A PROBABILISTIC VARIANT OF SPERNER'S THEOREM AND OF MAXIMAL r-COVER FREE FAMILIES

NOGA ALON, SHONI GILBOA, AND SHAY GUERON

ABSTRACT. A family of sets is called *r-cover free* if no set in the family is contained in the union of r (or less) other sets in the family. A 1-cover free family is simply an antichain with respect to set inclusion. Thus, Sperner's classical result determines the maximal cardinality of a 1-cover free family of subsets of an n-element set. Estimating the maximal cardinality of an r-cover free family of subsets of an n-element set for $r > 1$ was also studied. In this note we are interested in the following probabilistic variant of this problem. Let $S_0, S_1, \ldots, S_r$ be independent and identically distributed random subsets of an n-element set. Which distribution minimizes the probability that $S_0 \subseteq \bigcup_{i=1}^{r} S_i$? A natural candidate is the uniform distribution on an r-cover-free family of maximal cardinality. We show that for $r = 1$ such distribution is indeed best possible. In a complete contrast, we also show that for every $r > 1$ and n large enough, such distribution can be beaten by an exponential factor.

1. INTRODUCTION

For every positive integer n, let Ω_n be the set of all subsets of some fixed n-element set. For a positive integer r, a family $\mathcal{F} \subseteq \Omega_n$ is called *r-cover free* if no set in $\mathcal{F}$ is contained in the union of r (or less) other sets in $\mathcal{F}$. Let us denote by $g_r(n)$ the maximal cardinality of an r-cover free family in Ω_n. A 1-cover free family in Ω_n is just an antichain in Ω_n, with respect to set inclusion. Hence $g_1(n) = \binom{n}{\lfloor n/2 \rfloor}$, by the classical result of Sperner ([7]). For $r = 2$ it was shown in [2] that $1.134^n < g_2(n) < O(\sqrt{n}) \left(\frac{5}{4}\right)^n$ and in the subsequent paper [3], the same authors showed that for every r,

$$\left(1 + \frac{1}{4r^2}\right)^n < g_r(n) \leq \sum_{k=1}^{n} \frac{\binom{n}{\lceil k/r \rceil}}{\binom{k-1}{\lceil k/r \rceil - 1}}. \tag{1}$$

A different upper bound, which is better for large r, was obtained in [1]. In [6], this bound was given a simpler proof and the following, more explicit, form: for every $r \geq 2$ and n large enough,

$$g_r(n) \leq r^{8n/r^2}. \tag{2}$$

We will now describe a probabilistic variant of r-cover free families of maximal cardinality. Let $\mathcal{P}_n := \{p : \Omega_n \to [0, \infty) : \sum_{A \in \Omega_n} p(A) = 1\}$ be the family of probability distributions on Ω_n. For a positive integer r and $p \in \mathcal{P}_n$, let $\tau_r(p)$ be the probability that $S_0 \subseteq \bigcup_{i=1}^{r} S_i$, where $S_0, S_1, \ldots, S_r$ are random sets, drawn independently from Ω_n according to the distribution p. Natural candidates to minimize τ_r are distributions in the set

Date: June 1, 2020.
1991 *Mathematics Subject Classification.* 60C05.
Key words and phrases. cover free families, Sperner's theorem.

$\mathcal{CF}_{n,r} := \{p \in \mathcal{P}_n : p$ is supported on an r-cover free family$\}$ (in which case, one only has to worry about choosing the same set twice).

Clearly, $\min_{p\in\mathcal{CF}_{n,1}} \tau_1(p) = \frac{1}{\binom{n}{\lfloor n/2\rfloor}}$ where the minimum is attained for any distribution which is uniformly supported on a maximal antichain in Ω_n. Our first result is that for $n \geq 2$ this is indeed the minimum of τ_1 over all $\mathcal{P}_n$.

