Forum of Mathematics, Sigma (2026), **Vol. 14**:e20 1–22
doi:10.1017/fms.2025.10163

CAMBRIDGE UNIVERSITY PRESS

RESEARCH ARTICLE

On the number of $\mathcal{H}$-free hypergraphs

Tao Jiang [1] and Sean Longbrake [2]

[1]Department of Mathematics, Miami University, 100 Bishop Circle, Oxford, Ohio, USA;
E-mail: jiangt@miamioh.edu (Corresponding author).
[2]Department of Mathematics, Emory University, 400 Dowman Dr, Atlanta, Georgia, USA; E-mail: sean.longbrake@emory.edu.

Received: 16 March 2025; **Revised:** 18 September 2025; **Accepted:** 18 September 2025

2020 Mathematics Subject Classification: *Primary* – 05D05; *Secondary* – 05C35

Abstract
Two central problems in extremal combinatorics are concerned with estimating the number $\mathrm{ex}(n, \mathcal{H})$, the size of the largest $\mathcal{H}$-free hypergraph on n vertices, and the number $\mathrm{forb}(n, \mathcal{H})$ of $\mathcal{H}$-free hypergraph on n vertices. It is well known that $\mathrm{forb}(n, \mathcal{H}) = 2^{(1+o(1))\mathrm{ex}(n,\mathcal{H})}$ for k-uniform hypergraphs that are not k-partite. In a recent breakthrough, Ferber, McKinley, and Samotij proved that for many k-partite (or *degenerate*) hypergraphs $\mathcal{H}$, $\mathrm{forb}(n, \mathcal{H}) = 2^{O(\mathrm{ex}(n,\mathcal{H}))}$. However, there are few known instances of degenerate hypergraphs $\mathcal{H}$ for which $\mathrm{forb}(n, \mathcal{H}) = 2^{(1+o(1))\mathrm{ex}(n,\mathcal{H})}$ holds.

In this paper, we show that $\mathrm{forb}(n, \mathcal{H}) = 2^{(1+o(1))\mathrm{ex}(n,\mathcal{H})}$ holds for a wide class of degenerate hypergraphs known as 2-contractible hypertrees. This is the first known infinite family of degenerate hypergraphs $\mathcal{H}$ for which $\mathrm{forb}(n, \mathcal{H}) = 2^{(1+o(1))\mathrm{ex}(n,\mathcal{H})}$ holds. As a corollary of our main results, we obtain a sharp estimate of $\mathrm{forb}(n, C_\ell^{(k)}) = 2^{(\lfloor \frac{\ell-1}{2} \rfloor + o(1))\binom{n}{k-1}}$ for the k-uniform linear ℓ-cycle, for all pairs $k \geq 5, \ell \geq 3$, thus settling a question of Balogh, Narayanan, and Skokan affirmatively for all $k \geq 5, \ell \geq 3$. Our methods also lead to some sharp results on the related random Turán problem.

As a key ingredient of our proofs, we develop a novel supersaturation variant of the delta systems method for set systems, which may be of independent interest.

Contents

1. Introduction

1.1. History

In extremal combinatorics, the problem of enumerating the number of discrete structures that avoid given substructures has a very rich history. One of the most natural questions one may ask is as follows: given a fixed graph H, how many n-vertex (labelled) graphs are there that contain no copy of H? Formally, given a fixed graph H, we say that a graph G is *H-free* if it does not contain H as a subgraph. For each natural number n, we let $\mathrm{Forb}(n, H)$ denote the family of all labeled H-free graphs on the vertex set $[n] := \{1, \dots, n\}$ and let $\mathrm{forb}(n, H) = |\mathrm{Forb}(n, H)|$. The problem is to determine or estimate $\mathrm{forb}(n, H)$. This function is closely related to another classic function studied in extremal graph theory, namely the *extremal number* $\mathrm{ex}(n, H)$, defined as the maximum number of edges in an n-vertex H-free graph. Indeed, if we take a maximum n-vertex H-free graph G and take all the subgraphs of it we get $2^{\mathrm{ex}(n,H)}$ many H-free graphs. On the other hand, for each $0 \le i \le \mathrm{ex}(n, H)$ there are at most $\binom{\binom{n}{2}}{i}$ many n-vertex H-free graphs with i edges. Hence, we trivially have

$$2^{\mathrm{ex}(n,H)} \le \mathrm{forb}(n, H) \le \sum_{i \le \mathrm{ex}(n,H)} \binom{\binom{n}{2}}{i} = n^{O(\mathrm{ex}(n,H))}. \tag{1}$$

For non-bipartite graphs H, the upper bound in (1) was significantly sharpened by Erdős, Frankl, and Rödl [10], extending the earlier seminal work of Erdős, Kleitman, and Rothschild [11] for complete graphs, showing that $\mathrm{forb}(n, H) \le 2^{\mathrm{ex}(n,H)+o(n^k)}$. Therefore, for any non-bipartite H,

$$\mathrm{forb}(n, H) = 2^{(1+o(1))\mathrm{ex}(n,H)}. \tag{2}$$

For bipartite graphs H, estimating $\mathrm{forb}(n, H)$ is much more difficult, even for the few bipartite H whose extremal numbers $\mathrm{ex}(n, H)$ are relatively well-understood. The first breakthrough in this area was made by Kleitman and Winston [27], who showed that the number of C_4-free graphs was $2^{O(\mathrm{ex}(n,C_4))}$. For complete bipartite graphs, this was extended by Balogh and Samotij first for symmetric [4] and then asymmetric [5] versions, showing that the number of $K_{s,t}$-free graphs is no more than $2^{O(n^{2-1/s})}$, a near optimal result for t sufficiently large compared to s. For even cycles, this was extended by Morris and Saxton [32] in a breakthrough work, showing that the number of $C_{2\ell}$-free graphs is no more than $2^{O(n^{1+\frac{1}{\ell}})}$. In a more recent breakthrough, Ferber, McKinley, and Samotij [14] showed that for all bipartite graphs satisfying a mild condition (see their Theorem 5), $\mathrm{forb}(n, H) = 2^{O(\mathrm{ex}(n,H))}$ holds. Until recently, it was believed that in fact $\mathrm{forb}(n, H) = 2^{(1+o(1))\mathrm{ex}(n,H)}$ should hold for all bipartite graphs that contain a cycle just as it does for non-bipartite H. But this was disproved by Morris and Saxton [32], who showed that $\mathrm{forb}(n, C_6) \ge 2^{(1+c)\mathrm{ex}(n,C_6)}$ for some positive c. Thus, unlike for non-bipartite graphs, for bipartite H, the trivial lower bound of $2^{\mathrm{ex}(n,H)}$ is not always asymptotically tight.

One may consider the natural extension of the problem to the setting of uniform hypergraphs. For an integer $k \ge 2$, a *k-uniform hypergraph* (or k-graph) is pair (V, E) of finite sets, where the *edge set* E is a family of k-element subsets of the *vertex set* V. For a fixed k-graph $\mathcal{H}$, one defines the extremal number $\mathrm{ex}(n, \mathcal{H})$ and $\mathrm{forb}(n, \mathcal{H})$ analogously as in the graph setting. As for graphs, for any k-graph $\mathcal{H}$, we trivially have

$$2^{\mathrm{ex}(n,\mathcal{H})} \le \mathrm{forb}(n, \mathcal{H}) \le \sum_{i \le \mathrm{ex}(n,\mathcal{H})} \binom{\binom{n}{k}}{i} = n^{O(\mathrm{ex}(n,\mathcal{H}))}. \tag{3}$$

A k-graph $\mathcal{H}$ is *k-partite* (or *degenerate*) if its vertex can be partitioned into k parts $X_1, \dots, X_k$ so that each edge contains exactly one vertex in each part. We call $(X_1, \dots, X_k)$ a *k-partition* of $\mathcal{H}$. It is easy to see that when $\mathcal{H}$ is not k-partite, $\mathrm{ex}(n, \mathcal{H}) = \Theta(n^k)$. When $\mathcal{H}$ is k-partite, it follows from a result of Erdős [8] that $\mathrm{ex}(n, \mathcal{H}) = O(n^{k-c})$, for some constant $c > 0$. Extending the work of Erdős, Frankl and Rödl [10],

https://doi.org/10.1017/fms.2025.10163 Published online by Cambridge University Press

Nagle, Rödl, and Schacht [37] showed that for any fixed k-graph $\mathcal{H}$, $\mathrm{forb}(n,\mathcal{H}) \le 2^{\mathrm{ex}(n,\mathcal{H})+o(n^k)}$. Hence, for any k-graph $\mathcal{H}$ that is not k-partite, we have

$$\mathrm{forb}(n,\mathcal{H}) = 2^{(1+o(1))\mathrm{ex}(n,\mathcal{H})}. \tag{4}$$

When $\mathcal{H}$ is k-partite, however, the bound $\mathrm{forb}(n,\mathcal{H}) \le 2^{\mathrm{ex}(n,\mathcal{H})+o(n^k)}$ becomes too weak. To improve this trivial upper bound for k-partite k-graphs, a natural approach is to first study the problem for some prototypical examples of k-partite k-graphs. It follows from the work of Balogh et al [1] on the typical structures of t-intersecting families that $\mathrm{forb}(n,\mathcal{F}_{k,t}) = 2^{(1+o(1))\mathrm{ex}(n,\mathcal{F}_{k,t})}$, where $\mathcal{F}_{k,t}$ is the family of k-graphs each of which consists of two edges sharing at least t vertices in common. (Their actual estimate is in fact even sharper than stated, see their Theorem 1.4). Mubayi and Wang [34] investigated $\mathrm{forb}(n,C_\ell^{(k)})$ for $k \ge 3$, where $C_\ell^{(k)}$ is the so-called *k-uniform linear cycle* of length ℓ which is the k-graph obtained from a graph ℓ-cycle by expanding each edge with $k-2$ degree 1 vertices. It follows from the work of Füredi and Jiang [21] and of Kostochka, Mubayi and Verstraëte [29] that for all $k,\ell \ge 3, \mathrm{ex}(n,C_\ell^{(k)}) \sim \lfloor \frac{\ell-1}{2} \rfloor \binom{n}{k-1}$. Mubayi and Wang [34] showed that $\mathrm{forb}(n,C_\ell^{(3)}) = 2^{O(n^{k-1})}$ for all even $\ell \ge 4$ and further conjectured that $\mathrm{forb}(n,C_\ell^{(k)}) = 2^{O(n^{k-1})}$ for all $k,\ell \ge 3$. Their conjecture was subsequently settled by Balogh, Narayanan and Skokan [3], who then posed the natural question of whether $\mathrm{forb}(n,C_\ell^{(k)}) = 2^{(1+o(1))\mathrm{ex}(n,C_\ell^{(k)})}$ holds for all $k,\ell \ge 3$. More recently, in the same paper mentioned earlier, Ferber, McKinley, and Samotij [14] established the very general result that for all k-partite k-graphs H that satisfy a mild condition (see their Theorem 9)

$$\mathrm{forb}(n,\mathcal{H}) \le 2^{O(\mathrm{ex}(n,\mathcal{H}))}, \tag{5}$$

thus significantly sharpening the trivial upper bound in (3), while also retrieving the results of Mubayi and Wang and of Balogh, Narayan and Skokan on $\mathrm{ex}(n,C_\ell^{(k)})$. In spite of this remarkable progress, it remains an intriguing question whether the bounds can be further sharpened, and in particular, in view of (4) and the question of Balogh, Narayanan and Skokan [3] whether there exists some nontrivial infinite family of k-partite k-graphs for which $\mathrm{forb}(n,\mathcal{H}) = 2^{1+o(1))\mathrm{ex}(n,\mathcal{H})}$ holds.

In this paper, we establish a large and the first known family of degenerate k-graphs, for which $\mathrm{forb}(n,\mathcal{H}) = 2^{(1+o(1))\mathrm{ex}(n,\mathcal{H})}$ holds. As an immediate consequence of our main result, we settled the question of Balogh, Narayanan and Skokan [3] in the affirmative for all $k \ge 5, \ell \ge 3$.

Theorem 1.1. *For all integers, $k \ge 5, \ell \ge 3$, we have*

$$\mathrm{forb}(n,C_\ell^{(k)}) = 2^{(1+o(1))\lfloor \frac{\ell-1}{2} \rfloor \binom{n}{k-1}}.$$

To state our main results, we need a few definitions and history, which we detail next.

1.2. Main Result

First, we define hypertrees recursively as follows. A single edge E is a hypertree. In general, a hypergraph $\mathcal{H}$ with at least two edges is a hypertree if there exists an edge E such that $\mathcal{H}' := \mathcal{H} \setminus E$ is a hypertree and there exists an edge F in $\mathcal{H}'$ such that $E \cap V(\mathcal{H}') = E \cap F$; we call such an edge E a *leaf edge* of $\mathcal{H}$ and call F a *parent edge* of E in $\mathcal{H}'$. It follows from the definition above that if $\mathcal{H}$ is a hypertree with $m \ge 2$ edges, there exists an ordering of its edges as $E_1,\dots,E_m$ such that for each $i = 2,\dots,m$, $\mathcal{H}_i := \{E_1,\dots,E_i\}$ is a tree and E_i is a leaf edge of $\mathcal{H}_i$. We call such an edge-ordering a *tree-defining* ordering for $\mathcal{H}$. If k is a positive integer, we will call a k-uniform hypertree a *k-tree*. As a simple example, a k-uniform matching is a k-tree. It is easy to show by induction that every k-tree is k-partite. Given a positive integer $t \le k-1$, we say that a k-graph $\mathcal{H}$ is *t-contractible* if each edge of $\mathcal{H}$ contains t vertices of degree 1 and a *t-contraction* of $\mathcal{H}$ is the $(k-t)$-uniform multi-hypergraph obtained by deleting t degree 1 vertices from each edge of $\mathcal{H}$. The notion of t-contractibility is more general than that of the

https://doi.org/10.1017/fms.2025.10163 Published online by Cambridge University Press

k-uniform expansion of a graph. The latter has been quite extensively studied, as detailed in the survey by Mubayi and Verstraëte [33]. Specifically, the *k-expansion* of a graph G is the k-graph $G^{(k)}$ obtained from G by expanding each edge into a k-set by adding $k-2$ degree 1 vertices. Thus, for instance, the k-uniform linear cycle $C_\ell^{(k)}$ is the k-expansion of a graph ℓ-cycle. Hence, the k-expansion of a graph G is a $(k-2)$-contractible k-graph whose $(k-2)$-contraction is a simple hypergraph with no repeated edges, whereas for a general $(k-2)$-contractible k-graph $\mathcal{H}$, its $(k-2)$-contraction is allowed to be a multigraph.

Given a hypergraph $\mathcal{H}$, a set $S \subseteq V(\mathcal{H})$ is called a *cross-cut* of $\mathcal{H}$ if each edge of $\mathcal{H}$ intersects S in exactly one vertex. If $\mathcal{H}$ has a cross-cut then we denote the minimum size of a cross-cut of it by $\sigma(\mathcal{H})$, and call it the *cross-cut number* of $\mathcal{H}$. Note that every k-partite k-graph $\mathcal{H}$ has a cross-cut, for instance by taking any part in a k-partition of $\mathcal{H}$. Generalizing a long line of work [20, 21, 23, 29, 30] for k-graphs with $k \geq 5$, Füredi and Jiang [22] established the following sharp results on the extremal number of any subgraph of a 2-contractible k-tree for $k \geq 5$.

