COVERING THE HYPERCUBE, THE UNCERTAINTY PRINCIPLE, AND AN INTERPOLATION FORMULA

PAATA IVANISVILI, OHAD KLEIN, AND ROMAN VERSHYNIN

Abstract. We show that the minimal number of skewed hyperplanes that cover the hypercube $\{0,1\}^n$ is at least $\frac{n}{2}+1$, and there are infinitely many n's when the hypercube can be covered with $n-\log_2(n)+1$ skewed hyperplanes. The minimal covering problems are closely related to uncertainty principle on the hypercube, where we also obtain an interpolation formula for multilinear polynomials on $\mathbb{R}^n$ of degree less than $\lfloor n/m \rfloor$ by showing that its coefficients corresponding to the largest monomials can be represented as a linear combination of values of the polynomial over the points $\{0,1\}^n$ whose hamming weights are divisible by m.

2020 Mathematics Subject Classification. Primary: 06E30; Secondary: 42C10, 68Q32.

Key words. Discrete hypercube, Uncertainty principle, minimal covering, Littlewood–Offord

1. Introduction

1.1. **Covering the hypercube.** How many affine hyperplanes are needed to cover the hypercube $\{-1,1\}^n$? Notice that two affine hyperplanes $x_1=-1$ and $x_1=1$ cover the hypercube, and clearly this is the minimal number. However, if one requires that the affine hyperplanes are skewed, i.e., $a_1x_1+\ldots+a_nx_n+b=0$ with all $a_1,\ldots,a_n\neq 0$, then the problem becomes challenging[1].

It follows from Littlewood–Offord inequalities that any skewed hyperplane covers at most $n^{-1/2}$ fraction of points in $\{-1,1\}^n$ (up to a universal constant factor), therefore, one needs at least $\Omega(n^{1/2})$ skewed hyperplanes to cover the hypercube. In [7], this lower bound was improved to $\Omega(n^{0.51})$, and recently [4] to $\Omega(n^{2/3}\log(n)^{-4/3})$ by the second named author of the present paper.

The family of $n+1$ hyperplanes, $x_1+\ldots+x_n=2k-n$ for all $k=0,\ldots n$, covers the hypercube. In fact, if n is even one can cover with n skewed hyperplanes just by replacing the two hyperplanes corresponding to $k=0$ and $k=n$ in the previous example with one hyperplane $x_1+\ldots+x_{n/2}-x_{n/2-1}-\ldots-x_n=0$. Moreover, it follows from [1] that for even n, the upper bound n on the minimal cover is also a lower bound if one restricts the covering to the family of "regular" hyperplanes, i.e., the ones $\varepsilon_1x_1+\ldots+\varepsilon_nx_n+b=0$ with $\varepsilon_j=\pm1$ for all $j=1,\ldots,n$.

Looking at the results for the case of "regular" hyperplane cover in [1], one may suspect that in analogy to Littlewood–Offord problem the sharp lower bound on the minimal skew hyperplane cover should be n. Surprisingly, one can cover the hypercube $\{-1,1\}^5$ with the following 4 skewed hyperplanes

$$\begin{aligned}
x_1+x_2+x_3+x_4+2x_5&=0,\\
x_1+x_2+x_3-x_4+2x_5&=0,\\
x_1+x_2+x_3+x_4-2x_5&=0,\\
x_1+x_2+x_3-x_4-2x_5&=0.
\end{aligned}$$

Also the hypercube $\{-1,1\}^6$ can be covered with 5 skewed hyperplanes (see Section 2.3). In fact, one can cover the hypercube $\{-1,1\}^n$ with $n-\log_2(n)+1$ skewed hyperplanes for infinitely many n's.

[1] In what follows we will omit the word affine and we will be referring to such hyperplanes as skewed hyperplanes.

