1 Introduction

For a graph H, we let $\mathrm{ex}(n, H)$ denote the *extremal number* of H, that is, the maximum number of edges in an H-free graph on n vertices. When $\chi(H) > 2$, the Erdős–Stone–Simonovits theorem gives us a precise asymptotic for $\mathrm{ex}(n, H)$, namely

$$\mathrm{ex}(n, H) = \left(1 - \frac{1}{\chi(H) - 1} + o(1)\right)\binom{n}{2}.$$

However, when H is bipartite, all this tells us is that $\mathrm{ex}(n, H) = o(n^2)$, and much of the modern study of extremal numbers is concerned with understanding $\mathrm{ex}(n, H)$ for bipartite H.

Perhaps the most basic result about the extremal numbers of bipartite graphs is the Kővári–Sós–Turán theorem, which gives an upper bond on $\mathrm{ex}(n, K_{s,t})$.

Theorem 1.1 (Kővári, Sós, Turán 1954)**.** *For all integers $t \geq s \geq 1$, we have that $\mathrm{ex}(n, K_{s,t}) = O_t(n^{2-1/s})$.*

It is known that this bound is asymptotically tight if $t \gg s$. The Kővári–Sós–Turán theorem also immediately gives an upper bound on $\mathrm{ex}(n, H)$ for an arbitrary bipartite graph H. Namely, if H has parts of sizes s and t, then $\mathrm{ex}(n, H) \leq \mathrm{ex}(n, K_{s,t}) = O(n^{2-1/\min\{s,t\}})$.

However, we expect that in general, this bound will be very weak. For example, if $H = C_{2k}$ is an even cycle, then Bondy and Simonovits proved that $\mathrm{ex}(n, C_{2k}) = O(n^{1+1/k})$, whereas the reduction to the Kővári–Sós–Turán theorem above would only yield the much weaker bound $\mathrm{ex}(n, C_{2k}) = O(n^{2-1/k})$. The issue, it seems, is that even if a graph has the same vertex set as $K_{s,t}$, it may have many fewer edges, and thus we would expect its extremal number to be much smaller.

One of the most important results confirming this intuition is the following result of Füredi.

Theorem 1.2 (Füredi 1991)**.** *Let H be a bipartite graph in which all the vertices on one side have degree at most r. Then $\mathrm{ex}(n, H) = O_H(n^{2-1/r})$.*

This is a huge improvement over the Kővári–Sós–Turán bound in case H is sparse, as the exponent on n depends on the *maximum degree* of a side, rather than the number of vertices in that side. Also, as the Kővári–Sós–Turán bound on $\mathrm{ex}(n, K_{s,t})$ is known to be sharp when $t \gg s$, we see that Theorem 1.2 is best possible in general.

About a decade later, Theorem 1.2 was reproved by Alon, Krivelevich, and Sudakov, using the technique of dependent random choice, and this is the proof that I will present. We begin with the following lemma, which is a standard consequence of the dependent random choice technique.

Lemma 1.3. *For all positive integers a, b, there exists some constant $C > 0$ such that the following holds. Let G be an n-vertex graph with average degree $d \geq Cn^{1-1/r}$. Then there exists $U \subseteq V(G)$ with $|U| \geq a$ so that every r-tuple of vertices in U has at least $a+b$ common neighbors.*

Proof. Let $x_1, \ldots, x_r$ be iid uniformly random vertices of G, and let $\mathbf{X}$ be the common neighborhood of $x_1, \ldots, x_r$. By linearity of expectation and Jensen's inequality, we have

$$\mathbb{E}[|\mathbf{X}|] = \sum_{u \in V(G)} \Pr(u \in \mathbf{X}) = \sum_{u \in V(G)} \left(\frac{\deg(u)}{n}\right)^r \geq n \left(\frac{d}{n}\right)^r = \frac{d^r}{n^{r-1}} \geq C^r.$$

Now, let $\mathbf{Y}$ count the number of r-tuples of vertices in $\mathbf{X}$ with fewer than $a+b$ common neighbors. There are at most $\binom{n}{r}$ total such r-tuples, and each one appears in $\mathbf{X}$ with probability less than $((a+b)/n)^r$. So we have that

$$\mathbb{E}[\mathbf{Y}] < \binom{n}{r} \left(\frac{a+b}{n}\right)^r < (a+b)^r.$$

Therefore, we see that $\mathbb{E}[|\mathbf{X}| - \mathbf{Y}] > C^r - (a+b)^r \geq a$, by picking C sufficiently large with respect to a and b. Therefore, there exists some set $X \subseteq V(G)$ so that if Y denotes the number of r-tuples in X with fewer than $a+b$ common neighbors, then $|X| - Y \geq a$. By deleting one vertex from every "bad" r-tuple in X, we obtain a set U with the desired properties. □

Using this lemma, we can give the short proof of Theorem 1.2 due to Alon, Krivelevich, and Sudakov.

