Traces of Hypergraphs

Noga Alon∗ Guy Moshkovitz† Noam Solomon‡

April 28, 2018

Abstract

Let Tr(n, m, k) denote the largest number of distinct projections onto k coordinates guaranteed in any family of m binary vectors of length n. The classical Sauer-Perles-Shelah Lemma implies that Tr(n, n^r, k) = 2^k for k ≤ r. While determining Tr(n, n^r, k) precisely for general k seems hopeless even for constant r, estimating it, and more generally estimating the function Tr(n, m, k) for all range of the parameters, remains a widely open problem with connections to important questions in computer science and combinatorics. Here we essentially resolve this problem when k is linear and m = n^r where r is constant, proving that, for any constant α > 0, Tr(n, n^r, αn) = ˜Θ(n^µ) with µ = µ(r, α) = \frac{r+1−\log(1+\alpha)}{2^{−\log(1+\alpha)}}.

For the proof we establish a “sparse” version of another classical result, the Kruskal-Katona Theorem, which gives a stronger guarantee when the hypergraph does not induce dense subhypergraphs. Furthermore, we prove that the parameters in our sparse Kruskal-Katona theorem are essentially best possible. Finally, we mention two simple applications which may be of independent interest.

1 Introduction

For a hypergraph (or a set system) F, the trace of F on a vertex subset I is defined as the set of projections of the edges of F onto I, namely, F_I = \{e ∩ I : e ∈ F\}. The shatter function, or trace function, of F is Tr(F, k) = max_I |F_I| with I a set k vertices. The focus of this paper is the following important extremal function; for integers n ≥ k and 0 ≤ m ≤ 2^n, let Tr(n, m, k) denote the largest number of distinct projections onto k vertices guaranteed in any n-vertex m-edge hypergraph:

Tr(n, m, k) = \min_{F \subseteq 2[n], |F| = m} \max_{I \subseteq [n], |I| = k} |F_I|.

There is a considerable number of results, in various areas of discrete mathematics, determining or estimating this function for certain values of the parameters. The most famous result is arguably the Sauer-Perles-Shelah Lemma ([16], [17], see also Vapnik and Chervonenkis [18] for a slightly weaker estimate).

Theorem 1.1 (Sauer-Perles-Shelah). Tr(n, m, k) = 2^k for m > \sum_{i=0}^{k-1} \binom{n}{i}.
The **VC-dimension** of a hypergraph $\mathcal{F}$ is the largest $k$ so that $\text{Tr}(\mathcal{F}, k) = 2^k$, i.e., it is the largest number $k$ so that $\mathcal{F}$ has a full projection on some $k$ vertices. The VC-dimension is a basic combinatorial measure of the complexity of a hypergraph; understanding the shatter function beyond the case of full projections is a very natural direction. Shatter functions and the VC-dimension are extensively studied in combinatorial and computational geometry, as well as in machine learning (see the survey of Matoušek [15] for several geometric and algorithmic applications of shatter functions, and the survey of Angulin [3] for the role VC-dimension is playing in computational learning theory).

In [6], Bondy proved that $\text{Tr}(n, n, n - 1) = n$, and a remark in [1] and in [9] is that $\text{Tr}(n, m, 3) = 7$ for $m = 1 + n + \lceil n^2/4 \rceil + 1$, and the same argument implies that determining the smallest $m$ for which $\text{Tr}(n, m, 4) = 15$ is equivalent to determining the maximum possible number of edges of a 3-uniform hypergraph on $n$ vertices with no complete hypergraph on 4 vertices—a well-known open problem of Turán. Additional results that can all be formulated in terms of the function $\text{Tr}(n, m, k)$ appear in [5], [12], [4], [8] and more.

Recently, Bukh and Goaoc were able to estimate $\text{Tr}(n, n^{O(1)}, k)$ for constant values of $k$ that are not too small (and also improved an earlier lower bound of [8]). In the other extreme regime, a classical paper of Kahn et al. [12] proves that for every $0 < \alpha < 1$,

$$\text{Tr}(n, 2^{n-1}, \alpha n) \geq (1 - n^{-c(\alpha)}) 2^{\alpha n},$$

where $c(\alpha) > 0$ depends on $\alpha$. Benny Chor conjectured in the 80s that one can in fact make the error term exponentially rather than polynomially small in $n$. This conjecture was recently disproved by Bourgain et al. [4]. In fact, Bourgain et al. prove several additional results, in particular strengthening those of [12].

In [5], Bollobás and Radcliffe considered the case where $m$ is polynomial and $k$ is linear. For the lower bound they were able to prove the following.

**Theorem 1.2** (Bollobás and Radcliffe [5, Theorem 7]). For constants $r \geq 2$ and $0 < \alpha \leq 1$ it holds that $\text{Tr}(n, n^r, \alpha n) \geq \Omega(n^{\lambda r})$ with

$$\lambda = \begin{cases} \log(1 + \alpha) & \alpha \in [\sqrt{2} - 1, 1] \\ \log(1 + \alpha)/H(\log(1 + \alpha)) & \alpha \in (0, \sqrt{2} - 1) \end{cases}$$

For the upper bound, it would seem that among hypergraphs on $n$ vertices with a given number of edges $m$, a hypergraph $\mathcal{F}$ with $\text{Tr}(\mathcal{F}, k) = \text{Tr}(n, m, k)$ should be very symmetric, when $k$ is not too small or too large. A natural candidate for such an extremal hypergraph is thus the hypergraph containing all edges up to the appropriate size. Bollobás and Radcliffe were able to show that this is in fact not the case.

**Theorem 1.3** (Bollobás-Radcliffe [5, Theorem 11]). For every constant $r \geq 2$,

$$\text{Tr} \left( n, \sum_{i=0}^{r} \binom{n}{i}, n/2 \right) \leq o(n^r).$$

---

$^1$Here $H(x) = -x \log(x) - (1 - x) \log(1 - x)$ is the binary entropy function, and the logarithms are in base 2.
1.1 Our results

Our main result in this paper determines the value of $\text{Tr}(n, n^r, \alpha n)$, for constant $r$ and $\alpha$, up to logarithmic factors, thus closing the gap between the lower and upper bounds in Theorems 1.2 and 1.3. We henceforth use the following standard notation: for two functions $f(n)$ and $g(n)$, by $f = \tilde{O}(g)$ we mean $f = O(g \log^c(g))$ for some absolute constant $c > 0$; $f = \tilde{\Omega}(g)$ and $f = \tilde{\Theta}(g)$ are defined analogously. The main result of this paper is as follows.

**Theorem 1.4** (Main result). *For constants $r > 1$ and $0 < \alpha \leq 1$ we have $\text{Tr}(n, n^r, \alpha n) = \tilde{\Theta}(n^\mu)$, where

$$\mu = \mu(r, \alpha) = \frac{r + 1 - \log(1 + \alpha)}{2 - \log(1 + \alpha)}. \quad (1)$$

It is perhaps instructive to consider one representative special case: $r = 2$, $\alpha = 1/2$. In this case, the proofs in [5] bound $\text{Tr}(n, n^2, n/2)$ as follows;

$$\Omega(n^{1.169925..}) = \Omega(n^{2\log_2 3/2}) \leq \text{Tr}(n, n^2, n/2) \leq \frac{n^2(\log \log n)^O(1)}{\log n} = o(n^2),$$

whereas we prove that

$$\text{Tr}(n, n^2, n/2) = \tilde{\Theta}(n^{1+1/(3-\log_2 3)}) = \tilde{\Theta}(n^{1.706695..}).$$

**A new Kruskal-Katona-type theorem.** As it turns out, our main result can be readily deduced from a new version of the well-known Kruskal-Katona Theorem. Recall that the Kruskal-Katona Theorem gives a lower bound on the number of $i$-sets contained within the edges of a uniform hypergraph. Formally, for a hypergraph $F$ and $i \in \mathbb{N}$ we denote

$$\left( F^i \right) = \{ S \mid |S| = i \text{ and } \exists e \in F: S \subseteq e \}.$$

The following classical version of the Kruskal-Katona Theorem was given by Lovász [14]. Henceforth, for real $y > 0$ we denote $\binom{y}{i} = \frac{y(y-1)\cdots(y-i+1)}{i!}$. We use the abbreviation that $F$ is a $k$-graph to mean that $F$ is a $k$-uniform hypergraph.

**Theorem 1.5** (Kruskal-Katona Theorem, Lovász [14]). *Let $F$ be a $k$-graph. If $|F| = \binom{y}{k}$ with real $y > 0$ then for every $0 \leq i \leq k$ we have $|\left( F^i \right)| \geq \binom{y}{i}$.*

Our new version of the Kruskal-Katona theorem gives a stronger lower bound depending on the sparsity of the hypergraph $F$. As is standard, we denote the sub-hypergraph of a hypergraph $F$ induced on a vertex subset $I$ by $F[I] = (I, \{ e \mid e \in F \text{ and } e \subseteq I \})$. We denote the largest number of edges in an induced sub-hypergraph on $i$ vertices by

$$\text{span}(F, i) := \max_{I \subseteq [n], |I| = i} |F[I]|.$$

We give here the new version of the Kruskal-Katona Theorem. (See Theorem 2.7 for a slightly stronger form.) For the parameters relevant to our applications here, it provides a significantly stronger estimate than the classical theorem, using an appropriate sparseness assumption.
**Theorem 1.6** ("Sparse Kruskal-Katona Theorem"). Let $F$ be a $k$-graph with $n$ vertices and $|F| = n^r$ edges, $r > 1$. If

$$\text{span}(F, \alpha n) \leq \min \left\{ \left( \frac{k}{x - \lceil r \rceil} \right) n, \frac{1}{2} |F| \right\}$$

with real $x \geq 2k$ then for every $r + 1 \leq i \leq k$ we have

$$\left| \binom{F}{i} \right| \geq \frac{1}{C} \cdot \frac{\binom{x}{i}}{\binom{k}{i}} |F|,$$

with $C = \left( \frac{8k}{\alpha} \right)^{\lceil 3r \rceil} \log n$.

