# Positive linear combination free families*

Zoltán Füredi †, and Miklós Ruszinkó ‡

August 9, 2002, *Discrete Applied Mathematics* **156** (2008), 1536–1541.

Dedicated to our good friend Levon Khachatrian who's sudden death is a great loss to our community.

**Abstract**

A family of subsets of $[n]$ is *positive linear combination free* if the characteristic vector of neither edge is the positive linear combination of the characteristic vectors of some other ones. We construct a positive linear combination free family which contains $(1-o(1))2^n$ subsets of $[n]$ and we give tight bounds on the $o(1)2^n$ term. The problem was posed by the late Levon Khachatrian and the result has geometric consequnces.

## 1 Positive linear combination free families

The following question was posed by Levon Khachatrian on EuroComb01 in Barcelona. How many edges may a hypergaph 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? He added that they worked on this problem with Rudolf Ahlswede, and got a construction with $(1/2 + c)2^n$ sets. They wanted to know if such a family may contain almost all edges or significantly less. Here we give an explicit cosntruction for such a family which contains $(1 - o(1))2^n$ edges and tight bounds for the $o(1)2^n$ 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 $i^{th}$ coordinate iff $i \in A$. (We use the same notation for sets and characteristic vectors.) $A$ is the

*2000 Mathematics Subject Classification: 05D05, 52B05

†Department of Mathematics, University of Illinois, Urbana, IL61801, USA, and Rényi Institute of Mathematics of the Hungarian Academy of Sciences, Budapest, P. O. Box 127, Hungary-1364. Email: `z-furedi@math.uiuc.edu`. Research supported in part by the Hungarian National Science Foundation under grant OTKA T 032452, and by the National Science Foundation under grant 1-5-29066 NSF DMS 00-70312.

‡Computer and Automation Research Institute of the Hungarian Academy of Sciences, Budapest, P. O. Box 63, Hungary-1518. Email: `ruszinko@lutra.sztaki.hu`. Research supported in part by OTKA Grants T029074, T030059, T038198, Bolyai Fellowship, and KFKI-ISYS Datamining Lab.

positive linear combination of $A_1, \dots A_k$ iff $A = c_1A_1 + \dots + c_kA_k$ and $\forall i: \; c_i > 0$. $\mathcal{F} \subseteq 2^{[n]}$ is positive linear combination free iff no set (vector) is the positive positive linear combination of some other sets from $\mathcal{F}$, i.e., for arbitrary choice of positive coefficients and $\mathcal{F}' \subseteq \mathcal{F}$

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

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

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

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

The result has the following straightforward geometric interpretation.

**Corollary 1.1** *It is possible to construct a convex cone with generating vectors $\mathcal{F} \subseteq \{0,1\}^n$ such that*

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

## 2 The construction

We shall give a construction similar to the ones given in [1], [2], [3]. Partition $n$ evenly into parts $P_1, \dots, P_m$ of size, say, $k = \log n - \log\log\log n$. (In order to make the calculations more transparent we omit the use of the upper and lower integer parts, i.e. we assume that $n = km$.) Let $\mathcal{F}$ contain all the sets $A$ which intersect all parts in at least one element and one part in exactly one element, i.e.,

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

Observe that $\mathcal{F}$ is positive linear combination free. Indeed, assume to the contrary that $A = c_1A_1 + \dots + c_kA_k$, $c_i > 0$. Let $x = A \cap P_j$ the one element intersection. Clearly, $A_i \cap P_j \subseteq \{x\}$ $\forall 1 \leq i \leq k$, else $A_i \not\subseteq A$. On the other hand – by the defintion of $\mathcal{F}$ – $|A_i \cap P_j| \geq 1$, so $A_i \cap P_j = \{x\}$ $\forall 1 \leq i \leq k$. This means that every vector $A_i$ has one in the coordinate identified by $x$, so $c_1 + \dots + c_k = 1$. But, say, $A_1 \neq A$ and, therefore, there is a coordinate $\ell$ where $A$ has 1 and $A_1$ has 0. Thus in the weighted sum of the $\ell^{th}$ coordinates $c_2 + \dots + c_k < c_1 + c_2 \dots + c_k = 1$, a contradiction.

