Defect Sauer Results

Béla Bollobás
University of Cambridge and Louisiana State University

A.J. Radcliffe
University of Nebraska-Lincoln

Abstract: In this paper we present a unified account of various results concerning traces of set systems, including the original lemma proved independently by Sauer [14], Shelah [15], and Vapnik and Chervonenkis [16], and extend these results in various directions. Included are a new criterion for a set system to be extremal for the Sauer inequality and upper and lower bounds, obtained by random methods, for the trace of a set system of size n^r guaranteed on some αn sized subset of $\{1, 2, \ldots, n\}$.

Notation

A *set system* is a subset of $\mathcal{P}(n)$, the power set of $[n] = \{1, 2, \ldots, n\}$. The complement of a set $I \subset [n]$ is written I^c, all other set differences are written out explicitly, e.g., $\mathcal{P}(n) \setminus \mathcal{A}$. The system consisting of all sets of size k is written $[n]^{(k)}$, while $\mathcal{A}^{(k)} = \mathcal{A} \cap [n]^{(k)}$, for any set system $\mathcal{A}$. We define $[n]^{(\leq k)}$, $\mathcal{A}^{(\leq k)}$, etc. similarly. Given $\mathcal{A} \subset \mathcal{P}(n)$ and $I \subset [n]$ we write $I + \mathcal{A}$ for $\{A \cup I : A \in \mathcal{A}\}$.

Introduction

A striking extremal result concerning set systems was proved in the early 70s by Sauer [14], Shelah [15] and Vapnik and Chervonenkis [16].

In order to state the result, we say that a set $I \subset [n]$ is *traced* by a set system $\mathcal{A} \subset \mathcal{P}(n)$ if the collection of intersections $\mathcal{A}|_I = \{\mathcal{A} \cap I : A \in \mathcal{A}\}$ is the whole of $\mathcal{P}(I)$. Sauer's lemma (as it has become known) states that any set system $\mathcal{A}$ traces at least $|\mathcal{A}|$ sets in $\mathcal{P}(n)$. (This formulation is due to Pajor [12].) In particular, if $|\mathcal{A}| > \sum_{i=0}^{k} \binom{n}{i}$, then $\mathcal{A}$ must trace some set of size k.

This theorem leads naturally in several directions. Both Frankl [9] and Dudley [7] have characterised maximal systems which cover no 2-set, and several authors have considered 'defect Sauer' results. Such results address the problem of determining how large a set system $\mathcal{A} \subset \mathcal{P}(n)$ must be before one is guaranteed a trace of at least M on some k-set. We write $\mathrm{tr}\,(\mathcal{A}, k)$ for the maximum size of $\mathcal{A}|_I$ over all $I \in [n]^{(k)}$ and we shall use the arrow notation $(N, n) \to (M, k)$ to mean that whenever $\mathcal{A} \subset \mathcal{P}(n)$ has $|\mathcal{A}| \geq N$, then $\mathrm{tr}\,(\mathcal{A}, k) \geq M$. Thus the Sauer lemma states that if $N > \sum_{i=0}^{k} \binom{n}{i}$ then $(N, n) \to (2^k, k)$.

In this paper we proceed along both lines. The first section is concerned with extremal systems. There are two natural meanings of extremality; either that $\mathcal{A}$ is maximal tracing no k-set or that $\mathcal{A}$ traces exactly $|\mathcal{A}|$ sets. It is the latter we shall be most concerned with. Extending the work of Bollobás, Leader and Radcliffe [4] we present a new criterion for a system to be extremal and analyse the relationship between Sauer's lemma and the Reverse Kleitman Inequality of [4].

In the later sections we consider various cases of the defect Sauer problem. First we prove an extension of some of the results of Frankl's [9] and in the final section we turn to the case where N is a polynomial function of n and k is proportional to n. In the positive direction we show that if r is fixed, $\alpha \in (0, 1)$ and $n \to \infty$ then

$$(n^r, n) \to \big((1 - o(1))n^{\lambda r}, \alpha n\big).$$

Here λ is a function of α only. We also construct examples, using random techniques, to show that

$$\left(\sum_{i=0}^{r}\binom{n}{i}, n\right) \not\to \left(\sum_{i=0}^{r}\binom{n/2}{i} - (1-o(1))2^{-r}\binom{n}{r}, n/2\right)$$

§1 Extremal cases for Sauer's inequality.

The statement of the Sauer-Shelah lemma given in the introduction – that a set system $\mathcal{A} \subset \mathcal{P}(n)$ must trace at least $|\mathcal{A}|$ sets – was first made explicit by Pajor ([12]). The same result, in disguised form, was proved by Leader and the current authors in [4]. We say that a set $I \subset [n]$ is *strongly traced* by $\mathcal{A} \subset \mathcal{P}(n)$ when a full copy of $\mathcal{P}(I)$ can be found somewhere in $\mathcal{A}$, i.e., when there is some $B \subset I^c$ such that $B + \mathcal{P}(I) \subset \mathcal{A}$. The reverse Kleitman inequality from [4] states that a set system $\mathcal{A}$ strongly traces at most $|\mathcal{A}|$ sets. (It seems clear now that a better name for this result would be the reverse Sauer inequality, and that is how we shall refer to it from now on.) These two results are rather easily seen to be equivalent. Indeed, the sets that $\mathcal{A}$ fails to trace are exactly the complements of those which $\mathcal{P}(n) \setminus \mathcal{A}$ strongly traces. So writing

$$T(\mathcal{A}) = \{I \subset [n] \,:\, \mathcal{A}|_I = \mathcal{P}(I)\}$$
$$S(\mathcal{A}) = \{I \subset [n] \,:\, \exists\, J \subset I^c \text{such that } J + \mathcal{P}(I) \subset \mathcal{A}\},$$

one has

$$\mathcal{P}(n) \setminus T(\mathcal{A}) = \{I^c \,:\, I \in S(\mathcal{P}(n) \setminus \mathcal{A})\}.$$

Therefore $|T(\mathcal{A})| + |S(\mathcal{P}(n) \setminus \mathcal{A})| = 2^n$ and $|T(\mathcal{A})| \geq |\mathcal{A}|$ iff $|S(\mathcal{P}(n) \setminus \mathcal{A})| \leq |\mathcal{P}(n) \setminus \mathcal{A}|$. This equivalence allows one to take advantage of the various criterion given in [4] for a set system $\mathcal{A}$ to be *extremal for reverse Sauer*, i.e., for $|S(\mathcal{A})| = |\mathcal{A}|$ to hold.