Proof of Theorem 1.2. Let the bipartition of H be $A \cup B$ where $|A| = a, |B| = b$, and suppose that every vertex in B has degree at most r. Let C be the constant from Lemma 1.3, and let G be an n-vertex graph with at least $Cn^{2-1/r}$ edges. Then G has average degree $d \geq Cn^{1-1/r}$, so by Lemma 1.3, we may find a set $U \subseteq V(G)$ with $|U| \geq a$ and so that every r-tuple of vertices in U has at least $a+b$ common neighbors. We can now greedily embed H in G, as follows. We first arbitrarily embed the vertices of A into U. Now, we go through the vertices of B one by one. For a given vertex $v \in B$, it has at most r neighbors in A, and the embeddings of those neighbors have at least $a+b$ common neighbors. So there are at least $a+b$ options for how to embed v, and fewer than $a+b$ of these have been used by previously-embedded vertices. Thus, there is at least one valid choice for v, and by iterating this argument, we can embed H in G. This shows that $\mathrm{ex}(n,H) \leq Cn^{2-1/r}$, as claimed. □

Recall that a graph H is called *r-degenerate* if every subgraph of H has a vertex of degree at most r. Equivalently, H is r-degenerate if and only if there is a linear ordering of its vertices so that every vertex has at most r neighbors which precede it. The *degeneracy* of H is the minimum r for which H is r-degenerate. Erdős conjectured the following strengthening of Theorem 1.2.

Conjecture 1.4 (Erdős)**.** *If H is bipartite and r-degenerate, then* $\mathrm{ex}(n,H) = O_H(n^{2-1/r})$.

If true, then Conjecture 1.4 immediately implies Theorem 1.2, as a bipartite graph in which one side has maximum degree r is clearly r-degenerate. However, it is much stronger; for example, a double star (a tree consisting of an edge plus many pendant edges on each of its endpoints) is 1-degenerate, but can have arbitrarily high degrees on both sides.

Conjecture 1.4 remains open, but the following approximate form of it was proved by Alon, Krivelevich, and Sudakov.

Theorem 1.5 (Alon, Krivelevich, Sudakov 2003)**.** *If H is bipartite and r-degenerate, then* $\mathrm{ex}(n, H) = O(n^{2-1/(4r)})$.

The proof of Theorem 1.5 is similar to the proof of Theorem 1.2 presented above, but with an extra twist. Namely, one applies the dependent random choice technique *twice* in order to find in G two sets U_1, U_2 so that every r-tuple of vertices in U_1 has many common neighbors in U_2, and similarly every r-tuple of vertices in U_2 has many common neighbors in U_1. We can then greedily embed H in G by arranging the vertices of H according to the degenerate ordering, and one by one placing vertices of A in U_1 and vertices of B in U_2. However, the fact that we have to apply dependent random choice twice means that we lose something, which is why the exponent in Theorem 1.5 does not match Conjecture 1.4.

Before moving on, we remark that Conjecture 1.4 and Theorem 1.5 are essentially best possible for every graph H. This follows from the following lower bound, which is proved using the method of alterations. Recall that the 2-density of a graph H is defined by

$$m_2(H) = \max_{F \subseteq H} \frac{e(F) - 1}{v(F) - 2},$$

where the maximum is taken over all subgraphs F of H with at least three vertices.

