Bounds for Optimal Coverings

Carlos Gustavo T. de A. Moreira [a] and Yoshiharu Kohayakawa [b]

[a] *IMPA – Estrada Dona Castorina 110, 22460–320 Rio de Janeiro, RJ, Brazil – e-mail: gugu@impa.br*

[b] *Instituto de Matemática e Estatística, USP – Rua do Matão 1010, 05508–900 São Paulo, SP, Brazil – e-mail: yoshi@ime.usp.br*

Abstract

We give bounds for optimal coverings of finite sets by elements of regular families of subsets, and show that both upper and lower bounds are asymptotically sharp for some families of examples.

Key words: Coverings, set systems, hypergraphs, the probabilistic method

We study coverings of finite sets by subsets belonging to a regular family. By an (a, b)*-regular* family of subsets of a finite set F, or an (a, b)*-regular hypergraph on* F, we mean a family $\mathcal{C}$ for which there are integers a and b such that for each $C \in \mathcal{C}$ we have $|C| = a$, and, for each $x \in F$, we have $\deg(x) = b$, where as usual $\deg(x) = \left|\{C \in \mathcal{C} \mid x \in C\}\right|$. Thus (a, b)-regular families are what are sometimes called *a-uniform*, *b-regular* hypergraphs. The problem of constructing small covers for such hypergraphs and of estimating the minimum possible size for such covers are common, and appear in many contexts, as in the study of covering codes (see [2]), in particular, in football pool problems (see [5]), in the study of vertex covers of graphs, and others. Somewhat surprisingly, some sharp results are known for problems of this kind, see, e.g., [1], [3, Theorem 8.11], [4], and [8]. In this note, we prove another result that shows that one may require strong regularity conditions and still obtain hypergraphs that behave asymptotically 'badly' with respect to the size of covers that they admit.

If $\mathcal{C}$ is a hypergraph on F, let $\alpha(F, \mathcal{C})$ be the minimal integer r for which there exist $C_1, \ldots, C_r \in \mathcal{C}$ such that $\bigcup_{1 \le j \le r} C_j = F$. In this note, we shall be

⋆ Partially supported by MCT/FINEP/CNPq (ProNEx Project 416/96 and 107/97), FAPESP (Proc. 96/04505–2), CNPq (Proc. 300647/95-6 and 300334/93–1), and by FAPERJ.

⋆⋆ The authors gratefully acknowledge the careful work of the referees.

concerned with estimating $\alpha(F,\mathcal{C})$. Our first result is a slight improvement of classical results by Johnson, Stein, and Lovász (see [6], [7], and [9]).

Proposition 1 *Let F be an n-element set and $\mathcal{C}$ an (a,b)-regular hypergraph on F. Let also $\ell = \min\{a, \lceil n/a \rceil\}$. Then*

$$\left\lceil \frac{n}{a} \right\rceil \le \alpha(F,\mathcal{C}) \le \frac{\ln(a/\ell)}{\ln \dfrac{1}{1-a/n}} + \frac{n}{a}\sum_{r=1}^{\ell}\frac{1}{r} < \frac{n}{a}\Big(\ln(a+1)+\gamma\Big) - \frac{1}{2}\ln\frac{a}{\ell} + \frac{1}{2}, \quad (1)$$

where $\gamma = 0.5772156649\ldots$ is Euler's constant.

The proof of Proposition 1 is based on the following remark: given a subset $A \subset F$, there is always a set $C \in \mathcal{C}$ such that $|A \cap C| \ge (b/|\mathcal{C}|)|A| = (a/n)|A|$. (This remark follows from an averaging argument.) We omit the details.

Remark 2 *If we just suppose that $|C| \le a$ for all $C \in \mathcal{C}$ and $\deg(x) \ge b$ for all $x \in F$, then we can set $m = |\mathcal{C}|$ and conclude that $n/a \le \alpha(F,\mathcal{C}) \le f(a,b,m,n)$, where*

$$f(a,b,m,n) = \begin{cases} \dfrac{n}{a} + \dfrac{m}{b}\displaystyle\sum_{r=2}^{a}\frac{1}{r} \le \dfrac{m}{b}(\ln(a+1)+\gamma) & \text{if } b \le m/\sqrt{n} \\ \dfrac{m}{b}\left(\ln\left(\dfrac{bn}{m}\right)+\gamma\right) - \dfrac{1}{2}\ln\left(\dfrac{b^2 n}{m^2}\right) + \dfrac{1}{2} & \text{if } b \ge m/\sqrt{n}. \end{cases}$$

Our main result, given in the following proposition, shows that the estimates of the previous proposition are asymptotically sharp for certain regular families of subsets of finite sets. We work with families of translations of a-element subsets of $\mathbb{Z}/n\mathbb{Z}$. If $A = \{\bar{0}, \bar{1}, \ldots, \overline{a-1}\} \subset \mathbb{Z}/n\mathbb{Z}$ and $\mathcal{C} = \{A + t \mid t \in \mathbb{Z}/n\mathbb{Z}\}$, where $A + t = \{x + t \mid x \in A\}$, then $\alpha(\mathbb{Z}/n\mathbb{Z}, \mathcal{C}) = \lceil n/a \rceil$. In the other direction, we have the following proposition.