Theorem 1.2 [22]. *Let $k \geq 4$ be an integer. For any k-graph $\mathcal{H}$ that is a subgraph of a 2-contractible k-tree, we have*

$$\mathrm{ex}(n, \mathcal{H}) = (\sigma(\mathcal{H}) - 1 + o(1))\binom{n}{k-1}.$$

In particular, Theorem 1.2 implies that for $k \geq 5$ and $\ell \geq 3$, $\mathrm{ex}(n, C_\ell^{(k)}) = \lfloor \frac{\ell-1}{2} \rfloor \binom{n}{k-1}$. Indeed, for $k \geq 5$, it is easy to see that $C_\ell^{(k)}$ is a subgraph of a k-tree and $\sigma(C_\ell^{(k)}) = \lfloor \frac{\ell+1}{2} \rfloor$ (see [21] for instance for details). The authors of [22] also demonstrated existence of 1-contractible hypertrees for which the conclusion is no longer valid. The main method used in [22] is the so-called *Delta system method*, which is a powerful tool for studying extremal problems on set systems.

In this paper, we develop a supersaturation variant of the Delta system method and use it in conjunction with the container method to establish the following sharp enumeration result.

Theorem 1.3 (Main Theorem). *Let $k \geq 4$ be an integer. For every 2-contractible k-tree $\mathcal{H}$, we have*

$$\mathrm{forb}(n, \mathcal{H}) = 2^{(\sigma(\mathcal{H})-1+o(1))\binom{n}{k-1}}.$$

This gives the first known family of degenerate k-graphs $\mathcal{H}$ for which $\mathrm{forb}(n, \mathcal{H}) = 2^{(1+o(1))\mathrm{ex}(n,\mathcal{H})}$ holds.

Furthermore, as a by-product of the supersaturation variant of the Delta-system method, we also get an optimal supersaturation result for the family of 2-contractible k-trees for $k \geq 4$, which may be of independent interest (see Theorem 4.4).

As an immediate corollary of Theorem 1.3, we obtain Theorem 1.1, which answers the question of Balogh, Narayanan and Skokan [3] affirmatively for all $k \geq 5, \ell \geq 3$. Indeed, as mentioned earlier, $\sigma(C_\ell^{(k)}) = \lfloor \frac{\ell+1}{2} \rfloor$, and by considering ℓ even and ℓ odd cases separately it is not hard to construct a k-tree $T_\ell^{(k)}$ with cross-cut number $\lfloor \frac{\ell+1}{2} \rfloor$ that contains $C_\ell^{(k)}$. Thus, by Theorem 1.3, $\mathrm{forb}(n, C_\ell^{(k)}) \leq \mathrm{forb}(n, T_\ell^{(k)}) \leq 2^{(\lfloor \frac{\ell-1}{2} \rfloor + o(1))\binom{n}{k-1}}$, for all $k \geq 5$ and $\ell \geq 3$, from which Theorem 1.1 follows.

1.3. Applications to the random Turán problem

The methods used in establishing our main theorem also readily yield some sharp results on the so-called *random Turán problem*. Let $k \geq 2$ be an integer. We let $G_{n,p}^{(k)}$ be the so-called *Erdős-Rényi random graph* formed by keeping each edge of $K_n^{(k)}$ uniformly at random with probability p. The random Turán problem is to study the random variable $\mathrm{ex}(G_{n,p}^{(k)}, H)$ that counts the maximum number of edges in an H-free subgraph of $G_{n,p}^{(k)}$. For a thorough introduction to the random Turán problem, the reader is referred to the excellent survey by Rödl and Schacht [41], though we will offer a brief summary here.

https://doi.org/10.1017/fms.2025.10163 Published online by Cambridge University Press

The random Turán problem for non-k-partite k-graphs was essentially solved in breakthrough works by Conlon and Gowers [6] and by Schacht [43], who showed that when $p \gg n^{-1/m_k(H)}$, almost surely $\mathrm{ex}(n, G(n,p)^{(k)}) = p(\mathrm{ex}(n,H) + o(n^k))$, as $n \to \infty$, where $m_k(H) = \max_{F \subseteq H, e(F) \geq 2} \frac{e(F)-1}{v(F)-k}$. The theorem was then reproved using the container method by Balogh, Morris and Samotij [2] and independently by Saxton and Thomason [42].

For k-partite k-graphs (which are also called *degenerate k-graphs*) much less is known. For degenerate graphs, early results due to Haxell, Kohayakawa, and Łuczak [25] and Kohayakawa, Kreuter, and Steger [28] essentially solved the problem for even graph cycles for small values of p. For dense ranges, Morris and Saxton [32] solved the problem for cycles and complete bipartite graphs, and McKinley and Spiro [36] extended their result for theta graphs. For dense p, some general upper bounds are obtained in [26].

Even less is known for degenerate hypergraphs. For k-uniform even linear cycles, Mubayi and Yepremyan [35] and independently Nie [39] proved

Theorem 1.4 [35, 39]. *For every $\ell \geq 2$ and $k \geq 4$ with high probability, the following holds:*

$$\mathrm{ex}\left(G_{n,p}^{(k)}, C_{2\ell}^{(k)}\right) = \begin{cases} \Theta(pn^{k-1}), & \text{if } p \geq n^{-(k-2)+\frac{1}{2\ell-1}+o(1)} \\ n^{1+\frac{1}{2\ell-1}+o(1)}, & \text{if } n^{-(k-1)+\frac{1}{2\ell-1}+o(1)} \leq p \leq n^{-(k-2)+\frac{1}{2\ell-1}+o(1)} \\ (1-o(1))pn^k, & \text{if } n^{-k} \ll p \ll n^{-(k-1)+\frac{1}{2\ell-1}}. \end{cases}$$

For other classes of hypergraphs, near optimal results were obtained by Nie [38] for k-expansions of subgraphs of tight p-trees and of K_p^{p-1}, where $k \geq p \geq 3$, which includes both odd and even linear cycles. Nie and Spiro [40] also were able to get near optimal bounds for expansions of theta-graphs. Furthermore, for any k-uniform hypergraph H, which satisfies that for every k-uniform G there are at least $e(G)^{e(H)}v(G)^{v(H)-ke(H)}$ many homomorphisms from $H \to G$, they proved that there is some $r_0 \geq k$ such that for every $r \geq r_0$, tight bounds on the random Turán number of r-expansions of H hold.

As an immediate byproduct of our main result, for dense p and for $C_\ell^{(k)}$ with $k \geq 5$, we are able to sharpen the bound $pn^{k-1+o(1)}$ to the correct bound $p(\lfloor \frac{\ell-1}{2} \rfloor + o(1))\binom{n}{k-1} = p(1+o(1))\mathrm{ex}(n, C_\ell^{(k)})$, albeit for a slightly more restricted range of p. This comes as an immediate corollary to the following more general theorem. A hypergraph $\mathcal{H}$ is ℓ-overlapping if for any two edges E, F in $\mathcal{H}$, $|E \cap F| \leq \ell$.

Theorem 1.5. *Let $k \geq 4$ be an integer. Let $\mathcal{H}$ be a 2-contractible ℓ-overlapping k-tree. When $p \gg \frac{\log(n)^2}{n^{k-\ell-1}}$, with high probability*

$$\mathrm{ex}(G(n,p)^{(k)}, \mathcal{H}) = (\sigma(\mathcal{H}) - 1 + o(1))p\binom{n}{k-1} = p(1+o(1))\mathrm{ex}(n, \mathcal{H}).$$

As mentioned earlier, for all $k \geq 5, \ell \geq 3$, there exists a 2-contractible k-tree $T_\ell^{(k)}$ containing $C_\ell^{(k)}$ with $\sigma(T_\ell^{(k)}) = \sigma(C_\ell^{(k)}) = \lfloor \frac{\ell+1}{2} \rfloor$. Hence, Theorem 1.5 immediately implies

Corollary 1.6. *For all $k \geq 5, \ell \geq 3$ and $p \gg \log(n)^2 n^{-(k-3)}$, with high probability,*

$$\mathrm{ex}(G(n,p)^{(k)}, C_\ell^{(k)}) = p\left(\left\lfloor \frac{\ell-1}{2} \right\rfloor + o(1)\right)\binom{n}{k-1} = p(1+o(1))\mathrm{ex}(n, C_\ell^{(k)}).$$

1.4. Overiew of methodology and organization of the paper

At the heart of our work is the so-called optimal balanced supersaturation at the Turán threshold. Several long-standing conjectures of Erdős and Simonovits [12] addressed the question of how many copies of a graph H we can guarantee in a dense enough host graph G. Loosely speaking, the conjectures say that the number of copies of H we expect in G should be at least on the same order of magnitude as the number of copies of H we expect in a random graph with the same edge-density as G. These are referred to as supersaturation conjectures. The strongest of the Erdős-Simonovits supersaturation conjectures says that the conjectured bound on the number of copies of H should already hold as soon as an n-vertex graph G has just barely asymptotically a bit more edges than what is enough to guarantee a single copy of H, namely when $e(G) \geq (1+\varepsilon)\mathrm{ex}(n, H)$ for any small real $\varepsilon > 0$. We will refer to this phenomenon as *optimal supersaturation at the Turán threshold*. Establishing optimal supersaturation at the Turán threshold turns out to be a very difficult task for degenerate (i.e., bipartite) graphs, with the problem being unsolved except for very few bipartite graphs.

In the last decade or so, the development of the container method has brought enhanced importance to the supersaturation problem. Specifically, the container method allows one to obtain tight enumeration results for $\mathrm{forb}(n, H)$ once one is able to obtain supersaturation of H with the additional feature that the copies of H found are evenly distributed in a sense. While this paved the way for several breakthroughs mentioned earlier, naturally developing optimal supersaturation at the Turán threshold with the added balanced feature is an even more difficult task than the one without the additional balanced requirement, as witnessed by the fact this has not been done even for the 4-cycle, whose extremal number is very well-understood [19] and for whom supersaturation at the Turán threshold has been achieved [13].

Our main results crucially build on the optimal balanced supersaturation at the Turán threshold for 2-contractible hypertrees. Given the difficulty with the optimal balanced supersaturation at the Turán threshold for graphs, it does come as a surprise that one is able to establish it for a large family of degenerate hypergraphs. This is in a strong sense attributed to the power of the method we use, known as the Delta system method for set-systems. However, while the Delta system method has been successfully used on Turán type problem for hypergraphs, it has not been tailored for supersaturation problems before. In that regard, the most important innovative aspect of our work is the development of a supersaturation variant of the Delta system method and applying it successfully with the container method to get tight enumeration results. We believe that this variant of the Delta system method will find future applications.

We organize the rest of the paper as follows. In Section 2, we give some notation. In Section 3, we develop a supersaturation variant of the Delta system method and develop a structural dichotomy for all k-graphs with $\Theta(n^{k-1})$ edges, both of which may be of independent interest. In Section 4, we develop optimal supersaturation at the Turán threshold for 2-contractible hypertrees. In Section 5, we develop optimal balanced supersaturation at the Turán threshold for 2-contractible hypertrees. In Section 6, we prove our main theorem, Theorem 1.3, as well as Theorem 1.5.

2. Notation

Let $\mathcal{F}$ be a hypergraph on $V = V(\mathcal{F})$. For each integer $i \geq 0$, we define the *i-shadow* of $\mathcal{F}$ to be

$$\partial_i(\mathcal{F}) := \{D : |D| = i, \exists F \in \mathcal{F}, D \subseteq F\}.$$

The Lovász' [31] version of the Kruskal-Katona theorem states that if $\mathcal{F}$ is a k-graph of size $|\mathcal{F}| = \binom{x}{k}$, where $x \geq k-1$ is a real number, then for all i with $1 \leq i \leq k-1$ one has

$$|\partial_i(\mathcal{F})| \geq \binom{x}{i}. \tag{6}$$

Given $D \subseteq V(\mathcal{F})$, we define the *link of D* in $\mathcal{F}$ to be

$$\mathcal{L}_{\mathcal{F}}(D) = \{F \setminus D : F \in \mathcal{F}, D \subseteq F\}.$$

Note that we allow $\emptyset$ to be a member of $\mathcal{L}_{\mathcal{F}}(D)$. We define the *degree of D* in $\mathcal{F}$ to be $d_{\mathcal{F}}(D) := |\mathcal{L}_{\mathcal{F}}(D)|$.

Given a k-graph $\mathcal{F}$ where $k \geq 2$ is an integer and integer i with $1 \leq i \leq k-1$, let

$$\Delta_i(\mathcal{F}) = \max\{d_{\mathcal{F}}(D) : D \in \partial_i(\mathcal{F})\} \text{ and } \delta_i(\mathcal{F}) := \min\{d_{\mathcal{F}}(D) : D \in \partial_i(\mathcal{F})\}.$$

https://doi.org/10.1017/fms.2025.10163 Published online by Cambridge University Press

We call $\delta_i(\mathcal{F})$ the *proper minimum i-degree* of $\mathcal{F}$. By definition, every i-set in $\mathcal{F}$ either has degree 0 or has degree at least $\delta_i(\mathcal{F})$. Let $\mathcal{F}$ be a k-graph and $S \subseteq V(\mathcal{F})$, we let $\mathcal{F} - S := \{F \setminus S : F \in \mathcal{F}\}$

3. A variant of the Delta system method and a structural dichotomy for k graphs of size $\Theta(n^{k-1})$

The *delta system method*, originated by Deza, Erdős and Frankl [7] and others, is a powerful tool for solving extremal set problems. A particularly versatile tool within the method, which one may call *the intersection semilattice lemma* was developed by Füredi [18] (Theorem 1'). The delta system method, particularly aided by the semilattice lemma, has been very successfully used to obtain a series of sharp results on extremal set problems (see for instance [16, 17, 18, 20, 21, 23]). However, despite its effectiveness in determining the threshold on the size of a hosting hypergraph beyond which a certain subgraph occurs, it does not readily allow us to effectively count the number of such subgraphs (known as the supersaturation problem). In this section, we develop a variant of Füredi's intersection semilattice lemma, Theorem 3.1 below, to also address hypergraph supersaturation.

Let $k \geq 2$ be an integer. Let $\mathcal{F}$ be a k-partite k-graph $\mathcal{F}$ with a fixed k-partition $(X_1, \dots, X_k)$. For each $J \subseteq [k]$ and $F \in \mathcal{F}$, we define the *J-projection* of F, denoted by F_J to be

$$F_J := F \cap (\bigcup_{i \in J} X_i).$$

Conversely, for any $D \in \bigcup_{i=1}^k \partial_i(\mathcal{F})$, we define the *pattern* of D, denoted by $\pi(D)$, to be

$$\pi(D) := \{i \in [k] : D \cap X_i \neq \emptyset\}.$$

Note that since $\mathcal{F}$ is k-partite, $|\pi(D)| = |D|$ for each $D \in \bigcup_{i=1}^k \partial_i(\mathcal{F})$.

Given a positive integer $s \geq 2$, a nonempty hypergraph $\mathcal{H}$ is called *s-diverse* if

$$\forall v \in V(\mathcal{H}), d_{\mathcal{H}}(v) < (1/s)|\mathcal{H}|.$$

Note that implicitly an s-diverse hypergraph necessarily contains more than s edges.

