On extremal problems concerning the traces of sets

Piga, Simón; Schülke, Bjarne

[Link to publication on Research at Birmingham portal](#)

ON EXTREMAL PROBLEMS CONCERNING THE TRACES OF SETS

SIMÓN PIGA AND BJARNE SCHÜLKE

ABSTRACT. Given two non-negative integers n and s, define $m(n, s)$ to be the maximal number such that in every hypergraph $\mathcal{H}$ on n vertices and with at most $m(n, s)$ edges there is a vertex x such that $|\mathcal{H}_x| \geqslant |E(\mathcal{H})| - s$, where $\mathcal{H}_x = \{H \smallsetminus \{x\} : H \in E(\mathcal{H})\}$. This problem has been posed by Füredi and Pach and by Frankl and Tokushige. While the first results were only for specific small values of s, Frankl determined $m(n, 2^{d-1} - 1)$ for all $d \in \mathbb{N}$ with $d \mid n$. Subsequently, the goal became to determine $m(n, 2^{d-1} - c)$ for larger c. Frankl and Watanabe determined $m(n, 2^{d-1} - c)$ for $c \in \{0, 2\}$. Other general results were not known so far.

Our main result sheds light on what happens further away from powers of two: We prove that $m(n, 2^{d-1} - c) = \frac{n}{d}(2^d - c)$ for $d \geqslant 4c$ and $d \mid n$ and give an example showing that this equality does not hold for $c = d$. The other line of research on this problem is to determine $m(n, s)$ for small values of s. In this line, our second result determines $m(n, 2^{d-1} - c)$ for $c \in \{3, 4\}$. This solves more instances of the problem for small s and in particular solves a conjecture by Frankl and Watanabe.

§1. INTRODUCTION

A hypergraph $\mathcal{H}$ is a pair $(V, \mathcal{F})$ where V is the set of vertices and $\mathcal{F} \subseteq 2^V$ is the set of edges. In the literature, the problems we consider in this article are often presented in the context of families rather than hypergraphs. If not necessary, it is then not distinguished between the family $\mathcal{F} \subseteq 2^V$ and the hypergraph $(V, \mathcal{F})$. We will follow this notational path.

Let V be an n-element set and let $\mathcal{F}$ be a family of subsets of V. For a subset T of V define the *trace* of $\mathcal{F}$ on T by $\mathcal{F}_{|T} = \{F \cap T : F \in \mathcal{F}\}$. For integers n, m, a, and b, we write

$$(n, m) \rightarrow (a, b)$$

if for every family $\mathcal{F} \subseteq 2^V$ with $|\mathcal{F}| \geqslant m$ and $|V| = n$ there is an a-element set $T \subseteq V$ such that $|\mathcal{F}_{|T}| \geqslant b$ (we also say that (n, m) *arrows* (a, b)).

The first type of question that was asked for this arrowing notation is similar to the spirit of the classic Turán problem: For a fixed number of vertices n, how many edges are needed such that there is a subset of vertices such that all its subsets lie in the trace. The following result on this question was conjectured by Erdős [5] and was proved independently by Sauer [8], Shelah and Perles [9], and Vapnik and Červonenkis [10]. It states that for a large family $\mathcal{F}$ on n vertices, there is an s-set of vertices such that all its subsets lie in the trace of $\mathcal{F}$. More precisely, they showed that $(n, m) \rightarrow (s, 2^s)$ whenever $m > \sum_{0 \leqslant i < s} \binom{n}{i}$.

Another fundamental question that was raised in the area is how large a family can be at most so that there will still be a vertex v such that the trace on $V \smallsetminus \{v\}$ is not much smaller than the original family. More precisely, the following problem was posed by Füredi and Pach [5] and, more recently, by Frankl and Tokushige as Problem 3.8 in their monograph [3][1]:

Problem 1.1. *Given non-negative integers n and s, what is the maximum value $m(n, s)$ such that for every $m \leqslant m(n, s)$ we have*

$$(n, m) \rightarrow (n - 1, m - s).$$

As described in the abstract, this problem can also be formulated as finding the maximal number $m(n, s)$ such that the following holds. In every hypergraph $\mathcal{H}$ with some n-set V as vertex set and with at most $m(n, s)$ edges there is a vertex x such that $|\mathcal{H}_x| \geqslant |\mathcal{H}| - s$, where $\mathcal{H}_x = \mathcal{H}_{|V \smallsetminus \{x\}} = \{H \smallsetminus \{x\} : H \in \mathcal{H}\}$.

A family $\mathcal{F}$ is *hereditary* if for every $F' \subseteq F \in \mathcal{F}$ we have that $F' \in \mathcal{F}$. In [2] Frankl proves that among families with a fixed number of edges and vertices, the trace is minimised by hereditary families. Thus, the problems considered here, and in particular Problem 1.1, can be reduced to hereditary families (see Lemma 2.1). Note that in hereditary families, Problem 1.1 is asking for the maximum number of edges such that there is always a vertex of small degree (as usual, we define the degree of a vertex v as the number of edges that contain v).

The investigation of this problem started with Bondy [1] and Bollobás [6] determining $m(n, 0)$ and $m(n, 1)$, respectively. Later Frankl [2] and Frankl and Watanabe [4] proved part (1) and (2), respectively, of the following theorem.

Theorem 1.2. *For $d, n \in \mathbb{N}$ and $d|n$, we have*

(1) $m(n, 2^{d-1} - 1) = \frac{n}{d}(2^d - 1)$,

(2) $m(n, 2^{d-1} - 2) = \frac{n}{d}(2^d - 2)$.

---

[1]There have been slightly different versions in use for the arrowing notation and for what we denote by $m(n, s)$. In this work, we follow the notation in [3].

Consider a family consisting of a set of size d and all possible subsets, and take n/d vertex disjoint copies of it. The resulting family has minimum degree 2^{d-1} and $\frac{n}{d}(2^d - 1) + 1$ edges. Thus, this family is an extremal construction for (1). By taking out all sets of size d, we obtain an extremal construction for (2).

Our main result makes further progress on Problem 1.1, solving it for general $s = 2^{d-1} - c$ as long as c is linearly small in d.

Theorem 1.3 (Main theorem). *Let $d, c, n \in \mathbb{N}$ with $d \geqslant 4c$ and $d|n$. Then*

$$m(n, 2^{d-1} - c) = \frac{n}{d}(2^d - c).$$

Remark 1.4. In fact, our proof of Theorem 1.3 yields that for $d \geqslant 4c$ and $m \leqslant \frac{n}{d}(2^d - c)$ we have $(n, m) \to (n - 1, m - (2^{d-1} - c))$ without any divisibility conditions on n. The assumption $d|n$ is only necessary for the extremal constructions showing the maximality of $\frac{n}{d}(2^d - c)$. Analogous remarks hold for Theorem 1.2 above and Theorem 1.5 below. In Section 5 we provide a construction showing that the equality in Theorem 1.3 does not hold for $d = c$ (see Construction 5.2).

