The Saturation Function of Complete Partite Graphs

Tom Bohman[1,3]   Maria Fonoberova[2]   Oleg Pikhurko[1,4]

August 30, 2010

Abstract

A graph G is called F-*saturated* if it is F-free but the addition of any missing edge to G creates a copy of F. Let the *saturation function* $\mathrm{sat}(n, F)$ be the minimum number of edges that an F-saturated graph on n vertices can have. We determine this function asymptotically for every fixed complete partite graph F as n tends to infinity and give some structural information about almost extremal F-saturated graphs. If the two largest parts of F have different sizes, then we can reduce the error term to an additive constant.

1   Introduction

Let K_r denote the complete graph on r vertices. For $r \geq 2$ and positive integers $s_1, \ldots, s_r$, let $K_{s_1, \ldots, s_r}$ denote the complete r-partite graph with part sizes $s_1, \ldots, s_r$, that is, its edges are precisely those pairs that connect two different parts. The *join* $F * G$ of graphs F and G is obtained by taking the disjoint union of F and G and adding all edges between them. Let $\overline{G}$ denote the complement of a graph G. Thus $K_r = K_{1, \ldots, 1}$ and $K_{s_1, \ldots, s_r} = \overline{K}_{s_1} * \cdots * \overline{K}_{s_r}$.

A graph G is F-*saturated* if G does not contain F as a subgraph (not necessarily induced) but the addition of any new edge to G creates at least one copy of F. In other words, G is F-saturated if and only if it is a maximal F-free graph.

The *saturation function* sat(n, F) is the minimum number of edges that an F-saturated graph on n vertices can have. This can be viewed as the dual function of the celebrated *Turán function* ex(n, F), the maximum number of edges in an F-saturated graph of order n. One crucial difference is that, for any fixed F, sat(n, F) is bounded by a linear function of n (as was shown by Kászonyi and Tuza [KT86]) while ex(n, F) may be quadratic in n.

The structure and extremal properties of clique-saturated graphs were studied already in the 1940s by Zykov [Zyk49]. The *saturation problem* (to determine the saturation function and describe all extremal graphs) for K_r was completely solved by Erdős, Hajnal, and Moon [EHM64]. They used their result to prove a conjecture of Erdős and Gallai [EG61] about graphs that are critical with respect to edge domination. It turns out that $K_{r-2} * \overline{K}_{n-r+2} \cong K_{1,\ldots,1,n-r+2}$ is the unique minimum K_r-saturated graph for every $n \geq r$. Thus

$$\text{sat}(n, K_r) = \binom{r-2}{2} + (r-2)(n-r+2), \qquad n \geq r. \tag{1}$$

One direction of extending the above result is to view K_3 as the 3-cycle and consider the general m-cycle C_m. Ollmann [Oll72] determined the saturation function for the 4-cycle and described all extremal graphs. (A shorter proof was found by Tuza [Tuz89].) Very recently, Chen [Che09, Che10] completely solved the case of C_5. The value of sat(n, C_m) is unknown (even asymptotically) for any other fixed m; various bounds are proved in [BCE$^+$96, GŁS06]. Also, the aggregate of papers [Bon72, Isa75, CE83, CCES86, CES92, XWCY97], with the final gaps filled by computer search, determines sat(n, C_n), the Hamilton cycle case, for all $n \geq 3$.

Another direction of research motivated by the result of Erdős, Hajnal, and Moon [EHM64] is to view the clique K_r as $K_{1,\ldots,1}$ and consider the saturation problem for complete partite graphs. The case of a *star* $K_{1,s}$ is easy, see [KT86]. Pikhurko [Pik04] computed the saturation function of $K_t * \overline{K_s} = K_{1,\ldots,1,s}$ exactly when $n \geq n_0(s, t)$. Unaware of this result, Chen, Faudree, and Gould [CFG08] reproved it, with explicit estimates on $n_0(s, t)$. Pikhurko and Schmitt [PS08] determined the asymptotic of sat$(n, K_{2,3})$. Gould and Schmitt [GS07] stated a conjecture about sat$(n, K_{2,\ldots,2})$ and proved it under some additional assumption.

We should also mention a related conjecture of Erdős, Hajnal, and Moon

[EHM64] about the minimum number of edges in a bipartite $K_{s,t}$-free graph $G \subseteq K_{n,m}$ such that the addition of any new edge between the parts creates at least one copy of $K_{s,t}$. The conjecture was proved independently by Wessel [Wes66] and Bollobás [Bol67b], while all minimum graphs were characterized in [Bol67a, Wes67].

In this paper we determine the asymptotic of $\text{sat}(n, K_{s_1,\ldots,s_r})$ as $n \to \infty$ for **every** sequence $s_1 \le \cdots \le s_r$. In order to state our results, it will be convenient to denote

$$p = s_1 + \cdots + s_{r-1} - 1. \tag{2}$$

Theorem 1 (Asymptotic Result) *Let* $r \ge 2$, $s_r \ge \cdots \ge s_1 \ge 1$, *and* $F = K_{s_1,\ldots,s_r}$. *Define* p *by (2). Then, for all large* n,

$$\left(p + \frac{s_r - 1}{2}\right) n \;-\; O(n^{3/4}) \;\le\; \text{sat}(n, F)$$

$$\le\; \binom{p}{2} + p(n-p) + \left\lceil \frac{(s_r-1)(n-p)}{2} - \frac{s_r^2}{8} \right\rceil. \tag{3}$$

In particular,

$$\text{sat}(n, K_{s_1,\ldots,s_r}) = \left(s_1 + \cdots + s_{r-1} + \frac{s_r - 3}{2}\right) n + O(n^{3/4}).$$

An example of an F-saturated graph can be obtained by taking the join of K_p and a K_{1,s_r}-saturated graph H, see Construction 1 in Section 3 for details.

Remarkably, our proof (with some extra work) shows not only that the above construction is asymptotically best possible but also allows us to describe the structure of all almost extremal graphs when $s_{r-1} < s_r$. It turns out that any F-saturated graph on n vertices with at most $\text{sat}(n, F) + o(n)$ edges can be transformed into the above example $K_p * H$ by adding and removing at most $o(n)$ edges. Informally speaking, this says that all almost extremal graphs have almost the same structure. This *stability property* is important on its own and is often very helpful for getting more precise estimates or in applications. Also, we obtain some structural information when $s_{r-1} = s_r$. It is convenient to partition all sequences $s_1 \le \cdots \le s_r$ into three families:

Case I: $s_r > s_{r-1}$,

Case II: $s_r = 1$ (and thus $s_1 = \cdots = s_{r-1} = 1$),

Case III: $s_r = s_{r-1} \geq 2$.

Theorem 2 (Stability) *Let $r \geq 2$ and $s_1 \leq \cdots \leq s_r$. Let $F = K_{s_1,\ldots,s_r}$ and define p by (2). Let $n \to \infty$ and let G be an arbitrary F-saturated graph of order n and size $\mathrm{sat}(n, F) + o(n)$.*

Then G has a set P of $O(1)$ vertices such that all but at most $o(n)$ vertices of G have exactly p neighbors in P and exactly $s_r - 1$ neighbors in the complement of P. In Cases I and II, we can additionally require that $|P| = p$.

By refining further the structural information about G in Cases I and II, we obtain a very sharp version of stability, which implies that if the size of the graph is within $O(1)$ from $\mathrm{sat}(n, F)$, then we need to change at most $O(1)$ edges.