Let

$$\mathrm{Int}(\mathcal{F}) = \{\pi(E \cap F) : E, F \in \mathcal{F}, E \neq F\}.$$

Theorem 3.1 (Super-homogeneous Subfamily Lemma). *Let $s, k \geq 2$ be integers where $s \geq 2k$. Let $c(k,s) = \frac{k!}{k^k(2s(1+2^k))^{2^k}}$. Let $\mathcal{F}$ be a k-graph on $[n]$. Then there exists a k-partite subgraph $\mathcal{F}' \subseteq \mathcal{F}$ with some k-partition $(X_1, \dots, X_k)$ such that the following holds:*

1. $|\mathcal{F}'| \geq c(k,s)|\mathcal{F}|$.
2. *For every $F \in \mathcal{F}'$ and $J \in \mathrm{Int}(\mathcal{F}')$, $\mathcal{L}_{\mathcal{F}'}(F_J)$ is s-diverse.*
3. *For every $F \in \mathcal{F}'$ and $J \in \mathrm{Int}(\mathcal{F}')$, $d_{\mathcal{F}'}(F_J) \geq \max\{s, \frac{1}{2k}\frac{|\mathcal{F}'|}{n^{|J|}}\}$.*

We call such a k-graph $\mathcal{F}'$ s-super-homogeneous.

Proof. By a well-known result of Erdős and Kleitman [9], $\mathcal{F}$ contains a k-partite subgraph $\mathcal{F}_0$ with $|\mathcal{F}_0| \geq \frac{k!}{k^k}|\mathcal{F}|$. Let $(X_1, X_2, \dots X_k)$ be a fixed k-partition of $\mathcal{F}_0$. If $\mathcal{F}_0$ satisfies conditions 2 and 3, then the theorem holds with $\mathcal{F}' = \mathcal{F}_0$. Otherwise, let $\mathcal{G}_0 = \mathcal{F}_0$. We perform a so-called *filtering process* on $\mathcal{G}_0$ as follows. Let $\mathcal{W}_0 = \emptyset$. For each $J \in \mathrm{Int}(\mathcal{G}_0)$, let $\mathcal{A}_J = \emptyset$. We iteratively modify $\mathcal{G}_0$, $\mathcal{W}_0$ and the $\mathcal{A}_J$'s for $J \in \mathrm{Int}(\mathcal{G}_0)$ as follows. Whenever there is an edge $F \in \mathcal{G}_0$ and a $J \in \mathrm{Int}(\mathcal{G}_0)$ such that $\mathcal{L}_{\mathcal{G}_0}(F_J)$ is nonempty and not s-diverse, we remove all the edges of $\mathcal{G}_0$ containing F_J and add them to $\mathcal{A}_J$. Whenever there is an edge $F \in \mathcal{G}_0$ and a $J \in \mathrm{Int}(G_0)$ such that $\mathcal{L}_{\mathcal{G}_0}(F_j)$ is s-diverse but $d_{\mathcal{G}_0}(F_J) < \frac{1}{2k}\frac{|\mathcal{F}_0|}{n^{|J|}}$, we remove all the edges of $\mathcal{G}_0$ containing F_J and add them to $\mathcal{W}_0$. Let $\mathcal{G}_0^*$ denote the final $\mathcal{G}_0$ at the end of the filtering process. By definition, if $\mathcal{G}_0^*$ is nonempty, then $\mathcal{G}_0^*$ satisfies conditions 2 and 3.

Note that $|\mathcal{W}_0| \le |\mathcal{F}_0|/2$. This is because for any fixed $j = 1, \ldots, k-1$, there are at most $\binom{n}{j}$ different F_J's. When all edges containing some F_J are moved to $\mathcal{W}_0$, by definition, fewer than $\frac{1}{2k}\frac{|\mathcal{F}_0|}{n^{|J|}}$ edges are moved. Hence $|\mathcal{F}_0 \setminus \mathcal{W}_0| \ge |\mathcal{F}_0|/2$. If $|\mathcal{G}_0^*| \ge \frac{1}{1+2^k}|\mathcal{F}_0 \setminus \mathcal{W}_0|$, we let $\mathcal{F}' = \mathcal{G}_0^*$. Otherwise, by the pigeonhole principle, there exists some $J \in \mathrm{Int}(\mathcal{G}_0)$ such that $|\mathcal{A}_J| \ge \frac{1}{1+2^k}|\mathcal{F}_0 \setminus \mathcal{W}_0|$. Note that edges were added to $\mathcal{A}_J$ in batches, with each batch consisting of edges F with the same J-projection and different batches have different J-projections. Consider any batch $\mathcal{B}$ added to $\mathcal{A}_J$. Let D denote the common J-projection of the edges in $\mathcal{B}$. By definition, $\mathcal{L}_{\mathcal{B}}(D) = \mathcal{L}_{\mathcal{G}_0}(F_J)$ is nonempty and not s-diverse at the moment $\mathcal{B}$ was added to $\mathcal{A}_J$. By definition, there exists a vertex v in $\mathcal{L}_{\mathcal{B}}(D)$ that lies in at least $(1/s)|\mathcal{L}_{\mathcal{B}}(D)|$ of the edges. Let $\mathcal{B}'$ denote the subset of edges in $\mathcal{B}$ that also contain v. We now remove $\mathcal{B}$ from $\mathcal{A}_J$ and replace it with $\mathcal{B}'$. We do this for each batch of edges that were added to $\mathcal{A}_J$, and denote the resulting subgraph of $\mathcal{A}_J$ by $\mathcal{A}'_J$. Then $|\mathcal{A}'_J| \ge (1/s)|\mathcal{A}_J|$. Furthermore, it is easy to see that $J \notin \mathrm{Int}(\mathcal{A}'_J)$. We let $\mathcal{F}_1 = \mathcal{A}'_J$. Then

$$|\mathcal{F}_1| \ge \frac{1}{s(1+2^k)}|\mathcal{F}_0 \setminus \mathcal{W}_0| \ge \frac{1}{2s(1+2^k)}|\mathcal{F}_0| \quad \text{and} \quad |\mathrm{Int}(\mathcal{F}_1)| \le |\mathrm{Int}(\mathcal{F}_0)| - 1.$$

Now, let $\mathcal{G}_1 = \mathcal{F}_1$, let $\mathcal{W}_1 = \emptyset$ and set $\mathcal{A}_J = \emptyset$ for all $J \in \mathrm{Int}(\mathcal{G}_1)$. We then perform the same filtering on $\mathcal{G}_1$ to iteratively modify $\mathcal{G}_1$, $\mathcal{W}_1$ and the $\mathcal{A}_J$'s for $J \in \mathrm{Int}(\mathcal{G}_1)$. Let $\mathcal{G}_1^*$ denote the final $\mathcal{G}_1$ at the end of the filtering process. By definition, $\mathcal{G}_1^*$ satisfies conditions 2 and 3. If $|\mathcal{G}_1^*| \ge \frac{1}{1+2^k}|\mathcal{F}_1 \setminus \mathcal{W}_1|$, we let $\mathcal{F}' = \mathcal{G}_1^*$. Otherwise, as before, there exists some $J \in \mathrm{Int}(\mathcal{G}_1)$ and a subgraph $\mathcal{A}'_J \subseteq \mathcal{A}_J$ with $|\mathcal{A}'_J| \ge \frac{1}{2s(1+2^k)}|\mathcal{F}_1|$ such that $|\mathrm{Int}(\mathcal{A}'_J)| \le |\mathrm{Int}(\mathcal{G}_1)| - 1$. We let $\mathcal{F}_2 = \mathcal{A}'_J$.

We continue like this, obtaining a sequence $\mathcal{F}_0, \mathcal{F}_1, \ldots$. Since $|\mathrm{Int}(\mathcal{F}_i)|$ strictly decreases with i, the sequence must end with $\mathcal{F}_m$ for some $m \le 2^k - 1$. Since $\mathcal{F}_{m+1}$ is undefined, this must mean that

$$|\mathcal{G}_m^*| \ge \frac{1}{1+2^k}|\mathcal{F}_m \setminus \mathcal{W}_m| \ge \frac{1}{2(1+2^k)}|\mathcal{F}_m| \ge \frac{1}{[2s(1+2^k)]^{2^k}}|\mathcal{F}_0| \ge c(k,s)|\mathcal{F}|,$$

where $c(k,s) = \frac{k!}{k^k} \cdot \frac{1}{[2s(1+2^k)]^{2^k}}$. Let $\mathcal{F}' = \mathcal{G}_m^*$. Then $|\mathcal{F}'| \ge c(k,s)|\mathcal{F}|$ and $\mathcal{F}'$ also satisfies conditions 2 and 3, by the definition of $\mathcal{G}_m^*$. □

Throughout the rest of the paper, whenever we consider an s-super-homogeneous k-graph $\mathcal{F}$, we always implicitly fix a k-partition associated with $\mathrm{Int}(\mathcal{F})$.

Next, we collect some useful facts about s-super-homogeneous families, for which we need the following definition.

Definition 3.2. Let k be a positive integer. Given a family $\mathcal{J}$ of proper subsets of $[k]$, let

$$r(\mathcal{J}) := \min\{|D| : D \subseteq [k], \nexists J \in \mathcal{J} \text{ such that } D \subseteq J\}.$$

We call $r(\mathcal{J})$ the *rank* of $\mathcal{J}$.

Lemma 3.3. *Let $s, k \ge 2$ be integers with $s \ge 2k$. Let $\mathcal{F}$ be an s-super-homogeneous k-partite k-graph on $[n]$ with some fixed k-partition $(X_1, \ldots, X_k)$. Then, the following hold.*

1. *For each $J \subsetneq [k]$ where $J \notin \mathrm{Int}(\mathcal{F})$ and each $F \in \mathcal{F}$, there is no $F' \in \mathcal{F}$ satisfying $F \cap F' = F_J$.*
2. *For all $J, J' \in \mathrm{Int}(\mathcal{F})$, $J \cap J' \in \mathrm{Int}(\mathcal{F})$, that is, $\mathrm{Int}(\mathcal{F})$ is closed under intersection.*
3. *If $\mathrm{Int}(\mathcal{F})$ has rank m, then $|\mathcal{F}| \le \binom{n}{m}$.*

Proof. Statement 1 follows from definition of $\mathrm{Int}(\mathcal{F})$. For statement 2, let $J, J' \in \mathrm{Int}(\mathcal{F})$ with $J \ne J'$. Let $F \in \mathcal{F}$. By our assumption $\mathcal{L}_{\mathcal{F}}(F_J)$ is s-diverse. So the vertices in $F \setminus F_J$ block fewer than $k(1/s)|\mathcal{L}_{\mathcal{F}}(F_J)| \le |\mathcal{L}_{\mathcal{F}}(F_J)|$ of the edges in $\mathcal{L}_{\mathcal{F}}(F_J)$. So there exists $F' \in \mathcal{F}$ with $F \cap F' = F_J$. By a similar reasoning, since $2k(1/s)|\mathcal{L}_{\mathcal{F}}(F'_{J'})| \le |\mathcal{L}_{\mathcal{F}}(F'_{J'})|$ there exists $F'' \in \mathcal{F}$ containing $F'_{J'}$ that avoids vertices in $(F \setminus F') \cup (F' \setminus F'_{J'})$. Now, $\pi(F \cap F'') = J \cap J'$. Hence, $J \cap J' \in \mathrm{Int}(\mathcal{F})$.

For statement 3, suppose $\mathrm{Int}(\mathcal{F})$ has rank m. Then $\exists D \subseteq [k]$ with $|D| = m$ such that D is not contained in any member of $\mathrm{Int}(\mathcal{F})$. Consider any $F, F' \in \mathcal{F}$, where $F \neq G$. If $F[D] = F'[D]$, then $\pi(F \cap F')$ is a member of $\mathrm{Int}(\mathcal{F})$ that contains D, a contradiction. So, the D-projections of members of $\mathcal{F}$ are all distinct. This implies that $|\mathcal{F}| \leq \binom{n}{m}$. □

The following structural lemma strengthens Lemma 7.1 of [17] (see also Lemma 4.2 of [22]) and leads to a structural dichotomy theorem (Theorem 3.5) that is important for our main arguments.

Lemma 3.4. *Let $k \geq 3$ be an integer. Let $\mathcal{J} \subseteq 2^{[k]}$ be a family of proper subsets of $[k]$ that is closed under intersection. Suppose $\mathcal{J}$ has rank at least $k-1$. Then one of the following must hold.*

1. *There exists $B \subseteq [k]$, with $|B| = k-2$, such that $2^B \subseteq \mathcal{J}$.*
2. *There exists a unique $i \in [k]$ such that $\forall D \subsetneq [k]$ with $i \in D$ we have $D \in \mathcal{J}$ and for every $D \subseteq [k] \setminus \{i\}$ with $|D| \geq k-2$ we have $D \notin \mathcal{J}$. We call i the central index for $\mathcal{J}$.*

If condition 1 holds for $\mathcal{J}$, we say that $\mathcal{J}$ is of type 1. If condition 2 holds for $\mathcal{J}$, we say that $\mathcal{J}$ is of type 2.

Proof. If $r(\mathcal{J}) = k$ then $[k] \setminus \{j\} \in \mathcal{J}$ for each $j \in [k]$. Since $\mathcal{J}$ is closed under intersection, we see that $S \in \mathcal{J}$ for each proper subset S of $[k]$. Hence statement 1 clearly holds. Next, suppose $r(\mathcal{J}) = k-1$. Then some $(k-1)$-subset of $[k]$ is not in $\mathcal{J}$. Without loss of generality, suppose $[k] \setminus \{j\} \notin \mathcal{J}$ for $j = 1, \dots, t$ and $[k] \setminus \{j\} \in \mathcal{J}$ for $j = t+1, \dots, k$, for some $1 \leq t \leq k$.

First, suppose $t = 1$. Then $[k] \setminus \{1\} \notin \mathcal{J}$ and $[k] \setminus \{2\}, \dots, [k] \setminus \{k\} \in \mathcal{J}$. Since $\mathcal{J}$ is closed under intersection, every proper subset of $[k]$ that contains 1 is in $\mathcal{J}$. We already have $[k] \setminus \{1\} \notin \mathcal{J}$. Suppose there is a $(k-2)$-subset B of $[k] \setminus \{1\}$ that is in $\mathcal{J}$. Let S be any subset of B. By earlier discussion $S \cup \{1\} \in \mathcal{J}$. Since $\mathcal{J}$ is closed under intersection, we have $S = (S \cup \{1\}) \cap B \in \mathcal{J}$. Hence, $2^B \subseteq \mathcal{J}$. So statement 1 holds in this case. Hence, we may assume that no $(k-2)$-subset of $[k] \setminus \{1\}$ is in $\mathcal{J}$. Then statement 2 holds for $i = 1$. It is easy to see that if an i satisfies the requirements, it can only be 1.

Next, suppose $t \geq 2$. Observe that for any $1 \leq i < j \leq t$, we must have $[k] \setminus \{i, j\} \in \mathcal{J}$, as otherwise $[k] \setminus \{i, j\}$ is not contained in any member of $\mathcal{J}$, contradicting $r(\mathcal{J}) \geq k-1$. Also, by our assumption for every $t+1 \leq j \leq k$ we have $[k] \setminus \{j\} \in \mathcal{J}$. Since $\mathcal{J}$ is closed under intersection, we see that every subset of $[k] \setminus \{1, 2\}$ is in $\mathcal{J}$. So statement 1 holds. □

Now, we describe our structural theorem that gives a dichotomy of n-vertex k-graphs with $\Theta(n^{k-1})$ edges. We believe that it is also of independent interest.

