ADVANCES IN COMBINATORICS, 2024:3, 26 pp.
www.advancesincombinatorics.com

Balanced Supersaturation and Turán Numbers in Random Graphs

Tao Jiang* Sean Longbrake†

Received 12 March 2023; Published 15 July 2024

Abstract: In a ground-breaking paper solving a conjecture of Erdős on the number of n-vertex graphs not containing a given even cycle, Morris and Saxton [45] made a broad conjecture on so-called balanced supersaturation property of a bipartite graph H. Ferber, McKinley, and Samotij [29] established a weaker version of this conjecture and applied it to derive far-reaching results on the enumeration problem of H-free graphs.

In this paper, we show that Morris and Saxton's conjecture holds under a very mild assumption about H, which is widely believed to hold whenever H contains a cycle. We then use our theorem to obtain enumeration results and general upper bounds on the Turán number of a bipartite H in the random graph $G(n,p)$, the latter being the first of its kind.

Key words and phrases: supersaturation, extremal-graph-theory

1 Introduction

Given a graph H, we say that a graph G is *H-free* if it doesn't contain H as a subgraph. For a family $\mathcal{F}$ of graphs, we say G is *$\mathcal{F}$-free* if it doesn't contain any member of $\mathcal{F}$ as a subgraph. For a given positive integer n and a graph H, the *extremal number* $\mathrm{ex}(n,H)$ denotes the maximum number of edges in an H-free n-vertex graph. (For a family $\mathcal{F}$, $\mathrm{ex}(n,\mathcal{F})$ is analogously defined for $\mathcal{F}$-free graphs.) A central problem in extremal graph theory is to determine the extremal number $\mathrm{ex}(n,H)$ and the typical structure of H-free graphs. Such a study was initiated by Turán [53] in the 1940s, who determined precisely the extremal number of the

*Research supported by National Science Foundation grant DMS-1855542.
†Research partially supported by National Science Foundation grant DMS-1855542.

DOI: 10.19086/aic.2024.3

complete graph. The problem of studying $\mathrm{ex}(n,H)$ is therefore also referred to as the *Turán problem*. Erdős and Stone [28] (see also [24]) determined asymptotically the extremal number $\mathrm{ex}(n,H)$ for any non-bipartite graph H, thus leaving estimating $\mathrm{ex}(n,H)$ for a bipartite graph H the main remaining challenge in the field. Kővári, Sós and Turán [43] showed that $\mathrm{ex}(n,K_{s,t}) = O(n^{2-1/s})$, where $K_{s,t}$ denotes the complete bipartite graph with part sizes s and t. This was later shown to be asymptotically tight when $t > s!$ by Kollár, Rónyai, and Szabó [44] and when $t > (s-1)!$ by Alon, Rónyai, and Szabó [2]. Both of these were obtained via algebraic constructions. More recently, an innovative random algebraic approach had led to many new tight lower bound constructions, see [12, 13, 14] for instance. In particular, Bukh and Conlon [14] applied the method to settle a long-standing conjecture that asserts that for every rational number α in $[1,2)$ there is a family $\mathcal{F}$ of bipartite graphs for which $\mathrm{ex}(n,\mathcal{F}) = \Theta(n^{\alpha})$. A series of recent progresses have also been made on the related exponent conjecture for single bipartite graphs H (rather than forbidding a family $\mathcal{F}$), resulting in many rationals $\alpha \in [1,2]$ and bipartite graphs H for which $\mathrm{ex}(n,H) = \Theta(n^{\alpha})$, see for instance [18, 19, 34, 35, 37, 38]. Another general result is due to Füredi [30], and independently to Alon, Krivelevich, and Sudakov [1] that asserts that $\mathrm{ex}(n,H) = O(n^{2-1/s})$ for every bipartite graph H where vertices in one part have degree at most s (see [32, 36] for recent generalizations of this result). Alon, Krivelevich, Sudakov [1], in addition, showed that $\mathrm{ex}(n,H) = O(n^{2-1/4s})$ for every s-degenerate bipartite graph H. Despite these substantial progresses on the bipartite Turán problem, it remains to be the case that for most bipartite graphs H there exist substantial gaps between best known lower and upper bounds on $\mathrm{ex}(n,H)$, even for even cycles $C_{2\ell}$, where $\ell \neq 2,3,5$. For more background, the reader is referred to the excellent survey by Füredi and Simonovits [31].

Erdős, Kleitman and Rothschild [22] introduced the problem of counting H-free graphs on n vertices. They showed that there are $2^{(1+o(1))\,\mathrm{ex}(n,K_r)}$ K_r-free graphs, and furthermore that almost all triangle-free graphs are bipartite. Erdős, Frankl and Rödl [21] generalized the former result, showing that for any non-bipartite H there are $2^{(1+o(1))\,\mathrm{ex}(n,H)}$ H-free graphs. Kolaitis, Prömel and Rothschild [42] generalized the latter result, showing that almost all K_r-free graphs are $(r-1)$-partite. This was further extended by Prömel and Steger [49] to all r-critical graphs. Using hypergraph regularity method, Nagle, Rödl and Schacht [48] showed that there are $2^{(1+o(1))\,\mathrm{ex}(n,H)}$ H-free k-uniform hypergraphs for any non-k-partite k-uniform hypergraph H. Balogh, Morris and Samotij [7] and Saxton and Thomason [51] reproved this result using the hypergraph container method. Balogh, Bollobás and Simonovits [3, 4, 5] proved more precise counting and structural results for graphs.

For bipartite H, Kleitman and Winston [40] made the first breakthrough by showing that there are at most $2^{(1+c)n^{3/2}}$ C_4-free graphs on n vertices, where $c \approx 0.0819$ and resolving a long-standing question of Erdős. No further progress was made until the recent significant work of Balogh and Samotij [8, 9], who showed for every $2 \leq s \leq t$ that there are at most $2^{(O(n^{2-1/s}))}$ $K_{s,t}$-free graphs on n vertices. The next major breakthrough was made by Morris and Saxton [45], who showed that the number of $C_{2\ell}$-free graphs on n vertices is at most $2^{O(n^{1+1/\ell})}$, confirming a conjecture of Erdős. There are two key ingredients in Morris and Saxton's work. One is the framework of the container method and the other is the so-called balanced supersaturation property of $C_{2\ell}$, stated in one of their main theorems as below.

Theorem 1.1 (Morris-Saxton [45])**.** *For every $\ell \geq 2$, there exist constants $C > 0, \delta > 0$ and $k_0 \in \mathbb{N}$ such that the following holds for every $k \geq k_0$ and every $n \in \mathbb{N}$. Given a graph G with n vertices and $kn^{1+1/\ell}$ edges,*

there exist a collection $\mathcal{H}$ of copies of $C_{2\ell}$ in G, satisfying:

(a) $|\mathcal{H}| \geq \delta k^{2\ell} n^2$, *and*

(b) $d_{\mathcal{H}}(\sigma) \leq C \cdot k^{2\ell-|\sigma|-\frac{|\sigma|-1}{\ell-1}} n^{1-1/\ell}$ *for every* $\sigma \subset E(G)$ *with* $1 \leq |\sigma| \leq 2\ell-1$, *where* $d_{\mathcal{H}}(\sigma) = |\{A \in \mathcal{H} : \sigma \subset A\}|$ *denotes the 'degree' of the set* σ *in* $\mathcal{H}$.

Using their container method framework and Theorem 1.1, Morris and Saxton [45] were able to not only obtain the enumeration result on the number of $C_{2\ell}$-free graphs on n vertices, but also make further progress on the Turán number of $C_{2\ell}$ in the Erdős-Rényi random graph $G(n,p)$, where as usual $G(n,p)$ denotes the random graph on $[n]$ where each pair ij is included as an edge independently with probability p. Given a graph H, let

$$\mathrm{ex}(G(n,p),H) := \max\{e(G) : G \subset G(n,p) \text{ and } G \text{ is } H\text{-free}\}.$$

Note that both $G(n,p)$ and $\mathrm{ex}(G(n,p),H)$ are random variables.

The problem of determining the threshold function for the maximum number of edges in an H-free subgraph of $G(n,p)$ has received much attention. For more thorough discussion, the reader is referred to the excellent survey by Rödl and Schacht [50]. The most significant work was the following breakthrough, first independently obtained by Conlon and Gowers [17] (under the assumption that H is strictly 2-balanced, see Definition 1.6) and by Schacht [52], then reproved using the container method by Balogh, Morris and Samotij [7] and independently by Saxton and Thomason [51].

Theorem 1.2 (Conlon-Gowers [17], Schacht [52])**.** *Let H be a graph with $\Delta(H) \geq 2$ and chromatic number $\chi(H)$. Let $m_2(H) = \max\{(e(F)-1)/(v(F)-2) : F \subseteq H, v(F) \geq 3\}$. For every $\varepsilon > 0$ there exists a constant $C > 0$ such that if $p \geq Cn^{-1/m_2(H)}$, then*

$$\mathrm{ex}(G(n,p),H) \leq (1 - \frac{1}{\chi(H)-1} + \varepsilon)\binom{n}{2}p,$$

with high probability, as $n \to \infty$.

While Theorem 1.2 gives a satisfactory answer for non-bipartite H, much less is known for bipartite H. For $C_{2\ell}$, Haxell, Kohayakawa and Łuczak [33] showed that if $p \gg n^{-1+1/(2\ell-1)}$ then $\mathrm{ex}(G(n,p),C_{2\ell}) \ll e(G(n,p))$, whereas if $p = o(n^{-1+1/(2\ell-1)})$ then $\mathrm{ex}(G(n,p),C_{2\ell}) = (1+o(1))e(G(n,p))$. For $p = \alpha n^{-1+1/(2\ell-1)}$ and $2 \leq \alpha \leq n^{1/(2\ell-1)^2}$, Kohayakawa, Kreuter and Steger [41] obtained the tight result that with high probability,

$$\mathrm{ex}(G(n,p),C_{2\ell}) = \Theta\left(n^{1+1/(2\ell-1)}(\log\alpha)^{1/(2\ell-1)}\right).$$

For recent work on the analogous problem for linear cycles in random r-uniform hypergraphs, see Mubayi and Yepremyan [47].

Applying the container method to Theorem 1.1, Morris and Saxton obtained the following.

Theorem 1.3 (Morris-Saxton [45])**.** *For every $\ell \geq 2$, there exists a constant $C = C(\ell) > 0$ such that*

$$\mathrm{ex}(G(n,p),C_{2\ell}) \le \begin{cases} Cn^{1+1/(2\ell-1)}(\log n)^2 & \text{if } p \le n^{-(\ell-1)/(2\ell-1)} \cdot (\log n)^{2\ell} \\ Cp^{1/\ell}n^{1+1/\ell} & \text{otherwise} \end{cases}$$

with high probability as $n \to \infty$.

A well-known conjecture of Erdős and Simonovits (see [31]) states that $\mathrm{ex}(n,\{C_3,C_4,\ldots,C_{2\ell}\}) = \Theta(n^{1+1/\ell})$. Morris and Saxton [45] further showed that with high probability as $n \to \infty$, $\mathrm{ex}(n,G(n,p)) = \Omega(p^{1/\ell}n^{1+1/\ell})$ for each ℓ for which the Erdős-Simonovits conjecture is true. The successful applications of Theorem 1.1 to both the enumeration problem and the random Turán problem on $C_{2\ell}$ motivated Morris and Saxton [45] to make a general conjecture about all bipartite graphs.

Conjecture 1.4 (Morris-Saxton [45])**.** *Given a bipartite graph H containing a cycle,*[1] *there exist constants $C > 0, \varepsilon > 0$ and $k_0 \in \mathbb{N}$ such that the following holds. Let $k \ge k_0$, and suppose that G is a graph on n vertices with $k \cdot \mathrm{ex}(n,H)$ edges. Then there exists a (non-empty) collection*[2] *$\mathcal{H}$ of copies of H in G, satisfying*

$$d_{\mathcal{H}}(\sigma) \le \frac{C \cdot |\mathcal{H}|}{k^{(1+\varepsilon)(|\sigma|-1)}e(G)} \text{ for every } \sigma \subset E(G) \text{ with } 1 \le |\sigma| \le e(H).$$

An important aspect of this conjecture is that its truth would immediately yield desired enumeration results on H-free graphs, in the following sense.