Proposition 1. *For any integer $m \geq 1$ the hypercube $\{-1,1\}^{2^m+m-1}$ can be covered with 2^m skewed hyperplanes.*

This proposition shows that the skew hyperplane covering problem is genuinely different from the original "regular" problem solved in [1].

Question 2. *What is the minimal number of skewed hyperplane cover of the hypercube $\{-1,1\}^n$?*

We prove the following lower bound.

Theorem 3. *The minimal number of skewed hyperplane cover of $\{-1,1\}^n$ is at least $\frac{n}{2}+1$.*

There is a close relation between minimal hyperplane covering problem and the uncertainty principle on the hypercube. Let $p(x)$ be a polynomial on $\mathbb{R}^n$, and let $\mathrm{supp}(p) = \{x \in \mathbb{R}^n : p(x) \neq 0\}$. Under what conditions on $\mathrm{supp}(p) \cap \{-1,1\}^n$ and $\deg(p)$ does it follows that $p(x) \equiv 0$ on $\{-1,1\}^n$?

It turns out that the support of a nonzero low degree polynomial cannot be contained in a skewed hyperplane:

Theorem 4 (Linial–Radhakrishnan [5])**.** *If $\deg(p) < \frac{n}{2}$ and $\mathrm{supp}(p) \cap \{-1,1\}^n$ belongs to a skewed hyperplane then $p(x) \equiv 0$ on $\{-1,1\}^n$.*

Observe that Theorem 3 follows from Theorem 4. Indeed, let $H_1, \ldots, H_{k+1}$ be a minimal skew hyperplane cover of $\{-1,1\}^n$. If H_j's are given via equations $\ell_j(x) = a_{1j}x_1 + \ldots + a_{nj}x_n + b_j = 0$, for all $j = 1, \ldots, k+1$, then it follows that $p(x)\ell_{k+1}(x) \equiv 0$ on $\{-1,1\}^n$,where $p(x) = \prod_{j=1}^{k} \ell_j(x)$ is not identically zero polynomial on $\{-1,1\}^n$ of degree at most k. Hence $\mathrm{supp}(p) \cap \{-1,1\}^n$ belongs to H_{k+1} and Theorem 4 implies that $k \geq n/2$.

After the current paper was completed, independently and concurrently the paper [6] appeared on arXiv where Theorem 3 was derived from Theorem 4 proved in [5] (see Lemma 2 in [5]). The proof of Theorem 4 in [5] in turn is based on either Combinatorial Nullstellensatz or the spectral properties of the Johnson graph (the authors [5] attribute the nonsingularity of the Johnson graph to [3]). Our proof of Theorem 4, given in Section 2.1, is simple and self-contained.

1.2. **An interpolation formula.** In [1] sharp lower bound n on the minimal number of "regular" hyperplane cover of the n dimensional hypercube (for even n) was based on the following technical observation: if a multilinear polynomial $p(x)$ vanishes on all those points of $\{-1,1\}^n$ which have even number of 1's in its coordinates, and $\deg(p) < n/2$ then p is identically zero (see Lemma 2.1 in [1]). This observation suggests that perhaps the coefficients of the multilinear polynomials of small degree can be reconstructed by its values at sparse points of $\{-1,1\}^n$. The goal of this section is to obtain such an interpolation formula.

Recall that any function $f : \{-1,1\}^n \mapsto X$, where X is a normed space, has Fourier–Walsh representation

$$f(x) = \sum_{S \subset \{1,\ldots,n\}} \widehat{f}(S) x^S,$$

for some $\widehat{f}(S) \in X$, where $x = (x_1, \ldots, x_n)$, $x^S = \prod_{j \in S} x_j$ and $x^{\emptyset} = 1$. We say that f has degree $\deg(f)$ if $\widehat{f}(S) = 0$ for all $S \subset \{1, \ldots, n\}$ with $|S| > \deg(f)$, and there exists a subset S of cardinality $\deg(f)$ such that $\widehat{f}(S) \neq 0$.

