Available online at www.sciencedirect.com
ScienceDirect
Discrete Applied Mathematics 156 (2008) 1536–1541
DISCRETE APPLIED MATHEMATICS
www.elsevier.com/locate/dam

Large convex cones in hypercubes

Zoltán Füredi[a, b, 1], Miklós Ruszinkó[c, 2]

[a] *Department of Mathematics, University of Illinois, Urbana, IL 61801, USA*
[b] *Rényi Institute of Mathematics of the Hungarian Academy of Sciences, Budapest, P.O. Box 127, 1364, Hungary*
[c] *Computer and Automation Research Institute of the Hungarian Academy of Sciences, Budapest, P.O. Box 63, 1518, Hungary*

Received 9 August 2002; received in revised form 3 March 2004; accepted 2 November 2006
Available online 27 August 2007

Dedicated to the memory of our good friend and colleague Levon Khachatrian

Abstract

A family of subsets of $[n]$ is *positive linear combination free* if the characteristic vector of neither member is the positive linear combination of the characteristic vectors of some other ones. We construct a positive linear combination free family which contains $(1-\text{o}(1))2^n$ subsets of $[n]$ and we give tight bounds on the $\text{o}(1)2^n$ term. The problem was posed by Ahlswede and Khachatrian [Cone dependence—a basic combinatorial concept, Preprint 00-117, Diskrete Strukturen in der Mathematik SFB 343, Universität Bielefeld, 2000] and the result has geometric consequences.

MSC: 05D05; 52B05

Keywords: Hypercube; Convex cone

1. Positive linear combination free families

We address a question which was formulated in [1]. How many edges may a hypergraph on n vertices contain such that the characteristic vector of neither edge is the positive linear combination of the characteristic vectors of some other ones? In [1] a construction with $(\frac{1}{2}+c)2^n$ sets was given and it was asked if such a family could contain almost all edges or significantly less. Here we give an explicit construction of such a family which contains $(1-\text{o}(1))2^n$ edges and tight bounds for the $\text{o}(1)2^n$ term.

Let $[n]=\{1,\ldots,n\}$. The characteristic vector of $A \subseteq [n]$ is the vector A in $\{0,1\}^n$ which has 1 in the ith coordinate iff $i \in A$. (We use the same notation for sets and characteristic vectors.) A is the positive linear combination of $A_1,\ldots,A_k$ iff $A = c_1A_1+\cdots+c_kA_k$ and $\forall i : c_i > 0$. $\mathscr{F} \subseteq 2^{[n]}$ is positive linear combination free iff no set (vector) is the positive linear combination of some other sets from $\mathscr{F}$, i.e., for arbitrary choice of positive

[1] Research supported in part by the Hungarian National Science Foundation under Grant OTKA T 032452, and by the National Science Foundation Grant DMS 0140692.
[2] Research supported in part by OTKA Grants T029074, T030059, T038198.
E-mail address: ruszinko@sztaki.hu (M. Ruszinkó).

0166-218X/$ - see front matter © 2007 Elsevier B.V. All rights reserved.
doi:10.1016/j.dam.2006.11.018

coefficients and $\mathscr{F}' \subseteq \mathscr{F}$

$$A \neq \sum_{\substack{A \neq A_i \in \mathscr{F}' \subseteq \mathscr{F} \\ c_i > 0}} c_i A_i .$$

Let $f(n)$ be the maximum size of a positive linear combination free family. In the next section we construct a positive linear combination free family of size

$$2^n \left(1 - \mathrm{O}\left(\frac{\log\log n}{\log n}\right)\right)$$

which already shows that $f(n) = (1 - \mathrm{o}(1))2^n$. Then we give tight bounds on the $\mathrm{o}(1)2^n$ term. (Here O and o—and later Ω—are used in conventional sense, i.e., for sequences $f(m)$ and $g(m)$ $f(m) = \mathrm{O}(g(m))$ if $f(m) \leqslant cg(m)$ holds for some constant $c > 0$ and every sufficiently large m, $f(m) = \Omega(g(m))$ if $g(m) = \mathrm{O}(f(m))$ and $f(m) = \mathrm{o}(g(m))$ if $f(m)/g(m) \to 0$).