Proposition 1.6. *For any graph H, we have*

$$\mathrm{ex}(n, H) = \Omega_H(n^{2-1/m_2(H)}).$$

In particular, if H has degeneracy r, then

$$\Omega_H\left(n^{2-\frac{2}{r}}\right) \le \mathrm{ex}(n, H) \le O_H\left(n^{2-\frac{1}{4r}}\right).$$

Proof sketch. Let $p = cn^{-1/m_2(H)}$ for a constant $c > 0$ to be chosen later, and let $G \sim G(n, p)$ be an Erdős–Rényi random graph. With high probability, G has $\Theta(cn^{2-1/m_2(H)})$ edges. Let $F \subseteq H$ be a subgraph so that

$$m_2(H) = \frac{e(F) - 1}{v(F) - 2}.$$

The expected number of copies of F in G is at most $n^{v(F)}p^{e(F)} = c^{e(F)}n^{v(F)-e(F)/m_2(H)}$. Note that

$$\begin{aligned}
v(F) - \frac{e(F)}{m_2(H)} &= v(F) - \frac{e(F)(v(F) - 2)}{e(F) - 1} \\
&= \frac{v(F)(e(F) - 1) - e(F)(v(F) - 2)}{e(F) - 1} \\
&= \frac{2e(F) - v(F)}{e(F) - 1} \\
&= 2 - \frac{v(F) - 2}{e(F) - 1} \\
&= 2 - \frac{1}{m_2(H)}.
\end{aligned}$$

Therefore, the expected number of copies of F in G is at most $c^{e(F)}n^{2-1/m_2(H)}$. By picking c sufficiently small, we see that this quantity is at most half the number of edges of G. So by deleting one edge from every copy of F in G, we obtain an F-free graph G' with $\Theta(n^{2-1/m_2(H)})$ edges. Since G' is F-free and $F \subseteq H$, we see that G' is H-free, and thus $\mathrm{ex}(n, H) = \Omega_H(n^{2-1/m_2(H)})$.

For the second claim, the upper bound is simply Theorem 1.5. For the lower bound, note that if H has degeneracy r, then there is a subgraph F of H with minimum degree at least r. This subgraph F then has at least $\frac{r}{2}v(F)$ edges. This implies that

$$m_2(H) \geq \frac{e(F)-1}{v(F)-2} \geq \frac{\frac{r}{2}v(F)-1}{v(F)-2} = \frac{r}{2} \cdot \frac{v(F)-\frac{2}{r}}{v(F)-2} \geq \frac{r}{2}.$$

Thus,

$$\mathrm{ex}(n, H) = \Omega_H(n^{2-1/m_2(H)}) = \Omega_H(n^{2-2/r}),$$

as claimed. □

2 Subdivisions

Recall Theorem 1.2, which says that $\mathrm{ex}(n, H) = O(n^{2-1/r})$ if H is bipartite and every vertex on one side of H has degree at most r. As mentioned, this bound is tight in general, since $\mathrm{ex}(n, K_{s,t}) = \Theta(n^{2-1/s})$ if $t \gg s$. In fact, it has long been conjectured that the Kővári–Sós–Turán bound is tight even for $s = t$, i.e. that $\mathrm{ex}(n, K_{s,s}) = \Theta(n^{2-1/s})$. The truth of this has been known for decades for $s \in \{2, 3\}$, and nowadays, some people think that it may be false in general. But in any case, one might expect that "the reason" why Theorem 1.2 is tight in general is the presenence of copies of $K_{r,r}$. This is the motivation for the following conjecture of Conlon and Lee.

Conjecture 2.1 (Conlon, Lee 2021)**.** *Suppose H is a bipartite graph with every vertex on one side having degree at most r. If $K_{r,r} \nsubseteq H$, then* $\mathrm{ex}(n, H) = O(n^{2-1/r-\varepsilon})$ *for some* $\varepsilon > 0$.

This conjecture remains open for all $r \geq 3$, but Conlon and Lee proved it for $r = 2$. Before stating their result, let us think about which $K_{2,2}$-free bipartite graphs have maximum degree 2 on one side. Let H be such a graph with bipartition $A \cup B$, and let every vertex in B have degree at most 2. By adding dummy vertices to A, we may assume that actually every vertex in B has degree 2. Then the fact that H is $K_{2,2}$-free means that every pair of vertices in A has at most a single common neighbor. If we define a graph Γ on A by connecting those pairs with a common neighbor in B, we see that H is a subgraph of the 1-*subdivision* of Γ. Since Γ is a subgraph of $K_{|A|}$, we conclude that every $K_{2,2}$-free bipartite graph in which one side has maximum degree 2 is a subgraph of $\widehat{K_t}$ for some t, where $\widehat{\Gamma}$ denotes the 1-subdivision of Γ.