Theorem 1. *Suppose that $n \geq 2$. Then $\tau_1(p) \geq \frac{1}{\binom{n}{\lfloor n/2\rfloor}}$ for every $p \in \mathcal{P}_n$ and consequently, $\min_{p\in\mathcal{P}_n} \tau_1(p) = \min_{p\in\mathcal{CF}_{n,1}} \tau_1(p)$.*

We note that the weaker statement that $\Pr(S_0 \subseteq S_1 \text{ or } S_0 \supseteq S_1) \geq \frac{1}{\binom{n}{\lfloor n/2\rfloor}}$ for every independent identically distributed random sets S_0, S_1 in Ω_n, readily follows from the fact that Ω_n may be covered by $\binom{n}{\lfloor n/2\rfloor}$ chains (with respect to set inclusion). This symmetrized version of Theorem 1 may be generalized as follows. For a property P of families of sets, let $ex(n, P)$ denote the maximum possible cardinality of a family of sets in Ω_n satisfying P and let $ex(n, k, P)$, for $0 \leq k \leq n$, denote the maximum possible cardinality of a family of k-element sets in Ω_n satisfying P. Thus, for example, if P_1 is the property of being an antichain then $ex(n, P_1) = \binom{n}{\lfloor n/2\rfloor}$ by Sperner's Theorem, if P_2 is the property of being an intersecting family and $n \geq 2k$ then $ex(n, k, P_2) = \binom{n-1}{k-1}$ by the Erdős-Ko-Rado Theorem [4], and if P_3 is the property of not containing two sets whose symmetric difference has cardinality smaller than d, then $ex(n, P_3)$ is the maximum possible cardinality of an error correcting code with length n and minimum distance d. Similarly, $ex(n, k, P)$ is the maximum cardinality of the corresponding constant weight code.

Theorem 2. *Let $\mathcal{H}$ be a family of unordered pairs of distinct sets in Ω_n and let $P_{\mathcal{H}}$ be the property of containing no pair from $\mathcal{H}$. For $p \in \mathcal{P}_n$, let $\tau_{\mathcal{H}}(p) := \Pr(\{S_0, S_1\} \in \mathcal{H} \text{ or } S_0 = S_1)$, where S_0, S_1 are random sets, drawn independently from Ω_n according to the distribution p. Then $\min_{p\in\mathcal{P}_n} \tau_{\mathcal{H}}(p) = \frac{1}{ex(n,P_{\mathcal{H}})}$. Similarly, for every $0 \leq k \leq n$, the minimum of $\tau_{\mathcal{H}}(p)$ over distributions $\mathcal{P}_n$ whose support is a subset of $\{A \in \Omega_n : |A| = k\}$ is $\frac{1}{ex(n,k,P_{\mathcal{H}})}$.*

The examples mentioned above provide several specific applications of the theorem, and it is not difficult to describe others.

In a complete contrast to Theorem 1, we show that for every $r > 1$ (and n large enough), the minimum of τ_r on $\mathcal{P}_n$ is much smaller than the minimum of τ_r over $\mathcal{CF}_{n,r}$. For every $0 \leq \ell \leq n$, let p_ℓ be the probability distribution in $\mathcal{P}_n$ uniformly supported on the family of all ℓ-element sets in Ω_n.

Theorem 3. *Suppose that $r \geq 2$. There is $0 < \mu_r < 1$ such that for every n large enough, $\min_{0<\ell<\frac{n}{r}} \tau_r(p_\ell) < \mu_r^n \min_{p\in\mathcal{CF}_{n,r}} \tau_r(p)$ and consequently, $\min_{p\in\mathcal{P}_n} \tau_r(p) < \mu_r^n \min_{p\in\mathcal{CF}_{n,r}} \tau_r(p)$.*

For every $r \geq 2$, Theorem 3 shows that $\min_{p\in\mathcal{P}_n} \tau_r(p)$ is (much) smaller than $\min_{p\in\mathcal{CF}_{n,r}} \tau_r(p)$, which is at most $1 - \left(1 - \frac{1}{g_r(n)}\right)^r < \frac{r}{g_r(n)}$, as shown by considering any probability distribution uniformly supported on an r-cover free family of maximal cardinality. A lower bound for $\min_{p\in\mathcal{P}_n} \tau_r(p)$ is given in the following theorem.

