On forbidding graphs as traces of hypergraphs

Dániel Gerbner[*]

Michael E. Picollelli[†]

Abstract

We say that a hypergraph $\mathcal{H}$ contains a graph H as a trace if there exists some set $S \subset V(\mathcal{H})$ such that $\mathcal{H}|_S = \{h \cap S : h \in E(\mathcal{H})\}$ contains a subhypergraph isomorphic to H. We study the largest number of hyperedges in 3-uniform hypergraphs avoiding some graph F as trace. In particular, we improve a bound given by Luo and Spiro in the case $F = C_4$, and obtain exact bounds for large n when F is a book graph.

1 Introduction

A fundamental theorem in extremal Combinatorics is due to Turán [29] and determines the largest number of edges in n-vertex K_k-free graphs (the case $k = 3$ was proved earlier by Mantel [20]). More generally, given a family $\mathcal{F}$ of graphs, $\mathrm{ex}(n, \mathcal{F})$ denotes the largest number of edges in n-vertex graphs that do not contain any member of $\mathcal{F}$ as a (not necessarily induced) subgraph, and if there is one forbidden subgraph, we use the simpler notation $\mathrm{ex}(n, F)$ instead of $\mathrm{ex}(n, \{F\})$. The Erdős-Stone-Simonovits theorem [7, 9] determines the asymptotics of $\mathrm{ex}(n, \mathcal{F})$ in the case $\mathcal{F}$ does not contain any bipartite graphs. The bipartite case is much less understood and is the subject of extensive research, see [12] for a survey.

There is a natural analogue of this problem for hypergraphs and was already asked by Turán. Given a family $\mathcal{F}$ of hypergraphs, we denote be $\mathrm{ex}_r(n, \mathcal{F})$ the largest number of edges in an r-uniform hypergraph that does not contain any member of $\mathcal{F}$. This problem is much more complicated, for example we still do not know the asymptotics in the next obvious question, when the complete 4-vertex 3-uniform hypergraph is forbidden. A relatively recent line of research is to consider *graph-based hypergraphs*. This is an informal common name of hypergraph classes that are obtained from graphs by enlarging their edges according to some set of rules. Extremal results concerning such hypergraphs were collected in Section 5.2. in [17].

---

[*]Alfréd Rényi Institute of Mathematics, HUN-REN, E-mail: `gerbner@renyi.hu`.

[†]Department of Mathematics, California State University San Marcos, San Marcos, CA 92096. E-mail: `mpicollelli@csusm.edu`

The most studied graph-based hypergraphs are the following. The *expansion* $F^{(r)+}$ of F is obtained by adding $r-2$ new vertices to each edge such that each new vertex is added to only one edge, see [21] for a survey on expansions. A *Berge copy* of F is obtained by adding $r-2$ new vertices to each edge arbitrarily. The new vertices may be already in F or not, and they may be added to any number of hyperedges. More precisely, we say that a hypergraph $\mathcal{F}$ is a Berge copy of F if there is a bijection f between the edges of F and the hyperedges of $\mathcal{F}$ such that for each edge e we have $e \subset f(e)$. Observe that there can be several non-isomorphic r-uniform Berge copies of F, the expansion being one of them. We denote by Berge-F the family of Berge copies of F. Each Berge copy of F is defined by a graph copy of F on a subset of the vertices, we call that the *core* of the Berge-F. Berge hypergraphs were defined (generalizing the notion of hypergraph cycles due to Berge [2]) by Gerbner and Palmer [15].

Here we study a third type of graph based hypergraphs. We denote by $\mathrm{Tr}(F)$ the family of Berge copies of F where the vertices added to the edges of F are each outside $V(F)$. In other words, the *trace* of these Berge copies is F on $V(F)$, i.e., $f(e) \cap V(F) = e$ for each edge e of F. These were called *induced Berge* in [11].

The maximum number of hyperedges in hypergraphs with some forbidden traces have long been studied. For example, the celebrated Sauer Lemma [25, 26, 30] deals with the case $\mathcal{H}$ does not contain the power set of a t-element set as a trace.

The Turán problem for these graph-based hypergraphs is closely related to the so-called *generalized Turán problems*. Given two graphs H and G, we denote by $\mathcal{N}(H, G)$ the number of copies of H contained in G. Given an integer n and graphs H and F, we let $\mathrm{ex}(n, H, F) = \max\{\mathcal{N}(H, G) : G$ is an n-vertex F-free graph$\}$. After several sporadic results, the systematic study of this function was initiated by Alon and Shikhelman [1].

It is easy to see that if we take the vertex sets of r-cliques in an F-free graph as hyperedges, the resulting graph is Berge-F-free. Therefore, we have $\mathrm{ex}(n, K_r, F) \leq \mathrm{ex}_r(n, \text{Berge-}F) \leq \mathrm{ex}_r(n, \mathrm{Tr}(F)) \leq \mathrm{ex}_r(n, F^{(r)+})$, where the second and third inequality follows from $F^{(r)+} \in \mathrm{Tr}(F) \subset$ Berge-F. Stronger connection was established for these cases: $\mathrm{ex}_r(n, \text{Berge-}F) \leq \mathrm{ex}(n, K_r, F) + \mathrm{ex}(n, F)$ [16], $\mathrm{ex}_r(n, \mathrm{Tr}(F)) = \Theta(\max_{s \leq k} \mathrm{ex}(n, K_s, F))$ [11] and $\mathrm{ex}_r(n, F^{(r)+}) = \mathrm{ex}(n, K_r, F) + O(n^{r-1})$ [13].