One might also try to solve Problem 1.1 for small values of s. Apart from the aforementioned results by Bondy and Bollobás, progress was made by Frankl [2], Watanabe [11, 12], and by Frankl and Watanabe [4]. In [4] they conjectured that $m(n, 12) = (28/5 + o(1))n$. Theorem 1.3 does not consider cases for which d is very small in terms of c. The following results extend Theorem 1.2 to $c = 3$ and 4 and every $d \geqslant 3$ (for smaller d the respective $m(n, s)$ is not defined). In particular, it proves the conjecture of Frankl and Watanabe for $s = 12$ in a strong sense.

Theorem 1.5. *Let $d, n \in \mathbb{N}$ with $d \geqslant 3$ and $d|n$. Then*
  (1) $m(n, 2^{d-1} - 3) = \frac{n}{d}(2^d - 3)$ and
  (2) $m(n, 2^{d-1} - 4) = \frac{n}{d}(2^d - 4)$. In particular, $m(n, 12) = \frac{28}{5}n$.

Note that for larger d, this theorem is of course a special case of Theorem 1.3.

1.1. **Idea of the proof.** To show the maximality of $\frac{n}{d}(2^d - c)$ we give a construction similar to the one presented after Theorem 1.2. As mentioned above, Lemma 2.1 reduces the problem to a problem in hereditary families. We need to show that for every hereditary hypergraph $\mathcal{F}$ on n vertices with minimum degree at least $2^{d-1} - c + 1$ we have that

$$|\mathcal{F}| \geqslant \frac{n}{d}(2^d - c) + 1.$$

In the proof of Theorem 1.2 [2, 4] the equality $|\mathcal{F} \smallsetminus \{\varnothing\}| = \sum_{v \in V} \sum_{H \in L_v} \frac{1}{|H|+1}$, where L_v is *the link of the vertex v*, was used, which comes from a simple double counting argument.

Subsequently, they used a generalised form of the Kruskal-Katona Theorem (see Theorem 2.2 below) to obtain a general lower bound for $\sum_{H \in L_v} \frac{1}{|H|+1}$ for every v. Due to the aforementioned double counting this in turn yields a lower bound on the number of edges.

For $c \geqslant 3$ there exist extremal families which show that a general bound on $\sum_{H \in L_v} \frac{1}{|H|+1}$ for every vertex v is not sufficient to provide the desired bound on the number of edges. To overcome this difficulty first observe that the double counting argument can be generalised by interpreting $\sum_{H \in L_v} \frac{1}{|H|+1}$ as the weight $w_{\mathcal{F}}(v)$ of a vertex v. We will refer to this weight as *uniform weight* since it can be imagined as uniformly distributing the unit weight of an edge to each of its vertices. In contrast, to prove Theorem 1.3 and Theorem 1.5, we will use a non-uniform weight. Moreover, instead of bounding the weight of single vertices we will bound the weight of sets of vertices.

To this aim take the maximum possible set of vertices with small uniform weight such that their neighbourhoods are pairwise disjoint. Call these vertices together with their neighbours *clusters*. Note that in this way the neighbourhood of every vertex with small uniform weight needs to intersect some cluster. For bounding the weight of vertices whose neighbourhood does not intersect any cluster (and therefore have a large uniform weight), we introduce a "local" lemma (see Lemma 3.1) which is a close relative to a general form of the Kruskal-Katona theorem. Given a vertex of fixed degree, it provides a lower bound on the uniform weight and furthermore, the minimum weight surplus if its link deviates enough from the minimising link. Since the link of vertices whose neighbourhood does not intersect any cluster indeed deviates enough from the minimising link, the lemma then gives that these will have a large weight.

The next step is to bound the average weight of the vertices in each such cluster. Even if the number of edges inside a cluster is not large enough, $\mathcal{F}$ being hereditary and the minimum degree of $\mathcal{F}$ still provide some lower bound for the number of edges in a cluster. Then a second local lemma (Lemma 3.2) yields that there are several vertices within that cluster whose degree (with respect to the cluster) is not the minimum degree in $\mathcal{F}$. Therefore, there exist several crossing edges, i.e., edges containing vertices from both the inside and the outside of the cluster. If we use the uniform weight, these crossing edges will contribute enough to the weight of the cluster, even more than needed.

At this point we still need to bound the weight of vertices with small uniform weight lying outside of any cluster. As mentioned above, the neighbourhood of every such vertex intersects some cluster, meaning every such vertex is contained in a crossing edge. Recall that in fact, a uniform weight on crossing edges would contribute more weight than needed for the inside of a cluster. Now the second idea comes into play: the unit weight of these edges will be distributed non-uniformly among its vertices. Hence, when splitting the unit

weight of such a crossing edge according to the above mentioned imbalance, both sides will get a share that is big enough.

Note that this strategy is compatible with the extremal constructions in so far as that those are composed of disjoint copies of locally optimal families.

§2. PRELIMINARIES

In this work we consider the set of natural numbers $\mathbb{N}$ to start with 1 and the logarithms considered are to the base 2. Further, for $i \in \mathbb{Z}$ we set as usual $[i] = \{1, \ldots, i\}$, and it is also convenient to define $[i]_0 = \{0, \ldots, i\}$. Given a set $F \subseteq \mathbb{N}$ and some $i \in \mathbb{N}$, we denote by $F + i$ the set $\{j + i : j \in F\}$. For our considerations isolated vertices, i.e., vertices that are contained in the vertex set of a hypergraph but do not lie in any edges, usually do not play an important rôle. This will lead to a few easy peculiarities in notation. For two hypergraphs $\mathcal{H}$ and $\mathcal{H}'$ we write $\mathcal{H} \cong \mathcal{H}'$ if they are isomorphic up to isolated vertices, more precisely, if there are vertex sets V disjoint to $V(\mathcal{H})$ and V' disjoint to $V(\mathcal{H}')$ such that the hypergraph $(V(\mathcal{H}) \cup V, E(\mathcal{H}))$ is isomorphic to $(V(\mathcal{H}') \cup V', E(\mathcal{H}'))$.

For a hypergraph $\mathcal{H} = (V, E)$ and $v \in V$ we define the link L_v of v to be the hypergraph on V with edge set $\{F \smallsetminus \{v\} : F \in E\}$. Further, we write

$$V_v = \{w \in V : \text{there is an } e \text{ with } \{v, w\} \subseteq e \in E\},$$

note that if v is not an isolated vertex, then $v \in V_v$. This notation will be useful in the proof of Theorem 1.3 when defining the clusters mentioned in the overview of the proof.

The following lemma due to Frankl [2] provides the aforementioned reduction of Problem 1.1 to hereditary families.

Lemma 2.1. *For $n, m, a, b \in \mathbb{N}$ the following statements are equivalent.*

    (1) For every n-set V and every hereditary family $\mathcal{F} \subseteq 2^V$ with $|\mathcal{F}| \geq m$, there exists a set $T \subseteq V$ with $|T| = a$ such that $|\mathcal{F}_{|T}| \geq b$.