It is of course natural to ask whether the bound in Theorem 1.6 is essentially best possible. Our third result in this paper proves that this is indeed the case.

**Theorem 1.7** (Upper bound for Sparse Kruskal-Katona). Let $n, k, x \in \mathbb{N}^+$, $r \geq 1$ and $0 < \alpha \leq 1$ with $3r \leq k \leq x \leq n^{1/6}$ and $n \leq \alpha k n^r \leq \binom{x}{r} n$. There exists a $k$-graph $F$ with $n$ vertices, $|F| = n^r$ edges, and $\text{span}(F, \alpha n) \leq O\left( \binom{x}{r} n \right)$ such that for every $0 \leq i \leq k$ we have $|\binom{F}{i}| \leq \frac{\binom{x}{i}}{\binom{k}{i}} |F|$.

**Applications.** We end the paper with two simple applications of our main result, in geometry and in graph theory. We next describe the geometric application. Let $\mathcal{H}$ be a family of halfspaces in $\mathbb{R}^d$, and let $P$ be a set of points in $\mathbb{R}^d$. We say that $P$ separates $\mathcal{H}$ if for every pair of distinct halfspaces $H_1 \neq H_2 \in \mathcal{H}$ there is a point in $P$ that lies in one and outside the other.

**Proposition 1.8.** Fix $r > 1$, let $\mathcal{H}$ be a family of $n^r$ halfspaces in $\mathbb{R}^d$ and let $P \subset \mathbb{R}^d$ be a set of $n$ points such that $P$ separates $\mathcal{H}$. Then there is a subset $P' \subseteq P$ of at most $\frac{n}{\log n}$ points and a subset $\mathcal{H}' \subseteq \mathcal{H}$ of $\tilde{\Omega}(n^{r+1}/ \alpha^{r+1})$ halfspaces such that $P'$ separates $\mathcal{H}'$.

**Organization.** In Section 2 we prove Theorem 1.6 using an appropriate hypergraph decomposition method. We use it in Subsection 2.2 to deduce the lower bound in Theorem 1.4. In Subsection 3.1 we prove that the parameters of Theorem 1.6 are essentially best possible, and in Subsection 3.2 we prove a matching upper bound for Theorem 1.4, using a probabilistic construction. Our applications, Proposition 4.1 and Proposition 4.3, are given in Section 4.

**Proofs overview.** For the proof of the sparse Kruskal-Katona Theorem (Theorem 1.6, see also Theorem 2.7 below) we proceed as follows. In the first part of the proof we apply a new approximate hypergraph decomposition method, relying on the sparseness of the input hypergraph, into links. The decomposition is performed iteratively, in each step finding many vertices of high degree within the current link and restricting the next links to them. The final step of these iterations consists of “cleaning” each link by iteratively removing vertices of low degree. We then prove, using the sparseness of the hypergraph, that in fact most of the parts in our decomposition have few edges. In the second part of the proof we find, by applying the classical Kruskal-Katona Theorem, $i$-subsets within the edges of each (sub)link separately. We then argue that, since the links approximately decompose the hypergraph, we may essentially collect the $i$-subsets from all links without much overcounting. The proof of our lower bound for traces (in Theorem 1.4) follows quite easily from the sparse Kruskal-Katona theorem by applying it on the most “popular layer” of the hypergraph.

---

2 A halfspace in $\mathbb{R}^d$ consist of all points above a hyperplane.
(i.e., the uniform hypergraph with the most edges contained in our hypergraph, which we may assume is down-closed) and projecting onto a random subset of \( \alpha n \) vertices.

For the proof that the parameters in the sparse Kruskal-Katona Theorem are essentially best possible (Theorem 3.1 below) we give a randomized construction of a uniform hypergraph whose edges contain few \( i \)-subsets. The construction is fairly simple: the union of a carefully chosen number of cliques on random subsets, such that it simultaneously holds that there are many cliques and yet they are nearly edge disjoint. We show in particular that the expected number of edges induced on subsets of \( \alpha n \) vertices is sufficiently small so as to allow taking a union bound over all cliques. The proof of the upper bound for traces (in Theorem 1.4) follows by taking the down-closed hypergraph generated by the uniform hypergraph above, and then upper bounding the expected trace on a random subset of \( \alpha n \) vertices.

Throughout the paper we assume, whenever needed, that \( n \) is sufficiently large. All logarithms are in base 2 unless otherwise specified. To simplify the presentation we omit all floor and ceiling signs whenever these are not crucial.

2 Sparse Kruskal-Katona and Traces Lower Bound

In this section we prove Theorem 1.6. Henceforth, for a hypergraph \( \mathcal{F} \) on \( V \) and for a vertex subset \( I \subseteq V \) we denote by \( \mathcal{F}(I) \) the link of \( I \) in \( \mathcal{F} \), that is,

\[
\mathcal{F}(I) = (V \setminus I, \{ e \setminus I \mid I \subseteq e \in \mathcal{F} \}).
\]

Note that if \( \mathcal{F} \) is a \( k \)-graph then \( \mathcal{F}(I) \) is a \((k - |I|)\)-graph. For a tuple \( U \) of vertices in \( V \) we denote by \( |U| \) the number of distinct vertices in \( U \), and by \( \mathcal{F}(U) \) the link \( \mathcal{F}(I) \) where \( I \) is the set of (distinct) vertices in \( U \) (and so \( \mathcal{F}(U) \) is a \((k - |U|)\)-graph).

For the proof we will need several lemmas which we state and prove below. We begin with the following simple “hypergraph regularization” lemma.

**Lemma 2.1.** Every hypergraph \( \mathcal{F} = (V, E) \) has an induced sub-hypergraph \( \mathcal{F}' = (V', E') \) satisfying:

1. \(|E'|/|V'| \geq |E|/|V|\).

2. The degree of each vertex \( v \in V' \) in \( \mathcal{F}' \) is at least \(|E'|/2|V'| \ln |V'|\).

3. \(|E'| > |E|/2\).

**Proof.** Put \( n = |V| \), \( V_0 = V \) and \( E_0 = E \). Starting with \( i = 0 \), as long as the hypergraph \( \mathcal{F}_i = (V_i, E_i) \) in which \(|V_i| = n - i| \) does not satisfy (2), let \( V_{i+1} \) be the set obtained from \( V_i \) by removing a vertex of minimum degree in \( \mathcal{F}_i \), and let \( \mathcal{F}_{i+1} \) be the induced subhypergraph on this set. It is easy to see that \(|E_{i+1}|/|V_{i+1}| > |E_i|/|V_i|\) and hence this process must terminate with a nonempty hypergraph \( \mathcal{F}_j = (V_j, E_j) \). Define \( V' = V_j \), \( E' = E_j \). Then (1) holds as the quantity \(|E_i|/|V_i|\) keeps increasing during the process, (2) holds by the definition of \( j \), and (3) holds since

\[
|E'| = |E_j| \geq |E_0| \prod_{i=0}^{j-1} \left( 1 - \frac{1}{2(n - i) \log n} \right) \geq |E_0| \left( 1 - \sum_{i=0}^{j-1} \frac{1}{2(n - i) \log n} \right) \geq |E| \left( 1 - \frac{\ln n}{2 \log n} \right) > \frac{|E|}{2}.
\]

\(\square\)

**Lemma 2.2.** Let \( \mathcal{F} \) be a hypergraph on \( n \) vertices with \( \text{span}(\mathcal{F}, i) \leq \frac{|F|}{2} \), for some integer \( 0 < i < n \). Then \( \mathcal{F} \) has at least \( i \) vertices of degree at least \( \frac{|F|}{2n} \).
Proof. Let $I \subseteq V(F)$ denote the set of vertices of $F$ of degree at least $\frac{|F|}{2n}$. Then $|F[I]| > |F| - n \cdot \frac{|F|}{2n} = \frac{|F|}{2}$. By the assumption on $\text{span}(F, i)$ we thus have $|I| > i$. 

By an iterative application of Lemma 2.2 we obtain the following.

Lemma 2.3. Let $s \in \mathbb{N}^+$ and let $F$ be a hypergraph on $V$ with $\text{span}(F, i) \leq \frac{|F|}{2n^{i+1}}$. Then there are at least $i^s$ $s$-tuples $U \subseteq V$ with $|F(U)| \geq \frac{|F|}{(2n)^s}$. 

