Kolmogorov Width of Discrete Linear Spaces:
an Approach to Matrix Rigidity

Alex Samorodnitsky[*]
Hebrew University
salex@cs.huji.ac.il

Ilya Shkredov[†]
Steklov Mathematical Institute
ilya.shkredov@gmail.com

Sergey Yekhanin
Microsoft
yekhanin@microsoft.com

Abstract

A square matrix V is called rigid if every matrix V' obtained by altering a small number of entries of V has sufficiently high rank. While random matrices are rigid with high probability, no explicit constructions of rigid matrices are known to date. Obtaining such explicit matrices would have major implications in computational complexity theory. One approach to establishing rigidity of a matrix V is to come up with a property that is satisfied by any collection of vectors arising from a low-dimensional space, but is not satisfied by the rows of V even after alterations. In this paper we propose such a candidate property that has the potential of establishing rigidity of combinatorial design matrices over the field $\mathbb{F}_2$.

Stated informally, we conjecture that under a suitable embedding of $\mathbb{F}_2^n$ into $\mathbb{R}^n$, vectors arising from a low dimensional $\mathbb{F}_2$-linear space always have somewhat small Kolmogorov width, i.e., admit a non-trivial simultaneous approximation by a low dimensional Euclidean space. This implies rigidity of combinatorial designs, as their rows do not admit such an approximation even after alterations. Our main technical contribution is a collection of results establishing weaker forms and special cases of the conjecture above.

1 Introduction

The notion of matrix rigidity was introduced by Leslie Valiant in 1977 [23]. In this paper we say that an $n \times n$ matrix A defined over a field is (r, d)-rigid, if it is not possible to reduce the rank of A below r by arbitrarily altering each row of A in up to d coordinates. Explicit rigid matrices are known to imply lower bounds for computational complexity of explicit functions.

The most prominent reduction of this nature is due to Valiant [23] who showed that for each $(\Omega(n), n^\epsilon)$-rigid matrix $A \in \mathbb{F}^{n \times n}$ the linear transformation induced by A cannot be computed by a linear circuit that simultaneously has size $O(n)$ and depth $O(\log n)$. Two other reductions that call for explicit (r, d)-rigid matrices with a sub-linear value of the remaining rank r, are given in [15, 18]. Reductions above naturally lead to the challenge of constructing rigid matrices explicitly. After more than three decades of efforts, however, this challenge remains elusive [13].

None of the existing techniques for constructing rigid matrices [11, 12, 3, 17, 5, 2] surpasses the basic *untouched minor* argument of [19] that amounts to taking a matrix where every minor has full rank, and using the bound from the Zarankiewicz problem [9, p. 25] to show that after up to d arbitrary changes per row there remains a somewhat large minor that has not been touched. Quantitatively, this yields explicit

$$\left(r, \Omega\left(\frac{n}{r} \log \frac{n}{r}\right)\right)$$

rigid matrices over fields of size $\Omega(n)$, when $\log^2 n \le r \le n/2$. Similar parameters are known to be attainable over small finite fields [6]. Note that for $r \le n/2$ random matrices over any field are at least

$$\left(r, \Omega\left(\frac{(n-r)^2}{n \log n}\right)\right)$$

rigid with a very high probability [23, Theorem 6.4].

1.1 Combinatorial designs

A family $\mathcal{F}$ of w-subsets of a universe of size n is called an (n, w, λ) design if every pair of distinct elements of $[n]$ belongs to exactly λ sets in $\mathcal{F}$. A combinatorial design is symmetric if $|\mathcal{F}| = n$. Geometric designs are a well studied class of symmetric combinatorial designs. A geometric design is defined by the incidence relation between points and hyperplanes in an m-dimensional projective space $\mathbb{PG}(m+1, q)$ over the finite field $\mathbb{F}_q$. Such a relation yields (n, w, λ) symmetric designs, where

$$n = \frac{q^{m+1} - 1}{q - 1} \qquad w = \frac{q^m - 1}{q - 1} \qquad \lambda = \frac{q^{m-1} - 1}{q - 1}. \tag{1}$$

With a slight abuse of notations we write $G_{m,q}$ or just G_m to denote both geometric designs and their incidence matrices.