    (2) $(n, m) \to (a, b)$

In particular, this means that in the proof of our results we only need to consider hereditary families.

Let $n \in \mathbb{N}$, for $A, B \in 2^{[n]}$ we say that $A <_{col} B$ or A precedes B in the *colexicographic order* if $\max(A \triangle B) \in B$. Let $m \in \mathbb{N}$ with $m \leq 2^n$ and define $\mathcal{R}_n(m)$ to be the family on n vertices containing the first m sets of $2^{[n]}$ according to the colexicographic order. Note that for $n \leq n'$ and $m \leq 2^n$, we have $\mathcal{R}_n(m) \cong \mathcal{R}_{n'}(m)$ and hence, we will not distinguish between $\mathcal{R}_n(m)$ and $\mathcal{R}_{n'}(m)$ and we will omit the subscript. The following theorem due to Katona [7] is a generalisation of the well-known Kruskal-Katona theorem.

Theorem 2.2. *Let $f : \mathbb{N}_0 \to \mathbb{R}$ be a monotone non-increasing function and let $\mathcal{F}$ be a hereditary family with $|\mathcal{F}| = m$. Then*

$$\sum_{F \in \mathcal{F}} f(|F|) \geqslant \sum_{R \in \mathcal{R}(m)} f(|R|).$$

In the proofs of Theorems 1.3 and 1.5, normally Theorem 2.2 is applied with $\mathcal{F}$ being the link of a vertex. Moreover, as we usually consider the uniform weight mentioned in Section 1.1, the function f will often be $f(k) = \frac{1}{k+1}$. The weight of $\mathcal{R}(m)$ with respect to this f will come up repeatedly and hence, for brevity we set $W(m) := \sum_{R \in \mathcal{R}(m)} \frac{1}{|R|+1}$. Note that we have $W(2^{d-1}) = \frac{2^d - 1}{d}$ and further the following estimate[2] for $W(2^{d-1} - c)$ for a $c \in [2^{d-2}]$:

$$W(2^{d-1} - c) \geqslant \frac{2^d - 1}{d} - \frac{c}{d - \log c} \tag{2.1}$$

Indeed, if $A \in 2^{[d-1]} \smallsetminus \mathcal{R}(2^{d-1} - c)$, then there are at least $2^{d-1-|A|}$ sets in $2^{[d-1]} \smallsetminus \mathcal{R}(2^{d-1} - c)$. Thus, it follows that for every $A \in 2^{[d-1]} \smallsetminus \mathcal{R}(2^{d-1} - c)$ we have $|A| \geqslant d - 1 - \log c$. This gives that $W(2^{d-1}) - W(2^{d-1} - c) \leqslant \frac{c}{d - \log c}$ and thereby (2.1).

§3. Proof of Theorem 1.3

As mentioned in Section 1.1 for proving Theorem 1.3 we introduce two "local" lemmas. The first lemma says that if a family deviates enough from $\mathcal{R}(m)$, the weight of this family will have a surplus with respect to $W(m)^2$.

Lemma 3.1. *Let $d \geqslant 4$ and $c \leqslant 2^d$ be integers. For a hereditary family $\mathcal{H}$, with $|\mathcal{H}| \geqslant 2^d - c$ the following holds.*

(1) $\sum_{H \in \mathcal{H}} \frac{1}{|H|+1} \geqslant W(2^d - c)$.

(2) If there are at least $d + 1$ non isolated vertices in $\mathcal{H}$, then

$$\sum_{H \in \mathcal{H}} \frac{1}{|H| + 1} \geqslant W(2^d - c) + \frac{1}{6}.$$

(3) If $c \in \{2, 3\}$ and $\mathcal{H} \not\cong \mathcal{R}(2^d - c)$, then we have

$$\sum_{H \in \mathcal{H}} \frac{1}{|H| + 1} \geqslant W(2^d - c) + \min\left(\frac{1}{6}, \frac{1}{d}\right).$$

Proof. Let d, n, c, and $\mathcal{H}$ be given as in the statement. The first part follows by applying Theorem 2.2 with $f(k) = \frac{1}{k+1}$.

---

[2]To have a clearer presentation of our main results and their proofs, we refrained from striving for optimal bounds.

In order to prove part (2) and (3) we need some preparation. Denote by h_i and r_i the number of i-sets in $\mathcal{H}$ and $\mathcal{R}(2^d - c)$, respectively. Given $s \in [d]_0$ set $g(k) = 1$ for $k \leqslant s$ and $g(k) = 0$ for $k > s$. Then Theorem 2.2 applied with $f = g$ yields

$$\sum_{i \in [s]_0} h_i \geqslant \sum_{i \in [s]_0} r_i. \tag{3.1}$$

Next, let $H_1, \ldots, H_{|\mathcal{H}|}$ be an enumeration of the elements of $\mathcal{H}$ such that $|H_j| \leqslant |H_{j+1}|$. Given $i \in [d-1]$ let $\varphi(i)$ be the number of edges of size at most i in the family $\mathcal{R}(2^d - c)$, i.e., $\varphi(i) = \sum_{j \in [i]_0} r_j$. Let $\mathcal{H}_0 = \{H_1\} = \{\varnothing\}$ and for $i \in [d-1]$ consider the following set of edges $\mathcal{H}_i = \{H_{\varphi(i-1)+1}, \ldots, H_{\varphi(i)}\}$ and observe that its size is r_i. Inequality (3.1) implies that for $H \in \mathcal{H}_i$, where $i \in [d-1]_0$, we have $|H| \leqslant i$. Thus,

$$\sum_{i \in [d-1]_0} \sum_{H \in \mathcal{H}_i} \frac{1}{|H| + 1} \geqslant \sum_{i \in [d-1]_0} \frac{r_i}{i + 1} = W(2^d - c). \tag{3.2}$$

If now at least $d + 1$ vertices are contained in edges of $\mathcal{H}$, then even for $H_{d+2} \in \mathcal{H}_2$ it holds that $|H_{d+2}| = 1$. Hence, (3.2) now becomes $\sum_{H \in \mathcal{H}} \frac{1}{|H|+1} \geqslant \frac{1}{2} - \frac{1}{3} + W(2^d - c)$ and (2) is proved.

For proving (3), let $c \in \{2, 3\}$ and note that if there are at least $d + 1$ non isolated vertices in $\mathcal{H}$, then the result follows from (2). Thus, assume that there are only d non isolated vertices in $\mathcal{H}$. Observe that $r_i = \binom{d}{i}$ for $i \in [d-2]$, $r_d = 0$ and $r_{d-1} = d - (c-1)$. Hence, due to (3.1) we have $h_i = \binom{d}{i}$ for $i \in [d-2]$ and because of $\mathcal{H}$ being hereditary and the size of $\mathcal{H}$, further $h_{d-1} \geqslant d - (c-1)$. In fact, $h_{d-1} > r_{d-1} = d - (c-1)$ has to hold since $\mathcal{H} \not\cong \mathcal{R}(2^d - c)$. Together with (3.2) the result follows. $\qquad\square$

The following is the second local lemma mentioned in the overview. Part (2) states that a family on d vertices with high minimum degree contains many edges and therefore, considering Lemma 2.1, this is a local version of Theorem 1.3. Moreover, Part (1) states that if a family has not enough edges, then there are several vertices of low degree.