The result has the following straightforward geometric interpretation.

Corollary 1. *It is possible to construct a convex cone with generating vectors $\mathscr{F} \subseteq \{0, 1\}^n$ such that*

- $|\mathscr{F}| = (1 - \mathrm{o}(1))2^n$ *and*
- *every vector in $\mathscr{F}$ is a generator of the cone.*

2. The construction

We shall give a construction similar to the ones given in [2,3,5]. Partition n evenly into parts $P_1, \ldots, P_m$ of size, say, $t = \log n - \log\log\log n$. (In order to make the calculations more transparent we omit the use of the upper and lower integer parts, and we assume that $n = tm$.) Let $\mathscr{F}$ contain all the sets A which intersect all parts in at least one element and one part in exactly one element, i.e.,

$$\mathscr{F} = \{A \subseteq [n] : \forall i, A \cap P_i \neq \emptyset \text{ and } \exists j, |A \cap P_j| = 1\}.$$

Observe that $\mathscr{F}$ is positive linear combination free. Indeed, assume on the contrary that $A = c_1 A_1 + \cdots + c_k A_k$, $c_i > 0$. Let $x = A \cap P_j$ a one element intersection. Clearly, $A_i \cap P_j \subseteq \{x\}, \forall 1 \leqslant i \leqslant k$, else $A_i \nsubseteq A$. On the other hand, by definition of $\mathscr{F}$—, $|A_i \cap P_j| \geqslant 1$, so $A_i \cap P_j = \{x\}$, $\forall 1 \leqslant i \leqslant k$. This means that every vector A_i has one in the coordinate identified by x, so $c_1 + \cdots + c_k = 1$. But, say, $A_1 \neq A$ and, therefore, there is a coordinate ℓ where A has 1 and A_1 has 0. Thus in the weighted sum of the ℓth coordinates $c_2 + \cdots + c_k < c_1 + c_2 + \cdots + c_k = 1$, a contradiction.

It remains to show that the defined family is as large as it is stated. Let $\mathscr{F}_0 \subseteq 2^{[n]}$ be the collection containing all the sets which do not intersect at least one of the parts, i.e.,

$$\mathscr{F}_0 = \{A \subseteq [n] : \exists i, A \cap P_i = \emptyset\},$$

and $\mathscr{F}_2 \subseteq 2^{[n]}$ is the collection containing all the sets which do intersect every part in at least two elements, i.e.,

$$\mathscr{F}_2 = \{A \subseteq [n] : \forall i, |A \cap P_i| \geqslant 2\}.$$

Clearly,

$$|\mathscr{F}| \geqslant 2^n - |\mathscr{F}_0| - |\mathscr{F}_2|. \tag{1}$$

By the choice of t

$$|\mathscr{F}_0| \leqslant \frac{n}{t} 2^{n-t} = 2^n \mathrm{O}\left(\frac{\log\log n}{\log n}\right),$$

and the lower bound holds, since $|\mathscr{F}_2|$ is the smaller term in (1):

$$|\mathscr{F}_2| \leqslant (2^t - t - 1)^{n/t} \leqslant 2^n (1/e)^{n/2^t} = 2^n (1/e)^{\log\log n} < 2^n \frac{\log\log n}{\log n}.$$

3. Tight bounds on the $o(1)2^n$ term

Theorem 1.

$$2^n\left(1-\Omega\left(\frac{(\log n)^{3/2}}{\sqrt{n}}\right)\right)\leqslant f(n)\leqslant 2^n\left(1-\mathrm{O}\left(\frac{1}{\sqrt{n}}\right)\right).$$

The proof of the upper bound is—generally speaking—a typical example of the permutation method and it is quite similar to the proof of Theorem 15 in [4].