The first to study forbidden graphs as traces were Mubayi and Zhao [22]. They studied the case $F = K_k$ (in fact they considered the more general case when complete hypergraphs are forbidden as traces). They observed that in the case $r < k$ we have $\mathrm{ex}_r(n, \mathrm{Tr}(K_k)) = \mathrm{ex}_r(n, F^{(r)+})$ for sufficiently large n, and the exact value of that was determined by Pikhurko [23]. In the case $r \geq k$, Mubayi and Zhao conjectured that the extremal construction for sufficiently large n is the following. We take a K_k-free graph on $n - r + k - 1$ vertices with $\mathrm{ex}(n, K_{k-1}, K_k)$ copies of K_{k-1}. Note that this is the so-called *Turán graph* by a theorem of Zykov [31]. Then we take a set U of $r - k + 1$ new vertices, and pick as hyperedges the union of U with the vertex set of any $(r-1)$-clique of the Turán graph. Mubayi and Zhao proved this conjecture asymptotically if $k = 3$, i.e., they showed that $\mathrm{ex}_r(n, \mathrm{Tr}(K_3)) = n^2/4 + o(n^2)$. They proved this exactly if $r = 3$. An earlier version of this paper contained a proof of the $k = 3$ case of their conjecture for any r, i.e., the proof of $ex_r(n, \mathrm{Tr}(K_3)) = \lfloor (n - r + 2)^2/4 \rfloor$,

if n is sufficiently large. However, this statement follows from an old theorem of Frankl and Pach [10] for every n, which also implies the extremal examples are constructed by adding a common set of $r-2$ vertices to every edge of maximum complete bipartite graph on $n-r+2$ vertices.

We extend this result to book graphs in the 3-uniform case. The *book graph* B_t consists of t triangles sharing an edge, i.e., $B_t = K_{1,1,t}$.

Theorem 1.1. *For every t, if n is sufficiently large, then we have $\mathrm{ex}_3(n, \mathrm{Tr}(B_t)) = \lfloor (n-1)^2/4 \rfloor$, with equality only for the 3-uniform hypergraph formed by adding a common vertex to every edge of a maximum complete bipartite graph on $n-1$ vertices.*

We remark that for any non-star F and any r and n we have $\mathrm{ex}_r(n, \mathrm{Tr}(F)) \geq \mathrm{ex}(n-r+2, F)$, since if we add the same $r-2$ new vertices to each edge of an F-free graph, we obtain a $\mathrm{Tr}(F)$-free hypergraph.

Other specific graphs F such that $\mathrm{ex}_r(n, \mathrm{Tr}(F))$ has been studied include stars [11, 24] and $K_{2,t}$ [19, 24]. In particular, Luo and Spiro [19] showed $n^{3/2}/2 + o(n^{3/2}) \leq \mathrm{ex}_3(n, \mathrm{Tr}(C_4)) \leq 5n^{3/2}/6 + o(n^{3/2})$. We improve the constant factor in the main term of the upper bound.

Theorem 1.2. $\mathrm{ex}_3(n, \mathrm{Tr}(C_4)) \leq (1+\sqrt{2})n^{3/2}/4 + o(n^{3/2})$.

Finally, we show a connection of $\mathrm{ex}_r(\mathrm{Tr}(n, F))$ and generalized supersaturation. Given a graph F and a positive integer m, the supersaturation problem deals with the minimum number of copies of F in n-vertex graphs with at least m edges. In the generalized version, we are also given a graph H and the n-vertex graphs contain at least m copies of H. Such problems were studied in [3, 14, 18]. Note that this is equivalent to studying the most number of copies of H when we are given an upper bound on the number of copies of F.

Proposition 1.3. *There is an n-vertex graph G with $O(n^{|V(F)|-1})$ copies of F such that $\mathrm{ex}_r(n, \mathrm{Tr}(F)) \leq \mathcal{N}(K_r, G)$.*

Obviously this connection to generalized Turán problems is less useful than the previously mentioned ones, partly because there are less results on generalized supersaturation problems. In the non-degenerate case when $\chi(F) > r$, a special case of a result of Halfpap and Palmer [18] states that $\mathcal{N}(F, G) = o(n^{|V(F)|})$ implies $\mathcal{N}(K_r, G) \leq (1+o(1))\mathrm{ex}(n, K_r, F)$. Therefore, with Proposition 1.3 we obtain the (already known) asymptotics of $\mathrm{ex}(n, \mathrm{Tr}(F))$ in this case. We could not find any application of Proposition 1.3 that gives new results, although it plays a small role in the proof of Theorem 1.1.

2 Proofs

The *shadow graph* of a hypergraph $\mathcal{H}$ has vertex set $V(\mathcal{H})$ and uv is an edge if and only if there is a hyperedge in $\mathcal{H}$ containing both u and v. If $\mathcal{H}$ contains a $\mathrm{Tr}(F)$ (or any Berge-F), then the core of that is a copy of F in the shadow graph. The converse is not true, for example if $\mathcal{H}$ is r-uniform, then the shadow graph contains copies of K_r, even if $\mathcal{H}$ is Berge-K_r-free.

3

Note that in the shadow graph of a $\mathrm{Tr}(F)$-free hypergraph, each copy of F has to contain an edge uv with the property that each hyperedge containing uv contains another vertex from that copy of F. In multiple proofs below, we will pick such an edge for each copy of F and call it the *special edge* of that copy of F.

We can show the following generalization of Proposition 1.3.

Proposition 2.1. *If $\mathcal{H}$ is $\mathrm{Tr}(F)$-free, then the shadow graph G of $\mathcal{H}$ contains $O(n^{|V(F)|-1})$ copies of F.*

Note that hyperedges of $\mathcal{H}$ create distinct copies of K_r in G, thus $|E(\mathcal{H})| \leq \mathcal{N}(K_r, G)$, hence the above proposition implies Proposition 1.3.

Proof. Let us fix an edge uv and count the number of copies of F that have uv as special edge. Let h be a hyperedge containing u and v. Then each copy of F containing uv contains also at least one of the other $r-2$ vertices of h. Therefore, we can count the copies of F containing uv by picking a non-empty subset of the other vertices of h ($O(1)$ ways), then picking the rest of the vertices of F ($O(n^{|V(F)|-3})$ ways) and then picking a copy of F on those $|V(F)|$ vertices ($O(1)$ ways). As there are $O(n^2)$ special edges, the proof is complete. ∎

Let us continue with the proof of Theorem 1.2. We will use the following lemma due to Luo and Spiro (Lemma 3.1 in [19]). We say that a subset of a hyperedge is *light* if exactly one hyperedge contains it, and *heavy* otherwise.