Lemma 3.2. *Let $d, c \in \mathbb{N}$, V be a d-set and let $\mathcal{H} \subseteq 2^V$ be hereditary.*

(1) If $|\mathcal{H}| \leqslant 2^d - c - 1$, then $\deg(v) \leqslant 2^{d-1} - c - 1$ for at least $d - c$ vertices v.

(2) If $d \geqslant c + 1$ and $\delta(\mathcal{H}) \geqslant 2^{d-1} - c$, then $|\mathcal{H}| \geqslant 2^d - c$.

Proof. (1): By $\bar{\mathcal{H}}$ denote the family $\{V \smallsetminus F : F \in 2^V \smallsetminus \mathcal{H}\}$. The bound on $|\mathcal{H}|$ implies that $c + 1 \leqslant |\bar{\mathcal{H}}|$ and observe that since $\mathcal{H}$ is hereditary, $\bar{\mathcal{H}}$ is hereditary. Consider some ordering $\bar{\mathcal{H}} = \{H_1, \ldots, H_{|\bar{\mathcal{H}}|}\}$ with $|H_i| \leqslant |H_{i+1}|$. Note that because $\bar{\mathcal{H}}$ is hereditary, we know that if some vertex $v \in V$ is contained in one of the edges $H_1, \ldots, H_j$, then in fact $\{v\} = H_i$ for some $i \in [j]$. Thus, there are $d - c$ vertices that do not lie in any

of $H_1, \ldots, H_{c+1}$. Note that these vertices lie in at least $c + 1$ sets of $2^V \smallsetminus \mathcal{H}$ and therefore, for each such v we have $\deg_{\mathcal{H}}(v) \leqslant 2^{d-1} - c - 1$.

(2): Assume for contradiction that $|\mathcal{H}| \leqslant 2^d - c - 1$. Then (1) gives the contradiction. $\qquad\square$

Now we are ready to prove Theorem 1.3.

Proof of Theorem 1.3. Let n, d, and c be given as in the theorem. First note that

$$\mathcal{F}_0 = \left\{ F + (i - 1) : F \in \mathcal{R}(2^d - (c - 1)) \text{ and } i \in \left[\frac{n}{d} \right] \right\} \subseteq 2^{[n]}$$

shows that for $m = \frac{2^d - c}{d} n + 1$, we have $(n, m) \twoheadrightarrow \left(n - 1, m - (2^{d-1} - c) \right)$.

In an hereditary family on n vertices with m edges the existence of a set of size $n - 1$ on which the trace of the family has size at least $m - (2^{d-1} - c)$ is equivalent to the existence of a vertex with degree at most $2^{d-1} - c$. Therefore, Lemma 2.1 implies that it is sufficient to show that for every hereditary family $\mathcal{F}$ on n vertices with minimum degree at least $2^{d-1} - c + 1$ we have $|\mathcal{F}| \geqslant \frac{2^d - c}{d} n + 1$. Let now $\mathcal{F} \subseteq 2^V$ be such a hereditary family on some n-set V in which every vertex has degree at least $2^{d-1} - c + 1$.

To prove the lower bound on the number of edges, we will define a weight function w on V with the property that $1 + \sum_{v \in V} w_{\mathcal{F}}(v) \leqslant |\mathcal{F}|$. Subsequently, it will be enough to show that $\sum_{v \in V} w_{\mathcal{F}}(v) \geqslant \frac{2^d - c}{d} n$. Indeed, for $c = 1$ the weight function $\sum_{H \in L_v} \frac{1}{|H|+1}$ together with Lemma 3.1 provide this, so from now on we assume $c \geqslant 2$. Note however, that for this uniform weight and c large, in $\mathcal{F}_0$ there are vertices with weight below and above $\frac{2^d - c}{d}$. As mentioned in the overview, we overcome this difficulty by using non-uniform weights and by bounding the average weight of sets of vertices instead of bounding the weight of every single vertex.

To that aim, we will in the following consider a partition of V. Let us call a vertex $v \in V$ *light* if $|V_v| = d$. Further, let $\mathcal{L}$ be a maximum set of light vertices such that $V_v \cap V_{v'} = \varnothing$ for all $v, v' \in \mathcal{L}$ and call the sets V_v with $v \in \mathcal{L}$ *clusters*. Later, the weight of a vertex will be defined depending on how it relates to to these clusters. Moreover, call the vertices $u \in V \smallsetminus \bigcup_{v \in \mathcal{L}} V_v$ with $|V_u| > d$ *heavy* vertices and let $\mathfrak{H}$ be the set of all heavy vertices. The vertices in $\mathcal{L}$ will be distinguished further into two different types $\mathcal{L}_1$ and $\mathcal{L}_2$ as follows. Let $\mathcal{L}_1$ be the set of those vertices $v \in \mathcal{L}$ for which every vertex in V_v is only contained in edges of 2^{V_v}, that is

$$\mathcal{L}_1 = \{ v \in \mathcal{L} : \text{there is no } e \in \mathcal{F} \smallsetminus 2^{V_v} \text{ with } e \cap V_v \neq \varnothing \}.$$

Furthermore, let $\mathcal{L}_2$ be the set of those vertices $v \in \mathcal{L}$ for which there exists an $x \in V_v$ that is contained in an edge of $\mathcal{F} \smallsetminus 2^{V_v}$, in other words,

$$\mathcal{L}_2 = \{ v \in \mathcal{L} : \text{there is an } e \in \mathcal{F} \smallsetminus 2^{V_v} \text{ with } e \cap V_v \neq \varnothing \}. \tag{3.3}$$

Note that we have $\mathcal{L} = \mathcal{L}_1 \cup \mathcal{L}_2$. Lastly, we collect the remaining vertices in the set $\bar{\mathcal{L}} = V \setminus (\mathfrak{H} \cup \bigcup_{v \in \mathcal{L}} V_v)$. Thus, we have $V = \mathfrak{H} \cup \bigcup_{v \in \mathcal{L}_1} V_v \cup \bigcup_{v \in \mathcal{L}_2} V_v \cup \bar{\mathcal{L}}$.

Next, for each of the partition classes $\mathfrak{H}$, $\bigcup_{v \in \mathcal{L}_1} V_v$, $\bigcup_{v \in \mathcal{L}_2} V_v$, and $\bar{\mathcal{L}}$ the weights will be defined and we will show that the average weight in each partition class is bounded from below by $\frac{2^d - c}{d}$.