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

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

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

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

Clearly,

$$|\mathcal{F}| \geq 2^n - |\mathcal{F}_0| - |\mathcal{F}_2|. \tag{1}$$

By the choice of $k$

$$|\mathcal{F}_0| \leq \frac{n}{k} 2^{n-k} = 2^n O\left(\frac{\log\log n}{\log n}\right),$$

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

$$|\mathcal{F}_2| \leq \left(2^k - k - 1\right)^{n/k} \leq 2^n (1/e)^{n/2^k} = 2^n (1/e)^{\log\log n} < 2^n \frac{\log\log n}{\log n}.$$

□

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

**Theorem 3.1** $2^n \left(1 - \Omega\left(\frac{(\log n)^{3/2}}{\sqrt{n}}\right)\right) \leq f(n) \leq 2^n \left(1 - O\left(\frac{1}{\sqrt{n}}\right)\right).$

The proof of the upper bound is – generally speaking – an averaged version of the permutation method and it is quite similar to the proof of Theorem [15] in [4].

Let $\mathcal{F}$ be a positive linear combination free family on $[n]$ and denote by $f_k$ the size of $\mathcal{F}_k$ (i.e., the size of $\{F \in \mathcal{F} : |F| = k\}$). For positive integers $p > q$ call $p$ sets $A_1, A_2, \ldots, A_p$ of $[n]$ 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 $A_1, A_2, \ldots, A_p$ together with $A = \cup_{i=1}^p A_i$, because $A = 1/q \sum_{i=1}^p A_i$.

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

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

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

$$\sum_{\pi\in S_n}|\pi(\mathcal{H})\cap\mathcal{F}|\leq(p-1)\sum_{\substack{\pi\in S_n\\ \pi(A)\in\mathcal{F}}}1+p\sum_{\substack{\pi\in S_n\\ \pi(A)\not\in\mathcal{F}}}1\leq c(p-1)n!+cpn=(p-c)n!.$$

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

$$\sum_{E\in\mathcal{H}}f_k|E|!(n-|E|)!=\sum_{k\in K}\alpha_k f_k k!(n-k)!\leq(p-c)n!.$$

Rearranging we get (2).

Now Choose, say, $c=1/2$. If $f_{n/2}<1/2\binom{n}{n/2}$ then – by Stirling formula – we are ready. Else we explicitly construct some $(p_i,\{0,q_i\})$-systems $\mathcal{H}_i$ with $q_i=p_i-1$ on vertex set $[n/2]$ and then will apply (2) to it. For $\sqrt{n}/4\leq i\leq\sqrt{n}/2$ 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 $0\leq r_i<p_i$. Clearly, in this range of $i$,

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

The edges of $\mathcal{H}_i$ $\{A_1,\dots,A_{p_i}\}$ 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 $p_i<x\leq n/2$ the element $x$ belongs to the edges $A_{q_ix+j}$ for $1\leq j\leq q_i$ (again indices are taken modulo $p_i$). Then $\mathcal{H}_i$ consists of edges $A_i$ of sizes $\lfloor n/2-2i\rfloor$ and $\lfloor n/2-2i\rfloor+1$ only and $|A|=|\cup_{i=1}^{p}A_i|=n/2$. Since $f_{n/2}\geq 1/2\binom{n}{n/2}$, it follows from (2) that for every $\sqrt{n}/4\leq i\leq\sqrt{n}/2$

$$\begin{aligned}\text{either}\quad f_{\frac{n}{2}-2i}\Big/\binom{n}{\frac{n}{2}-2i}\quad&\leq\quad\frac{p_i-1/2}{p_i}\leq 1-\frac{1}{\sqrt{n}}\\ \text{or}\quad f_{\frac{n}{2}-2i+1}\Big/\binom{n}{\frac{n}{2}-2i+1}\quad&\leq\quad\frac{p_i-1/2}{p_i}\leq 1-\frac{1}{\sqrt{n}}\end{aligned}$$

Let

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

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

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

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 [2]. First of all observe that in our construction we do not need necessarily a partition.

**Claim 3.2** *Let $\mathcal{G} \subseteq 2^{[n]}$ and $\mathcal{F}$ contain all the sets $A$ which intersect every $B \in \mathcal{G}$ and one part $B \in \mathcal{G}$ in exactly one element, i.e.,*

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

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

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

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

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

.

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

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

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

**Claim 3.3** *For arbitrary $\mathcal{F} \subseteq 2^{[n]}$ $\mathcal{N}(\mathcal{I}(\mathcal{F}))$ is positive linear combination free.*

**Proof.** Indeed, $\mathcal{N}(\mathcal{I}(\mathcal{F})) \cap \mathcal{I}(\mathcal{F}) = \emptyset$, so edges of $\mathcal{N}(\mathcal{I}(\mathcal{F}))$ intersect every edge in $\mathcal{F}$. Take an arbitrary set $A \in \mathcal{N}(\mathcal{I}(\mathcal{F}))$. It is neighbor of some set $A' \in \mathcal{I}(\mathcal{F})$ and there is a set $A^* \in \mathcal{F}$ such that $A' \cap A^* = \emptyset$. Observe that $|A \cap A^*| = 1$. Indeed, $A \notin \mathcal{I}(\mathcal{F})$ so $|A \cap A^*| \geq 1$ and $A' \cap A^* = \emptyset$. But $A$ differs from $A'$ only in one element, i.e., $|A \cap A^*| \leq 1$. By Claim 3.2 $\mathcal{N}(\mathcal{I}(\mathcal{F}))$ is positive linear combination free. □

In view of Claim 3.3, all that we need is to construct a suitable family $\mathcal{F}$ that has an ideal $\mathcal{I}(\mathcal{F})$ with a neighborhood of size

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

for some positive constant $c$.