In his original paper [23, Problem 4] Valiant proposed matrices G_2 defined above as natural candidates for $(\Omega(n), n^\epsilon)$-rigidity over the field $\mathbb{F}_2$. Taken literally, this conjecture is not true as some matrices G_2 have low rank over $\mathbb{F}_2$. In fact, the rank of geometric designs is a well studied quantity in design theory. We have [20]:

$$\operatorname{rank}_p G_{m,q} = \begin{cases} n & \text{if } q \ne p^e, \ w + (n-1)\lambda \ne 0 \bmod p; \\ n - 1 & \text{if } q \ne p^e, \ w + (n-1)\lambda = 0 \bmod p; \\ \binom{p+m-1}{m}^e & \text{if } q = p^e. \end{cases} \tag{2}$$

Thus in some cases the rank of geometric designs turns out to be surprisingly low, e.g., when $\operatorname{char} \mathbb{F}_q = 2$, for fixed m and growing q we have

$$\operatorname{rank}_2 G_{m,q} = O\left(n^{\frac{\log_2(m+1)}{m}}\right). \tag{3}$$

Identity (3) implies that any proof of (r, d)-rigidity of matrices G_m with $r = \Omega(n)$ cannot just rely on the combinatorial structure of these matrices as this structure does not seem to change much with the characteristic of the field underlying the projective space. Thus any rigidity proof that relies solely on the design properties of G_m (and thus applies to all designs with the parameters of geometric designs) has to be aiming at the regime of polynomially low remaining rank $r = O(n^\delta)$. In Section 1.3 we outline our approach to proving a result like this.

1.2 Hamada conjecture

In what follows let V_m denote an incidence matrix (or the set of rows of an incidence matrix) of a combinatorial (n, w, λ) design that has the parameters of the geometric design $G_{m,q}$. Clearly, any proof of (r, d)-rigidity of V_m has to imply that matrices V_m have rank at least r when no alterations are allowed. Bounding the rank of matrices V_m over finite fields has received some attention in design theory.

It is not hard to show that when $q \neq p^e$ we have, $\mathrm{rank}_p V_m \geq n - 1$. A conjecture [7] due to Noboru Hamada from 1973, asserts that when $q = p^e$, geometric designs $G_{m,q}$ have the lowest possible $\mathbb{F}_p$-rank among all designs V_m with the same parameters. Relatively little is known about the validity of Hamada conjecture [10, Section 4]. (A stronger version of Hamada's conjecture that asserts that every design V_m whose $\mathbb{F}_p$-rank equals that of $G_{m,q}$ has to be isomorphic to $G_{m,q}$ is known to be false [10].)

One easier natural question to ask that fits well with our approach to rigidity is whether one can prove any nontrivial lower bounds on the rank of design matrices V_m. We are particularly interested in the asymptotic setting of fixed m and growing q. Hamada conjecture and identity (2) suggest that

$$\mathrm{rank}_p V_m \geq \Omega\left(n^{\frac{\log_p\left(\binom{p+m-1}{m}\right)}{m}}\right). \tag{4}$$

The trivial lower bound is $\mathrm{rank}_p V_m \geq n^{\frac{1}{m}}$. We are not aware of any better bound.

1.3 Our approach

In order to establish rigidity of matrices V_m over the field $\mathbb{F}_2$ we propose a certain property that is not satisfied by the rows of V_m even after alterations, yet that we conjecture to hold for any collection of vectors arising from a low-dimensional $\mathbb{F}_2$-linear space.

As a first step of our argument we consider a natural embedding of the space $\mathbb{F}_2^n$ into $\mathbb{R}^n$. We treat elements of $\mathbb{F}_2^n$ as real $\{0, 1\}$-vectors and normalize them to have L_2 norm one. Thus a non-zero $x \in \mathbb{F}_2^n$ gets mapped to $\frac{x}{\|x\|}$. In what follows we assume that this embedding is implied and treat vectors in $\mathbb{F}_2^n$ as real vectors.