Definition 5. *For any integer $m > 1$ the symbol $W(m)$ denotes the subset of $\{-1,1\}^n$ consisting of all points $x = (x_1, \ldots, x_n) \in \{-1,1\}^n$ such that $\#\{j : x_j = -1\}$ is divisible by m.*

For any integer $m > 1$ let $k(m)$ be the smallest positive integer greater than one that divides m. In this section we obtain the following interpolation formula

Theorem 6. *Let $f : \{-1,1\}^n \mapsto X$ and let $m > 1$ be an integer such that*

$$\deg(f) + 1 \le \frac{n}{m} + \frac{1}{k(m)}. \tag{1}$$

Then for any $\widehat{f}(S)$ with $|S| = d$, there exists a probability measure supported on $W(m)$ such that

$$\widehat{f}(S) = \int_{W(m)} h(x) f(x) d\mu(x)$$

for some $h : W(m) \mapsto \{-1,1\}$. Both the measure $d\mu$ and h depend only on $S, m, \deg(f), n$.

The next corollary follows from the theorem

Corollary 7. *If $f : \{-1,1\}^n \mapsto X$ vanishes on a set $W(m)$ for some integer m satisfying (1) then $f \equiv 0$.*

Remark 8. *Notice that condition (1) always holds if $1 < m \le \frac{n+1}{\deg(f)+1}$. If m is even then (1) is the same as $\deg(f) \le \frac{n}{m} - \frac{1}{2}$. In particular, if $m = 2$ Corollary 7 recovers a result of Lemma 2.1 in* [1].

Remark 9. *In the proof of Theorem 6 both the measure $d\mu$ and $h(x)$ will be constructed explicitly.*

Notice that since $\widehat{f}(S) = \mathbb{E} f(x) x^S$ then $|\widehat{f}(S)| \le \max_{x \in \{-1,1\}^n} |f(x)|$. However, if f has low degree then $\max_{x \in \{-1,1\}^n} |f(x)|$ can be replaced by a maximum over sparse family of points of $\{-1,1\}^n$ provided that $|S| = \deg(f)$. Indeed, Theorem 6 gives the following

Corollary 10. *If the degree of $f : \{-1,1\}^n \mapsto X$, where X is a normed space, satisfies (1), then*

$$\|\widehat{f}(S)\| \le \max_{x \in W(m)} \|f(x)\|$$

for all $S \subset \{,\dots,n\}$ with $|S| = \deg(f)$.

2. Proofs

2.1. **The proof of Theorem 4.** Denote $[n] := \{1,\dots,n\}$. Every polynomial $p(x)$ of degree d, when restricted to $\{-1,1\}^n$, can be written as $f(x) = \sum_{|S| \le d} b_S x^S$ for some $b_S \in \mathbb{R}$. The assumptions in Theorem 4 imply that

$$(a_1 x_1 + \ldots + a_n x_n + b) \sum_{|S| \le d} b_S x^S = 0 \quad \text{on} \quad \{-1,1\}^n.$$

After expanding this expression we see that the coefficients in front of all monomials x^T must vanish, and in particular this must happen for any degree $d+1$ monomial, i.e., x^S with $|S| = d+1$. This means that

$$\sum_{j \in W} a_j b_{W \setminus j} = 0 \tag{2}$$

for all $W \subset [n]$ with $|W| = d+1$. We can view (2) as a system of linear equations $Ab = 0$ in $b = (b_S : S \subset [n], |S| = d)$. It suffices to prove

Lemma 11. *We have* $\mathrm{Ker}(A) = 0$ *as long as long as $n \ge 2d+1$.*

Proof. We will prove the lemma by induction on d. For the base case $d = 1$ and $n \ge 3$ let us write the equation (2) corresponding to the sets $\{2,3\}$, $\{1,3\}$ and $\{1,2\}$:

$$a_3 b_2 + a_2 b_3 = 0$$
$$a_1 b_3 + a_3 b_1 = 0$$
$$a_1 b_2 + a_2 b_1 = 0.$$

In matrix form, this becomes

$$\begin{bmatrix} 0 & a_3 & a_2 \\ a_3 & 0 & a_1 \\ a_2 & a_1 & 0 \end{bmatrix} \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix} = 0.$$

The determinant of this 3×3 matrix equals $2a_1a_2a_3$, which is nonzero by assumption.