Suppose that $n$ is divisible by 8, and let $B_1 \cup \ldots B_{n/2}$ be a partition of the underlying set into pairs. Let $k$ be an integer $k \sim \sqrt{n/\log n}$. For every $K \in \binom{[n/2]}{k}$ let $\xi_K$ be a random variable with

$$\begin{aligned} \mathbf{Pr}(\xi_K = 1) &= \frac{(1000 \log n)^{3/2}}{\sqrt{n}} \binom{n/8}{k}^{-1} = p \\ \mathbf{Pr}(\xi_K = 0) &= 1 - p. \end{aligned}$$

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

$$\mathcal{F} = \{\cup_{i \in K} B_i : \xi_K = 1\}.$$

We next show that the expected size of $\mathcal{N}(\mathcal{I}(\mathcal{F}))$ is as large as it was given in Theorem 3.1.

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 $N_2 = \{i : B_i \subset N\}$. 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 $\mathcal{N}(\mathcal{I}(\mathcal{F}))$. It is easy to check that $N$ is in $\mathcal{N}(\mathcal{I}(\mathcal{F}))$ 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)
- $\forall K$ s.t. $K \cap (N_2 \cup N_1) = \emptyset \quad \xi_K = 0$ (to make sure that $N$ is not in the ideal, i.e., it intersects every set in $\mathcal{F}$)

Since the variables $\xi_K$ are independent, we obtain that

$$\begin{aligned} \mathbf{Pr}(N \in \mathcal{N}(\mathcal{I}(\mathcal{F}))) &= (1-p)^{\binom{n-n_1-n_2}{k}} \left(1 - (1-p)^{n_1 \binom{n-n_1-n_2}{k-1}}\right) && (4) \\ &\geq \left(1 - p\binom{n-n_1-n_2}{k}\right) \left(1 - \exp\left[-pn_1 \binom{n-n_1-n_2}{k-1}\right]\right) && (5) \end{aligned}$$

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

$$\mathcal{T} = \left\{ N \in 2^{[n]} : \left| n_2(N) - \frac{n}{8} \right| < \sqrt{n \log n} \text{ and } \left| n_1(N) - \frac{n}{4} \right| < \sqrt{n \log n} \right\}. \qquad (6)$$

The well-known de Moivre-Laplace formula (see, e.g. in [5], p. 151) gives that if $A = np + a\sqrt{npq} + 1/2$ and $B = np + b\sqrt{npq} - 1/2$ then

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

where

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

i.e.,

$$|\mathcal{T}| > 2^n \left(1 - \frac{1}{n}\right). \tag{7}$$

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

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

and

$$pn_1\binom{n-n_1-n_2}{k-1} = \frac{(1000\log n)^{3/2}}{\sqrt{n}} \frac{kn_1}{n-n_1-n_2-k+1} \frac{\binom{n-n_1-n_2}{k}}{\binom{n/8}{k}} > 2\log n. \tag{9}$$

Then (8) and (9) imply the following lower bound in (5). If $N \in \mathcal{T}$ then

$$\mathbf{Pr}(N \in \mathcal{N}(\mathcal{I}(\mathcal{F}))) > 1 - c\frac{(\log n)^{3/2}}{\sqrt{n}}. \tag{10}$$

Then (7) and (10) give that the expected size $E\,|\mathcal{N}(\mathcal{I}(\mathcal{F}))|$ fulfils the lower bound in Theorem 3.1, 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] 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.

[2] Z. Füredi, J. Griggs, D.J. Kleitman, A minimal cutset of the Boolean lattice with almost all members. *Graphs Combin.* **5** (1989), 327–332.

[3] Z. Füredi, J. Kahn, D.J. Kleitman, Sphere coverings of the hypercube with incomparable centers, *Discrete Math* **83** (1990), 129–134.

[4] Z. Füredi, A. Gyárfás, M. Ruszinkó, On the maximum size of $(p, Q)$–free families, *Discrete Mathematics*, to appear.

[5] A. Rényi, *Probability Theory*, 1970, North-Holland, Amsterdam.