A GENERALIZED TURÁN PROBLEM IN RANDOM GRAPHS

WOJCIECH SAMOTIJ AND CLARA SHIKHELMAN

Abstract. We study the following generalization of the Turán problem in sparse random graphs. Given graphs T and H, let $\mathrm{ex}\big(G(n,p),T,H\big)$ be the random variable that counts the largest number of copies of T in a subgraph of $G(n,p)$ that does not contain H. We study the threshold phenomena arising in the evolution of the typical value of this random variable, for every H and an arbitrary 2-balanced T.

Our results in the case when $m_2(H) > m_2(T)$ are a natural generalization of the Erdős–Stone theorem for $G(n,p)$, which was proved several years ago by Conlon and Gowers and by Schacht; the case $T = K_m$ has been recently resolved by Alon, Kostochka, and Shikhelman. More interestingly, the case when $m_2(H) \leqslant m_2(T)$ exhibits a more complex and subtle behavior. Namely, the location(s) of the (possibly multiple) threshold(s) are determined by densities of various coverings of H with copies of T and the typical value(s) of $\mathrm{ex}\big(G(n,p),T,H\big)$ are given by solutions to deterministic hypergraph Turán-type problems that we are unable to solve in full generality.

1. Introduction

The well-known Turán function is defined as follows. For a fixed graph H and an integer n, we let $\mathrm{ex}(n,H)$ be the maximum number of edges in an H-free[1] subgraph of K_n. This function has been studied extensively and generalizations of it were offered in different settings (see [34] for a survey). Erdős and Stone [11] determined $\mathrm{ex}(n,H)$ for any nonbipartite graph H up to lower order terms.

Theorem 1.1 ([11]). *For every fixed nonempty graph H,*

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

Note that if H is bipartite, then the theorem only tells us that $\mathrm{ex}(n,H) = o(n^2)$. In fact, the classical result of Kővári, Sós, and Turán [26] implies that in this case $\mathrm{ex}(n,H) = O(n^{2-c})$ for some $c > 0$ that depends only on H.

Two natural generalizations of Theorem 1.1 have been considered in the literature. First, instead of maximizing the number of edges in an H-free subgraph of the complete graph with n vertices, one can consider only H-free subgraphs of some other n-vertex graph G. One natural choice is to let G be the random graph $G(n,p)$, that is, the random graph on n vertices whose each pair of vertices forms an edge independently with probability p. This leads to the study of the random variable $\mathrm{ex}\big(G(n,p),H\big)$, the maximum number of edges in an H-free subgraph of $G(n,p)$. Considering the intersection between the largest H-free subgraph of K_n and the random graph $G(n,p)$, one can show that if $p \gg \mathrm{ex}(n,H)^{-1}$, then w.h.p.[2]

$$\mathrm{ex}\big(G(n,p),H\big) \geqslant (1+o(1)) \cdot \mathrm{ex}(n,H)p. \tag{1}$$

Research supported in part by the Israel Science Foundation grant 1147/14 (WS) and grants from the Israel Science Foundation and the German-Israeli Foundation for Scientific Research and Development (CS).

[1] A graph is H-free if it does not contain H as a (not necessarily induced) subgraph.

[2] We write w.h.p. as an abbreviation of with high probability, that is, with probability tending to one as the number of vertices n tends to infinity.

The above bound is not always best-possible. If p decays sufficiently fast so that the expected number of copies of (some subgraph of) H that contain a given edge of $G(n,p)$ is $o(1)$, then (1) can be strengthened to $\mathrm{ex}\big(G(n,p),H\big) \geqslant (1+o(1)) \cdot \binom{n}{2}p$. Indeed, one can remove all copies of H from $G(n,p)$ by arbitrarily removing an edge from each copy of (some subgraph H' of) H and the assumption on p implies that w.h.p. only a tiny proportion of the edges will be removed this way. Such considerations naturally lead to the notion of *2-density* of H, denoted by $m_2(H)$, which is defined by

$$m_2(H) = \max\left\{\frac{e_H - 1}{v_H - 2} : H' \subseteq H,\, e_{H'} \geqslant 2\right\}.$$

Moreover, we say that H is 2-*balanced* if H itself is one of the graphs achieving the maximum above, that is, if $m_2(H) = (e_H - 1)/(v_H - 2)$. It is straightforward to verify that the expected number of copies of (some subgraph H' of) H that contain a given edge of $G(n,p)$ tends to zero precisely when $p \ll n^{-1/m_2(H)}$.

Haxell, Kohayakawa, Łuczak, and Rödl [18, 24] conjectured that if the opposite inequality $p \gg n^{-1/m_2(H)}$ holds, then the converse of (1) must (essentially) be true. (The case when H is bipartite is much more subtle; see, e.g., [23, 28].) This conjecture was proved by Conlon and Gowers [7], under the additional assumption that H is 2-balanced, and, independently, by Schacht [33]; see also [5, 8, 12, 31, 32].

Theorem 1.2 ([7, 33])**.** *For any fixed graph H with at least two edges, the following holds w.h.p.*

$$\mathrm{ex}\big(G(n,p),H\big) = \begin{cases} \left(1 - \frac{1}{\chi(H)-1} + o(1)\right)\binom{n}{2}p & \text{if } p \gg n^{-1/m_2(H)}, \\ (1+o(1)) \cdot \binom{n}{2}p & \text{if } n^{-2} \ll p \ll n^{-1/m_2(H)}. \end{cases}$$

The second generalization of the Turán problem is to fix two graphs T and H and ask to determine the maximum number of copies of T in an H-free subgraph of K_n. Denote this function by $\mathrm{ex}(n,T,H)$ and note that $\mathrm{ex}(n,H) = \mathrm{ex}(n,K_2,H)$, so this is indeed a generalization. Erdős [9] resolved this question in the case when both T and H are complete graphs, proving that the balanced complete $(\chi(H)-1)$-partite graph has the most copies of T. Another notable result was recently obtained by Hatami, Hladký, Kráľ, Norine, and Razborov [17] and, independently, by Grzesik [14], who determined $\mathrm{ex}(n,C_5,K_3)$, resolving an old conjecture of Erdős. The systematic study of the function $\mathrm{ex}(n,T,H)$ for general T and H, however, was initiated only recently by Alon and Shikhelman [3].

Determining the function $\mathrm{ex}(n,T,H)$ asymptotically for arbitrary T and H seems to be a very difficult task and a generalization of Theorem 1.1 to this broader context has yet to be discovered. On the positive side, a nowadays standard argument can be used to derive the following generalization of the Erdős–Stone theorem to the case when T is a complete graph from the aforementioned result of Erdős.

Theorem 1.3. *For any fixed nonempty graph H and any integer $m \geqslant 2$,*

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

Analogously to Theorem 1.1, in the case $\chi(H) \leqslant m$, the above theorem only tells us that $\mathrm{ex}(n,K_m,H) = o(n^m)$. The following simple proposition generalizes this fact. A *blow-up* of a graph T is any graph obtained from T by replacing each of its vertices with an independent set and each of its edges with a complete bipartite graph between the respective independent sets.

Proposition 1.4 ([3])**.** *Let T be a fixed graph with t vertices. Then $\mathrm{ex}(n,T,H) = \Omega(n^t)$ if and only if H is not a subgraph of a blow-up of T. Otherwise, $\mathrm{ex}(n,T,H) \leqslant n^{t-c}$ for some $c > 0$ that depends only on T and H.*

We remark that both the problems of (i) determining the limit of $\mathrm{ex}(n,T,H)\cdot n^{-t}$ for general T and H such that H is not contained in a blow-up of T and (ii) computing $\mathrm{ex}(n,T,H)$ up to a constant factor for arbitrary T and H such that H is contained in a blow-up of T seem extremely difficult. Even the case $T=K_2$ of (ii) alone, that is, determining the order of magnitude of the Turán function $\mathrm{ex}(n,H)$ for an arbitrary bipartite graph H is a notorious open problem, see [13].

The common generalization of Theorems 1.2 and 1.3 was considered in [2]. Let $\mathrm{ex}\big(G(n,p),T,H\big)$ be the random variable that counts the maximum number of copies of T in an H-free subgraph of $G(n,p)$. Generalizing the easy argument that yields (1), one can show that the inequality

$$\mathrm{ex}\big(G(n,p),T,H\big) \geqslant \big(\mathrm{ex}(n,T,H)+o\big(n^{v_T}\big)\big)\,p^{e_T}$$

holds (w.h.p.) whenever $p \gg n^{-v_{T'}/e_{T'}}$ for every nonempty $T' \subseteq T$; it is well-known that if $p=O(n^{-v_{T'}/e_{T'}})$ for some $T'\subseteq T$, then $G(n,p)$ contains no copies of T with probability $\Omega(1)$. It seems natural to guess that the opposite inequality holds as soon as $p \gg n^{-1/m_2(H)}$. The case $T=K_m$ was studied in [2], where the following generalization of Theorem 1.2 was proved.

Theorem 1.5 ([2])**.** *Let $m\geqslant 2$ be an integer and let H be a fixed graph with $m_2(H)>m_2(K_m)$ and $\chi(H)>m$. If p is such that $\binom{n}{m}p^{\binom{m}{2}}$ tends to infinity with n, then w.h.p.*

$$\mathrm{ex}\big(G(n,p),K_m,H\big) = \begin{cases} (1+o(1))\cdot\binom{\chi(H)-1}{m}\left(\frac{n}{\chi(H)-1}\right)^m p^{\binom{m}{2}} & \text{if } p\gg n^{-1/m_2(H)},\\ (1+o(1))\cdot\binom{n}{m}p^{\binom{m}{2}} & \text{if } p\ll n^{-1/m_2(H)}.\end{cases}$$

Let us draw the reader's attention to the assumption that $m_2(H)>m_2(K_m)$ in the statement of the theorem. No such assumption was (explicitly) present in the statement of Theorem 1.2 and it is natural to wonder whether it is really necessary. Since we assume that H is not m-colorable, then it must contain a subgraph whose average degree is at least m, larger than the average degree of K_m. In particular, it is natural to guess that this implies that the 2-density of H is larger than the 2-density of K_m. Perhaps surprisingly, this is not true and only the weaker inequality $m_2(H)>m_2(K_{m-1})$ does hold for every non-m-colorable graph H. A construction of a graph H such that $\chi(H)=4$ and $m_2(H)<m_2(K_3)$ was given in [1]. Subsequently, constructions of graphs H such that $\chi(H)=m+1$ but $m_2(H)<m_2(K_m)$ were given for all m in [2]. It was also shown there that for such graphs H, the typical value of $\mathrm{ex}\big(G(n,p),K_m,H\big)$ does not change at $p=n^{-1/m_2(H)}$, as in Theorem 1.5. More precisely, if $p=n^{-1/m_2(H)+\delta}$ for some small but fixed $\delta=\delta(H)>0$, then still $\mathrm{ex}\big(G(n,p),K_m,H\big)=(1+o(1))\cdot\binom{n}{m}p^{\binom{m}{2}}$. This led to the following open questions:

(i) Where does the 'phase transition' of $\mathrm{ex}\big(G(n,p),K_m,H\big)$ take place if $m_2(H)\leqslant m_2(K_m)$?
(ii) How does the function $p\mapsto \mathrm{ex}\big(G(n,p),T,H\big)$ grow for general T and H?

In this paper we answer both of these questions under the assumptions that T is 2-balanced and H is not contained in a blow-up of T. Answering question (ii) in the case when H is contained in a blow-up of T seems extremely challenging, as even the order of magnitude of $\mathrm{ex}(n,T,H)$, which corresponds to setting $p=1$ above, is not known for general graphs T and H, see the comment below Proposition 1.4.

The case when $m_2(H)>m_2(T)$ holds no surprises, as the following extension of Theorem 1.5 is valid. We denote by $\mathcal{N}_T(K_n)$ the number of copies of a graph T in the complete graph K_n.