Now assume that the lemma holds for $d-1$, and let us prove it for d. Let $n \geq 2d+1$. Then $n-2 \geq 2(d-1)+1$. Thus, by the induction hypothesis, the following system is well posed:

$$a_{i_1}\bar{b}_{i_2,i_3,\ldots,i_d} + a_{i_2}\bar{b}_{i_1,i_3,\ldots,i_d} + \cdots + a_{i_d}\bar{b}_{i_1,i_2,\ldots,i_{d-1}} = 0 \quad \forall\{i_1,\ldots,i_d\} \subset [n-2]. \tag{3}$$

Assume that the system of equations (2) holds for some b. To complete the induction step, it suffices to show that $b = 0$.

Fix $j := i_{d+1} \in \{n-1, n\}$, and let us look only at those equations in (2) where $\{i_1,\ldots,i_d\} \subset [n-2]$. Moving the last term into the right hand side, we can rewrite those equations as

$$a_{i_1}b_{i_2,i_3,\ldots,i_d,j} + a_{i_2}b_{i_1,i_3,\ldots,i_d,j} + \cdots + a_{i_d}b_{i_1,i_2,\ldots,i_{d-1},j} = -a_j b_{i_1,i_2,\ldots,i_d} \quad \forall\{i_1,\ldots,i_d\} \subset [n-2].$$

Compare this system with the one in (3). The variables are called differently and the right hand side is not zero, but the matrix is the same and thus the system is well posed. So if we express it as

$$A\bar{b} = -a_j\tilde{b},$$

the matrix A is invertible and thus

$$\bar{b} = -a_j A^+\tilde{b},$$

where A^+ denotes the pseudoinverse of A.

Next, let $j \in \{n-1, n\}$ vary while keeping $\{i_1,\ldots,i_d\}$ fixed. Note that neither A nor $\tilde{b}$ depend on j. Hence $A^+\tilde{b}$ does not depend on j either. Thus $\bar{b}$ must be proportional to a_j. Thus all coordinates of $\bar{b}$ must be proportional to a_j, so in particular

$$b_{i_1,i_2,\ldots,i_{d-1},j} = a_j \cdot c_{i_1,i_2,\ldots,i_d} \quad \forall j \in \{n-1, n\},$$

where $c_{i_1,i_2,\ldots,i_d}$ is determined by $\{i_1, i_2, \ldots, i_d\}$ and may not depend on j. Dividing the two equations by each other, we arrive at the following conclusion.

For any distinct indices $i_1, i_2, \ldots, i_{d-1} \in [n-2]$ and distinct indices $j, j' \in \{n-1, n\}$, either

$$b_{i_1,i_2,\ldots,i_{d-1},j} = b_{i_1,i_2,\ldots,i_{d-1},j'} = 0$$

or

$$\frac{b_{i_1,i_2,\ldots,i_{d-1},j}}{b_{i_1,i_2,\ldots,i_{d-1},j'}} = \frac{a_j}{a_{j'}}.$$

By relabeling the indices, the same conclusion holds for any distinct indices $i_1, i_2, \ldots, i_{d-1}, j, j' \in [n]$.