Theorem 3.5. *Let a, ε be fixed positive reals with $\varepsilon < \min\{1, a^{k-2}\}$. Let k, s be fixed positive integers with $s \geq 2k$. Let n be sufficiently large in terms of a, ε, k, s. Let $\mathcal{F}$ be a k-graph on $[n]$ with $|\mathcal{F}| = a\binom{n}{k-1}$. Then one of the following two holds:*

1. *$\mathcal{F}$ contains an s-super-homogeneous k-partite subgraph $\mathcal{F}'$ such that $|\mathcal{F}'| \geq \frac{1}{2}c(k,s)\varepsilon|\mathcal{F}|$ and $\mathrm{Int}(\mathcal{F}')$ is of type 1, where $c(k,s)$ is the constant as in Theorem 3.1.*
2. *There exist a $W \subseteq [n]$ and a subgraph $\mathcal{F}'' \subseteq \mathcal{F}$ satisfying the following:*
   (a) $|W| \leq \left(\frac{10}{c(k,s)\varepsilon^2}\right)^{k-1}$.
   (b) $|\mathcal{F}''| \geq (1-\varepsilon)|\mathcal{F}|$.
   (c) *Every $F \in \mathcal{F}''$ satisfies $|W \cap F| = 1$.*

Proof. First we apply Lemma 3.1 to $\mathcal{F}$ to get an s-super-homogeneous subgraph $\mathcal{F}_1$ of size at least $c(k,s)|\mathcal{F}|$. If $|\mathcal{F} \backslash \mathcal{F}_1| \leq \frac{1}{2}\varepsilon|\mathcal{F}|$, then we stop. Otherwise, we apply the lemma to $\mathcal{F} \backslash \mathcal{F}_1$ to find an s-super-homogeneous subgraph $\mathcal{F}_2$ of size at least $c(k,s)|\mathcal{F} \setminus \mathcal{F}_1|$. In general, as long as $|\mathcal{F} \setminus (\mathcal{F}_1 \cup \cdots \cup \mathcal{F}_i)| > \frac{1}{2}\varepsilon|\mathcal{F}|$, we let $\mathcal{F}_{i+1}$ be an s-super-homogeneous subgraph of $\mathcal{F} \setminus (\mathcal{F}_1 \cup \cdots \cup \mathcal{F}_i)$ of size at least $c(k,s)|\mathcal{F} \setminus (\mathcal{F}_1 \cup \cdots \cup \mathcal{F}_i)|$, which exists by Lemma 3.1. Suppose the sequence of subgraphs we constructed are $\mathcal{F}_1, \dots, \mathcal{F}_m$.

By our algorithm, for each $i \in [m]$, $|\mathcal{F}_i| \geq \frac{1}{2}c(k,s)\varepsilon|\mathcal{F}|$. Note that this implies that $m \leq \frac{2}{c(k,s)\varepsilon}$. For each $i \in [m]$, by Lemma 3.3 part 3, $r(\mathrm{Int}(\mathcal{F}_i)) \geq k-1$ and by Lemma 3.3 part 2 and Lemma 3.4, $\mathrm{Int}(\mathcal{F}_i)$ is either of type 1 or is of type 2. If for some $i \in [m]$, $\mathrm{Int}(\mathcal{F}_i)$ is of type 1, then we let $\mathcal{F}' = \mathcal{F}_i$

https://doi.org/10.1017/fms.2025.10163 Published online by Cambridge University Press

for such an i and Statement 1 holds. Hence, we may assume that for each $i \in [m]$, $\mathrm{Int}(\mathcal{F}_i)$ is of type 2. Let $\mathcal{F}^* = \bigcup_{j=1}^m \mathcal{F}_j$. Then

$$|\mathcal{F}^*| \geq (1 - \frac{\varepsilon}{2})|\mathcal{F}|.$$

Set $h = \left\lfloor \left(\frac{5m}{\varepsilon}\right)^{k-1} \right\rfloor$. Since $m \leq \frac{2}{c(k,s)\varepsilon}$, we have

$$h \leq \left(\frac{10}{c(k,s)\varepsilon^2}\right)^{k-1}. \tag{7}$$

For each $j \in [m]$, let $i_j \in [k]$ denote the central index for $\mathrm{Int}(\mathcal{F}_j)$ and for each $F \in \mathcal{F}_j$, let $c(F) = F_{\{i_j\}}$. We call $c(F)$ *the central vertex* of F. We partition $\mathcal{F}^*$ according to $c(F)$. For each $i \in [n]$, let

$$\mathcal{A}_i = \{F \in \mathcal{F}^* : c(F) = i\} \text{ and } \mathcal{A}'_i = \{F \setminus \{i\} : F \in \mathcal{A}_i\}.$$

Clearly, for each $i \in [n]$ we have $|\mathcal{A}_i| = |\mathcal{A}'_i|$. For each $j \in [m]$, let

$$\mathcal{F}'_j = \{F \setminus \{c(F)\} : F \in \mathcal{F}_j\} = \{F_{[k]\setminus\{i_j\}} : F \in \mathcal{F}_j\}.$$

Then, in particular:

$$\bigcup_{i=1}^n \mathcal{A}'_i = \bigcup_{j=1}^m \mathcal{F}'_j.$$

The next claim is crucial to our argument.

Claim 3.6. For each $j \in [m]$, we have

$$|\partial_{k-2}(\mathcal{F}'_j)| = \sum_{i=1}^n |\partial_{k-2}(\mathcal{A}'_i \cap \mathcal{F}'_j)|.$$

Proof of Claim 3.6. Let $(X_1, \ldots, X_k)$ denote the k-partition associated with $\mathrm{Int}(\mathcal{F}_j)$. Recall that $i_j \in [k]$ denotes the central index for $\mathrm{Int}(\mathcal{F}_j)$. Without loss of generality, we may assume $i_j = 1$. To prove the claim, it suffices to show that for any $a, b \in X_1$ where $a \neq b$, we have $\partial_{k-2}(\mathcal{A}'_a \cap \mathcal{F}'_j) \cap \partial_{k-2}(\mathcal{A}'_b \cap \mathcal{F}'_j) = \emptyset$. Suppose for contradiction that there exists $S \in \partial_{k-2}(\mathcal{A}'_a \cap \mathcal{F}'_j) \cap \partial_{k-2}(\mathcal{A}'_b \cap \mathcal{F}'_j)$. Then there exist $F_a \in \mathcal{A}_a \cap \mathcal{F}_j, F_b \in \mathcal{A}_b \cap \mathcal{F}_j$ such that $F_a \cap F_b \supseteq S$. By definition, $\pi(F_a \cap F_b) \in \mathrm{Int}(\mathcal{F}_j)$. But $\pi(F_a \cap F_b)$ is a subset of $[k] \setminus \{1\}$ with size at least $k-2$. This contradicts $\mathrm{Int}(\mathcal{F}_j)$ being of type 2 (see Lemma 3.4 (2)). The claim immediately follows. □

Note that trivially $|\partial_{k-2}(\mathcal{F}'_j)| \leq \binom{n}{k-2}$. Now, we have

$$\sum_{i=1}^n |\partial_{k-2}(\mathcal{A}'_i)| \leq \sum_{i=1}^n \sum_{j=1}^m |\partial_{k-2}(\mathcal{A}'_i \cap \mathcal{F}'_j)| = \sum_{j=1}^m \sum_{i=1}^n |\partial_{k-2}(\mathcal{A}'_i \cap \mathcal{F}'_j)| = \sum_{j=1}^m |\partial_{k-2}(\mathcal{F}'_j)| \leq m\binom{n}{k-2} \tag{8}$$

For each $i \in [n]$, let x_i be the real such that $|\partial_{k-2}(\mathcal{A}'_i)| = \binom{x_i}{k-2}$, where without loss of generality, $x_1 \geq x_2 \cdots \geq x_n$. By (6), for each $i \in [n]$ we have $|\mathcal{A}'_i| \leq \binom{x_i}{k-1}$.

By (8), $\sum_{i=1}^{n}\binom{x_i}{k-2} \le m\binom{n}{k-2}$. Thus, $\binom{x_{h+1}}{k-2} \le \frac{m}{h+1}\binom{n}{k-2}$, which yields

$$x_{h+1} - k + 3 \le \left(\frac{m}{h+1}\right)^{1/(k-2)} n. \tag{9}$$

Thus, we have

$$\begin{aligned}\sum_{i>h}|\mathcal{A}_i| = \sum_{i>h}|\mathcal{A}'_i| &\le \sum_{i=h+1}^{n}\binom{x_i}{k-1}\\ &= \sum_{h+1}^{n}\frac{x_i-k+2}{k-1}\binom{x_i}{k-2}\\ &\le \frac{x_{h+1}-k+2}{k-1}\sum_{i=h+1}^{n}\binom{x_i}{k-2}\\ &\le \left(\frac{m}{h+1}\right)^{1/(k-2)}\frac{nm}{k-1}\binom{n}{k-2} \qquad \text{(by (8) and (9))}\end{aligned}$$

By our choice of h, $h+1 \ge (\frac{5m}{\varepsilon})^{k-1}$. So we have $\sum_{i>h}|\mathcal{A}_i| < \left(\frac{\varepsilon}{5}\right)^{1+\frac{1}{k-2}}\frac{n}{k-1}\binom{n}{k-2} \le \frac{\varepsilon}{4}a\binom{n}{k-1}$, for sufficiently large n and by our assumption $a \ge \varepsilon^{\frac{1}{k-2}}$. Let $W := [h]$. Let $\mathcal{G}_1 := \{F \in \mathcal{F}^* : c(F) \notin W\}$ and $\mathcal{G}_2 := \{F \in \mathcal{F}^* : |F \cap W| \ge 2\}$.

Then by the argument above, $|\mathcal{G}_1| = \sum_{i>h}|\mathcal{A}_i| \le \frac{\varepsilon}{4}a\binom{n}{k-1}$. Also, for n sufficiently large, we have $|\mathcal{G}_2| \le \binom{W}{2}\binom{n}{k-2} \le \frac{\varepsilon}{4}a\binom{n}{k-1}$. Let $\mathcal{F}'' := \mathcal{F}^* \setminus (\mathcal{G}_1 \cup \mathcal{G}_2)$. We can readily check that $\mathcal{F}''$ satisfy all of conditions 1,2,3. □

Let us close this section by noting that a structural dichotomy like Theorem 3.5 was used in earlier works such as [20, 21, 22]. However, in those papers, such a dichotomy was made possible only under the extra assumption that $\mathcal{F}$ is $\mathcal{H}$-free, where $\mathcal{H}$ is a 2-contractible tree. Here, Theorem 3.5 is applicable to any k-graph with $\Theta(n^{k-1})$ edges. Since we will aim to count number of copies of $\mathcal{H}$ in a hosting graph $\mathcal{F}$, the removal of the $\mathcal{H}$-free condition is essential for our overall arguments.

4. Optimal Supersaturation at the Turán threshold

In this section, we develop an asymptotically tight bound on the number of copies of a 2-contractible tree $\mathcal{H}$ in a host graph with size just beyond the Turán threshold for $\mathcal{H}$, which provides the most important foundation on which the rest of our work is built. It may also be of independent interest as a supersaturation result. We start with the following two folklore lemmas. Recall the definition of $\delta_i(\mathcal{F})$ from Section 2.

Lemma 4.1. *Let $k \ge 2$ be an integer. Let $\mathcal{F}$ be a k-graph. Then $\mathcal{F}$ contains a nonempty subgraph $\mathcal{F}'$ with at least $\frac{1}{2}|\mathcal{F}|$ edges such that for each $i = 1, \dots, k-1$, $\delta_i(\mathcal{F}') \ge \frac{1}{2k}\frac{|\mathcal{F}|}{\binom{n}{i}}$.*

Proof. We iteratively remove edges of $\mathcal{F}$ to form $\mathcal{F}'$ in accordance with the following process. Whenever there is some $B \subseteq V(\mathcal{F})$ with $|B| \le k-1$ such that B is contained in at least one but fewer than $\frac{1}{2k}\frac{|\mathcal{F}|}{\binom{n}{|B|}}$ edges, we mark B and remove all edges containing B from $\mathcal{F}'$. Clearly at some point this process will terminate. We denote the final graph by $\mathcal{F}'$.

Let us count now how many edges $\mathcal{F}'$ has. For each $i = 1, \dots, k-1$, there are at most $\binom{n}{i}$ marked i-sets. For each marked i-set, we have removed at most $\frac{1}{2k}\frac{|\mathcal{F}|}{\binom{n}{i}}$ edges from $\mathcal{F}$. Summing over $i = 1, \dots, k-1$, we see that we have removed at most $\frac{1}{2}|\mathcal{F}|$ edges. Hence, $|\mathcal{F}'| \ge \frac{1}{2}|\mathcal{F}|$. Since the process terminates

https://doi.org/10.1017/fms.2025.10163 Published online by Cambridge University Press

with a nonempty $\mathcal{F}'$, it must be the case that for each $i = 1, \ldots, k-1$ and any i-set contained in an edge of $\mathcal{F}'$ it is in at least $\frac{1}{2k}\frac{|\mathcal{F}|}{\binom{n}{i}}$ edges, completing the proof. □

Lemma 4.2. *Let k, s, t be positive integers, where $k, s \geq 2$, $t \geq 1$. Let $\mathcal{H}$ be a k-tree on s vertices with t edges. Let $\mathcal{F}$ be a k-graph on $[n]$, where $|\mathcal{F}| = a\binom{n}{k-1}$ and $a \geq 8sk!$. Then there exists a constant $\eta > 0$ such that $\mathcal{F}$ contains at least $\eta a^t n^{s-t}$ copies of $\mathcal{H}$.*

Proof. By picking η sufficiently small, we may assume n is sufficiently large in terms of s and k. By Lemma 4.1, $\mathcal{F}$ contains a subgraph $\mathcal{F}'$ such that $|\mathcal{F}'| \geq \frac{1}{2}|\mathcal{F}|$ and for each $i \in [k-1]$, $\delta_i(\mathcal{F}') \geq \frac{1}{2k}\frac{|\mathcal{F}|}{\binom{n}{i}} = \frac{a}{2k}\frac{\binom{n}{k-1}}{\binom{n}{i}}$. Let $E_1, \ldots, E_t$ be a tree-defining ordering of $\mathcal{H}$. For each $j \in [t]$, let $\mathcal{H}_j = \{E_1, \ldots, E_j\}$. We prove by induction that for each $j = 1, \ldots, t$, $\mathcal{F}$ contains at least $(\frac{a}{8k!})^j n^{v(\mathcal{H}_j)-j}$ copies of $\mathcal{H}_j$.

