Discrete Mathematics 83 (1990) 129–134
North-Holland

NOTE

SPHERE COVERINGS OF THE HYPERCUBE WITH INCOMPARABLE CENTERS

Zoltán FÜREDI*
Math. Inst. Hungarian Acad. Sci. 1364 Budapest, P.O.B. 127, Hungary

Jeff KAHN**
Dept. Math. and RUTCOR, Rutgers University, New Brunswick, NJ 08903, USA

Daniel J. KLEITMAN***
Dept. Math., Massachusetts Institute of Technology, Cambridge, MA 02139, USA

Received 1 March 1988

It is shown that the shadow of a Sperner family can cover 10 percent of the Boolean algebra. Whether this can be improved to $(100 - o(1))\%$ remains open.

1. Shadows of Sperner families

Let $[n]$ denote the set of the first n integers, $2^{[n]}$ its power set. The collection of all k-subsets of a set S is denoted by $\binom{S}{k}$. Let $\mathscr{F}$ be a subfamily of $2^{[n]}$. The *neighborhood* of $\mathscr{F}$, $N(\mathscr{F})$, is defined as the family of sets in $[n]$ whose Hamming distance is exactly 1 from $\mathscr{F}$, i.e. $N(\mathscr{F}) = \{N \subset [n] : N \notin \mathscr{F}$ and there exists an $F \in \mathscr{F}$ such that $|N \triangle F| = 1\}$. (If we identify the subsets of $[n]$ with the vertices of the n-dimensional unit-cube, then $N(\mathscr{F})$ is the usual neighborhood in the graph Q^n.) The *shadow* of $\mathscr{F}$, $\partial\mathscr{F}$, consists of those members of $N(\mathscr{F})$ which are covered by a member of $\mathscr{F}$, i.e. $\partial\mathscr{F} = \{S : S \notin \mathscr{F}$ and there exists an $F \in \mathscr{F}$ such that $S \subset F$, $|F \backslash S| = 1\}$.

The family $\mathscr{F}$ is a *Sperner family* if no two of its members contain each other. One of the oldest results in the theory of finite sets states that the size of the largest Sperner family is $\binom{n}{\lfloor n/2 \rfloor}$ and the extremal family consists of all members of $2^{[n]}$ of size either $\lfloor n/2 \rfloor$ or $\lceil n/2 \rceil$ (Sperner [13]). The size of the shadow of such a family is again a binomial coefficient, so it is not more than $\binom{n}{\lfloor n/2 \rfloor}$. Engel [2] and independently Zuev [14] conjectured that there exists a positive real C such that

$$|\partial\mathscr{F}| < C\binom{n}{n/2} < C' \frac{2^n}{\sqrt{n}} \tag{1.1}$$

holds for every Sperner family $\mathscr{F}$. This was disproved by Kospanov [8] who

* Research supported in part by Airforce Grant OSR-86-0076.
** Research supported in part by NSF Grant MCS 83-01867, Airforce Grant OSR 0271 and a Sloan Research Fellowship
*** Research supported in part by NSF grant DMS 86-06225 and Airforce Grant OSR-86-0076.

0012-365X/90/$03.50 © 1990 — Elsevier Science Publishers B.V. (North-Holland)

showed that

$$\max |\partial \mathscr{F}| > cn^{-\frac{1}{6}}2^n.$$

Griggs [3] also constructed a family whose shadow was larger than $\log n\binom{n}{n/2}$. The aim of this note is to prove

Theorem. *There exists a Sperner family $\mathscr{S}$ over n elements such that $|\partial \mathscr{S}| > 0.1 \cdot 2^n$ (for all $n > n_0$).*

Conjecture. *There exists a $c < 1$ such that $|\partial S| < c2^n$ holds for every Sperner family $\mathscr{S}$.*

A theorem of Kostochka [9] implies that

$$|\partial \mathscr{S}| < \left(1 - \frac{(\log n)^{\frac{3}{2}}}{100\sqrt{n}}\right)2^n,$$

which is the best upper bound we know.