Assign the uniform weight $w_{\mathcal{F}}(u) = \sum_{H \in L_u} \frac{1}{|H|+1}$ to every heavy vertex $u \in \mathfrak{H}$. This definition and (2) from Lemma 3.1 give that every heavy vertex has weight at least

$$\frac{1}{6} + W(2^{d-1} - c + 1) \geqslant \frac{1}{6} + \frac{2^d - 1}{d} - \frac{c - 1}{d - \log(c-1)} \geqslant \frac{2^d - c}{d}, \tag{3.4}$$

where we used the bound (2.1) for the first inequality and $d \geqslant 4c$ and $\log x \leqslant \frac{2}{3}x$ for $x \geqslant 1$ for the second (recall that we can assume $c \geqslant 2$).

Given $v \in \mathcal{L}_1$, we have that $\mathcal{F}[V_v]$ is a family on d vertices with minimum degree at least $2^{d-1} - c + 1$. Thus, from Lemma 3.2 (2) (with $c - 1$ here in place of c there) it follows that $|\mathcal{F}[V_v]| \geqslant 2^d - c + 1$. Since summing the uniform vertex weights of all vertices of a family amounts to the number of non-empty edges in that family, assigning the uniform weight $w_{\mathcal{F}}(x) = \sum_{H \in L_x} \frac{1}{|H|+1}$ to every $x \in V_v$ yields

$$\frac{1}{d} \sum_{x \in V_v} w_{\mathcal{F}}(x) = \frac{|\mathcal{F}[V_v] \setminus \{\varnothing\}|}{d} \geqslant \frac{2^d - c}{d}. \tag{3.5}$$

Given $v \in \mathcal{L}_2$, the idea is that the vertices in V_v already have a relatively large uniform weight just taking into account the edges on V_v. Thus, they only need a smaller proportion of the weight of an edge that includes vertices outside of V_v. More precisely, we assign the weight

$$w_{\mathcal{F}}(x) = \sum_{H \in L_x} \frac{1}{|H| + 1} - |V_x \setminus V_v| \left(\frac{1}{2} - \frac{c - 1}{d - c} \right)$$

to every vertex $x \in V_v$. This definition can be understood as vertices in V_v basically having the uniform weight but then renouncing part of their uniform share of 2-uniform edges that cross from the inside of a cluster to the outside. Later, these crossing edges will contribute more than their uniform share to the outside vertex.

Of course, if $|\mathcal{F}[V_v]| \geqslant 2^d - c + 1$, then again the bound (3.5) follows for v directly by double counting and thus, we may assume that $|\mathcal{F}[V_v]| \leqslant 2^d - c$. Define the set C as the set of vertices $x \in V_v$ for which there exists some F_x with $x \in F_x \in \mathcal{F} \setminus 2^{V_v}$. Note that in fact, since $\mathcal{F}$ is hereditary, we may assume $|F_x| = 2$. Considering the minimum degree condition in $\mathcal{F}$ and applying Lemma 3.2 (1) to $\mathcal{F}[V_v]$ (with $c - 1$ here instead of c there) it follows that

$$|C| \geqslant d - c + 1. \tag{3.6}$$

Moreover, the minimum degree of $\mathcal{F}$ implies $d(v) \geqslant 2^{d-1} - c + 1$ and hence, $\mathcal{F}$ being hereditary gives that $|2^{V_v} \smallsetminus \mathcal{F}| \leqslant 2(c-1)$. Therefore, double counting the non-empty edges in $\mathcal{F}[V_v]$ yields

$$|\mathcal{F}[V_v] \smallsetminus \{\varnothing\}| = \sum_{x \in V_v} \sum_{H \in L_x \cap 2^{V_v}} \frac{1}{|H| + 1} \geqslant 2^d - 2c + 1. \tag{3.7}$$

Now, observe that the definition of the weight together with (3.7) and (3.6) give

$$\begin{aligned}
\frac{1}{d} \sum_{x \in V_v} w_{\mathcal{F}}(x) &\geqslant \frac{1}{d} \left(|\mathcal{F}[V_v] \smallsetminus \{\varnothing\}| + |C| \frac{c-1}{d-c} \right) \\
&\geqslant \frac{1}{d} \left(2^d - 2c + 1 + (d - c + 1) \frac{c-1}{d-c} \right) \\
&\geqslant \frac{2^d - c}{d}.
\end{aligned} \tag{3.8}$$

Lastly consider vertices from $\bar{\mathcal{L}}$. Recall that in particular, these vertices are light and could potentially have a too low weight if the uniform weight would be used. Note that by the maximality of $\mathcal{L}$, for every vertex $a \in \bar{\mathcal{L}}$ we can pick a $v(a) \in \mathcal{L}_2$ such that there exists an edge containing a and a vertex of $V_{v(a)}$. Since the vertices in $\bigcup_{v \in \mathcal{L}_2} V_v$ renounced their full share of some of those edges, the vertices in $\bar{\mathcal{L}}$ can be given a larger fraction. To be precise, the weight for $a \in \bar{\mathcal{L}}$ is defined as

$$w_{\mathcal{F}}(a) = \sum_{H \in L_a} \frac{1}{|H| + 1} + |V_a \cap V_{v(a)}| \left(\frac{1}{2} - \frac{c-1}{d-c} \right).$$

Lemma 3.1 (1) yields that

$$w_{\mathcal{F}}(a) \geqslant W(2^{d-1} - c + 1) + \frac{1}{2} - \frac{c-1}{d-c} \geqslant W(2^{d-1} - c + 1) + \frac{1}{6} \geqslant \frac{2^d - c}{d}, \tag{3.9}$$

where the second inequality follows from $d \geqslant 4c$ and the third follows as in (3.4). Observe that the definition of $w_{\mathcal{F}}$ implies $\sum_{x \in V} w_{\mathcal{F}}(x) \leqslant 1 + |\mathcal{F}|$ because the left-hand side counts every edge of $\mathcal{F}$ apart from the empty set at most once. Since (3.4), (3.5), (3.8), and (3.9) say that the average weight per vertex in $\mathcal{F}$ is at least $\frac{2^d-c}{d}$, the proof is complete. $\qquad \square$

§4. Proof of Theorem 1.5

This section is dedicated to the proof of Theorem 1.5. The proof is very similar to the proof of the main theorem just with some adaptions to obtain more precise bounds at certain points. Hence, we will omit some details that already appeared in the last section.

Proof of Theorem 1.5. Let $c \in \{3, 4\}$ and note that the cases $d = 3$ and 4 have been solved before, see [4] and [11], so assume $d \geqslant 5$. Firstly, the family $\mathcal{F}_0$ from the proof of Theorem 1.3 shows that for $m = \frac{n}{d}(2^d - c) + 1$, we have $(n, m) \nrightarrow (n - 1, m - (2^{d-1} - c))$.

Let now $\mathcal{F} \subseteq 2^V$ be a hereditary family on some n-set V in which every vertex has degree at least $2^{d-1} - c + 1$. In the following we will show that $|\mathcal{F}| \geq (2^d - c)\frac{n}{d} + 1$.