For the base case, $\mathcal{F}'$ contains at least $\frac{1}{2}|\mathcal{F}| = \frac{a}{2}\binom{n}{k-1} \geq \frac{a}{8k!}n^{v(\mathcal{H}_1)-1}$ copies of $\mathcal{H}_1$. So the claim holds. For the induction step, let $1 \leq j \leq t-1$ and suppose $\mathcal{F}'$ contains at least $(\frac{a}{8k!})^j n^{v(\mathcal{H}_j)-j}$ copies of $\mathcal{H}_j$. Let F denote a parent edge of E_{j+1} in $\mathcal{H}_j$. Let $L = F \cap E_{j+1} = V(\mathcal{H}_j) \cap E_{j+1}$ and $\ell = |L|$. Let $\mathcal{H}'$ denote a copy of $\mathcal{H}_j$ in $\mathcal{F}'$ and L' the image of L in $\mathcal{H}'$. By our assumptions of $\mathcal{F}'$, $\deg_{\mathcal{F}'}(L') \geq \frac{a}{2k}\frac{\binom{n}{k-1}}{\binom{n}{\ell}}$. For each vertex $v \in V(\mathcal{H}')$, we have that v is contained in at most $\binom{n-\ell-1}{k-\ell-1}$ edges which contain L'. As there are at most s vertices in $\mathcal{H}_j$, we have that for sufficiently large n the number of edges of $\mathcal{F}'$ that contain L' but do not intersect $\mathcal{H}'$ somewhere else is at least

$$\frac{a}{2k}\frac{\binom{n}{k-1}}{\binom{n}{\ell}} - s\binom{n-\ell-1}{k-\ell-1} \geq \frac{a\ell!}{4k!}n^{k-1-\ell} - \frac{s}{(k-\ell-1)!}n^{k-\ell-1} \geq \frac{a}{8k!}n^{k-1-\ell},$$

where we used the fact that $a \geq 8sk!$. Hence $\mathcal{H}'$ can be extended to at least $\frac{a}{8k!}n^{k-1-\ell}$ copies of $\mathcal{H}_{j+1}$ in $\mathcal{F}'$. Therefore, $\mathcal{F}'$ contains at least $(\frac{a}{8k!})^{j+1}n^{v(\mathcal{H}_j)-j+k-1-\ell} = \frac{a}{8k!}^{j+1}n^{v(\mathcal{H}_{j+1})-(j+1)}$ copies of $\mathcal{H}_{j+1}$. This completes the induction and the proof. □

Before we prove the supersaturation theorem, we first deal with one of the cases.

Lemma 4.3. *Let $k \geq 4$ be an integer. Let $\mathcal{H}$ be a 2-contractible k-tree with s vertices and t edges. Let $\mathcal{F}$ be an s-super-homogeneous k-graph on $[n]$ such that $|\mathcal{F}| \geq \varepsilon n^{k-1}$ and* Int$(\mathcal{F})$ *is of type 1. Then, $\mathcal{F}$ contains at least $(\frac{\varepsilon}{2ks})^t n^{s-t}$ many copies of $\mathcal{H}$.*

Proof. Let $(X_1, \ldots, X_k)$ be a k-partition associated with $\mathrm{Int}(\mathcal{F})$. Without loss of generality, suppose $2^{[k-2]} \subseteq \mathrm{Int}(\mathcal{F})$. Since $\mathcal{H}$ is 2-contractible, each edge E contains at least two vertices of degree 1. We will designate two of them as *expansion vertices* for E.

Let $E_1, \cdots, E_t$ be a tree-defining ordering of $\mathcal{H}$. For each $i \in [t]$, let $\mathcal{H}_i = \{E_1, \ldots, E_i\}$, $v_i = |V(\mathcal{H}_i)|$, and let Π_i denote the set of injective homomorphisms φ^i of $\mathcal{H}_i$ into $\mathcal{F}$ that maps all the expansion vertices in $\mathcal{H}_i$ into $X_{k-1} \cup X_k$. We will prove by induction the stronger statement that for each $i \in [t]$, $|\Pi_i| \geq (\frac{\varepsilon}{2ks})^i n^{v_i - i}$. The base step $i = 1$ is trivial since $\mathcal{F}$ has at least εn^{k-1} edges E and for each E we can define an embedding φ^1 of $\mathcal{H}_1$ to E where the expansion vertices are mapped to $X_{k-1} \cup X_k$. For the induction step, let $1 \leq i \leq t-1$ and suppose $|\Pi_i| \geq (\frac{\varepsilon}{2ks})^i n^{v_i - i}$. Let F be a parent edge of E_{i+1} in $\mathcal{H}_i$. Consider any $\varphi^i \in \Pi_i$. Since vertices in $F \cap E_{i+1}$ have degree at least two in $\mathcal{H}$ and are not expansion vertices in $\mathcal{H}_i$, we have $\varphi^i(F \cap E_{i+1}) \subseteq \bigcup_{j=1}^{k-2} X_j$. Since $2^{[k-2]} \subseteq \mathrm{Int}(\mathcal{F})$, we have $\pi(\varphi^i(F \cap E_{i+1})) \in \mathrm{Int}(\mathcal{F})$. So, $\mathcal{L}_{\mathcal{F}}(\varphi^i(F \cap E_{i+1}))$ is s-diverse. Therefore, there are at least $\frac{s-v_i}{s}|\mathcal{L}_{\mathcal{F}}(\varphi^i(F \cap E_{i+1}))| \geq \frac{1}{s}|\mathcal{L}_{\mathcal{F}}(\varphi^i(F \cap E_{i+1}))|$ edges of $\mathcal{F}$ that contain $\varphi^i(F \cap E_{i+1})$ and do not intersect the rest of $\varphi^i(\mathcal{H}_i)$. Let $\ell = |F \cap E_{i+1}|$. Since $\mathcal{F}$ is s-super-homogeneous with $\mathrm{Int}(\mathcal{F}) \supseteq 2^{[k-2]}$ and $\pi(\varphi^i(F \cap E_{i+1})) \in \mathrm{Int}(\mathcal{F})$, by the conditions of s-super-homogeneous given in Theorem 3.1, $|\mathcal{L}(\varphi^i(F \cap E_{i+1}))| \geq \frac{|\mathcal{F}|}{2kn^\ell} \geq \frac{\varepsilon}{2k}n^{k-\ell-1}$. Hence each mapping φ_i in Π_i can be extended to at least $\frac{\varepsilon}{2ks}n^{k-\ell-1}$ injective mappings φ^{i+1} of $\mathcal{H}_{i+1}$, by mapping E_{i+1} to an edge F' of $\mathcal{F}$ that contains $\varphi^i(F \cap E_{i+1})$ and avoids the rest of $\varphi^i(\mathcal{H}_i)$. Furthermore, we can require φ^{i+1} to map the two designated expansion vertices of E_{i+1} to $X_{k-1} \cup X_k$. Hence, $|\Pi_{i+1}| \geq |\Pi_i| \cdot \frac{\varepsilon}{2ks} n^{k-\ell-1} \geq (\frac{\varepsilon}{2ks})^{i+1} n^{v_{i+1}-(i+1)}$. This completes the induction step and the proof. □

We are now ready to prove our first supersaturation theorem.

Theorem 4.4. *Let $k \geq 4$ be an integer. Let $0 < \gamma < 1$, and let $\mathcal{H}$ be a 2-contractible k-tree with s vertices and t edges. Then, there exists $\beta > 0$, depending on $\mathcal{H}$ and γ, such that the following holds. Let $\mathcal{F}$ be a k-graph on $[n]$ with $|\mathcal{F}| = a\binom{n}{k-1}$, where $a \geq \sigma(\mathcal{H}) - 1 + \gamma$. Then, $\mathcal{F}$ contains at least $\beta a^t n^{s-t}$ many copies of $\mathcal{H}$.*

Proof. For convenience, let $\sigma = \sigma(\mathcal{H})$. Note that if $a \geq 8sk!$, we may apply Lemma 4.2, so at the cost of a constant factor on β, we will assume $a = \sigma - 1 + \gamma$. We apply Theorem 3.5 to $\mathcal{F}$ with $\varepsilon = \frac{\gamma}{4\sigma}$ if $\sigma \geq 2$ and $\varepsilon = \frac{1}{4}\gamma^{k-2}$ if $\sigma = 1$. Then one of the following two cases holds.

Case 1. $\mathcal{F}$ contains an s-super-homogeneous subgraph $\mathcal{F}'$ such that $|\mathcal{F}'| \geq \frac{1}{2}c(k,s)\varepsilon|\mathcal{F}|$ and $\mathrm{Int}(\mathcal{F}')$ is of type 1.

By Lemma 4.3, $\mathcal{F}'$ contains $\left(\frac{\varepsilon c(k,s)}{2ks}\right)^t n^{s-t}$ copies of $\mathcal{H}$. So the claim holds.

Case 2. There exists a $W \subseteq [n]$ and $\mathcal{F}'' \subseteq \mathcal{F}$ satisfying the following:

(a)

$$|W| \leq \left(\frac{160\sigma^2}{c(k,s)\gamma^{2k-4}}\right)^{k-1}.$$

(b)

$$|\mathcal{F}''| \geq \left(1 - \frac{\gamma}{4\sigma}\right)|\mathcal{F}|.$$

(c) Every $F \in \mathcal{F}''$ satisfies $|W \cap F| = 1$.

Observe that in $\sigma = 1$ case, Theorem 3.5 gives the upper bound $W \leq \left(\frac{160}{c(k,s)\gamma^{2k-4}}\right)^{k-1}$ and in $\sigma \geq 2$ case, gives the bound $W \leq \left(\frac{160\sigma^2}{c(k,s)\gamma^2}\right)^{k-1}$. In both cases, the written upper bound holds. For simplicity, observe that for both choices of ε, $|\mathcal{F}''| \geq (1-\varepsilon)|\mathcal{F}| \geq \left(1 - \frac{\gamma}{4\sigma}\right)|\mathcal{F}|$.

For each $S \in \binom{W}{\sigma}$, let $\mathcal{L}^*(S) = \bigcap_{v \in S} \mathcal{L}_{\mathcal{F}''}(v)$. We say a set $S \in \binom{W}{\sigma}$ is *good*, if $|\mathcal{L}^*(S)| \geq 8s(k-1)!\binom{n}{k-2}$; otherwise we say S is *bad*. Let $\mathcal{D}$ denote the set of $(k-1)$-sets D in $[n] \setminus W$ with $d_{\mathcal{F}''}(D) \geq \sigma$. Since each edge of $\mathcal{F}''$ contains one vertex in W and a $(k-1)$-set in $[n] \setminus W$, via double counting, we get

$$\sum_{S \in \binom{W}{\sigma}} |\mathcal{L}^*(S)| = \sum_{D \in \mathcal{D}} \binom{d_{\mathcal{F}''}(D)}{\sigma}. \tag{10}$$

By our assumption, $|\mathcal{F}''| \geq (1 - \frac{\gamma}{4\sigma})|\mathcal{F}| \geq (1 - \frac{\gamma}{4\sigma})(\sigma - 1 + \gamma)\binom{n}{k-1} \geq (\sigma - 1 + \frac{\gamma}{2})\binom{n}{k-1}$. This implies that $m := \sum_{D \in \mathcal{D}} d_{\mathcal{F}''}(D) \geq \frac{\gamma}{2}\binom{n}{k-1}$. Observe that for all $D \in \mathcal{D}$, $\binom{d_{\mathcal{F}''}(D)}{\sigma} \geq \frac{d_{\mathcal{F}''}(D)}{\sigma}$.

Thus,

$$\sum_{S \in \binom{W}{\sigma}} |\mathcal{L}^*(S)| \geq \frac{m}{\sigma} \geq \frac{\gamma}{2\sigma}\binom{n}{k-1}.$$

On the other hand, the contribution to the left-hand side of (10) from the bad σ-sets of W is at most $\binom{|W|}{\sigma} 8s(k-1)!\binom{n}{k-2} < \frac{\gamma}{4\sigma}\binom{n}{k-1}$, for sufficiently large n. Hence, we have

https://doi.org/10.1017/fms.2025.10163 Published online by Cambridge University Press

$$\sum_{S\in\binom{W}{\sigma},\, S \text{ good}} |\mathcal{L}^*(S)| \geq \frac{\gamma}{4\sigma}\binom{n}{k-1}.$$

Fix any $S \in \binom{|W|}{\sigma}$ that is good, let $d(S) = \frac{|\mathcal{L}^*(S)|}{\binom{n}{k-2}}$, in particular, since S is good, $d(S) \geq 8s(k-1)!$. Let R be a minimum cross-cut of $\mathcal{H}$. Let $\mathcal{H}' = \mathcal{H} - R$. It is easy to see that $\mathcal{H}'$ is a $(k-1)$-tree. Observe that any copy of $\mathcal{H}'$ in $\mathcal{L}^*(S)$ corresponds to a copy of $\mathcal{H}$ in $\mathcal{F}$ by joining it to S and different copies of $\mathcal{H}'$ in $\mathcal{L}^*(S)$ give rise to different copies of $\mathcal{H}$. Let $t' = |\mathcal{H}'|$ and note that $v(\mathcal{H}') = s - \sigma$. By Lemma 4.2, for every good S, there are at least $\eta d(S)^{t'} n^{s-\sigma-t'}$ many copies of $\mathcal{H}'$ in $\mathcal{L}^*(S)$. Thus, the number of copies of $\mathcal{H}$ in $\mathcal{F}$ is at least:

$$\begin{aligned}\sum_{S\in\binom{W}{\sigma},\, S \text{ good}} \eta d(S)^{t'} n^{s-\sigma-t'} &\geq \eta n^{s-\sigma-t'-(k-2)t'} \sum_{S\in\binom{W}{\sigma},\, S \text{ good}} |\mathcal{L}^*(S))|^{t'} \\ &\geq \eta n^{s-\sigma-t'-(k-2)t'} |W|^{\sigma} \left(\frac{1}{|W|^{\sigma}} \sum_{S \text{ good}} |\mathcal{L}^*(S)|\right)^{t'} \\ &\geq \eta n^{s-\sigma-t'-(k-2)t'} |W|^{\sigma} \left(\frac{1}{|W|^{\sigma}} \frac{\gamma}{4\sigma}\binom{n}{k-1}\right)^{t'} \\ &\geq \beta a^t n^{s-\sigma} \geq \beta a^t n^{s-t},\end{aligned}$$

for some constant $\beta > 0$, where we used the fact that $a = \sigma - 1 + \gamma$ is a constant and $|W|$ is bounded by a constant. □

Note that the bound obtained in Theorem 4.4 is tight up to a constant factor, as shown by considering $G(n,p)^{(k)}$, with $p = |\mathcal{F}|/\binom{n}{k}$.

5. Optimal Balanced Supersaturation at the Turán Threshold

In this section, we develop a balanced version of the supersaturation theorem of Section 4. In the next section, we will use this balanced supersaturation theorem to obtain almost tight bounds on the number of $\mathcal{H}$-free graphs.