Theorem 4. *Suppose that $r \geq 2$. There is $C_r > 0$ such that $\min_{p\in\mathcal{P}_n} \tau_r(p) \geq \frac{C_r}{(g_r(n))^r}$ and hence, for n large enough, by (2), $\min_{p\in\mathcal{P}_n} \tau_r(p) \geq \frac{C_r}{r^{8n/r}}$.*

We prove Theorems 1 and 2 in Section 2 and Theorems 3 and 4 in Section 3.

2. The case $r = 1$

Proof of Theorem 1. Let $p \in \mathcal{P}_n$. Let $\mathcal{C}$ be the set of all maximal chains in Ω_n, with respect to set inclusion. Every $A \in \Omega_n$ belongs to exactly $\frac{|\mathcal{C}|}{\binom{n}{|A|}}$ maximal chains. Therefore, $\frac{1}{|\mathcal{C}|}\sum_{C\in\mathcal{C}}\sum_{A\in C}\binom{n}{|A|}p(A) = \sum_{A\in\Omega_n} p(A) = 1$ and since $\binom{n}{k} \le \binom{n}{\lfloor n/2\rfloor}$ for every $0 \le k \le n$,

$$\binom{n}{\lfloor \frac{n}{2}\rfloor}\sum_{A\in\Omega_n} p(A)^2 \ge \sum_{A\in\Omega_n}\binom{n}{|A|}p(A)^2 = \frac{1}{|\mathcal{C}|}\sum_{C\in\mathcal{C}}\sum_{A\in C}\binom{n}{|A|}^2 p(A)^2. \tag{3}$$

Similarly, every pair $A_0 \subsetneq A_1$ of sets in Ω_n belong to exactly $\frac{|\mathcal{C}|}{\binom{n}{|A_1|}\binom{|A_1|}{|A_0|}}$ maximal chains. Therefore, since $\frac{\binom{n}{k}}{\binom{\ell}{k}} \le \frac{1}{2}\binom{n}{\lfloor \frac{n}{2}\rfloor}$ for every $0 \le k < \ell \le n$,

$$\begin{aligned}\binom{n}{\lfloor \frac{n}{2}\rfloor}\sum_{\substack{(A_0,A_1)\in\Omega_n^2\\ A_0\subsetneq A_1}} p(A_0)p(A_1) &\ge 2\sum_{\substack{(A_0,A_1)\in\Omega_n^2\\ A_0\subsetneq A_1}} \frac{\binom{n}{|A_0|}}{\binom{|A_1|}{|A_0|}}p(A_0)p(A_1)\\ &= \frac{1}{|\mathcal{C}|}\sum_{C\in\mathcal{C}}\sum_{\substack{(A_0,A_1)\in C^2\\ A_0\neq A_1}}\binom{n}{|A_0|}\binom{n}{|A_1|}p(A_0)p(A_1).\end{aligned} \tag{4}$$

Summing up (3) and (4) yields

$$\begin{aligned}\binom{n}{\lfloor \frac{n}{2}\rfloor}\tau_1(p) = \binom{n}{\lfloor \frac{n}{2}\rfloor}\sum_{\substack{(A_0,A_1)\in\Omega_n^2\\ A_0\subseteq A_1}} p(A_0)p(A_1) &\ge \frac{1}{|\mathcal{C}|}\sum_{C\in\mathcal{C}}\sum_{(A_0,A_1)\in C^2}\binom{n}{|A_0|}\binom{n}{|A_1|}p(A_0)p(A_1)\\ = \frac{1}{|\mathcal{C}|}\sum_{C\in\mathcal{C}}\left(\sum_{A\in C}\binom{n}{|A|}p(A)\right)^2 &\ge \left(\frac{1}{|\mathcal{C}|}\sum_{C\in\mathcal{C}}\sum_{A\in C}\binom{n}{|A|}p(A)\right)^2 = 1.\end{aligned}$$

as claimed. ☐

Proof of Theorem 2. Let G be the complement of the graph $(\Omega_n, \mathcal{H})$. The size of the maximum clique in G is clearly $ex(n, P_{\mathcal{H}})$. Therefore, by a theorem of Motzkin and Straus [5, Theorem 1],

$$\min_{p\in\mathcal{P}_n}\tau_{\mathcal{H}}(p_{\mathcal{H}}) = 1 - 2\max_{p\in\mathcal{P}_n}\sum_{\{A_0,A_1\}\text{ is an edge of }G} p(A_0)p(A_1) = \frac{1}{ex(n, P_{\mathcal{H}})}.$$