Proof. Put $n = |V|$. For a tuple $U$ of vertices in $V$ we denote by $F^U$ the sub-hypergraph of $F$ on $V$ with edge set $\{e \in F : U \subseteq e\}$. Note that $|F^U| = |F(U)|$. We proceed by induction on $s$, noting that the induction basis $s = 1$ is Lemma 2.2. For the induction step, let $F$ be as in the statement, and note that by the induction hypothesis there are at least $i^{s-1}$ $(s-1)$-tuples $U \subseteq V^{s-1}$ with $|F^U| = |F(U)| \geq \frac{|F|}{(2n)^{s-1}}$. Fix one such $U = (v_1, \ldots, v_{s-1})$ and apply Lemma 2.2 on the hypergraph $F^U$, noting that, as required,

$$\text{span}(F^U, i) \leq \text{span}(F, i) \leq \frac{|F|}{2s^{n-1}} \leq \frac{|F^U|}{2},$$

where the first inequality uses the fact that $F^U$ is a sub-hypergraph of $F$ on $V$. Thus, $F^U$ has at least $i$ vertices $v$ of degree at least $\frac{|F^U|}{2n^s} \geq \frac{|F|}{(2n)^{s+1}}$. This means that for each such $v$, the $s$-tuple $U' = (v_1, \ldots, v_{s-1}, v)$ satisfies $|F(U')| = |F^U| \geq \frac{|F|}{(2n)^{s+1}}$. Going over all $i^{s-1}$ $(s-1)$-tuples $U$ in a similar fashion, we deduce that the total number of $s$-tuples $U'$ as above is at least $i^{s-1} \cdot i$. This completes the induction step and the proof.

The following lemma gives a unified lower bound for the summation $\sum_{i=0}^k \binom{x}{i} \gamma^i$ that is independent of the ratio between $k$ and $x$. See Section A in the Appendix for a proof of this lemma.

Lemma 2.4. For every $k \in \mathbb{N}^+$ and real $0 \leq \gamma \leq 1$, $x \geq k$ we have

$$\sum_{i=0}^k \binom{x}{i} \gamma^i \geq \frac{1}{4} \left( \sum_{i=0}^k \binom{x}{i} \right)^{\log(1+\gamma)}.$$

Finally, we have the following well-known bounds.

Claim 2.5. We have $e^{-2x} \leq 1 - x \leq e^{-x}$, where the upper bound holds for every real $x$ and the lower bound holds for every $0 \leq x \leq 1/2$.

Proof. The upper bound is well known, and the lower bound follows from it since we have $1 - x = \left(1 + \frac{x}{1-x}\right)^{-1} \geq e^{-\frac{x}{1-x}} \geq e^{-2x}$, where the last inequality uses $0 \leq x \leq 1/2$. 

2.1 Sparse Kruskal-Katona Theorem

In this subsection we prove Theorem 2.7, which is a more precise version of Theorem 1.6. First, we will need a lemma which extends the classical Kruskal-Katona Theorem 1.5 by collecting $i$-subsets from the hypergraph’s links.

Lemma 2.6. Let $F$ be a $k$-graph on $V$, and let $t \in \mathbb{N}$. For every $t \leq i \leq k$ we have

$$\binom{|F|}{i} \geq i^{-t} \sum_{U \subseteq V^t} \binom{x_U}{i - |U|},$$

where $x_U$ is given by $|F(U)| = \binom{x_U}{k - |U|}$.

6
Proof. Put \( \mathcal{I} = V^t \), and let \( t \leq i \leq k \). Apply the Kruskal-Katona Theorem (Theorem 1.5) on each link \( \mathcal{F}(U) \) with \( U \in \mathcal{I} \). Since \( \mathcal{F}(U) \) is a \((k - |U|)\)-graph and \( 0 \leq i - |U| \leq k - |U| \) (using \( |U| \leq t \leq i \) for the lower bound), Theorem 1.5 implies that \( |\binom{\mathcal{F}(U)}{i-|U|}| \geq \binom{x_{i-|U|}}{i-|U|} \). Now, for every \( U \in \mathcal{I} \) denote

\[
\left( \binom{\mathcal{F}(U)}{i} \right)^* = \left\{ f \cup U \mid f \in \left( \binom{\mathcal{F}(U)}{i-|U|} \right) \right\}.
\]

We have that

\[
\left( \binom{\mathcal{F}(U)}{i} \right)^* \subseteq \left( \binom{\mathcal{F}}{i} \right) \quad \text{and} \quad \left| \left( \binom{\mathcal{F}(U)}{i} \right)^* \right| = \left| \left( \binom{\mathcal{F}(U)}{i-|U|} \right) \right| \geq \binom{x_U}{i-|U|}.
\]

We therefore deduce that

\[
\left| \left( \binom{\mathcal{F}}{i} \right) \right| \geq \left| \bigcup_{U \in \mathcal{I}} \left( \binom{\mathcal{F}(U)}{i} \right)^* \right| \geq i^{-t} \sum_{U \in \mathcal{I}} \left| \left( \binom{\mathcal{F}(U)}{i} \right)^* \right| \geq i^{-t} \sum_{U \in \mathcal{I}} \binom{x_U}{i-|U|},
\]

where, crucially, the penultimate inequality uses the fact that if an \( i \)-set \( g \) appears in \( \left( \binom{\mathcal{F}(U)}{i} \right)^* \) then \( U \in g^t \), implying that \( g \) appears in at most \( i \) families \( \left( \binom{\mathcal{F}(U)}{i} \right)^* \). This completes the proof. \( \square \)

Note that Lemma 2.6 recovers Theorem 1.5 by taking \( t = 0 \).

We prove the following stronger form of Theorem 1.6, our sparse Kruskal-Katona Theorem.

**Theorem 2.7.** Let \( \mathcal{F} \) be a \( k \)-graph with \( n \) vertices and \( |\mathcal{F}| = n^r \) edges, \( r > 1 \). Let \( s \in \mathbb{N} \) be the smallest satisfying \( \frac{|\mathcal{F}|}{(2n)^s} < c \cdot \text{span}(\mathcal{F}, an) \) with \( c = (8k)^r / \alpha^{[r]} - 1 \) \(^3\) and put \( t = s + 1 \). If

\[
\text{span}(\mathcal{F}, an) \leq \min \left\{ \binom{x}{k-t}n, \frac{1}{2} |\mathcal{F}| \right\}
\]

with real \( x > 0 \) then for every \( t \leq i \leq k \) we have

\[
\left| \left( \binom{\mathcal{F}}{i} \right) \right| \geq \frac{1}{C} \cdot \binom{x}{k-t} |\mathcal{F}|,
\]

with \( C = (8k/\alpha)^{[2r]} \log n \).

**Remark 2.8.** The parameters in Theorem 2.7 satisfy the following relations:

\[
t \leq [r] \leq k.
\]

For the first inequality, note that otherwise \( s = [r] \) and so, by the definition of \( s \), \( c \cdot \text{span}(\mathcal{F}, an) \leq |\mathcal{F}|/(2n)^{s-1} \leq n \), implying that \( \text{span}(\mathcal{F}, an) < an \) and thus, by averaging, \( |\mathcal{F}| < n \), contradicting the statement’s assumption \( r > 1 \). For the second inequality, notice \( n^r = |\mathcal{F}| \leq \binom{k}{r} \leq n^k \).

Note that the error term \( C \) increases with the quotient \( k/\alpha \) and with \( r \). This precludes us from taking hypergraphs of large uniformity, with many edges, or from inducing on too few vertices. More formally, we have the following corollary.

**Corollary 2.9.** Under the assumptions of Theorem 2.7:

\(^3\)One may think of \( |\mathcal{F}|/(2n)^s \) as an approximation (from below) to the average degree of a vertex \( s \)-tuple in \( \mathcal{F} \). Alternatively, \( s \) can be defined as \( s = \left\lceil \log \left( \frac{|\mathcal{F}|}{\log(2n)} \right) \right\rceil \) with \( \sigma = c \cdot \text{span}(\mathcal{F}, an) \).
1. If \((k/\alpha)^r \leq \text{poly}(\log n)\) then \(|\mathcal{F}|/C \geq \tilde{\Omega}(|\mathcal{F}|)\).

2. If \((k/\alpha)^r \leq n^{o(1)}\) then \(|\mathcal{F}|/C \geq |\mathcal{F}|^{1-o(1)}\).

We now show how to deduce the sparse Kruskal-Katona Theorem from Theorem 2.7.