Let $d > 0$ be a real. A t-uniform hypergraph $\mathcal{K}$ is called *d-graded* if for all sets $S \subseteq V(\mathcal{K})$ with $1 \leq |S| \leq t$, we have

$$d_{\mathcal{K}}(S) \leq d^{t-|S|}.$$

We say $S \subseteq V(\mathcal{K})$ is *saturated* by $\mathcal{K}$ if $d_{\mathcal{K}}(S) = \lfloor d^{t-|S|} \rfloor$. We say S is *admissible* if no subset of it is saturated, and *inadmissible* otherwise. For each admissible $S \subseteq V(\mathcal{K})$, let

$$\mathcal{Z}_{\mathcal{K}}(S) := \{v \in V(\mathcal{K}) : \{v\} \cup S \text{ is inadmissible and } \{v\} \text{ is not saturated}\}.$$

Lemma 5.1. *Let $d \geq 2$ be a real and $\mathcal{K}$ be a d-graded t-uniform hypergraph of size less than $d^{t-1}M$. Then, the set of vertices $v \in V(\mathcal{K})$ satisfying $d_{\mathcal{K}}(v) = \lfloor d^{t-1} \rfloor$ is less than $2tM$. Furthermore, any admissible set S of size $|S| \leq t-1$, satisfies*

$$|\mathcal{Z}_{\mathcal{K}}(S)| \leq t2^{t+1}d.$$

Proof. Let $B = \{v \in V(\mathcal{K}) : d_{\mathcal{K}}(v) = \lfloor d^{t-1} \rfloor\}$. Then clearly,

$$|B|\lfloor d^{t-1} \rfloor \leq t|\mathcal{K}|.$$

Since $|\mathcal{K}| \leq d^{t-1}M$, we have $|B| \leq 2tM$.

https://doi.org/10.1017/fms.2025.10163 Published online by Cambridge University Press

Now, consider any admissible set $S \subseteq V(\mathcal{K})$ with $|S| \leq t-1$. For every $D \subseteq S$, let $\mathcal{B}(D) := \{v \in V(\mathcal{K}) : D \cup \{v\} \text{ is saturated}\}$. By definition, $\mathcal{Z}_{\mathcal{K}}(S) = \bigcup_{D\subseteq S, D\neq\emptyset} \mathcal{B}(D)$. Consider a fixed nonempty $D \subseteq S$. Note that

$$\sum_{v\in\mathcal{B}(D)} d_{\mathcal{K}}(D\cup\{v\}) \leq t d_{\mathcal{K}}(D), \tag{11}$$

since each edge of $\mathcal{K}$ containing D is counted at most t times in the sum on the left-hand side. Since S is admissible, D is not saturated. Thus, $d_{\mathcal{K}}(D) \leq d^{t-|D|}$. On the other hand, for every $v \in \mathcal{B}(D)$, $D \cup \{v\}$ is saturated. So, $d_{\mathcal{K}}(D \cup \{v\}) = \lfloor d^{t-|D|-1} \rfloor$. By (11), we get $|\mathcal{B}(D)|\lfloor d^{t-|D|-1}\rfloor \leq td^{t-|D|}$, implying

$$|\mathcal{B}(D)| \leq 2td.$$

Summing over all $D \subseteq S$ with $D \neq \emptyset$, we get $|\mathcal{Z}(S)| \leq t2^{t+1}d$. □

We say that a hypergraph $\mathcal{H}$ is *ℓ-overlapping* if every two edges $E, F \in \mathcal{H}$ satisfy that $|E \cap F| \leq \ell$. The next lemma says for any 2-contractible, ℓ-overlapping k-tree $\mathcal{H}$, a sufficiently dense k-graph $\mathcal{F}$ always contains a dense balanced collection of copies of $\mathcal{H}$.

Lemma 5.2. *Let $k \geq 4$ be an integer. Let $\mathcal{H}$ be a 2-contractible, ℓ-overlapping k-tree with t edges. There exists a_0 and β such that the following holds if n is sufficiently large. Let $\mathcal{F}$ be a k-graph on $[n]$ that has size $a\binom{n}{k-1}$, with $a \geq a_0$. Then, there exists a collection $\mathcal{K}$ of copies of $\mathcal{H}$ in $\mathcal{F}$ such that:*

1.

$$|\mathcal{K}| \geq \beta a n^{k-1}(\beta a n^{k-\ell-1})^{t-1}$$

2. *For all $S \subseteq \mathcal{F}$ with $1 \leq |S| \leq t$,*

$$d_{\mathcal{K}}(S) \leq (\beta a n^{k-\ell-1})^{t-|S|}.$$

Proof. Let β be sufficiently small and a_0 be sufficiently large. Let $\mathcal{K}$ be maximal collection of copies of $\mathcal{H}$ in $\mathcal{F}$ that satisfies condition 2. Since any collection that consists of a single copy of $\mathcal{H}$ satisfies condition 2, $\mathcal{K}$ exists. We show that $\mathcal{K}$ must also satisfy condition 1. Suppose on the contrary that $|\mathcal{K}| < \beta^t n^{k-1}(an^{k-\ell})^{t-1}$. To derive a contradiction, it suffices to show that $\mathcal{F}$ contains a copy $\mathcal{H}^*$ of $\mathcal{H}$ that is admissible relative to $\mathcal{K}$. Indeed, if we found such an $\mathcal{H}^*$, in particular, it would not be saturated by $\mathcal{K}$. Hence, $d_{\mathcal{K}}(\mathcal{H}^*) < (\beta a n^{k-\ell-1})^{t-t} = 1$. So, $\mathcal{H}^* \notin \mathcal{K}$. Furthermore, for all $S \subseteq \mathcal{H}^*$, we would have $d_{\mathcal{K}}(S) \leq \lfloor(\beta a n^{k-\ell})^{t-|S|}\rfloor - 1$. Thus, $d_{\mathcal{K}\cup\mathcal{H}^*}(S) \leq (\beta a n^{k-\ell})^{t-|S|}$. Therefore, $\mathcal{K} \cup \{\mathcal{H}^*\}$ would still satisfy condition 2, which would contradict the maximality of $\mathcal{K}$ and complete the proof. Thus, it suffices to find such an admissible $\mathcal{H}^*$.

Since $\mathcal{K}$ satisfies condition 2, $\mathcal{K}$ is $\beta a n^{k-\ell-1}$-graded. Recall that a set $S \subseteq \mathcal{F}$ is saturated by $\mathcal{K}$ if $d_{\mathcal{K}}(S) = \lfloor(\beta a n^{k-\ell})^{t-|S|}\rfloor$. Let $\mathcal{F}_{\text{heavy}} := \{F \in \mathcal{F} : d_{\mathcal{K}}(F) = \lfloor(\beta a n^{k-\ell})^{t-1}\rfloor\}$, so that $\mathcal{F}_{\text{heavy}}$ consists of all the saturated edges of $\mathcal{F}$. By Lemma 5.1, $|\mathcal{F}_{\text{heavy}}| \leq 2t\beta a n^{k-1}$. Let $\mathcal{F}^* = \mathcal{F} \setminus \mathcal{F}_{\text{heavy}}$. Then $|\mathcal{F}^*| \geq \frac{1}{2}|\mathcal{F}|$, as long as $\beta < \frac{1}{8t(k-1)!}$. By definition, every edge in $\mathcal{F}^*$ is admissible relative to $\mathcal{K}$. By Lemma 4.1, there exists $\mathcal{F}' \subseteq \mathcal{F}^*$ such that $|\mathcal{F}'| \geq \frac{1}{4}|\mathcal{F}|$ and for each $i \in [k-1]$, $\delta_i(\mathcal{F}') \geq \frac{a}{4k}\frac{\binom{n}{k-1}}{\binom{n}{i}}$.

For each $i \in [t]$, let $\mathcal{H}_i = \{E_1, \ldots, E_i\}$. We will inductively find embeddings φ^i of $\mathcal{H}_i$ into $\mathcal{F}'$, such that $\varphi^i(\mathcal{H}_i)$ is admissible relative to $\mathcal{K}$. First, let F_1 be any edge of $\mathcal{F}'$ and let φ^1 be any bijection from the vertices of E_1 to the vertices of F_1. Since F_1 is admissible relative to $\mathcal{K}$, $\varphi^1(\mathcal{H}_1)$ is admissible. Let $1 \leq i \leq t-1$ and suppose we have defined an embedding φ^i of $\mathcal{H}_i$ into $\mathcal{F}'$ such that $\varphi^i(\mathcal{H}_i)$ admissible relative to $\mathcal{K}$. Let F denote a parent edge of E_{i+1} in $\mathcal{H}_i$. Let $p = |F \cap E_{i+1}|$. Since $\mathcal{H}$ is ℓ-overlapping, $p \leq \ell$. By our conditions on $\mathcal{F}'$, we have $d_{\mathcal{F}'}(\varphi^i(F \cap E_{i+1})) \geq \frac{a}{4k}\frac{\binom{n}{k-1}}{\binom{n}{p}}$. The number of edges of $\mathcal{F}'$ that contain $\varphi^i(F \cap E_{i+1})$ and some other vertex of $\varphi^i(\mathcal{H}_i)$ is at most $kt\binom{n}{k-p-1}$. Since $\mathcal{K}$ is $\beta a n^{k-\ell-1}$-graded,

https://doi.org/10.1017/fms.2025.10163 Published online by Cambridge University Press

by Lemma 5.1, we have $\mathcal{Z}_{\mathcal{K}}(\varphi^i(\mathcal{H}_i))$, the set of edges E of $\mathcal{F}'$ such that $\{E\} \cup \varphi^i(\mathcal{H}_i)$ is inadmissible relative to $\mathcal{K}$, has size at most $t2^{t+1}\beta a n^{k-\ell-1}$.

Therefore as long as β and a satisfy the following inequality:

$$\frac{a}{4k}\frac{\binom{n}{k-1}}{\binom{n}{p}} > kt\binom{n}{k-p-1} + 2^{t+1}t\beta a n^{k-\ell-1},$$

there is some edge E of $\mathcal{F}' \setminus \mathcal{Z}_{\mathcal{K}}(\varphi^i(\mathcal{H}_i))$ that contains $\varphi^i(F \cap E_{i+1})$ but does not intersect the rest of $\varphi^i(\mathcal{H}_i)$. Picking some a_0 sufficiently large in terms of k, ℓ, t and β sufficiently small in terms of k, ℓ, t, the inequality holds and such E exists. We extend φ^i to an embedding φ^{i+1} of $\mathcal{H}_{i+1}$ into $\mathcal{F}'$ by mapping the vertices of $E_{i+1} \setminus (F \cap E_{i+1})$ injectively to the vertices of $E \setminus \varphi^i(F \cap E_{i+1})$ so that $\varphi^{i+1}(E_{i+1}) = E$. Note that $\varphi^{i+1}(\mathcal{H}_{i+1}) = \varphi^i(\mathcal{H}_i) \cup \{E\}$. Since $E \notin \mathcal{Z}_{\mathcal{K}}(\varphi^i(\mathcal{H}_i))$ and $\varphi^{i+1}(\mathcal{H}_{i+1})$ is admissible relative to $\mathcal{K}$. This completes the induction. Now, $\varphi^t(\mathcal{H}_t)$ is a copy of $\mathcal{H}$ that is admissible relative to $\mathcal{K}$, completing our proof. □

Next, we seek to prove a similar result for $|\mathcal{F}| = a\binom{n}{k-1}$ with $\sigma(\mathcal{H}) - 1 + \gamma \le a \le a_0$, where γ is any given small positive real. At the cost of a constant, we will prove a balanced supersaturation lemma only for $a = \sigma(\mathcal{H}) - 1 + \gamma$.

Lemma 5.3. *Let $k \ge 4$ be an integer. Let $\mathcal{H}$ be a 2-contractible ℓ-overlapping k-tree with t edges. Let $\gamma > 0$ be a positive real. There exists a $\beta > 0$ such that the following holds if n is sufficiently large. Let $\mathcal{F}$ be a k-graph on $[n]$ with $|\mathcal{F}| \ge (\sigma(\mathcal{H}) - 1 + \gamma)\binom{n}{k-1}$. Then, there exists a collection $\mathcal{K}$ of copies of $\mathcal{H}$ in $\mathcal{F}$ such that:*

1.

$$|\mathcal{K}| \ge \beta n^{k-1}(\beta n^{k-\ell-1})^{t-1}$$

2. *For all $S \subseteq \mathcal{F}$ with $1 \le |S| \le t$,*

$$d_{\mathcal{K}}(S) \le (\beta n^{k-\ell-1})^{t-|S|}.$$

Proof. For convenience, let $\sigma = \sigma(\mathcal{H})$. Let β be sufficiently small and a_0 be sufficiently large. Let $\mathcal{K}$ be maximal collection of copies of $\mathcal{H}$ in $\mathcal{F}$ that satisfies condition 2. Since any collection consisting of a single copy of $\mathcal{H}$ satisfies condition 2, $\mathcal{K}$ exists. We show that $\mathcal{K}$ must also satisfy condition 1. Suppose on the contrary that $|\mathcal{K}| < \beta^t n^{k-1}(n^{k-\ell})^{t-1}$. To derive a contradiction, as in the proof Lemma 5.2, it suffices to show that $\mathcal{F}$ contains a copy $\mathcal{H}^*$ of $\mathcal{H}$ that is admissible relative to $\mathcal{K}$.

By definition, $\mathcal{K}$ is $\beta n^{k-\ell-1}$-graded. Recall that a set $S \subseteq \mathcal{F}$ is saturated with respect to $\mathcal{K}$ if $\deg_{\mathcal{K}}(S) = \lfloor(\beta n^{k-\ell})^{t-|S|}\rfloor$. Let $\mathcal{F}_{\text{heavy}} := \{F \in \mathcal{F} : \deg_{\mathcal{K}}(F) = \lfloor(\beta n^{k-\ell})^{t-1}\rfloor\}$. Note that $\mathcal{F}_{\text{heavy}}$ consists of all the saturated edges. By Lemma 5.1, $|\mathcal{F}_{\text{heavy}}| \le 2t\beta n^{k-1} < \frac{\gamma}{2}\binom{n}{k-1}$, if β is sufficiently small. Let $\widehat{\mathcal{F}} := \mathcal{F} - \mathcal{F}_{\text{heavy}}$. Then $|\widehat{\mathcal{F}}| \ge (\sigma - 1 + \frac{\gamma}{2})\binom{n}{k-1}$. By definition, all edges of $\widehat{\mathcal{F}}$ are admissible relative to $\mathcal{K}$. Let $s = v(\mathcal{H})$. Now we apply Theorem 3.5 to $\widehat{\mathcal{F}}$ with

$$a = (\sigma - 1 + \frac{\gamma}{2}) \text{ and } \varepsilon = \frac{1}{\sigma}\left(\frac{\gamma}{2}\right)^{k-2}.$$

There are two cases to consider:

Case 1. $\widehat{\mathcal{F}}$ satisfies case (1) of Theorem 3.5.

In this case, we find an $\mathcal{F}' \subseteq \widehat{\mathcal{F}}$ which is s-super-homogeneous, $|\mathcal{F}'| \ge \frac{c(k,s)\gamma^{k-2}}{2^{k-1}\sigma}\binom{n}{k-1}$, and $\text{Int}(\mathcal{F}')$ is of type 1, where without loss of generality $2^{[k-2]} \subseteq \text{Int}(\mathcal{F}')$. Let $(X_1, \dots, X_k)$ be the k-partition associated with $\text{Int}(\mathcal{F}')$. Since $\mathcal{H}$ is 2-contractible, each edge contains at least two vertices of degree 1. We will designate two of them as *expansion vertices* for E.

https://doi.org/10.1017/fms.2025.10163 Published online by Cambridge University Press

Let $E_1, \cdots, E_t$ be a tree-defining ordering of the edges of $\mathcal{H}$. For each $i \in [t]$, let $\mathcal{H}_i = \{E_1, \dots, E_i\}$. Let Π_i denote the set of embeddings φ^i of $\mathcal{H}_i$ into $\mathcal{F}'$ satisfying that $\varphi^i(\mathcal{H}_i)$ is admissible relative to $\mathcal{K}$ and φ^i maps all expansion vertices in $\mathcal{H}_i$ to $X_{k-1} \cup X_k$. We use induction on i to prove that for each $i \in [t]$, $\Pi_i \neq \emptyset$.

For the base step, we let F_1 be any edge of $\mathcal{F}'$. Let φ^1 be a bijection that maps the vertices of E_1 to the vertices of F_1 such that the two expansion vertices are mapped to $F_1 \cap X_{k-1}$ and $F_1 \cap X_k$, respectively. Since all edges of $\mathcal{F}'$ are admissible, $\varphi^1(\mathcal{H}_1)$ is admissible. Hence, $\Pi_1 \neq \emptyset$.