Proposition 1.5 (Morris-Saxton [45])**.** *Let H be a bipartite graph. If Conjecture 1.4 holds for H, then there are at most $2^{O(\mathrm{ex}(n,H))}$ H-free graphs on n vertices.*

The work of Morris and Saxton and Conjecture 1.4 generated a lot of interest in the field. In a recent breakthrough, Ferber, McKinley and Samotij [29] were able to establish a weaker version of Conjecture 1.4 and applied it to obtain very general enumeration results on H-free hypergraphs for all r-partite r-uniform hypergraphs (r-graphs in short) H that satisfy a very mild assumption which is widely believed to hold for all r-partite r-graphs. The weaker version of Conjecture 1.4 that Ferber, McKinley and Samotij [29] proved is a bit technical to state here and does not seem to immediately apply to the random Turán problem. However, the enumeration results it yields are very general and significant.

Definition 1.6 (r-density and proper r-density)**.** Let $r \ge 2$ be an integer. Given an r-graph H with $v(H) \ge r+1$, we define its *r-density* to be

$$m_r(H) = \max\{\frac{e(F)-1}{v(F)-r} : F \subseteq H, v(F) > r\}.$$

We define its *proper r-density* to be

$$m_r^*(H) = \max\{\frac{e(F)-1}{v(F)-r} : F \subsetneq H, v(F) > r\}.$$

[1] While the phrase "containing a cycle" does not appear in Morris-Saxton's conjecture, context makes it clear that it was intended. It is easy to see the conjecture is false for forests.

[2] For such a collection to exist, it is neccessary that $|\mathcal{H}| \ge C^{-1}k^{(1+\varepsilon)(e(H)-1)}e(G)$, as can be seen by letting σ be a member of $\mathcal{H}$.

If $m_r(H) > m_r^*(H)$, we say that H is *strictly r-balanced*.

For instance, for the even cycle $C_{2\ell}$, we have $m_2(C_{2\ell}) = \frac{2\ell-1}{2\ell-2}$ and $m_2^*(C_{2\ell}) = 1$. Throughout the paper, we will tacitly assume that H has at least two edges. The main results of Ferber, McKinley and Samotij are

Theorem 1.7 (Ferber-McKinley-Samotij [29])**.** *Let H be an r-uniform hypergraph and let α and A be positive constants. Suppose that $\alpha > r - \frac{1}{m_r(H)}$ and that $\mathrm{ex}(n,H) \le An^\alpha$ for all n. Then there exists a constant C depending only on α, A, and H such that for all n, there are at most 2^{Cn^α} H-free r-uniform hypergraphs on n vertices.*

Theorem 1.8 (Ferber-McKinley-Samotij [29])**.** *Let H be an r-uniform hypergraph and assume that $\mathrm{ex}(n,H) \ge \varepsilon n^{r-\frac{1}{m_r(H)}+\varepsilon}$ for some $\varepsilon > 0$ and all n. Then there exists a constant C depending only on ε and H such that for all n, there are at most $2^{C\,\mathrm{ex}(n,H)}$ H-free r-uniform hypergraphs on n vertices.*

To get a sense of where the condition on H in Theorem 1.8 came from, observe that a simple first moment argument shows that for any r-uniform hypergraph H, $\mathrm{ex}(n,H) \ge \Omega(n^{r-1/m_r(H)})$ holds. In the 2-uniform case, it is widely believed that this simple probabilistic lower bound is not asymptotically tight for any H that contains a cycle. The $r \ge 3$ case is expected to be similar. Ferber, McKinley and Samotij made the following conjecture.

Conjecture 1.9 (Ferber-McKinley-Samotij [29])**.** *Let H be an arbitrary graph that is not a forest. There exists an $\varepsilon > 0$ such that $\mathrm{ex}(n,H) \ge \varepsilon n^{2-1/m_2(H)+\varepsilon}$.*

Conjecture 1.9 is known to hold for quite a few families of bipartite graphs, including complete bipartite graphs, even cycles, the cube graph, and etc (see [31] and [29] for some discussions). In particular, the work of Bukh and Conlon [14] on the Turán exponent of a bipartite family provides a large family of bipartite graphs H for which Conjecture 1.9 holds, namely graphs H that are obtained by gluing enough copies of a so-called balanced tree at the leaves. There is also other strong evidence that Conjecture 1.9 should be true. Bohman and Keevash [11] showed that if H is a bipartite graph that is strictly 2-balanced (see Defintion 1.6) then $\mathrm{ex}(n,H) \ge \Omega(n^{2-\frac{1}{m_2(H)}}(\log n)^{1/(e(H)-1)})$. Bennett and Bohman [10] later generalized this result for hypergraphs (See also [29] for another generalization). On the other hand, a well-known conjecture of Erdős and Simonovits [20] says that for any bipartite graph H, there exist constants $\alpha \in [1,2), c_1, c_2 > 0$ such that $c_1 n^\alpha \le \mathrm{ex}(n,H) \le c_2 n^\alpha$. Hence, if the Erdős-Simonovits conjecture were true, then the result of Bohman and Keevash would imply Conjecture 1.9.

2 Main results

As our main result of the paper, we prove Conjecture 1.4 of Morris and Saxton in a more explicit form under the same mild condition about H assumed by Ferber, McKinley and Samotij [29].

Theorem 2.1 (Balanced Supersaturation)**.** *Let r be an integer with $r \ge 2$. Let H be an r-partite r-graph with h vertices and ℓ edges. Let α and A be positive reals satisfying that $A \ge r^{2r}$, $\alpha > r - \frac{1}{m_r(H)}$ and that*

$\mathrm{ex}(n,H) \le An^\alpha$ *for all n, There exist constants* $k_0, C > 0$ *such that the following holds. Let G be an n-vertex r-graph with* $m = kn^\alpha$ *edges where* $k \ge k_0$*. Then G contains a non-empty family* $\mathcal{H}$ *of copies of H such that ,*

$$d_{\mathcal{H}}(S) \le Ck^{-\lambda(\alpha,H)(|S|-1)}\frac{|\mathcal{H}|}{e(G)}, \text{ for every } S \subseteq E(G), 1 \le |S| \le e(H),$$

where $\lambda(\alpha,H) = \frac{1}{m_r(H)(r-\alpha)}$.

Since $\lambda(\alpha,H) > 1$, Theorem 2.1 resolves Conjecture 1.4 in a more explicit form under the mild assumption that $\alpha > r - 1/m_r(H)$. Our general approach for establishing Theorem 2.1 is inspired by the approach used by Ferber, McKinley and Samotij [29]. However, we also added some crucial new twists, in particular establishing a stronger supersaturation theorem for general bipartite graphs then was currently known. Theorem 2.1 allows one to retrieve Theorem 1.7 and Theorem 1.8.

Theorem 2.1 does not explicitly describe how dense the family $\mathcal{H}$ is. Futhermore, we generalized the balanced supersaturation result in Ferber, McKinley, Samotij [29] to turn any dense family of copies of H in G into a balanced one that is almost as dense. In view of that, we can obtain an even stronger version of Theorem 2.1 for those H that satisfy the following well-known conjecture of Erdős and Simonovits, which roughly says that one can expect to find asymptotically as many copies of H in G as one would expect in a random graph with the same edge-density as G.

Conjecture 2.2 (Erdős-Simonovits)**.** *Let H be a bipartite graph with h vertices and* ℓ *edges. Let* A, α *be positive reals satisfying that* $\mathrm{ex}(n,H) \le An^\alpha$ *for all* $n \in \mathbb{N}$*. There exists constant* C, c*, depending on H such that for all sufficiently large n if G is an n-vertex graph with* $e(G) > Cn^\alpha$ *edges then G contains at least* $c[e(G)]^\ell/n^{2\ell-h}$ *copies of H.*

Given a bipartite graph H, we say that H is *Erdős-Simonovits good* if it satisfies Conjecture 2.2. There are quite a few known Erdős-Simonovits good graphs for appropriate values of α, for instance, even cycles [27] (see also [45]), complete bipartite graphs [25], bipartite graphs that have a vertex complete to the other part [16], tree blowups [32], tree degenerate graphs [36], etc. For Erdős-Simonovits good H, we obtain the following stronger theorem.

Theorem 2.3 (Balanced supersaturation for Erdős-Simonovits good graphs)**.** *Let H be a bipartite graph with h vertices and* ℓ *edges. Suppose that H is Erdős-Simonovits good. Let* α *and A be positive reals satisfying that* $A \ge 16$*,* $\alpha > 2 - \frac{1}{m_2(H)}$ *and that* $\mathrm{ex}(n,H) \le An^\alpha$ *for all n. There exist constants* $\delta_H, k_0, C > 0$ *such that the following holds. Let G be an n-vertex graph with* kn^α *edges where* $k \ge k_0$*. Then G contains a family* $\mathcal{H}$ *of copies of H satisfying that*

1. $|\mathcal{H}| \ge \delta_H[e(G)]^\ell/n^{2\ell-h}$,
2. $\forall S \subseteq E(G), 1 \le |S| \le \ell$,
$$d_{\mathcal{H}}(S) \le C\beta^{|S|-1}\frac{|\mathcal{H}|}{e(G)},$$
where
$$\beta = \max\{(1/k)n^{-\phi(\alpha,H)}, k^{-\lambda^*(\alpha,H)}\},$$

$\phi(\alpha,H)=\frac{\alpha\ell-\alpha+h-2\ell}{\ell-1}$, $\lambda^*(\alpha,H)=\frac{1}{m_2^*(H)(2-\alpha)}$.

Equipped with Theorem 2.1 and Theorem 2.3, we then apply them under the framework of Morris and Saxton [45] to obtain general bounds on $\mathrm{ex}(G(n,p),H)$ as follows.

Theorem 2.4. *Let H be a bipartite graph with h vertices and ℓ edges. Let A,α be positive reals such that $\mathrm{ex}(n,H)\le An^\alpha$ for every $n\in\mathbb{N}$. There exists a constant $C=C(H)$ such that*

$$\mathrm{ex}(G(n,p),H)\le\begin{cases}Cn^{2-\frac{1}{m_2(H)}} & \text{if } p\le n^{-\frac{1}{m_2(H)}},\\ Cp^{1-\frac{1}{\lambda(\alpha,H)}}n^\alpha & \text{otherwise}\end{cases}$$

with high probability as $n\to\infty$.

Theorem 2.5. *Let H be a bipartite graph with h vertices and ℓ edges such that H is Erdős-Simonovits good. Let A,α be positive reals such that $\mathrm{ex}(n,H)\le An^\alpha$ for every $n\in\mathbb{N}$. There exists a constant $C=C(H)$ such that*

$$\mathrm{ex}(G(n,p),H)\le\begin{cases}Cn^{\alpha-\phi(\alpha,H)}(\log(n))^2 & \text{if } p\le n^{-\frac{\phi(\alpha,H)\lambda^*(\alpha,H)}{\lambda^*(\alpha,H)-1}}\cdot(\log n)^{\frac{2\lambda^*(\alpha,H)}{\lambda^*(\alpha,H)-1}}\\ Cp^{1-\frac{1}{\lambda^*(\alpha,H)}}n^\alpha & \text{otherwise}\end{cases}$$

with high probability as $n\to\infty$, where $\phi(\alpha,H)=\frac{\alpha\ell-\alpha+h-2\ell}{\ell-1}$, $\lambda^*(\alpha,H)=\frac{1}{m_2^*(H)(2-\alpha)}$.

Theorem 2.5 implies Morris and Saxton's result on $\mathrm{ex}(G(n,p),C_{2\ell})$ (see Corollary 4.11). To the best of our knowledge, Theorem 2.4 and Theorem 2.5 appear to be the first general results on $\mathrm{ex}(G(n,p),H)$ for bipartite H.

The rest of the paper is organized as follows. In Section 3, we derive our results on balanced supersaturation. In Section 4, we apply our supersaturation results to derive general bounds on the random Turán problem.

3 Balanced supersaturation

We start with a standard estimation lemma.

Lemma 3.1. *Let $n\ge w\ge h$ be positive integers. Then $\binom{n-h}{w-h}/\binom{n}{w}\le(w/n)^h$. Furthermore, if $w\ge h^2$ then $\binom{n-h}{w-h}/\binom{n}{w}\ge(1/2)(w/n)^h$.*

Proof. We have

$$\frac{\binom{n-h}{w-h}}{\binom{n}{w}}=\frac{w(w-1)\dots(w-h+1)}{n(n-1)\cdots(n-h+1)}.$$

Hence, $\binom{n-h}{w-h}/\binom{n}{w}<(w/n)^h$. If $w\ge h^2$, then

$$\frac{\binom{n-h}{w-h}}{\binom{n}{w}}>\frac{w(w-1)\cdots(w-h+1)}{n^h}>\frac{w^h-\binom{h}{2}w^{h-1}}{n^h}>\frac{w^h[1-\binom{h}{2}(1/w)]}{n^h}>\frac{1}{2}\frac{w^h}{n^h}.$$

□

The following simple lemma is folklore. We include a proof for completeness.