**Proof of Theorem 1.6.** We have that \((\begin{array}{c} x \\ k-r \end{array}) \leq (\begin{array}{c} x \\ k \end{array})\) using (2) and since, by assumption, \(x \geq 2k\). Thus, the condition here implies the condition in Theorem 2.7. As for the guarantee in Theorem 2.7, note that

\[
\frac{(x-t)}{(x)} = \frac{(k-t)!}{(i-t)!} \left( \prod_{j=i}^{k} (x-j+t) \right)^{-1}
\]

\[
= \frac{(k-t)!}{(i-t)!} \left( \prod_{j=i}^{k} (x-j) \left( 1 + \frac{t}{x-j} \right) \right)^{-1} \geq \frac{k^{-t}k!}{i!} \left( \left( 1 + \frac{t}{x-(k-1)} \right)^{k-i} \prod_{j=i}^{k} (x-j) \right)^{-1}
\]

\[
\geq \frac{k^{-t}k!}{i!} \left( 1 + \frac{t}{k} \left( \prod_{j=i}^{k} (x-j) \right)^{-1} \right)^{-1} \geq (ek)^{-t} \frac{k!}{i!} \left( \prod_{j=i}^{k} (x-j) \right)^{-1} = (ek)^{-t} \frac{(x)}{(k)}
\]

where the second inequality uses the statement’s assumption \(x \geq 2k\), and the third inequality uses the upper bound in Claim 2.5. Thus, multiplying \(C\) from Theorem 2.7 by \((8k)^{r} \geq (ek)^{t}\) (recall (10)) completes the proof. \(\Box\)

**Proof of Theorem 2.7.** We will prove the implication that if the stronger condition

\[
\text{span}(\mathcal{F}, \alpha n) \leq \min \left\{ \frac{1}{c} \left( \begin{array}{c} x \\ k-t \end{array} \right) \alpha n, \frac{1}{2} |\mathcal{F}| \right\},
\]

(where \(c\) is as in the statement of the theorem) holds then for every \(t \leq i \leq k\) we in fact have

\[
\left| \left( \begin{array}{c} \mathcal{F} \\ i \end{array} \right) \right| \geq \frac{1}{c \log n} |\mathcal{F}| \left( \begin{array}{c} x \\ k-t \end{array} \right).
\]

To see why this would complete the proof, let \(\tilde{x}\) satisfy

\[
\left( \begin{array}{c} \tilde{x} \\ k-t \end{array} \right) = \frac{c}{\alpha} \left( \begin{array}{c} x \\ k-t \end{array} \right),
\]

so that if \(\mathcal{F}\) satisfies the statement’s original assumption that \(\text{span}(\mathcal{F}, \alpha n) \leq \min \left\{ (\begin{array}{c} x \\ k-t \end{array})n, \frac{1}{2} |\mathcal{F}| \right\}\) then it satisfies (3) with \(\tilde{x}\) replacing \(x\). Thus, from (4),

\[
\left| \left( \begin{array}{c} \mathcal{F} \\ i \end{array} \right) \right| \geq \frac{1}{c \log n} |\mathcal{F}| \left( \begin{array}{c} \tilde{x} \\ k-t \end{array} \right) \geq \frac{\alpha}{c^{2} \log n} |\mathcal{F}| \left( \begin{array}{c} x \\ k-t \end{array} \right) \geq \frac{1}{C} |\mathcal{F}| \left( \begin{array}{c} x \\ k-t \end{array} \right),
\]

where the second inequality uses (5) for both the denominator and the numerator, as (5) implies \(\tilde{x} \geq x\) since \(c/\alpha \geq 1\).

We henceforth assume (3), and our goal is to prove (4). By the definition of \(s\) we have

\[
\frac{|\mathcal{F}|}{(2n)^{s}} < c \cdot \text{span}(\mathcal{F}, \alpha n) \leq \frac{|\mathcal{F}|}{(2n)^{s-1}}.
\]

8
The upper bound in (6) implies in particular that \( \text{span}(\mathcal{F}, \alpha n) \leq \frac{|\mathcal{F}|}{2^n} \). Thus, by Lemma 2.3, there is a family \( \mathcal{U} \subseteq V(\mathcal{F})^s \) of \( s \)-tuples \( U \) of vertices of \( \mathcal{F} \) satisfying

\[
|\mathcal{F}(U)| \geq \frac{|\mathcal{F}|}{(2n)^s} =: b, \tag{7}
\]
such that \( |U| \geq \alpha^s n^s \). For each \( U \in \mathcal{U} \) apply Lemma 2.1 on the \((k - |U|)\)-graph \( \mathcal{F}(U) \) to obtain an induced subgraph \( \mathcal{F}(U)_{\text{reg}} \) with

\[
|\mathcal{F}(U)_{\text{reg}}| > \frac{1}{2} |\mathcal{F}(U)| \geq \frac{1}{2} b, \tag{8}
\]
such that \( \mathcal{F}(U)_{\text{reg}} \) has \( n_U \) vertices and minimum degree at least

\[
\frac{1}{4 \log n} \cdot \frac{b}{n_U} =: \left( \frac{x_U}{k - (|U| + 1)} \right) \tag{9}
\]
with \( x_U > 0 \). Let \( t \leq i \leq k \). Applying Lemma 2.6 with our \( t \) to the links of the vertices of each \( \mathcal{F}(U)_{\text{reg}} \), we deduce that

\[
\left| \binom{\mathcal{F}}{i} \right| \geq \frac{1}{i!} \sum_{U \in \mathcal{U}} n_U \left( i - (|U| + 1) \right) \geq \frac{1}{4i \log n} \cdot \frac{b}{n_U} \sum_{U \in \mathcal{U}} \binom{i - (|U| + 1)}{x_U} \geq \frac{1}{4(2k)^t \log n} \cdot \frac{1}{n} \cdot \frac{1}{s} \sum_{U \in \mathcal{U}} \binom{x_U}{k - t} |\mathcal{F}| \tag{10}
\]
where the second inequality uses (9), the third relies (crucially) on the decreasing monotonicity of the function \( z \mapsto \binom{y}{a-z}/\binom{y}{b-z} \) (with \( z \leq a \leq b \leq y \) and \( a, b, z \in \mathbb{N} \)) as well as (7) and the fact that \( i \leq k \). Let

\[
\mathcal{U}' = \left\{ U \in \mathcal{U} : \left( \frac{x_U}{k - (|U| + 1)} \right) \leq \left( \frac{x}{k - t} \right) \right\}. \tag{11}
\]
We will show that

\[
|\mathcal{U}'| \geq \frac{1}{2} \alpha^s n^s. \tag{11}
\]
By (10), this would imply that

\[
\left| \binom{\mathcal{F}}{i} \right| \geq \frac{\alpha^s}{8(2k)^t \log n} \cdot \frac{\binom{x}{k-t}}{\binom{x}{k-t}} |\mathcal{F}| \geq \frac{1}{c \log n} \cdot \frac{\binom{x}{k-t}}{\binom{x}{k-t}} |\mathcal{F}|,
\]
where the last inequality uses \( c = (8k)^{t/r} |\alpha|^{r-1} \leq 8(2k)^t / \alpha^s \) (recall (10)). Thus, proving (11) would imply (4) and complete the proof.

Put \( S = \binom{x}{k-t} \alpha n \). It remains to prove (11). Assume for contradiction that \(|\mathcal{U} \setminus \mathcal{U}'| \geq \frac{1}{2} (\alpha n)^s\). Note that by definition of \( \mathcal{U}' \) together with (9) we deduce that for every \( U \in \mathcal{U} \setminus \mathcal{U}' \) we have

\[
n_U \leq \frac{1}{4 \log n} \cdot \frac{b}{\binom{x}{k-t}} \leq \frac{1}{2} \cdot \frac{b}{\binom{x}{k-t}} = \frac{b}{S} \cdot \frac{1}{2} \alpha n =: n_0.
\]
Note that from the lower bound in (6) together with (3) we deduce that \( b \leq S \). Since \( n_0 \geq n_U \geq 1 \), we have that \( n_0 \) satisfies

\[
1 \leq n_0 \leq \frac{1}{2} \alpha n. \quad \tag{12}
\]
Put

\[
\ell = \frac{1}{2} \left\{ \begin{array}{ll}
\frac{\alpha n}{n_0 + s} & \text{if } s \geq 1 \\
1 & \text{if } s = 0
\end{array} \right. \tag{13}
\]
where the penultimate inequality uses the lower bound $\ell$ on $n$. Let $I$ denote the union of these sets of vertices. Then $I$ satisfies that

$$|I| \leq \sum_{U \in \mathcal{U}^*} (n_U + s) \leq [\ell] (n_0 + s) \leq 2\ell(n_0 + s) \leq \alpha n,$$

where the penultimate inequality uses the lower bound $\ell \geq \frac{1}{2}$ from (14). Moreover, $I$ satisfies that

$$|\mathcal{F}[I]| \geq \left| \bigcup_{U \in \mathcal{U}^*} \mathcal{F}[I_U] \right| \geq k^{-s} \sum_{U \in \mathcal{U}^*} |\mathcal{F}[I_U]| = k^{-s} \sum_{U \in \mathcal{U}^*} |\mathcal{F}(U)_{\text{reg}}| > k^{-s} \cdot \frac{1}{2} b [\ell],$$

where the second inequality uses the fact that $e \in \mathcal{F}[I_U]$ for at most $k^s$ $s$-tuples $U$ (by (15), $e \in \mathcal{F}[I_U]$ implies $U \subseteq e$), and the third inequality uses (8). Now, if $s = 0$ then $[\ell] = 1$ and $b = |\mathcal{F}|$, hence we get $|\mathcal{F}[I]| > |\mathcal{F}|/2 \geq \text{span}(\mathcal{F}, \alpha n)$ using (3), a contradiction. Otherwise, we get

$$|\mathcal{F}[I]| > \frac{1}{8k^{s+1}} \cdot \frac{b \alpha n}{n_0} = \frac{1}{4k^{s+1}} S \geq \frac{1}{c} S \geq \text{span}(\mathcal{F}, \alpha n),$$

where the first inequality uses (13) and bounds $n_0 + s \leq 2n_0 k$ (as $n_0 \geq 1$ by (12) and $s \leq k$ by (10)), the equality uses (12), and the last inequality uses (3). We thus again obtain a contradiction. This completes the proof.