Lemma 2.2 (Luo, Spiro [19]). *Let $\mathcal{H}$ be a $\mathrm{Tr}(C_4)$-free 3-uniform hypergraph. Let $\mathcal{H}_2$ denote the subhypergraph of $\mathcal{H}$ consisting of the hyperedges that do not contain light edges. Then every edge of the shadow graph is in at most two hyperedges of $\mathcal{H}_2$.*

We restate Theorem 1.2 here for convenience.

Theorem. $\mathrm{ex}_3(n, \mathrm{Tr}(C_4)) \leq (1 + \sqrt{2})n^{3/2}/4 + o(n^{3/2})$.

We denote hyperedges consisting of vertices u, v, w by uvw. Note that there are 4-uniform hyperedges in the proof below, but we do not use the analogous notation, because we use that for 4-cycles. We let $uvwx$ denote the 4-cycle with edges uv, vw, wx, xu.

Proof. Let $\mathcal{H}$ be a 3-uniform $Tr(C_4)$-free hypergraph, and let G denote its shadow graph. For vertices u, v of $\mathcal{H}$, we denote by $d_3(u, v)$ the number of hyperedges containing both u and v, and by $d_2(u, v)$ the number of common neighbors of u and v in G.

Claim 2.3. $d_2(u, v) \leq d_3(u, v) + 3$.

Proof of Claim. Assume we have four vertices w_1, w_2, w_3, w_4 such that for every i, uvw_i is not a hyperedge in $\mathcal{H}$, yet uw_i and vw_i are in G. Consider the 4-cycle uw_1vw_2 in G. Since it is not the core of a $Tr(C_4)$ in $\mathcal{H}$, there is a hyperedge inside these four vertices. Since no hyperedge contains u, v and w_i, we have that either uw_1w_2 or vw_1w_2 is in $\mathcal{H}$. This holds for each pair w_iw_j, which implies that the 4-cycle $w_1w_2w_3w_4$ is the core of a $Tr(C_4)$, where each edge is extended by u or v to a hyperedge. ∎

Let us return to the proof of the theorem. Consider now the triangles in G such that the vertices do not form a hyperedge of $\mathcal{H}$. For each edge of G, there are at most three such triangles, and each triangle has three edges, thus there are at most $|E(G)|$ such triangles. This implies that the total number of triangles in G is at most $|E(G)|+|E(\mathcal{H})|= O(n^{3/2})$.

Consider now two non-adjacent vertices u and v. By Claim 2.3 we have $d_2(u,v) \leq 3$. Observe that if w_1 and w_2 are in the common neighborhood of u and v, then at least one of uw_1w_2 and vw_1w_2 is in $\mathcal{H}$ (because of the 4-cycle uw_1vw_2). In particular, w_1 and w_2 are adjacent in G. Assume that $d_2(u,v) = 3$ and let w_1, w_2, w_3 be the common neighbors of u and v. Then these three vertices form a triangle, which we denote by $T(u,v)$.

Assume that uw_1w_2, uw_1w_3 and uw_2w_3 are each in $\mathcal{H}$. Since uw_1vw_2 is not a core of a $Tr(C_4)$, one of the edges incident to v, say vw_1 is contained only in hyperedges of $\mathcal{H}$ inside this 4-cycle, which must be vw_1w_2 since uv is not in the shadow graph of $\mathcal{H}$. Consider now the 4-cycle uw_1vw_3. By similar reasoning, either vw_1 or vw_3 is only in hyperedges inside this 4-cycle, thus $vw_1w_3 \in \mathcal{H}$. This contradicts that vw_1 is contained only in vw_1w_2 in $\mathcal{H}$. Therefore, uw_iw_j is not in $\mathcal{H}$ for some $1 \leq i < j \leq 3$.

We obtained that u (and analogously v) is the common neighbor of the endpoints of an edge w_iw_j in G such that the triangle uw_iw_j does not form a hyperedge in $\mathcal{H}$, and w_iw_j is in a triangle in G. For each edge of G, there are at most 3 such vertices, thus for each triangle in G, there are at most 27 such pairs. Therefore, the number of pairs (u,v) such that uv is not an edge in G and u and v have three common neighbors is at most $27\mathcal{N}(K_3,G) = O(n^{3/2})$. Let E' denote the set of non-adjacent pairs u,v with $d_2(u,v) = 3$. We have

$\sum_{uv\in E(G)} d_2(u,v) \leq \sum_{uv\in E(G)} d_3(u,v)+3 = 3|\mathcal{H}|+3|E(G)|= O(n^{3/2})$, $\sum_{(u,v)\in E'} d_2(u,v) = 3|E'|= O(n^{3/2})$ and $\sum_{(u,v)\notin E(G)\cup E'} d_2(u,v) \leq \sum_{(u,v)\notin E(G)\cup E'} 2 \leq 2\binom{n}{2}$. Therefore,

$$\sum_{u,v\in V(G)} d_2(u,v) = \sum_{uv\in E(G)} d_2(u,v) + \sum_{(u,v)\in E'} d_2(u,v) + \sum_{(u,v)\notin E(G)\cup E'} d_2(u,v) \leq (2+o(1))\binom{n}{2}.$$

This implies that $\sum_{v\in V(G)} \binom{d(v)}{2} \leq (2 + o(1))\binom{n}{2}$, and then $|E(G)|\leq (1 + o(1))\frac{\sqrt{2}}{2}n^{3/2}$ follows by Jensen's inequality.