Theorem 1.6. *If H and T are fixed graphs such that T is 2-balanced and that $m_2(H)>m_2(T)$, then w.h.p.*

$$\mathrm{ex}\big(G(n,p),T,H\big) = \begin{cases} \big(\mathcal{N}_T(K_n)+o\big(n^{v_T}\big)\big)\,p^{e_T} & \text{if } n^{-v_T/e_T}\ll p\ll n^{-1/m_2(H)},\\ \big(\mathrm{ex}\big(n,T,H\big)+o\big(n^{v_T}\big)\big)\,p^{e_T} & \text{if } p\gg n^{-1/m_2(H)}.\end{cases}$$

As already hinted at by [2], the case when $m_2(H) \leqslant m_2(T)$ exhibits a more complex behavior. We find that there are several potential 'phase transitions' and we relate their locations to a measure of density of various coverings of H with copies of T that generalizes the notion of the 2-density of H. Moreover, we show that the (typical) asymptotic value of $\mathrm{ex}\big(G(n,p),T,H\big)$ is determined, for every p that does not belong to any of the constantly many 'phase transition windows', by a solution of a deterministic hypergraph Turán-type problem. Unfortunately, we were unable to solve this Turán-type problem in full generality. Worse still, we do not understand it sufficiently well to either show that for some pairs of T and H, the function $p \mapsto \mathrm{ex}\big(G(n,p),T,H\big)$ undergoes more than one 'phase transition' or to rule out the existence of such pairs. We leave these questions as a challenge for future work.

In order to make the above discussion formal and state the main theorem, we will require several definitions.

1.1. **Notations and definitions.** A *T-covering* of H is a minimal collection $F = \{T_1, \dots, T_k\}$ of pairwise edge-disjoint copies of T (in a large complete graph) whose union contains a copy of H.[3] Given two T-coverings $F = \{T_1, \dots, T_k\}$ and $F' = \{T_1', \dots, T_k'\}$, a map f from the union of the vertex sets of the T_is to the union of the vertex sets of the T_i's is an *isomorphism* if it is a bijection and for every $T_i \in F$, the graph $f(T_i)$ belongs to F'. We can then say that the *type* of a T-covering of H is just the isomorphism class of this covering. Observe that there are only finitely many types of T-coverings of H. One special type of a T-covering of H that will be important in our considerations is the covering of H with e_H copies of T such that each copy of T intersects H in a single edge and is otherwise completely (vertex) disjoint from the remaining $e_H - 1$ copies of T that constitute this covering. We denote this covering by $F^e_{T,H}$ and note that the union of all members of $F^e_{T,H}$ is a graph with $v_H + e_H(v_T - 2)$ vertices and $e_H e_T$ edges.

For a collection F' of copies of T, denote by $U(F')$ the *underlying graph* of F', that is, the union of all members of F'. We define the *T-density* of a T-covering F, which we shall denote by $m_T(F)$, as follows:

$$m_T(F) = \max\left\{\frac{e_{U(F')} - e_T}{v_{U(F')} - v_T} : F' \subseteq F, |F'| \geqslant 2\right\}.$$

Note that this generalizes the notion of 2-density of a graph. Indeed, the 2-density of H is the K_2-density of (the edge set of) H. The notion of T-density is motivated by the following observation. For graphs G and T, we let $T(G)$ denote the collection of copies of T in G and let $\mathcal{N}_T(G) = |T(G)|$.

Remark 1.7. For every collection F of at least two copies of T,

$$\mathbb{E}\left[\mathcal{N}_{U(F')}\big(G(n,p)\big)\right] \ll \mathbb{E}\left[\mathcal{N}_T\big(G(n,p)\big)\right] \text{ for some } F' \subseteq F \iff p \ll n^{-1/m_T(F)}.$$

Even though we are interested in maximizing $\mathcal{N}_T(G)$ in an H-free subgraph $G \subseteq K_n$, we shall be considering (more general) abstract collections of T-copies in K_n that do not contain a T-covering of H of a certain type (or a set of types). In particular, if $G \subseteq K_n$ is H-free, then $T(G)$ is one such collection of T-copies, as it does not contain any T-covering of H (since the underlying graph of every T-covering of H contains H as a subgraph). However, not all the collections we shall consider will be 'graphic', that is, of the form $T(G)$ for some graph G.

The aforementioned Turán-type problem for hypergraphs asks to determine the following quantity. For a given family $\mathcal{F}$ of T-coverings of H, we let $\mathrm{ex}^*(n,T,\mathcal{F})$ be the maximum size of a collection of copies of T in K_n that does not contain any member of $\mathcal{F}$. Note that

[3]The collection $F = \{T_1, \dots, T_k\}$ is minimal in the sense that for every $i \in [k]$, the union of all graphs in $F \setminus \{T_i\}$ no longer contains a copy of H.

for any collection $\mathcal{F}$ of T-coverings of H, one has that $\mathrm{ex}^*(n,T,\mathcal{F}) \geqslant \mathrm{ex}(n,T,H)$. Indeed, if G is an H-free graph with n vertices such that $\mathrm{ex}(n,T,H) = \mathcal{N}_T(G)$, then $T(G)$ is $\mathcal{F}$-free. However, this inequality can be strict, as not every collection of T-copies is of the form $T(G)$ for some graph G. Having said that, we shall show in Lemma 3.5 that at least $\mathrm{ex}^*(n,T,F^e_{T,H}) = \mathrm{ex}(n,T,H) + o(n^{v_T})$. We are now equipped to formulate the key definition needed to state our main result.

Definition 1.8. Suppose that T and H are fixed graphs and assume that T is 2-balanced. The T-*resolution* of H is the sequence $F_1,\dots,F_k$ of all types of T-coverings of H whose T-density does not exceed $m_T(F^e_{T,H})$, ordered by their T-density (with ties broken arbitrarily). The *associated threshold sequence* is the sequence $p_0,p_1,\dots,p_k$, where $p_0 = n^{-v_T/e_T}$ and $p_i = n^{-1/m_T(F_i)}$ for $i \in [k]$.

1.2. **Statement of the main theorem.** The following theorem is the main result of this paper.

Theorem 1.9. *Suppose that H and T are fixed graphs and assume that T is 2-balanced and that $m_2(H) \leqslant m_2(T)$. Let $F_1,\dots,F_k$ be the T-resolution of H and let $p_0,p_1,\dots,p_k$ be the associated threshold sequence. Then the following hold for every $i \in [k]$:*

(i) If $p_0 \ll p \ll p_i$, then w.h.p.

$$\mathrm{ex}\big(G(n,p),T,H\big) \geqslant \big(\mathrm{ex}^*\big(n,T,\{F_1,\dots,F_{i-1}\}\big) + o\big(n^{v_T}\big)\big)\, p^{e_T}.$$

(ii) If $p \gg p_i$, then w.h.p.

$$\mathrm{ex}\big(G(n,p),T,H\big) \leqslant \big(\mathrm{ex}^*\big(n,T,\{F_1,\dots,F_{i}\}\big) + o\big(n^{v_T}\big)\big)\, p^{e_T}.$$

Even though the above theorem determines the typical values of $\mathrm{ex}\big(G(n,p),T,H\big)$ for almost all p, these values remain somewhat of a mystery as we do not know how to compute $\mathrm{ex}^*\big(n,T,\{F_1,\dots,F_i\}\big)$ in general. One thing that we do know how to prove is that $\mathrm{ex}^*(n,T,\mathcal{F}) = \mathrm{ex}(n,T,H) + o(n^{v_T})$ for every family $\mathcal{F}$ of T-coverings of G that contains the special covering $F^e_{T,H}$, see Lemma 3.5. Moreover, it is not hard to verify that

$$m_T(F^e_{T,H}) = \frac{e_T}{v_T - 2 + 1/m_2(H)},$$

which, when T is 2-balanced, is equal to the so-called *asymmetric* 2*-density of T and H*, a quantity that arises in the study of asymmetric Ramsey properties of $G(n,p)$, see [15, 22, 25, 29]. Note that if T is 2-balanced and $m_2(H) < m_2(T)$, then $m_2(H) < m_T(F^e_{T,H}) < m_2(T)$. An 'abbreviated' version of Theorem 1.9 can be now stated as follows.

Corollary 1.10. *Suppose that H and T are fixed graphs and assume that T is 2-balanced and that $m_2(H) \leqslant m_2(T)$. There is an integer $t \geqslant 1$ and rational numbers $\mu_0 < \dots < \mu_t$, where*

$$\mu_0 = \frac{e_T}{v_T} \qquad \text{and} \qquad \mu_k \leqslant \frac{e_T}{v_T - 2 + 1/m_2(H)},$$

and real numbers $\pi_0 > \dots > \pi_t$, where

$$\pi_0 = \frac{1}{|\mathrm{Aut}(T)|} \qquad \text{and} \qquad \pi_t = \lim_{n\to\infty} \mathrm{ex}(n,T,H)\cdot n^{-v_T},$$

such that w.h.p.

$$\mathrm{ex}\big(G(n,p),T,H\big) = \begin{cases} (\pi_i + o(1)) n^{v_T} p^{e_T}, & n^{-1/\mu_i} \ll p \ll n^{-1/\mu_{i+1}} \quad \text{for } i \in \{0,\dots,t-1\}, \\ (\pi_t + o(1)) n^{v_T} p^{e_T}, & p \gg n^{-1/\mu_t}. \end{cases}$$

A rather disappointing feature of Corollary 1.10 (and thus of Theorem 1.9) is that we are unable to determine whether or not there exists a pair of graphs H and T for which the typical value of $\mathrm{ex}\big(G(n,p),T,H\big)$ undergoes more than one 'phase transition' (that is, the integer t from the statement of the corollary is strictly greater than one). If one was allowed to replace H with a finite family of forbidden graphs, then one can see an arbitrary (finite) number of 'phase transitions' even in the case when $T = K_2$, see [27, Theorem 6.4].

Even though we were able to construct pairs of H and T which admit T-coverings of H whose T-density is strictly smaller than the T-density of the special covering of H with e_H copies of T, for no such T-covering F we were able to show that $\mathrm{ex}^*(n,T,F) \geqslant \mathrm{ex}(n,T,H) + \Omega(n^{v_T})$. On the other hand, if one removes the various (important) assumptions on the densities of H, T, and F, then one can find such triples. A simple example is $H = K_7$, $T = K_3$, and F being a decomposition of K_7 into edge-disjoint triangles (the Fano plane). Indeed, in this case $\mathrm{ex}^*(n,K_3,F) \geqslant (\frac{3}{4} - o(1))\binom{n}{3}$ as witnessed by the family of all triangles in K_n that have at least one vertex in each of the parts of some partition of $V(K_n)$ into two sets of (almost) equal size (the Fano plane is not 2-colorable). On the other hand, Theorem 1.3 implies that $\mathrm{ex}(n,K_3,K_7) \leqslant (\frac{5}{9} + o(1))\binom{n}{3}$. We thus pose the following question.

Question 1.11. *Do there exist pairs of graphs H and T such that $m_2(H) \leqslant m_2(T)$, T is 2-balanced, and the family $\mathcal{F}$ of all T-coverings of H that have the smallest T-density (among all T-coverings of H) satisfies $\mathrm{ex}^*(n,T,\mathcal{F}) \geqslant \mathrm{ex}(n,T,H) + \Omega(n^{v_T})$?*

Let us point out that answering Question 1.11 is equivalent to determining whether or not there is a pair of graphs H and T, where T is 2-balanced, for which $\mathrm{ex}\big(G(n,p),T,H\big)$ undergoes multiple 'phase transitions' in the sense described above. Indeed, suppose that H and T are fixed graphs and assume that T is 2-balanced and that $m_2(H) \leqslant m_2(T)$. Let $F_1,\dots,F_k$ be the T-resolution of H and let $p_0,p_1,\dots,p_k$ be the associated threshold sequence. The numbers $\pi_0,\pi_1,\dots,\pi_t$ from the statement of Corollary 1.10 are precisely all numbers π satisfying

$$\pi = \lim_{n\to\infty} \mathrm{ex}^*(n,T,\{F_1,\dots,F_i\}) \cdot n^{-v_T}$$

for some $i \in \{0,\dots,k\}$ such that either $p_{i+1} \neq p_i$ or $i = k$. Standard averaging arguments can be used to show that $\mathrm{ex}^*(n,T,\mathcal{F}) \leqslant \mathcal{N}_T(K_n) - \Omega(n^{v_T})$ for every nonempty family $\mathcal{F}$ of T-coverings of H whereas the aforementioned Lemma 3.5 yields

$$\mathrm{ex}(n,T,H) \leqslant \mathrm{ex}^*(n,T,\{F_1,\dots,F_k\}) \leqslant \mathrm{ex}^*(n,T,F^e_{T,H}) \leqslant \mathrm{ex}(n,T,H) + o(n^2).$$

Thus $t > 1$ is and only if there exists some $i \in \{1,\dots,k-1\}$ such that

$$p_{i+1} \gg p_i \qquad \text{and} \qquad \mathrm{ex}^*(n,T,\{F_1,\dots,F_i\}) \geqslant \mathrm{ex}(n,T,H) + \Omega(n^{v_T}).$$

If the latter condition is satisfied, then it also holds when i is the largest index such that $p_1 = p_i$. But then $\{F_1,\dots,F_i\}$ is precisely the family $\mathcal{F}$ defined in Question 1.11.