Next for sets $X \subseteq \mathbb{R}^n$ we consider the quantity $A_r(X)$ that we call the approximability measure. $A_r(X)$ is defined to be the maximum over all r-dimensional Euclidian linear spaces W of the square of the smallest projection of a vector from X onto W. Thus sets with large value of A_r are precisely those that are well approximated by some r-dimensional Euclidian linear space, i.e., the ones that have small Kolmogorov width.

Now let $m = \frac{1}{\epsilon}$. We argue that for all values of $r = \omega(n^{1-\epsilon})$, the approximability measure $A_r(V_m) \approx A_r(\mathbb{F}_2^n)$. Thus for sufficiently large r, approximating rows of an incidence matrix of a combinatorial design is no easier than approximating all of the Hamming space. The claim remains true even if we allow rows of V_m be altered in up to $O\left(n^{1-2\epsilon}\right)$ coordinates.

Finally, we conjecture that for any $\mathbb{F}_2$-linear space L, $\dim L \leq n^{2\epsilon+\delta}$, and some $r = \omega(n^{1-\epsilon})$, the approximability measure $A_r(L) \geq (1+\alpha)A_r(\mathbb{F}_2^n)$, for some positive α. In other words, we conjecture that low dimensional $\mathbb{F}_2$-linear spaces keep some tiny amount of resemblance to Euclidian linear spaces after the embedding, and can be approximated better than all of the Hamming cube. It is easy to see that this conjecture implies $(n^{2\epsilon+\delta}, n^{1-2\epsilon})$-rigidity of matrices V_m, since if one of these matrices had low rank after alterations, its rows would belong to a low dimensional $\mathbb{F}_2$-linear space, and thus admit a non-trivial Euclidian approximation.

We measure our progress towards the conjecture by looking at the largest value of dimension k for that we are indeed able to prove that all k-dimensional $\mathbb{F}_2$-linear spaces L satisfy $A_r(L) \geq (1 + \alpha) A_r(\mathbb{F}_2^n)$, for some $r = \omega(n^{1-\epsilon})$. Currently, our main Theorem 15 gives this for all $k = o(n^\epsilon \log n)$. Apart from this result, we also establish the conjecture for a certain restricted class of linear spaces called cut-spaces. While substantial further progress is needed to establish rigidity of matrices V_m, our current results (Corollary 16) already suffice to get the bound

$$\text{rank}_p V_m \geq \Omega\left(n^{\frac{1}{m}} \log_2 n\right),\tag{5}$$

for all values of p, a results that seems to be new.

From the technical viewpoint our main contribution is a new relation between discrete ($\mathbb{F}_2$) linear spaces and Euclidian linear spaces, yielding some insight into combinatorics of low weight codewords in linear codes.

1.4 Organization

In Section 3 we formally introduce the approximability measure A_r. We argue that for sufficiently large values of dimension r, we have $A_r(V_m) \approx A_r(\mathbb{F}_2^n)$. We establish a similar result for perturbed matrices V_m. Next, we introduce our main conjecture stating that low-dimensional $\mathbb{F}_2$-linear spaces L always have a somewhat large value of A_r. We show how this conjecture implies rigidity of matrices V_m.

In Sections 4 through 6 we prove our main results regarding approximability of low-dimensional $\mathbb{F}_2$-linear spaces L. In Section 4 we deal with low-dimensional approximations and obtain a bound for $A_1(L)$. In Section 5 we deal with high dimensional approximations and state the implications of our results for the Hamada conjecture. All our results in Sections 4 and 5 apply not just to $\mathbb{F}_2$-linear spaces but to all families of vectors that have bounded triangular rank [14].

In Section 6 we establish our main approximability conjecture for a certain class a linear spaces called cut-spaces and give a simpler proof of a slightly weaker version of the results from Section 5. Our constructions of approximating real spaces use low weight vectors in the dual space of L. Finally, in Section 7 we discuss the relation of our approach to matrix rigidity to the natural proofs lower bounds barrier of Razborov and Rudich [16].