In this language, Conlon and Lee proved that $\mathrm{ex}(n, \widehat{K_t}) = O(n^{3/2-6^{-t}})$, which implies Conjecture 2.1 in the case $r = 2$. Their result was shortly afterwards strengthened by Janzer, who proved the following theorem.

Theorem 2.2 (Janzer 2019)**.** *For every $t \geq 3$, we have* $\mathrm{ex}(n, \widehat{K_t}) = O\left(n^{\frac{3}{2}-\frac{1}{4t-6}}\right) = O\left(n^{1+\frac{t-2}{2t-3}}\right)$.

Proof. We pick an absolute constant $C > 0$ to be chosen later. Let G be a bipartite graph with bipartition $A \cup B$, where $|A| = |B| = n$, and suppose that G is d-regular, where $d = Cn^{\frac{t-2}{2t-3}}$. We will show that G contains a copy of $\widehat{K_t}$. Note that we made two simplifying assumptions, namely that G is balanced bipartite, and that G is d-regular. Both of these assumptions are essentially without loss of generality: every graph contains a balanced bipartite subgraph with at least half its edges, and a useful lemma of Erdős and Simonovits shows that for extremal problems, we may assume the host graph is *almost regular*, meaning that the minimum and maximum degrees are within a constant factor of each other. But for simplicity, I will just assume that G is d-regular.

For vertices $u, v \in A$, let $\operatorname{codeg}(u, v)$ denote the number of common neighbors of u and v in B. For a pair of vertices $u, v \in A$, we say that (u, v) is *heavy* if $\operatorname{codeg}(u, v) \geq \binom{t}{2}$, and we say that (u, v) is *light* if $1 \leq \operatorname{codeg}(u, v) < \binom{t}{2}$. The first key observation is that if we can find $u_1, \ldots, u_t \in A$ so that (u_i, u_j) is heavy for all $i \neq j$, then G contains a copy of $\widehat{K_t}$. Indeed, each pair (u_i, u_j) has at least $\binom{t}{2}$ common neighbors, so by the same greedy embedding argument as above, we can find a copy of $\widehat{K_t}$.

To say this differently, let's define an edge-weighted graph W_G with vertex set A, in which two vertices are adjacent if they have at least one common neighbor in B. Additionally, we assign the edge (u, v) weight $\operatorname{codeg}(u, v)$. Then the argument above shows that if W_G contains a copy of K_t consisting of heavy edges, then $\widehat{K_t} \subseteq G$. Because of this, we henceforth assume that there is no heavy K_t in W_G (i.e. a K_t containing only heavy edges). The idea now is that a *light* K_t in W_G can also be helpful for us: two different light edges both have few common neighbors, so we can hope to find pairs of light edges whose common neighborhoods are disjoint. By doing this carefully, we can find a light K_t in W_G such that distinct pairs of vertices have disjoint common neighborhoods, and then we can embed $\widehat{K_t}$ by using such a light K_t. In other words, while a heavy K_t is very helpful for us, a light K_t is also helpful for an essentially complementary reason. We now show have to find such a "good" light K_t. The key lemma which enables us to do this is the following, which shows that any large subset of A contains many light edges.

Lemma 2.3. *If $U \subseteq A$ satisfies $|U| \geq 2tn/d$, then the number of light edges in U is at least $\frac{d^2}{t^3 n}\binom{|U|}{2}$.*

Proof. Fix some $b \in B$. Every pair of vertices in $N_U(b)$ has at least one common neighbor (namely b), so every pair in $N_U(b)$ is either a light or a heavy edge. Additionally, there can be no heavy K_t in $N_U(b)$. So by Turán's theorem, the number of light edges in $N_U(b)$ is at least

$$(t-1)\binom{\frac{\deg_U(b)}{t-1}}{2}.$$

Now note that every light edge can appear in $N_U(b)$ for at most $\binom{t}{2}$ choices of b, by the definition of light. Therefore, the total number of light edges is at least

$$\frac{t-1}{\binom{t}{2}} \sum_{b \in B} \binom{\frac{\deg_U(b)}{t-1}}{2} \geq \frac{2n}{t}\binom{\frac{1}{n}\sum_{b \in B}\frac{\deg_U(b)}{t-1}}{2} \geq \frac{2n}{t}\binom{\frac{d|U|}{nt}}{2} \geq \frac{n}{t} \cdot \left(\frac{d}{nt}\right)^2 \binom{|U|}{2} = \frac{d^2}{t^3 n}\binom{|U|}{2},$$