1.3. **Structure of the paper.** The rest of the paper is structured as follows. In Section 2, we give a high level overview of the proofs of Theorems 1.6 and 1.9. In Section 3, we introduce the main tools, the hypergraph container lemma and Harris's and Janson's inequalities, and prove a few useful lemmas and corollaries concerning extremal and random graphs. In Section 4, we give the proofs of the main theorems, starting with the simpler Theorem 1.6 and then continuing to the more difficult Theorem 1.9. Finally, in Section 5, we give concluding remarks and offer open problems.

2. Proof outline

Before diving into the details of the proofs of Theorems 1.6 and 1.9, let us briefly go over the main steps we take, highlighting the main ideas.

The proofs of the lower bounds on $\mathrm{ex}\big(G(n,p),T,H\big)$ are rather standard. Suppose that $G \sim G(n,p)$. In the setting of Theorem 1.6, the H-free subgraph of G with a large number of copies of T is obtained by arbitrarily removing from G one edge from every copy of (some subgraph of) H. In the setting of Theorem 1.9, we remove from G all edges that are either (i) not contained in a copy of T that belongs to a fixed extremal $\{F_1,\dots,F_{i-1}\}$-free collection $\mathcal{T} \subseteq T(K_n)$, or (ii) contained in a copy of T that constitutes some T-covering of H in $T(G)\cap\mathcal{T}$, or (iii) contained in more than one copy of T in G. Note that all copies of H in G are removed this way. Our assumption on p guarantees that in steps (ii) and (iii) above we lose only a negligible proportion of $T(G)$.

The upper bound implicit in the equality $\mathrm{ex}\big(G(n,p),T,H\big) = \big(\mathcal{N}_T(K_n)+o\big(n^{v_T}\big)\big)\,p^{e_T}$ in Theorem 1.6 follows from a standard application of the second moment method; see, e.g., [20]. The proofs of the remaining upper bounds, in both Theorems 1.6 and 1.9, utilize the method of hypergraph containers [5, 32]; see also [6]. Roughly speaking, the hypergraph container theorems state that the family of all independent sets of a uniform hypergraph whose edges are distributed somewhat evenly can be covered by a relatively small family of subsets, called *containers*, each of which is 'almost independent' in the sense that it contains only a negligible proportion of the edges of the hypergraph.

In the setting of Theorem 1.6, a standard application of the method yields a collection $\mathcal{C}$ of $\exp\big(O(n^{2-1/m_2(H)}\log n)\big)$ subgraphs of K_n (the containers), each with merely $o(n^{v_H})$ copies of H, that cover the family of all H-free subgraphs of K_n. Suppose that $G\sim G(n,p)$ and let G_0 be an H-free subgraph of G and note that G_0 has to be a subgraph of one of the containers. A standard supersaturation result states that each graph in $\mathcal{C}$ can have at most $\mathrm{ex}(n,T,H)+o(n^{v_H})$ copies of H. It follows that for each *fixed* container $C\in\mathcal{C}$, the intersection of G with C can have no more than $\big(\mathrm{ex}(n,T,H)+o(n^{v_T})\big)p^{e_T}$ copies of T. At this point, one would normally take the union bound over all containers and conclude that w.h.p. the number of copies of T in $G\cap C$ is small simultaneously for all $C\in\mathcal{C}$ and hence also G_0 has this property, as $G_0\subseteq G\cap C$ for some $C\in\mathcal{C}$.

Unfortunately, we cannot afford to take such a union bound as the rate of the upper tail of the number of copies of T in $G(n,p)$ is much too slow to allow this, see [21]. Luckily, the rate of the *lower tail* of the number of copies of T in $G(n,p)$ is sufficiently fast, see Lemma 3.9, to allow a union bound over all containers. Therefore, what we do is first prove that w.h.p. $\mathcal{N}_T(G)=(1+o(1))\mathcal{N}_T(K_n)p^{e_T}$ and then show that w.h.p. the number of copies of T in G that are not fully contained in C is at least $\big(\mathcal{N}_T(K_n)-\mathrm{ex}(n,T,H)-o(n^{v_T})\big)p^{e_T}$ simultaneously for all $C\in\mathcal{C}$. This implies that w.h.p. $\mathcal{N}_T(G_0)\leqslant \max_{C\in\mathcal{C}}\mathcal{N}_T(G\cap C)\leqslant\big(\mathrm{ex}(n,T,H)+o(n^{v_T})\big)p^{e_T}$.

In the setting of Theorem 1.9, instead of building containers for all possible graphs G_0, we build containers for all possible collections $T(G_0)$, exploiting the fact that $T(G_0)$ cannot contain any T-covering of H, as G_0 is H-free. More precisely, we work with hypergraphs $\mathcal{H}_1,\dots,\mathcal{H}_i$, each with vertex set $T(K_n)$, whose edges are copies of the T-coverings $F_1,\dots,F_i$, respectively. A version of the container theorem presented in Corollary 3.2 provides us with a small collection $\mathcal{C}$ of subsets of $T(K_n)$ such that (i) each $\{F_1,\dots,F_i\}$-free collection $\mathcal{T}\subseteq T(K_n)$ is contained in some member of $\mathcal{C}$ and (ii) each $C\in\mathcal{C}$ has only $o(n^{v_{U(F_j)}})$ copies of F_j, for each $j\in[i]$, and thus (by a standard averaging argument) it comprises at most $\mathrm{ex}^*(n,T,\{F_1,\dots,F_i\})+o(n^{v_T})$ copies of T. The key parameter q from the statement of Corollary 3.2, which determines the size of $\mathcal{C}$, exactly matches our definitions of $m_T(F_1),\dots,m_T(F_i)$. Now, since the underlying graph of each F_j contains H as a subgraph and G_0 is H-free, $T(G_0)$ must be contained in some member of $\mathcal{C}$. The rest of the argument is similar to the proof of Theorem 1.6 – we first bound the upper tail of $\mathcal{N}_T(G)$ and then the lower tail of $|T(G) \setminus C|$ for all $C \in \mathcal{C}$ simultaneously.

3. Tools and Preliminary Results

3.1. **Hypergraph container lemma.** The first key ingredient in our proof is the following version of the hypergraph container lemma, proved by Balogh, Morris, and Samotij [5]. An essentially equivalent statement was obtained independently by Saxton and Thomason [32]. We first introduce the relevant notions. Suppose that $\mathcal{H}$ is a k-uniform hypergraph. For a set $B \subseteq V(\mathcal{H})$, we let $\deg_{\mathcal{H}}(B) = \big|\{A \in E(\mathcal{H}) : B \subseteq A\}\big|$ and for each $\ell \in [k]$, we let

$$\Delta_\ell(\mathcal{H}) = \max\big\{ \deg_{\mathcal{H}}(B) : B \subseteq V(\mathcal{H}),\, |B| = \ell\big\}.$$

We denote by $\mathcal{I}(\mathcal{H})$ the collection of independent sets in $\mathcal{H}$.

Theorem 3.1 ([5])**.** *For every positive integer k and all positive K and ε, there exists a positive constant C such that the following holds. Let $\mathcal{H}$ be a k-uniform hypergraph and assume that $q \in (0,1)$ satisfies*

$$\Delta_\ell(\mathcal{H}) \leqslant K q^{\ell-1} \frac{e(\mathcal{H})}{v(\mathcal{H})} \qquad \text{for all } \ell \in [k]. \tag{2}$$

There exist a family $\mathcal{S} \subseteq \binom{V(\mathcal{H})}{\leqslant Cqv(\mathcal{H})}$ and functions $f \colon \mathcal{S} \to \mathcal{P}(V(\mathcal{H}))$ and $g \colon \mathcal{I}(\mathcal{H}) \to \mathcal{S}$ such that:

(i) For every $I \in \mathcal{I}(\mathcal{H})$, $g(I) \subseteq I$ and $I \setminus g(I) \subseteq f(g(I))$.

(ii) For every $S \in \mathcal{S}$, $e(\mathcal{H}[f(S)]) \leqslant \varepsilon e(\mathcal{H})$.

(iii) If $g(I) \subseteq I'$ and $g(I') \subseteq I$ for some $I, I' \in \mathcal{I}(\mathcal{H})$, then $g(I) = g(I')$.

Let us make two remarks here. First, condition (ii) in the above statement is equivalent to the condition that the image of the function f from the statement of [5, Theorem 2.2] is $\overline{\mathcal{F}}$, where $\mathcal{F}$ is the (increasing) family of all subsets of $V(\mathcal{H})$ that induce more than $\varepsilon e(\mathcal{H})$ edges. Second, that the final assertion of the statement of Theorem 3.1 is not present in the original statement of [5, Theorem 2.2]. It is, however, proved in the final claim of the proof of [5, Theorem 2.2].

Since the hypergraphs we shall be working with in the proof of Theorem 1.9 are not necessarily uniform, we shall be actually invoking the following (rather straightforward) corollary of Theorem 3.1.

Corollary 3.2. *For all positive integers $k_1, \dots, k_i$ and all positive K and ε, there exists a positive constant C such that the following holds. Suppose that $\mathcal{H}_1, \dots, \mathcal{H}_i$ are hypergraphs with the same vertex set V and that $\mathcal{H}_j$ is k_j-uniform, for each $j \in [i]$. Assume that $q \in (0,1)$ is such that for all $j \in [i]$,*

$$\Delta_\ell(\mathcal{H}_j) \leqslant K q^{\ell-1} \frac{e(\mathcal{H}_j)}{v(\mathcal{H}_j)} \qquad \text{for all } \ell \in [k_j]. \tag{3}$$

There exist a family $\mathcal{S} \subseteq \binom{V}{\leqslant Cq|V|}$ and functions $f \colon \mathcal{S} \to \mathcal{P}(V)$ and $g \colon \bigcap_{j=1}^i \mathcal{I}(\mathcal{H}_j) \to \mathcal{S}$ such that:

(i) For every $I \in \bigcap_{j=1}^i \mathcal{I}(\mathcal{H}_j)$, $g(I) \subseteq I$ and $I \setminus g(I) \subseteq f(g(I))$.

(ii) For every $S \in \mathcal{S}$, $e(\mathcal{H}_j[f(S)]) \leqslant \varepsilon e(\mathcal{H}_j)$ for every $j \in [i]$.

(iii) If $g(I) \subseteq I'$ and $g(I') \subseteq I$ for some $I, I' \in \bigcap_{j=1}^i \mathcal{I}(\mathcal{H}_j)$, then $g(I) = g(I')$.

Proof. For each $j \in [i]$, let C_j be the constant given by Theorem 3.1 with $k \leftarrow k_j$. Assume that $q \in (0,1)$ is such that the hypergraphs $\mathcal{H}_1, \dots, \mathcal{H}_i$ satisfy (3). For each $j \in [i]$, we may apply Theorem 3.1 to the hypergraph $\mathcal{H}_j$ to obtain a family $\mathcal{S}_j \subseteq \binom{V}{\leqslant C_j|V|}$ and functions $f_j \colon \mathcal{S}_j \to \mathcal{P}(V)$ and $g_j \colon \mathcal{I}(\mathcal{H}_j) \to \mathcal{S}_j$ as in the assertion of the theorem.