Let $\mathscr{F}$ be a positive linear combination free family on $[n]$ and denote by f_k the size of $\mathscr{F}_k$ (i.e., the size of $\{F\in\mathscr{F}:|F|=k\}$). For positive integers $p>q$, p sets $A_1,A_2,\dots,A_p$ of $[n]$ form a $(p,\{0,q\})$-system if the number of sets A_i containing x is either 0 or q for every $x\in[n]$ (i.e., they cover every element of their union exactly q times). Notice that a positive linear combination free family may not contain a $(p,\{0,q\})$-system $A_1,A_2,\dots,A_p$ together with $A=\bigcup_{i=1}^p A_i$, because $A=1/q\sum_{i=1}^p A_i$.

Let $\mathscr{H}$ be a $(p,\{0,q\})$-system $A_1,A_2,\dots,A_p$ on $[n]$, $A=\bigcup_{i=1}^p A_i$, $K=\{|H|:H\in\mathscr{H}\}$, $\alpha_k=|\{H\in\mathscr{H}:|H|=k\}|$. If $|A|=m$ and $f_m\geqslant c\binom{n}{m}$ then

$$\sum_{k\in K}\frac{\alpha_k f_k}{\binom{n}{k}}\leqslant p-c. \tag{2}$$

Indeed, consider a permutation π of $[n]$ and apply it to $\mathscr{H}$ and consider $\pi(\mathscr{H})\cap\mathscr{F}$. It consists of at most $p-1$ hyperedges for every $\pi(A)\in\mathscr{F}_m\subseteq\mathscr{F}$. Therefore,

$$\begin{aligned}\sum_{\pi\in S_n}|\pi(\mathscr{H})\cap\mathscr{F}|&\leqslant(p-1)\sum_{\substack{\pi\in S_n\\ \pi(A)\in\mathscr{F}}}1+p\sum_{\substack{\pi\in S_n\\ \pi(A)\notin\mathscr{F}}}1\\ &=(p-1)\frac{|\mathscr{F}_m|}{\binom{n}{m}}n!+p\left(1-\frac{|\mathscr{F}_m|}{\binom{n}{m}}\right)n!=\left(p-\frac{|\mathscr{F}_m|}{\binom{n}{m}}\right)n!\\ &\leqslant(p-c)n!.\end{aligned}$$

On the other hand, every edge $E\in\mathscr{H}_k$ appears exactly $f_k|E|!(n-|E|)!$ times on the left-hand side. We obtain

$$\sum_{k\in K}\alpha_k f_k k!(n-k)!\leqslant(p-c)n!.$$

Rearranging we get (2).

Now choose, say, $c=\frac{1}{2}$. If $f_{n/2}<\frac{1}{2}\binom{n}{n/2}$ then—by Stirling's formula—we are ready. Else we explicitly construct some $(p_i,\{0,q_i\})$-systems $\mathscr{H}_i$ with $q_i=p_i-1$ on the vertex set $[n]$ and then apply (2) to it. For $\sqrt{n}/4\leqslant i\leqslant\sqrt{n}/2$ let p_i be the positive integer so that

$$p_i\left\lfloor\frac{n}{2}-2i\right\rfloor+r_i=(p_i-1)\frac{n}{2},$$

for some $0\leqslant r_i<p_i$. Clearly, in this range of i,

$$\frac{\sqrt{n}}{2}\leqslant p_i\leqslant\sqrt{n}. \tag{3}$$

The edges of $\mathscr{H}_i$ $\{A_1,\dots,A_{p_i}\}$ are defined as follows. A_j meets $[p_i]$ in q_i vertices, $A_j\cap[p_i]=\{j,j+1,\dots,j+q-1\}$ (we have to take the elements here modulo p_i), and for $p_i<x\leqslant n/2$ the element x belongs to the edges A_{q_ix+j} for $1\leqslant j\leqslant q_i$ (again indices are taken modulo p_i). Then $\mathscr{H}_i$ consists of edges A_j of sizes $\lfloor n/2-2i\rfloor$ and $\lfloor n/2-2i\rfloor+1$