Theorem 3 (Refined Stability) *Let $r \geq 2$ and $s_1 \leq \cdots \leq s_r$ such that $s_{r-1} < s_r$ or $s_r = 1$. Let $F = K_{s_1,\ldots,s_r}$ and define p by (2). Let $n \to \infty$ and let G be an arbitrary F-saturated graph of order n and size at most $\left(p + \frac{s_r - 1}{2}\right) n + O(1)$.*

Then G has a set P of p vertices such that all but at most $O(1)$ vertices of G are connected to every vertex of P and have exactly $s_r - 1$ neighbors in the complement of P.

Clearly, Theorems 1 and 3 determine the saturation function within an additive error $O(1)$ in Case I.

Corollary 4 *Fix $r \geq 2$ and $s_1 \leq \cdots \leq s_r$ with $s_{r-1} < s_r$. Let $n \to \infty$. Then*

$$\mathrm{sat}(n, K_{s_1,\ldots,s_r}) = \left(s_1 + \cdots + s_{r-1} + \frac{s_r - 3}{2}\right) n + O(1).$$

Note that in Case II we know the saturation function exactly by (1) while in Case III the error term is $O(n^{3/4})$ by Theorem 1

Section 5 provides constructions showing that some of the stability results cannot be strengthened in certain ways.

4

2   Notation

We use the standard notation and terminology of graph theory. In particular, $v(G)$ and $e(G)$ denote respectively the number of vertices and edges of a graph G. All undefined terms can be found in e.g. [Bol98]. The following less standard shorthands are helpful in our proofs.

Suppose that we have a graph $G = (V, E)$ (which will typically be the $K_{s_1,\dots,s_r}$-saturated graph that we study). For $A \subseteq V$ and $x \in V$, we use the following notation:

- $\overline{A} = V \setminus A$ is the complement of A,

- $G - A = G[V \setminus A]$ is the subgraph of G induced by $\overline{A}$,

- $\Gamma(A)$ is the set of all vertices $y \in V \setminus A$ that are connected to *every* vertex of A,

- $\Gamma(x) = \Gamma(\{x\})$ is the neighborhood of x,

- $d(x) = |\Gamma(x)|$ is the degree of x,

- $d_A(x) = |\Gamma(x) \cap A|$ is the number of neighbors of x in A.

Also, let $G[A, B]$ denote the bipartite subgraph of G induced by disjoint vertex sets $A, B \subseteq V$.

If the addition of xy to G creates a copy of $K_{s_1,\dots,s_r}$, we denote its parts by $S_1, \dots, S_r$ with $|S_i| = s_i$ for each $i = 1, \dots, r$. In addition to (2), we will also use the shorthand

$$q = p + s_r - 1 = s_1 + \cdots + s_r - 2. \tag{4}$$

In order to improve the readability of the paper and decrease the number of variables, we use the asymptotic notation (such as $O(1)$ and $o(1)$) which hides constants that depend on $(s_1, \dots, s_r)$ but not on n.

Also, we denote $[n] = \{1, \dots, n\}$ and $[m, n] = \{m, \dots, n\}$.

3   Proof of Theorem 1

Let us recall the following easy result (see [KT86]).

Lemma 5 *Let $s \geq 2$ and $n \geq s + \lfloor s/2 \rfloor$. Then*

$$\text{sat}(n, K_{1,s}) = \left\lceil \frac{(s-1)n}{2} - \frac{s^2}{8} \right\rceil. \tag{5}$$

Proof. A graph G is $K_{1,s}$-saturated if and only if the maximal degree of G is at most $s - 1$ and $G[A]$ is a clique, where A consists of vertices of degree at most $s - 2$ in G. Thus we have to minimize $\lceil f(a) \rceil$, where

$$f(a) = \binom{a}{2} + \frac{(s-1)(n-a)}{2} = \frac{(s-1)n}{2} - \frac{a(s-a)}{2}.$$

The maximum of $a(s-a)/2$ is attained for e.g. $a = \lfloor s/2 \rfloor$. If s is even, this gives the claimed bound (5). If $s = 2k + 1$ is odd, then $f(k) = kn - \frac{s^2}{8} + \frac{1}{8}$, giving the required. ∎

The following construction exhibits the upper bound of Theorem 1.

Construction 1 Let F be as in Theorem 1. Define p and q by (2) and (4). Let G' be a K_{1,s_r}-saturated graph on $[n - p]$ obtained from a clique on a set $A \subseteq [n - p]$ of size $\lfloor s_r/2 \rfloor$ by adding at most one edge connecting A to $B = [n - p] \setminus A$ and adding edges inside B so that $G'[B]$ does not contain any K_{t,s_r+1-t} with $1 < t < s_r$ and every vertex of B has degree $s_r - 1$ in G'. (Such G' exists for all large n.) Finally, let $G = K_p * G'$. Then G is F-saturated.

Proof. First, let us check that G is F-free. When we injectively map $V(F)$ into $V(G)$, then a set X of at least $v(F) - p = s_r + 1$ vertices of F goes inside $V(G')$. Since X cannot entirely fit inside some part S_i of F, it follows that $F[X]$ contains a complete bipartite subgraph K_{t,s_r+1-t} for some $t \in [s_r]$. If $t = 1$ or $t = s_r$, then $F[X]$ has a vertex of degree s_r and thus it is not a subgraph of G'. Otherwise $s_r \geq 3$ and $F[X]$ is bridgeless. Thus, if $F[X]$ embeds into G', then X has to go inside A. However, $|A| = \lfloor s_r/2 \rfloor < |X| = s_r + 1$ is too small. Hence, $F \not\subseteq G$.

On the other hand, consider the addition of any missing edge xy to G. Since xy lies inside $V(G')$, it creates K_{1,s_r} in G'. Then $G + xy$ contains $K_p * K_{1,s_r} \cong K_{p+1} * \overline{K}_{s_r} \supseteq F$, as required. ∎

Next, we prove the lower bound of Theorem 1. Let r, s_i, F be as in Theorem 1. Let n be large and let G be an arbitrary F-saturated graph on

$V = [n]$ with $e(G) = O(n)$. Some properties of G that we establish here will be used later when we work on the stability property. In particular, the following definitions apply to the whole article:

$$
\begin{aligned}
A &= \{x \in V : d(x) \geq n^{1/4}\}, \\
B &= \{x \in V \setminus A : d_A(x) \leq p - 1\}, \\
C &= \{x \in V : d(x) \leq q - 1\}.
\end{aligned}
\tag{6}
$$

Since $e(G) \geq |A|\, n^{1/4}/2$, we have

$$
|A| = O(n^{3/4}).
\tag{7}
$$

Lemma 6 $|B| = O(n^{3/4})$.

Proof. Take any two non-adjacent vertices $x, y \in B$. The addition of xy to G creates a copy of F on $V(F) = S_1 \cup \cdots \cup S_r$, say, with $x \in S_i$ and $y \in S_j$. Since each part S_t is an independent set in F, we have $i \neq j$. Let $S = V(F) \setminus (S_i \cup S_j)$.

Let us show that $G - A$ contains a path of length at most 3 connecting x to y. Assume that A contains S, the common neighborhood of x and y in F, for otherwise we are done. Since

$$
d_{A \setminus S}(x) \leq p - 1 - |S| \leq |S_j| - 2,
$$

and x is connected in F to every vertex of S_j, there is $y' \in (S_j \setminus \{y\}) \setminus A$. Likewise, there is $x' \in (S_i \setminus \{x\}) \setminus A$. Then $xy', y'x', x'y \in E(F)$ give the required 3-path connecting x to y in $G - A$.