Lemma 3.2. *Let $r \geq 2$. Let H be an r-graph. Let G be an n-vertex r-graph with $e(G) > \mathrm{ex}(n,H)$. Then G contains at least $e(G) - \mathrm{ex}(n,H)$ different copies of H.*

Proof. Let m be the number of copies of H in G. Then we can find a set S of at most m edges whose removal destroys all the copies of H in G. Hence $e(G) - m \leq \mathrm{ex}(n,H)$ and thus $m \geq e(G) - ex(n,H)$. □

We next give a nontrivial lower bound on the number of copies of any given r-partite r-graph in a dense enough r-graph. This lower bound may be of independent interest.

Lemma 3.3. *Let r be an integer with $r \geq 2$. Let H be an r-partite r-graph with h vertices. Let α and A be positive reals satisfying that $A \geq r^{2r}$ and that $\mathrm{ex}(n,H) \leq An^{\alpha}$ for all n. There exists a constant $c_H > 0$ such that the following holds. Let G be an n-vertex graph with kn^{α} edges where $k \geq 2^{3r}A$. Then G contains at least $c_H[e(G)]^{\frac{h-\alpha}{r-\alpha}}/n^{\frac{\alpha(h-r)}{r-\alpha}} = c_H k^{\frac{h-\alpha}{r-\alpha}} n^{\alpha}$ copies of H.*

Proof. Let p be a real such that

$$(8A/k)^{\frac{1}{r-\alpha}} \leq p \leq 2(8A/k)^{\frac{1}{r-\alpha}} \quad \text{and} \quad np \in \mathbb{Z}^+.$$

Such p exists since

$$(8A/k)^{\frac{1}{r-\alpha}} \geq (8A/n^{r-\alpha})^{\frac{1}{r-\alpha}} \geq \frac{r^2}{n},$$

by our condition on A. Since $k \geq 2^{3r}A$, $p \in (0,1)$. Let W be a uniform random subset of $V(G)$ of size $w = np$. By our choice of p, we have $w \geq r^2$. By Lemma 3.1,

$$\mathbb{E}[e(G[W])] = e(G)\binom{n-r}{w-r}/\binom{n}{w} \geq \frac{1}{2}e(G)(w/n)^r = \frac{1}{2}e(G)p^r. \tag{3.1}$$

Let $t = \binom{n}{w}$ and $W_1, W_2, \ldots, W_t$ be all the $\binom{n}{w}$ subsets of $V(G)$ of size w. For each $i \in [t]$, let $G_i = G[W_i]$. Let $\mathcal{I}_{good}$ be the set of $i \in [t]$ such that $e(G_i) \geq \frac{1}{4}mp^r$ and $\mathcal{I}_{bad} = [t] \setminus \mathcal{I}_{good}$. Then $\sum_{i \in \mathcal{I}_{bad}} e(G_i) \leq \binom{n}{w}\frac{1}{4}e(G)p^r$ and hence by (3.1)

$$\sum_{i \in \mathcal{I}_{good}} e(G_i) \geq \binom{n}{w}\frac{1}{4}e(G)p^r. \tag{3.2}$$

For each $i \in \mathcal{I}_{good}$, we have

$$e(G_i) \geq \frac{1}{4}e(G)p^r = \frac{1}{4}kn^{\alpha}p^r = \frac{1}{4}kp^{r-\alpha}(np)^{\alpha} = \frac{1}{4}kp^{r-\alpha}w^{\alpha} \geq 2Aw^{\alpha},$$

where the last inequality holds since $p \geq (\frac{8A}{k})^{\frac{1}{r-\alpha}}$. Hence $e(G_i) \geq 2Aw^{\alpha} > 2\,\mathrm{ex}(w,H)$. By Lemma 3.2, G_i contains at least $e(G_i) - \mathrm{ex}(n,H) \geq \frac{1}{2}e(G_i)$ copies of H. Let λ denote the number of copies of H in G. Then, using (3.2), we have

$$\lambda \geq \frac{1}{\binom{n-h}{w-h}} \sum_{i \in \mathcal{I}_{good}} \frac{1}{2}e(G_i) \geq \frac{1}{8}\frac{\binom{n}{w}}{\binom{n-h}{w-h}}e(G)p^r \geq \frac{1}{8}(n/w)^h e(G)p^r = \frac{1}{8}e(G)p^{r-h} = \frac{1}{8}e(G)(1/p)^{h-r}.$$

Since $k = e(G)/n^\alpha$ and $p \le 2(8A/k)^{\frac{1}{r-\alpha}}$, we have

$$\lambda \ge \frac{1}{8} e(G)(1/2)^{h-r}(k/8A)^{\frac{h-r}{r-\alpha}} \ge c_H[e(G)]^{\frac{h-\alpha}{r-\alpha}}/n^{\frac{\alpha(h-r)}{r-\alpha}},$$

for some constant $c_H > 0$. □

While Lemma 3.3 gives a reasonably dense family of copies of H in a dense enough host graph G, it is generally not as dense as what is conjectured in Conjecture 2.2. Next, we present our key lemma, which is the basis of our main results in this paper.

Lemma 3.4 (Key Lemma)**.** *Let r, h, ℓ be positive integers, where $h \ge r \ge 2$. Let H be an r-partite r-graph with h vertices and ℓ edges. Let α, A be positive reals such that for each n every n-vertex graph G with $m \ge An^\alpha$ edges contains at least $f(n,m)$ copies of H, where f is a function satisfying the following.*

1. *There is a constant $\delta > 0$ such that for all $p \in (0,1]$ and all positive reals n, m*

$$f(n,m) \ge \delta m^{\frac{h-\alpha}{r-\alpha}}/n^{\frac{\alpha(h-r)}{r-\alpha}} \quad \text{and} \quad f(np, mp^r) \ge \delta f(n,m)p^h.$$

2. *For fixed n, $f(n,m)$ is increasing and convex in m.*

Then there exist constants $k_0 = k_0(H)$ and $C = C(H) \ge 1$ such that if G is an n-vertex r-graph with kn^α edges, where $k \ge k_0$, then G contains a family $\mathcal{F}$ of copies of H satisfying that

1. $|\mathcal{F}| \ge \delta f(n, \frac{1}{8}e(G))$.
2. $\forall S \subseteq E(G), 1 \le |S| \le \ell - 1$,
$$d_{\mathcal{F}}(S) \le Ck^{-\lambda^*(\alpha,H)(|S|-1)} \frac{|\mathcal{F}|}{e(G)},$$
$\lambda^*(\alpha,H) = \frac{1}{m_r^*(H)(r-\alpha)}$ *and $m_r^*(H)$ is the proper r-density of H.*

Proof. First we define some constants. Let k_0 be a sufficiently large constant, depending on H and A, α, such that the statement after (3.9) holds. Let $N = \delta f(n, \frac{1}{8}e(G))$. By our assumption about f and that $n^\alpha = e(G)/k$,

$$N \ge \delta \cdot \delta(\frac{1}{8}e(G))^{\frac{h-\alpha}{r-\alpha}}/n^{\frac{\alpha(h-r)}{r-\alpha}} = \delta^2(\frac{1}{8})^{\frac{h-\alpha}{r-\alpha}} k^{\frac{h-r}{r-\alpha}} e(G) = \delta' k^{\frac{h-r}{r-\alpha}} e(G), \tag{3.3}$$

where $\delta' = \delta^2(\frac{1}{8})^{\frac{h-\alpha}{r-\alpha}}$. Let

$$C = \max\{2^{\ell+h+2} r^{2h} (8A)^{\frac{h}{r-\alpha}} \ell, \frac{4}{\delta'}\}. \tag{3.4}$$

Trivially, $C \ge 1$. For convenience, let

$$\beta = k^{-\lambda^*(\alpha,H)} \tag{3.5}$$

where $\lambda^*(\alpha,H)$ is as defined in condition 2.

To prove Lemma 3.4, it suffices to find a family $\mathcal{F}$ with $|\mathcal{F}| = N$ such that the following condition holds

$$\forall S \subseteq E(G), 1 \le |S| \le \ell - 1, d_{\mathcal{F}}(S) \le C\beta^{|S|-1}\frac{N}{e(G)}. \tag{3.6}$$

To build such a family, let $\mathcal{F} = \emptyset$. We show that as long as $|\mathcal{F}| < N$, we can find a new copy of H in G to add to $\mathcal{F}$ so that (3.6) holds.

Clearly, initially $\mathcal{F}$ satisfies (3.6). Given a set $S \subseteq E(G)$, where $1 \le |S| \le \ell - 1$, we call S *saturated* if $d_{\mathcal{F}}(S) \ge C\beta^{|S|-1}\frac{N}{2e(G)}$.

Recall that $m_r^*(H) = \max_{F \subsetneq H}\{\frac{e(F)-1}{v(F)-r}\} \ge \frac{\ell-2}{h-r}$. Hence, we have that $\lambda^*(\alpha, H) \le \frac{(h-r)}{(\ell-2)(r-\alpha)}$. Combining this with (3.3), we have

$$C\beta^{|S|-1}\frac{N}{2e(G)} \ge Ck^{-\lambda^*(\alpha,H)(\ell-2)}\frac{\delta' k^{\frac{h-r}{r-\alpha}} e(G)}{2e(G)} \ge \frac{C\delta'}{2} \ge 2. \tag{3.7}$$

For each $i = 1, \ldots, \ell - 1$, let B_i denote the family of saturated i-subsets of $E(G)$. We will call a copy of H in G a *good* copy of H if it does not contain any member of $\bigcup_{i=1}^{\ell-1} B_i$.

Claim 1. For each $i = 1, \ldots, \ell - 1$, $|B_i| \le (2^\ell/C)(1/\beta)^{i-1}e(G)$.

Proof of Claim 1. Let μ denote the number of pairs (H', S) where H' is a member of $\mathcal{F}$ and $S \in B_i$ and $S \subseteq E(H')$. If we count μ by S, then by definition,

$$\mu \ge |B_i| C\beta^{i-1}\frac{N}{2e(G)}.$$

On the other hand, if we count μ by H', then

$$\mu \le |\mathcal{F}|\binom{\ell}{i} < 2^{\ell-1}N.$$

The claim follows by combining the last two inequalities and solving for $|B_i|$. □

Let $q = (8A)^{\frac{1}{r-\alpha}} r^2 \beta^{m_r^*(H)}$. Since $\beta = k^{-\lambda^*(\alpha,H)}$, we have

$$q \ge (8A)^{\frac{1}{r-\alpha}} r^2 k^{-\lambda^*(\alpha,H)m_r^*(H)} = (8A)^{\frac{1}{r-\alpha}} r^2 k^{-\frac{1}{r-\alpha}} \ge r^2 (n^{r-\alpha})^{-\frac{1}{r-\alpha}} = \frac{r^2}{n}. \tag{3.8}$$

Let p be a positive real such that

$$q \le p \le 2q \quad \text{and} \quad np \in \mathbb{Z}^+. \tag{3.9}$$

Since $q \ge r^2/n$, it is easy to see such a p exists. Because $k \ge k_0$, by choosing k_0 to be large enough constant depending on H and A, we can ensure that $2q < 1$ and hence $p \in (0, 1)$.