Proposition 3 *There is $a_0 \in \mathbb{N}$ such that if $n > a \ge a_0$ then, for some $A \subset F := \mathbb{Z}/n\mathbb{Z}$ with $|A| = a$, the family $\mathcal{C} = \{A + t \mid t \in \mathbb{Z}/n\mathbb{Z}\}$ is such that*

$$\alpha(F,\mathcal{C}) > k_0 = \left(1 - \frac{12\ln\ln a}{\ln a}\right)\frac{\ln a}{\ln\dfrac{1}{1-a/n}} > \left(1 - \frac{12\ln\ln a}{\ln a}\right)\left(\frac{n}{a} - 1\right)\ln a. \quad (2)$$

An interesting feature of Proposition 3 is that it claims the existence of uniform, regular hypergraphs with large α. More importantly, the parameters $n > a$ are free and (2) gives good estimates regardless of the relation between them. The reader is invited to compare the bounds in (1) and (2) for the cases in which *(i)* $a \sim \ln n$, *(ii)* $a \sim n/\ln n$, and *(iii)* $a \sim n/2$. In the course of answering a question of Tuza, Alon [1] obtained sharp bounds for case (ii), although, strictly speaking, the hypergraphs in [1] are not precisely a-uniform (the hyperedges have average cardinality a). We also observe that, in Proposition 3 above, for simplicity, we restrict ourselves to (a, b)-regular hypergraphs with $a = b$.

We now prove Proposition 3. Because of space limitations, we consider the case $a \geq n/\ln^3 n$ only (the other case requires an additional idea). We consider all the a-element subsets of $\mathbb{Z}/n\mathbb{Z}$, taken with equal probability. Let us estimate the probability that such a set A has k translations that cover F, where k is a given positive integer, i.e., let us estimate the probability $p(n, a, k)$ that there should exist $t_1, \dots, t_k \in \mathbb{Z}/n\mathbb{Z}$ such that $(A+t_1) \cup \cdots \cup (A+t_k) = \mathbb{Z}/n\mathbb{Z}$. Note that $1-p(n, a, k)$ is the probability that $(A+t_1) \cup \cdots \cup (A+t_k) \neq \mathbb{Z}/n\mathbb{Z}$ for any $t_1, \dots, t_k \in \mathbb{Z}/n\mathbb{Z}$, i.e., $(A^c+t_1) \cap \cdots \cap (A^c+t_k) \neq \emptyset$ for any $t_1, \dots, t_k \in \mathbb{Z}/n\mathbb{Z}$. Observe that, given $t_1, \dots, t_k \in \mathbb{Z}/n\mathbb{Z}$, we have $(A^c+t_1) \cap \cdots \cap (A^c+t_k) \neq \emptyset$ if and only if there is $x \in \mathbb{Z}/n\mathbb{Z}$ such that $\{x - t_1, \dots, x - t_k\} \subset A^c$.

Fix $T = \{t_1, \dots, t_k\} \subset \mathbb{Z}/n\mathbb{Z}$. We claim that, for $r = \lceil n/k^2 \rceil$, there are $x_1, \dots, x_r \in \mathbb{Z}/n\mathbb{Z}$ such that the sets $B_i = x_i - T = \{x_i - t_1, \dots, x_i - t_k\}$ $(1 \leq i \leq r)$ are pairwise disjoint. To prove our claim, take $x_1 = 0$, and suppose that we have chosen $x_1, \dots, x_s$, with $s < n/k^2$, such that $B_1, \dots, B_s$ are pairwise disjoint. Since $s < n/k^2$, clearly $U = \bigcup_{1 \leq i \leq s}(x_i - T)$ has fewer than n/k elements. Thus the average cardinality of the intersection of U with $x - T$, for $x \in \mathbb{Z}/n\mathbb{Z}$, is smaller than one, whence there is $x_{s+1} \in \mathbb{Z}/n\mathbb{Z}$ such that $x_{s+1} - T$ is disjoint from U. Our claim follows by induction. Let us now proceed with the proof of our proposition.

