conjecture fails:

Theorem 3. *Let $\mathcal{O}_2$ be the graph with a single edge and a loop at each end. Then $\mathcal{E}_k(\mathcal{O}_2) = \frac{3k}{2}$, whereas $\mathcal{E}_k^*(\mathcal{O}_2) \sim \phi k$, where ϕ is the golden ratio.*

In our study of $\mathcal{E}_k(\mathcal{H})$ and optimal host graphs, we will show also that:

1. Cliques are best at forcing cliques,
2. Cliques are best at forcing a cycle,
3. Complements of matchings are best at forcing $\{P_3, K_3\}$,
4. Cliques with pendant edges are best at forcing P_3,
5. Two disjoint cliques or a modified power of a cycle are best at forcing $P_1 \cup P_2$,
6. For uniform hypergraphs, $\mathcal{E}_k(H)$ is only infinite for sunflowers,
7. For 1-uniform multigraphs, $\mathcal{E}_k^*(H)$ is quadratic in k.