Let W be a uniform random subset of $V(G)$ of size $w = np$. By (3.8), $w \geq r^2$. By Lemma 3.1,

$$\mathbb{E}[e(G[W])] = e(G)\frac{\binom{n-r}{np-r}}{\binom{n}{np}} \geq \frac{1}{2}e(G)p^r. \tag{3.10}$$

For each $i = 1,\dots,\ell-1$, let $Y_i(W)$ denote the set of members of B_i that are contained in W. Fix any $i = 1,\dots,\ell-1$. Consider any member S of B_i. Suppose S spans v_S vertices. Then by the definition of $m_r^*(H)$, we have $\frac{i-1}{v_S-r} \leq m_r^*(H)$ and hence $v_S \geq r + \frac{i-1}{m_r^*(H)}$. Hence,

$$\mathbb{P}[S \subseteq W] = \binom{n-v_S}{np-v_S}\Big/\binom{n}{np} \leq p^{v_S} \leq p^{r+\frac{i-1}{m_r^*(H)}}.$$

This, along with Claim 1, implies that for each $i = 1,\dots,\ell-1$,

$$\mathbb{E}[|Y_i(W)| \leq |B_i|p^{r+\frac{i-1}{m_r^*(H)}} \leq (2^\ell/C)(1/\beta)^{i-1}p^{r+\frac{i-1}{m_r^*(H)}}e(G) = (2^\ell/C)(1/\beta)^{i-1}p^{\frac{i-1}{m_r^*(H)}} \cdot e(G)p^r. \tag{3.11}$$

Hence, by (3.9) and (3.11), for each $i = 1,\dots,\ell-1$, we have

$$\mathbb{E}[|Y_i(W)|] \leq (2^\ell/C)[2(8A)^{\frac{1}{r-\alpha}}r^2]^{\frac{i-1}{m_r^*(H)}} \cdot e(G)p^r < (2^\ell/C)[2(8A)^{\frac{1}{r-\alpha}}r^2]^h \cdot e(G)p^r.$$

By our choice of C, given in (3.4), we have

$$\mathbb{E}[|\bigcup_{j=1}^{\ell-1} Y_j(W)|] \leq (\ell-1)(2^\ell/C)[2(8A)^{\frac{1}{r-\alpha}}r^2]^h \cdot e(G)p^r \leq \frac{1}{4}e(G)p^r. \tag{3.12}$$

By (3.10) and (3.12),

$$\mathbb{E}[e(G[W]) - |\bigcup_{j=1}^{\ell-1} Y_j(W)|] \geq \frac{1}{2}e(G)p^r - \frac{1}{4}e(G)p^r \geq \frac{1}{4}e(G)p^r, \tag{3.13}$$

Let $t = \binom{n}{np}$ and $W_1,\dots,W_t$ be all the np-subsets of $V(G)$. Let $\mathcal{I}_{good}$ be the set of $i \in [t]$ such that $e(G[W_i]) - |\bigcup_{j=1}^{\ell-1} Y_j(W_i)| \geq \frac{1}{8}e(G)p^r$. Let $\mathcal{I}_{bad} = [t] \setminus \mathcal{I}_{good}$. Then

$$\sum_{i\in\mathcal{I}_{bad}} [e(G[W_i]) - |\bigcup_{j=1}^{\ell-1} Y_j(W_i)|] \leq \binom{n}{np}\frac{1}{8}e(G)p^r.$$

Hence, by (3.13),

$$\sum_{i\in\mathcal{I}_{good}} e(G[W_i]) - |\bigcup_{j=1}^{\ell-1} Y_j(W_i)| \geq \binom{n}{np}\frac{1}{8}e(G)p^r. \tag{3.14}$$

For each $i \in \mathcal{I}_{good}$, let G_i' be a subgraph of $G[W_i]$ obtained by deleting an edge from each member of $\bigcup_{j=1}^{\ell-1} Y_j[W_i]$.

By (3.14),

$$\sum_{i\in\mathcal{I}_{good}} e(G_i') \geq \binom{n}{np}\frac{1}{8}e(G)p^r. \tag{3.15}$$

By the definition of $\mathcal{I}_{good}$, for each $i \in \mathcal{I}_{good}$, we have

$$e(G_i') \geq \frac{1}{8}e(G)p^r = \frac{1}{8}kn^\alpha p^r = \frac{1}{8}kp^{r-\alpha}(np)^\alpha = \frac{1}{8}kp^{r-\alpha}w^\alpha. \tag{3.16}$$

By (3.8) and (3.9),

$$p \geq q \geq (8A)^{\frac{1}{r-\alpha}} r^2 k^{-\frac{1}{r-\alpha}}.$$

Hence, by (3.16), for each $i \in \mathcal{I}_{good}$,

$$e(G_i') \geq \frac{1}{8}k(8A)r^{2(r-\alpha)}k^{-1}w^\alpha \geq Aw^\alpha.$$

Hence, by our assumption about H, for each $i \in \mathcal{I}_{good}$, G_i' contains at least $f(w, e(G_i')) = f(np, e(G_i'))$ copies of H in G. Now, the crucial observation is that any copy H' of H in G_i', where $i \in \mathcal{I}_{good}$, is a good copy of H in G. Indeed, suppose H' contains a member S of B_j for some $j = 1, \ldots, \ell-1$, then $S \subseteq E(G_i') \subseteq E(G[W_i])$. So $S \in Y_j[W_i]$. But in forming G_i' from $G[W_i]$ we have removed an edge from each member of $Y_j[W_i]$ and hence $S \not\subseteq E(G_i')$, a contradiction.

Let $\mathcal{H}_{good}$ denote the family of good copies of H in G. By our discussions above, (3.15) and our assumptions about the function f, we have

$$\begin{aligned} |\mathcal{H}_{good}| &\geq \frac{1}{\binom{n-h}{np-h}} \sum_{i\in\mathcal{I}_{good}} f(np, e(G_i')) \geq \frac{1}{\binom{n-h}{np-h}} \binom{n}{np} f(np, \frac{1}{8}e(G)p^r) \\ &\geq (1/p)^h \delta f(n, \frac{1}{8}e(G))p^h = \delta f(n, \frac{1}{8}e(G)). \end{aligned}$$

Hence, $|\mathcal{H}_{good}| \geq N > |\mathcal{F}|$. So, there must exist a member H' of $\mathcal{H}_{good}$ that is not in $\mathcal{F}$. Let us add H' to $\mathcal{F}$. Consider any subset $S \subseteq E(G)$, where $1 \leq |S| \leq \ell-1$. If S is not contained in H' then $d_{\mathcal{F}}(S)$ is unchanged. If $S \subseteq E(H')$, then since H' is good, S is unsaturated prior to the addition of H' and hence now satisfies $d_{\mathcal{F}}(S) \leq C\beta^{|S|-1}\frac{N}{2e(G)} + 1 \leq C\beta^{|S|-1}\frac{N}{e(G)}$, by (3.7). Hence (3.6) still holds for the new family $\mathcal{F}$. Thus, we can iterate the process to find a family $\mathcal{F}$ that satisfies (3.6) and such that $|\mathcal{F}| \geq N$. □

For convenience, we will refer to the β associated with the family $\mathcal{F}$ as *the codegree ratio* of $\mathcal{F}$.

By Lemma 3.3 and Lemma 3.4, we obtain the following general balanced supersaturation theorem, which implies our first main theorem, Theorem 2.1.

Theorem 3.5. *Let r be an integer with $r \geq 2$. Let H be an r-partite r-graph with h vertices and ℓ edges. Let α and A be positive reals satisfying that $A \geq r^{2r}$, $\alpha > r - \frac{1}{m_r(H)}$ and that $\mathrm{ex}(n,H) \leq An^\alpha$ for all n, There exist constants $\delta_H, k_0 = k_0(H), C = C(H) > 0$ such that the following holds. Let G be an n-vertex graph with kn^α edges where $k \geq k_0$. Then G contains a family $\mathcal{F}$ of copies of H satisfying that*

1. $|\mathcal{F}| \geq \delta_H [e(G)]^{\frac{h-\alpha}{r-\alpha}} / n^{\frac{\alpha(h-r)}{r-\alpha}}$,
2. $\forall S \subseteq E(G), 1 \leq |S| \leq \ell$
$$d_{\mathcal{F}}(S) \leq C k^{-\lambda(\alpha,H)(|S|-1)} \frac{|\mathcal{F}|}{e(G)},$$
where $\lambda(\alpha,H) = \frac{1}{m_r(H)(r-\alpha)}$ *and* $m_r(H)$ *is the r-density of* H.

Proof. By choosing k_0 to be at least 2^{3r}, by Lemma 3.3, H has the property that every n-vertex graph with kn^α edges, where $k \geq k_0$ contains at least $c_H [e(G)]^{\frac{h-\alpha}{r-\alpha}} / n^{\frac{\alpha(h-r)}{r-\alpha}}$ copies of H. Let $f(n,m) = c_H m^{\frac{h-\alpha}{r-\alpha}} / n^{\frac{\alpha(h-r)}{r-\alpha}}$. It is straightforward to see that $f(np, mp^r) = f(n,m)p^h$, for any p. Furthermore, for fixed n, $f(n,m)$ is clearly increasing and convex in m. By Lemma 3.4, there exist constants k_0 and C' such that for every n if G is an n-vertex r-graph with kn^α edges, where $k \geq k_0$, then G contain a family $\mathcal{F}$ of copies of H satisfying

(a) $|\mathcal{F}| \geq c_H f(n, \frac{1}{8} e(G))$.

(b) $\forall S \subseteq E(G), 1 \leq |S| \leq \ell$,
$$d_{\mathcal{F}}(S) \leq C' \beta^{|S|-1} \frac{|\mathcal{F}|}{e(G)},$$
where
$$\beta = \max\left\{ \left(\frac{e(G)}{|\mathcal{F}|} \right)^{\frac{1}{\ell-1}}, k^{-\lambda(\alpha,H)} \right\},$$
$\lambda(\alpha,H) = \frac{1}{m_r(H)(r-\alpha)}$. Note that here we used the fact that $\lambda(\alpha,H) \leq \lambda^*(\alpha,H)$, as to make the conclusion work for sets of size ℓ, we switch to λ.

Since $e(G) = kn^\alpha$, by condition (a) above,
$$|\mathcal{F}| \geq c_H [e(G)]^{\frac{h-\alpha}{r-\alpha}} / n^{\frac{\alpha(h-r)}{r-\alpha}} = c_H k^{\frac{h-r}{r-\alpha}} e(G). \tag{3.17}$$
For convenience, we may further assume that $c_H < 1$. By definition of r-density, $\frac{\ell-1}{h-r} \leq m_r(H)$. Hence, by (3.17),
$$\left(\frac{e(G)}{|\mathcal{F}|} \right)^{\frac{1}{\ell-1}} \leq (c_H^{-1} k^{-\frac{h-r}{r-\alpha}})^{\frac{1}{\ell-1}} \leq (c_H)^{-\frac{1}{\ell-1}} k^{-\lambda(\alpha,H)}. \tag{3.18}$$
By the definition of β in condition (b) and (3.18),
$$\beta \leq c_H^{-\frac{1}{\ell-1}} k^{-\lambda(\alpha,H)}.$$
Let $C = C'(c_H)^{-1}$. By condition (b) above, we have $\forall S \subseteq E(G), 1 \leq |S| \leq \ell$,
$$d_{\mathcal{F}}(S) \leq C' \beta^{|S|-1} \frac{|\mathcal{F}|}{e(G)} \leq C' (c_H^{-\frac{|S|-1}{\ell-1}} k^{-\lambda(\alpha,H)(|S|-1)|}) \frac{|\mathcal{F}|}{e(G)} \leq C k^{-\lambda(\alpha,H)(|S|-1)|} \frac{|\mathcal{F}|}{e(G)}.$$
So, the theorem holds with $\delta_H = c_H$. □

If one applies Theorem 3.5 to an adaption of Proposition 1.5, one can retrieve one of the enumeration results of Ferber, McKinley and Samotij [29] as below. The difference is that Ferber, McKinley and Samotij used a weaker version of balanced supersaturation.

Corollary 3.6. *Let H be an r-uniform hypergraph and let α and A be positive constants. Suppose that $\alpha > r - \frac{1}{m_r(H)}$ and that $\mathrm{ex}(n,H) \le An^\alpha$ for all n. Then there exists a constant C depending only on α, A, and H such that for all n, there are at most 2^{Cn^α} H-free r-uniform hypergraphs on n vertices.*

Next, we give a stronger balanced supersaturation theorem for Erdős-Simonovits good bipartite graphs. This implies our second main result, Theorem 2.3.