Alon's [1] proof of the Sauer lemma, Pajor's proof of the stronger inequality and one of the proofs of the reverse Sauer inequality from [4] proceed by the technique of down compression, which we outline here. The idea is to take a set system $\mathcal{A}$ about which one knows nothing and shift it around until it is in some canonical position, all the time retaining control over the size of the system and other relevant features. The *down compressions* $C_i, i \in [n]$, are defined by

$$C_i(\mathcal{A}) = \{A - \{i\} \,:\, A \in \mathcal{A}\} \cup \{A \,:\, i \in A \in \mathcal{A}, A - \{i\} \in \mathcal{A}\}.$$

In other words, from every pair A, $A\Delta\{i\}$ the system $C_i(\mathcal{A})$ takes the same number as does $\mathcal{A}$, choosing the set without i for preference. Given a sequence $\sigma = (\sigma_j)_1^m$ of elements of $[n]$ we write $C_\sigma(\mathcal{A})$ for the system obtained by applying compressions to $\mathcal{A}$ in the order specified by σ. So $C_\sigma(\mathcal{A}) = C_{\sigma(m)}(C_{\sigma(m-1)}(\ldots C_{\sigma(1)}(\mathcal{A})\ldots))$. If σ is a permutation of $[n]$ then the end result of applying all these compressions is a *down-set*; a system $\mathcal{D}$ such that $J \subset I \in \mathcal{D}$ implies $J \in \mathcal{D}$. The crucial lemma regarding down compressions follows.

Lemma 1 ([1], [9], [4]). *If $\mathcal{A} \subset \mathcal{P}(n)$ and $i \in [n]$ then*

$$|C_i(\mathcal{A})| = |\mathcal{A}| \qquad T(C_i(\mathcal{A})) \subset T(\mathcal{A}) \qquad S(C_i(\mathcal{A})) \supset S(\mathcal{A}). \qquad \blacksquare$$

With this in hand the proof of, e.g., the reverse Sauer inequality is immediate. Given $\mathcal{A} \subset \mathcal{P}(n)$, pick any permutation $\pi \in \Sigma_n$ and let $\mathcal{D} = C_\pi(\mathcal{A})$. This system is a down set satisfying $|\mathcal{D}| = |\mathcal{A}|$ and $S(\mathcal{D}) \supset S(\mathcal{A})$. For down sets however it is clear that $S(\mathcal{D}) = \mathcal{D}$ so $|S(\mathcal{A})| \leq |S(\mathcal{D})| = |\mathcal{D}| = |\mathcal{A}|$.

One should notice that the symmetry of $\mathcal{P}(n)$ which, for fixed $I \subset [n]$, sends A to $A \Delta I$ preserves both $S(\mathcal{A})$ and $T(\mathcal{A})$. We shall use this frequently, usually just appealing to 'symmetry'.

The next result shows that the two conditions of extremality for Sauer's inequality and the reverse Sauer inequality are equivalent (which is *not* a trivial consequence of the observation above that the two inequalities are equivalent).

Theorem 2. *Given a set system $\mathcal{A} \subset \mathcal{P}(n)$ the following are equivalent:*

(1) $\mathcal{A}$ is extremal for Sauer

(2) $\mathcal{P}(n) \setminus \mathcal{A}$ is extremal for reverse Sauer

(3) $\mathcal{A}$ has a unique down compression. In other words for all permutations $\pi, \sigma \in \Sigma_n$, $C_\pi(\mathcal{A}) = C_\sigma(\mathcal{A})$

(4) $\mathcal{A}$ is extremal for reverse Sauer.

Proof. (1) and (2) are clearly equivalent, simply from the proof that the two inequalities are equivalent. Concerning (3) notice that if $|T(\mathcal{A})| = |\mathcal{A}|$ then since $T(\mathcal{A}) \subset T(C_\pi(\mathcal{A})) = C_\pi(\mathcal{A})$ for any permutation $\pi \in \Sigma_n$, one gets $C_\pi(\mathcal{A}) = T(\mathcal{A}) = C_\sigma(\mathcal{A})$ for any pair of permuations π and σ. Thus (1) implies (3). On the other hand if $I = \{i_1, i_2, \ldots, i_k\}$ and $I^c = \{j_1, j_2, \ldots, j_{n-k}\}$ then I belongs to $C_{j_1 j_2 \ldots j_{n-k} i_1 i_2 \ldots i_k}(\mathcal{A})$ iff $I \in T(\mathcal{A})$. Similarly I belongs to $C_{i_1 i_2 \ldots i_k j_1 j_2 \ldots j_{n-k}}(\mathcal{A})$ exactly if I is a member of $S(\mathcal{A})$. So condition (3) tells us that $T(\mathcal{A}) = S(\mathcal{A})$, and therefore that $\mathcal{A}$ is extremal for both Sauer and reverse Sauer. In other words (3) implies (1), (2) and (4), and the circle is complete. $\blacksquare$

The next result explicates to some extent the condition of being extremal (for Sauer and reverse Sauer) by describing it in a manner which is invariant under complementation. However a structural description of extremal systems is still sorely lacking.

Following [4] we call a set of the form $\mathcal{C} = \{A \in \mathcal{A} : A \cap I = J\}$ a *chunk* of $\mathcal{A}$. It is simply the intersection of $\mathcal{A}$ with some k-dimensional face of the cube $\mathcal{P}(n)$, where $k = n - |I|$. This chunk is *self-complementary* if $\{C \Delta I^c : C \in \mathcal{C}\} = \mathcal{C}$. In other words, having fixed the behaviour inside I we check and see whether on I^c our chunk looks like a self-complementary set system. A chunk is *trivial* if it is either empty or of maximal size, i.e. $\mathcal{C} = J + \mathcal{P}(I^c)$.

Given two elements A_1 and A_2 of $\mathcal{P}(n)$ we say that they *span* the interval $[A_1 \cap A_2, A_1 \cup A_2] = \{B : A_1 \cap A_2 \subset B \subset A_1 \cap A_2\} \subset \mathcal{P}(n)$. If both A_1 and A_2 belong to a system $\mathcal{A}$ then we call $\mathcal{A} \cap [A_1 \cap A_2, A_1 \cup A_2]$ the *chunk of $\mathcal{A}$ they span*. Note that this is a chunk of $\mathcal{A}$; it's $\{A \in \mathcal{A} : A \cap (A_1 \Delta A_2)^c = A_1 \cap A_2\}$.

If $\mathcal{A} \subset \mathcal{P}(n)$ and $B \subset [n]$ write, again as in [4], $\mathcal{A}(B)$ for $\{I \in \mathcal{A} \cap \mathcal{P}(B^c) : I + \mathcal{P}(B) \subset \mathcal{A}\}$. We say that $\mathcal{A}$ is connected if it spans a connected subgraph of the graph on $\mathcal{P}(n)$ with edges $(A, A\Delta\{i\})$ for $A \in \mathcal{P}(n)$ and $i \in [n]$. We collect together in the following theorem the results from [4] of which we make use.

Theorem 3 ([4])**.** *Suppose $\mathcal{A} \subset \mathcal{P}(n)$.*

(i) If $\mathcal{A} \subset \mathcal{P}(n)$ is extremal then so are all chunks of $\mathcal{A}$.

(ii) $\mathcal{A}$ is extremal iff $\mathcal{A}(B)$ is connected for every $B \subset [n]$.