Let $\mathcal{H}_1$ be the subhypergraph consisting of the hyperedges that contain a light edge. Let us pick a light edge from each hyperedge of $\mathcal{H}_1$ and let G_1 denote the resulting graph. Then G_1 is C_4-free, thus $|E(\mathcal{H}_1)|= |E(G_1)|\leq \mathrm{ex}(n, C_4) = (1 + o(1))n^{3/2}/2$.

Let $\mathcal{H}_2$ denote the rest of the hyperedges of $\mathcal{H}$ and G_2 denote the shadow graph of $\mathcal{H}_2$. Then G_2 may contain copies of C_4, but each copy $uvwx$ of C_4 contains an edge uv such that each hyperedge in $\mathcal{H}$ that contains u and v also contains w or x. We say that uv is a *special edge* for this C_4. In particular, u and v are contained together in at most two hyperedges of $\mathcal{H}$, and since $uv \in E(G_2)$, we have that $uvw, uvx \in \mathcal{H}$ (with at least one of them in $\mathcal{H}_2$). This also implies that there is a K_4 on u, v, w, x in G.

Let us consider a 4-cycle $uvwx$ in G_2 with special edge uv, and assume that both uvw and uvx are in $\mathcal{H}_2$. Then the K_4 on u, v, w, x is in G_2 and the 4-cycle $uwvx$ also has a special edge, without loss of generality uw. This shows that $uwx \in \mathcal{H}$ (since uw is in G_2). Then $uwx \in \mathcal{H}_2$ since each subedge is in G_2. Observe that we have found two hyperedges

5

in $\mathcal{H}_2$ containing ux. Now we have two possibilities. Either there are no further hyperedges containing ux, or there are some hyperedges containing ux (they must be in $\mathcal{H}_1$ because of Lemma 2.2).

Now we move the hyperedges uvw, uvx and uwx to $\mathcal{H}_1$ and the edges uv and uw to G_1. In the first case, if ux is not contained in any hyperedges of $\mathcal{H}_1$, we also move ux to G_1. We repeat this as long as we can. We denote the hypergraph we obtain this way from $\mathcal{H}_1$ by $\mathcal{H}'_1$ and the graph we obtain from G_1 by G'_1.

Claim 2.4. G'_1 *is C_4-free.*

Proof of Claim. First we show that each hyperedge of $\mathcal{H}$ contains at most one edge of G'_1, except for the hyperedges added to $\mathcal{H}'_1$. This holds for G_1 since the edges are light and we picked only one from each hyperedge. The hyperedges of $\mathcal{H}_1$ do not contain any of the edges uv in $E(G'_1) \setminus E(G_1)$, since the only hyperedges of $\mathcal{H}$ that contain uv and uw (and ux if that is also added) are uvw, uwx and uvx, and they are in $\mathcal{H}_2$.

Observe that the special edge ab of a C_4 $abcd$ in G'_1 would be contained in a hyperedge that contains two edges of G'_1, thus in a hyperedge that was added to $\mathcal{H}'_1$. Since the only hyperedges containing ab are abc and abd, we have that ab is moved from G_2 to G'_1 together with, say ad (and potentially ac). This implies that $abc \in \mathcal{H}_2$. Then $bc \in E(G_2)$, thus bc was also moved to G'_1. But that movement cannot be at the same time when we moved ab and ad (and potentially ac) to G'_1, thus cannot be at the same time when we moved abc to $\mathcal{H}'_1$. But we move each edge together with the two hyperedges of $\mathcal{H}_2$ that contain that edge, thus we have to move ab at the same time when we move abc, a contradiction. ∎

Let us return to the proof of the theorem. Let $f = |E(G'_1) \setminus E(G_1)|$. The number of hyperedges in $\mathcal{H}'_1$ is at most the number of edges in G_1 plus $3/2$ times the number of new edges in G'_1, i.e., $|E(\mathcal{H}'_1)| \leq (1 + o(1))n^{3/2}/2 + f/2$. We also have $|E(\mathcal{H}_1)| = |E(G_1)| \leq (1 + o(1))n^{3/2}/2 - f$.

Let $\mathcal{H}'_2$ denote the hyperedges of $\mathcal{H}$ that are not in $\mathcal{H}'_1$, thus $\mathcal{H}'_2$ is a subhypergraph of $\mathcal{H}_2$. Let G'_2 denote the graph of we obtain from G_2 by deleting the edges we moved to G'_1. We have that $|E(G'_2)| = |E(G_2)| - f$. Recall that by Lemma 2.2, each edge of G_2 is in at most two hyperedges of $\mathcal{H}_2$.

Let Y denote the set of edges in G'_2 that are in exactly two hyperedges of $\mathcal{H}_2$ and $g := |Y|$. We next provide an upper bound on g in terms of f and edges of G'_2 that are not in Y.

Let $uv \in Y$, then the two corresponding hyperedges uvw and uvz create a C_4 with a chord uv in G_2. That C_4 has a special edge, say uw, then uwz is also in $\mathcal{H}$ (if there are multiple special edges and one of them was moved to G'_1, we pick that as uw). If $uwz \notin \mathcal{H}_2$, then $uw \in G'_2$ and is contained in less than two hyperedges of $\mathcal{H}_2$. We say that uw *belongs* to uv.

Claim 2.5. *Each edge uw not in Y belongs to at most one edge uv.*

Proof of Claim. We have that uw is a special edge of a 4-cycle $uwvz$, and $uwz \notin \mathcal{H}_2$ implies that wz must be a light edge. Observe that uw can belong only to an edge in the unique hyperedge in $\mathcal{H}_2$ that contains u, w, thus to uv or wv. If uw belongs to wv as well, then uw

is the special edge of a 4-cycle of the form $uwz'v$ in G_2. We have $z' \neq z$ since wz is light. But then uwz' must be in $\mathcal{H}$ because of $uwzv'$, but cannot be in $\mathcal{H}$ because uw is only in hyperedges inside $uvwz$. ∎

Assume now that $uwz \in \mathcal{H}_2$ and we have not moved uw to G_1'. The only reasons can be that we have already moved at least one of the edges of the K_4 on u, v, w, z to G_1', or we have already moved at least one of the three hyperedges we would have moved with it to $\mathcal{H}_1'$, because of another subedge, another K_4. Let x be the fourth vertex of this other K_4. Note that if we have moved one of those edges, we have also moved all hyperedges of $\mathcal{H}_2$ containing that edge, in particular we have moved at least one of the hyperedges uwv, uwz and uvz.