Theorem 3.7. *Let H be a bipartite graph with h vertices and ℓ edges. Suppose that H is Erdős-Simonovits good. Let α and A be positive reals satisfying that $A \ge 16$, $\alpha > 2 - \frac{1}{m_2(H)}$ and that $\mathrm{ex}(n,H) \le An^\alpha$ for all n. There exist constants $\delta_H, k_0 = k_0(H), C = C(H) > 0$ such that the following holds. Let G be an n-vertex graph with kn^α edges where $k \ge k_0$. Then G contains a family $\mathcal{F}$ of copies of H satisfying that*

1. $|\mathcal{F}| \ge \delta_H [e(G)]^\ell / n^{2\ell-h}$,
2. $\forall S \subseteq E(G), 1 \le |S| \le \ell$,
$$d_{\mathcal{F}}(S) \le C\beta^{|S|-1} \frac{|\mathcal{F}|}{e(G)},$$
where
$$\beta = \max\{k^{-1} n^{-\phi(\alpha,H)}, k^{-\lambda^*(\alpha,H)}\},$$
$\phi(\alpha,H) = \frac{\alpha\ell - \alpha + h - 2\ell}{\ell - 1}$, $\lambda^*(\alpha,H) = \frac{1}{m_2^*(H)(2-\alpha)}$ *and $m_2^*(H)$ is the proper 2-density of H.*

Proof. Since H is Erdős-Simonovits good, there exist constants $c_H, k_1 > 0$ such that n-vertex graph with $m = kn^\alpha$ edges, where $k \ge k_1$ contains at least $c_H m^\ell / n^{2\ell-h}$ copies of H. Let $f(n,m) = c_H m^\ell / n^{2\ell-h}$. Using the facts $\alpha \ge 2 - \frac{1}{m_r} \ge 2 - \frac{h-2}{\ell-1}$ and $k \le 2 - \alpha$, one can show that $f(n,m) \ge c_H m^{\frac{h-\alpha}{2-\alpha}} / n^{\frac{\alpha(h-2)}{2-\alpha}}$. Also, it is straightforward to see that $f(np, mp^2) = f(n,m)p^h$, for any p. Furthermore, for fixed n, $f(n,m)$ is clearly increasing and convex in m. By Lemma 3.4, there exist constants k_0 and C' such that if G is an n-vertex r-graph with kn^α edges, where $k \ge k_0$, then G contain a family $\mathcal{F}$ of copies of H satisfying

(a) $|\mathcal{F}| \ge c_H f(n, \frac{1}{8}e(G))$.

(b) $\forall S \subseteq E(G), 1 \le |S| \le \ell$,
$$d_{\mathcal{F}}(S) \le \hat{C}\hat{\beta}^{|S|-1} \frac{|\mathcal{F}|}{e(G)},$$
where
$$\hat{\beta} = \max\left\{ \left(\frac{e(G)}{|\mathcal{F}|} \right)^{\frac{1}{\ell-1}}, k^{-\lambda^*(\alpha,H)} \right\},$$
$\lambda^*(\alpha,H) = \frac{1}{m_2^*(H)(2-\alpha)}$ and $m_2^*(H)$ is the proper 2-density of H.

By condition (a) above,

$$|\mathcal{F}| \geq c_H[e(G)]^{\ell}/n^{2\ell-h}. \tag{3.19}$$

For convenience, we may further assume that $c_H < 1$. Hence,

$$|\mathcal{F}|/e(G) \geq c_H[e(G)]^{\ell-1}/n^{2\ell-h} = c_H k^{\ell-1} n^{\alpha\ell-\alpha+h-2\ell} = c_H k^{\ell-1} n^{\phi(\alpha,H)(\ell-1)}.$$

Hence, by the definition of $\hat{\beta}$ given in condition (b) above,

$$\hat{\beta} \leq \max\{c_H^{-\frac{1}{\ell-1}} k^{-1} n^{-\phi(\alpha,H)}, k^{-\lambda^*(\alpha,H)}\}.$$

Let

$$\beta = \max\{k^{-1}n^{-\phi(\alpha,H)}, k^{-\lambda^*(\alpha,H)}\} \quad \text{and} \quad C = \hat{C} \cdot c_H^{-1}.$$

It is straightforward to verify that condition 2 holds for these choices of β and C. So the theorem holds with $\delta_H = c_H$. □

The advantage of Theorem 3.7 over Theorem 3.5 is that for Erdős-Simonovits good H, the former produces a β value that is no larger than the former and in many cases produces a smaller β, at least for a suitable range of k. Next, we show that the method developed in this section allows us to give a short proof of Theorem 1.1 of Morris and Saxton [45].

Theorem 3.8 (Restatement of Theorem 1.1)**.** *For every $\ell \geq 2$, there exist constants $C > 0, \delta > 0$ and $k_0 \in \mathbb{N}$ such that for each $n \in \mathbb{N}$ if G is an n-vertex graph with $kn^{1+1/\ell}$ edges, where $k_0 \leq k$, then there exists a collection $\mathcal{F}$ of copies of $C_{2\ell}$ in G such that*

(a) $|\mathcal{F}| \geq \delta k^{2\ell} n^2$,

(b) $\forall S \subseteq E(G), 1 \leq |S| \leq 2\ell - 1$, $d_{\mathcal{F}}(S) \leq Ck^{2\ell-|S|-\frac{|S|-1}{\ell-1}} n^{1-1/\ell}$.

Proof. It is well known that $C_{2\ell}$ is Erdős-Simonovits good for $\alpha = 1 + 1/\ell$ and some $A > 0$, so we can apply Lemma 3.4 with $f(n,m) = c\left(\frac{m}{n}\right)^{2\ell}$ with c some positive constant depending only on $C_{2\ell}$. Let G be an n-vertex graph with kn^{α} edges, with $k_0 \leq k$. Let $\mathcal{F}$ be the family of copies of H in G guaranteed by Lemma 3.4. One can check that $m_2^*(C_{2\ell}) = 1$ and hence $\lambda^*(\alpha, C_{2\ell}) = \frac{1}{1\cdot(2-\alpha)} = \frac{\ell}{\ell-1}$. Condition (a), (b) readily follow from conditions (1),(2) of $\mathcal{F}$ guaranteed in Lemma 3.4. □

4 Applications to the Turán problem in random graphs

In this section, we apply our balanced supersaturation results to obtain some general bounds on the Turán number of a bipartite graph H in the Erdős-Rényi random graph $G(n,p)$. In fact, once we have Theorem 3.5 and Theorem 3.7, the corresponding random Turán results will readily follow using the framework set up by Morris and Saxton [45]. Nevertheless, for completeness, we will include all the technical details. The framework is based on the container method pioneered by Balogh, Morris, and Samotij [7] and independently by Saxton and Thomason [51].

Definition 4.1. Given an r-graph $\mathcal{F}$, define the *co-degree function* of $\mathcal{F}$

$$\delta(\mathcal{F},\tau)=\frac{1}{|\mathcal{F}|}\sum_{j=2}^{r}\frac{1}{\tau^{j-1}}\sum_{v\in V(\mathcal{F})}d^{(j)}(v),$$

where

$$d^{(j)}(v)=\max\{d_{\mathcal{F}}(S):v\in S\subseteq v(\mathcal{F}) \text{ and } |S|=j\}.$$

We need the following theorem from Morris and Saxton [45], which is a quick consequence of analogous theorems of Balogh, Morris and Samotij [51] Theorem 6.2 and of Saxton and Thomason [7] Proposition 3.1.

Theorem 4.2. *Let $r\geq 2$ be an integer. Let $0<\delta<\delta_0(r)$ be a sufficiently small real. Let $\mathcal{F}$ be an r-graph with N vertices. Suppose that $\delta(\mathcal{F},\tau)\leq\delta$ for some $\tau>0$. Then there exists a collection $\mathcal{C}$ of subsets of $V(\mathcal{F})$ and a function $f:V(\mathcal{F})^{(\leq\tau N/\delta)}\to\mathcal{C}$ such that*

(a) *For every independent set I in $\mathcal{F}$ there exists $T\subset I$ with $|T|\leq\tau N/\delta$ and $I\subset f(T)$.*

(b) *$e(\mathcal{F}[C])\leq(1-\delta)e(\mathcal{F})$ for every $C\in\mathcal{C}$.*

Proposition 4.3. *Let H be a bipartite graph with h vertices and ℓ edges. Let α,A be positive reals satisfying that for each $n\in\mathbb{N}$ every n-vertex graph G with $m\geq An^{\alpha}$ edges contains at least $f(n,m)$ copies of H, where f is a function satisfying the following.*

1. *There is a constant $\delta>0$ such that for all $p\in(0,1]$ and all positive reals n,m*

$$f(n,m)\geq\delta m^{\frac{h-\alpha}{2-\alpha}}/n^{\frac{\alpha(h-2)}{2-\alpha}} \quad \text{and} \quad f(np,mp^2)\geq\delta f(n,m)p^h.$$

2. *For fixed n, $f(n,m)$ is increasing and convex in m.*

Let $k_0,C,\mathcal{F}$ and β be as guaranteed by Lemma 3.4. There exist $k_0^*\in\mathbb{N}$ and a real $\varepsilon>0$ such that the following holds for every $k\geq k_0^*$ and every $n\in\mathbb{N}$. Set

$$\mu=\frac{k\beta}{\varepsilon}.$$

Given a graph G with n vertices and kn^{α} edges, there exists a function f_G that maps subgraphs of G to subgraphs of G such that for every H-free subgraph $I\subset G$,

(a) *There exists a subgraph $T=T(I)\subset I$ with $e(T)\leq\mu n^{\alpha}$ and $I\subset f_G(T)$, and*

(b) *$e(f_G(T(I)))\leq(1-\varepsilon)e(G)$.*

Proof. Note that condition 1 in the proposition still holds if we replace δ with an even smaller positive real. Hence in the our proof, we may assume δ to be sufficiently small. Let $\mathcal{F}$ be the family guaranteed by Lemma 3.4 with codegree ratio β. Let $N=v(\mathcal{F})=e(G)=kn^{\alpha}$. Since we will view $\mathcal{F}$ as a hypergraph, we will write $e(\mathcal{F})$ for $|\mathcal{F}|$. Set

$$\frac{1}{\tau}=\frac{\delta^2}{\beta} \quad \text{and} \quad \varepsilon=\delta^3.$$

Since $\forall S \subseteq E(G), 1 \le |S| \le \ell-1, d_{\mathcal{F}}(S) \le C\beta^{|S|-1}e(\mathcal{F})/N$ holds, we have

$$\frac{1}{e(\mathcal{F})}\sum_{j=2}^{\ell-1}\frac{1}{\tau^{j-1}}\sum_{v\in V(\mathcal{F})}d^{(j)}(v) \le \frac{1}{e(\mathcal{F})}\sum_{j=2}^{\ell-1}(\frac{\delta^2}{\beta})^{j-1}\cdot N\left(C\beta^{j-1}\frac{e(\mathcal{F})}{N}\right) < C\sum_{j=2}^{\ell-1}\delta^{2j-2} \le 2C\delta^2. \tag{4.1}$$

Also, since $\forall v \in V(\mathcal{F})$, $d^{(\ell)}(v) \le 1$ and $C\beta^{\ell-1}e(\mathcal{F})/N \ge 1$, we have

$$\frac{1}{e(\mathcal{F})}\cdot\frac{1}{\tau^{\ell-1}}\sum_{v\in V(F)}d^{(\ell)}(v) \le \frac{N}{e(\mathcal{F})}\cdot(\frac{\delta^2}{\beta})^{\ell-1} \le C\delta^{2\ell-2} \le C\delta^2,$$

By our discussion above, we get

$$\delta(\mathcal{F},\tau) = \frac{1}{e(\mathcal{F})}\sum_{j=2}^{\ell}\frac{1}{\tau^{j-1}}\sum_{v\in V(\mathcal{F})}d^{(j))}(v) \le 2C\delta^2 + C\delta^2 \le \delta. \tag{4.2}$$

By Theorem 4.2, there exist a collection $\mathcal{C}$ of subsets of $V(\mathcal{F})$ and a function $f_G : V(\mathcal{F})^{(\le \tau N/\delta)} \to \mathcal{C}$ such that for every H-free subgraph $I \subset G$,

(a') there exists $T = T(I) \subset I$ with $e(T) \le \tau N/\delta$ and $I \subset f_G(T)$, and

(b') $e(\mathcal{F}[T(I)])) \le (1-\delta)e(\mathcal{F})$.