We have the following important corollary of Theorem 2.7.

**Corollary 2.10.** Let $\mathcal{F}$ be a $k$-uniform hypergraph with $n$ vertices and $|\mathcal{F}| = n^r$ edges, $r > 1$. If

$$\text{span}(\mathcal{F}, \alpha n) \leq \min \left\{ Bn, \frac{1}{2} |\mathcal{F}| \right\}$$

with real $B > 0$, then for every $0 \leq \gamma \leq 1$ the expected trace of $\mathcal{F}$ on a uniformly random subset of $\gamma n$ vertices is at least

$$\frac{1}{C'} \cdot \frac{|\mathcal{F}|}{B^{1-\log(1+\gamma)}},$$

with $C' = (8k/\alpha \gamma)^{3r} \log n$, provided $k \leq \sqrt{\gamma n}$.

---

4Otherwise $C \geq |\mathcal{F}|$, and so the statement’s lower bound on $|\binom{s}{r}|$ is trivially true since $\binom{s}{r} \leq \binom{k-s}{r-k}$, where the first inequality uses the decreasing monotonicity of the function $z \mapsto \binom{a}{b}/\binom{a}{c}$ with $a \leq b \leq z$. 

---

10
Proof. Write $B = \binom{x-t}{k-t}$ with $x > 0$ with $t \in \mathbb{N}$ as in Theorem 2.7, so that $\text{span}(\mathcal{F}, \alpha n) \leq \min \{ \binom{x}{k-t} n, \frac{1}{\sqrt{2}} |\mathcal{F}| \}$. The expected trace of $\mathcal{F}$ on a uniformly random set of $q = \gamma n$ vertices is, by Theorem 2.7,

$$
\sum_{i=0}^{k} \left| \binom{\mathcal{F}}{i} \right| \frac{(n-i)!}{q^i (n-q)^{i-1}} \geq \sum_{i=0}^{k} \left( \binom{\mathcal{F}}{i} \right) \gamma^{i-1} \geq \frac{|\mathcal{F}|}{eC(k-t)} \sum_{i=t}^{k} \binom{x}{i-t} \gamma^{i-1} \geq \frac{|\mathcal{F}| e^C}{4C(k-t)} \left( \frac{x}{k-t} \right)^{\log(1+\gamma)} = \left( \frac{\gamma}{4C} \right)^t \frac{|\mathcal{F}|}{B^{1-\log(1+\gamma)}}
$$

with $C = (8k/\alpha)^{[2r]} \log n$ as in Theorem 2.7, where the first inequality uses (recall the lower bound in Claim 2.5 and the statement’s assumption $k \leq \sqrt{q}$)

$$
\frac{(n-i)!}{q^i (n-q)^{i-1}} = \frac{(n-i)!}{n!} \frac{q!}{(q-i)!} \prod_{j=0}^{i-1} \frac{q-j}{n-j} \geq \left( \frac{q}{n} \right)^i \prod_{j=0}^{i-1} \left( 1 - \frac{j}{q} \right) \geq \gamma^{i} \prod_{j=0}^{i-1} e^{-2j/q} = \gamma^{i} e^{-2i/q} \geq \gamma^{i} e^{-1},
$$

and the last inequality uses Lemma 2.4. Using (2) to bound $t$, the proof follows. \hfill \Box

Note that the statement in Corollary 2.10 does not depend on $k$, the uniformity of the hypergraph, except in the error term $C'$. In fact, in order for this statement to be meaningful, $C'$ should be negligible, so $k$ should be relatively small, e.g., poly-logarithmic in $n$.

Remark 2.11. If the bound in Corollary 2.10 is tight, then, in the proof of Theorem 2.7, all the inequalities become (essentially) equalities. In the proof of Theorem 2.7, we showed that the number of $s$-tuples in $U'$ is at least $\frac{1}{2} (\alpha n)^{x}$. It is then easy to verify that for half of these $s$-tuples $U$, we have $t_U \approx t_0, x_U \approx x$ and then $x \approx 2k$. (Quantifying this statement precisely requires more details, which we omit here.) Note that this remark applies to Corollary 2.10 but not to Theorem 2.7 itself, as evident from the matching upper bound in Subsection 3.1.

2.2 Tight lower bound for traces

Theorem 2.12. For any constant $r > 1$, $0 < \alpha < 1$ we have $\text{Tr}(n, n^r, \alpha n) \geq \tilde{\Omega}(n^{r+1-\log(1+\alpha)/2\log(1+\alpha)})$.

Proof. Let $\mathcal{H}$ be a down-closed hypergraph\footnote{Also called monotone. A hypergraph $\mathcal{H}$ is down-closed if for every edge $\mathcal{H}$, all its subsets are also edges of $\mathcal{H}$.} on $n$ vertices with $|\mathcal{H}| = n^r$. We will show that $\text{Tr}(\mathcal{H}, \alpha n) \geq \tilde{\Omega}(n^{r+1-\log(1+\alpha)/2\log(1+\alpha)})$, which would complete the proof since we may assume down-closeness of $\mathcal{H}$ (this is standard, see, e.g., [1]). Put $\mathcal{H}_i = \{ e \in \mathcal{H} \mid |e| = i \}$, and let $\mathcal{F} = \mathcal{H}_k$ where $k$ maximizes $|\mathcal{H}_k|$. Observe that since $\mathcal{H}$ is down-closed we have $|\mathcal{H}| \geq 2^k$, which implies that $k \leq \log |\mathcal{H}|$, and therefore $|\mathcal{F}| \geq |\mathcal{H}|/(\log |\mathcal{H}| + 1)$. We assume $\text{Tr}(\mathcal{F}, \alpha n) \leq \frac{1}{2} |\mathcal{F}|$, as otherwise we are done since

$$
\text{Tr}(\mathcal{H}, \alpha n) \geq \text{Tr}(\mathcal{F}, \alpha n) \geq \frac{1}{2} |\mathcal{H}|/(\log |\mathcal{H}| + 1) = \tilde{\Omega}(n^r).
$$

Now, write $\text{Tr}(\mathcal{F}, \alpha n) = Bn$ with $B > 0$. Since $\text{span}(\mathcal{F}, \alpha n) \leq \text{Tr}(\mathcal{F}, \alpha n)$ and since $k \leq \sqrt{\alpha n}$ (for all large enough $n$), we apply Corollary 2.10 with $\gamma = \alpha$ to obtain that

$$
Bn = \text{Tr}(\mathcal{F}, \alpha n) \geq \frac{\gamma}{C'} \frac{|\mathcal{F}|}{B^{1-\log(1+\alpha)}}.
$$

11
with $C' = (8k/\alpha^2)^{3r} \log n \leq (\log n)^{O(1)}$, where the inequality uses the fact that the maximal trace on $\alpha n$ vertices is at least as large as the expected trace on a random subset of $\alpha n$ vertices. This gives a lower bound on $B$; we therefore deduce

$$
\text{Tr}(H, \alpha n) \geq \text{Tr}(F, \alpha n) = Bn \geq \left(\frac{1}{C''} |F| \cdot \frac{1}{n} \right)^{\frac{1}{2 - \log(1 + \alpha)}} \cdot n
$$

$$
= \left(\frac{1}{C''} |F| n^{1 - \log(1 + \alpha)} \right)^{\frac{1}{2 - \log(1 + \alpha)}} \geq \left(\frac{1}{C''} |H| n^{1 - \log(1 + \alpha)} \right)^{\frac{1}{2 - \log(1 + \alpha)}},
$$

with $C'' = C' \log |H| + 1 \leq (\log n)^{O(1)}$. This completes the proof.

**Remark.** In the proof of Theorem 2.12 it seems tempting to write $\text{Tr}(F, \alpha_0 n) = B_0 n$ with $\alpha_0 \ll \alpha$, so that $B_0 \ll B$. Then, Corollary 2.10 would mean that the expected trace on $\alpha_0 n$ random vertices would be roughly $\frac{|F|}{B^{1 - \log(1 + \alpha)}}$, which seemingly contradicts the fact that our lower bound is tight! The reason this cannot happen is that, for any $\alpha_0$ that is not too small, the average trace on $\alpha_0 n$ vertices is substantially smaller than that on $\alpha n$ vertices, the maximal traces can actually be the same. Indeed, in the upper bound construction one can show that $\text{Tr}(F, \alpha_0 n) \approx \text{Tr}(F, \alpha n)$ for any $\alpha_0 \leq \alpha$ that is not too small.

### 3 Upper Bounds

#### 3.1 Upper Bound for the Sparse Kruskal-Katona Theorem

In this subsection we show that the parameters in Theorem 2.7 are best possible up to the error term. Formally, we prove the following.

**Theorem 3.1** (Upper bound for sparse Kruskal-Katona). Let $n, k, x \in \mathbb{N}^+$, $r \geq 1$ and $0 < \alpha \leq 1$ with $3r \leq k \leq x \leq n^{1/6}$ and $n \leq \alpha^{k+1} n^r \leq \binom{n}{k} n$. There exists a $k$-graph $F$ with $n$ vertices, $|F| = n^r$ edges, and $\text{span}(F, \alpha n) \leq 6\binom{x}{k} n$ such that for every $0 \leq i \leq k$ we have $\left| \binom{n}{i} \right| \leq \binom{\binom{x}{k}}{i} |F|$.