Thus the diameter of B with respect to the graph $G - A$ is at most 3. By estimating the possible number of paths of length at most 3 in $G - A$ that start in B, we obtain

$$
\binom{|B|}{2} \leq |B|(n^{1/4} + n^{2/4} + n^{3/4}),
$$

which implies the lemma. ∎

Lemma 7 $|C| = O(1)$.

Proof. Let $m = v(F)$ be the number of vertices in F. Let the *Ramsey number* $R_{q-1}(m)$ be the smallest t such that any coloring of the edges of K_t with $q-1$ colors gives a monochromatic copy of K_m. By Ramsey's Theorem [Ram30], $R_{q-1}(m)$ exists (i.e. is finite).

We claim that $|C| < qR_{q-1}(m)$. Suppose that this is false. Since each vertex of C has degree at most $q-1$, the straightforward greedy algorithm gives us an independent set $C' \subseteq C$ with $|C'| \geq R_{q-1}(m)$. For each vertex $x \in C'$, fix an arbitrary ordering of its neighborhood in G, which we encode by a bijection $f_x : \Gamma(x) \to [d(x)]$.

For each pair of distinct vertices $x, y \in C'$, do the following. Fix a copy of F in $G + xy$ on $S_1 \cup \cdots \cup S_r$. Let $x \in S_i$ and $y \in S_j$. Then $i \neq j$. Let $X = S_j \setminus \{y\} \subseteq \Gamma(x)$ and $Y = V(F) \setminus (S_j \cup \{x\}) \subseteq \Gamma(y)$. Since $|X \cup Y| = |V(F)| - 2 > q - 1$ (while $d(x), d(y) \leq q - 1$) there are $x' \in X$ and $y' \in Y$ with $f_x(x')$ and $f_y(y')$ equal to the same index $t \in [q-1]$. But the bipartite graph $F[X, Y]$ is complete, so x' and y' are adjacent. Let $c(xy) = t$.

This defines an edge $(q-1)$-coloring c of all pairs of C'. Let an m-set $M \subseteq C'$ be monochromatic, of color t. Then all vertices $f_x^{-1}(t)$, $x \in M$, are pairwise adjacent in G (in particular, they are pairwise distinct). Thus G contains a clique of order $m = v(F)$, which contradicts the F-freeness of G. This contradiction proves the lemma. ∎

Now, armed with Lemmas 6 and 7, we are ready to finish the proof of Theorem 1. For each vertex of $V \setminus (A \cup B)$, select some p edges that connect it to A. This gives, by Lemma 6, $p(n - |A| - |B|) = pn - O(n^{3/4})$ edges. The number of the remaining edges can be bounded from below by

$$\frac{1}{2} \sum_{x \in \overline{A}} (d(x) - p) \geq \frac{(n - |A|)(q - p) - q|C|}{2} \geq \frac{q - p}{2} n - O(n^{3/4}), \quad (8)$$

where we used (7) and Lemma 7. Putting these bounds together, we get the lower bound in (3).

4 Stability

4.1 Proof of Theorem 2

Let us prove Theorem 2. Fix $r \geq 2$ and $s_1 \leq \cdots \leq s_r$. Let $n \to \infty$. Given an arbitrary F-saturated graph G on $V = [n]$ with at most $\mathrm{sat}(n, F) + o(n)$

edges, we prove a sequence of lemmas, each giving a more accurate information about the structure of G. If a lemma applies only to some of Cases I–III, we explicitly specify this in its statement.

As before, p and q are defined by (2) and (4). Recall that $\text{sat}(n, F) = (\frac{p+q}{2} + o(1)) n$ by Theorem 1. As in Section 3, we define C to be the set of vertices of degree less than q in G. By Lemma 7, $|C| = O(1)$. Let $C' = \{x \in V : d(x) > q\}$. Clearly, we can add an extra summand $|C'|/2$ to to (8). Thus we have

$$|C'| = o(n). \tag{9}$$

The following auxiliary lemma about our almost minimum F-saturated graph G is very useful in our proof.

Lemma 8 *If a set $L \subseteq V$ is such that $|L| = n - o(n)$ and all but $o(n)$ vertices of L have at least p neighbors in $\overline{L} = V \setminus L$, then*

1. *$G[L]$ can be obtained from some graph H of maximum degree at most $q - p$ by adding $o(n)$ edges,*

2. *every vertex of L, with at most $o(n)$ exceptions, has exactly p neighbors in $\overline{L}$ and exactly $q - p$ neighbors in L.*

Proof. We modify the edge counting argument at the end of the proof of Theorem 1. For each vertex x of L select a set of $\min(p, d_{\overline{L}}(x))$ edges connecting x to $\overline{L}$. The total number of selected edges is $pn - o(n)$. Let R be the set of the remaining edges and let $D = \{x \in L : d_L(x) > q - p\}$. Then

$$2|R| \geq \sum_{x \in L \setminus (C \cup D)} (q - p) + \sum_{x \in D} (q - p + 1) \geq (q - p)n + |D| - o(n).$$

Thus $e(G) \geq \frac{p+q}{2} n + \frac{1}{2} |D| - o(n)$. Since G is almost extremal, we conclude that $|D| = o(n)$. Hence, we also have

$$2|R| \geq \sum_{x \in L \setminus (C \cup D)} (q - p) + \sum_{x \in D} (d(x) - p) \geq (q - p)n + \sum_{x \in D} d(x) - o(n).$$

Thus the number of edges incident to D is at most $\sum_{x \in D} d(x) = o(n)$. The graph $H = G[L \setminus D]$ satisfies the first part of the lemma.

Also, if a vertices of L have more than p neighbors in $\overline{L}$, then by selecting $p+1$ edges into $\overline{L}$ per each of these vertices in the above argument, we would improve the final lower bound on $e(G)$ by $a/2$. Thus $a = o(n)$ and almost every vertex of L has exactly p neighbors in $\overline{L}$.

Among the vertices of L that have $\overline{L}$-degree exactly p, those that fail to have exactly $q - p$ neighbors in L are restricted to $C \cup C'$. This set is small by Lemma 7 and (9). The lemma is proved. ∎

Let $L_0 = \{x \in V : d(x) = q\} = V \setminus (C \cup C')$ and $K = C \cup C' = V \setminus L_0$. For $i = 1, 2$, define

$$L_i = \{x \in L_{i-1} : d_{L_{i-1}}(x) = q - p\}. \tag{10}$$

By Lemma 6 (since $A \subseteq K$, where A is defined by (6)), almost every vertex of L_0 has at least p neighbors in K. By Lemma 7 and (9), we have $|L_0| = n - o(n)$. Thus Lemma 8 applies to the set $L = L_0$ and shows that $|L_1| = n - o(n)$. Since $|L_0 \setminus L_1| = o(n)$, we can apply Lemma 8 again, this time to the set L_1 and conclude that $|L_2| = n - o(n)$.

Thus each vertex $x \in L_2$ has exactly p neighbors in $\overline{L_1}$ and $q - p$ neighbors in L_1. Also, the vertex $x \in L_2 \subseteq L_1$ has degree p and $q - p$ into $K \subseteq \overline{L_1}$ and $L_0 \supseteq L_1$ respectively. So, each $x \in L_2$ as well as every $y \in \Gamma(x) \cap L_0 = \Gamma(x) \cap L_1$ has exactly $q - p$ and p neighbors in L_0 and K respectively.