Since all $a_i \neq 0$, there exists $c \in \mathbb{R}$ such that

$$b_{1,\ldots,d} = ca_1 \cdots a_d.$$

Fix any $\{i_1,\ldots,i_d\} \subset [n]$. By changing one index at a time using the fact above (first $1 \to i_1$, next $2 \to i_2$, etc.) the corresponding factors in this product change (first $a_1 \to a_{i_1}$, next $a_2 \to a_{i_2}$, etc.) After d swaps, we conclude that

$$b_{i_1,\ldots,i_d} = ca_{i_1} \cdots a_{i_d} \quad \forall\{i_1,\ldots,i_d\} \subset [n].$$

Plugging this into (2), we get

$$(d+1) \cdot ca_{i_1} \cdots a_{i_{d+1}} = 0.$$

Since all $a_i \neq 0$, it follows that $c = 0$, so all $b_{i_1,\ldots,i_d} = 0$. The induction step is proved, and the theorem is established.

□

2.2. **The proof of Theorem 6.** Let $d = \deg(f)$. Let $\widehat{f}(S)$ be such that $S = \{m, 2m, ...dm\}$. The general case of S with $|S| = d$ will be discussed at the end of the proof. We remind that for any $u = (u_1, \dots, u_n) \in \{-1,1\}^n$ we have

$$f(u) = \sum_{|S| \le d} \widehat{f}(S) u^S, \quad u^S = \prod_{j \in S} u_j.$$

Consider

$$\begin{aligned} g_y(x) := f(&y_1 x_1, y_1 x_2, \dots, y_1 x_m, \; y_2 x_{m+1}, y_2 x_{m+2}, \dots, y_2 x_{2m}, \; \dots, \\ &y_d x_{(d-1)m+1}, y_d x_{(d-1)m+2}, \dots, y_d x_{md}, \\ &x_{md+1}, x_{md+2}, \dots, x_{md+m-\frac{m}{k(m)}}, \dots, x_n), \end{aligned}$$

where $(y_1, \dots, y_d) \in \{-1,1\}^d$, and $(x_1, \dots, x_n) \in \{-1,1\}^n$. In other words, we take a long vector

$$x = (x_1, \dots, x_d, \dots, x_{2d}, \dots, x_{md}, \dots, x_{md+m-\frac{m}{k(m)}}, \dots, x_n)$$

and we split it into disjoint groups $G_1, G_2, \dots, G_d, G_{\text{extra}}, G_{\text{rest}}$, where

$$\begin{aligned} &G_1 = (x_1, \dots, x_m), \\ &G_2 = (x_{m+1}, \dots, x_{2m}), \\ &\dots \\ &G_d = (x_{(d-1)m+1}, \dots, x_{md}), \\ &G_{\text{extra}} = (x_{md+1}, \dots, x_{md+m-\frac{m}{k(m)}}), \\ &G_{\text{rest}} = (x_{md+m-\frac{m}{k(m)}+1}, \dots, x_n). \end{aligned}$$

The group G_{rest} can be empty (for example if we have equality in (1)). Each group G_j is multiplied by the variable y_j (termwise) for each $j = 1, \dots, d$. The variables in groups G_{extra} and G_{rest} are unchanged. We will see that the group G_{rest} will not be important to us, but the group G_{extra} will be helpful to us, it will play some kind of a "counter" role in the proof.

It will be helpful to think about $g_y(x)$ as follows: if we define

$$y_j G_j = (y_j x_{(j-1)m+1}, y_j x_{(j-1)m+2}, \dots, y_j x_{jm}),$$

then

$$g_y(x) = f(y_1 G_1, \dots, y_d G_d, G_{\text{extra}}, G_{\text{rest}}).$$

Next we will construct a distribution $\mathcal{D}$ on $\{-1,1\}^n$ such that

$$\widehat{f}(S) = \mathbb{E}_{x \sim \mathcal{D}} \, \mathbb{E}_{y \sim \text{unif}(\{-1,1\}^d)} \, g_y(x) y_1 \dots y_d,$$

moreover, we will have

$$(y_1 G_1, \dots, y_d G_d, G_{\text{extra}}, G_{\text{rest}}) \in W(m)$$

for all $y \in \{-1,1\}^d$, and any x from the support of the distribution $\mathcal{D}$. This will finish the proof of the theorem.