For the induction step, let $1 \leq i \leq t-1$ and suppose that $\Pi_i \neq \emptyset$. Let $\varphi^i \in \Pi_i$. Let F be a parent edge of E_{i+1} in $\mathcal{H}_i$. Since vertices in $F \cap E_{i+1}$ have degree at least two in $\mathcal{H}$, they are not expansion vertices in $\mathcal{H}_i$. Hence, $\varphi^i(F \cap E_{i+1}) \subseteq \bigcup_{j=1}^{k-2} X_j$. Since $2^{[k-2]} \subseteq \mathrm{Int}(\mathcal{F})$, we have $\pi(\varphi^i(F \cap E_{i+1})) \in \mathrm{Int}(\mathcal{F})$. So, in particular that $\mathcal{L}_{\mathcal{F}}(\varphi^i(F \cap E_{i+1}))$ is s-diverse. Therefore, we have that at least $1 - \frac{v(H_i)}{s} \geq \frac{1}{s}$ proportion of the edges in $\mathcal{L}_{\mathcal{F}}(\varphi^i(F \cap E_i))$ do not intersect the rest of $\varphi^i(\mathcal{H}_i)$. Furthermore, by Lemma 5.1, with $d = \beta n^{k-\ell-1}$, $|\mathcal{Z}_{\mathcal{K}}(\varphi^i(\mathcal{H}_i))| \leq t2^{t+1}\beta n^{k-\ell-1}$. Let $p = |F \cap E_{i+1}|$. Since $\mathcal{H}$ is ℓ-overlapping, $p \leq \ell$. Since $\mathcal{F}'$ is s-super-homogoneus and $\pi(F \cap E_{i+1}) \in \mathrm{Int}(\mathcal{F}')$,

$$|\mathcal{L}_{\mathcal{F}'}(\varphi^i(F \cap E_{i+1}))| \geq \frac{|\mathcal{F}'|}{2kn^p} \geq \frac{c(k,s)\gamma^{k-2}}{2^k \sigma k!} n^{k-p-1} \geq \frac{c(k,s)\gamma^{k-2}}{2^k \sigma k!} n^{k-\ell-1}.$$

By picking β sufficiently small in terms of γ and $\mathcal{H}$, we can ensure that

$$\frac{1}{s}|\mathcal{L}'_{\mathcal{F}}(\varphi^i(F \cap E_{i+1}))| > |\mathcal{Z}_{\mathcal{K}}(\varphi^i(\mathcal{H}_i))|.$$

Hence, $\mathcal{F}' \setminus \mathcal{Z}(\varphi^i(\mathcal{H}_i))$ contains an edge E that contains $\varphi^i(F \cap E_{i+1})$ but does not intersect the rest of $\varphi^i(\mathcal{H}_i)$. We extend φ^i to an embedding of $\mathcal{H}_{i+1}$ in $\mathcal{F}'$ by mapping the vertices of $E_{i+1} \setminus (F \cap E_{i+1})$ injectively into $E \setminus \varphi^i(F \cap E_{i+1})$, so that $\varphi^{i+1}(E_{i+1}) = E$ and the expansion vertices of E are mapped into X_{k-1} and X_k. Since $E \notin \mathcal{Z}_{\mathcal{K}}(\varphi^i(\mathcal{H}_i))$, $\varphi^{i+1}(\mathcal{H}_{i+1})$ is admissible relative to $\mathcal{K}$. This completes the induction. Now, $\varphi^t(\mathcal{H}_t)$ is a copy of $\mathcal{H}$ in $\mathcal{F}'$ that is admissible relative to $\mathcal{K}$. This completes the proof for Case 1.

Case 2. $\widehat{\mathcal{F}}$ satisfies case (2) of Theorem 3.5.

In this case, there exist a $W \subseteq [n]$ and $\mathcal{F}'' \subseteq \widehat{\mathcal{F}}$ satisfying

(a) $|W| \leq \left(\frac{5 \cdot 2^{2k-2}\sigma^2}{c(k,s)\gamma^{2k-4}}\right)^{k-1}$.
(b) $|\mathcal{F}''| \geq (1 - \frac{\gamma^{k-2}}{2^{k-2}\cdot\sigma})|\widehat{\mathcal{F}}| \geq (1 - \frac{\gamma}{4\sigma})(\sigma - 1 + \frac{\gamma}{2})\binom{n}{k-1} \geq (\sigma - 1 + \frac{\gamma}{8})\binom{n}{k-1}$.
(c) Every $F \in \mathcal{F}''$ satisfies $|W \cap F| = 1$.

For each $S \in \binom{W}{\sigma}$, let $\mathcal{L}^*(S) = \bigcap_{v \in S} \mathcal{L}_{\mathcal{F}''}(v)$. Let $\mathcal{D}$ denote the set of $(k-1)$-sets D in $[n] \setminus W$ with $d_{\mathcal{F}''}(D) \geq \sigma$. Since each edge of $\mathcal{F}''$ contains one vertex in W and a $(k-1)$-set in $[n] \setminus W$, via double counting, we get

$$\sum_{S \in \binom{W}{\sigma}} |\mathcal{L}^*(S)| = \sum_{D \in \mathcal{D}} \binom{d_{\mathcal{F}''}(D)}{\sigma}.$$

By our assumption, $|\mathcal{F}''| \geq (\sigma - 1 + \frac{\gamma}{8})\binom{n}{k-1}$. This implies that $m := \sum_{D \in \mathcal{D}} d_{\mathcal{F}''}(D) \geq \frac{\gamma}{8}\binom{n}{k-1}$. Thus, by observing that if $d_{\mathcal{F}''}(D) \geq \sigma$, $\binom{d_{\mathcal{F}''}(D)}{\sigma} \geq \frac{d_{\mathcal{F}''}(D)}{\sigma}$, the right-hand side is at least $\frac{\gamma}{8\sigma}\binom{n}{k-1}$.

Hence, there exists a $S \in \binom{W}{\sigma}$, such that

$$|\mathcal{L}^*(S)| \geq \frac{\gamma}{8\sigma|W|^\sigma}\binom{n}{k-1}. \tag{12}$$

https://doi.org/10.1017/fms.2025.10163 Published online by Cambridge University Press

We fix such an S. By Lemma 4.1, $\mathcal{L}^*(S)$ contains a subgraph $\mathcal{L}'$ such that $|\mathcal{L}'| \geq \frac{1}{2}|\mathcal{L}^*(S)|$ and for each $i \in [k-1]$, $\delta_i(\mathcal{L}') \geq \frac{1}{2k}\frac{|\mathcal{L}^*(S)|}{\binom{n}{i}}$. Note that since $\mathcal{L}' \subseteq \mathcal{L}^*(S)$, for each $D \in \mathcal{L}'$ and each $v \in S$, $D \cup \{v\} \in \mathcal{F}''$.

Let $E_1, \dots, E_t$ be a tree-defining order of $\mathcal{H}$. For each $i \in [t]$, let $\mathcal{H}_i = \{E_1, \dots, E_i\}$. Let R be a minimum cross-cut of $\mathcal{H}$. For each $i \in [t]$, let $\{v_i\} = E_i \cap R$ and $E'_i = E_i \setminus \{v_i\}$. Let $\mathcal{H}' = \mathcal{H} - R$. It is easy to see by definition that $\mathcal{H}'$ is a $(k-1)$-tree for which $E'_1, \dots, E'_t$ is a tree-defining ordering of $\mathcal{H}'$. Furthermore, since each E_i contains two degree 1 vertices, $E'_1, \dots, E'_t$ are all distinct.

For each $i \in [t]$, let Π_i denote the collection of embeddings of φ^i of $\mathcal{H}_i$ into $\mathcal{F}'$ such that $\varphi(\mathcal{H}_i)$ is admissible relative to $\mathcal{K}$. We prove by induction on i that $\Pi_i \neq \emptyset$. Let g be any bijection from R to S. For the base step, let D_1 be any edge of $\mathcal{L}'$. We define φ^1 to be any mapping from $V(\mathcal{H}_1)$ to $V(\mathcal{F}'')$ that maps v_1 to $g(v_1)$ and E'_1 to D_1. By earlier discussion, $\varphi^1(E_1) = D_1 \cup \{v_1\}$ is an edge in $\mathcal{F}''$. Since all edges in $\widehat{\mathcal{F}}$ are admissible, all edges in $\mathcal{F}''$ are admissible. Hence, $\varphi^1(\mathcal{H}_1) = \varphi^1(E_1)$ is admissible relative to $\mathcal{K}$. For the induction step, let $1 \leq i \leq t-1$ and suppose $\Pi_i \neq \emptyset$. Let $\varphi^i \in \Pi_i$. Suppose a parent edge of E'_{i+1} in $\mathcal{H}'_i$ is E'_j, for some $j \leq i$. Let $p = |E_j \cap E'_{i+1}|$. Since $\mathcal{H}$ is ℓ-overlapping, $\mathcal{H}'$ is ℓ-overlapping. So $p \leq \ell$. By (12) and subsequent discussion, $d_{\mathcal{L}'}(\varphi^i(E'_j \cap E'_{i+1})) \geq \frac{\gamma}{16\sigma|W|^\sigma}\frac{\binom{n}{k-1}}{\binom{n}{p}}$. The number of edges of $\mathcal{L}'$ containing $\varphi^i(E'_j \cap E'_{i+1})$ and intersect the rest of $\varphi^i(\mathcal{H}_i)$ is at most $kt\binom{n}{k-p-2}$. Furthermore, by Lemma 5.1, with $d = \beta n^{k-\ell-1}$, $|\mathcal{Z}_{\mathcal{K}}(\varphi^i(\mathcal{H}_i))| \leq 2^{t+1}t\beta n^{k-\ell-1}$. Let $\mathcal{Z}' = \{E \setminus R : E \in \mathcal{Z}_{\mathcal{K}}(\varphi^i(\mathcal{H}_i))\}$. Then $|\mathcal{Z}'| \leq 2^{t+1}t\beta n^{k-\ell-1}$. By choosing β small enough, for sufficiently large n, we have

$$\frac{\gamma}{16k\sigma|W|^\sigma}\frac{\binom{n}{k-1}}{\binom{n}{p}} > kt\binom{n}{k-p-2} + 2^{t+1}t\beta n^{k-\ell-1},$$

where we used the fact that $|W|$ is bounded by a constant.

Hence, $\mathcal{L}' \setminus Z'$ contains an edge D that contains $\varphi^i(E'_j \cap E'_{i+1})$ but does not intersect the rest of $\varphi^i(\mathcal{H}_i)$. We extend φ^i to an embedding of $\mathcal{H}_{i+1}$ in $\mathcal{F}'$ by mapping the vertices of $E'_{i+1} \setminus (E'_j \cap E'_{i+1})$ injectively into $D \setminus \varphi^i(E'_j \cap E'_{i+1})$ and mapping v_{i+1} to $g(v_{i+1})$ so that $\varphi^{i+1}(E_{i+1}) = D \cup \{g(v_{i+1})\}$. Since $D \notin \mathcal{Z}'$, $\varphi^{i+1}(E_{i+1}) \notin \mathcal{Z}_{\mathcal{K}}(\varphi^i(\mathcal{H}_i))$, and hence $\varphi^{i+1}(\mathcal{H}_{i+1})$ is admissible relative to $\mathcal{K}$. This completes the induction. Now, $\varphi^t(\mathcal{H}_t)$ is a copy of $\mathcal{H}$ in $\mathcal{F}''$ that is admissible relative to $\mathcal{K}$. This completes the proof for Case 2. □

Lemma 5.2 and Lemma 5.3 together imply the following:

Theorem 5.4. *Let $k \geq 4$. Let $\varepsilon > 0$. Let $\mathcal{H}$ be a 2-contractible ℓ-overlapping k-tree with t edges. There exist $c = c(\varepsilon, \mathcal{H}), c' = c'(\varepsilon, \mathcal{H}) > 0$ such that the following holds. Let $\mathcal{F}$ be a k-graph on $[n]$ with $|\mathcal{F}| = a\binom{n}{k-1}$, where $a \geq \sigma(\mathcal{H}) - 1 + \varepsilon$. Then, there exists a collection $\mathcal{K}$ of copies of $\mathcal{H}$ in $\mathcal{F}$ such that $|\mathcal{K}| \geq c'an^{k-1}(an^{k-\ell-1})^{t-1}$ and for all $1 \leq b \leq t$:*

$$\Delta_b(\mathcal{K}) \leq c \cdot [\tau(\mathcal{F})]^{b-1}\frac{|\mathcal{K}|}{|\mathcal{F}|},$$

where $\tau(\mathcal{F}) := \frac{1}{an^{k-\ell-1}}$.

Observe that the size of the collection $\mathcal{K}$ returned by Theorem 5.4 matches the bound in Theorem 4.4 if the intersection of each edge of $\mathcal{H}$ and its parent edge has size exactly ℓ. Thus, it can be viewed as a direct strengthening of Theorem 4.4 for such 2-contractible hypertrees.

6. Number of $\mathcal{H}$-free hypergraphs

In this section, we prove Theorem 1.3 and Theorem 1.5. First, we recall the celebrated Container Lemma as follows.

https://doi.org/10.1017/fms.2025.10163 Published online by Cambridge University Press

Lemma 6.1 [2, 42]. *For every $t \in \mathbb{N}$ and every $c > 0$, there exists a $\delta > 0$ such that the following holds for all $N \in \mathbb{N}$. Let $\mathcal{K}$ be a t-uniform hypergraph on N vertices such that for all $1 \leq b \leq t$,*

$$\Delta_b(\mathcal{K}) \leq c\tau^{b-1}\frac{|\mathcal{K}|}{N}.$$

Then there exists a collection $\mathcal{C}$ of subsets of $V(\mathcal{K})$ such that the following holds:

1. *For every $I \in \mathcal{I}(\mathcal{K})$, there is a $C \in \mathcal{C}$ such that $I \subseteq C$.*
2. $|\mathcal{C}| \leq \binom{N}{\leq t\tau N}$.
3. *For every $C \in \mathcal{C}$, $|V(C)| \leq (1-\delta)|V(\mathcal{K})|$.*

Applying Lemma 6.1 with Theorem 5.4, we have the following:

Lemma 6.2. *Let $k \geq 4$ be an integer. Let $\mathcal{H}$ be a 2-contractible ℓ-overlapping k-tree with t edges and cross-cut number σ. Then, for every $\varepsilon > 0$, there exists a $\delta > 0$ such that the following holds. Let $\mathcal{F}$ be a k-graph on $[n]$ with $|\mathcal{F}| \geq (\sigma - 1 + \varepsilon)\binom{n}{k-1}$. Then, there is a collection $\mathcal{C} \subseteq 2^{\mathcal{F}}$ satisfying*

1. *For every $\mathcal{H}$-free subgraph $I \subseteq \mathcal{F}$, there is a $C \in \mathcal{C}$ such that $I \subseteq C$.*
2. $|\mathcal{C}| \leq \binom{\binom{n}{k}}{\leq \frac{t}{n^{k-\ell-1}}\binom{n}{k-1}}$.
3. *For every $C \in \mathcal{C}$, $|V(C)| \leq (1-\delta)|\mathcal{F}|$.*

Proof. Given $\mathcal{F}$, let $\mathcal{K}$ be the collection of copies of $\mathcal{H}$ returned by Theorem 5.4. We will view $\mathcal{K}$ as a t-uniform hypergraph with $V(\mathcal{K}) = E(\mathcal{F})$ whose hyperedges are the copies of $\mathcal{H}$ in the collection. By the conditon on $\mathcal{K}$, we have that for all $1 \leq b \leq t$

$$\Delta_b(\mathcal{K}) \leq c\tau(\mathcal{F})^{b-1}\frac{|\mathcal{K}|}{|\mathcal{F}|}$$

where $\tau(\mathcal{F}) := \frac{1}{n^{k-\ell-1}}\frac{\binom{n}{k-1}}{|\mathcal{F}|}$.