We now let $C = C_1 + \ldots + C_i$ and define

$$\mathcal{S} = \big\{ S_1 \cup \ldots \cup S_i : S_j \in \mathcal{S}_j \text{ for each } j \in [i] \big\} \subseteq \binom{V}{\leqslant Cq|V|}$$

and, given an $I \in \bigcap_{j=1}^i \mathcal{I}(\mathcal{H}_j)$,

$$g(I) = g_1(I) \cup \ldots \cup g_i(I).$$

Suppose that $g(I) \subseteq I'$ and $g(I') \subseteq I$ for some $I, I' \in \bigcap_{j=1}^i \mathcal{I}(\mathcal{H}_j)$. Then also $g_j(I) \subseteq g(I) \subseteq I'$ and, similarly, $g_j(I') \subseteq g(I') \subseteq I$ for each $j \in [i]$. Assertion (iii) of Theorem 3.1 implies that that $g_j(I) = g_j(I')$ for each j and thus $g(I) = g(I')$. Since $g(I) \subseteq I$, we may also conclude that if $g(I) = g(I')$, then also $g_j(I) = g_j(I')$ for each $j \in [i]$. In particular, we may define, for each $I \in \bigcap_{j=1}^i \mathcal{I}(\mathcal{H}_j)$,

$$f(g(I)) = f_1(g_1(I)) \cap \ldots \cap f_i(g_i(I)).$$

It is routine to verify that $I \setminus g(I) \subseteq f(g(I))$ and that $e\big(\mathcal{H}_j[f(g(I))]\big) \leqslant \varepsilon e(\mathcal{H}_j)$ for every $j \in [i]$. $\square$

3.2. **Supersaturation results.** The following two statements can be proved using a standard averaging argument in the spirit of the classical supersaturation theorem of Erdős and Simonovits [10].

Lemma 3.3. *Given graphs H and T and a $\delta > 0$, there exists an $\varepsilon > 0$ such that the following holds. Every n-vertex graph G with $\mathcal{N}_T(G) \geqslant \mathrm{ex}(n, T, H) + \delta n^{v_T}$ contains more than εn^{v_H} copies of H.*

Lemma 3.4. *Given graphs H and T, a (finite) family $\mathcal{F}$ of T-coverings of H, and a $\delta > 0$, there exists an $\varepsilon > 0$ such that the following holds. For every collection $\mathcal{T} \subseteq T(K_n)$ with $|\mathcal{T}| \geqslant \mathrm{ex}^*(n, T, \mathcal{F}) + \delta n^{v_T}$, there exists an $F \in \mathcal{F}$ such that $\mathcal{T}$ contains more than $\varepsilon n^{v_{U(F)}}$ copies of F.*

Our final lemma states that the extremal function $\mathrm{ex}^*(n, T, \mathcal{F})$ corresponding to a family $\mathcal{F}$ of T-coverings of H can be approximated by $\mathrm{ex}(n, T, H)$ at least when $\mathcal{F}$ contains the special T-covering $F^e_{T,H}$ of H with e_H copies of T.

Lemma 3.5. *Given graphs H and T, let $F^e = F^e_{T,H}$ be the T-covering of H with e_H copies of T defined in Section 1.1. Then*

$$\mathrm{ex}^*(n, T, F^e) = \mathrm{ex}(n, T, H) + o(n^{v_T}).$$

Proof. Since the underlying graph of F^e contains a copy of H, then $\mathrm{ex}^*(n, T, F^e) \geqslant \mathcal{N}_T(G)$ for every H-free graph G. This shows that $\mathrm{ex}^*(n, T, F^e) \geqslant \mathrm{ex}(n, T, H)$. For the opposite inequality, fix an arbitrary $\varepsilon > 0$ and suppose that $\mathcal{T}$ is a collection of $\mathrm{ex}(n, T, H) + \varepsilon n^{v_T}$ copies of T in K_n. Let E be the set of all edges of K_n that belong to fewer than $\varepsilon n^{v_T - 2}$ copies of T from $\mathcal{T}$ and let $\mathcal{T}'$ comprise only those copies of T from $\mathcal{T}$ that contain no edge from E. Observe that

$$|\mathcal{T}'| \geqslant |\mathcal{T}| - |E| \cdot \varepsilon n^{v_T - 2} \geqslant |\mathcal{T}| - \binom{n}{2} \cdot \varepsilon n^{v_T - 2} > \mathrm{ex}(n, T, H).$$

Let $G \subseteq K_n$ be the union of all copies of T in $\mathcal{T}'$. Since $\mathcal{N}_T(G) \geqslant |\mathcal{T}'| > \mathrm{ex}(n, T, H)$, the graph G contains a copy of H. As each edge of G is contained in at least $\varepsilon n^{v_T - 2}$ copies of T from $\mathcal{T}$, each copy of H in G must be covered by a copy of F^e that is contained in $\mathcal{T}$. Indeed, given a copy of H in G, one may construct such an F^e greedily by considering the edges of H ordered arbitrarily as $f_1, \ldots, f_{e_H}$ and then finding some $T_i \in \mathcal{T}$ that contains f_i and whose remaining

$v_T - 2$ vertices lie outside of $V(T_1) \cup \ldots \cup V(T_{i-1})$, for each $i \in \{1, \ldots, e_H\}$ in turn. One is guaranteed to find such a T_i since the number of copies of T in $T(K_n)$ that contain f_i and have at least one more vertex in $V(T_1) \cup \ldots \cup V(T_{i-1})$ is only $O(n^{v_T-3})$. $\square$

3.3. **Properties of graph densities.** Here, we establish several useful facts relating the three notions of graph density that we consider in this work – the density, the 2-density, and the T-density. Our first lemma partially explains why the two cases $m_2(H) \leqslant m_2(T)$ and $m_2(H) > m_2(T)$, which we consider separately while studying the typical value of $\mathrm{ex}\big(G(n,p), T, H\big)$, are so different.

Lemma 3.6. *Suppose that H and T are fixed graphs and assume that T is 2-balanced.*

(i) If $m_2(H) \leqslant m_2(T)$, then the T-covering $F^e = F^e_{T,H}$ of H with e_H edges satisfies $m_T(F^e) \leqslant m_2(T)$.

(ii) If $m_2(H) > m_2(T)$, then every T-covering F of H satisfies $m_T(F) > m_2(T)$.

Proof. To see (i), assume that $m_2(H) \leqslant m_2(T)$ and fix some $F' \subseteq F^e$. Since F' is a T-covering of some subgraph $H' \subseteq H$ with $|F'|$ edges by pairwise edge-disjoint copies of T, then

$$\begin{aligned}\frac{e_{U(F')} - e_T}{v_{U(F')} - v_T} &= \frac{e_{H'}e_T - e_T}{v_{H'} + e_{H'}(v_T - 2) - v_T} = \frac{(e_{H'} - 1)(e_T - 1) + e_{H'} - 1}{(e_{H'} - 1)(v_T - 2) + v_{H'} - 2} \\ &\leqslant \max\left\{\frac{e_T - 1}{v_T - 2}, \frac{e_{H'} - 1}{v_{H'} - 2}\right\} \leqslant \max\{m_2(T), m_2(H)\} = m_2(T).\end{aligned}$$

To see (ii), assume that $m_2(H) > m_2(T)$ and let F be an arbitrary T-covering of H. Let $H' \subseteq H$ be any subgraph of H satisfying $\frac{e_{H'}-1}{v_{H'}-2} > \frac{e_T-1}{v_T-2}$ and denote by $T_1, \ldots, T_k$ all those elements of F that intersect H'. For each $i \in [k]$, denote by v_i and e_i the numbers of vertices and edges of $T_i \cap H'$, respectively, and note that $e_i - 1 \leqslant m_2(T)(v_i - 2)$. One easily verifies that

$$\begin{aligned}m_T(F) \geqslant \frac{e_{U(F')} - e_T}{v_{U(F')} - v_T} &= \frac{e_{H'} + \sum_{i=1}^k (e_T - e_i) - e_T}{v_{H'} + \sum_{i=1}^k (v_T - v_i) - v_T} \\ &= \frac{e_{H'} - 1 + (k-1)(e_T - 1) - \sum_{i=1}^k (e_i - 1)}{v_{H'} - 2 + (k-1)(v_T - 2) - \sum_{i=1}^k (v_i - 2)} > \frac{e_T - 1}{v_T - 2} = m_2(T),\end{aligned}$$

as claimed. $\square$

Our next lemma computes the rate of the lower tail of the number of copies of a 2-balanced graph T in $G(n,p)$, which Lemma 3.9, stated below, provides in a somewhat implicit form.

Lemma 3.7. *If T is a 2-balanced graph, then*

$$\min\left\{n^{v(T')}p^{e(T')} : \emptyset \neq T' \subseteq T\right\} = \begin{cases} n^{v_T}p^{e_T} & \text{if } p \leqslant n^{-1/m_2(T)}, \\ n^2 p & \text{if } p \geqslant n^{-1/m_2(T)}. \end{cases}$$

Proof. Let T be a 2-balanced graph. Suppose first that $p \geqslant n^{-1/m_2(T)}$ and fix some $T' \subseteq T$ with at least two edges. Since T is 2-balanced, then $m_2(T) \geqslant \frac{e_{T'}-1}{v_{T'}-2}$ and hence $p \geqslant n^{-\frac{v_{T'}-2}{e_{T'}-1}}$. It follows that

$$n^{v_{T'}}p^{e_{T'}} = n^2 p \cdot n^{v_{T'}-2}p^{e_{T'}-1} \geqslant n^2 p \cdot n^{v_{T'}-2}\left(n^{-\frac{v_{T'}-2}{e_{T'}-1}}\right)^{e_{T'}-1} = n^2 p.$$

Suppose now that $p \leqslant n^{-1/m_2(T)}$ and fix a nonempty $T' \subseteq T$. Since $m_2(T) \geqslant \frac{e_{T'}-1}{v_{T'}-2}$, then

$$\frac{e_T - e_{T'}}{m_2(T)} = \frac{(e_T - 1) - (e_{T'} - 1)}{m_2(T)} \geqslant (v_T - 2) - (v_{T'} - 2) = v_T - v_{T'}.$$

It follows that

$$n^{v_{T'}}p^{e_{T'}} = n^{v_T}p^{e_T} \cdot n^{v_{T'}-v_T}p^{e_{T'}-e_T} \geqslant n^{v_T}p^{e_T} \cdot n^{v_{T'}-v_T}\left(n^{1/m_2(T)}\right)^{e_T-e_{T'}} \geqslant n^{v_T}p^{e_T}, \qquad (4)$$

as required. $\square$

3.4. **Small subgraphs in $G(n,p)$.** Our proofs will require several properties of the distribution of the number of copies of a given fixed graph T in the random graph $G_{n,p}$. Following the classical approach of Ruciński and Vince [30], we first prove that if T is 2-balanced, then the number of copies of T in $G_{n,p}$ is concentrated around its expectation, provided that this expectation tends to infinity with n. Moreover, we show that if $p \ll n^{-1/m_2(T)}$, then copies of T in $G_{n,p}$ are essentially pairwise edge-disjoint.

Lemma 3.8. *Suppose that T is a fixed 2-balanced graph, assume that $p \gg n^{-v_T/e_T}$, and let $G \sim G(n,p)$. Then w.h.p. $\mathcal{N}_T(G) = (1+o(1)) \cdot \mathbb{E}[\mathcal{N}_T(G)]$. Moreover, if $p \ll n^{-1/m_2(T)}$, then w.h.p. G contains a subgraph G^* with the following two properties:*

(i) $\mathcal{N}_T(G^*) = (1+o(1)) \cdot \mathbb{E}[\mathcal{N}_T(G)]$.

(ii) Every edge of G^* belongs to exactly one copy of T.