For $x \in L_2$, let us call the set $\Gamma(x) \cap K$ the *stem* of x and any set of the form $\Gamma(y) \cap K$ for some $y \in \Gamma(x) \cap L_0$ a *branch* of x. The above properties of L_2 ensure that each vertex $x \in L_2$ has exactly $q - p$ branches (when we take into account their multiplicity) and that each stem and branch is a p-subset of K. These definitions also apply when $s_r = 1$ (when we are in Case II): then there are no branches.

Lemma 9 *Let $x, y \in L_2$ be at distance at least 4 in $G - K$. Then the following holds, depending on which case we are in.*

Case I: *The stem of one vertex is equal to a branch of the other vertex.*

Case II: *The stem of x equals to the stem of y.*

Case III: *The stem of one vertex is equal to a branch of the other vertex or we have $|\Gamma(x) \cap \Gamma(y) \cap K| = s_1 + \cdots + s_{r-2}$ and G contains all edges between the non-empty sets $(\Gamma(x) \setminus \Gamma(y)) \cap K$ and $(\Gamma(y) \setminus \Gamma(x)) \cap K$.*

Proof. Suppose that the addition of xy to G creates a copy of F on $S_1 \cup \cdots \cup S_r$. Let i and j be the indices such that $x \in S_i$ and $y \in S_j$. Let $S = V(F) \setminus (S_i \cup S_j)$. Since each of x and y is connected to everything in S, we have

$$S \subseteq \Gamma(x) \cap \Gamma(y) \subseteq K. \tag{11}$$

The number of neighbors of x (or y) in $K \setminus S$ is $p - |S| = s_i + s_j - s_r - 1$. Also, since there is is no 3-path connecting x to y in $G - K$, at least one of $S_i \setminus \{x\}$ and $S_j \setminus \{y\}$ is a subset of K. Suppose that it is the latter. Then the size considerations show that $s_i = s_r$ and $\Gamma(x) \cap K = S \cup (S_j \setminus \{y\})$.

Suppose that we are in Case I (that is, $s_r > s_{r-1}$). We necessarily have $i = r$ and $s_j < s_r$. Thus $s_i - 1 > s_i + s_j - s_r - 1$ and the set $S_i \setminus \{x\}$ intersects L_0. Pick any vertex z in this intersection. Since $yz \in E(F)$ and $y \in L_2$, we have in fact $z \in L_1$. Also, z has to be connected to all of $S \cup (S_j \setminus \{y\})$ in G. Since this set is a subset of K, this identifies p K-neighbors of z. But the vertex $z \in L_1$ has exactly p neighbors in K. Thus the set $\Gamma(z) \cap K$, which is a branch of y, is equal to the stem $\Gamma(x) \cap K$ of x, giving the required.

If we are in Case II, then x and y have $|S| = p$ common neighbors in F and thus (by $x, y \in L_2$) we have $\Gamma(x) \cap K = \Gamma(x) \cap K$, as required.

Finally, suppose that we are in Case III. If $S_i \setminus \{x\}$ intersects L_0 in some vertex z, then we argue in the same way as in Case I that a branch of y (defined by this $z \in \Gamma(y)$) is equal to the stem of x. Suppose that $S_i \setminus \{x\} \subseteq K$. Then $s_j = s_r$ and $\Gamma(y) \cap K = S \cup (S_i \setminus \{x\})$. Now, $(\Gamma(x) \setminus \Gamma(y)) \cap K = S_j \setminus \{y\} \neq \emptyset$ and $(\Gamma(y) \setminus \Gamma(x)) \cap K = S_i \setminus \{x\} \neq \emptyset$ while F contains all edges between these two sets, giving the required. ∎

Lemma 10 *There is a collection $\mathcal{S}$ of p-subsets of K such that $|\mathcal{S}| = O(1)$ and the number of vertices $x \in L_2$ whose stem belongs to $\mathcal{S}$ is $n - o(n)$.*

Proof. We use the following *Selection Procedure.* We inductively choose vertices $x_1, x_2, \ldots \in L_2$ so that each new vertex x_{i+1} is at distance more than 3 in $G - K$ from each of $x_1, \ldots, x_i$ and the stem of x_{i+1} is not equal to the stem or some branch of a previous vertex. Suppose that the procedure stops (no more suitable vertices exist) after $x_1, \ldots, x_m$ have been selected.

We show first that m is bounded. Although we can write a formal proof that applies to all the cases, the argument is more readable when we split it case by case.

If we are in Case I, then $m \leq s_r$. Indeed, if x_{s_r+1} exists, then by Lemma 9 the stem of each $x_1, \ldots, x_{s_r}$ has to be equal to a branch of x_{s_r+1}. But these stems are all pairwise distinct by our construction while $x_{s_r+1} \in L_2$ has at most $s_r - 1$ different branches, a contradiction.

By Lemma 9, $m \leq 1$ in Case II.

Suppose that we are in Case III. Let t be an integer such that any p-uniform hypergraph with t edges contains a *delta-system* of size $v(F)$, that is $v(F)$ hyperedges whose pairwise intersections are all the same. The celebrated result of Erdős and Rado [ER60] states that such t exists (i.e. is finite). We claim that $m \leq 2(s_r - 1)(t - 1) + 1$. Suppose on the contrary that this is false. Define a graph H on $[m]$ by putting an edge ij if and only if the second alternative in Case III of Lemma 9 holds for the pair $x = x_i$ and $y = x_j$ of vertices. The argument of Case I shows that for each $i \in [m]$ at most $s_r - 1$ pairs ij with $j \in [i-1]$ are missing from H. Thus

$$e(H) \geq \binom{m}{2} - (s_r - 1)m > \frac{t-2}{t-1}\binom{m}{2}.$$

By Turán's theorem [Tur41], H contains a copy of K_t, say with vertices $h_1, \ldots, h_t$. By the choice of t, the p-uniform hypergraph $\mathcal{H} = \{\Gamma(x_{h_i}) \cap K : i \in [t]\}$, contain a delta-system with $v(F)$ hyperedges, say $\mathcal{F} = \{\Gamma(y_i) \cap K : i \in [v(F)]\}$. By Lemma 9 we know that the common intersection of $\mathcal{F}$ has $s_1 + \cdots + s_{r-2} < p$ vertices. Hence, for each hyperedge $\Gamma(y_i) \cap K$ of $\mathcal{F}$ we can pick a vertex z_i which belongs only to this hyperedge. The vertices $z_1, \ldots, z_{v(F)}$ span a clique in G. This contradicts the F-freeness of G.

Thus, in each of the three cases, the size of the set $X = \{x_1, \ldots, x_m\} \subseteq L_2$ produced by the Selection Procedure is bounded by a constant. Since $G - K$ has maximum degree at most q, the number of vertices of $L_2 \subseteq L_0$ at distance at most 3 from X in $G - K$ is $O(1)$. The stem of each remaining vertex of L_2 already occurs as the stem or a branch of some $x \in X$ by the maximality of X. Thus the family of p-sets

$$\mathcal{S} = \bigcup_{x \in X} \Big(\{\Gamma(x) \cap K\} \cup \{\Gamma(y) \cap K : y \in \Gamma(x) \cap L_0\}\Big),$$

whose size is at most $s_r m = O(1)$, establishes the validity of the lemma. ∎

Fix the family $\mathcal{S}$ given by Lemma 10. By removing p-sets $X \in \mathcal{S}$ with $|\Gamma(X)| = o(n)$ from $\mathcal{S}$, we can additionally assume that $|\Gamma(X)| = \Omega(n)$ for

each $X \in \mathcal{S}$. Recall that $\Gamma(X)$ consists of those vertices that are connected to everything in X.