No matter which of the hyperedges uwv, uwz and uvz we have moved earlier, it contains two subedges containing u. Each such subsedge is contained by two of the hyperedges uwv, uwz and uvz. Therefore, none of the hyperedges in $\mathcal{H}_2$ containing such a subedge contains x. This implies that at least two 3-sets containing x inside the K_4 are not in $\mathcal{H}_2$. But we only move things to $\mathcal{H}_1'$ if at least three hyperedges inside the K_4 are in $\mathcal{H}_2$, a contradiction. This implies that we have not moved any of these three hyperedges uwv, uwz and uvz earlier. But these hyperedges in $\mathcal{H}_2$ and whenever we move an edge to G_1', we also move the hyperedges that contain that edge from $\mathcal{H}_2$ to $\mathcal{H}_1'$, thus each subedges, each edges inside u, v, w, z have not been moved, therefore, we could move uw to G_1', a contradiction.

Finally, assume that $uwz \in \mathcal{H}_2$ and we have moved uw to G_1'. It was because of the K_4 on u, v, w, z since the hyperedges containing u, w are uwv and uwz. Then we also moved at least one adjacent edge to G_1' at the same time. If we moved three edges to G_1', then there are three edges inside the K_4 that remain in G_2'. If we moved two edges to G_1', then there are four edges inside the K_4 that remain in G_2'. This can happen at most $f/2$ times, thus there are at most $2f$ edges in Y where this occurs.

We have obtained for each edge of G_2' in Y a unique edge not in Y, with at most $2f$ exceptions. This implies that $g = |Y| \leq |E(G_2')|/2 + f = |E(G_2)|/2 + f/2$. Then in G_2, we have at most $g + f \leq |E(G_2)|/2 + 3f/2$ edges that are contained in exactly two hyperedges of $\mathcal{H}_2$. This shows that the total number of edge-hyperedge incidences between G_2 and $\mathcal{H}_2$ is at most $3|E(G_2)|/2 + 3f/2$, hence $|E(\mathcal{H}_2)| \leq |E(G_2)|/2 + f/2$. Recall that $|E(G_1)| + f/2 = |E(G_1')| - f/2 \leq (1 + o(1))n^{3/2}/2 - f/2$ and $|E(G_2)| \leq |E(G)| - |E(G_1)| \leq (1 + o(1))\frac{\sqrt{2}}{2}n^{3/2} - |E(G_1)|$.

Combining the upper bounds we obtained on $|E(\mathcal{H}_1)|$ and $|E(\mathcal{H}_2)|$, we obtain that $|E(\mathcal{H})| = |E(\mathcal{H}_1)| + |E(\mathcal{H}_2)| \leq |E(G_1)| + |E(G_2)|/2 + f/2 \leq \frac{|E(G_1)|}{2} + \frac{|E(G)|}{2} + f/2 \leq \frac{|E(G_1)|}{2} + (1 + o(1))\frac{\sqrt{2}}{4}n^{3/2} + f/2 = (1 + o(1))n^{3/2}/4 + (1 + o(1))\frac{\sqrt{2}}{4}n^{3/2}$. ∎

We finish the paper by proving Theorem 1.1. Recall that it states $\mathrm{ex}_3(n, \mathrm{Tr}(B_t)) = \lfloor (n-1)^2/4 \rfloor$ for sufficiently large n.

Proof of Theorem 1.1. The lower bound is given by taking a complete bipartite graph on $n - 1$ vertices, and adding the same new vertex to each edge.

Let $\mathcal{H}$ be a $\mathrm{Tr}(B_t)$-free hypergraph and let $\mathcal{H}_1$ be the subhypergraph consisting of the hyperedges having a light subedge. Let G_1 be a graph obtained by taking a light subedge

for each hyperedge in $\mathcal{H}_1$. Let $\mathcal{H}_2$ denote the rest of the hyperedges of $\mathcal{H}$ and G_2 denote the shadow graph of $\mathcal{H}_2$. Let G denote the shadow graph of $\mathcal{H}$.

Claim 2.6. $|E(\mathcal{H}_2)|= o(n^2)$.

Proof of Claim. G_2 contains $O(n^{t+1})$ copies of B_t by Proposition 2.1. Hence by the removal lemma [6] there is a set A of $o(n^2)$ edges in G_2 such that each copy of B_t contains an edge from A.

 We claim that each edge of G_2 is in at most $3t-3$ hyperedges of $\mathcal{H}_2$. We use a lemma of Luo and Spiro [19], who showed the analogous statement in the case of forbidden $\mathrm{Tr}(K_{2,t})$. More precisely, they showed that if an edge xy is in at least $3t-2$ hyperedges of $\mathcal{H}_2$, then there is a $\mathrm{Tr}(K_{2,t})$ in $\mathcal{H}$ such that x and y are the vertices in the smaller part of the core $K_{2,t}$. To find a $\mathrm{Tr}(B_t)$ in $\mathcal{H}$, we need to find a hyperedge containing x, y such that the third vertex of that hyperedge is not in this core $K_{2,t}$. This is doable if $t > 1$, since we can pick any of the $2t-2$ hyperedges containing x, y and avoiding the other t vertices of the core $K_{2,t}$.