In condition (a') we have

$$e(T) \le \tau N/\delta = (\beta/\delta^5)e(G) = (\beta/\varepsilon)kn^{\alpha} = \mu n^{\alpha},$$

condition (a') is equivalent to condition (a). To complete the proof, it suffices to show that if I is an independent set in $\mathcal{F}$ (i.e. if I is an H-free subgraph of G) we have $e(f_G(T(I))) \le (1-\varepsilon)e(G)$. Let $D = f_G(T(I))$. By condition (b'), $e(\mathcal{F}[D]) \le (1-\delta)e(\mathcal{F})$. Hence, if we delete $v(\mathcal{F})\setminus D$ from $\mathcal{F}$ we lose at least $\delta e(\mathcal{F})$ edges of $\mathcal{F}$. On the other hand, by our assumption about $\mathcal{F}$, each vertex of $\mathcal{F}$ lies in at most $Ce(\mathcal{F})/e(G)$ edges of $\mathcal{F}$. Hence if we delete $V(\mathcal{F})\setminus C$ from $\mathcal{F}$, we lose at most $(v(\mathcal{F})-|D|)Ce(\mathcal{F})/e(G)$ edges of $\mathcal{F}$. Hence, we have

$$\delta e(\mathcal{F}) \le (v(\mathcal{F})-|D|)Ce(\mathcal{F})/e(G).$$

Solving for $|D|$ and using $v(\mathcal{F}) = e(G)$, we have

$$|D| \ge v(\mathcal{F}) - (\delta/C)e(G) = [1-(\delta/C)]e(G) \ge (1-\varepsilon)e(G).$$

In other words, we have $e(f_G(T(I)) \ge (1-\varepsilon)e(G)$, as desired. □

Remark 4.4. When we apply Proposition 4.3, we can take $\beta = k^{-\lambda(\alpha,H)}$ as in Theorem 3.5 and for Erdős-Simonovits good H, we can take $\beta = \max\{k^{-1}n^{-\phi(\alpha,H)}, k^{-\lambda^*(\alpha,H)}\}$ as in Theorem 3.7.

We need an estimation lemma from [45].

Lemma 4.5 ([45]). *Let $M > 0, s > 0$ and $0 < \delta < 1$. If $a_1, \ldots, a_m$ are reals that satisfy $s = \sum_j a_j$ and $1 \le a_j \le (1-\delta)^j M$ for each $j \in [m]$, then*

$$s \log s \le \sum_{j=1}^{m} a_j \log a_j + O(M).$$

In what follows, by a *colored graph*, we mean a graph together with a labelled partition of its edge set. Next, we prove two similar theorems, by adapting the arguments given in Section 6 of Morris-Saxton [45] to fit our general balanced supersaturation results. The former applies to all bipartite graphs that contain a cycle and the latter applies to Erdős-Simonovits good bipartite graphs that contain a cycle.

Theorem 4.6. *Let H be a bipartite graph with h vertices and ℓ edges that contains a cycle. Let A, α be positive reals such that $\mathrm{ex}(n,H) \le An^\alpha$ holds for every $n \in \mathbb{N}$. There exists a constant C such that the following holds for all sufficiently large $n \in \mathbb{N}$ and $k \in \mathbb{R}^+$. Let $\mathcal{I}(n)$ denote all the H-free graphs on $[n]$ and $\mathcal{G}(n,k)$ the collection of all graphs on $[n]$ with at most kn^α edges. There exists a collection $\mathcal{S}$ of colored graphs with n vertices and at most $Ck^{1-\lambda(\alpha,H)}n^\alpha$ edges and functions*

$$g : \mathcal{I} \to \mathcal{S} \quad \text{and} \quad h : \mathcal{S} \to \mathcal{G}(n,k)$$

with the following properties

(a) $\forall s \ge 1$ *the number of colored graphs in $\mathcal{S}$ with s edges is at most*

$$\left(\frac{Cn^\alpha}{s}\right)^{\frac{\lambda(\alpha,H)}{\lambda(\alpha,H)-1}\cdot s} \cdot \exp\left(Ck^{1-\lambda(\alpha,H)}n^\alpha\right).$$

(b) $\forall I \in \mathcal{I}(n)\ g(I) \subset I \subset h(g(I)) \cup g(I)$.

Proof. Note that since H contains a cycle, $m_2(H) \ge 1$. Let $I \in \mathcal{I}(n)$. We will apply Proposition 4.3 repeatedly (with $\beta = k^{-\lambda(\alpha,H)}, \mu = k\beta/\varepsilon = (1/\varepsilon)k^{1-\lambda(\alpha,H)}$). Let $G_0 = K_n$. For sufficiently large n, G_0 clearly satisfies the condition on G in Proposition 4.3. Apply Proposition 4.3, with G_0 playing the role of G to obtain the function f_{G_0} and a subset T_1 of I with $T_1 \subset I \subset f_{G_0}(T_1)$ where T_1 and $f_{G_0}(T_1)$ satisfy the additional properties described in Proposition 4.3. Now, let $G_1 = f_{G_0}(T_1) \setminus T_1$ and $I_1 = I \cap G_1 = I \setminus T_1$. Apply Proposition 4.3 again, with G_1 playing the role of G and I_1 playing the role of I to obtain the function f_{G_1} and a subset T_2 of I_1 with $T_2 \subset I_1 \subset f_{G_1}(T_2)$. We continue like this until we arrive at a graph $G_{m(I)}$ with at most kn^α edges.

Let $g(I) = T_1 \cup T_2 \cup \cdots \cup T_{m(I)}$, where elements of T_i are colored with color i. Let $\mathcal{S}(s) = \{g(I) : |g(I)| = s\}$. Let $h(g(I)) = G_{m(I)}$. Note that h is well-defined (see [7, 45, 51] for detailed discussion). Furthermore, as $g(I) \subset I \subset h(g(I)) \cup g(I)$, conditions (b) is fulfilled. It remains to show (a). We begin by partitioning $\mathcal{S}(s)$ into sets $\mathcal{S}_m(s)$ where $\mathcal{S}_m(s) = \{S \in \mathcal{S}(s) :$ the edges of S are colored with m colors$\}$.

For each $m \in \mathbb{N}$, let

$$\mathcal{K}(m) = \{\mathbf{k} = (k_1, \cdots k_m) : k_j \in \mathbb{R}, (1-\varepsilon)^{j-m}k \le k_j \le (1-\varepsilon)^j n^{2-\alpha} \text{ and } k_j n^\alpha \in \mathbb{N}\}$$

And for each $\mathbf{k} \in \mathcal{K}(m)$, let

$$\mathcal{A}(\mathbf{k}) = \{\mathbf{a} = (a_1, \dots a_m) : a_j \in \mathbb{N}, a_j \leq \frac{1}{\varepsilon} k_j^{1-\lambda(\alpha,H)} \text{ and } \sum_j a_j = s\}$$

By definition each sequence in $\mathbf{k} \in \mathcal{K}(m)$ corresponds to a potential sequence $(G_1, \cdots G_m)$ where $e(G_j) = k_j n^\alpha$, as the edges of $(1-\varepsilon)^{j-m} k n^\alpha \leq e(G_j) \leq (1-\varepsilon)^j n^2$ and by Proposition 4.3, we have $e(T_j) \leq \frac{1}{\varepsilon} k_j^{1-\lambda(\alpha,H)} n^\alpha$. Note further that our algorithm returns pairs (G_i, T_i), such that each sequence of $(T_1, \cdots T_m)$ is uniquely identified with a sequence $(G_1, \cdots G_m)$. Thus it suffices to only count the choices of T_i. The sequence $\mathbf{a} \in \mathcal{A}(\mathbf{k})$ corresponds to a sequence of sizes for T_i. Thus, we have that for a fixed m

$$|\mathcal{S}_m(s)| \leq \sum_{\mathbf{k} \in \mathcal{K}(m)} \sum_{\mathbf{a} \in \mathcal{A}(\mathbf{k})} \prod_{j=1}^{m} \binom{k_j n^\alpha}{a_j}.$$

Since for each $j \in [m]$

$$a_j \leq (1/\varepsilon) k_j^{1-\lambda(\alpha,H)} \leq (1/\varepsilon)[(1-\varepsilon)^{-j+1} k]^{1-\lambda(\alpha,H)} = (1/\varepsilon)(1-\varepsilon)^{(\lambda(\alpha,H)-1)(j-1)} k^{1-\lambda(\alpha,H)}.$$

So

$$k_j \leq (1/\varepsilon)^{\frac{1}{\lambda(\alpha,H)-1}} n / a_j^{\frac{1}{\lambda(\alpha,H)-1}}.$$

Hence

$$\binom{k_j n^\alpha}{a_j} \leq \left(\frac{e k_j n^\alpha}{a_j}\right)^{a_j} \leq \left(\frac{n^\alpha}{\varepsilon a_j}\right)^{\frac{\lambda(\alpha,H)}{\lambda(\alpha,H)-1} \cdot a_j}$$

Applying Lemma 4.5 with $M = (1/\varepsilon) k^{1-\lambda(\alpha,H)} n^\alpha, 1-\delta = (1-\varepsilon)^{\lambda(\alpha,H)-1}$, the product over $j = 1, \dots, m$ is at most

$$\left(\frac{C' n^\alpha}{s}\right)^{\frac{\lambda(\alpha,H)}{\lambda(\alpha,H)-1} \cdot s} \cdot \exp\left(C' k^{1-\lambda(\alpha,H)} n^\alpha\right),$$

for some $C' = C'(H)$.

Thus,

$$|\mathcal{S}(s)| \leq \sum_{m=1}^{\infty} \sum_{\mathbf{k} \in \mathcal{K}(m)} \sum_{\mathbf{a} \in \mathcal{A}(\mathbf{k})} \left(\frac{C' n^\alpha}{s}\right)^{\frac{\lambda(\alpha,H)}{\lambda(\alpha,H)-1} \cdot s} \cdot \exp\left(C' k^{1-\lambda(\alpha,H)} n^\alpha\right)$$

Note that $|\mathcal{K}(m)| = 0$ for values of $m \gg \log n$ as $(1-\varepsilon)^m n^2 < k$. For any fixed $m = O(\log n)$, $|\mathcal{K}(m)| = n^{O(\log n)}$, as when picking any $\mathbf{k}$, we have $O(n^2)$ choices for each coordinate and $O(\log n)$ coordinates to choose for. For any fixed $\mathbf{k}$, similarly, $|\mathcal{A}(\mathbf{k})| = n^{O(\log n)}$. Hence,

$$\sum_{m=1}^{\infty} \sum_{\mathbf{k} \in \mathcal{K}(m)} |\mathcal{A}(\mathbf{k})| = n^{O(\log n)} = \exp(O(\log^2 n)).$$

Since $m_2(H) \geq 1$, $\lambda(\alpha, H) \leq 1/(2-\alpha)$. This imples $(2-\alpha)(1-\lambda) \geq 1-\alpha$. Since $k \leq n^{2-\alpha}$, we have

$k^{1-\lambda(\alpha,H)}n^{\alpha} \geq n \gg \log^2 n$. Hence, the theorem holds by choosing an appropriate $C > C'$. □

Remark 4.7. The proof of Theorem 4.6 can be adapted for r-partite r-graphs as well. However, as the random Turán problem in hypergraphs is not a focus of this paper, we opt to state Theorem 4.6 just for bipartite graphs.

Theorem 4.8. *Let H be a bipartite graph with h vertices and ℓ edges that contain a cycle. Let A, α be positive reals such that $\mathrm{ex}(n,H) \leq An^{\alpha}$ holds for every $n \in \mathbb{N}$. Suppose that H is Erdős-Simonovits good. There exists a constant C such that the following holds for all sufficiently large $n \in \mathbb{N}$ and $k \in \mathbb{R}^+$ with*

$$k \leq n^{\frac{\phi(\alpha,H)}{\lambda^*(\alpha,H)-1}}/(\log n)^{\frac{2}{\lambda^*(\alpha,H)-1}}.$$

Let $\mathcal{I}(n)$ denote all the H-free graphs on $[n]$ and $\mathcal{G}(n,k)$ the collection of all graphs on $[n]$ with at most kn^{α} edges. There exists a collection $\mathcal{S}$ of colored graphs with n vertices and at most $Ck^{1-\lambda^*(H)}n^{\alpha}$ edges and functions

$$g : \mathcal{I} \to \mathcal{S} \quad \text{and} \quad h : \mathcal{S} \to \mathcal{G}(n,k)$$

with the following properties

(a) $\forall s \geq 1$ *the number of colored graphs in $\mathcal{S}$ with s edges is at most*

$$\left(\frac{Cn^{\alpha}}{s}\right)^{\frac{\lambda^*(\alpha,H)}{\lambda^*(\alpha,H)-1}\cdot s} \cdot \exp\left(Ck^{1-\lambda^*(\alpha,H)}n^{\alpha}\right).$$

(b) $\forall I \in \mathcal{I}(n)\ g(I) \subset I \subset h(g(I)) \cup g(I)$.