where the first inequality uses the convexity of the binomial coefficient, the second inequality uses the fact that $\sum_{b\in B}\deg_U(b) = d|U|$ since both count the number of edges incident to U, and the third inequality uses the observation that if $\lambda x \geq 2$, then $\binom{\lambda x}{2} \geq \frac{\lambda^2}{2}\binom{x}{2}$, as well as the assumption $|U| \geq 2nt/d$. $\square$

We are now ready to complete the proof. The idea is to find a collection $u_1, \ldots, u_t$ of vertices in A with the property that all pairs (u_i, u_j) are light, and such that $N(u_i) \cap N(u_j) \cap N(u_k) = \varnothing$ for all distinct i, j, k. Because all pairs are light, every pair u_i, u_j has at least one common neighbor in B, and these common neighbors are distinct across all $\binom{t}{2}$ pairs, thanks to the second condition. Therefore, by choosing one such common neighbor for every pair (u_i, u_j), we find a copy of $\widehat{K_t}$ in G.

In order to do this, we will pick out vertices one by one. We will also maintain a sequence $A = S_1 \supseteq S_2 \supseteq \cdots$ of "candidate vertices", where S_i is the set of all possible vertices we can choose to be u_i. Namely, S_i is the set of vertices v so that (u_j, v) is light for all $1 \leq j \leq i-1$. We will inductively maintain the property that

$$|S_i| \geq \left(\frac{d^2}{4t^3 n}\right)^{i-1} n.$$

Note that this holds trivially for $i = 1$, since $|S_1| = |A| = n$. To begin the induction, note that by Lemma 2.3, the number of light edges in S_1 is at least $\frac{d^2}{t^3 n}\binom{|S_1|}{2}$, so there is a vertex $u_1 \in S_1$ incident to at least $\frac{d^2}{t^3 n}(|S_1| - 1) \geq (\frac{d^2}{4t^3 n})^1 n$ light edges. We let S_2 be this set of light neighbors of u_1.

Inductively, suppose we've defined $u_1, \ldots, u_{i-1}$, and have a set S_i of candidate vertices, i.e. (u_j, v) is light for all $v \in S_i$ and $j \leq i-1$. Let $U \subseteq S_i$ consist of those vertices v with $N(u_j) \cap N(u_k) \cap N(v) = \varnothing$ for all $1 \leq j < k \leq i-1$. Since (u_j, u_k) is light, we know that $\operatorname{codeg}(u_j, u_k) < \binom{t}{2}$. Additionally, if b is a common neighbor of u_j and u_k, then b has at most d neighbors in S_i. So the total number of vertices in S_i that are "ruled out" by (u_j, u_k) is at most $\binom{t}{2}d$. Adding this up over all pairs j, k, we see that

$$|S_i \setminus U| \leq \binom{i-1}{2}\binom{t}{2} d = O_t(d).$$

However, we also know that

$$|S_i| \geq \left(\frac{d^2}{4t^3 n}\right)^{i-1} n \geq \left(\frac{d^2}{4t^3 n}\right)^{t-1} n = \Omega_t\left(\frac{d^{2t-2}}{n^{t-2}}\right).$$

Since $d = Cn^{\frac{t-2}{2t-3}}$, by choosing C large enough, we can guarantee that $|U| \geq \frac{1}{2}|S_i|$ for sufficiently large n.

If $i = t$, then we may simply pick an arbitrary $u_t \in U$ and be done. If $i \leq t-1$, then the same computation shows that

$$|U| \geq \frac{1}{2}|S_i| \geq \left(\frac{d^2}{4t^3 n}\right)^{t-2} n = \Omega_t\left(\frac{d^{2t-4}}{n^{t-3}}\right).$$

On the other hand, $2tn/d = O_t(n/d)$, so by picking C large, we can ensure that $|U| \geq 2tn/d$, so we are in the position to apply Lemma 2.3. We conclude that U contains at least $\frac{d^2}{t^3 n}\binom{|U|}{2}$ light edges, so there is a vertex $u_i \in U$ incident to at least $\frac{d^2}{t^3 n}(|U|-1) \geq \frac{d^2}{4t^3 n}|S_i|$ light edges. By letting S_{i+1} be this set of light neighbors of u_i, we conclude the induction. $\square$