We will need the following lemma relating hypergeometric and binomial random variables.

**Lemma 3.2.** Let $H$ be a hypergeometric random variable with parameters $(n, q, x)$, and let $B$ be a binomial random variable with parameters $(x, q/n)$. If $x \leq \sqrt{n}$ then for every $0 \leq h \leq x$ we have $\text{Pr}[H = h] \leq 2 \text{Pr}[B = h]$.

**Proof.** Put $\alpha = q/n$. We have

$$
\text{Pr}[H = h] = \frac{\binom{x}{h} \binom{n-x}{q-h}}{\binom{n}{q}} = \frac{\binom{x}{h} \binom{n-x}{n-h}}{\binom{n}{q}} = \frac{(x)_h (n-x)!}{h!} \frac{q!}{(q-h)!} \frac{(n-q)!}{(n-q-(x-h))!},
$$

$$
= \frac{x}{h} \prod_{t=0}^{h-1} q-t \cdot \prod_{t=0}^{x-h-1} n-q-t \cdot \prod_{t=0}^{x-h-1} n-h-t \leq \frac{x}{h} \left( \frac{q}{n} \right)^h \frac{(n-q)^{x-h}}{(n-h)^{x-h}} \leq 2 \left( \frac{x}{h} \right) \alpha^h (1 - \alpha)^{x-h} = 2 \text{Pr}[B = h],
$$

where the first inequality uses that $h \leq q$ (as otherwise $\text{Pr}[H = h] = 0$ and there is nothing to prove), the second inequality uses the lower bound in Claim 2.5 as $h \leq x \leq \sqrt{n} \leq n/2$ (for the last inequality we assume $\sqrt{n} \geq 2$, as otherwise $x \leq 1$ in which case $H = B$ so there is nothing to prove), and the third inequality uses the fact that $h(x-h) \leq x^2/4 \leq n/4$ as $x \leq \sqrt{n}$. \[\square\]

---

6I.e., $H = |X \cap I|$ where $X \subseteq [n]$ is a uniformly random subset of size $x$ and $I \subseteq [n]$ is a fixed subset of size $q$.\[\]
We will make use of the following version of Chernoff’s bound (c.f., e.g., [2], Appendix A).

Claim 3.3. Let \(X_1, \ldots, X_n\) be mutually independent random variables with \(X_i \in [0, 1]\), and put \(X = \sum_{i=1}^n X_i\), \(\mu = \mathbb{E}[X]\). Then \(\Pr(X \geq 6x) \leq \exp(-x)\) for every \(x \geq \mu/3\).

Proof. Put \(y = 3x \geq \mu\). Then \(\Pr(X \geq 6x) = \Pr(X \geq 2y) \leq \Pr(X - \mu \geq y) \leq \exp(-y/3) = \exp(-x)\), where the second inequality is Chernoff’s large-deviation bound (again using \(y \geq \mu\)). \(\square\)

Proof of Theorem 3.1. Let \(\ell \in \mathbb{N}\) satisfy

\[
\frac{n^r}{\ell} \leq \ell \leq \frac{n}{\alpha^k},
\]

which is well defined by the statement’s upper bound on \(\alpha^kn^r\). Let \(S_1, \ldots, S_\ell\) be \(\ell\) independent uniformly random size-\(x\) subsets of \([n]\). We let \(F\) be the \(k\)-graph on \([n]\) consisting of a complete \(k\)-graph on each \(S_j\), that is, \(F = ([n], \bigcup_{j=1}^{\ell} \binom{S_j}{k})\).

We next analyze the random \(k\)-graph \(F\) constructed above. Let \(E_1\) be the event that every two sets \(S_j, S_{j'}\) \((1 \leq j \neq j' \leq \ell)\) intersect in fewer than \(t := 3r\) elements, and let \(E_2\) be the event that \(\text{span}(F, \alpha n) \leq 6(\frac{x}{k})n\). We first show that the proof would follow by proving that

\[
\Pr(E_1 \text{ and } E_2) > 0. \tag{17}
\]

To see this first note that, by construction, \(\left|\binom{\binom{F}{i}}{\ell}\right| \leq \sum_{j=1}^{\ell} \binom{|S_j|}{i} = \ell^i\) for every \(0 \leq i \leq k\). Furthermore, note that the event \(E_1\) implies that the cliques \(S_{j,k}\) are (edge-)disjoint by the statement’s assumption \(k \geq t\), and so

\[
E_1 \text{ implies } |F| = \ell^k. \tag{18}
\]

Therefore, (17) implies the existence of an \(n\)-vertex \(k\)-graph \(F\) satisfying:

- \(|F| = \ell^k \geq n^r\), using (18) and the lower bound in (16),
- \(\text{span}(F, \alpha n) \leq 6(\frac{x}{k})n\), and
- \(\left|\binom{\binom{F}{i}}{\ell}\right| \leq \ell^i\) = \(\binom{\ell}{i}\) \(|F|\), using (18),

from which the proof immediately follows by taking an arbitrary subgraph of \(F\) with \(n^r\) edges.

To prove (17) we first claim that

\[
\Pr(E_1) \geq \frac{1}{2}. \tag{19}
\]

Denote by \(B\) the binomial random variable with parameters \((x, p)\) where \(p = x/n\). For every \(j \neq j'\) we have

\[
\Pr(|S_j \cap S_{j'}| \geq t) \leq 2 \Pr(B \geq t) \leq 2 \cdot \left(\frac{x}{t}\right)^t \leq (xp)^t = x^{2t}/n^t = x^{6r}/n^{3r} \leq n^{-2r} \leq \ell^{-2},
\]

where the first inequality uses Lemma 3.2 using the statement’s upper bound on \(x\), the penultimate inequality again uses the statement’s upper bound on \(x\), and the last inequality uses the upper bound in (16) together with the statement’s lower bound on \(\alpha kn^r\). This implies, by taking the union bound over all \(\binom{\ell}{2}\) unordered pairs \(1 \leq j \neq j' \leq \ell\), that with probability at least \(\frac{1}{2}\) all set pairs \(S_j, S_{j'}\) with \(j \neq j'\) intersect at fewer than \(3r\) elements. This proves (19).
We now show that \( \Pr(E_2) \geq 1/2 \) (that is, that span(\( \mathcal{F} \), \( \alpha n \)) \( \leq 6(\frac{n}{k})^2 n \) except with probability smaller than \( 1/2 \)), which would prove (17). Fix \( I \subseteq [n] \) of size \( q = \alpha n \). Note that \( \mathcal{F}[I] = \bigcup_{j=1}^{\ell} (S_j \cap I) \). For each \( 1 \leq j \leq \ell \) consider the random variable \( X_j = (|S_j \cap I|) \), let \( X = \sum_{j=1}^{\ell} X_j \), and note that \( |\mathcal{F}[I]| \leq X \). We have

\[
\mathbb{E}(X_j) = \sum_{h=k}^{x} \binom{h}{k} \Pr[|S_j \cap I| = h] \leq 2 \sum_{h=k}^{x} \binom{x}{h} \binom{h}{k} \alpha^h (1 - \alpha)^{x-h} = 2 \sum_{h=k}^{x} \binom{x}{k} \binom{x-k}{h-k} \alpha^h (1 - \alpha)^{x-h} = 2 \alpha^k \binom{x}{k} \sum_{j=0}^{x-k} \binom{x-k}{j} \alpha^j (1 - \alpha)^{(x-k)-j} = 2 \alpha^k \binom{x}{k},
\]

where the first inequality follows from Lemma 3.2 using the statement’s upper bound on \( x \). Thus, by linearity of expectation,

\[
\mathbb{E}(X) \leq \ell \cdot 2 \alpha^k \binom{x}{k} \leq 2 \binom{x}{k} n, \tag{20}
\]

where the second inequality uses the upper bound in (16). Since \( X_j \leq \binom{x}{k} \) for every \( 1 \leq j \leq \ell \), we have that \( X/\binom{x}{k} \) is a sum of mutually independent random variables each in \([0, 1]\). We thus apply Claim 3.3 on \( X/\binom{x}{k} \), using the fact that \( n \geq \frac{1}{3} \mathbb{E}(X/\binom{x}{k}) \) by (20), to deduce that

\[
\Pr \left[ X \geq 6 \binom{x}{k} n \right] = \Pr \left[ X/\binom{x}{k} \geq 6n \right] \leq \exp(-n) < 2^{-n}.
\]

Using the union bound over all \( \binom{n}{q} \leq 2^{2^n} \) choices of \( I \subseteq [n] \) with \( |I| = \alpha n \) we deduce that, except with probability smaller than \( 1/2 \), for every \( I \subseteq [n] \) with \( |I| = \alpha n \) it holds that \( |\mathcal{F}[I]| \leq 6(\frac{n}{k})^2 n \). As mentioned before, together with (19) this proves (17) and so we are done. 

\[ \square \]

### 3.2 Traces upper bound

In this subsection we complete the proof of Theorem 1.4 by proving the upper bound \( \text{Tr}(n, \alpha n, \alpha n) \leq O(n^\mu) \), with \( \mu \) as in (1), for every \( 0 \leq a \leq 1 \) and constant \( r \). A proof can be obtained from the proof of Theorem 3.1 with some effort. For completeness, we give here a self-contained proof.