only and $|A| = |\cup_{j=1}^{p_i} A_j| = n/2$. Since $f_{n/2} \geqslant \frac{1}{2}\binom{n}{n/2}$, it follows from (2) that for every $\sqrt{n}/4 \leqslant i \leqslant \sqrt{n}/2$

$$f_{\frac{n}{2}-2i} \Big/ \binom{n}{\frac{n}{2}-2i} \leqslant \frac{p_i - 1/2}{p_i} \leqslant 1 - \frac{1}{2\sqrt{n}}$$

or

$$f_{\frac{n}{2}-2i+1} \Big/ \binom{n}{\frac{n}{2}-2i+1} \leqslant \frac{p_i - 1/2}{p_i} \leqslant 1 - \frac{1}{2\sqrt{n}}.$$

Let

$$I = \left\{ \frac{n}{2} - \sqrt{n} \leqslant i \leqslant \frac{n}{2} - \frac{\sqrt{n}}{2} : f_i \leqslant 1 - \frac{1}{2\sqrt{n}} \binom{n}{i} \right\},$$

we have that I has no large gap, $I \cap \{k, k+1\} \neq \emptyset$ for every $n/2 - \sqrt{n} \leqslant k < n/2 - \sqrt{n}/2$. Therefore, $|I| \geqslant \sqrt{n}/4$. Then

$$\begin{aligned} |\mathscr{F}| &= \sum_{i=0}^{n} f_i \leqslant \sum_{i=0}^{n} \binom{n}{i} - \frac{1}{2\sqrt{n}} \sum_{i \in I} \binom{n}{i} \\ &\leqslant 2^n - \frac{1}{2\sqrt{n}} |I| \binom{n}{\frac{n}{2} - \sqrt{n}} = 2^n \left(1 - \mathrm{O}\left(\frac{1}{\sqrt{n}}\right)\right), \end{aligned}$$

which gives the upper bound.

We shall get the tight lower bound using a very similar random approach to the one given in the proof of Theorem 4.1 in [3]. First of all observe that in our construction we do not necessarily need a partition.

Claim 1. *Let $\mathscr{G} \subseteq 2^{[n]}$ and $\mathscr{F}$ contain all the sets A which intersect every $B \in \mathscr{G}$ and one part $B \in \mathscr{G}$ in exactly one element*, *i.e.*,

$$\mathscr{F} = \{A \subseteq [n] : \forall B \in \mathscr{G}, A \cap B \neq \emptyset \text{ and } \exists B' \in \mathscr{G}, |A \cap B'| = 1\}.$$

Then $\mathscr{F}$ is positive linear combination free.

Proof. The proof is exactly the same as the one given in the construction: we did not utilize there that there was a partition. □

For an arbitrary family $\mathscr{F} \subseteq 2^{[n]}$ we associate an *ideal* $\mathscr{I}(\mathscr{F})$ *induced* by $\mathscr{F}$ as follows:

$$\mathscr{I}(\mathscr{F}) = \{I \subseteq [n] : \exists A \in \mathscr{F} \text{ such that } I \cap A = \emptyset\}.$$

The *neighborhood* $\mathscr{N}(\mathscr{G})$ of a family $\mathscr{G}$ is defined as the family of those subsets in $[n]$ whose Hamming distance from $\mathscr{G}$ is exactly 1, i.e.,

$$\mathscr{N}(\mathscr{G}) = \{N \subseteq [n] : N \notin \mathscr{G} \text{ and } \exists G \in \mathscr{G} \text{ such that } |N \triangle G| = 1\}.$$

Note that $\mathscr{G} \cap \mathscr{N}(\mathscr{G}) = \emptyset$.

Claim 2. *For arbitrary $\mathscr{F} \subseteq 2^{[n]}$ $\mathscr{N}(\mathscr{I}(\mathscr{F}))$ is positive linear combination free*.