Though Janzer's bound is somewhat strange, there is actually good reason to expect that it is tight. One reason is that it is tight in case $t = 3$, as $\widehat{K_3} = C_6$, and it is known that $\mathrm{ex}(n, C_6) = \Theta(n^{4/3})$. Another reason is a little harder to explain, since it requires setup. A famous result of Alon is that there exist "optimally pseudorandom triangle-free graphs". Namely, there exists a family of n-vertex triangle-free graphs with average degree $\Theta(n^{2/3})$ and second eigenvalue $\Theta(n^{1/3})$. This is called optimally pseudorandom, because in any graph, the second eigenvalue must be at least roughly the square root of the average degree, and the smaller the second eigenvalue is, the more pseudorandom the graph is. Additionally, this is optimal in another way: one can prove that a graph with average degree $\Theta(n^{2/3+\varepsilon})$ and second eigenvalue $\Theta(n^{1/3+\varepsilon/2})$ must contain a triangle, for any $\varepsilon > 0$. So Alon's graphs are as dense as possible, given the constraints that they be triangle-free and optimally pseudorandom.

It is a major open problem to construct, analogously, optimally pseudorandom K_t-free graphs for any $t \geq 4$. If such graphs exist, they would have numerous applications. Perhaps most notably, due to a result of Mubayi and Verstraëte, if an optimally K_t-free pseudorandom graph exists, then the off-diagonal Ramsey number $r(t, n)$ grows as $n^{t-1-o(1)}$, matching up to lower-order terms the Erdős–Szekeres upper bound from 1935. It turns out that an optimally pseudorandom K_t-free graph, if it exists, has average degree $\Theta(n^{1-1/(2t-3)})$; any denser one must contain a K_t, as above.

In addition to Alon's original construction, we now know of several other ways of constructing optimally pseudorandom triangle-free graphs. One of the ways, due to Conlon, is as follows. We begin with a C_6-free bipartite graph G with $\Theta(n^{4/3})$ edges, which is known to exist. In fact, since constructions for such graphs come from discrete geometry, we can even pick such a graph which is itself very pseudorandom. Now, if the bipartition of G is $A \cup B$, we build a new graph F on vertex set A as follows. For every $b \in B$, we randomly partition $N(b)$ in two, and put a complete bipartite graph between these two halves. Since G is C_6-free[1], one can check that F is triangle-free; basically, since $C_6 = \widehat{K_3}$, and since we are essentially "un-subdividing" G, the C_6-freeness of G guarantees the triangle-freeness of F. Additionally, since G was very pseudorandom and since we used these random partitions, one can check that F is pseudorandom[2] as well. Finally, for a given vertex $a \in A$, it has $\Theta(n^{1/3})$ neighbors $b \in B$, and each of these yield $\Theta(n^{1/3})$ neighbors of a in $F[N(b)]$, so F has average degree $\Theta(n^{2/3})$.

Now, suppose that Janzer's bound in Theorem 2.2 is tight, and suppose too that we had some construction of a bipartite graph G with $\Theta\left(n^{1+\frac{t-2}{2t-3}}\right)$ edges, no copy of $\widehat{K_t}$, and such that G is pseudorandom (as it likely is, if it somehow "algebraic" in nature). By "un-subdividing" G à la Conlon, we would obtain a pseudorandom K_t-free graph with average

[1]Really, for this to work, we need G to be $\{C_4, C_6\}$-free. But most of the constructions coming from discrete geometry are indeed $\{C_4, C_6\}$-free.

[2]Actually, Conlon's proof yields that F is slightly worse than optimally pseudorandom, but it is still pseudorandom enough for all the applications of such graphs (and he conjectured that the problem is his analysis, and that the construction actually is optimally pseudorandom).

degree $\Theta\left(n^{2\cdot\frac{t-2}{2t-3}}\right) = \Theta(n^{1-1/(2t-3)})$. Thus, a matching lower bound for Theorem 2.2 could yield optimally pseudorandom K_t-free graphs, and this is one reason to expect (or hope) that Theorem 2.2 is tight.