What we have achieved by now is that if we let $L = V \setminus (\cup_{X \in \mathcal{S}} X)$, then L is a set of size $n - O(1)$ such that almost every vertex of G has at least p neighbors in $\overline{L}$. But, since we need to apply Lemma 8 again, we prefer to keep shrinking L_i by moving all "exceptional" vertices to the complement of L_{i+1}. So we define

$$L_3 = \{x \in L_2 : \Gamma(x) \cap K \in \mathcal{S}\}.$$

We have $|L_3| = n - o(n)$. Define L_4 by (10) with $i = 4$. By Lemma 8, $|L_4| = n - o(n)$. Similarly as we argued about L_2, it follows that all $q - p$ L_0-neighbors of every $x \in L_4$ belong to L_3. Thus the stem and all branches of each $x \in L_4$ belong to the family $\mathcal{S}$.

Let L_5 consist of those vertices of L_4 all whose branches are the same as the stem:

$$L_5 = \big\{ x \in L_4 : \forall\, y \in \Gamma(x) \cap L_0 \;\; \Gamma(x) \cap K = \Gamma(y) \cap K \big\}.$$

Lemma 11 *In Case I (and vacuously in Case II), we have $|L_5| = n - o(n)$.*

Proof. Suppose on the contrary that $|L_4 \setminus L_5| = \Omega(n)$. Since $|\mathcal{S}| = O(1)$, we can find $X \in \mathcal{S}$ such that $|Y| = \Omega(n)$, where $Y = \Gamma(X) \cap (L_4 \setminus L_5)$.

Since $\Delta(G - K) \leq q$, we can find $x, y \in Y$ at distance at least 4 in $G - K$. Let the addition of xy to G create a copy of F on $S_1 \cup \cdots \cup S_r$. Let $x \in S_i$ and $y \in S_j$ with, say, $i > j$. As in the proof of Lemma 10, we argue that the set $S = V(F) \setminus (S_i \cup S_j)$ lies inside K, there is $z \in (S_i \setminus \{x\}) \cap L_0$, this vertex z belongs in fact to L_3, $S_j \setminus \{y\}$ is a subset of K, and $i = r$. By now, we have identified all p K-neighbors of x. By the definition of $Y \ni x, y$, the vertices x and y have the same neighborhood X in K. Thus the set $S_i \setminus \{x\}$ lies entirely in $\Gamma(y) \setminus K \subseteq L_0$. Since y is connected in F to all of S_i, we have that the set $S_i \setminus \{x\}$ includes all $q - p$ L_3-neighbors of Y; in particular, $S_i \setminus \{x\} \subseteq L_3$. However, each vertex $w \in S_i \setminus \{x\}$ is connected in F to everything in $S \cup (S_j \setminus \{y\}) = X$. This set has size p. Since $w \in L_3 \subseteq L_1$, we have that $\Gamma(w) \cap K = X$. Thus the stem and all branches of y are equal to X, which contradicts $y \in L_4 \setminus L_5$. The lemma is proved. ∎

Lemma 12 *If we are in Case I or II, then $|\mathcal{S}| = 1$.*

Proof. Suppose on the contrary that we have two distinct $X, Y \in \mathcal{S}$. By Lemma 11, we can pick $x, y \in L_5$ such that $\Gamma(x) \cap K = X$, $\Gamma(y) \cap K = Y$, and the distance between x and y in $G - K$ is at least 4.

Let the addition of xy to G create a copy of F on $S_1 \cup \cdots \cup S_r$ with $x \in S_i$ and $y \in S_j$. The set $S = V(F) \setminus (S_i \cup S_j)$ lies inside K. If we are in Case II, then $|S| = p$ and thus $X = Y = S$, a contradiction. So suppose that we are in Case I.

Without loss of generality assume that $i > j$. Then there is $z \in (S_i \setminus \{x\}) \cap L_0$, $z \in L_3$, $S_j \setminus \{y\} \subseteq K$, $i = r$, $X = \Gamma(x) \cap K = S \cup (S_j \setminus \{y\})$ and this is also the K-neighborhood of z: $\Gamma(z) \cap K = X$. However, z is a neighbor of $y \in L_5$, so $\Gamma(z) \cap K = Y$. This contradiction proves the lemma. ∎

Let $P = \cup_{X \in \mathcal{S}} X$. Now, Lemma 12 (in Cases I and II), Lemma 10 (in Case III), and one final application of Lemma 8 to $L = V \setminus P$ imply Theorem 2.

4.2 Proof of Theorem 3

Next, we move on to proving Theorem 3. As before, let G be an F-saturated graph on $V = [n]$ but we additionally assume that

$$e(G) \leq \frac{p+q}{2} n + O(1). \tag{12}$$

Let $P = \cup_{X \in \mathcal{S}} X$ be as before. Since we are in Case I or II, we have $|P| = p$. Let N consist of those vertices of $\overline{P}$ that are connected to every vertex of P. By the definition of $\mathcal{S}$, we have $|N| = n - o(n)$. For $i \in [0, p-1]$ let

$$Y_i = \{x \in \overline{P} : d_P(x) = i\}$$

and $y_i = |Y_i|$. Let $Y = \cup_{i=0}^{p-1} Y_i$ and $y = |Y|$. The sets N, P, and Y partition V. Also, let

$$\begin{aligned} L &= \{x \in \overline{P} : d(x) \geq \ln y\}, \\ M &= \{x \in N : d_N(x) \leq s_r - 2\}. \end{aligned}$$

If $y = O(1)$, then Theorem 3 follows from Lemma 7. So, suppose on the contrary that $y \to \infty$ or, more precisely, that we can select an increasing sequence of n and a graph G on $V = [n]$ for each selected n so that $y \to \infty$. (Note that $y = o(n)$.)

The following sequence of claims will give us more and more precise structure of G until we obtain the final contradiction.

By Lemma 7, the set C of vertices of degree less than q has size $O(1)$. The following simple initial estimate, where we take into account the number of edges between P and $\overline{P}$ plus the half of the sum of degrees of $G - P$, will be the basis for further refinements:

$$
\begin{aligned}
e(G) \ \geq \ & e(G[P, N \cup Y]) + \frac{1}{2} \sum_{x \in (N \cup Y) \setminus C} (q - d_P(x)) \\
\geq \ & p(n - y) + \sum_{i=0}^{p-1} i y_i + \frac{(n - y)(q - p)}{2} + \frac{1}{2} \sum_{i=0}^{p-1} (q - i) y_i - O(1) \\
\geq \ & \frac{p + q}{2} n + \sum_{i=0}^{p-1} \frac{i - p}{2} y_i - O(1).
\end{aligned}
\tag{13}
$$

So far this inequality is not very useful as we do not have much control over y_i. But if we can add e.g. an extra term py to the bound, this will imply the desired result $y = O(1)$. One way of improving (13) is to observe that, for each vertex $x \in \overline{P}$, the sum (13) counts at most q edges at x. Hence, we can add

$$
\frac{1}{2} \sum_{x \in N \cup Y} (d(x) - q)
\tag{14}
$$

to the right-hand side of (13). One immediate consequence is that

$$
|L| = O(y / \ln y) = o(y),
\tag{15}
$$

because $\ln y - q = (1 - o(1)) \ln y$ so the extra term (14) is at least $(1/2 - o(1))|L| \ln y$. But, this is $O(y)$ by (12) and (13), so (15) holds.