(iii) $\mathcal{A}$ is extremal iff every chunk of every $\mathcal{A}(B)$ is connected. In particular, if $\mathcal{A}$ is extremal and $B \subset [n]$ then any two elements of $\mathcal{A}(B)$ can be connected in the chunk of $\mathcal{A}(B)$ they span. ∎

We shall at least for the duration of the next proof, call any system of the form $I + \mathcal{P}(B)$, where I and B are disjoint subsets of $[n]$, a *B-cube* with *base* I, and say that two B-cubes $I_1 + \mathcal{P}(B)$ and $I_2 + \mathcal{P}(B)$ are *antipodal* if their bases are complements in B^c. (I.e. $I_1 = B^c \setminus I_2$.)

Theorem 4. *$\mathcal{A} \subset \mathcal{P}(n)$ is extremal iff it contains no non-trivial self-complementary chunks.*

Proof. To prove that no extremal set system contains any non-trivial self-complementary chunks it clearly suffices to prove that no extremal set system is non-trivial and self-complementary, since all chunks of extremal systems are extremal. Therefore suppose that

$\mathcal{A}$ is non-empty, self-complementary and extremal. We shall show that $\mathcal{A} = \mathcal{P}(n)$. Let B be maximal such that $\mathcal{A}$ contains some B-cube. (Such B exist since $\mathcal{A}$ is supposed non-empty.) We are done if $B = [n]$ so suppose it does not. We may assume by symmetry that $\emptyset \in \mathcal{A}(B)$. By self-complementarity $\{C^c : C \subset B\} = B^c + \mathcal{P}(B) \subset \mathcal{A}$. Now $\emptyset$ and B^c are both in $\mathcal{A}(B)$ so they are connected by a path in $\mathcal{A}(B)$. If the first vertex on that path is $\{i\}$ then $\mathcal{P}(B \cup \{i\}) \subset \mathcal{A}$, contradicting the maximality of B.

It remains to prove that if $\mathcal{A}$ contains no non-trivial self-complementary chunks then it is extremal. The proof is an inductive one, on n, with the base case $n = 1$ a trivial one.

First we'll show that $\mathcal{A}$ is connected. Suppose not, and let A_1 and A_2 belong to different components of $\mathcal{A}$ and be as close as poassible subject to that restriction. (To be precise we choose A_1 and A_2 with $|A_1 \Delta A_2|$ minimal. Then the chunk $\mathcal{B} = \mathcal{A} \cap [A_1 \cap A_2, A_1 \cup A_2]$ contains only A_1, A_2, else we could find a strictly closer pair of points from different components of $\mathcal{A}$ (any other point of $\mathcal{B}$ together with one of A_1, A_2). But A_1 and A_2 are antipodal in $\mathcal{B}$, contradicting the fact that $\mathcal{A}$ has no non-trivial self-complementary chunks.

Again let $\mathcal{A} \subset \mathcal{P}(n)$ be a system containing no non-trivial self-complementary chunks. We'll prove that $\mathcal{A}(B)$ is connected for every $B \subset [n]$, and deduce from Theorem 2 that $\mathcal{A}$ is extremal. Since the condition on $\mathcal{A}$ is inherited by chunks we know that every chunk of $\mathcal{A}$ (not equal to $\mathcal{A}$) is extremal, by the inductive hypothesis. This is enough to show that if $A_1 + \mathcal{P}(B)$ and $A_2 + \mathcal{P}(B)$ are non-antipodal B-cubes inside $\mathcal{A}$ then A_1 and A_2 belong to the same component in $\mathcal{A}(B)$, i.e., can be connected by a path in $\mathcal{A}(B)$. We need to remove the restriction on A_1, A_2. Suppose then that B, A_1, A_2 are such that $\mathcal{A}_1$ and A_2 belong to different components of $\mathcal{A}(B)$ and $d(A_1, A_2)$ is minimal over all such pairs. (By the remarks in the previous paragraph we may assume that $B \neq \emptyset$.) Without loss of generality we can take B to be $[k]$ where $k = |B|$. By restricting attention to the chunk of $\mathcal{P}([k+1, n])$ spanned by A_1 and A_2 we may suppose that they span it all. By applying the symmetry $A \mapsto A \Delta A_1$ to $\mathcal{A}$ we can in addition assume that $A_1 = \emptyset$ and therefore $A_2 = [k+1, n]$.

Claim. For all k-sets $K \subset [n]$ the system $\mathcal{A}$ contains exactly two K-cubes, which are antipodal. In other words $\mathcal{A}(K) = \{I_K, J_K\} \subset \mathcal{P}(K^c)$ and $I_K = K^c \setminus I_K$.

Proof of claim. We prove the claim by induction on $t = |K \setminus [k]|$. When $t = 0$ the claim simply describes the situation we have outlined above, with $I_{[k]} = \emptyset$ and $J_{[k]} = [k+1, n]$. Consider now any k-set $K \neq [k]$ and pick $i \in K \setminus [k]$ and $j \in [k] \setminus K$. Write K' for $(K \setminus \{i\}) \cup \{j\}$. Since $t' = |K' \setminus [k]| < t$ we know that $\mathcal{A}(K') = \{I', J'\} \subset \mathcal{P}(K'^c)$ and that I' and J' are complements inside K'^c. Now set $L = K \setminus \{i\} = K' \setminus \{j\}$ and write $\mathcal{A}_+$ for $\{A \in \mathcal{A} : i \in A\}$ and $\mathcal{A}_-$ for $\{A \in \mathcal{A} : i \notin A\}$. By the inductive hypothesis both $\mathcal{A}_+(L)$ and $\mathcal{A}_-(L)$ are connected so $\emptyset$ can is connected to $L^c \setminus \{i\}$ by a path entirely in $\mathcal{A}_-(L)$. Wherever there is a $\{j\}$-edge along the path there is an $(L \cup \{j\})$-cube, i.e., a K'-cube. Similarly there is a K'-cube inside $\mathcal{A}_+$. It only remains to establish that these are the only points of $\mathcal{A}(K')$ and that they are antipodal. Let us suppose then that $\mathcal{A}(K')$ contains some $I \in \mathcal{A}_+$ and $J \in \mathcal{A}_-$ which are not complements in K'^c. The chunk of $\mathcal{A}$ they span is not, therefore, the whole of $\mathcal{A}$ and hence is extremal (by the inductive hypothesis). In particular I and J are connected in $\mathcal{A}(K')$. At the i-edge of that path sits a $K' \cup \{i\} = K \cup \{j\}$ cube. This contradicts the inductive hypothesis concerning K, since it implies the existence of two non-antipodal K-cubes, and thereby establishes it for K'.

The proof of the theorem is now rather straightforward. Since we know where all the k-cubes in $\mathcal{A}$ belong we know in particular that $\mathcal{A} \subset [n]^{(\leq k)}$. Moreover every cube in $\mathcal{A}$ of size smaller than k is contained in some k-cube. This is an immediate consequence of the claim. Given a j-set C, with $j < k$, and a C-cube $J + \mathcal{P}(C) \subset \mathcal{A}$ we know from the claim that there is, inside $\mathcal{A}$ somewhere, a non-antipodal C-cube. By our earlier remarks we know that these cubes can be joined by a path in $\mathcal{A}(C)$ and therefore that $J + \mathcal{P}(C)$ is contained in some larger cube. Thus, by induction on $k - j$ we may deduce that $J + \mathcal{P}(C) \subset J' + \mathcal{P}(B)$ for some k-set B. Combining these facts we have shown that $\mathcal{A}$ is the union of its (antipodally placed) k-cubes and is therefore self-complementary. This contradicts our initial assumption concerning $\mathcal{A}$ and so the thoerem is proved. ■

§2. Defect Sauer for small k.