2 Notation

We use the following standard mathematical notation:

- $\|\cdot\|$ denotes the Euclidian norm;

- For an integer n, $[n] = \{1, \ldots, n\}$;

- For a vector $\mathbf{v}$, the set of non-zero coordinates of $\mathbf{v}$ is denoted $\text{supp}(\mathbf{v})$;

- We write $f(n) \approx g(n)$, if $f(n) = g(n)(1 + o(1))$. We adopt the same agreement for $\lesssim$ and $\gtrsim$;

3 The conjecture

We now introduce our approximability measure A_r. Following that, in Section 3.2 we argue that for sufficiently large values of dimension r, collections of rows of incidence matrices of combinatorial designs have essentially the smallest possible value of A_r even after alterations. Finally, in Section 3.3 we introduce our main conjecture stating that low-dimensional $\mathbb{F}_2$-linear spaces always have a somewhat large value of A_r. We show how this conjecture implies rigidity of incidence matrices of combinatorial designs.

3.1 The approximability measure

We consider a natural embedding of the space $\mathbb{F}_2^n$ into $\mathbb{R}^n$. We treat elements of $\mathbb{F}_2^n$ as real $\{0,1\}$-vectors and normalize them to have L_2 norm one. Thus a non-zero $\mathbf{v} \in \mathbb{F}_2^n$ gets mapped to $\frac{\mathbf{v}}{\|\mathbf{v}\|}$. Zero is mapped to zero. In what follows we assume that this embedding is implied and treat vectors in $\mathbb{F}_2^n$ as real vectors. Let V be an arbitrary subset of $\mathbb{F}_2^n$. Our approach is centered around the following approximability measure

$$A_r(V) = \max_{\dim W \leq r} \min_{\mathbf{v} \in V} \|\mathrm{Pr}_W(\mathbf{v})\|^2, \tag{6}$$

where the maximum is taken over all linear spaces $W \in \mathbb{R}^n, \dim W = r$, and minimum is taken over the non-zero elements of V. Observe that our notion of approximability measure is equivalent to the classical concept of Kolmogorov width $K_r(V) = \sqrt{1 - A_r(V)}$, also known as "poperechnik" of a family of vectors. See [21, 22]. We now derive a formula for approximability measure of the whole Boolean cube.

Lemma 1 *Let $r = o(n)$ be arbitrary. We have*

$$A_r\left(\mathbb{F}_2^n\right) \approx \frac{r}{n}. \tag{7}$$

Proof: Let $\mathbf{e}_i, i \in [n]$ denote the i-th unit vector. First we show that

$$A_r\left(\mathbb{F}_2^n\right) \leq A_r\left(\{\mathbf{e}_1, \ldots, \mathbf{e}_n\}\right) \leq \frac{r}{n}. \tag{8}$$

Let W be an arbitrary r-dimensional linear space with an orthonormal basis $\{\mathbf{w}_1, \ldots, \mathbf{w}_r\}$. Consider an $n \times r$ matrix M, where $M_{ij} = (\mathbf{e}_i, \mathbf{w}_j)^2$. Clearly, the sum of values in M is equal to r. Thus for some $i \in [n]$ we have $\sum_j (\mathbf{e}_i, \mathbf{w}_j)^2 \leq \frac{r}{n}$ and (8) follows.

We now exhibit a space W such that for all non-zero binary vectors $\mathbf{v}$, $\|\mathrm{Pr}_W(\mathbf{v})\|^2 \gtrsim \frac{r}{n}$. The space W is spanned by r unit vectors $\{\mathbf{w}_i\}$. These vectors have disjoint supports that partition $[n]$. Every support is of size $\lceil \frac{n}{r} \rceil$ or $\lfloor \frac{n}{r} \rfloor$. Each vector $\mathbf{w}_i$ is constant on its support. Let $\mathbf{v}$ be an arbitrary vector of weight w. Assume that the support of $\mathbf{v}$ intersects the supports of t different vectors $\{\mathbf{w}_i\}$, namely, $\mathbf{w}_{i_1}, \ldots, \mathbf{w}_{i_t}$. Clearly, $t \leq w$. For $j \leq t$, let $a_j = |\mathrm{supp}(\mathbf{v}) \cap \mathrm{supp}(\mathbf{w}_{i_j})|$. We have $\sum_{j=1}^t a_j = w$. We also have