The second statement follows similarly, by considering the graph induced by G on the vertex set $\{A \in \Omega_n : |A| = k\}$. ☐

3. The case $r > 1$

Note that if $p \in \mathcal{P}_n$ is supported on an r-cover free family $\mathcal{F}$, then

$$1 - \tau_r(p) = \sum_{F\in\mathcal{F}} p(F)\left(1 - p(F)\right)^r \le \sum_{F\in\mathcal{F}} p(F)\left(1 - p(F)\right) \le 1 - \frac{1}{|\mathcal{F}|},$$

and hence $\min_{p\in\mathcal{CF}_{n,r}}\tau_r(p)\geq\frac{1}{g_r(n)}$. Therefore, to prove Theorem 3 for some $r\geq 2$, it is enough to show that there is $0<\mu_r<1$ such that for n large enough,

$$\min_{0<\ell<\frac{n}{r}}\tau_r(p_\ell)<\mu_r^n\frac{1}{g_r(n)}. \tag{5}$$

For large r this may be easily deduced as follows. For $\ell:=\lfloor\frac{n}{er}\rfloor$, clearly

$$\tau_r(p_\ell)\leq\frac{\binom{r\ell}{\ell}}{\binom{n}{\ell}}\leq\left(\frac{r\ell}{n}\right)^\ell\leq\frac{1}{e^\ell}<e\frac{1}{e^{\frac{n}{er}}}=e\left(e^{-\frac{1}{e}}r^{\frac{8}{r}}\right)^{\frac{n}{r}}\frac{1}{r^{\frac{8n}{r^2}}}.$$

Therefore, by (2), for n large enough

$$\min_{0<\ell<\frac{n}{r}}\tau_r(p_\ell)<e\left(e^{-\frac{1}{e}}r^{\frac{8}{r}}\right)^{\frac{n}{r}}\frac{1}{r^{\frac{8n}{r^2}}}<e\left(e^{-\frac{1}{e}}r^{\frac{8}{r}}\right)^{\frac{n}{r}}\frac{1}{g_r(n)}. \tag{6}$$

It can be verified that $e^{-\frac{1}{e}}r^{\frac{8}{r}}<1$ for every $r\geq 101$. Thus, (6) confirms (5), and hence Theorem 3, for $r\geq 101$. We proceed to describe the proof Theorem 3 for general $r\geq 2$.

Proof of Theorem 3. Let ℓ be an integer in the interval $[0,\frac{n}{r})$ for which $\binom{n}{\ell+1}/\binom{r\ell}{\ell}$ is maximal. It is simple to verify that if n is large enough, then the sequence $\left(\binom{n}{j+1}/\binom{rj}{j}\right)_{j=0}^{\lfloor n/4r\rfloor+1}$ is increasing and hence $\ell>\frac{n}{4r}$.

Let $S_0,S_1,\ldots,S_r$ be random sets chosen, independently and uniformly, from all the ℓ-element sets in Ω_n.

Let $t:=\lfloor\ell^2/n\rfloor$ and let $\mathcal{E}$ be the event: $|\bigcup_{i=1}^r S_i|>r\ell-t$. It is easy to verify that the sequence $(\Pr(S_1\cup S_2=k))_{k=2\ell-t}^{2\ell}$ is decreasing, and hence

$$\Pr(\mathcal{E})\leq\Pr(|S_1\cup S_2|>2\ell-t)\leq t\Pr(|S_1\cup S_2|=2\ell-t)=t\frac{\binom{n-\ell}{\ell-t}\binom{\ell}{t}}{\binom{n}{\ell}}.$$