A wide variety of defect Sauer results have been proved. Bondy [5] showed that $(N,n) \to (N,n-1)$ provided that $N \leq n$. Bollobás (see [11] p.444) showed that $(N,n) \to (N-1,n-1)$ provided $N \leq \lceil 3n/2 \rceil$. Frankl generalised this and proved [9] that $(N,n) \to (N-2^t+1,n-1)$ on condition $N \leq \lceil n(2^{t+1}-1)/(t+1) \rceil$. This result was extended further in [13]. There it was shown that if one writes $\mathcal{B}(k)$ for the initial segment of $\mathcal{P}(n)$ of length k with respect to the colex (or binary) ordering ($A < B$ iff $\max(A\Delta B) \in B$), then

$$(N,n) \to (2^{n-m} - k + 1, n-m)$$

for

$$N > \sum_{i=m}^{n} \binom{n}{i} - \binom{n}{m} \sum_{B \in \mathcal{B}(k)} \binom{|B|+m}{m}^{-1}.$$

In a different direction Frankl [9] proved that if $N > \lfloor n^2/4 \rfloor + n + 1$ then $(N,n) \to (7,3)$. This result uses Turan's theorem and the technique of down compression. Similar arguments allow one to show that $(t_3(n)+n+2,n) \to (11,4)$, where $t_3(n)$ is the number of edges in the Turan graph $T_3(n)$. In this section we prove the 'next' natural result. The appearance here of an exceptional case is the first indication, confirmed in later sections, that the best examples are not neccessarily systems of the form $[n]^{(k)}$, even when such systems are natural candidates.

Theorem 5. *Let n an integer with $n \geq 4$ and $n \neq 6$. Then for $N > \binom{n}{2} + n + 1$ one has $(N,n) \to (12,4)$. In addition $(23,6) \not\to (12,4)$ and there is a unique (up to isomorphism) extremal system.*

Proof. If $n = 4,5,6,9$ the result is straighforward to check by hand. If $n = 6$ then the unique extremal example is generated by $\{123, 246, 345, 156\}$. Now suppose that n is minimal such that a counterexample to the theorem exists and let $\mathcal{A} \subset \mathcal{P}(n)$ by a system with $|\mathcal{A}| = N > \binom{n}{2}+n+1$ and $\mathrm{tr}\,(\mathcal{A},4) \leq 11$. Then $n \geq 7$ and $n \neq 9$. By Lemma 1 we may assume that $\mathcal{A}$ is a down set, and then it is clear that $\mathcal{A} \subset [n]^{(\leq 3)}$. Also every element of $[n]$ is covered by some pair in $\mathcal{A}$ for otherwise simple discard the point i which is not covered. By assumption $\left|\mathcal{A}|_{[n]-\{i\}}\right| \geq |\mathcal{A}|-1$ and, since n is minimal, $\left|\mathcal{A}|_{[n]-\{i\}}\right| \leq \binom{n-1}{2}+(n-1)+2$ (allowing for the case $n-1=6$). Therefore $|\mathcal{A}| \leq \binom{n}{2}+n+1$, a contradiction. One more condition on $\mathcal{A}$ is clear; that no two 3-sets in $\mathcal{A}$ intersect in a pair, for their union would be a 4-set on which $\mathcal{A}$ has trace 12.

Now suppose that A is a 3-set in $\mathcal{A}$. Set $\mathcal{B}_A = \{B \in \mathcal{A} : B \cap A \neq \emptyset\}$. Given $x \in [n] \setminus A$ it is joined by edges (2-sets) of $\mathcal{B}_A$ to at most 2 elements of A, since if it were joined to all of them then $\left|\mathcal{A}|_{A \cup \{x\}}\right| \geq 12$. Also since no pair of 3-sets shares an edge, the number of triples (3-sets) which intersect both A and $[n] \setminus A$ is at most half the number of edges which cross. In total

$$|\mathcal{B}_A| \leq 3(n-3) + 7 = 3n - 2.$$

If, in fact, $\mathcal{B}_A < 3n - 2$ then

$$\begin{aligned} |\mathcal{A}| &= \left|\mathcal{A}|_{[n]\setminus A}\right| + |\mathcal{B}_A| \\ &\leq \binom{n-3}{3} + n - 2 + 3n - 3 \\ &= \binom{n}{2} + n + 1. \end{aligned}$$

The bound on $\left|\mathcal{A}|_{[n]\setminus A}\right|$ is a consequence of the minimality of n; if $\left|\mathcal{A}|_{[n]\setminus A}\right|$ were stictly more than $\binom{n-3}{3} + n - 2$ then $\mathcal{A}|_{[n]\setminus A}$ would be a counter-example in $\mathcal{P}(n-3)$.

Thus it must be that for all 3-sets $A \in \mathcal{A}$ the size of $\mathcal{B}_A$ is $3n-2$. This is an extremely strong condition. In particular, given a triple $A \in \mathcal{A}$ and a point $x \notin \mathcal{A}$ there are exactly two edges from x to A, each in some triple. As a consequence, every 2-set in $\mathcal{A}$ is contained in some triple in $\mathcal{A}$.

To be specific suppose that $A = \{1, 2, 3\} \in \mathcal{A}$. Call an edge $\{i, j\} \subset [n] \setminus A$ a k-*base*, $k \in \{1, 2, 3\}$, if $\{i, j, k\} \in \mathcal{A}$. There are fairly severe restrictions on the collection of 1, 2 and 3 bases. Two 1-bases cannot intersect, else their associated triples would share an edge. Of course the same applies to 2- and 3-bases. Also, no 1-base $\{i, j\}$ can be incident with 2-bases at both ends else the trace of $\mathcal{A}$ on $\{i, j, 1, 2\}$ would be at least 12. In summary, the graph on $[n] \setminus A$ with edges all 1-, 2-, and 3-bases is a union of cycles in each of which the edges are cyclically labelled $1, 2, 3, 1, 2, 3, \ldots$. Certainly therefore each has length divisible by 3, and so n is divisible by 3.