2. The random construction

We use a random construction. The problem of finding an explicit construction giving a similar bound remains open. Let t be an integer, $t = (1 + o(1))\sqrt{n/2}$, and denote $\lfloor (n-t)/2 \rfloor$ by s. Then the size of the middle t levels of the Boolean lattice is

$$\sum_{a=s+1}^{s+t} \binom{n}{a} = (1 + o(1))2^n\left(\Phi\left(\frac{1}{\sqrt{2}}\right) - \Phi\left(-\frac{1}{\sqrt{2}}\right)\right) = (1 + o(1))0.520 \ldots \cdot 2^n. \quad (2.1)$$

Here $\Phi(x) = (2\pi)^{-\frac{1}{2}} \int_{-\infty}^{x} \mathrm{e}^{-y^2/2}\,\mathrm{d}y$, as usual. Let k be an integer, $k = (1 + o(1))\sqrt{n/2}$. We are going to define disjoint random families $\mathscr{K}(1), \ldots, \mathscr{K}(t)$ of k-sets. Let c be a fixed positive real (in the following calculations $c = 0.75$) and define p by the equation

$$tp\binom{s+t}{k} = c.$$

For every $K \in \binom{[n]}{k}$ let ξ_K be a random variable with

$$\mathrm{Prob}(\xi_K = 0) = 1 - tp$$
$$\mathrm{Prob}(\xi_K = i) = p$$

for $i = 1, \ldots, t$. These random variables are to be chosen totally independently. Let $\mathscr{K}(i)$ be the random family defined by $\mathscr{K}(i) = \{K \in \binom{[n]}{k} : \xi_K = i\}$. Finally, we define the family $\mathscr{S}$ as $\mathscr{S} = \mathscr{S}_1 \cup \cdots \cup \mathscr{S}_t$ where $\mathscr{S}_i$ is the family of those $s + i + 1$-element sets which contain a member of $\mathscr{K}(i)$ but do not contain any members of $\mathscr{K}(j)$ with $1 \leqslant j < i$. Obviously, $\mathscr{S}$ is a Sperner family.

We next show that the expected size of the shadow of $\mathscr{S}$ is greater than $0.1 \cdot 2^n$ (if $n > n_0$.) This implies the existence of a Sperner family with such a large shadow. To prove this we show that every a-element set A belongs to $\partial\mathcal{S}$ with a probability at least 0.2 if $s+1 \leqslant a \leqslant s+t$ and $A \subset [n]$, and then we use (2.1). For a family $\mathcal{F}$ and a set A we use the notation $\mathcal{F}_A$ for the induced subfamily, i.e. $\mathcal{F}_A = \{F \in \mathcal{F} : F \subset A\}$. Let $\mathcal{K}([i])$ denote $\mathcal{K}(1) \cup \cdots \cup \mathcal{K}(i)$.

$$\begin{aligned}
&\mathrm{Prob}(A \in \partial S_i)\\
&\quad \geqslant \mathrm{Prob}(\mathcal{K}([i])_A = \emptyset)\mathrm{Prob}(\exists x : A \cup \{x\} \in \mathcal{S}_i \mid \mathcal{K}([i])_A = \emptyset)\\
&\quad = \mathrm{Prob}(\mathcal{K}([i])_A = \emptyset)(1 - \mathrm{Prob}(\forall x \in [n]\backslash A : A \cup \{x\} \notin \mathcal{S}_i \mid \mathcal{K}([i])_A = \emptyset))\\
&\quad = \mathrm{Prob}(\mathcal{K}([i])_A = \emptyset)(1 - (\mathrm{Prob}(A \cup \{x\} \notin \mathcal{S}_i \mid \mathcal{K}([i])_A = \emptyset))^{n-a})\\
&\quad = \mathrm{Prob}(\mathcal{K}([i])_A = \emptyset)(1 - (1 - \mathrm{Prob}(A \cup \{x\} \in \mathcal{S}_i \mid \mathcal{K}([i])_A = \emptyset))^{n-a})\\
&\quad \geqslant \mathrm{Prob}(\mathcal{K}([i])_A = \emptyset)(1 - \exp[-\mathrm{Prob}(A \cup \{x\} \in \mathcal{S}_i \mid \mathcal{K}([i])_A = \emptyset)(n-a)]). \qquad (2.2)
\end{aligned}$$

Here we used the inequality $(1-x)^y \leqslant \exp[-xy]$ which holds for all reals $x \leqslant 1$ and $y \geqslant 0$. We estimate separately the two probabilities in the last line of (2.2).

$$\begin{aligned}
\mathrm{Prob}(\mathcal{K}([i])_A = \emptyset) = (1-ip)^{\binom{a}{k}} \geqslant (1-tp)^{\binom{s+t}{k}} &= (1+o(1))\exp\left[-tp\binom{s+t}{k}\right]\\
&= (1+o(1))\exp[-c]. \qquad (2.3)
\end{aligned}$$