Therefore, by (1),

$$\begin{aligned}\tau_r(p_\ell)&=\Pr\left(S_0\subseteq\bigcup_{i=1}^r S_i\right)\\&=\Pr(\mathcal{E})\Pr\left(S_0\subseteq\bigcup_{i=1}^r S_i\,\middle|\,\mathcal{E}\right)+\Pr(\Omega_n\setminus\mathcal{E})\Pr\left(S_0\subseteq\bigcup_{i=1}^r S_i\,\middle|\,\Omega_n\setminus\mathcal{E}\right)\\&\leq t\frac{\binom{n-\ell}{\ell-t}\binom{\ell}{t}}{\binom{n}{\ell}}\cdot\frac{\binom{r\ell}{\ell}}{\binom{n}{\ell}}+1\cdot\frac{\binom{r\ell-t}{\ell}}{\binom{n}{\ell}}=\left(t\frac{\binom{n-\ell}{\ell-t}\binom{\ell}{t}}{\binom{n}{\ell}}+\frac{\binom{r\ell-t}{\ell}}{\binom{r\ell}{\ell}}\right)\frac{n-\ell}{\ell+1}\cdot\frac{\binom{r\ell}{\ell}}{\binom{n}{\ell+1}}\\&\leq\left(t\frac{\binom{n-\ell}{\ell-t}\binom{\ell}{t}}{\binom{n}{\ell}}+\frac{\binom{r\ell-t}{\ell}}{\binom{r\ell}{\ell}}\right)\frac{(n-\ell)n}{\ell+1}\cdot\frac{1}{g_r(n)},\end{aligned}$$

and (5) follows by using standard estimates on binomial coefficients. This completes the proof of the theorem. □

Finally, we prove Theorem 4.

Proof of Theorem 4. Let $p \in \mathcal{P}_n$, let $N := 2g_r(n)$, let $S_1, \ldots, S_N$ be random sets, drawn independently from Ω_n according to the distribution p, and consider the random variable

$$I := \{i \in [N] : \text{there is } J \subset [N] \setminus \{i\} \text{ of cardinality } r \text{ such that } S_i \subseteq \bigcup_{j \in J} S_j\}.$$

The family $\{S_i\}_{i \in [N] \setminus I}$ is clearly r-cover free, therefore $N - |I| = |[N] \setminus I| \leq g_r(n)$ and hence $\mathbb{E}|I| \geq N - g_r(n) = g_r(n)$. On the other hand, clearly $\mathbb{E}|I| \leq N\binom{N-1}{r}\tau_r(p)$. Hence

$$\tau_r(p) \geq \frac{g_r(n)}{N\binom{N-1}{r}} \geq \frac{r!\, g_r(n)}{N^{r+1}} = \frac{r!}{2^{r+1} g_r(n)^r}$$

and the result follows. □

Acknowledgements. This research was partially supported by NSF grant DMS-1855464, BSF grant 2018267, the Simons Foundation, the Israel Science Foundation (grant No. 1018/16), The Center for Cyber Law and Policy at the University of Haifa, and a Google Research Award.

References

[1] A. G. Dyachkov and V. V. Rykov, Bounds on the length of disjunctive codes, Problemy Peredachi Informatsii 18, No. 3 (1982), 7-13.

[2] P. Erdős, P. Frankl and Z. Füredi, Families of finite sets in which no set is covered by the union of two others, J. Combin. Theory Ser. A **33** (1982), no. 2, 158–166.

[3] P. Erdős, P. Frankl and Z. Füredi, Families of finite sets in which no set is covered by the union of r others, Israel J. Math. **51** (1985), no. 1–2, 79–89.

[4] P. Erdős, C. Ko and R. Rado, Intersection theorems for systems of finite sets, Quart. J. Math. Oxford Ser. (2) **12** (1961), 313–320.

[5] T. S. Motzkin and E. G. Straus, Maxima for graphs and a new proof of a theorem of Turán, Canadian J. Math. 17 (1965), 533–540.

[6] M. Ruszinkó, On the upper bound of the size of the r-cover-free families, J. Combin. Theory Ser. A **66** (1994), no. 2, 302–310.

[7] E. Sperner, Ein Satz über Untermengen einer endlichen Menge, Math. Z. **27** (1928), no. 1, 544–548.

Princeton University, NJ 08544, USA and Tel Aviv University, Tel Aviv 69978, Israel

The Open University of Israel, Raanana 43107, Israel.

University of Haifa, Haifa 31905, Israel, and Amazon Web Services, USA.