We are now in a position to reconstruct the graph $G = ([n], \mathcal{A}^{(2)})$. If two vertices x and y are not joined in this graph then their neighbour sets are identical. This follows because any edge containing x is in some triple, and this triple does not contain y. Therefore y must be joined to the other two vertices of the triple. If on the other hand x and y are joined then they are elements of some triple $A \in \mathcal{A}$. By the remarks in the previous paragraph they each have degree $2n/3$ and they have $n/3$ neighbours in common. In particular G is $2n/3$ regular. If one takes any pair x, y which are neighbours in G and sets

$$I = [n] \setminus \Gamma(x) \qquad J = [n] \setminus \Gamma(y) \qquad K = [n] \setminus (I \cup J),$$

it is easy to see that G is the complete tripartite graph with tripartition (I, J, K). Every vertex in I, since it is not joined to x, is joined to exactly the vertices of $J \cup K = \Gamma(x)$. Similarly, if $z \in J$ then $\Gamma(z) = I \cup K = \Gamma(y)$. By the regularity of G, if $z \in K$ then $\Gamma(z) = I \cup J$. Therefore $\left|\mathcal{A}^{(2)}\right| = n^2/3$. Since $\left|\mathcal{A}^{(3)}\right| \leq \left|\mathcal{A}^{(2)}\right|/3$ we have in total

$$|\mathcal{A}| \leq 4n^2/9 + n + 1 \leq \binom{n}{2} + n + 1.$$

The last inequality depends on n being at least 9. This contradiction finishes the proof, though it is interesting to notice that the 2-sets from the exceptional example for $n = 6$ do indeed form the complete tripartite graph with equal parts. ■

§3. Lower bounds for the trace on αn-sets

We turn now to the problem of determining the best possible arrow relations $(N, n) \to (M, k)$ in the domain of the problem for which N is a polynomial in n and k is a positive proportion of n. In this section we give lower bounds; first a bound applicable in all ranges of the problem, then an argument using random methods tailored to the situation. Between the best lower bound in this section and the upper bound proved in the next there is quite a large gap. It seems likely that the correct answer is nearer to the upper bound.

Theorem 6. *For all $n \geq 1$, $k \in [n]$ and $1 \leq N \leq 2^n$,*

$$(N, n) \to (N^{k/n}, k).$$

Proof. Let $\mathcal{A} \subset \mathcal{P}(n)$ be a set system of size N and consider the following subset of $\mathbf{R}^n$. For any $A \subset [n]$ let $C_A = \{(x_i)_1^n \in [-1,1]^n : x_i \geq 0 \iff i \in A\}$ and let $C_{\mathcal{A}} = \bigcup_{A \in \mathcal{A}} C_A$. Then $(C_{\mathcal{A}}) = N$ and if P_I is the orthogonal projection onto $\operatorname{span}\{e_i : i \in I\}$ then $\operatorname{vol}_{|I|}(P_I(C_{\mathcal{A}})) = |\mathcal{A}|_I|$. The theorem then follows immediately from the standard estimate (see for instance [6]), valid for any measurable $C \subset \mathbf{R}^n$,

$$\operatorname{vol}(C)^{1/n} \leq \left(\prod_{I \in [n]^{(k)}} \operatorname{vol}_k(P_I(C))^{1/k} \right)^{1/\binom{n}{k}}.$$ ■

Though Theorem 6 is useful in a wide range of circumstances, it does not necessarily give a best possible lower bound. In particular if $N = n^r$ it shows that $(N, n) \to (n^{\alpha r}, \alpha n)$. The next result improves this bound by the use of random techniques. We will write $h(\lambda) = \lambda^\lambda (1-\lambda)^{1-\lambda}$ for $\lambda \in (0,1)$.

Theorem 7. *Fix $r \geq 2$ and $\alpha \in (0,1)$. Then*

$$(n^r, n) \to \left((1-o(1))n^{\lambda r}, \alpha n\right),$$

Where, writing $\lambda_0 = \log_2(1+\alpha)$,

$$\lambda = \begin{cases} \lambda_0 & \alpha \in [\sqrt{2}-1, 1) \\ -\lambda_0 \;/\; \log_2 h(\lambda_0) & \alpha \in (0, \sqrt{2}-1). \end{cases}$$

Proof. Set $k = r \log_2 n$, and let $\mathcal{A} \subset \mathcal{P}(n)$ have size n^r. In light of Lemma 1 we may assume that $\mathcal{A}$ is a down set. We may also assume that all the sets in $\mathcal{A}$ have size at most λk, for any set of larger size, together with all its subsets, would give a trace of more than $2^{\lambda k} = n^{\lambda r}$ to any αn set containing it.

Now choose a αn-set I randomly from the uniform distribution on $[n]^{(\alpha n)}$. The probability that a fixed i-set $A \subset [n]$ is contained in I is

$$\mathbf{P}(I \supset A) = \frac{\binom{n-i}{\alpha n - i}}{\binom{n}{pn}} = \frac{(\alpha n)_i}{(n)_i}.$$

where $(n)_i$ is the falling factorial $n(n-1)\dots(n-i+1)$. Since $\mathcal{A}$ is a down set, $\mathcal{A}|_I = \mathcal{A} \cap \mathcal{P}(I)$. Therefore, setting $t = |\mathcal{A}|_I|$ and $N_i = \left|\mathcal{A}^{(i)}\right|$ we have

$$\mathbf{E}t = \sum_{i=0}^{\lambda k} N_i \frac{(\alpha n)_i}{(n)_i}.$$

Katona [10] has shown that if $f : \mathbf{N} \to \mathbf{R}$ is decreasing and $\mathcal{A} \subset \mathbf{N}^{(\leq r)})$ is a finite down set then

$$\sum_{A \in \mathcal{A}} f(|A|) \geq \sum_{B \in \mathcal{B}(k)} f(|B|),$$

where $\mathcal{B}$ is the first $|\mathcal{A}|$ elements in the binary ordering on $\mathbf{N}^{(\leq r)}$. In similar spirit one can prove (see [13]) that

$$\sum_{A \in \mathcal{A}} f(|A|) \geq \sum_{i=0}^{r} \binom{x}{i} f(i)$$

provided that $|A| = \sum_{i=0}^{r} \binom{x}{i}$. Therofore $\mathbf{E}t$ is at least $\sum_{i=0}^{\lambda k} \binom{x}{i} (\alpha n)_i/(n)_i$. It remains to determine x. There are two cases, corresponding to the two ranges of α in the definition of λ.

Case 1. If $\alpha < \sqrt{2}-1$ then we apply a standard estimate on the weight in the tail of the binomial distribution (see [2] Chapter 1, Section 2). This states that for $\beta < 1/2$,

$$\sum_{i=0}^{\lfloor \beta n \rfloor} \binom{n}{i} \le \frac{2(1-\beta)}{(1-2\beta)} \binom{n}{\beta n} \le \frac{2(1-\beta)h(\beta)^{-n}}{(1-2\beta)\sqrt{2\pi\beta(1-\beta)n}}.$$

Applying this estimate we can deduce that $x \ge \lambda k/\lambda_0$ since

$$\begin{aligned}
\sum_{i=0}^{\lambda k} \binom{\lambda k/\lambda_0}{i} &\le \frac{2(1-\lambda_0)h(\lambda_0)^{-\lambda k/\lambda_0}}{(1-2\lambda_0)\sqrt{2\pi\lambda_0(1-\lambda_0)n}} \\
&= \frac{2(1-\lambda_0)h(\lambda_0)^{k/\log_2 h(\lambda_0)}}{(1-2\lambda_0)\sqrt{2\pi\lambda_0(1-\lambda_0)n}} \\
&= \frac{2(1-\lambda_0)2^k}{(1-2\lambda_0)\sqrt{2\pi\lambda_0(1-\lambda_0)n}} \\
&= \frac{2(1-\lambda_0)n^r}{(1-2\lambda_0)\sqrt{2\pi\lambda_0(1-\lambda_0)n}} \\
&= o(n^r).
\end{aligned}$$

Thus, in case 1, $x \ge \lambda k/\lambda_0$.

