The typical structure of sparse K_{r+1}-free graphs

Lutz Warnke
University of Cambridge

(joint work with József Balogh, Robert Morris, and Wojciech Samotij)

H-free graphs / Turán's theorem

Definition

Let H be a (small) fixed graph. A graph G is called H-free if it does not contain H as a (not necessarily induced) subgraph.

Definition

Given an integer n, we let the Turán number for H, denoted $\text{ex}(n, H)$, be the maximum number of edges in an n-vertex H-free graph.

Theorem (Turán [1941])

For every $n \geqslant r \geqslant 2$, the unique largest K_{r+1}-free graph on n vertices is the complete r-partite graph whose each color class has $\lfloor n/r \rfloor$ or $\lceil n/r \rceil$ elements, denoted $T_r(n)$. In particular,

$$\text{ex}(n, K_{r+1}) = e(T_r(n)) = \left(1 - \frac{1}{r}\right) \frac{n^2}{2} \pm O(n).$$

The Erdős–Kleitman–Rothschild theorem

Turán's theorem says something about the *extremal* H-free graphs. In this talk, we are interested in the properties of a *typical* H-free graph, as in the following classical result. Let $[n] = \{1, \ldots, n\}$ and let

$$\mathcal{F}_n(H) = \{H\text{-free graphs with the vertex set } [n]\},$$
$$\mathcal{G}_n(r) = \{r\text{-colorable graphs with the vertex set } [n]\}.$$

Theorem (Erdős, Kleitman, Rothschild [1976])

Almost all (a.a.) triangle-free (K_3-free) graphs are bipartite. More precisely,

$$\lim_{n\to\infty} \frac{|\mathcal{F}_n(K_3) \cap \mathcal{G}_n(2)|}{|\mathcal{F}_n(K_3)|} = 1.$$

In other words, if $F_n \in \mathcal{F}_n(K_3)$ is chosen uniformly at random (u.a.r.), then

$$\lim_{n\to\infty} \Pr(F_n \text{ is bipartite}) = 1.$$

The typical structure of H-free graphs

Theorem (Erdős, Frankl, Rödl [1986])

If $\chi(H) \geqslant 3$, then

$$2^{\mathrm{ex}(n,H)} \leqslant |\mathcal{F}_n(H)| \leqslant 2^{(1+o(1))\cdot \mathrm{ex}(n,H)}.$$

Theorem (Kolaitis, Prömel, Rothschild [1987])

For every $r \geqslant 2$, a.a. K_{r+1}-free graphs are r-colorable. That is,

$$\lim_{n\to\infty} \frac{|\mathcal{F}_n(K_{r+1}) \cap \mathcal{G}_n(r)|}{|\mathcal{F}_n(K_{r+1})|} = 1.$$

Theorem (Prömel, Steger [1992])

If H has a color-critical edge, then a.a. H-free graphs are $(\chi(H) - 1)$-col.

Several improvements and extensions of these results due to Balogh, Bollobás, and Simonovits [2004, 2009, 2011].

Motivation: Evolution of random graphs

The Erdős-Rényi random graph $G_{n,m}$ is the uniformly chosen random element of the family

$$\mathcal{G}_{n,m} = \{\text{graphs with the vertex set } [n] \text{ and exactly } m \text{ edges}\}.$$

The random graph $G_{n,m}$ shares a lot of properties with its better known "cousin" – the binomial random graph $G(n,p)$ – when $m = p\binom{n}{2}$.

A major part of the theory of random graph is concerned with:

Meta-question (Evolution of random graphs)

Let f be some graph parameter (e.g., f is the chromatic number or f is the characteristic function of some graph property, such as being connected, containing a Hamilton cycle, etc.).

"How does $f(G_{n,m})$ change as m increases from 0 to $\binom{n}{2}$?"

Evolution of H-free graphs

$\mathcal{F}_{n,m}(H) = \{H\text{-free graphs with the vertex set } [n] \text{ and exactly } m \text{ edges}\}$.