Proof. Assume that $p \gg n^{-v_T/e_T}$ and let $G \sim G(n,p)$. Let $X = \mathcal{N}_T(G)$ and write Y for the number of pairs of distinct copies of T in G that share at least one edge. A routine calculation (see, e.g., [20, Chapter 3]) shows that

$$\mathbf{Var}(X) \leqslant \mathbb{E}[X] + \mathbb{E}[Y] \quad \text{and} \quad \mathbb{E}[Y] \leqslant C \cdot \mathbb{E}[X]^2 \cdot \left(\min\left\{n^{v(T')}p^{e(T')} : \emptyset \neq T' \subsetneq T\right\}\right)^{-1}$$

for some constant C that depends only on T. Since $\mathbb{E}[X] = \Theta\big(n^{v_T}p^{e_T}\big)$, our assumption on p implies that $\mathbb{E}[X] \to \infty$ and, by Lemma 3.7, that $\mathbf{Var}(X) \ll \mathbb{E}[X]^2$. This proves the first assertion of the lemma. To see the second assertion, suppose further than $p \ll n^{-1/m_2(T)}$. We claim that in this case,

$$\min\left\{n^{v(T')}p^{e(T')} : \emptyset \neq T' \subsetneq T\right\} \gg n^{v_T}p^{e_T}.$$

To see this, one can repeat the calculation in the proof of Lemma 3.7 observing that under the assumption that $p \ll n^{-1/m_2(T)}$, the first '$\geqslant$' in (4) can be replaced with a '$\gg$' (because $e_{T'} < e_T$). This means, in particular, that $\mathbb{E}[Y] \ll \mathbb{E}[X]$ and thus w.h.p. $Y \ll X$. Finally, observe that if $X = (1+o(1))\mathbb{E}[X]$ and $Y \ll X$, then one may obtain a graph G^* with the claimed properties by first removing from G all edges that belong to more than one copy of T and subsequently removing all edges that are not contained in any copy of T. $\square$

The following optimal tail estimate for the number of copies of a fixed graph T from a given family $\mathcal{T} \subseteq T(K_n)$ that appear in $G(n,p)$ is a rather straightforward extension of the result of Janson, Łuczak, and Ruciński [19].

Lemma 3.9. *For every graph T and constant $\delta > 0$, there exists a constant $\beta > 0$ such that the following holds. For every p and each collection $\mathcal{T}$ of copies of T in K_n,*

$$\Pr\left(\big|\mathcal{T} \cap T\big(G(n,p)\big)\big| \leqslant \big(|\mathcal{T}| - \delta n^{v_T}\big) \cdot p^{e_T}\right) \leqslant \exp\left(-\beta \cdot \min\left\{n^{v_{T'}}p^{e_{T'}} : \emptyset \neq T' \subseteq T\right\}\right).$$

In particular, if T is 2-balanced, then

$$\Pr\left(\big|\mathcal{T} \cap T\big(G(n,p)\big)\big| \leqslant \big(|\mathcal{T}| - \delta n^{v_T}\big) \cdot p^{e_T}\right) \leqslant \begin{cases} \exp\left(-\beta n^{v_T}p^{e_T}\right) & \text{if } p \leqslant n^{-1/m_2(T)}, \\ \exp\left(-\beta n^2 p\right) & \text{if } p \geqslant n^{-1/m_2(T)}. \end{cases}$$

Note that the second assertion of the lemma follows immediately from the main assertion and Lemma 3.7. We shall derive Lemma 3.9 from the following well-known inequality (see, for example, [4, Chapter 8]).

Theorem 3.10 (Janson's inequality)**.** *Suppose that Ω is a finite set and let $B_1, \ldots, B_k$ be arbitrary subsets of Ω. Form a random subset $R \subseteq \Omega$ by independently keeping each $\omega \in \Omega$ with probability $p_\omega \in [0,1]$. For each $i \in [k]$, let X_i be the indicator of the event that $B_i \subseteq R$. Let $X = \sum_i X_i$ and define*

$$\mu = \mathbb{E}[X] = \sum_{i=1}^{k} \prod_{\omega \in B_i} p_\omega \qquad \text{and} \qquad \Delta = \sum_{\substack{i \neq j \\ B_i \cap B_j \neq \emptyset}} \mathbb{E}[X_i X_j] = \sum_{\substack{i \neq j \\ B_i \cap B_j \neq \emptyset}} \prod_{\omega \in B_i \cup B_j} p_\omega.$$

Then for any $0 \leqslant t \leqslant \mu$,

$$\Pr\big(X \leqslant \mu - t\big) \leqslant \exp\left(-\frac{t^2}{2(\mu + \Delta)}\right).$$

Proof of Lemma 3.9. Suppose that $\mathcal{T} = \{T_1, \ldots, T_k\}$ and for each $i \in [k]$, let X_i be the indicator of the event that T_i appears in $G(n,p)$, so that

$$X = \sum_{i=1}^{k} X_i = \big|\mathcal{T} \cap T\big(G(n,p)\big)\big|.$$

Let μ and Δ be as in the statement of Theorem 3.10 and observe that

$$\mu = \mathbb{E}[X] = |\mathcal{T}| \cdot p^{e_T} \leqslant n^{v_T} p^{e_T}$$

and that

$$\begin{aligned}\Delta = \sum_{i=1}^{k} \sum_{\substack{j \neq i \\ T_i \cap T_j \neq \emptyset}} \Pr\big(T_i \cup T_j \subseteq G(n,p)\big) &\leqslant |\mathcal{T}| \cdot \sum_{\emptyset \neq T' \subsetneq T} n^{v_T - v_{T'}} p^{2e_T - e_{T'}} \\ &\leqslant 2^{e_T} n^{2v_T} p^{2e_T} \cdot \left(\min\left\{n^{v_{T'}} p^{e_{T'}} : \emptyset \neq T' \subsetneq T\right\}\right)^{-1}.\end{aligned}$$

It thus follows from Theorem 3.10 that

$$\begin{aligned}\Pr\big(X \leqslant \mu - \delta n^{v_T} p^{e_T}\big) &\leqslant \exp\left(-\frac{\delta^2 n^{2v_T} p^{2e_T}}{2(\mu + \Delta)}\right) \leqslant \exp\left(-\delta^2 n^{2v_T} p^{2e_T} \cdot \min\left\{\frac{1}{4\mu}, \frac{1}{4\Delta}\right\}\right) \\ &\leqslant \exp\left(-2^{-e_T - 2} \delta^2 \cdot \min\left\{n^{v_{T'}} p^{e_{T'}} : \emptyset \neq T' \subseteq T\right\}\right),\end{aligned}$$

as claimed. □

3.5. **Harris's inequality.** Our proofs of Theorems 1.6 and 1.9 will use the well-known correlation inequality due to Harris [16, Lemma 4.1]. Below, we state a version of this inequality that is a slight rephrasing of [4, Theorem 6.3.2]. A family $\mathcal{G}$ of graphs is called *decreasing* if for every $G \in \mathcal{G}$, every subgraph of G belongs to $\mathcal{G}$. A family $\mathcal{G}$ of subgraphs of K_n is called *increasing* if for every $G \in \mathcal{G}$, every $H \subseteq K_n$ such that $H \supseteq G$ also belongs to $\mathcal{G}$.

Theorem 3.11. *Let $\mathcal{G}_1$ and $\mathcal{G}_2$ be two families of subgraphs of K_n and suppose that $G \sim G(n,p)$. If $\mathcal{G}_1$ is decreasing and $\mathcal{G}_2$ is increasing, then*

$$\Pr(G \in \mathcal{G}_1 \text{ and } G \in \mathcal{G}_2) \leqslant \Pr(G \in \mathcal{G}_1) \cdot \Pr(G \in \mathcal{G}_2).$$

4. Proof of Theorems 1.6 and 1.9

4.1. **Proof of Theorem 1.6.** Suppose that H and T are fixed graphs and assume that T is 2-balanced and that $m_2(H) > m_2(T)$.

Proof of the first assertion. Suppose that $n^{-v_T/e_T} \ll p \ll n^{-1/m_2(H)}$ and let $G \sim G(n,p)$. It follows from Lemma 3.8 that w.h.p. $\mathcal{N}_T(G) = (1+o(1))\mathbb{E}[\mathcal{N}_T(G)] = (1+o(1))\mathcal{N}_T(K_n)p^{e_T}$. Therefore, it suffices to show that for every positive constant δ, w.h.p. G contains an H-free subgraph with at least $\big(\mathcal{N}_T(K_n) - \delta n^{v_T}\big) \cdot p^{e_T}$ copies of T. We shall argue somewhat differently depending on whether or not $p \ll n^{-1/m_2(T)}$.

Case 1. $p \ll n^{-1/m_2(T)}$. Suppose that G satisfies both assertions of Lemma 3.8 and let G^* be the subgraph of G from the statement of the lemma. Since each edge of G^* is contained in exactly one copy of T, then each copy of H in G^* must correspond to some T-covering of H in $T(G^*)$.[4] Consider an arbitrary T-covering F of H. Since we have assumed that $m_2(H) > m_2(T)$, part (ii) of Lemma 3.6 yields $m_T(F) > m_2(T)$. Since $p \ll n^{-1/m_2(T)} \ll n^{-1/m_T(F)}$, Remark 1.7 implies that there is some $F' \subseteq F$ such that

$$\mathbb{E}\big[\mathcal{N}_{U(F')}(G)\big] \ll \mathbb{E}\big[\mathcal{N}_T(G)\big].$$

Since there are only $O(1)$ types of T-coverings of H, then w.h.p. one may remove from G^* some $o\big(\mathbb{E}[\mathcal{N}_T(G)]\big)$ edges to obtain an H-free graph G_0. Since each edge of G^* belongs to exactly one copy of T, then

$$\mathcal{N}_T(G_0) = \mathcal{N}_T(G^*) - o\big(\mathbb{E}[\mathcal{N}_T(G)]\big) \geqslant (1+o(1)) \cdot \mathcal{N}_T(K_n)p^{e_T}.$$

Case 2. $p = \Omega(n^{-1/m_2(T)})$. Suppose that G satisfies the assertion of Lemma 3.8. Since $p \ll n^{-1/m_2(H)}$, then there is some $H' \subseteq H$ such that $\mathbb{E}[\mathcal{N}_{H'}(G)] \ll n^2p$. In particular, w.h.p. one may delete $o(n^2p)$ edges from G to make it H-free. It suffices to show that w.h.p. for every set X of $o(n^2p)$ edges of G, the graph $G \setminus X$ contains at least $\big(\mathcal{N}_T(K_n) - \delta n^{v_T}\big) \cdot p^{e_T}$ copies of T. For a fixed $X \subseteq E(K_n)$, let $\mathcal{A}_X$ denote the event that

$$\mathcal{N}_T(G \setminus X) \leqslant \big(\mathcal{N}_T(K_n) - \delta n^{v_T}\big) \cdot p^{e_T}.$$

Since $|X| \ll n^2$, then $\mathcal{N}_T(K_n \setminus X) = \mathcal{N}_T(K_n) - o(n^{v_T})$ and thus Lemma 3.9 with $\mathcal{T} \leftarrow T(K_n \setminus X)$ together with Lemma 3.7 yield

$$\Pr(\mathcal{A}_X) \leqslant \exp\big(-\beta n^2 p\big)$$

for some positive constant β. Since for every $X \subseteq E(K_n)$, the event $X \subseteq G$ is increasing and the event $\mathcal{A}_X$ is decreasing, Harris's inequality (Theorem 3.11) implies that

$$\Pr\big(X \subseteq G \text{ and } \mathcal{A}_X\big) \leqslant \Pr\big(X \subseteq G\big) \cdot \Pr(\mathcal{A}_X).$$

Consequently,

$$\begin{aligned}\Pr\big(\mathcal{A}_X \text{ for some } X \subseteq G \text{ with } |X| = o(n^2p)\big) &\leqslant \sum_{\substack{X \subseteq E(K_n)\\ |X| \ll n^2 p}} p^{|X|} \cdot \exp\big(-\beta n^2 p\big)\\ \leqslant \sum_{x \ll n^2p} \binom{\binom{n}{2}}{x} p^x \cdot \exp\big(-\beta n^2p\big) &\leqslant \sum_{x \ll n^2p} \left(\frac{en^2p}{2x}\right)^x \cdot \exp\big(-\beta n^2p\big) \leqslant \exp\big(-\beta n^2 p/2\big),\end{aligned}$$

as the function $x \mapsto (ea/x)^x$ is increasing when $x \leqslant a$. □

Proof of the second assertion. Suppose that $p \gg n^{-1/m_2(H)}$ and let $G \sim G(n,p)$. Our aim is to show that for every positive constant δ, w.h.p. every H-free subgraph G_0 of G satisfies

$$\mathcal{N}_T(G_0) \leqslant \big(\mathrm{ex}(n,T,H) + 2\delta n^{v_T}\big) \cdot p^{e_T}.$$

Let $\mathcal{H}$ be the e_H-uniform hypergraph with vertex set $E(K_n)$ whose edges are all copies of H in K_n. Observe that

$$v(\mathcal{H}) = \Theta\big(n^2\big) \qquad \text{and} \qquad e(\mathcal{H}) = \Theta\big(n^{v_H}\big)$$

[4] Recall that T-coverings of H are collections of pairwise edge-disjoint copies of T.

and that there is a natural one-to-one correspondence between the independent sets of $\mathcal{H}$ and H-free subgraphs of K_n. As we shall be applying Theorem 3.1 to the hypergraph $\mathcal{H}$, we let $q = n^{-1/m_2(H)}$ and verify that $\mathcal{H}$ satisfies the main assumption of the theorem, provided that K is a sufficiently large constant.