Case 2. If $\alpha \ge \sqrt{2}-1$ then we certainly have $x \ge k$ since

$$\sum_{i=0}^{\lambda k} \binom{k}{i} \le 2^k = n^r.$$

So $x \ge k = \lambda k/\lambda_0$.

Now let us estimate $\sum_{i=0}^{\lambda k} \binom{x}{i}(\alpha n)_i/(n)_i$. Rather crudely, for $1 \le i \le \lambda k$ and N sufficiently large,

$$\begin{aligned}
\frac{(\alpha n)_i}{(n)_i} &\ge \alpha^i \left(1 - \frac{(i-1)-\alpha}{n-(i-1)}\right)^{i-1} \\
&\ge \alpha^i \left(1 - \frac{2\lambda^2 k^2}{n}\right).
\end{aligned}$$

As before we can bound the mean of t as follows:

$$\begin{aligned}
\mathbf{E}t &\ge \sum_{i=0}^{\lambda k} \binom{x}{i} \frac{(\alpha n)_i}{(n)_i} \\
&\ge \left(1 - \frac{2(\lambda k)^2}{n}\right) \sum_{i=0}^{\lambda k} \binom{x}{i} \alpha^i.
\end{aligned}$$

Finally note that even when $\alpha < \sqrt{2} - 1$ we have

$$\sum_{i=0}^{\lambda k} \binom{x}{i} \alpha^i = (1 - o(1))(1 + \alpha)^x.$$

This follows from the same estimate for the tail of the binomial distribution, this time applied to the distribution with parameters x and $p/(1 + \alpha)$. When $\alpha < \sqrt{2} - 1$ note that

$$\frac{\lambda k}{x} = \lambda_0 = \log_2(1 + \alpha) > \frac{\alpha}{1 + \alpha}.$$

Thus λk is 'past the hump' in the binomial distribution with these parameters. In conclusion,

$$\begin{aligned} \mathbf{E}t &\geq \left(1 - \frac{2(\lambda k)^2}{n}\right) \sum_{i=0}^{\lambda k} \binom{x}{i} \alpha^i \\ &= (1 - o(1))\ (1 + \alpha)^x \\ &\geq (1 - o(1))\ 2^{\lambda k} \\ &= (1 - o(1))\ n^{\lambda r}. \end{aligned}$$

Since on average t is at least $(1 - o(1))\ n^{\lambda r}$, there must be some αn-set I with $|\mathcal{A}|_I| \geq (1 - o(1))\ n^{\lambda r}$, and the theorem is proved. ■

§4. Upper bounds for traces on αn-sets.

In this section we present an upper bound for $\operatorname{tr}\left(\sum_{i=0}^{r} \binom{n}{i}, n/2\right)$. In other words we construct an example of a set system $\mathcal{A}$ with this size and with small trace on any $n/2$-set.

The most obvious example to consider is $\mathcal{A} = [n]^{(\leq r)}$. This example shows that $\operatorname{tr}\left(\sum_{i=0}^{r} \binom{n}{i}, n/2\right) \leq \sum_{i=0}^{r} \binom{n/2}{i}$. The thrust of this section is to show that this example is not the best possible, and indeed that for all (large) n there is a down set $\mathcal{A}_n$, of size $\sum_{i=0}^{r} \binom{n}{i}$, with $\operatorname{tr}(\mathcal{A}_n, n/2) = o(n^r)$.

In outline the proof proceeds as follows. Choose, at random, some sets of size $k = \log_2(\log_2 n)$. Estimate the size of the down system $\mathcal{A}$ generated by these sets, and also estimate the trace of that system on $[n/2]$. To be precise the estimates are on the means of these quantities. Of course just knowing an estimate for the mean trace on the particular $n/2$-set $[n/2]$ is not enough. What is crucial is that this random variable – $t = \left|\mathcal{A}|_{[n/2]}\right|$ – is highly concentrated around its mean. The probability that it takes a value far from this mean is exponentially small. The crux of the proof is showing that with probability at least 1/2 the system $\mathcal{A}$ has small trace on *all* $n/2$-sets.

The following result of Bollobás and Leader [3] plays a central rôle. It is an isoperimetric inequality for the 'weighted cube' $\mathcal{P}_p(n)$. This is the finite probability space $\mathcal{P}(n)$ with $\mathbf{P}(\{A\}) = p^{|A|}(1-p)^{n-|A|}$, for $p \in (0,1)$. It states that, among down sets of fixed probability, collections of the form $[n]^{(\leq r)}$ have the smallest vertex boundary. (Recall that the vertex boundary of a subset S of the vertices of a graph G is the set $S \cup \Gamma(S)$ consisting of S togther with all neighbours of vertices in S. Since the boundary of $[n]^{(\leq r)}$ is again of the same form the result also applies to the *t-boundary* of a system, defined to be $\mathcal{A}_{(t)} = \{B \in \mathcal{P}(n) \,:\, |A \Delta B| \leq t \text{ for some } A \in \mathcal{A}\}$.

Theorem 8. *Suppose $\mathcal{A} \subset \mathcal{P}_p(n)$ is a down set with $\mathbf{P}(\mathcal{A}) \geq 1/2$. If $r \in [n]$ satisfies $\mathbf{P}([n]^{(\leq r)}) \leq 1/2$ then for all $t \geq 0$,*

$$\mathbf{P}(\mathcal{A}_{(t)}) \geq \mathbf{P}([n]^{(\leq r+t)}). \qquad \blacksquare$$

To apply this Theorem it is clearly necessary to make use of bounds on the binomial distribution with parameters n and $p = p(n)$. For the remainder of this section we write $S_{n,p}$ for a random variable with this distribution, and also write q for $1-p$.

Recall that a *median* for a function $f : \mathcal{P}_p(n) \to \mathbf{R}$ is any $x \in \mathbf{R}$ for which both $\mathbf{P}(f(A) \leq x) \geq 1/2$ and $\mathbf{P}(f(A) \geq x) \geq 1/2$. Define such a function to be *increasing* if whenever $A \subset B \subset [n]$ we have $f(A) \leq f(B)$. The next two results are relatively standard consequences of the measure concentration phenomenon, and similar results appear in, e.g., [8], where they are applied to geometric functional analysis. The first states that an increasing Lipschitz function $f : \mathcal{P}_p(n) \to \mathbf{R}$ is very likely to take values close to its median, and the second that, as a consequence, its mean is very close to its median.

Lemma 9. *Given $\epsilon > 0$ there is an $n_0 = n_0(\epsilon)$ such that for all $n \geq n_0$ and increasing Lipschitz functions $f : \mathcal{P}_p(n) \to \mathbf{R}$ we have*

$$\mathbf{P}(|f(A) - M_f| \geq bt) \leq \mathbf{P}(|S_{n,p} - pn| \geq t - \epsilon\sqrt{pqn}). \qquad \blacksquare$$

where f has Lipschitz constant b, M_f is a median for f and t is an arbitrary positive integer.