To gain more precision later, this time we call a vertex $v \in V$ *light* if $L_v \cong \mathcal{R}(2^{d-1}-(c-1))$. Again, let $\mathcal{L}$ be a maximum set of light vertices such that $V_v \cap V_{v'} = \varnothing$ for all $v, v' \in \mathcal{L}$. Call the vertices $u \in V \smallsetminus \bigcup_{v \in \mathcal{L}} V_v$ with $L_u \not\cong \mathcal{R}(2^{d-1} - (c - 1))$ *heavy* vertices. The sets $\mathcal{L}_i$, $\mathfrak{H}$, $\bar{\mathcal{L}}$ are defined similarly as in the proof of Theorem 1.3, just according to the different definitions of light and heavy vertices here.

Again we assign the uniform weight to every heavy vertex of $\mathcal{F}$. Note that then, due to Lemma 3.1 (3) and the structure of $\mathcal{R}(2^{d-1} - c)$ for $c \leq 4$, every heavy vertex has weight at least

$$\min\left(\frac{1}{6}, \frac{1}{d}\right) + W(2^{d-1} - (c-1)) \geq \min\left(\frac{1}{6}, \frac{1}{d}\right) + \frac{2^d - 1}{d} - \frac{(c-1)d - 1}{(d-1)d} \geq \frac{2^d - c}{d}.$$
(4.1)

For $v \in \mathcal{L}_1$ and $x \in V_v$ the weight is again defined as the uniform weight and as in the proof of Theorem 1.3, we obtain

$$\frac{1}{d} \sum_{x \in V_v} w_{\mathcal{F}}(x) \geq \frac{2^d - c}{d}.$$
(4.2)

To write the next weight definitions in a compact way, we define the following set

$$\mathcal{S} = \left\{ H \in \mathcal{F} : |H| = 3 \text{ and } H \cap \bigcup_{v \in \mathcal{L}_2} V_v, H \cap \bar{\mathcal{L}} \neq \varnothing \right\}$$

Note that $\mathcal{S}$ is the set of those edges of size 3 in $\mathcal{F}$ crossing from the inside of some V_v with $v \in \mathcal{L}_2$ to its outside and contain a vertex from $\bar{\mathcal{L}}$. For $v \in \mathcal{L}_2$ and a vertex $x \in V_v$, assign the weight $w_{\mathcal{F}}(x) = \sum_{H \in L_x} \frac{1}{|H|+1} - \frac{1}{9}|\{H \in L_x : H \cup \{x\} \in \mathcal{S}\}|$.

Claim 4.1. *For $v \in \mathcal{L}_2$ we have $\frac{1}{d} \sum_{x \in V_v} w_{\mathcal{F}}(x) \geq \frac{2^d - c}{d}$.*

We postpone the proof of this claim to the end of the section and first finish the proof of Theorem 1.5 using the claim.

For a vertex define the weight $a \in \bar{\mathcal{L}}$ as $w_{\mathcal{F}}(a) = \sum_{H \in L_a} \frac{1}{|H|+1} + \frac{1}{18}|\{H \in L_a : H \cup \{a\} \in \mathcal{S}\}|$. Note that by the maximality of $\mathcal{L}$, there exists a $v(a) \in \mathcal{L}_2$ such that there are an edge F and a vertex $x_a \in V_{v(a)}$ with $a, x_a \in F$. In fact, it is easy to check that since $L_a \cong \mathcal{R}(2^{d-1}-(c-1))$, the number of 2-sets in L_a that contain x_a is at least $d - 2 \geq 3$. Thus, Lemma 3.1 (1) and the definition of the weight yield

$$w_{\mathcal{F}}(a) \geq W(2^d - (c-1)) + \frac{d-2}{18} \geq W(2^d - (c-1)) + \frac{1}{6} \geq \frac{2^d - c}{d}.$$
(4.3)

Now observe that the definition of $w_{\mathcal{F}}$ implies $\sum_{x \in V} w_{\mathcal{F}}(x) \leq 1 + |\mathcal{F}|$ because the left-hand side counts every edge of $\mathcal{F}$ apart from the empty set at most once. In particular,

for $H \in \mathcal{S}$ there are at least one $x \in H \cap \bigcup_{v \in \mathcal{L}_2} V_v$ and at most two $a, a' \in H \cap \bar{\mathcal{L}}$. Thus, H contributes at most 1 to $\sum_{x \in V} w_{\mathcal{F}}(x)$.

Since (4.1), (4.2), Claim 4.1, and (4.3) say that the average weight per vertex in $\mathcal{F}$ is at least $\frac{2^d - c}{d}$, the proof is complete. $\qquad\square$

Proof of Claim 4.1. Here, we will differ slightly depending on the value of c.

Case $c = 3$: If $\delta(\mathcal{F}[V_v]) \geqslant 2^{d-1} - 2$, then (4.2) holds for v as well and so we may assume $\delta(\mathcal{F}[V_v]) < 2^{d-1} - 2$ and thereby $|2^{V_v} \setminus \mathcal{F}| \geqslant 3$. On the other hand, since $d(v) \geqslant 2^{d-1} - 2$ and $\mathcal{F}$ is hereditary, $|2^{V_v} \setminus \mathcal{F}| \leqslant 4$. So we can assume that $|2^{V_v} \setminus \mathcal{F}| \in \{3, 4\}$. If $|2^{V_v} \setminus \mathcal{F}| = 3$, then $d(v) \geqslant 2^{d-1} - 2$ and $\mathcal{F}$ being hereditary imply that the sets in $2^{V_v} \setminus \mathcal{F}$ are V_v, $V_v \setminus \{v\}$, and some $A \in (V_v)^{(d-1)}$ with $v \in A$. Thus, each vertex $x \in A \setminus \{v\}$ lies in all three sets of $2^{V_v} \setminus \mathcal{F}$, and so there has to be an $F_x \in L_x \cap (V \setminus V_v)^{(1)}$ because of the minimum degree of $\mathcal{F}$. Thus, the definition of the weight and double counting the non-empty edges in $\mathcal{F}[V_v]$ implies

$$\sum_{x \in V_v} w_{\mathcal{F}}(x) \geqslant |\mathcal{F}[V_v] \setminus \{\varnothing\}| + \frac{|A \setminus \{v\}|}{2} \geqslant 2^d - 4 + \frac{d-2}{2} \geqslant 2^d - 3.$$

Similarly, if $|2^{V_v} \setminus \mathcal{F}| = 4$, then the sets in $2^{V_v} \setminus \mathcal{F}$ are V_v, $V_v \setminus \{v\}$, some $A \in (V_v)^{(d-1)}$ with $v \in A$, and $A \setminus \{v\}$. Hence, there are $d - 2$ vertices x (namely, the vertices in $A \setminus \{v\}$) for which there has to be an $F_x \in L_x \cap (V \setminus V_v)^{(1)}$ and at least one further $F'_x \in L_x$ with $F'_x \cap (V \setminus V_v) \neq \varnothing$ and $|F'_x| \leqslant 2$. Noting that each F_x contributes $1/2$ to $\sum_{x \in V_v} w_{\mathcal{F}}(x)$ and each F'_x at least $1/3 - 1/9 = 2/9$, we obtain in the usual way

$$\sum_{x \in V_v} w_{\mathcal{F}}(x) \geqslant 2^d - 5 + \frac{d-2}{2} + \frac{2(d-2)}{9} \geqslant 2^d - 3$$

and thereby the claim if $c = 3$.