3 Beyond $r = 2$

We have seen that Conjecture 2.1 is true for $r = 2$, namely that we can get a power saving on the bound $\mathrm{ex}(n, H) = O(n^{2-1/r})$ in case H has maximum degree r on one side but does not contain a copy of $K_{r,r}$. Conjecture 2.1 remains open for all $r \geq 3$, but in this section, I will present the following result which proves a weak version of Conjecture 2.1 for all r.

Theorem 3.1 (Sudakov, Tomon 2020)**.** *Let H be a bipartite graph with every vertex on one side having degree at most r, and suppose that $K_{r,r} \nsubseteq H$. Then $\mathrm{ex}(n, H) = o(n^{2-1/r})$.*

Proof. For $k \geq r$, let H_k be a bipartite graph whose first part X has k vertices, whose second part Y has $(r-1)\binom{k}{r}$ vertices, and where for every $S \in \binom{X}{r}$, there are exactly $r-1$ vertices in Y whose neighborhood is S. Every H as in the theorem statement is a subgraph of H_k for some k, so it suffices to prove that $\mathrm{ex}(n, H_k) = o(n^{2-1/r})$ for any fixed k. This is basically the same argument that allowed us to only work with $\widehat{K_t}$ when proving Conjecture 2.1 in the case $r = 2$.

Let G be a bipartite graph with partition $A \cup B$, where $|A| = |B| = n$, and suppose that G is d-regular where $d = \varepsilon n^{1-1/r}$. We think of $\varepsilon > 0$ as fixed, and want to show that if n is sufficiently large, then $H_k \subseteq G$. As before, our simplifying assumptions are essentially without loss of generality.

Our goal is now to "densify" the problem. To do so, we will pass to an induced subgraph by restricting A to some subset $U \subseteq A$, and we want to do this in such a way that on average, the common neighborhood in U of an $(r-1)$-tuple of vertices in B is a large constant. For a set $C \subseteq B$, let $N_U(C)$ denote the common neighborhood of C in U, and let $L = \sum_{C\in\binom{B}{r-1}} |N_U(C)|$. Our goal is to ensure that $L \geq Mn^{r-1}$, where M is the two-color hypergraph Ramsey number of $K_k^{(r)}$, i.e. every two-coloring of the edges of $K_M^{(r)}$ contains a monochromatic $K_k^{(r)}$.

To do so, for any $p \in (n^{-1/r}, 1)$, suppose we sample a p-random subset $U \subseteq A$. Note that

$$L = \sum_{u\in U} \binom{d}{r-1} = \sum_{u\in U} \binom{\varepsilon n^{1-1/r}}{r-1} \geq |U| \cdot \left(\frac{\varepsilon}{r-1}\right)^{r-1} n^{r-2+1/r} = c_{\varepsilon,r}|U|n^{r-2+1/r}.$$

Now, we pick $p = (2M/c_{\varepsilon,r})n^{-1/r}$. Then with high probability, we have that $|U| \approx pn = (2M/c_{\varepsilon,r})n^{1-1/r}$, and thus that $L \geq Mn^{r-1}$. We also have that with high probability, every vertex in B has degree $\approx pd$ into U. For simplicity, we assume that $|U| = pn$ and that every vertex in B has degree exactly pd into U.

Now, let W be the r-uniform hypergraph with vertex set U in which an r-tuple is an edge if and only if it has at least $r-1$ common neighbors in B. Let us call an edge of W *heavy* if it has at least $(r-1)\binom{k}{r}$ common neighbors in B, and *light* otherwise. As in the proof of Theorem 2.2, we see that if there is a copy of $K_k^{(r)}$ in W made up of heavy edges, then G

contains H_k, as we may embed one side of H_k in this $K_k^{(r)}$, and greedily embed the other side in the appropriate common neighborhoods in B. Therefore, we henceforth assume that there is no heavy $K_k^{(r)}$ in W.

Our goal now, as in the proof of Theorem 2.2, is to find very many light copies of $K_k^{(r)}$ in W. By doing so, we will be able to find a specific light copy of $K_k^{(r)}$ so that all pairs of edges have disjoint common neighborhoods. Once we find that, we again see that $H_k \subseteq G$.