Moreover

$$\begin{aligned}
\mathrm{Prob}(A \cup \{x\} \in S_i \mid \mathcal{K}([i])_A = \emptyset) &= (1-(i-1)p)^{\binom{a}{k-1}} - (1-ip)^{\binom{a}{k-1}}\\
&\geqslant p\binom{a}{k-1}(1-ip)^{\binom{a}{k-1}}. \qquad (2.4)
\end{aligned}$$

Here the last factor is $1-o(1)$, because

$$(1-ip)^{\binom{a}{k-1}} \geqslant 1 - ip\binom{a}{k-1} = 1 - ip\binom{a}{k}\frac{k}{a-k+1} \geqslant 1 - \frac{ck}{a-k+1}.$$

Moreover we have (see, e.g., in [10, p. 151]) that

$$\binom{a}{k} \geqslant \binom{s+t}{k}\exp[-tk/s](1-o(1)). \qquad (2.5)$$

Applying this to (2.4), we obtain

$$\begin{aligned}
&\mathrm{Prob}(A \cup \{x\} \in S_i \mid \mathcal{K}([i])_A = \emptyset) \geqslant p\binom{a}{k-1}(1-o(1))\\
&\quad = \frac{1-o(1)}{a-k+1}kp\binom{a}{k} \geqslant \frac{1-o(1)}{a-k+1}kp\binom{s+t}{k}\exp[-1+o(1)] = (1+o(1))\frac{c}{es}.
\end{aligned}$$

Using this result in (2.2) we obtain

$$\begin{aligned}
\mathrm{Prob}(A \in \partial S_i) &\geqslant (1-o(1))\exp[-c]\left(1 - \exp\left[-(n-a)\frac{c}{es}\right]\right)\\
&= (1-o(1))\exp[-c]\left(1 - \exp\left[-\frac{2c}{e}\right]\right) > 0.2003\ldots. \quad \square
\end{aligned}$$

Remark. See also [9] for a similar, though simpler, construction.

3. The complexity of the Boolean functions

The minimum number of conjunctions. Let $f(\boldsymbol{x})$ be a Boolean function of n variables, $f(x_1, \ldots, x_n)$: $\{0, 1\}^n \to \{0, 1\}$. Let $d(f)$ be the smallest integer d such that one can write f in a disjunctive normal form of d conjunctions, i.e. $d(f) =: \min\{d : \exists K_1 \cdots K_d$ such that $f(\boldsymbol{x}) = K_1 \vee \cdots \vee K_d\}$, where every term K has the form

$$K = x_{i_1}^{\varepsilon_1} \ldots x_{i_r}^{\varepsilon_r} \quad \text{where } x^{\varepsilon} = \begin{cases} x & \text{if } \varepsilon = 1, \\ \bar{x} & \text{if } \varepsilon = -1. \end{cases}$$

Korshunov [6] proved that there are positive reals c_1 and c_2 such that

$$c_1 \frac{2^n}{\log n \log \log n} < d(f) < c_2 \frac{2^n}{\log n \log \log n} \tag{3.1}$$

holds for almost all Boolean function f. Sapozhenko [12] gave a simple algorithm which provides a disjunctive normal form of length $c2^n/\log n$ for almost all Boolean function.

They also investigated the length of the longest irreducible normal form of f. A disjunctive normal form of the Boolean function f is called *irreducible* if by removal of a conjunction or of a letter one obtains a disjunctive normal form which does not generate f. Let $d_{\max}(f)$ denote the maximum number of conjunctions among all irreducible disjunctive normal forms which generate f. Sapozhenzo [11] proved that $d_{\max}(f) \sim 2^{n-1}$ for almost all f. For a short proof see Korshunov [7].

Representations by systems of linear inequalities. In [1] and [5] Balas and Jeroslow introduced the following notion. Let Z be a subset of $\{0, 1\}^n$, i.e. a finite point set in $\mathbb{R}^n$. Then let $l(Z)$ denote the minimum number of l of linear inequalities

$$\sum_{j=1}^{n} a_{ij} x_j \leq b_i \quad \text{where } i = 1, \ldots, l \tag{3.2}$$