Proof. Note that since H contains a cycle, $m_2^*(H) \geq 1$. Let $I \in \mathcal{I}(n)$. As in the proof of Theorem 4.6 we will apply Proposition 4.6 repeatedly. Since H is Erdős-Simonovits good, we can use

$$\beta = \max\{k^{-1}n^{-\phi(\alpha,H)}, k^{-\lambda^*(\alpha,H)}\} \text{ and } \mu = k\beta/\varepsilon = (1/\varepsilon)\max\{n^{-\phi(\alpha,H)}, k^{1-\lambda^*(\alpha,H)}\}.$$

let $G_0 = K_n$. We apply Proposition 4.3, with G_0 playing the role of G to obtain the function f_{G_0} and a subset T_1 of I with $T_1 \subset I \subset f_{G_0}(T_1)$ where T_1 and $f_{G_0}(T_1)$ satisfy the additional properties described in Proposition 4.3. Now, let $G_1 = f_{G_0}(T_1) \setminus T_1$ and $I_1 = I \cap G_1 = I \setminus T_1$. Apply Proposition 4.3 again, with G_1 playing the role of G and I_1 playing the role of I to obtain the function f_{G_1} and a subset T_2 of I_1 with $T_2 \subset I_1 \subset f_{G_1}(T_2)$. We continue like this until we arrive at a graph $G_{m(I)}$ with at most kn^{α} edges.

Let $g(I) = T_1 \cup T_2 \cup \cdots \cup T_{m(I)}$, where elements of T_i are colored with color i. Let $\mathcal{S}(s) = \{g(I) : |g(I)| = s\}$. Let $h(g(I)) = G_{m(I)}$. As in the proof of Theorem 4.6, h is well defined. Furthermore, as $g(I) \subset I \subset h(g(I)) \cup g(I)$, conditions (b) is fulfilled. It remains to show (a). We begin by partitioning $\mathcal{S}(s)$ into sets $\mathcal{S}_m(s)$ where $\mathcal{S}_m(s) = \{S \in \mathcal{S}(s) :$ the edges of S are colored with m colors$\}$.

For each $m \in \mathbb{N}$,

$$\mathcal{K}(m) = \{\mathbf{k} = (k_1, \cdots k_m) : k_j \in \mathbb{R}, (1-\varepsilon)^{j-m}k \leq k_j \leq (1-\varepsilon)^j n^{2-\alpha} \text{ and } k_j n^{\alpha} \in \mathbb{N}\}$$

And for each $\mathbf{k} \in \mathcal{K}(m)$,

$$\mathcal{A}(\mathbf{k}) = \{\mathbf{a} = (a_1, \dots a_m) : a_j \in \mathbb{N}, a_j \leq \frac{1}{\varepsilon} \max\{k_j^{1-\lambda^*(\alpha,H)}, n^{-\phi(\alpha,H)}\} n^\alpha \text{ and } \sum_j a_j = s\}$$

So each sequence in $\mathbf{k} \in \mathcal{K}(m)$ corresponds to a potential sequence $(G_1, \cdots G_m)$ where $e(G_i) = k_j n^\alpha$, as the edges of $(1-\varepsilon)^{j-m} k n^\alpha \leq e(G_j) \leq (1-\varepsilon)^j n^2$ and by Proposition 4.3 and Remark 4.4, we have $e(T_j) \leq \frac{1}{\varepsilon} \max\{k_j^{1-\lambda^*(\alpha,H)}, n^{-\phi(\alpha,H)}\} n^\alpha$. Note further that our algorithm returns pairs (G_i, T_i), such that each sequence of $(T_1, \cdots T_m)$ is uniquely identified with a sequence $(G_1, \cdots G_m)$. Thus it suffices to only count the choices of T_i. The sequence $\mathbf{a} \in \mathcal{A}(\mathbf{k})$ corresponds to a sequence of sizes for T_i. Thus, we have that for a fixed m

$$|\mathcal{S}_m(s))| \leq \sum_{\mathbf{k} \in \mathcal{K}(m)} \sum_{\mathbf{a} \in \mathcal{A}(\mathbf{k})} \prod_{j=1}^{m} \binom{k_j n^\alpha}{a_j}.$$

Now, given a $\mathbf{k} \in \mathcal{K}(m)$ and $\mathbf{a} \in \mathcal{A}(\mathbf{k})$, let us partition the product over j according to whether $k_j^{1-\lambda^*(\alpha,H)} < n^{-\phi(\alpha,H)}$ (call this type 1) or not (call this type 2). Because $|\mathcal{K}(m)| = 0$ for values of $m \gg \log(n)$ as $(1-\varepsilon)^m n^2 < k$, and some absolute constant C_1 the product over type 1 j's is at most

$$(n^2)^{\Sigma_j a_j} \leq \exp\left(C_1 \cdot n^{\alpha - \phi(\alpha,H)} (\log n)^2\right) \leq \exp\left(C_1 \cdot k^{1-\lambda^*(\alpha,H)} n^\alpha\right),$$

where in the last step, we used the fact that $k \leq n^{\frac{\phi(H)}{\lambda^*(\alpha,H)-1}} (\log n)^{\frac{2}{1-\lambda^*(\alpha,H)}}$. For each type 2 j, we have $k_j^{1-\lambda^*(\alpha,H)} \geq n^{-\phi(\alpha,H)}$ and thus $a_j \leq \frac{1}{\varepsilon} k_j^{1-\lambda^*(\alpha,H)} n^\alpha$. From this we get

$$k_j \leq (1/\varepsilon)^{\frac{1}{\lambda^*(\alpha,H)-1}} n / a_j^{\frac{1}{\lambda^*(\alpha,H)-1}}.$$

Hence

$$\binom{k_j n^\alpha}{a_j} \leq \left(\frac{e k_j n^\alpha}{a_j}\right)^{a_j} \leq \left(\frac{n^\alpha}{\varepsilon a_j}\right)^{\frac{\lambda^*(\alpha,H)}{\lambda^*(\alpha,H)-1} \cdot a_j}$$

Applying Lemma 4.5 with $M = k^{1-\lambda^*(\alpha,H)} n^\alpha, 1-\delta = (1-\varepsilon)^{\lambda^*(\alpha,H)-1}$, the product over type 2 j's is at most

$$\left(\frac{C_2 n^\alpha}{s}\right)^{\frac{\lambda^*(\alpha,H)}{\lambda^*(\alpha,H)-1} \cdot s} \cdot \exp\left(C_2 k^{1-\lambda^*(\alpha,H)} n^\alpha\right),$$

for some $C_2 = C(H)$.

Thus, letting $C' = 2\max\{C_1, C_2\}$,

$$|\mathcal{S}_m(s)| \leq \sum_{\mathbf{k} \in \mathcal{K}(m)} \sum_{\mathbf{a} \in \mathcal{A}(\mathbf{k})} \left(\frac{C' n^\alpha}{s}\right)^{\frac{\lambda^*(\alpha,H)}{\lambda^*(\alpha,H)-1} \cdot s} \cdot \exp\left(C' k^{1-\lambda^*(\alpha,H)} n^\alpha\right)$$

And thus,

$$|\mathcal{S}(s)| \le \sum_{m=1}^{\infty} \sum_{\mathbf{k}\in\mathcal{K}(m)} \sum_{\mathbf{a}\in\mathcal{A}(\mathbf{k})} \left(\frac{C'n^{\alpha}}{s}\right)^{\frac{\lambda^*(\alpha,H)}{\lambda^*(\alpha,H)-1}\cdot s} \cdot \exp\left(C'k^{1-\lambda^*(\alpha,H)}n^{\alpha}\right)$$

As in the proof of Theorem 4.6, we have $\sum_{m=1}^{\infty}\sum_{\mathbf{k}\in\mathcal{K}(m)}|\mathcal{A}(\mathbf{k})| = n^{O(\log n)}$. The theorem thus follows. □

Theorem 4.9 (Restatement of Theorem 2.4)**.** *Let H be a bipartite graph with h vertices and ℓ edges that contain a cycle. Let A,α be positive reals such that $\mathrm{ex}(n,H)\le An^{\alpha}$ for every $n\in\mathbb{N}$. There exists a constant $C=C(\alpha,H)$ such that*

$$\mathrm{ex}(G(n,p),H) \le \begin{cases} Cn^{2-\frac{1}{m_2(H)}} & \text{if } p \le n^{-\frac{1}{m_2(H)}}, \\ Cp^{1-\frac{1}{\lambda(\alpha,H)}}n^{\alpha} & \text{otherwise} \end{cases}$$

with high probability as $n\to\infty$.

Proof. Since $\mathrm{ex}(G(n,p),H)$ is an increasing function of $e(G)$, it suffices to prove the claimed bound in the case $p \ge n^{-\frac{1}{m_2(H)}}$. Given such a function $p=p(n)$, define $k=p^{-\frac{1}{\lambda(\alpha,H)}}$. Note that $k\le n^{2-\alpha}$. Suppose that there exists an H-free subgraph $I\subset G(n,p)$ with m edges. Then by Theorem 4.6 with our choice of k there exist functions g,h on the set of independent sets, with the property $g(I)\subset I\subset h(g(I))\cup g(I)$. Therefore, we know that $g(I)\subset G(n,p)$ and $G(n,p)$ has at least $m-e(g(I))$ edges of $h(g(I))$. The probability of this event is at most

$$\sum_{S\in\mathcal{S}} \binom{kn^{\alpha}}{m-e(S)}\cdot p^m \le \sum_{s=0}^{C'k^{1-\lambda(\alpha,H)}n^{\alpha}} \left(\frac{C'p^{\frac{\lambda(\alpha,H)-1}{\lambda(\alpha,H)}}n^{\alpha}}{s}\right)^{\frac{\lambda(\alpha,H)}{\lambda(\alpha,H)-1}\cdot s} \cdot \exp\left(C'k^{1-\lambda(\alpha,H)}n^{\alpha}\right)\cdot\left(\frac{3pkn^{\alpha}}{m-s}\right)^{m-s}$$

$$\le \exp\left(O(1)\cdot(p^{\frac{\lambda(\alpha,H)-1}{\lambda(\alpha,H)}}n^{\alpha}+k^{1-\lambda(\alpha,H)}n^{\alpha})\right)\cdot\left(\frac{4pkn^{\alpha}}{m}\right)^{m/2}\to 0,$$

as $n\to\infty$, as long as

$$m \ge C\cdot\max\{p^{\frac{\lambda(\alpha,H)-1}{\lambda(\alpha,H)}}n^{\alpha}, k^{1-\lambda(\alpha,H)}n^{\alpha}, pkn^{\alpha}\} = Cp^{\frac{\lambda(\alpha,H)-1}{\lambda(\alpha,H)}}n^{\alpha},$$

for some sufficiently large constant $C=C(\alpha,H)$, where we used the definition of $k=p^{-\frac{1}{\lambda(\alpha,H)}}$. □

Theorem 4.10 (Restatement of Theorem 2.5)**.** *Let H be a bipartite graph with h vertices and ℓ edges such that H contains a cycle and is Erdős-Simonovits good. Let A,α be positive reals such that $\mathrm{ex}(n,H)\le An^{\alpha}$ for every $n\in\mathbb{N}$. There exists a constant $C=C(\alpha,H)$ such that*

$$\mathrm{ex}(G(n,p),H) \le \begin{cases} Cn^{\alpha-\phi(\alpha,H)}(\log(n))^2 & \text{if } p \le n^{-\frac{\phi(\alpha,H)\lambda^*(\alpha,H)}{\lambda^*(\alpha,H)-1}}\cdot(\log n)^{\frac{2\lambda^*(\alpha,H)}{\lambda^*(\alpha,H)-1}} \\ Cp^{1-\frac{1}{\lambda^*(\alpha,H)}}n^{\alpha} & \text{otherwise} \end{cases}$$

with high probability as $n\to\infty$.