For any fixed set $T = \{t_1, \dots, t_k\}$, the probability that $x - T \subset A^c$ for some $x \in \mathbb{Z}/n\mathbb{Z}$ is at least the probability that $x_j - T \subset A^c$ for some $j \in [r] := \{1, \dots, r\}$, where the x_j $(1 \leq j \leq r)$ are fixed and are as given by our claim. This latter probability is $1 - \tilde{p}(n, a, k, r)$, where $\tilde{p}(n, a, k, r)$ is the probability that $x_j - T \not\subset A^c$ for each $j \in [r]$. To estimate $\tilde{p}(n, a, k, r)$, we consider random subsets $\tilde{A} \subset \mathbb{Z}/n\mathbb{Z}$ constructed as follows: let $y \in \tilde{A}$ with probability a/n, independently for all $y \in \mathbb{Z}/n\mathbb{Z}$. The probability that $\tilde{A}$ has m elements is $\binom{n}{m}(\frac{a}{n})^m(1 - \frac{a}{n})^{n-m}$, which is maximum for $m = a$, so the probability that $|\tilde{A}| = a$ is at least $1/n$. With this probability distribution, the events $B_j = x_j - T \not\subset \tilde{A}^c$ $(j \in [r])$ are independent (because the sets B_j are pairwise disjoint), and the probability of each of these events is $1 - (1 - a/n)^k$. Hence the probability that $B_j \not\subset \tilde{A}^c$ for all $j \in [r]$ is $(1 - (1 - a/n)^k)^r$. So we have

$$\begin{aligned} \tilde{p}(n, a, k, r) &= \mathbb{P}\left(\forall j \in [r] \text{ we have } x_j - T \not\subset \tilde{A}^c \,\middle|\, |\tilde{A}| = a\right) \\ &\leq n\,\mathbb{P}(\forall j \in [r] \text{ we have } x_j - T \not\subset \tilde{A}^c) = n\left(1 - \left(1 - \frac{a}{n}\right)^k\right)^r, \end{aligned}$$

and hence $p(n,a,k)$ is at most

$$\binom{n}{k}\tilde{p}(n,a,k,r) \leq \binom{n}{k} n \left(1-\left(1-\frac{a}{n}\right)^k\right)^r \leq n^{k+1}\left(1-\left(1-\frac{a}{n}\right)^k\right)^r. \quad (3)$$

Let $k = -\beta \ln a / \ln(1-a/n) \leq \beta(n/a)\ln n$, where $\beta \leq 1$. We have that the right-hand side of (3) is $n^{k+1}\left(1-a^{-\beta}\right)^r$, which is less than

$$n^{1+\beta\ln^4 n}\exp\left(-\frac{a^{-\beta}n}{k^2}\right) < \exp\left(-\frac{a^{2-\beta}}{\beta^2 n\ln^2 n}+\ln n\left(1+\ln^4 n\right)\right).$$

One may check that if $\beta \leq (1-12(\ln\ln a)/\ln a)$, then $a^{2-\beta}/\beta^2 n \ln^2 n \geq \ln^6 n$ for large n.

Since $\ln^6 n \gg \ln n(1+\ln^4 n)$, we have $p(n,a,k) \ll 1$. Therefore there is $A \subset \mathbb{Z}/n\mathbb{Z}$ with $|A| = a$ such that $(A+t_1)\cup\cdots\cup(A+t_k) \neq \mathbb{Z}/n\mathbb{Z}$ for any $t_1,\ldots,t_k \in \mathbb{Z}/n\mathbb{Z}$. This implies that, for some $A \subset \mathbb{Z}/n\mathbb{Z}$ with $|A| = a$ we have $\alpha(F,\mathcal{C}) > k_0$, where k_0 is as in (2).

References

[1] N. Alon: Transversal numbers of uniform hypergraphs, *Graphs Combin.* **6** (1990), 1–4.

[2] G. Cohen, I. Honkala, S. Litsyn and A. Lobstein: Covering Codes, *North-Holland*, Amsterdam, 1997.

[3] Z. Füredi: Matchings and covers in hypergraphs, *Graphs Combin.* **4** (1988), 115–206.

[4] Z. Füredi, G.J. Székely and Z. Zubor: On the lottery problem, *J. Combin. Des.* **4** (1996), 5–10.

[5] H. Hämäläinen, I. Honkala, S. Litsyn and P. Östergård: Football pools — a game for mathematicians, *Amer. Math. Monthly* **102** (1995), 579–588.

[6] D.S. Johnson: Approximation algorithms for combinatorial problems, *J. Comput. System Sciences* **9** (1974), 256–298.

[7] L. Lovász: On the ratio of optimal integral and fractional covers, *Discrete Math.* **13** (1975), 383–390.

[8] V. Rödl: On a packing and covering problem, *Europ. J. Combin.* **6** (1985), 69–78.

[9] S.K. Stein: Two combinatorial covering problems, *J. Comb. Th.*, Ser. A **16** (1974), 391–397.