A bit more subtle consequence is the following. For a vertex $x \in \overline{P}$ and an integer $i \geq 1$, let $\partial_i(x)$ consist of all vertices of $G - P$ that are at distance at most i (with respect to the graph $G - P$) from x. For a set $X \subseteq \overline{P}$, let $\partial_i(X) = \cup_{x \in X} \partial_i(x)$.

Lemma 13 *For every $x \in \overline{P}$, wee have $|\partial_3(x)| = O(y)$.*

Proof. Let $A \subseteq N \cup Y$ be an arbitrary set. Since each vertex of $\partial_1(A) \setminus A$ sends at least one edge to A and we cannot increase the right-hand side of (13) by

15

py, we have that

$$|\partial_1(A)| - |A| \le \sum_{x \in A} d(x) = \sum_{x \in A}(d(x) - q) + q|A| \le 2py + q|A|,$$

implying that $|\partial_1(A)| \le 2py + (q+1)|A|$. Iteratively applying this bound to $\partial_3(x) = \partial_1(\partial_1(\partial_1(x)))$, we obtain the required. ∎

Lemma 14 *The set Y_i is empty for each $i < s_1 + \cdots + s_{r-2}$.*

Proof. Suppose $Y_i \ni x$. By Lemma 13 there is $z \in N \setminus \partial_3(x)$. Let $S_1 \cup \cdots \cup S_r$ span a copy of F in $G + xz$. The common neighborhood of x and z in F has cardinality at least $s_1 + \cdots + s_{r-2}$. But x and z have at most $d_P(x) = i$ common neighbors in G, hence $i \ge s_1 + \cdots + s_{r-2}$, as required. ∎

If we are in Case II, then $p = r - 2$, $Y = \cup_{i=0}^{r-3} Y_i$ is empty by Lemma 14, and $y = 0$, which is the desired contradiction. So suppose that we are in Case I (that is, $s_{r-1} < s_r$).

Lemma 15 *For every $t \in [0, p-1]$, we have $|Y_t'| = o(y)$, where Y_t' consist of those vertices of Y_t that have at most $p - t - 1$ neighbors in L.*

Proof. Suppose on the contrary that $|Y_t'| = \Omega(y)$. The graph $G - P - L$ (by which we mean $G - (P \cup L)$) has maximum degree less than $\ln y$. Since $|L| = o(y)$, there are $x, z \in Y_t' \setminus L$ at distance at least 4 in $G - P - L$. Then each of x and z has at most $p - 1$ neighbors in $L \cup P$. Let the addition of xz to G create an F-subgraph on $S_1 \cup \cdots \cup S_r$ with $x \in S_i$ and $z \in S_j$. Let $S = V(F) \setminus (S_i \cup S_j)$. Then $S \subseteq L \cup P$ while both $S_i \setminus \{x\}$ and $S_j \setminus \{z\}$ intersect the complement of $L \cup P$ (because the degrees of z and x into $L \cup P$ are too small). This gives a 3-path connecting x to z in $G - P - L$, a contradiction. ∎

Lemma 16 *Let $M' = M \setminus \partial_1(L)$ consist of those vertices of M that are not in L nor have a neighbor in L. Then $|M'| = o(y)$.*

Proof. Suppose that the lemma is false. Since the graph $G - P - L$ has maximum degree at most $\ln y$ we can find $x, z \in M'$ at distance at least 4 in $G - P - L$. Since $x, z \notin L$ do not have any neighbors in L, their distance in

16

$G - P$ is also at least 4. Let $S_1 \cup \cdots \cup S_r$ span a copy of F in $G + xz$ with $x \in S_i$ and $z \in S_j$. Let $S = V(F) \setminus (S_i \cup S_j)$. Then S lies inside $\Gamma(x) \cap \Gamma(z) \subseteq P$. We have

$$|(\Gamma(x) \cap P) \setminus S| = |(\Gamma(y) \cap P) \setminus S| = s_i + s_j - s_r - 1.$$

However, we have $S_i \setminus \{x\} \subseteq (\Gamma(y) \cap P) \setminus S$ or $S_j \setminus \{z\} \subseteq (\Gamma(x) \cap P) \setminus S$ (for otherwise we find a 3-path in $G - P$ between x and z). Suppose that the latter inclusion holds. It follows that $s_i = s_r$ and $S \cup (S_j \setminus \{z\}) = \Gamma(x) \cap P$.

Every element of S_i is connected in F to everything in $S \cup (S_j \setminus \{z\}) = P$. Thus $S_i \subseteq N$ and z has at least $|S_i \setminus \{x\}| = s_r - 1$ neighbors in N. So $z \notin M$, a contradiction to our assumption $z \in M'$. ∎

Lemma 17 $|M| = o(y)$.

Proof. By (15) and Lemmas 15 and 16, the sum in (14) is at least

$$\frac{1}{2} \sum_{x \in L} (d(x) - q) = \frac{1}{2} \sum_{x \in L} d(x) - o(y) \geq \frac{|M|}{2} + \frac{1}{2} \sum_{i=0}^{p-1} (p - i) y_i - o(y).$$

When we add this to (13), the bound becomes

$$e(G) \geq \frac{p + q}{2} n + \frac{|M|}{2} - o(y).$$

Now, (12) implies that $|M| = o(y)$. ∎

Lemma 18 *Every vertex* $x \in Y$ *has at least* $s_r - 1$ *neighbors in* N.

Proof. Since $|\partial_3(x)| = O(y)$ is much smaller that $|N|$, there is $z \in N \setminus \partial_3(x)$. Let $S_1, \ldots, S_r$ be the parts of $F \subseteq G + xz$ with $x \in S_i$ and $z \in S_j$. The set $S = V(F) \setminus (S_i \cup S_j)$ is inside $\Gamma(x) \cap \Gamma(z) \subseteq P$. Moreover, $S_i \setminus \{x\} \subseteq (\Gamma(z) \cap P) \setminus S$ or $S_j \setminus \{z\} \subseteq (\Gamma(x) \cap P) \setminus S$ for otherwise x and z are at distance at most 3 in $G - P$. Since

$$|(\Gamma(x) \cap P) \setminus S| < |(\Gamma(z) \cap P) \setminus S| = s_i + s_j - s_r - 1 \leq \min(s_i - 1, s_j - 1), \quad (16)$$

we have $\Gamma(z) \cap P = S \cup (S_i \setminus \{x\})$. Also, the last inequality in (16) is in fact equality and $s_j = s_r$. Since F has all edges connecting S_j to $S \cup (S_i \setminus \{x\})$ and

17

the latter set has p elements, all belonging to P, we conclude that $S_j \subseteq N$. Thus x has at least $|S_j \setminus \{z\}| = s_r - 1$ neighbors in N, as required. ∎

Let $j = s_1 + \cdots + s_{r-2}$. We obtain the following lower bound on $e(G)$ by Lemmas 7, 14, 17, and 18:

$$
\begin{aligned}
e(G) \;\geq\;\; & \sum_{x \in N \setminus M} \left(d_P(x) + \frac{d_N(x)}{2} \right) + \sum_{x \in Y} \left(d_{N \cup P}(x) + \frac{d(x) - d_{N \cup P}(x)}{2} \right) \\
\geq\;\; & (n - |M| - y)\left(p + \frac{s_r - 1}{2} \right) + \sum_{i=j}^{p-1} y_i \, \frac{q + (s_r - 1) + i}{2} - q|C| \\
\geq\;\; & \frac{p+q}{2}\, n + \sum_{i=j}^{p-1} y_i \, \frac{i + s_r - s_1 - \cdots - s_{r-1}}{2} - o(y).
\end{aligned}
$$

Since $s_{r-1} < s_r$, the coefficient at y_i for each $i \in [j, p-1]$ is strictly positive. By (12), each y_i must be bounded by a constant, a contradiction to our assumption $y \to \infty$. This finishes the proof of Theorem 3.