Then, applying Lemma 6.1, there is a δ and a collection $\mathcal{C} \subseteq 2^{E(\mathcal{F})}$ satisfying:

1. For every independent set $I \subseteq \mathcal{H}$, there is a $C \in \mathcal{C}$ such that $I \subseteq C$.
2. $|\mathcal{C}| \leq \binom{n}{\leq \frac{t}{n^{k-\ell-1}}\binom{n}{k-1}}$.
3. For every $C \in \mathcal{C}$, $|V(C)| \leq (1-\delta)|\mathcal{F}|$.

Noting that every $\mathcal{H}$-free subfamily of $\mathcal{F}$ corresponds to an independent set in $\mathcal{K}$, the result follows. □

We now apply Lemma 6.2 repeatedly to derive the following lemma.

Lemma 6.3. *Let $k \geq 4$ be an integer. Let $\mathcal{H}$ be a 2-contractible ℓ-overlapping k-tree with t edges and cross-cut number σ. Let $\varepsilon > 0$. Then there exists a constant K such that there is a collection $\mathcal{C}$ of size at most*

$$\exp\left(\frac{K\log^2(n)}{n^{k-\ell-1}}\binom{n}{k-1}\right)$$

subgraphs of $K_n^{(k)}$ such that for every $\mathcal{H}$-free subgraph I of $K_n^{(k)}$, there is a $C \in \mathcal{C}$ with $I \subseteq C$ and every $C \in \mathcal{C}$ has size at most $(\sigma - 1 + \varepsilon)\binom{n}{k-1}$.

Proof. Let $\delta := \delta(\varepsilon, \mathcal{H})$ be the constant given by Lemma 6.2 and choose K large enough depending on ε, δ, t. Let n be sufficiently large.

We will build a sequence of auxiliary rooted trees $\mathcal{A}_1, \ldots, \mathcal{A}_m$ such that leaves of $\mathcal{A}_m$ will be our containers. We begin by fixing $\mathcal{A}_1$ to be the tree with a single vertex, the root $K_n^{(k)}$. To construct $\mathcal{A}_2$, we apply Lemma 6.2 to $K_n^{(k)}$ to find a collection $\mathcal{C}$ of subgraphs of $K_n^{(k)}$. We form $\mathcal{A}_2$ by adding each container $C \in \mathcal{C}$ as a vertex and connect it to the root $K_n^{(k)}$. Observe that in $\mathcal{A}_2$, $K_n^{(k)}$ has no more than

$$\binom{\binom{n}{k}}{\frac{2t}{n^{k-\ell-1}}\binom{n}{k-1}}$$

children, every $\mathcal{H}$-free subgraph of $K_n^{(k)}$ is contained in one of the children, and every child has size no more than $(1-\delta)\binom{n}{k-1}$.

To find $\mathcal{A}_{i+1}$ from $\mathcal{A}_i$, we first check to see if any leaf has size greater than $(\sigma - 1 + \varepsilon)\binom{n}{k-1}$. If not, we terminate the process then. For every leaf $\mathcal{F}$ of $\mathcal{A}_i$ which does have size greater than $(\sigma - 1 + \varepsilon)\binom{n}{k-1}$ (we call such $\mathcal{F}$ *dense*), we apply Lemma 6.2 to find a collection $\mathcal{C}_{\mathcal{F}} \subseteq 2^{\mathcal{F}}$ such that the following holds:

1. For every $\mathcal{H}$-free subgraph $I \subseteq \mathcal{F}$, there is there is a $C \in \mathcal{C}_{\mathcal{F}}$ such that $I \subseteq C$.
2. $|\mathcal{C}_{\mathcal{F}}| \le \binom{n}{\frac{2t}{n^{k-\ell-1}}\binom{n}{k-1}}$.
3. For every $C \in \mathcal{C}_{\mathcal{F}}$, $|V(C)| \le (1-\delta)|\mathcal{F}|$.

Then, from $\mathcal{A}_i$, we add every C in $\mathcal{C}_{\mathcal{F}}$ as a leaf below $\mathcal{F}$ for each dense $\mathcal{F}$ to form $\mathcal{A}_{i+1}$. Since for every $\mathcal{H}$-free graph $I \subseteq \mathcal{F}$, there is some $C \in \mathcal{C}_{\mathcal{F}}$ such that $I \subseteq C$, we have that every $\mathcal{H}$-free graph contained in $\mathcal{F}$ is contained in some child of $\mathcal{F}$ in $\mathcal{A}_{i+1}$. Furthermore, by construction, we have that there are no more than

$$\binom{\binom{n}{k}}{\frac{2t}{n^{k-\ell-1}}\binom{n}{k-1}}$$

children below each such $\mathcal{F}$.

We will terminate this process when every leaf has size less than $(\sigma - 1 + \varepsilon)\binom{n}{k-1}$. Let $\mathcal{A}_m$ be the final such tree. Observe that (3) implies the height of $\mathcal{A}_m$ is no more than $O(\log n)$, (1) implies that every $\mathcal{H}$-free subgraph of $K_n^{(k)}$ is contained in some leaf of $\mathcal{A}_m$, and the height condition with (2) implies that the number of leaves of $\mathcal{A}_m$ is no more than

$$\prod_{i=1}^{m} \binom{\binom{n}{k}}{\frac{2t}{n^{k-\ell-1}}\binom{n}{k-1}} \le \exp\left(\frac{K\log^2(n)}{n^{k-\ell-1}}\binom{n}{k-1}\right),$$

for some large enough constant K depending on ε, t, k, ℓ. □

We are now ready to prove Theorem 1.3 and Theorem 1.5.

Theorem 6.4 (Restatement of Theorem 1.3). *Let $k \ge 4$ be an integer. Let $\mathcal{H}$ be a 2-contractible k-tree with cross-cut number σ. The number of $\mathcal{H}$-free k-graphs on $[n]$ is at most $2^{(\sigma-1+o(1))\binom{n}{k-1}}$.*

Proof. Since $\mathcal{H}$ is 2-contractible, it is $(k-2)$-overlapping. Fix some $\varepsilon > 0$, and let $\mathcal{C}$ be the collection returned by Theorem 6.3 with $\ell = k-2$. Then every $\mathcal{H}$-free k-graph I on $[n]$ is a subgraph of some C in $\mathcal{C}$. The number of such I is at most $|\mathcal{C}|2^{(\sigma-1+\varepsilon)\binom{n}{k-1}}$.

Taking n sufficiently large, we have that $|\mathcal{C}| \le 2^{\varepsilon\binom{n}{k-1}}$, and thus the number of $\mathcal{H}$-free k-graphs on $[n]$ is at most $2^{(\sigma-1+2\varepsilon)\binom{n}{k-1}}$. As this holds for all ε, the result follows. □

Theorem 6.5 (Restatement of Theorem 1.5). *Let $k \ge 4$ be an integer. Let $\mathcal{H}$ be a 2-contractible ℓ-overlapping k-tree with cross-cut number σ. Then if $p \gg \frac{\log^2(n)}{n^{k-\ell-1}}$, with high probability* $\mathrm{ex}(G(n,p),\mathcal{H}) = (1-o(1))(\sigma-1+o(1))\binom{n}{k-1}$.

Proof. Fix some $\varepsilon > 0$, and let $\mathcal{C}$ be the collection returned by Theorem 6.3. Then, let X be the number of members $C \in \mathcal{C}$ such that $|C \cap G^{(k)}(n,p)| \ge (\sigma - 1 + 2\varepsilon)p\binom{n}{k-1}$.

The probability that a specific member of $C \in \mathcal{C}$ satisfies $|C \cap G^{(k)}(n,p)| \geq (\sigma - 1 + 2\varepsilon)p\binom{n}{k-1}$ is at most $\exp\left(-\frac{\varepsilon^2}{6\sigma}p\binom{n}{k-1}\right)$ by Chernoff's Inequality. Thus, the expected size $\mathbb{E}(X)$ of X is at most

$$\exp\left(\frac{K\log^2(n)}{n^{k-\ell-1}}\binom{n}{k-1}\right)\exp\left(-\frac{\varepsilon^2}{6\sigma}p\binom{n}{k-1}\right),$$

which goes to zero as $n \to \infty$ since $p \gg \frac{\log^2(n)}{n^{k-\ell-1}}$. Applying Markov's inequality, this implies that with high probability, $X = 0$. Hence, with high probability, the largest $\mathcal{H}$-free subgraph of $G^{(k)}(n,p)$ is of size at most $(\sigma - 1 + 2\varepsilon)p\binom{n}{k-1}$. As this holds for all ε, the upper bound follows.

A simple lower bound follows from the following: let $\mathcal{L}$ be the k-graph on $[n]$ vertices with the edge set $\{\{i_1, i_2, \ldots, i_k\} : i_i \in [\sigma - 1], \{i_2, \ldots i_k \in \binom{[n]\setminus[\sigma]}{k-1}\}$. Observe $\mathcal{L}$ contains no copy of $\mathcal{H}$ and furthermore has $(\sigma-1)\binom{n-\sigma+1}{k-1}$ edges. Note that as $|G^k(n,p) \cap \mathcal{L}| \geq (\sigma - 1 - \varepsilon)p\binom{n}{k-1}$ with probability $1 - \exp\left(-\frac{\varepsilon^2}{6\sigma}p\binom{n}{k-1}\right)$ by Chernoff's Inequality, the lower bound follows. □

Competing interests. The authors have no competing interest to declare.

Financial support. The first author's research was supported by the NSF grant DMS-1855542.

References

[1] J. Balogh, S. Das, M. Delcourt, H. Liu and M. Sharifzadeh, 'Intersecting families of discrete structures are typically trivial', *J. Combin. Th. Ser. A* **132** (2015), 224–245.

[2] J. Balogh, R. Morris and W. Samotij, 'Independent sets in hypergraphs', *J. Amer. Math. Soc.* **28** (2015), 669–709.

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

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

[5] J. Balogh and W. Samotij, 'The number of $K_{s,t}$-free graphs', *J. Lond. Math. Soc.* **83** (2011), 368–388.

[6] D. Conlon and T. Gowers, 'Combinatorial theorems in sparse random sets', *Ann. of Math.* **184** (2016), 367–454.

[7] M. Deza, P. Erdős and P. Frankl, 'Intersection properties of systems of finite sets', *Proc. Lond. Math. Soc. (3)* **36** (1978), 369–384.

[8] P. Erdős, 'On extremal problems of graphs and generalized graphs', *Israel Journal of Mathematics* **2** (1964), 183–190.

[9] P. Erdős and D. J. Kleitman, 'On coloring graphs to maximize the portion of multicolored k-edges', *J. Combin. Th* **5** (1968), 164–169.

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

[11] P. Erdős, D. J. Kleitman and B. L. Rothschild, Asymptotic enumeration of K_n-free graphs, Colloquio Internazionale sulle Teorie Combinatorie (Rome, 1973), Accad. Naz. Lincei, Rome, 1976, pp. 19–27.

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

[13] P. Erdős and M. Simonovits, 'Supersaturated graphs and hypergraphs', *Combinatorica* **3** (1983), 181–192.

[14] A. Ferber, G. McKinley and W. Samotij, 'Supersaturated sparse graphs and hypergraphs', *Int. Math. Res. Not.* (2) (2020), 378–402.

[15] P. Frankl, 'On families of finite sets no two of which intersect in a singleton', *Bull. Austral. Math. Soc.* **17** (1977), 125–134.

[16] P. Frankl and Z. Füredi, 'Forbidding just one intersection', *J. Combin. Th. Ser. A* **39** (1985), 160–176.

[17] P. Frankl and Z. Füredi, 'Exact solution of some Turán-type problems', *J. Combin. Th. Ser. A* **45** (1987), 226–262.

[18] Z. Füredi, 'On finite set-systems whose every intersection is a kernel of a star', *Discrete Math.* **47** (1983), 129–132.

[19] Z. Füredi, 'On the number of edges of quadrilateral-free graphs', *J. Combin. Th. Ser. B* **68** (1996), 1–6.

[20] Z. Füredi, 'Linear trees in uniform hypergraphs', *Euro. J. Combinatorics*, **35** (2014), 264–272.

[21] Z. Füredi and T. Jiang, 'Hypergraph Turán numbers of linear cycles', *J. Combin. Th. Ser. A* **123** (2014), 252–270.

[22] Z. Füredi and T. Jiang, 'Turán numbers of hypergraph trees', arXiv:1505.03210.

[23] Z. Füredi, T. Jiang and R. Seiver, 'Exact solution of the hypergraph Turán problem for k-uniform linear paths', *Combinatorica*, **34** (2014), 299–322.

[24] Z. Füredi and L. Özkahya, 'Unavoidable subhypergraphs: a-clusters', *J. Combin. Th. Ser. A* **118** (2011), 2246–2256.

[25] 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.

[26] T. Jiang and S. Longbrake, 'Balanced supersaturation and Turán numbers in random graphs', *Adv. Comb.* **3** (2024), 26 pp.

[27] D. J. Kleitman and K. J. Winston, 'On the number of graphs without 4-cycles', *Discrete Math.* **41** (1982), 167–172.
[28] 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.
[29] A. Kostochka, D. Mubayi and J. Verstraëte, 'Turán problems and shadows I: paths and cycles', *J. Combin. Th. Ser. A* **129** (2015), 57–79.
[30] A. Kostochka, D. Mubayi and J. Verstaëte, 'Turán problems and shadows II: trees', *J. Combin. Th. Ser. B* **122** (2017), 457–478.
[31] L. Lovász, *Combinatorial Problems and Exercises*, Problem 13.31 (Akadémiai Kiadó, Budapest and North Holland, Amsterdam, 1979).
[32] R. Morris and D. Saxton, 'The number of $C_{2\ell}$-free graphs', *Adv. Math.* **208** (2016), 534–580.
[33] D. Mubayi and J. Verstraëte, 'A survey of Turán problems for expansions', in *Recent trends in Combinatorics*, IMA Vol. Math. Appl., 159 (Springer, New York, 2016), 117–143.
[34] D. Mubayi and L. Wang, 'The number of triple systems without even cycles', *Combinatorica* **39** (2019), 679–704.
[35] D. Mubayi and L. Yepremyan, 'Random Turán theorem for hypergraph cycles', arXiv:2007.10320v1.
[36] G. McKinley and S. Spiro, 'The random Turán problem for theta graphs', arXiv:2305.16550.
[37] B. Nagle, V. Rödl and M. Schacht, 'Extremal hypergraph problems and the regularity method', *Topics in Discrete Mathematics*, Algorithms Combin., vol. 26 (Springer, Berlin, 2006), 247–278.
[38] J. Nie, 'Random Turán theorem for expansions of spanning subgraphs of tight trees', arXiv preprint, 2023, arXiv:2305.04193.
[39] J. Nie, 'Turán theorems for even cycles in random hypergraph', *J. Combin. Th. Ser. B* **167** (2024), 23–54.
[40] J. Nie and S. Spiro, 'Random Turán problems for hypergraph expansions', arXiv:2408.03406.
[41] V. Rödl and M. Schacht, 'Extremal results in random graphs, Erdős Centennial', *Bolyai Soc. Math. Stud.* **25** (2013), 535–583.
[42] D. Saxton and A. Thomason, 'Hypergraphs containers', *Invent. Math.* **201** (2015), 1–68.
[43] M. Schacht, 'Extremal results for random discrete structures', *Ann. of Math.* **184** (2016), 331–363.

https://doi.org/10.1017/fms.2025.10163 Published online by Cambridge University Press