Claim 4.1. *There is a constant K such that the hypergraph $\mathcal{H}$ satisfies (2) in Theorem 3.1 with $q = n^{-1/m_2(H)}$.*

Proof. Fix an arbitrary $\ell \in [e_H]$ and note that $\Delta_\ell(\mathcal{H})$ is the largest number of copies of H in K_n that contain some given set of ℓ edges. It follows that

$$\Delta_\ell(\mathcal{H}) \leqslant \sum_{H' \subseteq H, e_{H'} = \ell} n^{v_H - v_{H'}}$$

and hence

$$\frac{v(\mathcal{H})}{e(\mathcal{H})} \cdot \max_{\ell \in [e_H]} \frac{\Delta_\ell(\mathcal{H})}{q^{\ell-1}} \leqslant 2^{e_H} \cdot n^{2-v_H} \cdot \max_{\emptyset \neq H' \subseteq H} \frac{n^{v_H - v_{H'}}}{q^{e_{H'}-1}} = 2^{e_H} \cdot \left(\min_{\emptyset \neq H' \subseteq H} n^{v_{H'}-2} q^{e_{H'}-1} \right)^{-1}.$$

Finally, since $q = n^{-1/m_2(H)}$, then $n^{v_{H'}-2} q^{e_{H'}-1} \geqslant 1$ for every nonempty $H' \subseteq H$. $\square$

Denote by $\mathrm{Free}_n(H)$ the family of all H-free subgraphs of K_n and let ε be the constant given by Lemma 3.3 invoked with $\delta/4$ in place of δ. Apply Theorem 3.1 to the hypergraph $\mathcal{H}$ to obtain a constant C, a family $\mathcal{S} \subseteq \binom{E(K_n)}{\leqslant Cqn^2}$, and functions $g\colon \mathrm{Free}_n(H) \to \mathcal{S}$ and $f\colon \mathcal{S} \to \mathcal{P}(E(K_n))$ such that:

(i) For every $G_0 \in \mathrm{Free}_n(H)$, $g(G_0) \subseteq G_0$ and $G_0 \setminus g(G_0) \subseteq f(g(G_0))$.

(ii) For every $S \in \mathcal{S}$, the graph $f(S)$ contains at most εn^{v_H} copies of H.

Given an $S \in \mathcal{S}$, denote by $\mathcal{A}_S$ the event

$$\big|T(G) \setminus T\big(f(S) \cup S\big)\big| \leqslant \big(\mathcal{N}_T(K_n) - \mathrm{ex}(n,T,H) - \delta n^{v_T}\big) \cdot p^{e_T}.$$

Claim 4.2. *There is a constant $\beta > 0$ such that for every $S \in \mathcal{S}$,*

$$\Pr(\mathcal{A}_S) \leqslant \exp\left(-\beta n^2 p\right).$$

Proof. Fix an $S \in \mathcal{S}$ and let $\mathcal{T}_S$ denote the collection of all copies of T in K_n that are not completely contained in $f(S) \cup S$. Since $|S| \ll n^2$, then property (ii) above and Lemma 3.3 imply that

$$\begin{aligned} |\mathcal{T}_S| &= \mathcal{N}_T(K_n) - \mathcal{N}_T\big(f(S) \cup S\big) \geqslant \mathcal{N}_T(K_n) - \mathcal{N}_T\big(f(S)\big) - |S| \cdot n^{v_T-2} \\ &\geqslant \mathcal{N}_T(K_n) - \mathrm{ex}(n,T,H) - \delta n^{v_T}/2. \end{aligned}$$

Since T is 2-balanced and $p \gg n^{-1/m_2(H)} \geqslant n^{-1/m_2(T)}$, Lemma 3.9 implies that

$$\Pr(\mathcal{A}_S) \leqslant \exp\left(-\beta n^2 p\right)$$

for some positive constant β, as claimed. $\square$

Suppose that G satisfies the assertion of Lemma 3.8 and let $G_0 \subseteq G$ be an H-free subgraph of G that maximizes $\mathcal{N}_T(G_0)$. Since $G_0 \in \mathrm{Free}_n(H)$, then

$$g(G_0) \subseteq G_0 \subseteq f(g(G_0)) \cup g(G_0).$$

and hence

$$\begin{aligned} \mathcal{N}_T(G_0) &\leqslant \max\left\{\big|T(G) \cap T\big(f(S) \cup S\big)\big| : S \in \mathcal{S} \text{ and } S \subseteq G\right\} \\ &= \mathcal{N}_T(G) - \min\left\{\big|T(G) \setminus T\big(f(S) \cup S\big)\big| : S \in \mathcal{S} \text{ and } S \subseteq G\right\} \\ &= (1+o(1)) \cdot \mathbb{E}[\mathcal{N}_T(G)] - \min\left\{\big|T(G) \setminus T\big(f(S) \cup S\big)\big| : S \in \mathcal{S} \text{ and } S \subseteq G\right\}. \end{aligned}$$

We shall show that w.h.p. $\mathcal{A}_S$ does not hold for any $S \in \mathcal{S}$ such that $S \subseteq G$, which will imply that

$$\mathcal{N}_T(G_0) \leqslant \big(\mathrm{ex}(n,T,H) + 2\delta n^{v_T}\big) \cdot p^{e_T}.$$

Since for every $S \in \mathcal{S}$, the event $S \subseteq G$ is increasing and the event $\mathcal{A}_S$ is decreasing, Harris's inequality (Theorem 3.11) implies that

$$\Pr\big(S \subseteq G \text{ and } \mathcal{A}_S\big) \leqslant \Pr\big(S \subseteq G\big) \cdot \Pr(\mathcal{A}_S).$$

By Claim 4.2, in order to complete the proof in this case, it is sufficient to prove the following.

Claim 4.3.

$$\sum_{S \in \mathcal{S}} \Pr\big(S \subseteq G\big) \leqslant \exp\big(o(n^2 p)\big).$$

Proof. Since each $S \in \mathcal{S}$ is a graph with at most Cqn^2 edges and $q = n^{-1/m_2(H)} \ll p$, then

$$\sum_{S \in \mathcal{S}} \Pr\big(S \subseteq G\big) \leqslant \sum_{s \leqslant Cqn^2} \binom{n^2}{s} p^s \leqslant \sum_{s = o(pn^2)} \left(\frac{en^2 p}{s}\right)^s = \exp\big(o(n^2 p)\big),$$

as the function $s \mapsto (ea/s)^s$ is increasing when $s \leqslant a$. □

This completes the proof of the second assertion of Theorem 1.6. □

4.2. **Proof of Theorem 1.9.** Suppose that H and T are fixed graphs and assume that T is 2-balanced and that $m_2(H) \leqslant m_2(T)$. Recall Definition 1.8, let $F_1, \ldots, F_k$ be the T-resolution of H, and let $p_0, p_1, \ldots, p_k$ be the associated threshold sequence. For each $i \in \{0, \ldots, k\}$, denote by $\mathcal{F}_i$ the set $\{F_1, \ldots, F_i\}$. Finally, let $F^e = F^e_{T,H}$ be the minimal covering of H with e_H pairwise edge-disjoint copies of T.

Proof of part (i). Fix an $i \in [k]$, suppose that $p_0 \ll p \ll p_i$, and let $G \sim G(n,p)$. Our aim is to show that for every positive constant δ, w.h.p. G contains an H-free subgraph with at least $\big(\mathrm{ex}^*\big(n,T,\mathcal{F}_{i-1}\big) - \delta n^{v_T}\big) \cdot p^{e_T}$ copies of T. If $\mathrm{ex}^*\big(n,T,\mathcal{F}_{i-1}\big) = o\big(n^{v_T}\big)$, then the assertion is trivial (we may simply take the empty graph), so for the remainder of the proof we shall assume that $\mathrm{ex}^*\big(n,T,\mathcal{F}_{i-1}\big) \geqslant \gamma n^{v_T}$ for some positive constant γ.

It follows from part (i) of Lemma 3.6 that $p \ll p_i \leqslant n^{-1/m_T(F^e)} \leqslant n^{-1/m_2(T)}$, so we may assume that G satisfies both assertions of Lemma 3.8. Let G^* be the subgraph of G from the statement of the lemma and let $\mathcal{T}_{i-1}$ be an extremal collection of copies of T in K_n with respect to being $\mathcal{F}_{i-1}$-free. In other words, let $\mathcal{T}_{i-1}$ be a collection of $\mathrm{ex}^*(n,T,\mathcal{F}_{i-1})$ copies of T that does not contain any T-covering of either of the types $F_1, \ldots, F_{i-1}$. Let G' be the graph obtained from G^* by keeping only edges covered by $T(G^*) \cap \mathcal{T}_{i-1}$ and let G_0 be the graph obtained from G' by deleting all edges from every copy of H in G'. This graph is clearly H-free. Since each edge of G^* is contained in exactly one copy of T, then each copy of H in G^* must belong to some T-covering of H. Since $\mathcal{T}_{i-1}$ is $\mathcal{F}_{i-1}$-free, then the only T-coverings of H that we may find in G' are $F_i, \ldots, F_k$ and coverings whose T-density is strictly greater than $m_T(F^e)$. Since $p \ll p_i \leqslant n^{-1/m_T(F^e)}$ and there are only $O(1)$ types of T-coverings, then w.h.p. there are only $o\big(\mathbb{E}[\mathcal{N}_T(G)]\big)$ edges in $G' \setminus G_0$ and thus $\mathcal{N}_T(G') - \mathcal{N}_T(G_0) = o\big(\mathbb{E}[\mathcal{N}_T(G)]\big)$, as every edge of G' belongs to at most one copy of T. Now, Lemma 3.9 implies that w.h.p.

$$|T(G) \cap \mathcal{T}_{i-1}| \geqslant \big(\mathrm{ex}^*\big(n,T,\mathcal{F}_{i-1}\big) - \delta n^{v_T}/3\big) \cdot p^{e_T}.$$

Therefore,

$$\begin{aligned}\mathcal{N}_T(G_0) &\geqslant \mathcal{N}_T(G') - \delta n^{v_T} p^{e_T}/3 = |T(G^*) \cap \mathcal{T}_{i-1}| - \delta n^{v_T} p^{e_T}/3 \\ &\geqslant |T(G) \cap \mathcal{T}_{i-1}| - (\mathcal{N}_T(G) - \mathcal{N}_T(G^*)) - \delta n^{v_T} p^{e_T}/3 \geqslant \big(\mathrm{ex}^*\big(n,T,\mathcal{F}_{i-1}\big) - \delta n^{v_T}\big) \cdot p^{e_T},\end{aligned}$$

since $\mathcal{N}_T(G) = \mathcal{N}_T(G^*) + o\big(n^{v_T} p^{e_T}\big)$. □

Proof of part (ii). Fix an $i \in [k]$, suppose that $p \gg p_i$, and let $G \sim G(n,p)$. Our aim is to show that for every positive constant δ, w.h.p. every H-free subgraph G_0 of G satisfies

$$\mathcal{N}_T(G_0) \leqslant \bigl(\mathrm{ex}^*\bigl(n, T, \mathcal{F}_i\bigr) + 2\delta n^{v_T}\bigr) \cdot p^{e_T}. \tag{5}$$

For each $j \in [i]$, let $\mathcal{H}_j$ be the $|F_j|$-uniform hypergraph whose vertices are all copies of T in K_n and whose edges are all collections of $|F_j|$ copies of T in K_n that are isomorphic to the T-covering F_j. Observe that

$$v(\mathcal{H}_j) = \Theta\bigl(n^{v_T}\bigr) \qquad \text{and} \qquad e(\mathcal{H}_j) = \Theta\bigl(n^{v_{U(F_j)}}\bigr).$$

Since $U(F_j)$ contains a copy of H, then for every H-free graph G_0, the family $T(G_0)$ is an independent set in $\mathcal{H}_j$, for each $j \in [i]$. As we shall be applying Corollary 3.2 to the hypergraphs $\mathcal{H}_1, \ldots, \mathcal{H}_i$, we let $q = p_i^{e_T}$ and verify that all $\mathcal{H}_j$ satisfy the main assumption of the corollary, provided that K is a sufficiently large constant.