Theorem (Osthus, Prömel, Taraz [2003])

Let $m = m(n)$ and $F_{n,m} \in \mathcal{F}_{n,m}(K_3)$ be chosen u.a.r. For every $\varepsilon > 0$,

$$\lim_{n\to\infty} \Pr(F_{n,m} \text{ is bipartite}) = \begin{cases} 1 & \text{if } m = o(n), \\ 0 & \text{if } n/2 \leqslant m \leqslant (1-\varepsilon)m_2, \\ 1 & \text{if } m \geqslant (1+\varepsilon)m_2, \end{cases}$$

where

$$m_2 = m_2(n) = \frac{\sqrt{3}}{4} n^{3/2} \sqrt{\log n}.$$

An analogous result holds for odd cycles, where

$$m(C_{2\ell+1}) = \left(\frac{2\ell+1}{2\ell} \cdot \left(\frac{n}{2}\right)^{2\ell+1} \cdot \log n \right)^{\frac{1}{2\ell}}.$$

Main result

Theorem (Balogh, Morris, S., Warnke [2013+])

For every $r \geqslant 3$ and $\varepsilon > 0$, the following is true. Let $m = m(n)$ and let $F_{n,m} \in \mathcal{F}_{n,m}(K_{r+1})$ be chosen u.a.r. Then

$$\lim_{n\to\infty} \Pr(F_{n,m} \text{ is } r\text{-colorable}) = \begin{cases} 1 & \text{if } m \leqslant (1-\varepsilon)d_r, \\ 0 & \text{if } (1+\varepsilon)d_r \leqslant m \leqslant (1-\varepsilon)m_r, \\ 1 & \text{if } m \geqslant (1+\varepsilon)m_r, \end{cases}$$

where $d_r = d_r(n) = \Theta(n)$ and

$$m_r = m_r(n) = \frac{r-1}{2r} \cdot \left[r \cdot \left(\frac{2r+2}{r+2} \right)^{\frac{1}{r-1}} \right]^{\frac{2}{r+2}} \cdot n^{2-\frac{2}{r+2}} \cdot (\log n)^{\frac{1}{\binom{r+1}{2}-1}}.$$

- The case $r = 3$ was proved earlier by Steger and Warnke [2009].
- The first threshold is essentially due to Achlioptas and Friedgut [1999].

Related work: Turán's theorem in $G(n, p)$

Question (Babai, Simonovits, Spencer [1990])

For what p is the largest K_{r+1}-free subgraph of $G(n, p)$ (a.a.s.) r-colorable?

Theorem (Babai, Simonovits, Spencer [1990])

If $p > 1/2$, then a.a.s. the largest triangle-free subgraph of $G(n, p)$ is bipartite.

Theorem (Brightwell, Panagiotou, Steger [2012])

For every $r \geqslant 2$, there exists $c_r > 0$ such that if $p \geqslant n^{-c_r}$, then a.a.s. the largest K_{r+1}-free subgraph of $G(n, p)$ is r-colorable.

Related work: Turán's theorem in $G(n, p)$

Theorem (DeMarco, Kahn [2013+])

There exists a constant C such that if

$$p \geqslant C\sqrt{\frac{\log n}{n}},$$

then a.a.s. the largest triangle-free subgraph of $G(n, p)$ is bipartite.

Note that $\mathbb{E}[e(G(n,p))] \geqslant \sqrt{\log n/n} \cdot \binom{n}{2} = \Theta(m_2(n))$.

Theorem (DeMarco, Kahn [in preparation])

For every $r \geqslant 3$, there exists a cosntant C_r such that if

$$p \geqslant C_r m_r \binom{n}{2}^{-1},$$

then a.a.s. the largest K_{r+1}-free subgraph of $G(n, p)$ is r-colorable.

The first threshold

If $G_{n,m} \in \mathcal{G}_{n,m}$ is chosen u.a.r., then

$$\Pr(G_{n,m} \supseteq K_{r+1}) \leqslant \mathbb{E}[\#\text{copies of } K_{r+1} \text{ in } G_{n,m}] \approx \binom{n}{r+1}\left(\frac{2m}{n^2}\right)^{\binom{r+1}{2}}.$$

A simple calculation shows that if $m \ll n^{2-2/r}$, then the above is $o(1)$ and consequently a.a. graphs in $\mathcal{G}_{n,m}$ are K_{r+1}-free.