Let g_1 be indexes of the variables in G_1, i.e., $g_1 = \{1, \dots, d\}$. Similarly, g_j denotes the indexes of the variables in G_j for all $j = 1, \dots, d$. It is not difficult to see that for any x we have

$$\mathbb{E}_{y \sim \text{unif}(\{-1,1\}^d)} \, g_y(x) y_1 \dots y_d = \sum_{|T| = d, |T \cap g_j| = 1, \, j = 1, \dots, d} \widehat{f}(T) x^T. \tag{4}$$

Indeed, let us pick a monomial in front of $\widehat{f}(T)$. If $|T| < d$ then after multiplying the monomial by the factor $y_1 \dots, y_d$ and taking the expectation in y we get zero. If $|T| = d$, and T does not have the property that $|T \cap g_j| = 1$ for all $j = 1, \dots, d$, then the corresponding monomial will be missing at least one of the variables among $y_1 \dots, y_d$, so after multiplying by $y_1 \dots, y_d$ and taking the expectation in y we get zero.

Next, we will describe the distribution $\mathcal{D}$ for the random variable $x = (G_1, \dots, G_d, G_{\text{extra}}, G_{\text{rest}})$. Let $\ell = \frac{m}{k(m)}$, and let $\alpha = \frac{1}{2\binom{m-2}{\ell-1}}$. Set $\mathbb{P}(G_1 = (1,1,\dots,1)) = 1 - \binom{m-1}{\ell}\alpha$, and G_1 takes values $(x_1, \dots, x_{m-1}, 1)$ having exactly ℓ minus ones in it with equal probabilities α. Let $G_2, \dots, G_d$ be independent copies of G_1. Set $G_{\text{rest}} = (1,1,\dots,1)$.

To define the values for G_{extra} let us count the number of -1's in the list $(y_1 G_1, \dots, y_d G_d)$. If $y_1 = 1$ then G_1 has exactly ℓ number of -1's. If $y_1 = -1$ then G_1 has exactly $m-\ell$ number of -1's. So the list $(y_1 G_1, \dots, y_d G_d)$ has $n_1\ell + n_2(m-\ell)$ number of -1's for some nonnegative integers $n_1, n_2 \geq 0$. Our goal is to assign c number of -1's in the group G_{extra} so that $n_1\ell + n_2(m-\ell) + c = 0 \operatorname{mod}(m)$.

Since the size of the group G_{rest} is $m - \frac{m}{k(m)}$ we can choose any integer c satisfying $0 \leq c \leq m\left(\frac{k(m)-1}{k(m)}\right)$. We have

$$n_1\ell + n_2(m-\ell) + c = m\left(\frac{n_1 + n_2(k(m)-1)}{k(m)}\right) + c.$$

If $n_1 + n_2(k(m)-1) = 0 \operatorname{mod}(k(m))$ then we will set $c = 0$. Otherwise the choice $c = m\frac{k(m)-r}{k(m)}$ satisfies $n_1\ell + n_2(m-\ell) + c = 0 \operatorname{mod}(m)$, where r is the reminder of $n_1 + n_2(k(m)-1)$ after division by $k(m)$.

Finally, notice that for each $j = 1, \dots, m-1$ we have $\mathbb{P}(x_j = 1) = \mathbb{P}(x_j = -1) = \frac{1}{2}$. Thus if $T \neq S$, $|T| = d$ and $|T \cap g_j| = 1$ for all $j = 1, \dots, n$ then $\mathbb{E}x^T = \prod_{j \in T} \mathbb{E}x_j = 0$. If $S = T$, then $x_j \equiv 1$ for any $j \in T$ and hence $\mathbb{E}x^T = 1$. Thus taking the expectation $\mathbb{E}$ with respect to the distribution $\mathcal{D}$ in (4) finishes the proof for the case $S = \{m, 2m, \dots, dm\}$.