Claim 4.4. *There is a constant K such that for each $j \in [i]$, the hypergraph $\mathcal{H}_j$ satisfies* (3) *in Corollary 3.2 with $q = p_i^{e_T}$.*

Proof. Fix an arbitrary $\ell \in [|F_j|]$ and note that $\Delta_\ell(\mathcal{H}_j)$ is the largest number of copies of F_j in $T(K_n)$ that share the same set of ℓ copies of T. It follows that

$$\Delta_\ell(\mathcal{H}_j) \leqslant \sum_{F' \subseteq F_j, |F'| = \ell} n^{v_{U(F_j)} - v_{U(F')}}$$

and hence

$$\begin{aligned}\frac{v(\mathcal{H}_j)}{e(\mathcal{H}_j)} \cdot \max_{\ell \in [|F_j|]} \frac{\Delta_\ell(\mathcal{H}_j)}{q^{\ell-1}} &\leqslant 2^{|F_j|} \cdot \frac{n^{v_T}}{n^{v_{U(F_j)}}} \cdot \max_{\emptyset \neq F' \subseteq F_j} \frac{n^{v_{U(F_j)} - v_{U(F')}}}{p_i^{e_T \cdot (|F'|-1)}} \\ &= 2^{|F_j|} \cdot \left(\min_{\emptyset \neq F' \subseteq F_j} n^{v_{U(F')} - v_T} p_i^{e_{U(F')} - e_T} \right)^{-1}.\end{aligned}$$

Finally, since $p_i \geqslant p_j = n^{-1/m_T(F_j)}$, then

$$n^{v_{U(F')} - v_T} p_i^{e_{U(F')} - e_T} \geqslant 1$$

for every nonempty $F' \subseteq F_j$. □

Denote by $\mathrm{Free}_n(\mathcal{F}_i)$ the family of all subfamilies of $T(K_n)$ that do not contain any T-covering isomorphic to one of the members of $\mathcal{F}_i$ and let ε be the constant given by Lemma 3.4 invoked with $\delta/2$ in place of δ. Apply Corollary 3.2 to the hypergraphs $\mathcal{H}_1, \ldots, \mathcal{H}_i$ to obtain a constant C, a family $\mathcal{S} \subseteq \binom{T(K_n)}{\leqslant Cqn^{v_T}}$, and functions $g \colon \mathrm{Free}_n(\mathcal{F}_i) \to \mathcal{S}$ and $f \colon \mathcal{S} \to \mathcal{P}(T(K_n))$ such that:

(i) For every $\mathcal{T} \in \mathrm{Free}_n(\mathcal{F}_i)$, $g(\mathcal{T}) \subseteq \mathcal{T}$ and $\mathcal{T} \setminus g(\mathcal{T}) \subseteq f(g(\mathcal{T}))$.

(ii) For every $S \in \mathcal{S}$, the collection $f(S)$ has at most $\varepsilon n^{v_{U(F_j)}}$ copies of F_j for every $j \in [i]$.

(iii) If $g(\mathcal{T}) \subseteq \mathcal{T}'$ and $g(\mathcal{T}') \subseteq \mathcal{T}$ for some $\mathcal{T}, \mathcal{T}' \in \mathrm{Free}_n(\mathcal{F}_i)$, then $g(\mathcal{T}) = g(\mathcal{T}')$.

Given an $S \in \mathcal{S}$, denote by $\mathcal{A}_S$ the event

$$\bigl|T(G) \setminus f(S)\bigr| \leqslant \bigl(\mathcal{N}_T(K_n) - \mathrm{ex}^*(n, T, \mathcal{F}_i) - \delta n^{v_T}\bigr) \cdot p^{e_T}.$$

Claim 4.5. *There is a constant $\beta > 0$ such that for every $S \in \mathcal{S}$,*

$$\Pr(\mathcal{A}_S) \leqslant \exp\left(-\beta \cdot \min\left\{n^2 p, n^{v_T} p^{e_T}\right\}\right).$$

Proof. Fix an $S \in \mathcal{S}$ and let $\mathcal{T}_S$ denote the collection of all copies of T in K_n that do not belong to $f(S)$. Property (ii) above and Lemma 3.4 imply that

$$|\mathcal{T}_S| = \mathcal{N}_T(K_n) - |f(S)| \geqslant \mathcal{N}_T(K_n) - \mathrm{ex}^*(n, T, \mathcal{F}_i) - \delta n^{v_T}/2.$$

Since T is 2-balanced, Lemma 3.9 implies that

$$\Pr(\mathcal{A}_S) \leqslant \exp\left(-\beta \cdot \min\left\{n^2 p, n^{v_T} p^{e_T}\right\}\right)$$

for some positive constant β, as claimed. $\square$

We shall now argue somewhat differently depending on whether or not $p \ll n^{-1/m_2(T)}$.

Case 1. $p \ll n^{-1/m_2(T)}$. Suppose that G satisfies both assertions of Lemma 3.8 and let G^* be the subgraph of G from the statement of the lemma. Let $G_0 \subseteq G$ be an H-free subgraph of G that maximizes $\mathcal{N}_T(G_0)$ and let $G' = G_0 \cap G^*$. Since

$$\mathcal{N}_T(G_0) \leqslant \mathcal{N}_T(G') + \mathcal{N}_T(G) - \mathcal{N}_T(G^*) = \mathcal{N}_T(G') + o\big(\mathbb{E}[\mathcal{N}_T(G)]\big) = \mathcal{N}_T(G') + o\big(n^{v_T} p^{e_T}\big),$$

it suffices to show that (5) holds with G_0 replaced by G'. Since G' is H-free, then $T(G') \in \mathrm{Free}_n(\mathcal{F}_i)$ and hence

$$g\big(T(G')\big) \subseteq T(G') \subseteq f\big(g\big(T(G')\big)\big) \cup g\big(T(G')\big).$$

But this means that

$$\begin{aligned}
\mathcal{N}_T(G') &\leqslant \max\left\{|T(G) \cap f(S)| + |S| : S \in \mathcal{S} \text{ and } S \subseteq T(G')\right\} \\
&= \mathcal{N}_T(G) - \min\left\{|T(G) \setminus f(S)| - |S| : S \in \mathcal{S} \text{ and } S \subseteq T(G')\right\} \\
&\leqslant (1+o(1)) \cdot \mathbb{E}[\mathcal{N}_T(G)] + C p_i^{e_T} n^{v_T} - \min\left\{|T(G) \setminus f(S)| : S \in \mathcal{S} \text{ and } S \subseteq T(G')\right\} \\
&= (1+o(1)) \cdot \mathbb{E}[\mathcal{N}_T(G)] - \min\left\{|T(G) \setminus f(S)| : S \in \mathcal{S} \text{ and } S \subseteq T(G')\right\}.
\end{aligned}$$

Now, let $\mathcal{S}'$ comprise all the sets of T-copies $S \in \mathcal{S}$ that are pairwise edge-disjoint. Since $T(G')$ is a collection of pairwise edge-disjoint copies of T, then

$$\min\left\{|T(G) \setminus f(S)| : S \in \mathcal{S} \text{ and } S \subseteq T(G')\right\} = \min\left\{|T(G) \setminus f(S)| : S \in \mathcal{S}' \text{ and } S \subseteq T(G')\right\}.$$

We shall now show that w.h.p. $\mathcal{A}_S$ does not hold for any $S \in \mathcal{S}'$ such that $S \subseteq T(G)$, which will imply that

$$\mathcal{N}_T(G') \leqslant \left(\mathrm{ex}^*(n, T, \mathcal{F}') + 2\delta n^{v_T}\right) \cdot p^{e_T}.$$

Since for every $S \in \mathcal{S}$, the event $S \subseteq T(G)$ is increasing and the event $\mathcal{A}_S$ is decreasing, Harris's inequality (Theorem 3.11) implies that

$$\Pr\left(S \subseteq T(G) \text{ and } \mathcal{A}_S\right) \leqslant \Pr\left(S \subseteq T(G)\right) \cdot \Pr(\mathcal{A}_S).$$

Since we have assumed that $p \ll n^{-1/m_2(T)}$, then Claim 4.5 and Lemma 3.7 imply that

$$\Pr(\mathcal{A}_s) \leqslant \exp\left(-\beta n^{v_T} p^{e_T}\right)$$

and consequently, in order to complete the proof in this case, it is sufficient to prove the following.

Claim 4.6.

$$\sum_{S \in \mathcal{S}'} \Pr\left(S \subseteq T(G)\right) \leqslant \exp\left(o(n^{v_T} p^{e_T})\right).$$

Proof. Since each $S \in \mathcal{S}'$ consists of pairwise edge-disjoint copies of T, then

$$\Pr\left(S \subseteq T(G)\right) = \Pr\left(U(S) \subseteq G\right) = p^{e_{U(S)}} = p^{e_T \cdot |S|}.$$

Since $\mathcal{S}'$ contains only sets of at most $C p_i^{e_T} n^{v_T}$ copies of T in K_n and $p_i \ll p$, it now follows that

$$\begin{aligned}
\sum_{S \in \mathcal{S}'} \Pr\left(S \subseteq T(G)\right) &= \sum_{S \in \mathcal{S}'} p^{e_T \cdot |S|} \leqslant \sum_{s \leqslant C p_i^{e_T} n^{v_T}} \binom{n^{v_T}}{s} p^{e_T \cdot s} \\
&\leqslant \sum_{s = o(p^{e_T} n^{v_T})} \left(\frac{e n^{v_T} p^{e_T}}{s}\right)^s = \exp\left(o\big(n^{v_T} p^{e_T}\big)\right),
\end{aligned}$$

since the function $s \mapsto (ea/s)^s$ is increasing when $s \leqslant a$. $\square$

Case 2. $p = \Omega(n^{-1/m_2(T)})$. Suppose that G satisfies the assertion of Lemma 3.8 and let $G_0 \subseteq G$ be an H-free subgraph of G that maximizes $\mathcal{N}_T(G_0)$. Since G_0 is H-free, then $T(G_0) \in \mathrm{Free}_n(\mathcal{F}_i)$ and hence

$$g\big(T(G_0)\big) \subseteq T(G_0) \subseteq f\big(g\big(T(G_0)\big)\big) \cup g\big(T(G_0)\big).$$

Now, let $\mathcal{S}''$ comprise all the sets of T-copies $S \in \mathcal{S}$ that are of the form $g\big(T(G'')\big)$ for some H-free graph $G'' \subseteq K_n$ and observe that

$$\begin{aligned}\mathcal{N}_T(G_0) &\leqslant \max\big\{|T(G) \cap f(S)| + |S| : S \in \mathcal{S}'' \text{ and } S \subseteq T(G)\big\}\\ &= \mathcal{N}_T(G) - \min\big\{|T(G) \setminus f(S)| - |S| : S \in \mathcal{S}'' \text{ and } S \subseteq T(G)\big\}\\ &= (1+o(1)) \cdot \mathbb{E}[\mathcal{N}_T(G)] - \min\big\{|T(G) \setminus f(S)| : S \in \mathcal{S}'' \text{ and } S \subseteq T(G)\big\}.\end{aligned}$$

Analogously to Case 1, we shall show that w.h.p. $\mathcal{A}_S$ does not hold for any $S \in \mathcal{S}''$ such that $S \subseteq T(G)$, which will imply that

$$\mathcal{N}_T(G_0) \leqslant (\mathrm{ex}^*(n,T,\mathcal{F}_i) + 2\delta n^{v_T}) \cdot p^{e_T},$$

as claimed. As before, since for every $S \in \mathcal{S}$, the event $S \subseteq T(G)$ is increasing and the event $\mathcal{A}_S$ is decreasing, Harris's inequality (Theorem 3.11) implies that

$$\Pr\big(S \subseteq T(G) \text{ and } \mathcal{A}_S\big) \leqslant \Pr\big(S \subseteq T(G)\big) \cdot \Pr(\mathcal{A}_S).$$

Since we have assumed that $p = \Omega\big(n^{-1/m_2(T)}\big)$, then Claim 4.5 and Lemma 3.7 imply that

$$\Pr(\mathcal{A}_s) \leqslant \exp\big(-\beta n^2 p\big)$$

and consequently, in order to complete the proof in this case, it is sufficient to prove the following.