Therefore, if $m \ll n^{2-2/r}$ and $F_{n,m} \in \mathcal{F}_{n,m}(K_{r+1})$ is chosen u.a.r., then

$$\Pr(F_{n,m} \text{ is } r\text{-colorable}) = \Pr(G_{n,m} \text{ is } r\text{-colorable}) + o(1).$$

The existence of the first threshold now follows from the following result:

Theorem (Achlioptas, Friedgut [1999])

For every $r \geqslant 3$, the property of (not) being r-colorable has a sharp threshold in $G_{n,m}$ at $m = d_r$ for some $d_r = d_r(n) = \Theta(n)$.

About the second threshold

We expect that above the threshold a.a. K_{r+1}-free graphs are r-colorable.

If $m \gg n \log n$, then a.a. graphs in $\mathcal{G}_{n,m}(r)$ have a unique r-coloring whose all color classes have size about n/r.

Fix one such balanced coloring $\Pi \approx K(n/r, \ldots, n/r)$ and note that

$$\#\text{graphs properly colored by } \Pi = \binom{e(\Pi)}{m}.$$

We compare this with the number of graphs in $\mathcal{F}_{n,m}(K_{r+1})$ that are not r-colorable but are "almost" properly colored by Π.

We start with graphs with exactly one monochromatic edge. Fix an edge $uv \in \Pi^c$ and let P be the probability that, when we randomly choose $m-1$ edges of Π, the edge uv *does not* lie in a copy of K_{r+1}.

$$P \approx \left(1 - \left(\frac{m}{e(\Pi)}\right)^{\binom{r+1}{2}-1}\right)^{\left(\frac{n}{r}\right)^{r-1}} \approx \exp\left(-\left(\frac{n}{r}\right)^{r-1} \cdot \left(\frac{m}{\left(1-\frac{1}{r}\right)\frac{n^2}{2}}\right)^{\binom{r+1}{2}}\right).$$

About the second threshold

Observe that

$$\#\text{graphs in } \mathcal{F}_{n,m}(K_{r+1}) \text{ with exactly one monochromatic edge in } \Pi$$
$$= \binom{e(\Pi^c)}{1} \cdot \binom{e(\Pi)}{m-1} \cdot P = \Theta(m) \cdot P \cdot \binom{e(\Pi)}{m}$$

Calculation shows that $P = \Theta(1/m)$ exactly when $m = m_r$.

- If $m \leqslant (1-\varepsilon)m_r$, then $P \geqslant m^{-1+\delta}$.
- If $m \geqslant (1+\varepsilon)m_r$, then $P \leqslant m^{-1-\delta}$.

A rigorous version of the above heuristic establishes the 0-statement below the second threshold.

- the FKG inequality for the hypergeometric distribution,
- careful counting (employing some ideas of Prömel and Steger [1992]).

About the second threshold

One of the main tools in the proof of the 1-statement is the following:

Theorem (Balogh, Morris, S. / Saxton, Thomason [2012+])

For every $r \geqslant 2$ and $\delta > 0$, there exists a C such that if $m \geqslant Cn^{2-\frac{2}{r+2}}$, then a.e. graph in $\mathcal{F}_{n,m}(K_{r+1})$ can be made r-colorable by removing from it at most δm edges.

This was previously derived by Łuczak [2000] from the (then unproven) *KŁR conjecture* (proved by BMS and ST).

It follows that one only needs to estimate the number of graphs in $G \in \mathcal{F}_{n,m}(K_{r+1})$ with $o(m)$ monochromatic edges.

This is *more difficult* than counting graphs with *one* monochromatic edge.

Our two main tools are:

- A version of Janson's inequality for the hypergeometric distribution.
- A new concentration inequality for the number of edges induced by a random subset in sparse uniform hypergraphs.

Open problem(s)

We finish with a (natural) conjecture:

Conjecture

For every strictly 2-balanced graph H that contains a color-critical edge, there exists a constant C such that the following holds. If

$$m \geqslant C n^{2-1/m_2(H)} (\log n)^{\frac{1}{e(H)-1}},$$

then a.a. graphs in $\mathcal{F}_{n,m}(H)$ are $(\chi(H)-1)$-partite.

Thank you for your attention!