In general, if $|S| = d$ we still can split the coordinates of the vector $x = (x_1, \dots, x_n)$, into disjoint groups $G_1, \dots, G_d, G_{\text{extra}}, G_{\text{rest}}$ so that $|g_1| = \ldots = |g_d| = m$, $|g_j \cap S| = 1$ for all $j = 1, \dots, d$, and $|g_{\text{extra}}| = m - \frac{m}{k(m)}$. The rest of the proof proceeds as before. This finishes the proof in the general case. $\square$

2.3. **The proof of Proposition 1: how to cover the hypercube efficiently.** Consider 2^m skewed hyperplanes

$$\sum_{j=1}^{2^m-1} x_j + \sum_{j=0}^{m-1} \pm 2^j x_{2^m+j} = 0.$$

Since any odd integer k, $-(2^m-1) \leq k \leq 2^m-1$ can be written as a sum $\sum_{j=0}^{m-1} \pm 2^j$ for some choice of signs $\pm$ it follows that these hyperplanes cover the cube $\{-1,1\}^n$ with $n = 2^m + m - 1$. $\square$

There are other examples that are not produced by the construction above. In particular, for $n = 6$ the union of the following 5 skewed hyperplanes

$$\begin{aligned}
x_1 - x_2 + 2x_3 + x_4 + x_5 + 2x_6 &= 0,\\
x_1 - x_2 + x_3 + x_4 + x_5 - x_6 &= 0,\\
x_1 - x_2 - x_3 + 2x_4 - 2x_5 + x_6 &= 0,\\
x_1 + x_2 + x_3 + x_4 + x_5 - x_6 &= 0,\\
x_1 - x_2 - 3x_3 + x_4 + x_5 - x_6 &= 0.
\end{aligned}$$

cover the hypercube $\{-1,1\}^6$.

Acknowledgement

P.I. acknowledges support from NSF grant CAREER-DMS-2152401. O.K. was supported in part by a grant from the Israel Science Foundation (ISF Grant No. 1774/20), and by a grant from the US-Israel Binational Science Foundation and the US National Science Foundation (BSF-NSF Grant No. 2020643). R.V. acknowledges support from NSF DMS-1954233, NSF

DMS-2027299, U.S. Army 76649-CS, and NSF+Simons Research Collaborations on the Mathematical and Scientific Foundations of Deep Learning.

References

[1] N. Alon, E. .E. Bergmann, D. Coppersmith, A. M. Odlyzko, *Balancing sets of vectors*. IEEE Transactions on Information Theory, 34(1):128–130, 1988

[2] A. E. Brouwer, A. M. Cohen, A. Neumaier, *Distance-Regular Graphs*. (1989) Berlin, Heidelberg: Springer Berlin Heidelberg

[3] D. H. Gottlieb, *A certain class of incidence matrices*. Proc. Amer. Math. Soc. 17 (1966) 1233-1237.

[4] O. Klein, *Slicing all Edges of an n-cube requires $n^{2/3}$ hyperplanes*. (2022) arXiv: 2102.05536

[5] N. Linial, J. Radhakrishnan, *Essential covers of the cube by hyperplanes*. Journal of Combinatorial Theory, Series A 109.2 (2005), pp. 331-338.

[6] L. Saurman, Z. Xu, *Essential covers of the hypercube require many hyperplanes*. (2023) arXiv: 2310.05775

[7] G. Yehuda, A. Yehudayoff, *Slicing the hypercube is not easy*. (2021) arXiv: 2102.05536

(P. I.) Department of Mathematics, University of California, Irvine, Irvine, CA 92617, USA.
Email address: `pivanisv@uci.edu`

(O. K.) School of Computer Science and Engineering, Hebrew University of Jerusalem,, Jerusalem, Israel
Email address: `ohadkel@gmail.com`

(R. V.) Department of Mathematics, University of California, Irvine, Irvine, CA 92617, USA.
Email address: `rvershyn@uci.edu`