Case $c = 4$: In a similar way as in the beginning of the case $c = 3$, we observe that we may assume $|2^{V_v} \setminus \mathcal{F}| \in \{4, 5, 6\}$. Further observe that if $|2^{V_v} \setminus \mathcal{F}| = 4$, then since $d(v) = 2^{d-1} - 3$, the sets in $2^{V_v} \setminus \mathcal{F}$ are V_v, $V_v \setminus \{v\}$, A, and B for some distinct $A, B \in V_v^{(d-1)}$ which both contain v. Thus, there are at least $d - 3$ vertices (namely those in $A \cap B \setminus \{v\}$) that lie in four sets of $2^{V_v} \setminus \mathcal{F}$. Since for any such vertex x there has to be an $F_x \in L_x \cap (V \setminus V_v)^{(1)}$, we get $\sum_{x \in V_v} w_{\mathcal{F}}(x) \geqslant 2^d - 5 + \frac{d-3}{2} \geqslant 2^d - 4$.

Similarly, if $|2^{V_v} \setminus \mathcal{F}| = 5$, the sets in $2^{V_v} \setminus \mathcal{F}$ are V_v, $V_v \setminus \{v\}$, A, B, and $A \setminus \{v\}$ for some distinct $A, B \in V_v^{(d-1)}$ which both contain v. Hence, for the $d - 3$ vertices $x \in A \cap B \setminus \{v\}$ there have to be an $F_x \in L_x \cap (V \setminus V_v)^{(1)}$ and at least one further $F'_x \in L_x$ with $F'_x \cap (V \setminus V_v) \neq \varnothing$ and $|F'_x| \leqslant 2$. In addition, for the one vertex $x \in A \setminus B$ there has to be an $F_x \in L_x \cap (V \setminus V_v)^{(1)}$. For a vertex $x \in A \cap B \setminus \{v\}$ we observe the following. If $F'_x \notin \mathcal{S}$, then F'_x contributes at least $1/3$ to $\sum_{x \in V_v} w_{\mathcal{F}}(x)$. On the other hand, if $F'_x \in \mathcal{S}$,

then there is some $a \in \bar{\mathcal{L}}$ with $a \in F'_x$. Since for any $a \in \bar{\mathcal{L}}$ we have $L_a \cong \mathcal{R}(2^{d-1} - 3)$ (and $d \geqslant 5$), the number of 2-sets in L_a which contain x is at least $d - 2 \geqslant 3$. So in this case the edges in $\{H \in L_x : H \cup \{x\} \in \mathcal{S}\}$ contribute at least $\frac{2}{9} \cdot 3 = 2/3$. In either case, we derive

$$\sum_{x \in V_v} w_{\mathcal{F}}(x) \geqslant 2^d - 6 + \frac{d-2}{2} + \frac{d-3}{3} \geqslant 2^d - 4 \,.$$

Lastly, if $|2^{V_v} \smallsetminus \mathcal{F}| = 6$, then the sets in $2^{V_v} \smallsetminus \mathcal{F}$ are V_v, $V_v \smallsetminus \{v\}$, A, B, $A \smallsetminus \{v\}$, and $B \smallsetminus \{v\}$ for some distinct $A, B \in V_v^{(d-1)}$ which both contain v. Thus, for the $d - 3$ vertices $x \in A \cap B \smallsetminus \{v\}$ there is an $F_x \in L_x \cap (V \smallsetminus V_v)^{(1)}$ and at least two further $F_x^i \in L_x$ with $F_x^i \cap (V \smallsetminus V_v) \neq \varnothing$ and $|F_x^i| \leqslant 2$, $i \in [2]$. In addition, there are two further vertices $x \in A \triangle B$ for which there is at least one $F_x \in L_x \cap (V \smallsetminus V_v)^{(1)}$. For a vertex $x \in A \cap B \smallsetminus \{v\}$ we observe the following. If $F_x^i \notin \mathcal{S}$ for $i = 1, 2$, then these two edges together contribute at least $2/3$ to $\sum_{x \in V_v} w_{\mathcal{F}}(x)$. If $F_x^i \in \mathcal{S}$ for some $i \in \{1, 2\}$, then the edges in $\{H \in L_x : H \cup \{x\} \in \mathcal{S}\}$ contribute at least $2/3$ as noted above. Therefore the definition of the weight entails

$$\sum_{x \in V_v} w_{\mathcal{F}}(x) \geqslant 2^d - 7 + \frac{d-1}{2} + \frac{2(d-3)}{3} \geqslant 2^d - 4$$

and thereby the claim is proved if $c = 4$. $\qquad \square$

§5. Further Remarks and Open Problems

Consider $m(s)$ to be the following limit introduced in [4]

$$m(s) := \lim_{n \to \infty} \frac{m(n, s)}{n} \,.$$

It is not difficult to check that $m(s)$ is well-defined (see [4]). Rephrased by means of this definition, Theorem 1.3 implies that for $c \leqslant \frac{d}{4}$ we have that

$$m(2^{d-1} - c) = \frac{2^d - c}{d} \,. \tag{5.1}$$

The first open problem we would like to mention concerns finding a sharp relation between d and c such that (5.1) holds. More precisely, finding the maximum integer $c_0(d)$ such that the equality (5.1) holds for every $c \leqslant c_0$. In view of Theorem 1.3 we have that $c_0(d) \geqslant \lfloor \frac{d}{4} \rfloor$, and below we will give a construction that proves that $c_0(d) \leqslant d$ for $d \geqslant 5$.

Let $\mathcal{F} \subseteq 2^V$ with $|V| = n$ and d be a positive integer such that $d | n$. We say that $\mathcal{F}$ is d-*local* if there exists a partition of V into sets of size d such that every $F \in \mathcal{F}$ is a subset of one of the sets of the partition. Observe that the extremal construction presented in the proof of Theorem 1.3 is a d-local hypergraph with minimum degree $2^{d-1} - c + 1$ and with $m(n, 2^{d-1} - c) + 1$ edges. That construction can be generalised in the following way.