 Now we are ready to count the number of triangles in G_2. There are at most $(3t-3)|A|= o(n^2)$ triangles containing an edge in A. The rest of G_2 is B_t-free, thus contains $o(n^2)$ triangles by a result of Alon and Shikhelman [1]. As each hyperedge of $\mathcal{H}_2$ creates a triangle in G_2, we have $|E(\mathcal{H}_2)|= o(n^2)$, completing the proof. ∎

Claim 2.7. *For any edge uv in G, u and v have at most $3t-2$ common neighbors in G_1.*

Proof of Claim. Let us assume that $w_1, \ldots, w_{3t-1}$ are each adjacent to both u and v in G_1. There is a hyperedge in $\mathcal{H}$ containing u, v and a third vertex, which is not among $w_1, \ldots, w_{3t-2}$ without loss of generality. The hyperedge of $\mathcal{H}_1$ containing uw_1 has a third vertex w_1'. Then w_1' is not v, since then we would not pick vw_1 as an edge of G_1. If $w_1' = w_i$, we can assume without loss of generality that $i = 2$. Similarly, the third vertex of the hyperedge containing vw_1 is not u and if it is w_i, we can assume that $i \leq 3$.

 Then we consider the third vertices of the hyperedges containing uw_3 and vw_3, analogously we can assume that none of them are w_i with $i > 6$. We continue this way, the third vertices of the hyperedges containing uw_{3j-2} are not w_i with $i > 3j$. Note that those vertices are not of the form w_{3k-2} with $k < j$ either, since we have already studied the single hyperedge containing uw_{3k-2} and the single hyperedge containing uw_{3k-2}, and the third vertex is not w_{3j-2}. Then the B_t with vertices u, v and $w_1, w_4, \ldots, w_{3t-2}$ is the core of a $Tr(B_t)$, a contradiction completing the proof. ∎

 Let us return to the proof of the theorem. The above claim implies that G_1 is B_{3t-1}-free (it is easy to see that G_1 is actually B_t-free). Observe that $|E(\mathcal{H}_1)|= |E(G_1)|$, hence we are done by Claim 2.7 unless $|E(G_1)|\geq n^2/4 - o(n^2)$. Therefore, by the Erdős-Simonovits stability theorem [4, 5, 28] we can obtain a complete bipartite graph with parts A and B, by adding and deleting $o(n^2)$ edges. We may assume A and B form a maximum cut in G_1, and thus each vertex has no more neighbors in its own part than in the other part.

8

Our aim now is to refine the estimates on both $|E(\mathcal{H}_2)|$ and $|E(G_1)|$ sufficient to establish the theorem. With that in mind, we let ζ, ε be small but fixed positive constants, for which ζ is chosen to be small relative to ε.

We then partition the vertices of G into the following five sets: let $A' \subset A$ denote the set of vertices in A that are adjacent in G_1 to all but at most ζn vertices of B; we define $B' \subseteq B$ analogously. We next let $A'' \subseteq A \setminus A'$ consist of the vertices adjacent to at least $n/4 + \zeta n$ vertices of B' in G_1, and $B'' \subseteq B \setminus B'$ consist of the vertices adjacent to at least $n/4 + \zeta n$ vertices of A' in G_1. Finally, let $X = V(G) \setminus (A' \cup A'' \cup B' \cup B'')$ consist of all remaining vertices.

Since $|E(G_1)| \geq n^2/4 - o(n^2)$ and since $o(n^2)$ edges are missing between A and B or lie within A or B, we have that $|A| - |A'| = o(n)$, $|B| - |B'| = o(n)$, and both $|A'|$ and $|B'|$ lie within $o(n)$ of $n/2$. Consequently, every pair of vertices in $A' \cup A''$ and in $B' \cup B''$ has $\Omega(n)$ common neighbors in G_1, hence both are independent sets in G by Claim 2.7. Moreover, we may also assume that $|X \cup A'' \cup B''| \leq \zeta n$.

Let $x = |X|$.

Claim 2.8. $|E(G_1)| \leq n^2/4 + x(-n/4 + 3\zeta n)$.

Proof of Claim. Let $u \in X$: trivially u has at most ζn neighbors in $X \cup A'' \cup B''$. If u has a neighbor $v \in A'$ then u can be adjacent to the at most ζn non-neighbors of v in B' along with at most $3t - 2$ common neighbors by Claim 2.7. Similarly, if u has a neighbor in B' then it has at most $\zeta n + 3t - 2$ neighbors in A'. Consequently, if u has neighbors in both A' and B' then it has less than $3\zeta n$ such neighbors in total, whereas if it only has neighbors in one of A', B', it has at most $n/4 + \zeta n$ by definition of X. Combining these estimates gives u's degree is at most $n/4 + 2\zeta n$.

Since $G_1 - X$ is bipartite, it follows that

$$|E(G_1)| \leq \frac{(n-x)^2}{4} + x \cdot \left(\frac{n}{4} + 2\zeta n \right) = \frac{n^2}{4} + x \left(-\frac{n}{4} + \frac{x}{4} + 2\zeta n \right) < \frac{n^2}{4} + x \left(-\frac{n}{4} + 3\zeta n \right)$$

∎

Claim 2.9. *Provided ζ is suitably small compared to ε, $|E(\mathcal{H}_2)| \leq x \cdot 2t\varepsilon n$*

Proof. Every edge of $\mathcal{H}_2$ has at least one vertex in X: if we choose some $u \in X$ and some $v \in A'' \cup B'' \cup X$ adjacent to u in G_2, then the edge uv is contained in at most $3t - 2$ edges of $\mathcal{H}_2$. This gives us a bound of $x(3t - 2)\zeta n$ on the number of edges of $\mathcal{H}_2$ with at least two vertices in $A'' \cup B'' \cup X$.

To count the remaining hyperedges of $\mathcal{H}_2$, for each $u \in X$ let $\mathcal{H}_2(u)$ denote the family of edges of $\mathcal{H}_2$ containing u, a vertex in A', and a vertex in B'. We then let $A(u)$ denote the set of vertices in A' and $B(u)$ denote the set of vertices in B' that are in a hyperedge in $\mathcal{H}_2(u)$. Finally, let $u_1, \ldots, u_x$ be the vertices of X, and color every edge vw between $A' \cup B'$ with the subset of $\{1, \ldots, x\}$ for which $vwu_i \in \mathcal{H}$ (either $\mathcal{H}_1$ or $\mathcal{H}_2$).