Proof. This is a direct consequence of Claim 1. Indeed, $\mathscr{N}(\mathscr{I}(\mathscr{F})) \cap \mathscr{I}(\mathscr{F}) = \emptyset$, so edges of $\mathscr{N}(\mathscr{I}(\mathscr{F}))$ intersect every edge in $\mathscr{F}$. Take an arbitrary set $A \in \mathscr{N}(\mathscr{I}(\mathscr{F}))$. It is a neighbor of some set $A' \in \mathscr{I}(\mathscr{F})$ and there is a set $A^* \in \mathscr{F}$ such that $A' \cap A^* = \emptyset$. Observe that $|A \cap A^*| = 1$. Indeed, $A \notin \mathscr{I}(\mathscr{F})$ so $|A \cap A^*| \geqslant 1$ and $A' \cap A^* = \emptyset$. But A differs from A' only in one element, i.e., $|A \cap A^*| \leqslant 1$. By Claim 1 $\mathscr{N}(\mathscr{I}(\mathscr{F}))$ is positive linear combination free. □

In view of Claim 2, all that we need is to construct a suitable family $\mathscr{F}$ that has an ideal $\mathscr{I}(\mathscr{F})$ with a neighborhood of size

$$|\mathscr{N}(\mathscr{I}(\mathscr{F}))| > 2^n\left(1 - c\left(\frac{(\log n)^{3/2}}{\sqrt{n}}\right)\right)$$

for some positive constant c.

Suppose that n is divisible by 8, and let $B_1 \cup \cdots \cup B_{n/2}$ be a partition of the underlying set into pairs. (One can give a similar argument without the partitioning although in our view this simplifies the proof.) Let k be an integer $k \sim \sqrt{n/\log n}$. For every $K \in \binom{[n/2]}{k}$ let ξ_K be a random variable with

$$\mathbf{Pr}(\xi_K = 1) = \frac{(1000\log n)^{3/2}}{\sqrt{n}}\binom{n/8}{k}^{-1} = p,$$

$$\mathbf{Pr}(\xi_K = 0) = 1 - p.$$

These random variables are to be chosen totally independently. Let $\mathscr{F}$ be the random family defined by

$$\mathscr{F} = \left\{\bigcup_{i\in K} B_i : \xi_K = 1\right\}.$$

We next show that the expected size of $\mathscr{N}(\mathscr{I}(\mathscr{F}))$ is as large as it was given in Theorem 1.

Let N be an arbitrary but fixed member of $2^{[n]}$. Denote the number of blocks B_i which are contained in N by n_2, and let $N_2 = \{i : B_i \subseteq N\}$. Similarly, let $N_1 = \{i : |B_i \cap N| = 1\}$, and $|N_1| = n_1$. We give an exact formula for the probability that N belongs to $\mathscr{N}(\mathscr{I}(\mathscr{F}))$. It is easy to check that N is in $\mathscr{N}(\mathscr{I}(\mathscr{F}))$ if and only if

- $\exists K : K \cap N_2 = \emptyset$, $|K \cap N_1| = 1$ and $\xi_K = 1$ (to make sure the one element intersection).
- $\forall K$: $K \cap (N_2 \cup N_1) = \emptyset \Rightarrow \xi_K = 0$ (to make sure that N is not in the ideal, i.e., it intersects every set in $\mathscr{F}$).