$$\sum_{j=1}^t (\mathbf{v}, \mathbf{w}_{i_j})^2 \geq \sum_{j=1}^t \frac{a_j^2 r}{(n+r)w} = \frac{r}{(n+r)w} \sum_{j=1}^t a_j^2 \geq \frac{r}{w(n+r)} \left(\frac{w}{t}\right)^2 t \gtrsim \frac{r}{n}.$$

This concludes the proof. ∎

3.2 Inapproximability of combinatorial designs

In this Section we argue that approximating rows of a combinatorial design by a high-dimensional real space is as hard as approximating all of the Boolean cube. We also establish a robust version of this result.

Lemma 2 *Let $V \subseteq \mathbb{F}_2^n, |V| = n$. Let B be the $n \times n$ real matrix, where the rows of B are the normalized elements of V. Let $\lambda_1 \geq \ldots \geq \lambda_n \geq 0$ be the eigenvalues of BB^t; then for all r,*

$$A_r(V) \leq \frac{1}{n} \sum_{i \leq r} \lambda_i. \tag{9}$$

Proof: This is a simple corollary of a result in [8]. A special case of this result states that for all r

$$\frac{1}{n}\sum_{i\leq r}\lambda_i = \max_{\dim W\leq r} \mathbb{E}_{\mathbf{v}\in V}\|\Pr_W(\mathbf{v})\|^2, \tag{10}$$

where the expectation is taken with respect to the uniform distribution on V. Since the RHS of this equality is at least as large as $A_r(V)$, the claim of the Lemma follows. ∎

Let V_m be an incidence matrix of a combinatorial (n, w, λ) design with the parameters (1) of the geometric design $G_{m,q}$ for some value of q. Let $m = \frac{1}{\epsilon}$. We assume that m is fixed and q grows to infinity. Thus $w \approx n^{1-\epsilon}$ and $\lambda \approx n^{1-2\epsilon}$. Lemma 2 yields

Corollary 3 *With the notation above, we have*

$$A_r\left(V_{1/\epsilon}\right) \lesssim \frac{n^{1-\epsilon} + r}{n}. \tag{11}$$

Proof: Let B be the $n \times n$ matrix, where the rows of B are the normalized elements of V_m. Clearly, $B = \frac{1}{\sqrt{w}}V_m$. We have

$$BB^t = \frac{w - \lambda}{w}I + \frac{\lambda}{w}J,$$

where I denotes the identity matrix and J denotes the all-ones matrix. It is not hard to see that the eigenvalues of BB^t are given by

$$\lambda_1 \approx n^{1-\epsilon} \quad \text{and} \quad \lambda_2 = \ldots = \lambda_n \approx 1.$$

An application of Lemma 2 completes the proof. ∎

Combining (7) and (11), we conclude that for $r = \omega(n^{1-\epsilon})$ and $r = o(n)$,

$$A_r(V_{1/\epsilon}) \approx A_r(\mathbb{F}_2^n) \approx \frac{r}{n}. \tag{12}$$

Observe that identity (10) and the eigenvalue computation above can be used to show that matrices V_m are inapproximable on average and not just in the worst case. The following Lemma gives a stability result for A_r:

Lemma 4 *Let $V = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\} \subseteq \mathbb{F}_2^n$ be a set of vectors of Hamming weight w. Assume the new set $V' = \{\mathbf{v}'_1, \ldots, \mathbf{v}'_n\} \subseteq \mathbb{F}_2^n$ is obtained from V by altering at most d coordinates of each $\mathbf{v}_i$, where $d < w$; then for all r,*

$$A_r(V') \leq \left(\sqrt{A_r(V)} + \sqrt{\frac{d}{w}}\right)^2. \tag{13}$$

Proof: Let $\mathbf{v}, \mathbf{v}'$ be arbitrary binary vectors such that the Hamming weight of $\mathbf{v}$ is w and the Hamming distance between $\mathbf{v}$ and $\mathbf{v}'$ is at most d. It is not hard to see that after the embedding in the real space we have

$$(\mathbf{v}, \mathbf{v}') \geq \sqrt{1 - \frac{d}{w}}. \tag{14}$$

The minimum in (14) is attained by a vector $\mathbf{v}'$ of Hamming weight $w - d$, where the support of $\mathbf{v}'$ is a subset of the support of $\mathbf{v}$. Let W be an r-dimensional real space in $\mathbb{R}^n$ that attains the maximum in (6) for approximating the set V'. Let $\mathbf{w}_1, \ldots, \mathbf{w}_n$ be a family of unit vectors in W such that for all $i \in [n]$, $(\mathbf{v}'_i, \mathbf{w}_i)^2 \geq A_r(V')$. Let $A = \sqrt{A_r(V)}$. By definition of A there exists $i \in [n]$ such that

$$(\mathbf{v}_i, \mathbf{w}_i) \leq A. \tag{15}$$

We introduce notation for angles between vectors $\mathbf{v}_i, \mathbf{v}'_i$, and $\mathbf{w}_i$. Let

$$\alpha = \angle(\mathbf{v}'_i, \mathbf{w}_i), \quad \beta = \angle(\mathbf{v}_i, \mathbf{v}'_i), \quad \gamma = \angle(\mathbf{v}_i, \mathbf{w}_i).$$

Clearly $\alpha, \beta \in [0, \pi/2]$ and $\alpha \geq \gamma - \beta$. First suppose that $\gamma - \beta \geq 0$; then

$$(\mathbf{v}'_i, \mathbf{w}_i) = \cos\alpha \leq \cos\gamma\cos\beta + \sin\gamma\sin\beta \leq \max\{0, \cos\gamma\} + \sin\beta \leq A + \sqrt{\frac{d}{w}},$$

where the last inequality follows from (14) and (15). Now note that if $0 \leq \gamma \leq \beta \leq \pi/2$; then

$$(\mathbf{v}'_i, \mathbf{w}_i) \leq 1 \leq \cos\gamma + \sin\beta \leq A + \sqrt{\frac{d}{w}}.$$

The inequality (13) follows. ∎

The above Lemma and identity (12) yield

Proposition 5 *Let V_m be an $n \times n$ matrix of a combinatorial design with the parameters of a geometric design G_m. Assume $m = \frac{1}{\epsilon}$ is fixed and n grows to infinity. Let V'_m be obtained from V_m by altering each row in up to $O\left(n^{1-2\epsilon}\right)$ coordinates. Let $r = \omega(n^{1-\epsilon})$. We have*

$$A_r\left(V'_{1/\epsilon}\right) \approx \frac{r}{n}. \tag{16}$$

3.3 The conjecture and rigidity implications

We now introduce our main conjecture stating that low-dimensional $\mathbb{F}_2$-linear spaces L always have a somewhat large value of A_r and show how this conjecture implies rigidity of design matrices V_m. We begin with a formal definition of rigidity.

Definition 6 *Let V be an $n \times n$ matrix over a field $\mathbb{F}$. We say that V is (r, d)-rigid; if for every matrix V' that differs from V in at most d coordinates in each row, we have $\operatorname{rank}_{\mathbb{F}} V \geq r$.*

Conjecture 7 *There exists positive constants α, δ, and $\epsilon = \frac{1}{m}$ such that for all linear spaces $L \subseteq \mathbb{F}_2^n$ where $\dim L \leq n^{2\epsilon+\delta}$, for some $r = \omega(n^{1-\epsilon})$,*

$$A_r(L) \geq (1+\alpha)\frac{r}{n}. \tag{17}$$

The conjecture above trivially implies $\left(n^{2\epsilon+\delta}, n^{1-2\epsilon}\right)$-rigidity of design matrices V_m over the field $\mathbb{F}_2$. If some matrix V'_m had rank below $n^{2\epsilon+\delta}$ after $O(n^{1-2\epsilon})$ alterations in each row; then its rows would belong to a $n^{2\epsilon+\delta}$-dimensional linear space over $\mathbb{F}_2$, and thus have non-trivial approximation measure, contradicting Proposition 5. Currently we can only prove the conjecture for all linear spaces L with $\dim L = o\left(n^\epsilon \log n\right)$. (Theorem 15).