For a fixed $C \in \binom{B}{r-1}$, let $D = N_U(C)$. Note that every r-tuple of vertices in D has at least $r-1$ common neighbors in B (namely the set C), so D induces a clique in the hypergraph W. We partition D into $|D|/M$ sets of size M, each of which is itself a clique in W. Since M is the hypergraph Ramsey number of $K_k^{(r)}$, and since we have colored the edges of each such clique "light" or "heavy", and since there is no heavy $K_k^{(r)}$ in W, we conclude that D contains at least $|D|/M$ disjoint light copies of $K_k^{(r)}$. Let Z_C be the set of these light copies, and let $Z = \bigcup_C Z_C$ be the multiset of all light copies that arise in this way. Then we have that $|Z_C| \geq |D|/M$, so

$$|Z| = \sum_{C \in \binom{B}{r-1}} |Z_C| \geq \sum_{C \in \binom{B}{r-1}} \frac{|N_U(C)|}{M} = \frac{L}{M} \geq n^{r-1}.$$

Now, we form an auxiliary graph Γ with vertex set Z, where we connect $S, T \in Z$ if $|S \cap T| \geq r$. We now claim that the maximum degree in Γ is at most $\binom{k}{r}\binom{u}{k-1}$, where $u = (r-1)\binom{k}{r}$, i.e. that Γ has maximum degree $O_{k,r}(1)$. To see this, fix some $S \in Z$, and let $R \subseteq S$ be an r-subset. Since S spans a light clique, R must form a light edge in W, so the common neighborhood of R has size at most u. So there are at most $\binom{u}{r-1}$ sets C so that $R \subseteq N_U(C)$, and for each such C, at most one element of Z_C contains R. Since there are $\binom{k}{r}$ choices for $R \subseteq S$, the degree of S in Γ is at most $\binom{k}{r}\binom{u}{r-1}$, as claimed. Therefore, Γ contains an independent set Z' of order

$$\frac{|Z|}{\binom{k}{r}\binom{u}{r-1} + 1} \geq \delta_0 n^{r-1} \geq \delta_1 |U|^r,$$

for constants δ_0, δ_1 that do not depend on n. In the final step, we used the fact that $|U| = pn = \Omega(n^{1-1/r})$. Note too that since Z' is an independent set in Γ, every pair $S, T \in Z'$ intersect on fewer than r elements, and thus Z' consists of edge-disjoint copies of $K_k^{(r)}$ in W.

We now recall the hypergraph removal lemma, which says (in one of its equivalent forms) that if an r-uniform hypergraph W on vertex set U contains at least $\delta_1|U|^r$ edge-disjoint copies of $K_k^{(r)}$, then W must contain at least $\gamma|U|^k$ total copies of $K_k^{(r)}$, for some γ depending only on δ_1, k, and r. Therefore, we conclude that W contains at least $\gamma|U|^k$ copies of $K_k^{(r)}$.

Finally, let's say that a copy Q of $K_k^{(r)}$ in W is *bad* if it contains two r-tuples $R, R' \subseteq Q$ with $N(R) \cap N(R') \neq \varnothing$. If we can prove that not all copies are bad, then we have found a copy of $K_k^{(r)}$ in which all pairs of edges of have disjoint common neighborhoods, and thus we can embed a copy of H_k in G. So to conclude, it suffices to count the number of bad copies and show that it is less than $\gamma|U|^k$. Note that

$$\gamma|U|^k = \gamma(pn)^k = c(n^{1-1/r})^k = cn^{k-k/r}$$

for some constant c independent of n.

Fix a bad copy Q of $K_k^{(r)}$, and let $R, R' \subseteq Q$ be edges with non-disjoint common neighborhoods. Let $b \in N(R) \cap N(R')$, and note that $|N_U(b) \cap Q| \geq |R \cup R'| \geq r+1$. In other words, a bad copy Q of $K_k^{(r)}$ is witnessed by a vertex $b \in B$ with $|N_U(b) \cap Q| \geq r+1$. By summing over all witnessing vertices in B, we conclude that the total number of bad copies of $K_k^{(r)}$ in W is at most

$$\sum_{b \in B} \binom{pd}{r+1} |U|^{k-r-1} \leq n \cdot (pd)^{r+1} (pn)^{k-r-1} = p^k d^{r+1} n^{k-r} = Cn^{k-k/r-1/r}$$

for some constant C independent of n. For sufficiently large n, this is less than $cn^{k-k/r}$, showing that $H_k \subseteq G$ for sufficiently large n. $\square$