Put \( x = (\mu - 1) \log n \). Let \( S_1, \ldots, S_\ell \) be \( \ell \) independent uniformly random subsets of \([n]\), each of size \( x \), where

\[
\frac{1}{2} \ell := \frac{n^r}{2^x} = \frac{n^\mu}{(1+\alpha)^r}, \tag{21}
\]

Let the family \( \mathcal{F} \subseteq 2^{[n]} \) consist of the union over \( j \) of all subsets of the set \( S_j \); that is, \( \mathcal{F} = \bigcup_{j=1}^{\ell} 2^S_j \). Let \( E_1 \) be the event that \( |\mathcal{F}| \geq n^r \), and let \( E_2 \) be the event that \( \text{Tr}(\mathcal{F}, \alpha n) \leq 8n^\mu \). Note that the proof would follow by showing that

\[
\Pr(E_1 \text{ and } E_2) > 0. \tag{22}
\]

First, we claim that

\[
\Pr(E_1) \geq \frac{1}{2}. \tag{23}
\]

Put \( t = 3r \). Denote by \( B \) the binomial random variable with parameters \((x, p)\) where \( p = x/n \). For every \( j \neq j' \) we have

\[
\Pr(|S_j \cap S_{j'}| \geq t) \leq \Pr(B \geq t) \leq \binom{s}{t} p^t \leq (xp)^t = x^{6r}/n^{3r} \leq n^{-2r} \leq \ell^{-2},
\]
where the first inequality uses Lemma 3.2, and the last inequality uses (21). Conditioned on the above we have, by taking the union bound over all \( \binom{\ell}{2} \) pairs of sets, that

\[
|\mathcal{F}| \geq \ell \left( 2^x - \binom{x}{t} \right) \geq \ell \cdot \frac{1}{2} 2^x = n^r,
\]

where the second inequality uses \( 3r = t \leq x/2 \) (recall \( r \) is constant), and the equality uses (21). This proves (23).

We now show that \( \Pr(E_2) \geq \frac{1}{2} \) (that is, \( \text{Tr}(F, \alpha n) \leq 8n^\mu \) except with probability smaller than \( \frac{1}{2} \)), thus proving (22). Fix \( I \subseteq [n] \) of size \( q = \alpha n \). Note that

\[
|F_I| = \ell \bigcup_{j=1}^\ell \{S \cap I \mid S \subseteq S_j\} = \ell \{2^{S_j \cap I}\}.
\]

For each \( 1 \leq j \leq \ell \) consider the random variable \( X_j = |2^{S_j \cap I}| \), let \( X = \sum_{j=1}^\ell X_j \) and note that \( |F_I| \leq X \). We have

\[
\mathbb{E}(X_j) = \sum_{h=0}^x 2^h \Pr[|S_j \cap I| = h] \leq 2(1-\alpha)^x \sum_{h=0}^x \binom{x}{h} \left( \frac{2\alpha}{1-\alpha} \right)^h = 2(1-\alpha)^x \left( 1 + \frac{2\alpha}{1-\alpha} \right)^x = 2(1+\alpha)^x,
\]

where the first inequality uses Lemma 3.2. Thus, by linearity of expectation and by (21),

\[
\mathbb{E}(X) \leq \ell \cdot 2(1+\alpha)^x = 4n^\mu.
\]

Note that, by our choice of \( x \) at the beginning of the proof,

\[
n^\mu/2^x = n^\mu/2^{(\mu-1)\log n} = n.
\]

(25)

Since \( X_j \leq 2^x \) for every \( 1 \leq j \leq \ell \), we have that \( X/2^x \) is a sum of mutually independent random variables each in \([0,1]\). We thus apply Claim 3.3 on \( X/2^x \), using the fact that \( \mathbb{E}(X/2^x) \leq 4n \) by (24) and (25), to deduce that

\[
\Pr\left[ X \geq 8n^C \right] = \Pr\left[ X/2^x \geq 8n \right] \leq \exp \left( -\frac{4}{3}n \right) < 2^{-n}.
\]

Using the union bound over all \( \binom{n}{\alpha n} \leq \frac{1}{2} 2^n \) choices of \( I \subseteq [n] \) with \( |I| = \alpha n \) we deduce that, except with probability smaller than \( 1/2 \), for every \( I \subseteq [n] \) with \( |I| = \alpha n \) it holds that \( |F_I| \leq 8n^\mu \). As mentioned before, together with (19) this proves (17) and so we are done.

### 4 Applications

In this section we give two easy applications of our results on traces, in geometry and in graph theory.

#### 4.1 Separating halfspaces using few points

We recall the necessary definitions and the statement of this application. Let \( \mathcal{H} \) be a family of halfspaces\(^7\) in \( \mathbb{R}^d \), and let \( P \) be a set of points in \( \mathbb{R}^d \). We say that \( P \) separates \( \mathcal{H} \) if for every pair of distinct halfspaces \( H_1 \neq H_2 \in \mathcal{H} \) there is a point in \( P \) that lies in one and outside the other. Given \( P \) and \( \mathcal{H} \) such that \( P \) separates \( \mathcal{H} \), it is interesting to ask how few points in \( P \) can we choose while still separating many of the hyperplanes in \( \mathcal{H} \).

\(^7\)A halfspace in \( \mathbb{R}^d \) consist of all points above a hyperplane.
Proposition 4.1. Fix \( r > 1 \), let \( \mathcal{H} \) be a family of \( n^r \) halfspaces in \( \mathbb{R}^d \) and let \( P \subset \mathbb{R}^d \) be a set of \( n \) points such that \( P \) separates \( \mathcal{H} \). Then there is a subset \( P' \subseteq P \) of at most \( \frac{n}{\log n}^8 \) points and a subset \( \mathcal{H}' \subseteq \mathcal{H} \) of \( \tilde{\Omega}(n^{\frac{r+1}{2}}) \) halfspaces such that \( P' \) separates \( \mathcal{H}' \).

Proof. Let \( \mathcal{F} \) be the hypergraph on \( P \) with edge set \( \{ H \cap P \mid H \in \mathcal{H} \} \). By assumption, for any pair of distinct halfspaces \( H_1 \neq H_2 \in \mathcal{H} \) there is a point \( p \in P \) such that \( p \in H_1 \) and \( p \notin H_2 \), or \( p \in H_2 \) and \( p \notin H_1 \). In particular, \( H_1 \cap P \neq H_2 \cap P \), and therefore \( |\mathcal{F}| = |\mathcal{H}| = n^r \). Applying Theorem 2.12 with \( \alpha = \frac{1}{\log n} \), there exists a subset \( P' \subseteq P \) of size \( \frac{n}{\log n}^8 \) such that \( |\mathcal{F}_{P'}| = \tilde{\Omega}(n^{\frac{r+1}{2}}) \) (the choice of \( \alpha \) so small is allowed by Corollary 2.9). Note that \( \mathcal{F}_{P'} = \{ H \cap P' \mid H \in \mathcal{H} \} \). Let \( \mathcal{H}' \subseteq \mathcal{H} \) be obtained by assigning to each member \( Q \) of \( \mathcal{F}_{P'} \) an arbitrary halfspace \( H' \in \mathcal{H}' \) with \( H' \cap P' = Q \). By construction, \( |\mathcal{H}'| = |\mathcal{F}_{P'}| \), hence it remains to show that \( P' \) separates \( \mathcal{H}' \). By construction, for every pair of distinct halfspaces \( H_1 \neq H_2 \in \mathcal{H}' \) we have \( H_1 \cap P' \neq H_2 \cap P' \). This means that there exists a point \( p' \in P' \) such that either \( p' \in H_1 \) and \( p' \notin H_2 \), or \( p' \in H_2 \) and \( p' \notin H_1 \), thus completing the proof.

In fact, halfspaces can be replaced in Proposition 4.1 by any family of subsets of \( \mathbb{R}^d \), as long as the set of points \( P \) separates them. Indeed, in the proof of Proposition 4.1 we considered for each halfspace only the subset of points in \( P \) that are contained in it. The condition that \( P \) separates \( \mathcal{H} \) implies that all the corresponding subsets of \( P \) are distinct, and that is all that is really needed. However, it does seem reasonable to expect that for “favorable” geometric objects, such as halfspaces, better bounds hold—as we believe is the case.

Remark 4.2. It is worth noting that the number of halfspaces in Proposition 4.1 is indeed necessarily polynomial if the dimension \( d \) is constant; in fact, an exact formula is known, see [10], which in particular implies that the number of halfspaces in \( \mathbb{R}^d \) separated by \( n \) points is at most \( O(n^d) \). As mentioned above the assertion of the proposition holds also when we replace halfspaces by any family of \( n^r \) subsets, for example, general convex sets.

4.2 Retaining independent sets in induced subgraphs

An independent set in a graph is a vertex subset that spans no edges. Given a graph \( G \), one can ask how many independent sets are retained in small subgraphs of \( G \). Using our result for traces, we easily obtain that if \( G \) has a polynomial number of independent sets then it has a subset of \( o(n) \) vertices retaining asymptotically more than square root of the number of independent sets.