Lemma 10. *If $p = p(n)$ is such that $pqn \to \infty$ as $n \to \infty$ then for sufficiently large n any increasing function $f : \mathcal{P}_p(n) \to \mathbf{R}$ with median M_f and Lipschitz constant b satisifes*

$$|\mathbf{E}f - M_f| < 2b\sqrt{pqn}. \qquad \blacksquare$$

We are now in a position to present the random construction itself. The major point of this result is that the 'flat' example, $\mathcal{A} = [n]^{(\leq r)}$, is quite far from being best possible.

Theorem 11. *Fix an integer $r \geq 2$ and $\epsilon \in (0, 1/2)$. There is some $n_0 = n_0(r, \epsilon)$ such that for all $n \geq n_0$ there exists a down system $\mathcal{A} \in \mathcal{P}(n)$ with*

$$|\mathcal{A}| \geq \sum_{i=0}^{r} \binom{n}{i}$$

and

$$\mathrm{tr}\,(\mathcal{A}, n/2) \leq \sum_{i=0}^{r} \binom{n/2}{i} - (1-\epsilon)\, 2^{-r} \binom{n}{r}.$$

Proof. *Set $k = \log_2(\log_2 n)$ and let $\lambda \in (0,1)$ be the unique solution of*

$$\frac{e^{-\lambda}}{\lambda} = \left(1 - \frac{\epsilon}{2}\right) \binom{k}{r}^{-1} \sum_{i=r+1}^{k} \binom{k}{i}$$

and set $p = \lambda \,/\, \binom{n-r}{k-r}$. For sufficiently large n the following bounds for λ hold,

$$\frac{\binom{k}{r}}{\log_2 n} < \lambda < \frac{2k^r}{\log_2 n}.$$

We take $\mathcal{B}$ to be a random collection of k-sets of $[n]$, chosen independently with probability p. In other words $\mathcal{B}$ is a random point of $\mathcal{P}_p(N)$ where $N = \binom{n}{k}$. Write $\mathcal{A}$ for the down set generated by $\mathcal{B}$. The two parameters of $\mathcal{A}$ which are of interest are

$$S = |\mathcal{A}|$$

and

$$t = \big|\mathcal{A}|_{[n/2]}\big|.$$

These are to be regarded as random variables defined on the probability space $\mathcal{P}_p(N)$. Their respective means can be estimated relatively easily. Note that the probability of an i-set belonging to $\mathcal{A}$ is exactly $1 - (1-p)^{\binom{n-i}{k-i}}$. Therefore

$$\begin{aligned}
\mathbf{E}S &= \sum_{i=0}^{k} \binom{n}{i} \left[1 - (1-p)^{\binom{n-i}{k-i}}\right] \\
&= \sum_{i=0}^{r} \binom{n}{i} \left[1 - \left(1 - \frac{\lambda}{\binom{n-r}{k-r}}\right)^{\binom{n-i}{k-i}}\right] + \sum_{i=r+1}^{k} \binom{n}{i} \left[1 - (1-p)^{\binom{n-i}{k-i}}\right].
\end{aligned}$$

Write E_1 for the first part of the sum and E_2 for the second. Estimating separately:

$$E_1 \geq \sum_{i=0}^{r} \binom{n}{i} \left[1 - e^{-\lambda \binom{n-i}{k-i} / \binom{n-r}{k-r}}\right] = \sum_{i=0}^{r} \binom{n}{i} - \binom{n}{r} e^{-\lambda} + o(1).$$

In the second part of the sum, E_2, the above simple estimate breaks down. So

$$
\begin{aligned}
E_2 &\geq \sum i = r + 1^k \binom{n}{i} \left[p\binom{n-i}{k-i} - \frac{p^2}{2}\binom{n-i}{k-i}^2 \right] \\
&= \sum_{i=r+1}^{k} p\binom{n}{k} \left[\binom{k}{i} - \frac{p}{2}\binom{n-i}{k-i} \right] \\
&= \sum_{i=r+1}^{k} p\binom{n}{k}\binom{k}{i} \left[1 - \frac{\lambda\, i!}{2(k)_r (n-r)_{i-r}} \right] \\
&= (1-o(1))\, p\binom{n}{k} \sum_{i=r+1}^{k} \binom{k}{i}.
\end{aligned}
$$

Together we have

$$
\mathbf{E}S \geq \sum_{i=0}^{r} \binom{n}{i} - \binom{n}{r} e^{-\lambda} + (1-o(1))p\binom{n}{k} \sum_{i=r+1}^{k} \binom{k}{i}. \tag{1}
$$

What this shows, not too surprisingly, is that the system $\mathcal{A}$ will contain, with high probability,

i) almost all of $[n]^{(\leq r-1)}$ and

ii) a fraction $1 - e^{-\lambda}$ of level $[n]^{(r)}$.

In the layers higher than r the down sets generated by the elements of $\mathcal{B}$ will be essentially disjoint.

By the choice of λ the expression on the right hand side of (1) simplifies considerably. Since

$$
\begin{aligned}
\binom{n}{r} e^{-\lambda} &= (1 - \frac{\epsilon}{2})\lambda \frac{\binom{n}{r}}{\binom{k}{r}} \sum_{i=r+1}^{k} \binom{k}{i} \\
&= (1 - \frac{\epsilon}{2}) p\binom{n}{k} \sum_{i=r+1}^{k} \binom{k}{i},
\end{aligned}
$$

the bound on $\mathbf{E}S$ becomes

$$
\mathbf{E}S \geq \sum_{i=0}^{r} \binom{n}{i} + (\frac{\epsilon}{2} - o(1))p\binom{n}{k} \sum_{i=r+1}^{k} \binom{k}{i}.
$$

As we shall see a little later, Lemma 10, together with a straightforward estimate on the second term in the above expression, shows that $M_S \geq \sum_{i=0}^{r} \binom{n}{i}$. Therefore with probability at least 1/2 we have $S \geq \sum_{i=0}^{r} \binom{n}{i}$.