Proof. Since $\mathrm{ex}(G(n,p),H)$ is an increasing function of $e(G)$, it suffices to prove the claimed bound in the

case $p \geq n^{-\frac{\phi(H)\lambda^*(\alpha,H)}{\lambda^*(\alpha,H)-1}} \cdot (\log n)^{\frac{2\lambda^*(\alpha,H)}{\lambda^*(\alpha,H)-1}}$. Let $k = p^{-\frac{1}{\lambda^*(\alpha,H)}}$. Note that by our choice of p,

$$k \leq n^{\frac{\phi(\alpha,H)}{\lambda^*(\alpha,H)-1}} / (\log n)^{\frac{2}{\lambda^*(\alpha,H)-1}}.$$

Suppose that there exists an H-free subgraph $I \subset G(n,p)$ with m edges. Then by Theorem 4.8 with our choice of k there exist functions g,h such that $g(I) \subset G(n,p)$ and $G(n,p)$ has at least $m - e(g(I))$ edges of $h(g(I))$. The probability of this event is at most

$$\sum_{S \in \mathcal{S}} \binom{kn^\alpha}{m-e(S)} \cdot p^m \leq \sum_{s=0}^{C'k^{1-\lambda^*(\alpha,H)}n^\alpha} \left(\frac{C' p^{\frac{\lambda^*(\alpha,H)-1}{\lambda^*(\alpha,H)}} n^\alpha}{s} \right)^{\frac{\lambda^*(\alpha,H)}{\lambda^*(\alpha,H)-1} \cdot s} \cdot \exp\left(C'k^{1-\lambda^*(\alpha,H)}n^\alpha\right) \cdot \left(\frac{3pkn^\alpha}{m-s}\right)^{m-s}$$

$$\leq \exp\left(O(1) \cdot (p^{\frac{\lambda^*(\alpha,H)-1}{\lambda^*(\alpha,H)}} n^\alpha + k^{1-\lambda^*(\alpha,H)} n^\alpha)\right) \cdot \left(\frac{4pkn^\alpha}{m}\right)^{m/2} \to 0,$$

as $n \to \infty$, as long as

$$m \geq C \cdot \max\{p^{\frac{\lambda^*(\alpha,H)-1}{\lambda^*(\alpha,H)}} n^\alpha, k^{1-\lambda^*(\alpha,H)} n^\alpha, pkn^\alpha\} = Cp^{\frac{\lambda^*(\alpha,H)-1}{\lambda^*(\alpha,H)}} n^\alpha$$

for some sufficiently large $C(\alpha,H)$, where we used the definition of $k = p^{-\frac{1}{\lambda*(\alpha,H)}}$. □

As mentioned earlier, Theorem 4.10 implies Morris and Saxton's result on $C_{2\ell}$.

Corollary 4.11 ([45]). *For every $\ell \geq 2$, there exists a constant $C = C(\ell) > 0$ such that*

$$\mathrm{ex}(G(n,p), C_{2\ell}) \leq \begin{cases} Cn^{1+1/(2\ell-1)}(\log(n))^2 & \text{if } p \leq n^{-(\ell-1)/(2\ell-1)} \cdot (\log n)^{2\ell} \\ Cp^{1/\ell} n^{1+1/\ell} & \text{otherwise} \end{cases}$$

with high probability as $n \to \infty$.

Proof. It is known that $C_{2\ell}$ is Erdős-Simonovits good with $\alpha = 1 + 1/\ell$. Apply Theorem 4.10 with $H = \mathcal{C}_{2\ell}$ and $\alpha = 1 + 1/\ell$, noting that $\lambda^*(\alpha, C_{2\ell}) = \frac{\ell}{\ell-1}, \phi(\alpha,H) = \frac{\ell-1}{\ell(2\ell-1)}$. The corollary then follows directly. □

References

[1] N. Alon, M. Krivelevich, B. Sudakov, Turán numbers of bipartite graphs and related Ramsey-type questions, *Combin. Probab. Comput.* **12** (2003), 477-494. 2

[2] N. Alon, L. Rónyai and T. Szabó, Norm-graphs: variations and applications, *J. Combin. Theory Ser. B* **76** (1999), 280-290. 2

[3] J. Balogh, B. Bollobás, M. Simonovits, On the number of graphs without fobidden subgraph, *J. Combin. Th. Ser. B* **91** (2004), 1-24. 2

[4] J. Balogh, B. Bollobás, M. Simonovits, The typical structue of graphs without given excluded subgraphs, *Random Struct. Alg.* **34** (2009), 305-318. 2

[5] J. Balogh, B. Bollobás, M. Simonovits, The fine structure of octahedron-free graphs, *J. Combin. Th. Ser. B* **101** (2011), 67-84. 2

[6] J. Balogh, B. Narayanan, J. Skokan, The number of hypergraphs without linear cycles, *J. Combin. Th. Ser. B* **134** (2019), 309-321.

[7] J. Balogh, R. Morris, W. Samotij, Independent sets in hypergraphs, *J. Amer. Math. Soc.* **28** (2015), 669-709. 2, 3, 15, 16, 18

[8] J. Balogh, W. Samotij, The number of $K_{m,m}$-free graphs, *Combinatorica* **31** (2011), 131-150. 2

[9] J. Balogh, W. Samotij, The number of $K_{s,t}$-free graphs, *J. Lond. Math. Soc.* **83** (2011), 368-388. 2

[10] P. Bennett, T. Bohman A note on the random greedy independent set algorithm, *Random Structures & Algorithms*, **49** (2016), no. 3, 479-502. 5

[11] T. Bohman, P. Keevash, The early evolution of the H-free process, *Invent. Math.* **181** (2010), 291-336. 5

[12] P. Blagojević, B. Bukh, R. Karasev, Turán numbers for $K_{s,t}$-free graphs: topological obstructions and algebraic constructions, *Israel J. Math* **197** (2013), 199-214. 2

[13] B. Bukh, Random algebraic constructions of extremal graphs, *Bull. Lond. Math. Soc.* **47** (2013), 199-214. 2

[14] B. Bukh, D. Conlon, Rational exponents in extremal graph theory, *J. Eur. Math. Soc. (JEMS)* (2018), 1747-1757. 2, 5

[15] D. Conlon, Graphs with few paths of prescribed length between any two vertices, *Bull. Lond. Math. Soc.* **51** (2019), 1015-1021.

[16] D. Conlon, J. Fox, and B. Sudakov, An approximate version of Sidorenko's conjecture, *Geom. Funct. Anal.* **20** (2010), 1354-1366. 6

[17] D. Conlon, T. Gowers, Combinatorial theorems in sparse random sets, *Ann. of Math.* **184** (2016), 367-454. 3

[18] D. Conlon, O. Janzer, Rational exponents near two, *Advances in Combinatorics* (2022), Paper 9, 10pp. 2

[19] D. Conlon, O. Janzer and J. Lee, More on the extremal number of subdivisions, *Combinatorica* **41** (2021), 465-494. 2

[20] P. Erdős, On the combinatorial problems which I would most like to see solved, *Combinatorica* **1** (1981), 25-42. 5

[21] P. Erdős, P. Frankl, V. Rödl, The asymptotic number of graphs not containing a fixed subgraph and a problem for hypergraphs having no exponent, *Graph. Combin.* **2** (1986), 113-121. 2

[22] P. Erdős, D.J. Kleitman, B.L. Rothschild, Asymptotic enumeration of K_n-free graphs, in Colloquio Internazionale sulle Teorie Combinatorie, vol II, Rome, 1973, in Atti dei Convegni Lincei, vol 17, Acad. Naz. Lincei, Rome, 1976, pp.19-27. 2

[23] P. Erdős, Extremal problems in graph theory, in: M. Fiedler (Ed.), Theory of Graphs and Its Applications, 2nd ed., Proc. Symp., Smolenice, 1963, Academic Press, New York, 1965, pp29-36.

[24] P. Erdős, M. Simonovits, A limit theorem in graph theory, *Studia Sci. Math. Hungar.* **1** (1966), 215-235. 2

[25] P. Erdős, M. Simonovits, Supersaturated graphs and hypergraphs, *Combinatorica* **3** (1983), 181-192. 6

[26] P. Erdős, M. Simonovits, Cube-supersaturated graphs and related problems, *Progress in Graph theory (Waterloo, Ont., 1982)*, 203-218, Academic Press, Toronto, 1984.

[27] R. Faudree, M. Simonovits, Cycle-superaturated graphs, unpublished. 6

[28] P. Erdős, A. Stone, On the structure of linear graphs, *Bull. Amer. Math. Soc.* **52** (1946), 1087-1091. 2

[29] A. Ferber, G. McKinley, W. Samotij, Supersaturated sparse graphs and hypergraphs, *Int. Math. Res. Not.* **IMRN 2020** No. 2, 378-402. 1, 4, 5, 6, 14

[30] Z. Füredi, On a Turán type problem of Erdős, *Combinatorica* **11** (1991), 75-79. 2

[31] Z. Füredi and M. Simonovits, The history of the degenerate (bipartite) extremal graph problems, *Erdős centennial, Bolyai Soc. Math. Stud.* **25**, 169-264, János Bolyai Math. Soc., Budapest, 2013. See also arXiv:1306.5167. 2, 4, 5

[32] A. Grzesik, O. Janzer, Z. Nagy, The Turán number of blowups of trees, to appear in *J. Combin. Theory Ser. B*. See also arXiv:1904.07219v1. 2, 6

[33] P. Haxell, Y. Kohayakawa, T. Łuczak, Turán's extremal problem in random graphs: forbidding even cycles, *J. Combin. Th. Ser. B* **64** (1995), 273-287. 3

[34] O. Janzer, The extremal number of the subdivisions of the complete bipartite graph, *SIAM J. Discrete Math.* **34** (2020), 241-250. 2

[35] T. Jiang, Z. Jiang, J. Ma, Negligible obstructions and Turán exponents, *Ann. Appl. Math.* **38** (2022), 1-29. 2

[36] T. Jiang, S. Longbrake, Tree-degenerate graphs and nested dependent random choice, arXiv:2201.10699v1. 2, 6

[37] T. Jiang, J. Ma, and L. Yepremyan, On Turán exponents of bipartite graphs, *Combin. Probab. Comput.* **31** (2022), 333-344. 2

[38] T. Jiang, Y. Qiu, Many Turán exponents via subdivisions, *Combin. Probab. Comput.*, published online, July 2022, https://doi.org/10.1017/S0963548322000177 2

[39] D. Y. Kang, J. Kim, and H. Liu, On the rational Turán exponent conjecture, *J. Combin. Theory Ser. B* **148** (2021), 149-172.

[40] D. Kleitman, K. Winston, On the number of graphs without 4-cycles, *Discrete Math.* **41** (1982), 267-172. 2

[41] Y. Kohayakawa, B. Kreuter, A. Steger, An extremal problem for random graphs and the number of graphs with large even-girth, *Combinatorica* **18** (1998), 101-120. 3

[42] Ph.G. Kolaitis, H.J. Prömel, B.L. Rothschild, K_{t+1}-free graphs: asymptotic structure and a $0-1$ law, *Trans. Amer. Math. Soc.* **303** (1987), 637-671. 2

[43] T. Kővári, V.T. Sós and P. Turán, On a problem of K. Zarankiewicz, *Colloq. Math.* **3** (1954), 50-57. 2

[44] J. Kollár, L. Rónyai, T. Szabó, Norm graphs and bipartite Turán numbers, *Combinatorica* **16** (1996), 399-406. 2

[45] R. Morris, D. Saxton, The number of $C_{2\ell}$-free graphs, *Advances in Math.* **208** (2016), 534-580. 1, 2, 3, 4, 6, 7, 15, 16, 17, 18, 23

[46] D. Mubayi, L. Wang, The number of triple systems without even cycles, *Combinatorica* **39** (2019), 679-704.

[47] D. Mubayi, L. Yepremyan, Random Turán theorem for hypergraph cycles, arXiv:2007.10320v1. 3

[48] B. Nagle,V. Rödl, M. Schacht, Extremal hypergraph problems and the regularity method, in: Topics in Discrete Mathematics, in: Algorithms Combin., vol 26, 2006, pp. 247-278. 2

[49] H.J. Prömel, A. Steger, The asymptotic number of graphs not containing a fixed color-critical subgraph, *Combinatorica* **12** (1992), 463-473. 2

[50] V. Rödl, M. Schacht, Extremal results in random graphs, in Erdős Centennial, in: Bolyai Soc. Math. Stud., vol 25, 2013, pp. 535-583. 3

[51] D. Saxton, A. Thomason, Hypergraphs containers, *Invent. Math.* **201** (2015), 1-68. 2, 3, 15, 16, 18

[52] M. Schacht, Extremal results for random discrete structures, *Ann. of Math.* **184** (2016), 331-363. 3

[53] P. Turán, Eine Extremalauffgabe aus Graphentheorie, Mat. Fiz. Lapok **48** (1941), 436-452. 1

AUTHORS

Tao Jiang
Department of Mathematics
Miami University
Oxford, OH 45056, USA.
jiangt [at] miamioh [dot] edu

Sean Longbrake
Department of Mathematics
Miami University
Oxford, OH 45056, USA.
Currently At:
Department of Mathematics
Emory University
Atlanta, GA 30322, USA.
sean [dot] longbrake [at] emory [dot] edu