Claim 4.7.

$$\sum_{S \in \mathcal{S}''} \Pr\big(S \subseteq T(G)\big) \leqslant \exp\big(o(n^2p)\big).$$

Proof. We claim that the function U that maps a collection of copies of T to its underlying graph is injective when restricted to $\mathcal{S}''$. Indeed, suppose that $U\left(g\big(T(G_1)\big)\right) = U\left(g\big(T(G_2)\big)\right)$ for some H-free graphs G_1 and G_2. It follows that

$$g\big(T(G_1)\big) \subseteq T\left(U\left(g\big(T(G_1)\big)\right)\right) = T\left(U\left(g\big(T(G_2)\big)\right)\right) \subseteq T\left(U\big(T(G_2)\big)\right) = T(G_2)$$

and, vice-versa, $g\big(T(G_2)\big) \subseteq T(G_1)$. The consistency property of the function g, see (iii) above, implies that $g\big(T(G_1)\big) = g\big(T(G_2)\big)$.

Since $p_i \leqslant n^{-1/m_2(T)}$ by part (i) of Lemma 3.6, then Lemma 3.7 implies that $n^{v_T}p_i^{e_T} \leqslant n^2 p_i$. In particular, each $S \in \mathcal{S}''$ comprises at most Cn^2p_i copies of T and therefore $U(S)$ has at most $Ce_Tn^2p_i$ edges. Since the function U is injective when restricted to $\mathcal{S}''$, we may conclude that

$$\begin{aligned}\sum_{S \in \mathcal{S}''} \Pr\big(S \subseteq T(G)\big) &= \sum_{S \in \mathcal{S}''} \Pr\big(U(S) \subseteq G\big) = \sum_{U \in U(\mathcal{S}'')} \Pr\big(U \subseteq G\big) = \sum_{U \in U(\mathcal{S}'')} p^{e_U}\\ &\leqslant \sum_{u \leqslant Ce_Tn^2p_i} \binom{\binom{n}{2}}{u} p^u \leqslant \sum_{s = o(pn^2)} \left(\frac{en^2p}{2u}\right)^u = \exp\big(o(n^2p)\big),\end{aligned}$$

since the function $u \mapsto (ea/u)^u$ is increasing when $u \leqslant a$. $\square$

This completes the proof of part (ii) of Theorem 1.9. $\square$

5. Concluding remarks and open questions

In this paper, we have studied the random variable $\mathrm{ex}\big(G(n,p),T,H\big)$ that counts the largest number of copies of T in an H-free subgraph of the binomial random graph $G(n,p)$. We restricted our attention to the case when T is 2-balanced; the case when T is not 2-balanced poses further challenges and we were not able to resolve it using our methods. The threshold phenomena associated with the variable $\mathrm{ex}\big(G(n,p),T,H\big)$ are quite different depending on whether or not the inequality $m_2(H) > m_2(T)$ holds:

(i) If $m_2(H) > m_2(T)$, then our Theorem 1.6 offers a natural generalization of a sparse random analogue of the Erdős–Stone theorem that was proved several years ago by Conlon and Gowers [7] and by Schacht [33].

(ii) If $m_2(H) \leqslant m_2(T)$, then the 'evolution' of the random variable $\mathrm{ex}\big(G(n,p),T,H\big)$ as p grows from 0 to 1 exhibits a more complex behavior. Our Theorem 1.9 shows that there are several potential 'phase transitions' and that the typical values of the variable between these phase transitions are determined by solutions to deterministic hypergraph Turán-type problems which we were unable to solve in full generality.

There are several natural directions for further investigations that are suggested by this work:

- It would be interesting to study the variable $\mathrm{ex}\big(G(n,p),T,H\big)$ for general graphs T and H, that is, without assuming that T is 2-balanced.
- We have very little understanding of the Turán-type problems related to T-coverings of H that are described in Section 1.1, even in the case when T is a complete graph. A concrete problem that we found the most interesting is stated as Question 1.11. In short, we ask if there exists a pair of graphs T and H such that the variable $\mathrm{ex}\big(G(n,p),T,H\big)$ undergoes multiple 'phase transitions'.
- Given a family $\mathcal{H}$ of graphs, one may more generally ask to study the random variable $\mathrm{ex}\big(G(n,p),T,\mathcal{H}\big)$ that counts the largest number of copies of T in a subgraph of $G(n,p)$ that is free of every $H \in \mathcal{H}$. This problem is solved when $T = K_2$ and $\mathcal{H}$ is finite, see [27, Theorem 6.4], but not much is known, even in the deterministic case ($p = 1$), when $T \neq K_2$.

Acknowledgment: We are indebted to the two anonymous referees for their careful reading of the manuscript and many helpful suggestions. The second author thanks Orit Raz for helpful discussions.

References

1. P. Allen, J. Böttcher, S. Griffiths, Y. Kohayakawa, and R. Morris, *The chromatic thresholds of graphs*, Adv. Math. **235** (2013), 261–295.
2. N. Alon, A. Kostochka, and C. Shikhelman, *Many cliques in H-free subgraphs of random graphs*, J. Comb. **9** (2018), 567–597.
3. N. Alon and C. Shikhelman, *Many T copies in H-free graphs*, J. Combin. Theory Ser. B **121** (2016), 146–172.
4. N. Alon and J. H. Spencer, *The probabilistic method*, fourth ed., Wiley Series in Discrete Mathematics and Optimization, John Wiley & Sons, Inc., Hoboken, NJ, 2016.
5. J. Balogh, R. Morris, and W. Samotij, *Independent sets in hypergraphs*, J. Amer. Math. Soc. **28** (2015), 669–709.
6. ______, *The method of hypergraph containers*, Proceedings of the International Congress of Mathematicians—Rio de Janeiro 2018. Vol. 3, 2018, pp. 3045–3078.
7. D. Conlon and W. T. Gowers, *Combinatorial theorems in sparse random sets*, Ann. of Math. (2) **184** (2016), 367–454.
8. D. Conlon, W. T. Gowers, W. Samotij, and M. Schacht, *On the KLR conjecture in random graphs*, Israel J. Math. **203** (2014), 535–580.

9. P. Erdős, *On the number of complete subgraphs contained in certain graphs*, Magyar Tud. Akad. Mat. Kutató Int. Közl. **7** (1962), 459–464.
10. P. Erdős and M. Simonovits, *Supersaturated graphs and hypergraphs*, Combinatorica **3** (1983), 181–192.
11. P. Erdős and A. H. Stone, *On the structure of linear graphs*, Bull. Amer. Math. Soc. **52** (1946), 1087–1091.
12. E. Friedgut, V. Rödl, and M. Schacht, *Ramsey properties of random discrete structures*, Random Structures Algorithms **37** (2010), 407–436.
13. Z. Füredi and M. Simonovits, *The history of degenerate (bipartite) extremal graph problems*, Erdös centennial, Bolyai Soc. Math. Stud., vol. 25, János Bolyai Math. Soc., Budapest, 2013, pp. 169–264.
14. A. Grzesik, *On the maximum number of five-cycles in a triangle-free graph*, J. Combin. Theory Ser. B **102** (2012), 1061–1066.
15. L. Gugelmann, R. Nenadov, Y. Person, N. Škorić, A. Steger, and H. Thomas, *Symmetric and asymmetric Ramsey properties in random hypergraphs*, Forum Math. Sigma **5** (2017), e28, 47.
16. T. E. Harris, *A lower bound for the critical probability in a certain percolation process*, Proc. Cambridge Philos. Soc. **56** (1960), 13–20.
17. H. Hatami, J. Hladký, D. Král', S. Norine, and A. Razborov, *On the number of pentagons in triangle-free graphs*, J. Combin. Theory Ser. A **120** (2013), 722–732.
18. P. E. Haxell, Y. Kohayakawa, and T. Łuczak, *Turán's extremal problem in random graphs: forbidding even cycles*, J. Combin. Theory Ser. B **64** (1995), 273–287.
19. S. Janson, T. Łuczak, and A. Ruciński, *An exponential bound for the probability of nonexistence of a specified subgraph in a random graph*, Random graphs '87 (Poznań, 1987), Wiley, Chichester, 1990, pp. 73–87.
20. S. Janson, T. Łuczak, and A. Rucinski, *Random graphs*, Wiley-Interscience Series in Discrete Mathematics and Optimization, Wiley-Interscience, New York, 2000.
21. S. Janson, K. Oleszkiewicz, and A. Ruciński, *Upper tails for subgraph counts in random graphs*, Israel J. Math. **142** (2004), 61–92.
22. Y. Kohayakawa and B. Kreuter, *Threshold functions for asymmetric Ramsey properties involving cycles*, Random Structures Algorithms **11** (1997), 245–276.
23. Y. Kohayakawa, B. Kreuter, and A. Steger, *An extremal problem for random graphs and the number of graphs with large even-girth*, Combinatorica **18** (1998), 101–120.
24. Y. Kohayakawa, T. Łuczak, and V. Rödl, *On K^4-free subgraphs of random graphs*, Combinatorica **17** (1997), 173–213.
25. Y. Kohayakawa, M. Schacht, and R. Spöhel, *Upper bounds on probability thresholds for asymmetric Ramsey properties*, Random Structures Algorithms **44** (2014), 1–28.
26. T. Kővári, V. T. Sós, and P. Turán, *On a problem of K. Zarankiewicz*, Colloquium Math. **3** (1954), 50–57.
27. R. Morris, W. Samotij, and D. Saxton, *An asymmetric container lemma and the structure of graphs with no induced* 4*-cycle*, arXiv:1806.03706 [math.CO].
28. R. Morris and D. Saxton, *The number of $C_{2\ell}$-free graphs*, Adv. Math. **298** (2016), 534–580.
29. F. Mousset, R. Nenadov, and W. Samotij, *Towards the Kohayakawa–Kreuter conjecture on asymmetric Ramsey properties*, arXiv:1808.05070 [math.CO].
30. A. Ruciński and A. Vince, *Balanced graphs and the problem of subgraphs of random graphs*, Proceedings of the sixteenth Southeastern international conference on combinatorics, graph theory and computing (Boca Raton, Fla., 1985), vol. 49, 1985, pp. 181–190.
31. W. Samotij, *Stability results for random discrete structures*, Random Structures Algorithms **44** (2014), 269–289.
32. D. Saxton and A. Thomason, *Hypergraph containers*, Invent. Math. **201** (2015), 925–992.
33. M. Schacht, *Extremal results for random discrete structures*, Ann. of Math. (2) **184** (2016), 333–365.
34. M. Simonovits, *Paul Erdős' influence on extremal graph theory*, The mathematics of Paul Erdős, II, Algorithms Combin., vol. 14, Springer, Berlin, 1997, pp. 148–192.

School of Mathematical Sciences, Tel Aviv University, Tel Aviv 6997801, Israel
Email address: `samotij@post.tau.ac.il`

School of Mathematical Sciences, Tel Aviv University, Tel Aviv 6997801, Israel
Email address: `clarashk@mail.tau.ac.il`