Proposition 4.3. Let \( G = (V, E) \) be a graph, and assume that the number of independent sets in \( G \) is at most \( n^r \) for some fixed \( r > 1 \). Then there exists a subset \( V' \subseteq V \) of at most \( \frac{n}{\log n}^8 \) vertices such that the number of independent sets in the induced subgraph \( G[V'] \) is at least \( \tilde{\Omega}(n^{\frac{r+1}{2}}) \).

Proof. Let \( \mathcal{F} \) be the hypergraph on \( V \) whose edges are the independent sets in \( G \). By Theorem 2.12, there is a subset \( S \) of \( V(\mathcal{F}) = V \) with \( \alpha n = \frac{n}{\log n} \) vertices such that \( |\mathcal{F}_S| \), the number of projections of the edges of \( \mathcal{F} \) onto \( S \), is at least \( \tilde{\Omega}(n^{\frac{r+1}{2}}) \). The proof follows by observing that, by definition and construction,
\[
\mathcal{F}_S = \{ I \cap S : I \in \mathcal{F} \} = \{ I \cap S : I \text{ is an independent set of } G \} = \{ I : I \text{ is an independent set of } G[S] \}.
\]

\footnote{In fact, by part 2 of Corollary 2.9, we can find a set of \( n^{1-\epsilon} \) points and slightly increase the error term in the \( \tilde{\Omega}(n^{\frac{r+1}{2}}) \) notation.}
We note that, since Theorem 2.12 is such an abstract statement, we could have replaced the
notion of independent sets in a graph by any other monotone family (i.e., a family closed under
taking subsets) of vertex subsets. Thus, for example, a statement analogous to Proposition 4.3
holds for independent sets in hypergraphs, for subsets inducing a subgraph not containing some
fixed graph $H$, and more.

References


pp.


1409.3033.


[10] E. F. Harding, The number of partitions of a set of $N$ points in $k$ dimensions induced by

O. H., eds.), Akadémiai Kiadó, Budapest.


niques*, Univ. California Press, Berkeley (1963), 251-278.


[18] V. N. Vapnik and A. Ya. Chervonenkis, On the uniform convergence of relative frequencies of
A Proof of an Auxiliary Lemma

For the proof of Lemma 2.4 we will need the following claim, which follows easily from Newton’s generalized binomial formula.

Claim A.1. For every real $x > 0$ we have $2^{x-1} < \sum_{i=0}^{\lfloor x \rfloor} \binom{x}{i} \leq 2^x$.

Proof. We bound the summation in the statement from above and from below as follows;

$$2^{x-1} < 2^{|x|} \leq \sum_{i=0}^{|x|} \binom{|x|}{i} \leq \sum_{i=0}^\infty \binom{x}{i} = (1 + 1)^x = 2^x,$$

where the last inequality follows from

$$\sum_{i=0}^\infty \binom{x}{i} - \sum_{i=0}^{|x|} \binom{x}{i} = \sum_j \left(\binom{x}{j} + \binom{x}{j+1}\right) = \sum_j \binom{x+1}{j+1} = \sum_j \frac{(x+1)x\cdots(x-|x|)}{(j+1)!} \prod_{k=|x|+1}^{j-1} (x-k) \geq 0$$

with $j$ running over the values $|x| + 1, |x| + 3, \ldots$, and where the first equality follows from Newton’s generalized binomial formula.

Proof of Lemma 2.4. For real $z \geq 0$ we henceforth abbreviate $(\frac{x}{\leq z}) = \sum_{i=0}^{\lfloor x \rfloor} \binom{x}{i}$. If $k \geq \lceil \frac{x}{2} \rceil$ we are done since

$$2 \sum_{i=0}^k \binom{x}{i} \gamma^i \geq 2 \sum_{i=0}^{\lfloor \frac{x}{2} \rfloor} \binom{|x|}{i} \gamma^i = \sum_{i=0}^{\lfloor \frac{x}{2} \rfloor} \left(\binom{|x|}{i} + \binom{|x|}{|x|-i}\right) \gamma^i \geq \sum_{i=0}^{\lfloor \frac{x}{2} \rfloor} \left(\binom{|x|}{i} \gamma^i + \binom{|x|}{|x|-i} \gamma^{|x|-i}\right) \geq \sum_{i=0}^{\lfloor \frac{x}{2} \rfloor} \left(1 + \gamma\right)^{|x|} \geq \frac{1}{2} (1 + \gamma)^x \geq \frac{1}{2} \left(\frac{x}{\leq x}\right)^{\log(1+\gamma)} \geq \frac{1}{2} \left(\frac{x}{\leq k}\right)^{\log(1+\gamma)},$$

where the penultimate inequality follows from Claim A.1, and dividing over by 2 gives the desired result.

We thus assume $k \leq \lceil \frac{x}{2} \rceil - 1$ for the remainder of the proof. Put

$$y = \log \left(\frac{x}{\leq k}\right) \quad \text{and} \quad K = \lfloor y \rfloor + 1.$$

We have $0 \leq y \leq x - 1$, which follows using the fact that for $i \leq \lfloor \frac{x}{2} \rfloor$ we have $(\frac{x}{i}) \leq (\frac{x}{|x|-i})$ as follows;

$$1 \leq 2 \left(\frac{x}{\leq k}\right) \leq 2 \left(\lfloor \frac{x}{2} \rfloor - 1\right) \leq \sum_{i=0}^{\lfloor \frac{x}{2} \rfloor - 1} \left(\binom{x}{i} + \binom{x}{|x|-i}\right) \leq \left(\frac{x}{\leq x}\right) \leq 2^x,$$

where the first inequality follows since $1 \leq x$, the second inequality uses our assumption $k \leq \lfloor \frac{x}{2} \rfloor - 1$, and the last inequality uses Claim A.1. Furthermore, we have $K \geq k$, since otherwise $k \geq K + 1$ and thus

$$\left(\frac{x}{\leq k}\right) \geq \left(\frac{x}{\leq K + 1}\right) \geq \left(\frac{y + 1}{\leq K + 1}\right) = \left(\frac{y + 1}{\leq [y] + 2}\right) \geq \left(\frac{y + 1}{\leq [y] + 1}\right) > 2^y,$$

18
a contradiction, where the second and third inequalities used the fact that for a positive number $t$, \( \lfloor t \rfloor + 1 \) is non-negative (and $K = \lfloor y \rfloor + 1$), and the last inequality uses Claim A.1.

Let $a_0, \ldots, a_k$ and $b_0, \ldots, b_K$ be given by

$$a_i = \binom{x}{i} \text{ for } 0 \leq i \leq k, \quad b_i = \binom{y}{i} \text{ for } 0 \leq i \leq K - 1, \quad \text{and } b_K = \sum_{i=0}^{k} a_i - \sum_{i=0}^{K-1} b_i.$$ 

We have that $b_K \geq 0$ since

$$\sum_{i=0}^{K-1} b_i = \binom{y}{y} \leq 2^y = \binom{x}{k} = \sum_{i=0}^{k} a_i,$$

where the inequality follows from Claim A.1. Summarizing the properties of the sequences $(a_i)_{i=0}^{k}$ and $(b_i)_{i=0}^{K}$, we have $\sum_{i=0}^{k} a_i = \sum_{i=0}^{k} b_i$, and for every $0 \leq i \leq k - 1$ we have $a_i \geq b_i$ (recall $x > y \geq 0$). Denote $\Delta_i = a_i - b_i \geq 0$ for every $0 \leq i \leq k - 1$. We have

$$\sum_{i=0}^{\lfloor y \rfloor} \binom{y}{i} \gamma^i \leq \sum_{i=0}^{K-1} b_i \gamma^i = \sum_{i=0}^{k-1} (a_i - \Delta_i) \gamma^i + \sum_{i=k}^{K-1} b_i \gamma^i$$

$$\leq \sum_{i=0}^{k-1} (a_i - \Delta_i) \gamma^i + \left( \sum_{i=k}^{K-1} b_i \right) \gamma^k = \sum_{i=0}^{k-1} (a_i - \Delta_i) \gamma^i + \left( \sum_{i=0}^{k-1} \Delta_i + a_k \right) \gamma^k$$

$$\leq \sum_{i=0}^{k-1} a_i \gamma^i - \sum_{i=0}^{k-1} \Delta_i \gamma^k + \left( \sum_{i=0}^{k-1} \Delta_i + a_k \right) \gamma^k = \sum_{i=0}^{k} a_i \gamma^i = \sum_{i=0}^{k} \binom{x}{i} \gamma^i.$$

We thus showed that

$$\sum_{i=0}^{k} \binom{x}{i} \gamma^i \geq \sum_{i=0}^{\lfloor y \rfloor} \binom{y}{i} \gamma^i \geq \sum_{i=0}^{\lfloor y \rfloor} \binom{\lfloor y \rfloor}{i} \gamma^i = (1 + \gamma)^{\lfloor y \rfloor} \geq \frac{1}{2} (1 + \gamma)^y = \frac{1}{2} \left( \frac{x}{k} \right)^{\log(1+\gamma)},$$

which completes the proof.