Now assume that for $v \in A(u_i)$ there exists $w_1, w_2, \ldots, w_{t+1} \in B(u_i)$ such that vw_j's color is not $\{i\}$. We claim that there is a $Tr(B_t)$ with u_i and v as vertices in the smaller part. For the edge $u_i v$ we use an arbitrary hyperedge containing u_i and v, without loss of generality the third vertex is not among $w_1, \ldots, w_t$. For the edge vw_j, we use a hyperedge containing v and w_j and a vertex in X other than u_i, which exists since vw_j's color includes an index other than i. Finally, since each edge $u_i w_j$ lies in an edge of $\mathcal{H}_2$ (by definition of $B(u_i)$), $u_i w_j$ is heavy and is therefore contained in an edge of $\mathcal{H}$ that does not include v, and the third vertex of that edge cannot be in B' (thus cannot be w_ℓ) as B' is independent in G.

By construction, $|\mathcal{H}_2(u_i)|$ is bounded above by the number of heavy edges in the induced subgraph $G[A(u_i) \cup B(u_i)]$, and the argument in the preceding paragraph and symmetry show the heavy edges form a subgraph with maximum degree at most t. We now partition X into sets L and S as follows: let S contain all vertices u_i for which $\max\{|A(u_i)|, |B(u_i)|\} < \varepsilon n$, and let $L = X \setminus S$.

If $u_i \in S$ then $G[A(u_i) \cup B(u_i)]$ contains at most $t \cdot \min\{|A(u_i)|, |B(u_i)|\} \leq t \cdot \varepsilon n$ heavy edges, yielding that $|\bigcup_{u \in S} \mathcal{H}_2(u)| \leq |S| \cdot t \varepsilon n$. Similarly, if $u \in L$ then $G[A(u_i) \cup B(u_i)]$ contains at most $tn/2$ heavy edges, so $|\bigcup_{u \in L} \mathcal{H}_2(u)| \leq |L| \cdot tn/2$. The last step in our argument is to show that $|L| \leq \varepsilon x$: this is trivial if $L = \varnothing$ so we assume otherwise.

Let $u_i \in L$: each vertex $v \in A(u_i)$ is adjacent to at least $|B(u_i)| - t - \zeta n \geq \varepsilon n - t - \zeta n > \varepsilon n/2$ vertices of $B(u_i)$ with edges of color $\{i\}$, as $\zeta < \varepsilon/2$. Thus, the subgraph $G[A(u_i) \cup B(u_i)]$ contains at least $|A(u_i)| \cdot \varepsilon n/2 \geq \varepsilon^2 n^2/2$ edges with color $\{i\}$, which all lie inside G_1. Since $|E(G_1)| \leq n^2/4$, it follows that $|L| \leq (2\varepsilon^2)^{-1}$.

Fix $u_i \in L$ and $u_j \in X$, $j \neq i$: let vw be a heavy edge in $G[A(u_i) \cup B(u_i)]$ whose label includes j, where we assume $v \in A'$ and $w \in B'$. If $vwu_j \in \mathcal{H}_2$ then vw is an edge in $G[A(u_j) \cup B(u_j)]$. All such edges must lie in the intersection $G_{i,j} = G[(A(u_i) \cap A(u_j) \cup (B(u_i) \cap B(u_j))]$. The earlier arguments show the maximum degree in $G_{i,j}$ is at most $2t$ (else some vertex lies on $t + 1$ edges with a label other than i or other than j). The definition of A' and B' implies that $G_{i,j}$ has at most $2t + \zeta n$ vertices in each part, and therefore there are fewer than $(2t + \zeta n)(2t) < 4t\zeta n$ such heavy edges.

If $vwu_j \notin \mathcal{H}_2$, then it lies in $\mathcal{H}_1$ and one of vu_j, wu_j lies in G_1. But Claim 2.7 implies u_j has less than $\zeta n + 3t - 2 < 2\zeta n$ neighbors in A' and in B', since it has neighbors in both parts in G. Since the neighbor in G_1 determines the heavy edge, we have at most $4\zeta n$ such heavy edges.

Thus, j occurs in the color of at most $4(t + 1)\zeta n$ heavy edges in $G[A(u_i) \cup B(u_j)]$, and therefore there must be at least $|A(u_j)|/(4(t + 1)\zeta n) \geq \varepsilon/(4(t + 1)\zeta)$ indices $j \neq i$. In other words, $x - 1 \geq \varepsilon/(4(t + 1)\zeta)$. Thus, provided $\zeta \leq \varepsilon^4/(2(t + 1))$, we have

$$|L| \leq \frac{1}{2\varepsilon^2} \leq \varepsilon \cdot \frac{\varepsilon}{4(t + 1)\zeta} \leq \varepsilon x.$$

Combining the estimates, it follows that

$$|E(\mathcal{H}_2)| \leq x(3t - 2)\zeta n + |S| t \varepsilon n + |L| tn/2 \leq x\varepsilon n + xt\varepsilon n + \varepsilon xtn/2 \leq x \cdot 2t\varepsilon n.$$

∎

Provided ε and ζ are small enough, Claims 2.8 and 2.9 imply

$$|E(\mathcal{H})| \leq n^2/4 + x(-n/4 + 3\zeta n + 2t\varepsilon n) < (n-1)^2/4 - \Omega(n)$$

if $x > 2$. We suppose now that $x = 2$ and that $|E(\mathcal{H})| \geq (n-1)^2/4 - \varepsilon n$ and aim to produce a contradiction. Since $(n-1)^2/4 = (n-2)^2/4 + (2n-3)/4$ and ε is small, by Claim 2.9 we may assume that $|E(G_1)| > (n-2)^2/4 + 3n/8$. Since $G - X$ is bipartite, at least $3n/8$ edges of G_1 are incident with $X = \{u_1, u_2\}$, so without loss of generality we suppose that $d_{G_1}(u_2) \geq 3n/16$ and that $u_2 \in A$. Since $|X \cup A'' \cup B''| \leq \zeta n$, it follows that u_2 has $\Omega(n)$ neighbors in B' in G_1.