5   Further Constructions

Let us give an example which shows that we cannot generally require that $|P| = p$ in Theorem 2 in Case III (i.e. when $s_{r-1} = s_r \geq 2$).

Construction 2 Let $r \geq 2$, $1 \leq s_1 \leq \cdots \leq s_{r-1} = s_r$, $s_r \geq 2$, and $F = K_{s_1,\ldots,s_r}$. Define p by (2).

Let $n \to \infty$ be such that $n \equiv v(F) \pmod 2$. Take a partition $[n] = X \cup X' \cup Y \cup Y' \cup Z$, where the sizes of parts are $n/2 + o(n)$, $s_r - 1$, $n/2 + o(n)$, $s_r - 1$, and $s_1 + \cdots + s_{r-2}$ respectively, with $|X|$ and $|Y|$ even. Put a clique on $X' \cup Y' \cup Z$. Take arbitrary $(s_r - 1)$-regular C_4-free graphs on sets X and Y. Put complete bipartite graphs between X and $X' \cup Z$ and between Y and $Y' \cup Z$.

The obtained graph G is F-saturated of size $\frac{p+q}{2}\, n + O(1)$. However, G has only $|Z| = s_1 + \cdots + s_{r-2} < p$ vertices of degree $n - o(n)$.

Proof. First, suppose on the contrary that G contains a copy of F on $U = S_1 \cup \cdots \cup S_r$. If U intersects both X and Y, say $x \in U \cap X$ and $y \in U \cap Y$, then the vertices x and y, non-adjacent in G, belong to the same part S_i. But then

the number of common neighbors of x and y in F is at least $v(F) - s_i > |Z|$, a contradiction. Thus, by symmetry, we can assume that $U \subseteq X \cup X' \cup Y' \cup Z$. Suppose next that the set U intersects Y'. Since $v(F) > |X' \cup Y' \cup Z|$, the set U intersects X as well. As G has no edges between X and Y', some $r - 1$ parts of F have to lie entirely within $X' \cup Z$. But the latter set is too small for this, a contradiction. Hence, $U \subseteq X \cup X' \cup Z$. Now, the same case analysis as in Construction 1 gives a contradiction.

Next, let us show that the addition of any new edge xy to G creates an F-subgraph. Take an arbitrary partition $Z = S_1 \cup \cdots \cup S_{r-2}$ with part sizes $s_1, \ldots, s_{r-2}$ respectively. Up to symmetry, we have the following cases.

If $x \in X$ and $y \in Y$, then let $S_{r-1} = Y' \cup \{x\}$ and $S_r = X' \cup \{y\}$.

If $x \in X$ and $y \in X \cup Y'$, then let $S_{r-1} = X' \cup \{x\}$ and $S_r = (\Gamma(x) \cap X) \cup \{y\}$.

In all cases, $S_1 \cup \cdots \cup S_r$ spans the complete partite graph in $G + xy$. Hence, G is F-saturated. The remaining claims of Construction 2 are trivial to check. ∎

Remark. For some other constructions that violate $|P| = p$ in Theorem 2 when $s_{r-1} = s_r \geq 2$, see [Oll72] (when $F = K_{2,2}$) and [GS07] (when $s_1 = \cdots = s_r = 2$).

The following constructions shows that we cannot require in Theorem 3 that *every* vertex of $\overline{P}$ is connected to everything in P when $s_{r-1} \geq 2$.

Construction 3 Let $r \geq 2$, $1 \leq s_1 \leq \cdots \leq s_{r-1} \leq s_r$, $s_{r-1} \geq 2$, and $F = K_{s_1, \ldots, s_r}$. Define p by (2). Let n be large with $(n - p - 1)(s_r - 1)$ even.

Let $P = [p]$ and $X = [p + 1, n - 1]$. Define the graph G on $[n]$ as follows. Let the set X span an $(s_r - 1)$-regular C_4-free graph in G. The graphs $G[P]$ and $G[P, X]$ are complete. The remaining vertex n is connected to $[p - 1]$ as well as some set Y of $s_r - 1$ vertices of X such that the distance between any two vertices of Y in $G[X]$ is at least 3.

Then the constructed graph G is F-saturated of size $\frac{p+q}{2} n + O(1)$. It has exactly p vertices of degree $n - o(n)$ but the vertex n is not connected to all of them (namely, $\{n, p\}$ is not an edge).

Proof. Let us first check that G is F-free. The analysis of Construction 1 shows that $G - n$ is F-free. It remains to rule out the possibility of an F-subgraph on $U = S_1 \cup \cdots \cup S_r$ containing the special vertex n, say $n \in S_i$.

Every $y \in X \setminus Y$ has at most p common neighbors with n in G. Hence, S_i does not intersect $X \setminus Y$. Also, each S_j with $j \neq i$ is disjoint from $X \setminus Y$ because n sends no edge to $X \setminus Y$. Hence $U \subseteq P \cup Y \cup \{n\}$ but the latter set has only $p + (s_r - 1) + 1 < v(F)$ vertices, a contradiction.

Let us show that the addition of any new edge xy to G creates a copy of F. The analysis of Construction 1 takes care of all cases except when $n \in \{x, y\}$, say $x = n$.

If $y \in X \setminus Y$, then we can take $S_r = \{y\} \cup Y$ and take an arbitrary partition $P \cup \{n\} = S_1 \cup \cdots \cup S_{r-1}$ with $S_{r-1} \ni n, p$ and each $|S_i| = s_i$. (Recall that $s_{r-1} \geq 2$ by our assumption.) If $y = p$, then fix $z \in Y$ and take $S_r = \{n\} \cup (\Gamma(z) \cap X)$ and an arbitrary partition $P \cup \{z\} = S_1 \cup \cdots \cup S_{r-1}$ with the appropriate part sizes. In both cases, the sets $S_1, \ldots, S_r$ span a copy of F. Thus G is indeed F-saturated. All other claims of Construction 3 are trivial. ∎

Remark. Interestingly, if $s_{r-1} = 1$, then, under the assumptions of Theorem 3, the bipartite graph $G[P, \overline{P}]$ is necessarily complete for all large n. This follows from Lemma 14 (whose proof, with obvious modifications, also works when $y = O(1)$).

6   Concluding Remarks

The *exact* determination of the saturation function for complete partite graphs (at least when n is sufficiently large) is an interesting open problem. The upper bound in (3) gives the exact value of $\mathrm{sat}(n, F)$ when $s_r = 1$ (see [EHM64]) and when $s_1 = \cdots = s_{r-1} = 1$ and $n \geq n_0(r, s_r)$ is large (see [Pik04, CFG08]). However, it is not optimal when $F = K_{2,2}$ (see [Oll72]) and, more generally, when $s_1 = \cdots = s_r = 2$ (see [GS07]). The first open case is $K_{2,3}$. We believe that $\mathrm{sat}(n, K_{2,3}) = 2n - 3$ for all large n (in which case the upper bound in (3) is sharp).

Also, we do not know if Corollary 4 is valid in Case III.

It would be interesting to extend our results to k-uniform hypergraphs with $k \geq 3$. It comes as no surprise that the hypergraph case is much more complicated than the graph case. There are two natural ways of defining a complete r-partite k-graph on a vertex partition $S_1 \cup \cdots \cup S_r$:

Weak Version: $W_{s_1,\ldots,s_r}^k$ consists of all k-tuples that intersect k different parts,

Strong Version: $S_{s_1,\ldots,s_r}^k$ consist of all k-tuples that intersect at least two parts.