Take $d \geqslant 5$ and $c \in [2^{d-2}]$ and set $s = 2^{d-1} - c$, for simplicity let $d|n$. By definition of $m(\cdot, \cdot)$, there is a family on d vertices with $m(d, s) + 1$ edges such that all vertices have degree at least $s + 1$. Take n/d vertex disjoint copies of such a family. It is clear that in the resulting family all vertices have degree at least $s + 1$ and the number of edges is $m(d, s)\frac{n}{d} + 1$. This family minimises the number of edges for d-local families with minimum degree at least $s + 1$ and gives the following general upper bound on $m(2^{d-1} - c)$

$$m(2^{d-1} - c) \leqslant \frac{m(d, 2^{d-1} - c)}{d} . \tag{5.2}$$

Moreover, we observe that for $c = d + 1$ we have that $m(d, 2^{d-1} - (d+1)) < 2^d - (d+1)$. To see this, consider the family $\mathcal{F} \subseteq 2^{[d]}$ containing all sets with at most $d - 2$ vertices. Then $\mathcal{F}$ has $2^d - (d + 1)$ edges and minimum degree $2^{d-1} - d > 2^{d-1} - c$. Thus, from (5.2) it follows that

$$m(2^{d-1} - (d+1)) \leqslant \frac{2^d - (d+2)}{d} .$$

This means (5.1) does not hold for $c = d + 1$, and hence $c_0(d) \leqslant d$.

Note that this construction is also d-local. An interesting problem is to find the values of c for which there are no d-local extremal families.

Problem 5.1. *Given a positive integer $d \geqslant 2$, find the minimal $c_\star(d) \in [1, 2^{d-2}]$ such that for all $c \geqslant c_\star$ we have*

$$m(2^{d-1} - c) < \frac{m(d, 2^{d-1} - c)}{d} .$$

A solution to this problem would give an insight into the structural behaviour of the extremal families: For $c \geqslant c_\star$ and large n (possibly satisfying certain divisibility conditions) there is no d-local extremal family for $m(n, 2^{d-1} - c)$. Note that the results in [2, 4, 11] solved Problem 5.1 for $d \leqslant 4$.

In the following, given a vertex set of size n we describe a non d-local family that has less edges than any possible d-local family with the same minimum degree. More precisely, the construction below yields that, given $d \geqslant 5$ and $c = d$, we have

$$m(2^{d-1} - d) \leqslant \frac{2^d - d - \frac{1}{2}}{d} < \frac{m(d, 2^{d-1} - d)}{d} . \tag{5.3}$$

Construction 5.2. Let $d \geqslant 5$ and k a positive integer, set $n = 2dk$. Take V to be a set of n vertices. Consider $U_1, \dots, U_{2k}$ to be a partition of V into sets of size d, and for every

set U_i arbitrarily pick a vertex $x_i \in U_i$. Define

$$\mathcal{G} = \{S \subseteq V : \text{ there is an } i \text{ such that } S \subseteq U_i \text{ and } |S| \leqslant d - 2\}$$

$$\mathcal{H} = \{U_i \smallsetminus \{x_i\} : \text{ for } i \in \{1, 2, \ldots, 2k\}\}$$

$$\mathcal{I} = \{\{x_i, x_{i+1}\} : \text{ for } i \in \{1, 3, 5, \ldots, 2k - 1\}\}.$$

One can check that the number of edges of the family $\mathcal{F} = \mathcal{G} \cup \mathcal{H} \cup \mathcal{I}$ is given by

$$|\mathcal{G}| + |\mathcal{H}| + |\mathcal{I}| = \frac{2^d - d - 2}{d}n + 1 + \frac{n}{d} + \frac{n}{2d} = \frac{2^d - d - \frac{1}{2}}{d}n + 1.$$

Moreover, every vertex in V has degree $s = 2^{d-1} - d + 1$. This implies the first inequality of (5.3). Taking $d = c + 1$ in Lemma 3.2 (1) yields

$$2^d - d \leqslant m(d, 2^{d-1} - d),$$

and thereby the second inequality in (5.3).

For $s \leqslant 16$ (that is $d \leqslant 5$), considering the results from [2, 4, 11] and Theorem 1.5 all values of $m(s)$ are found, except $m(11)$. We recall the conjecture of Frankl and Watanabe [4], which states that Construction 5.2 is extremal for $d = 5$.

Conjecture 5.3 ([4]). $m(11) = 5.3$

A complementary approach than the one taken in this paper could be as follows.

Problem 5.4. *Given a positive integer d and an integer $c \in [0, 2^{d-1})$, find the value of $m(2^{d-1} + c)$.*

Naturally, for $c \geqslant 2^{d-1} - \frac{d}{4}$ Problem 5.4 is solved by Theorem 1.3. For $c \leqslant 2^{d-2}$, the only general result is given in [4], where it is shown that $m(2^{d-1}) = \frac{2^d - 1}{d} + \frac{1}{2}$. For other values of c Problem 5.4 is still open.

Observe that Theorems 1.3 and 1.5 and the results presented in [2, 4, 11] concern cases in which s is close to 2^d for some value of d. In general, there are still large intervals between powers of 2 for which the only bounds on $m(s)$ that are known are those that follow directly from the previously mentioned results. Finding a solution for Problem 5.1 might shed light on this problem by possibly providing a first understanding of the structural behaviour in those intervals.

References

[1] J. A. Bondy, *Induced subsets*, Journal of Combinatorial Theory, Series B **12** (1972), no. 2, 201–202. ↑1

[2] P. Frankl, *On the trace of finite sets*, Journal of Combinatorial Theory, Series A **34** (1983), no. 1, 41–45. ↑1, 1, 1.1, 2, 5, 5, 5

[3] P. Frankl and N. Tokushige, *Extremal problems for finite sets*, Vol. 86, American Mathematical Soc., 2018. ↑1, 1

[4] P. Frankl and M. Watanabe, *Some best possible bounds concerning the traces of finite sets*, Graphs and Combinatorics **10** (1994), no. 2-4, 283–292. ↑1, 1, 1.1, 4, 5, 5, 5, 5.3, 5

[5] Z. Füredi and J. Pach, *Traces of finite sets: extremal problems and geometric applications*, Extremal problems for finite sets **3** (1991), 255–282. ↑1

[6] L. Lovász, *Combinatorial Problems and Exercises. 1979*, North-Holland, Amsterdam. ↑1

[7] G. O. H. Katona, *Optimization for order ideals under a weight assignment*, Problèmes Comb, et Théorie des Graphes (1978), 257–258. ↑2

[8] N. Sauer, *On the density of families of sets*, Journal of Combinatorial Theory, Series A **13** (1972), no. 1, 145–147. ↑1

[9] S. Shelah, *A combinatorial problem; stability and order for models and theories in infinitary languages*, Pacific Journal of Mathematics **41** (1972), no. 1, 247–261. ↑1

[10] V. N. Vapnik and A. Y. Chervonenkis, *On the uniform convergence of relative frequencies of events to their probabilities*, Measures of complexity, 2015, pp. 11–30. ↑1

[11] M. Watanabe, *Some best possible bounds concerning the traces of finite sets II*, Graphs and Combinatorics **11** (1995), no. 3, 293–303. ↑1, 4, 5, 5, 5

[12] _____, *Arrow relations on families of finite sets*, Discrete mathematics **94** (1991), no. 1, 53–64. ↑1

Fachbereich Mathematik, Universität Hamburg, Hamburg, Germany
E-mail address: simon.piga@uni-hamburg.de

Fachbereich Mathematik, Universität Hamburg, Hamburg, Germany
E-mail address: bjarne.schuelke@uni-hamburg.de