Since the variables ξ_K are independent, we obtain that

$$\mathbf{Pr}(N \in \mathscr{N}(\mathscr{I}(\mathscr{F}))) = (1-p)^{\binom{n/2-n_1-n_2}{k}}\left(1 - (1-p)^{n_1\binom{n/2-n_1-n_2}{k-1}}\right) \tag{4}$$

$$\geqslant \left(1 - p\binom{n/2 - n_1 - n_2}{k}\right) \tag{5}$$

$$\times\left(1 - \exp\left[-pn_1\binom{n/2 - n_1 - n_2}{k-1}\right]\right). \tag{6}$$

Here we used the inequalities $1 - xy \leqslant (1-x)^y$ which holds for $0 \leqslant x \leqslant 1$ and $y \geqslant 1$ and $(1-x)^y \leqslant \exp[-xy]$ which holds for $-\infty \leqslant x \leqslant 1$ and $y \geqslant 0$. Now suppose that N is a *typical* subset of $[n]$. More exactly, define the collection $\mathscr{T}$ of typical sets N by

$$\mathscr{T} = \left\{N \in 2^{[n]} : \left|n_2(N) - \frac{n}{8}\right| < \sqrt{n\log n} \text{ and } \left|n_1(N) - \frac{n}{4}\right| < \sqrt{n\log n}\right\}. \tag{7}$$

The well-known de Moivre–Laplace formula (see, e.g. in [6, p. 151]) gives that if $A = np + a\sqrt{npq} + \frac{1}{2}$ and $B = np + b\sqrt{npq} - \frac{1}{2}$ then

$$\sum_{A \leqslant k \leqslant B}\binom{n}{k}p^k q^{n-k} = (1 + \mathrm{o}(1))\,(\Phi(b) - \Phi(a)),$$

where

$$\Phi(x) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x} \mathrm{e}^{-u/2}\,\mathrm{d}u,$$

i.e.,

$$|\mathscr{T}| > 2^n \left(1 - \frac{1}{n}\right). \tag{8}$$

There exists some positive constant c such that for every typical set N,

$$p\binom{n/2 - n_1 - n_2}{k} = \frac{(1000\log n)^{3/2}}{\sqrt{n}} \frac{\binom{n/2 - n_1 - n_2}{k}}{\binom{n/8}{k}} < c\frac{(\log n)^{3/2}}{\sqrt{n}} \tag{9}$$

and

$$pn_1\binom{n/2 - n_1 - n_2}{k-1} = \frac{(1000\log n)^{3/2}}{\sqrt{n}} \times \frac{kn_1}{n/2 - n_1 - n_2 - k + 1} \tag{10}$$

$$\times\frac{\binom{n/2 - n_1 - n_2}{k}}{\binom{n/8}{k}} > 2\log n. \tag{11}$$

Then (9) and (11) imply the following lower bound in (6). If $N \in \mathscr{T}$ then

$$\mathbf{Pr}(N \in \mathscr{N}(\mathscr{I}(\mathscr{F}))) > 1 - c\frac{(\log n)^{3/2}}{\sqrt{n}}. \tag{12}$$

Then (8) and (12) give that the expected size $E|\mathscr{N}(\mathscr{I}(\mathscr{F}))|$ fulfils the lower bound in Theorem 1, and hence there exists such a family.

Acknowledgment

We appreciate the fruitful discussions with Noga Alon, Tom Bohman, Colin Cooper, Alan Frieze, András Gyárfás, Levon Khachatrian and Ryan Martin.

References

[1] R. Ahlswede, L. Khachatrian, Cone dependence—a basic combinatorial concept, Preprint 00-117, Diskrete Strukturen in der Mathematik SFB 343, Universität Bielefeld, 2000.
[2] F.R.K. Chung, Z. Füredi, R.L. Graham, P. Seymour, On induced subgraphs of the cube, J. Combin. Theory Ser. A 49 (1988) 180–187.
[3] Z. Füredi, J. Griggs, D.J. Kleitman, A minimal cutset of the Boolean lattice with almost all members, Graphs Combin. 5 (1989) 327–332.
[4] Z. Füredi, A. Gyárfás, M. Ruszinkó, On the maximum size of (p, Q)—free families, Discrete Math. 257 (2000) 385–403.
[5] Z. Füredi, J. Kahn, D.J. Kleitman, Sphere coverings of the hypercube with incomparable centers, Discrete Math. 83 (1990) 129–134.
[6] A. Rényi, Probability Theory, North-Holland, Amsterdam, 1970.