such that the set of all $0-1$ solutions of (3.2) is exactly Z. If we identify the Boolean function f by its zero set, then this definition can be extended, i.e. let $Z(f) =: \{\boldsymbol{x} : f(\boldsymbol{x}) = 0\}$ and set $l(f) = l(Z(f))$. Denote by Q^n the graph of the *n-dimensional cube*, i.e. the vertex set of Q^n consists of all the $(0, 1)$-vectors of length n, and two vectors $\boldsymbol{x}, \boldsymbol{y} \in \{0, 1\}^n$ are adjacent if they differ from each other in exactly one component. For a graph $\mathcal{G}$ we denote the number of connected components by $c(\mathcal{G})$. Let $\bar{Z}$ denote the complement of Z in $\{0, 1\}^n$. Then it is easy to see [5, 4] that

$$c(Q^n_{\bar{Z}}) \leq l(Z) \leq 2^{n-1},$$

and that [14]

$$l(f) \leq d(f).$$

An asymptotic formula, analogous to (3.1), is not known for $l(f)$. It is possible, for example, that $l(f)=1$ while $d(f)=\binom{n}{\lfloor n/2\rfloor}$. Zuev [14] proved that for almost all Boolean function f, $l(f)\geq 2^n/n^2$ holds.

Monotone Boolean functions. A subset $Z\subset\{0,1\}^n$ is called *monotone* if $\boldsymbol{x}\in Z$ and $\boldsymbol{x}\leq\boldsymbol{y}$ imply $\boldsymbol{y}\in Z$. A Boolean function φ is monotone if $Z(\varphi)$ is monotone. Hammer, Ibaraki and Peled [4] proved that

$$\frac{1}{n}\binom{n}{\lfloor n/2\rfloor}\leq\max_{\varphi} l(\varphi)\leq\binom{n}{\lfloor n/2\rfloor}, \tag{3.3}$$

where φ runs over monotone functions. This was improved by Zuev [14]

$$l(\varphi)\leq N(n)\frac{1+\log n}{n}+1, \tag{3.4}$$

where $N(n)$ denotes the maximum size of the neighborhood of a Sperner family in $2^{[n]}$. (Actually, his proof was not completely clear for the authors of this paper.) Then (3.4) implies that $l(\varphi)\leq(c2^n\log n)/n$ holds for all monotone φ. He conjectures that the true order of the magnitude of $\max_\varphi l(\varphi)$ is given by the lower bound in (3.3).

References

[1] E. Balas and R.G. Jeroslow, Canonical cuts on the unit hypercube, SIAM J. Appl. Math. 23 (1972) 61–69.

[2] K. Engel, Lecture in the Meeting on Ordered Sets, Oberwolfach, February, 1985.

[3] J.R. Griggs, private communication (unpublished)

[4] P.L. Hammer, T. Ibaraki and U.N. Peled, Threshold numbers and threshold completions, in Studies on graphs and discrete programming (Brussels, 1979), North-Holland Math. Studies 59, Annals of Discrete Math. 11 (North-Holland, Amsterdam-New York, 1981) 125–145.

[5] R.G. Jeroslow, On defining sets of vertices of hypercube by linear inequalities, Discrete Math. 11 (1975) 119–124.

[6] A.D. Korshunov, Complexity of shortest disjunctive normal forms of Boolean functions (Russian), Metody Diskret. Analiz. 37 (1981) 9–41, 85.

[7] A.D. Korshunov, Complexity of shortest disjunctive normal forms of random Boolean functions (Russian). Metody Diskret. Analiz 40 (1983) 25–53, 100–101.

[8] E. Sh. Kospanov, On coverings by unit balls whose centers are incomparable, Metody Diskret. Analiz 44 (1986) 54–57.

[9] A. V. Kostochka, Maximum order of the boundary of a filter in an n-dimensional cube (Russian), Metody Discret. Analiz 41 (1984) 49–61, 109.

[10] A. Rényi, Probability Theory (North-Holland, Amsterdam, 1970).

[11] A. A. Sapozhenko, The greatest length of a dead-end disjunctive normal form for almost all Boolean functions (Russian), Mat. Zametki 4 (1968) 649–658.

[12] A. A. Sapozhenko, The complexity of disjunctive normal forms that are obtainable by means of the gradient algorithm (Russian), Diskret. Analiz 21 (1972) 62–71, 96.

[13] E. Sperner, Ein Satz über Untermengen einer endlichen Menge, Math. Z. 27 (1928) 544–548.

[14] Yu. A. Zuev, Representation of Boolean functions by systems of linear inequalities (Russian, English summary), Kibernetika (Kiev) (1985, no. 5) i, 7–9, 40, 133 (English translation: Cybernetics 21 (1985) 567–571.)