But then Claim 2.7 implies u_2 has no neighbors in A' in G. Consequently, every edge of G_1 in $A' \cup B'$ must form a triple in $\mathcal{H}$ with u_1, implying u_1 is adjacent in G to $A' \cup B'$. This, in turn, implies u_1 has fewer than, say, $2\zeta n$ neighbors in A' or in B' in G_1 (by Claim 2.7 again). Thus, $d_{G_1}(u_1) < 3\zeta n$, so u_2 has at least

$$d_{G_1}(u_2) - \zeta n \geq \frac{3n}{8} - 4\zeta n > \frac{n}{4} + \zeta n$$

neighbors in B'. But then $u_2 \in A''$ by definition rather than X, the desired contradiction.

Therefore, if $|E(\mathcal{H})| \geq (n-1)^2/4 - \varepsilon n$ then we must have $x = 1$, implying $\mathcal{H} = \mathcal{H}_1 \leq \lfloor (n-1)^2/4 \rfloor$, and that equality only holds when $\mathcal{H}$ is formed by adding a new vertex to every edge of a maximum bipartite graph on $n - 1$ vertices. ∎

References

[1] N. Alon, C. Shikhelman. Many T copies in H-free graphs. *Journal of Combinatorial Theory, Series B*, **121**, 146–172, 2016.

[2] C. Berge, Hypergraphes: combinatoire des ensembles finis, *Gauthier-Villars*, 1987.

[3] J. Cutler, J.D. Nir, A.J. Radcliffe, (2022). Supersaturation for subgraph counts. *Graphs and Combinatorics*, **38**(3), 65.

[4] P. Erdős. Some recent results on extremal problems in graph theory, *Theory of Graphs* (Internl. Symp. Rome), 118–123, 1966.

[5] P. Erdős. On some new inequalities concerning extremal properties of graphs, in Theory of Graphs (ed P. Erdős, G. Katona), Academic Press, New York, 77–81, 1968.

[6] P. Erdős, P. Frankl, 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.

[7] P. Erdős, M. Simonovits. A limit theorem in graph theory. *Studia Sci. Math. Hungar.* **1**, (1966), 51–57.

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

[9] P. Erdős, A. H. Stone. On the structure of linear graphs. *Bulletin of the American Mathematical Society* **52**, (1946), 1087–1091.

[10] P. Frankl, J. Pach. On disjointly representable sets, *Combinatorica* **4** (1984), 39-45.

[11] Z. Füredi, R. Luo. Induced Turán problems and traces of hypergraphs. *European Journal of Combinatorics*, 103692, 2023.

[12] Z.Füredi, M. Simonovits, (2013). The history of degenerate (bipartite) extremal graph problems. In Erdős centennial (pp. 169-264). Berlin, Heidelberg: Springer Berlin Heidelberg.

[13] D. Gerbner, On non-degenerate Turán problems for expansions, *European Journal of Combinatorics*, 124, 2025, 104071

[14] D. Gerbner, Z.L. Nagy, M. Vizer, (2022). Unified approach to the generalized Turán problem and supersaturation. *Discrete Mathematics*, 345(3), 112743.

[15] D. Gerbner, C. Palmer, Extremal Results for Berge hypergraphs. *SIAM Journal on Discrete Mathematics*, **31**, 2314–2327, 2017.

[16] D. Gerbner, C. Palmer. Counting copies of a fixed subgraph in F-free graphs. *European Journal of Combinatorics* **82**, 103001, 2019.

[17] D. Gerbner, B. Patkós, Extremal Finite Set Theory, 1st Edition, CRC Press, 2018.

[18] A. Halfpap, C. Palmer, (2021). On supersaturation and stability for generalized Turán problems. Journal of Graph Theory, 97(2), 232-240.

[19] R. Luo, S. Spiro, (2021). Forbidding $K_{2,t}$ Traces in Triple Systems. The Electronic Journal of Combinatorics, P2.4.

[20] W. Mantel, Problem28, Wiskundige Opgaven 10 (1907), 60–61.

[21] D. Mubayi, J. Verstraëte. A survey of Turán problems for expansions. *Recent Trends in Combinatorics*, 117–143, 2016.

[22] D. Mubayi, Y. Zhao, (2007). Forbidding complete hypergraphs as traces. Graphs and Combinatorics, 23(6), 667-679.

[23] O. Pikhurko. Exact computation of the hypergraph Turán function for expanded complete 2-graphs, *Journal of Combinatorial Theory, Series B*, **103**(2) 220–225, 2013.

[24] B. Qian, G. Ge, (2022). A note on the Turán number for the traces of hypergraphs. arXiv preprint arXiv:2206.05884.

[25] N.Sauer, On the density of families of sets, J. Combinatorial Theory Ser.A 13(1972), 145–147.

[26] S. Shelah, A combinatorial problem; stability and order for models and theories in infinitary languages, Pacifc J.Math. 41 (1972), 247–261.

[27] M. Simonovits, A method for solving extremal problems in graph theory, stability problems. In Theory of Graphs (Proc. Colloq., Tihany, 1966), pp. 279–319. 1968.

[28] M. Simonovits. Extremal graph problems with symmetrical extremal graphs. Additional chromatic conditions, *Discrete Math.* **7**, 349–376, 1974.

[29] P. Turán. Egy gráfelméleti szélsőértékfeladatról. *Mat. Fiz. Lapok*, **48**, 436–452, 1941.

[30] V.N. Vapnik, A.Ya. Chervonenkis, On the uniform convergence of relative frequencies of events to their probabilities, Theory Probab. Appl. 16 (1971), 264–280.

[31] A. A. Zykov. On some properties of linear complexes. *Matematicheskii Sbornik*,**66**(2), 163–188, 1949.