Now consider the mean of t. Since $\mathcal{A}$ is a down system, $\mathcal{A}|_{[n/2]} = \mathcal{A} \cap \mathcal{P}(n/2)$. Therefore the expected size of the restriction is of the same form as $\mathbf{E}S$. This time however the aim is to get an upper bound. Thus

$$\begin{aligned}\mathbf{E}t &= \sum_{i=0}^{k} \binom{n/2}{i} \left(1 - (1-p)^{\binom{n-i}{k-i}}\right) \\ &\leq \sum_{i=0}^{r} \binom{n/2}{i} - (1-o(1)) \binom{n/2}{r} e^{-\lambda} + \sum_{i=r+1}^{k} \binom{n/2}{i} \binom{n-i}{k-i} p.\end{aligned}$$

Again the sum is split into two parts and different estimates are applied to the different parts. Noting that $\binom{n/2}{r} 2^r / \binom{n}{r} \to 1$ as $n \to \infty$, and that the ratio is always at most 1, we have

$$\begin{aligned}\mathbf{E}t &\geq \sum_{i=0}^{r} \binom{n/2}{i} - (1-o(1)) \binom{n}{r} e^{-\lambda} 2^{-r} + \sum_{i=r+1}^{k} p \binom{n}{i} \binom{n-i}{k-i} 2^{-i} \\ &= \sum_{i=0}^{r} \binom{n/2}{i} - (1-o(1)) \binom{n}{r} e^{-\lambda} 2^{-r} + p \binom{n}{k} \sum_{i=r+1}^{k} \binom{k}{i} 2^{-i}.\end{aligned}$$

We now want to simplify this to obtain a workable expression for $\mathbf{E}t$. The second term, by the choice of λ, is

$$(1-o(1)) \binom{n}{r} e^{-\lambda} 2^{-r} = \left(1 - \frac{\epsilon}{2} - o(1)\right) 2^{-r} p \binom{n}{k} \sum_{i=r+1}^{k} \binom{k}{i}.$$

But this term dominates the next, for

$$\sum_{i=r+1}^{k} \binom{k}{i} 2^{-i} \leq (3/2)^k \qquad \text{and} \qquad \sum_{i=r+1}^{k} \binom{k}{i} = (1-o(1)) 2^k.$$

So, in sum,

$$\mathbf{E}t \leq \sum_{i=0}^{r} \binom{n/2}{i} - \left(1 - \frac{\epsilon}{2} - o(1)\right) p \binom{n}{k} 2^{k-r}. \tag{2}$$

Having estimated the means of S and t, we use Lemmas 9 and 10 to show that their values are highly concentrated about these means. Both S and t are increasing functions with Lipschitz constant $2^k = \log_2 n$, so, from (1),

$$\begin{aligned}M_S &\geq \sum_{i=0}^{r} \binom{n}{i} + \left(\frac{\epsilon}{2} - o(1)\right) p \binom{n}{k} 2^k - 2^{k+1} \sqrt{pq \binom{n}{k}} \\ &\geq \sum_{i=0}^{r} \binom{n}{i} + \left(\frac{\epsilon}{2} - o(1)\right) p \binom{n}{k} 2^k.\end{aligned}$$

Similarly, from (2),

$$M_t \leq \sum_{i=0}^{r} \binom{n/2}{i} - (1 - \frac{\epsilon}{2} - o(1))p\binom{n}{k}2^{k-r}.$$

At this point we can show that the probability of even quite a modest deviation from the median of t is $o(\binom{n}{n/2}^{-1})$. Indeed, by Lemma 10 and standard estimates on the expression therein we have, for sufficiently large n,

$$\begin{aligned}\mathbf{P}(t \geq M_t + (pqN)^{5/8}2^k) &\leq (pqN)^{-5/8}e^{-(pqN)^{1/4}/2} \\ &= o\left(\binom{n}{n/2}^{-1}\right).\end{aligned}$$

Thus, if n is sufficiently large,

$$\mathbf{P}(t \geq \sum_{i=0}^{r} \binom{n/2}{i} - (1-\epsilon)p\binom{n}{k}2^{k-r}) \leq \frac{1}{2}\binom{n}{n/2}^{-1}.$$

Clearly the distribution of $|\mathcal{A}|_I|$, where I is any $n/2$-set, is the same as that of t. So, with probability strictly greater than 1/2,

$$\operatorname{tr}(\mathcal{A}, n/2) \leq \sum_{i=0}^{r} \binom{n/2}{i} - (1-\epsilon)p\binom{n}{k}2^{k-r}.$$

Also, for sufficiently large n, the estimate above for M_S show that $\mathbf{P}(S \geq \sum_{i=0}^{r} \binom{n}{i}) > 1/2$. Therefore there is some down system $\mathcal{A}$ satisfying

$$\begin{aligned}|\mathcal{A}| &\geq \sum_{i=0}^{r} \binom{n}{i} \\ \operatorname{tr}(\mathcal{A}, n/2) &\leq \sum_{i=0}^{r} \binom{n/2}{i} - (1-\epsilon)p\binom{n}{k}2^{k-r}.\end{aligned}$$

To finish we unwind the definition of p and λ:

$$\begin{aligned}\operatorname{tr}(\mathcal{A}, n/2) &\leq \sum_{i=0}^{r} \binom{n/2}{i} - (1-\epsilon)\lambda\binom{n}{k}2^{k-r} \Big/ \binom{n-r}{k-r} \\ &\leq \sum_{I=0}^{r} \binom{n/2}{i} - (1-\epsilon)\binom{k}{r}\binom{n}{k}2^{k-r} \Big/ \binom{n-r}{k-r} \log_2 n \\ &= \sum_{i=0}^{r} \binom{n/2}{i} - (1-\epsilon)\binom{n}{r}2^{-r}.\end{aligned}$$

This completes the proof of the Theorem. ■

References

[1] N. Alon, On the desity of set of vectors, *Discrete Math.* **46** (1983), 199–202.

[2] Béla Bollobás, *Random Graphs*, Academic Press, London, (1985).

[3] B. Bollobás and I. Leader, Isoperimetric Inequalities and Fractional Set Systems, *J. Comb. Theory Ser. A* **56** (1991), 63–74.

[4] B. Bollobás, I. Leader and A.J. Radcliffe, Reverse Kleitman Inequalities, *Proc. London Math. Soc. (3)* **58** (1989), 153–168.

[5] J.A. Bondy, Induced subsets, *J. Comb. Theory Ser. B* **12** (1972), 201–202.

[6] Iu.D. Burago and V.A. Zalgaller, *Geometric Inequalities*, Springer-Verlag, Berlin, New York, (1986).

[7] R.M. Dudley, *The stucture of some Vapnik-Červonenkis Classes*, in *Proceedings of the Berkeley conference in honor of Jerzy Neyman and Jack Kiefer, Vol. II (Berkeley, California, 1983)*, 495–508, Wadsworth, Belmont, Calif. (1985).

[8] T. Figiel, J. Lindenstrauss and V.D. Milman, The dimension of almost spherical sections of convex bodies, *Acta Math.* **139** (1977), 53–94.

[9] P. Frankl, On the trace of finite sets, *J. Comb. Theory Ser. A* **34** (1983), 41–45.

[10] G.O.H. Katona, *Optimisation for order ideals under a weight assignment*, in *Problèmes Comb. et Théorie des Graphes*, 257–258, Colloque Int., CNRS Paris (1978).

[11] L. Lovász, *Combinatorial Problems and Exercises*, North-Holland, Amsterdam, New York, Oxford, (1979).

[12] A. Pajor, *Sous-espaces ℓ_1^n des espaces de Banach*, Travaux en Cours, Hermann, Paris, (1985).

[13] A.J. Radcliffe, *Sauer theorems for set systems*, Ph.D Dissertation, Cambridge University, (1989).

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

[15] S. Shelah, A combinatorial problem; Stability and order for models and theories in infinitary languages, *Pacific J. Math.* **41** (1972), 271–276.

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