Very few results about the saturation function for weak and strong complete partite k-graphs for $k \geq 3$ are known. Bollobás [Bol65] completely solved the problem for $W_{1,\ldots,1}^k \cong S_{1,\ldots,1}^k$. The saturation function $\mathrm{sat}(n, W_{1,1,m}^3)$ was asymptotically determined in [Pik99] with the result being exact for infinitely many values of n for every m (when Steiner triple systems with certain parameters exist). The paper [Pik01] relates $\mathrm{sat}(n, W_{1,1,1,m}^4)$ to the Turán function $\mathrm{ex}(n, W_{1,1,1,1}^3)$: in particular, if the famous conjecture of Turán [Tur41] that $\mathrm{ex}(n, W_{1,1,1,1}^3) = \left(\frac{5}{9} + o(1)\right)\binom{n}{3}$ holds, then $\mathrm{sat}(n, W_{1,1,1,m}^4) = \left(\frac{m-1}{9} + o(1)\right)\binom{n}{3}$. Erdős, Füredi, and Tuza [EFT91] determined the saturation function exactly for $S_{1,3}^3$ and asymptotically for $S_{1,k}^k$. The asymptotic of $\mathrm{sat}(n, S_{1,m}^k)$ for any fixed $3 \leq k < m$ was determined in [Pik00]. See [Pik04] for a conjecture about $\mathrm{sat}(n, S_{l,m}^k)$ when $l \leq k - 1$ and $l + m > k$.

The (almost) extremal saturated hypergraphs in the above problems exhibit a variety of patterns, so it is difficult even to make a general conjecture.

Acknowledgments

References

[BCE⁺96]   C. A. Barefoot, L. H. Clark, R. C. Entringer, T. D. Porter, L. A. Székely, and Z. Tuza. Cycle-saturated graphs of minimum size. *Discrete Math.*, 150:31–48, 1996.

[Bol65]   B. Bollobás. On generalized graphs. *Acta Math. Acad. Sci. Hung.*, 16:447–452, 1965.

[Bol67a]   B. Bollobás. Determination of extremal graphs by using weights. *Wiss. Z. Hochsch. Ilmenau*, 13:419–421, 1967.

[Bol67b]   B. Bollobás. On a conjecture of Erdős, Hajnal and Moon. *Amer. Math. Monthly*, 74:178–179, 1967.

[Bol98]   B. Bollobás. *Modern Graph Theory*. Springer-Verlag, Berlin, 1998.

[Bon72]   J. A. Bondy. Variations on the Hamiltonian scheme. *Can. Math. Bull.*, 15:57–62, 1972.

[CCES86]   L. H. Clark, R. P. Crane, R. C. Entringer, and H. D. Shapiro. On smallest maximally non-Hamiltonian graphs. In *Combinatorics, Graph Theory and Computing, Proc. 17th Southeast. Conf., Boca Raton*, volume 53 of *Congres. Numer.*, pages 215–220, 1986.

[CE83]   L. H. Clark and R. C. Entringer. Smallest maximally non-Hamiltonian graphs. *Period. Math. Hungar*, 14:57–68, 1983.

[CES92]   L. H. Clark, R. C. Entringer, and H. D. Shapiro. Smallest maximally non-Hamiltonian graphs II. *Graphs Combin.*, 8:225–231, 1992.

[CFG08]   G. Chen, R. J. Faudree, and R. J. Gould. Saturation numbers of books. *Electronic J. Combin.*, 15:12pp., 2008.

[Che09]   Y.-C. Chen. Minimum C_5-saturated graphs. *J. Graph Theory*, 61:111–126, 2009.

[Che10]   Y.-C. Chen. All minimum C_5-saturated graphs. Preprint, 2010.

[EFT91]   P. Erdős, Z. Füredi, and Z. Tuza. Saturated r-uniform hypergraphs. *Discrete Math.*, 98:95–104, 1991.

[EG61]   P. Erdős and T. Gallai. On the minimal number of vertices representing the edges of a graph. *Publ. Math. Inst. Hungar. Acad. Sci.*, 6:181–203, 1961.

[EHM64]   P. Erdős, A. Hajnal, and J. W. Moon. A problem in graph theory. *Amer. Math. Monthly*, 71:1107–1110, 1964.

[ER60]   P. Erdős and R. Rado. Intersection theorems for systems of sets. *J. Lond. Math. Soc.*, 35:85–90, 1960.

[GŁS06]   R. Gould, T. Łuczak, and J. Schmitt. Constructive upper bounds for cycle-saturated graphs of minimum size. *Electronic J. Combin.*, 13:19pp., 2006.

[GS07]   R. J. Gould and J. R. Schmitt. Minimum degree and the minimum size of K_2^t-saturated graphs. *Discrete Math.*, 307:1108–1114, 2007.

[Isa75]   R. Isaacs. Infinite families of non-trivial trivalent graphs which are not Tait colorable. *Amer. Math. Monthly*, 82:221–239, 1975.

[KT86]   L. Kászonyi and Z. Tuza. Saturated graphs with minimal number of edges. *J. Graph Theory*, 10:203–210, 1986.

[Oll72]   L. T. Ollmann. $K_{2,2}$-saturated graphs with a minimal number of edges. In *Combinatorics, Graph Theory and Computing, Proc. 13th Southeast. Conf., Boca Raton,* (Utilitas Math., Winnipeg), pages 367–392, 1972.

[Pik99]   O. Pikhurko. The minimum size of saturated hypergraphs. *Combin. Prob. Computing*, 8:483–492, 1999.

[Pik00]   O. Pikhurko. Asymptotic evaluation of the sat-function for r-stars. *Discrete Math.*, 214:275–278, 2000.

[Pik01]   O. Pikhurko. Constructing designs straightforwardly: Worst arising cases. *J. Combin. Designs*, 9:100–106, 2001.

[Pik04]   O. Pikhurko. Results and open problems on minimum saturated graphs. *Ars Combinatoria*, 72:111–127, 2004.

[PS08]   O. Pikhurko and J. Schmitt. A note on minimum $K_{2,3}$-saturated graphs. *Australas. J. Comb.*, 40:211–215, 2008.

[Ram30]   F. P. Ramsey. On a problem of formal logic. *Proc. London Math. Soc.*, 30:264–286, 1930.

[Tur41]   P. Turán. On an extremal problem in graph theory (in Hungarian). *Mat. Fiz. Lapok*, 48:436–452, 1941.

[Tuz89]   Z. Tuza. C_4-saturated graphs of minimum size. *Acta Univ. Carolin. Math. Phys.*, 30:161–167, 1989.

[Wes66]   W. Wessel. Über eine Klasse paarer Graphen, I: Beweis einer Vermutung von Erdős, Hajnal and Moon. *Wiss. Z. Hochsch. Ilmenau*, 12:253–256, 1966.

[Wes67]   W. Wessel. Über eine Klasse paarer Graphen, II: Bestimmung der Minimalgraphen. *Wiss. Z. Hochsch. Ilmenau*, 13:423–426, 1967.

[XWCY97]   L. Xiaohui, J. Wenzhou, Z. Chengxue, and Y. Yuansheng. On smallest maximally non-Hamiltonian graphs. *Ars Combinatoria*, 45:263–270, 1997.

[Zyk49]   A. A. Zykov. On some properties of linear complexes (in Russian). *Mat. Sbornik N.S.*, 24:163–188, 1949.