4 Low dimensional approximations from bounded triangular rank

In this and the following two Sections we prove our main results regarding approximability of low-dimensional $\mathbb{F}_2$-linear spaces L. In the current Section we deal with low-dimensional approximations and obtain a bound for $A_1(L)$. All results obtained in this Section apply not just to $\mathbb{F}_2$-linear spaces but to all families of vectors that have bounded triangular rank.

Definition 8 *Let $T = \{\mathbf{v}_1, \ldots, \mathbf{v}_t\}$ be a sequence of binary vectors of of dimension n. We say that T is a tower of height t if for all $j \leq t$:*

$$\operatorname{supp}(\mathbf{v}_j) \nsubseteq \bigcup_{s \leq j-1} \operatorname{supp}(\mathbf{v}_s).$$

Further, let V be an arbitrary collection of binary vectors. We define the triangular rank of V, denoted $\operatorname{trk}(V)$ to be the largest height of a tower that can be constructed from elements of V.

It is easy to see that for any subset V of an $\mathbb{F}_2$-linear space L we always have $\operatorname{trk}(V) \leq \dim L$. Let $L \subseteq \mathbb{F}_2^n$, $\operatorname{trk}(L) \leq k$. For $i \in [n]$, let

$$w_i = \min_{\mathbf{v} \in L \,:\, i \in \operatorname{supp}(\mathbf{v})} |\operatorname{supp}(\mathbf{v})|. \tag{18}$$

If $i \in [n]$ does not belong to the support of any vector in L; we define $w_i = \infty$. We set

$$\mu = \mu(L) = \sum_{i=1}^{n} w_i^{-1}. \tag{19}$$

Our proof of the following Lemma resembles some of the arguments in [14, Section 2].

Lemma 9 *Let $L \subseteq \mathbb{F}_2^n$, $\operatorname{trk}(L) \leq k$; then*

$$\mu(L) \leq k.$$

Proof: Assume $\mu > k$. We derive a contradiction by exhibiting a collection $V = \{\mathbf{v}_1, \ldots, \mathbf{v}_{k+1}\} \subseteq L$ such that for every $j \in [k+1]$, there exists $i_j \in \operatorname{supp}(\mathbf{v}_j)$, such that $i_j \notin \operatorname{supp}(\mathbf{v}_s)$, for all $s < j$. Consider an n-node hypergraph, where the hyperedges are the supports of the elements of L. Color all nodes white. Set $\Phi = \mu$, $V = \emptyset$. On the j-th step:

1. We choose a white node i whose w_i is the smallest among the white nodes;

2. We set $\mathbf{v}_j$ to be a weight-w_i element of L such that $i \in \operatorname{supp}(\mathbf{v}_j)$;

3. We set

$$\Delta = \sum_{s \in \operatorname{supp}(\mathbf{v}_j) \,|\, s \text{ is white}} w_s^{-1}.$$

   It is important to note that Δ is necessarily at most 1.

4. We reduce Φ by Δ and color all nodes $s \in \operatorname{supp}(\mathbf{v}_j)$ black.

On each step we reduce Φ by at most one and increase $\operatorname{trk}(V)$ by one. Thus after $k+1$ steps we necessarily have $\operatorname{trk}(V) > k$. ∎

Theorem 10 *Let $L \subseteq \mathbb{F}_2^n$, $\operatorname{trk}(L) \leq k$; then*

$$A_1(L) \geq \frac{1}{k}.$$

Proof: Let $\{w_i\}$ and μ be as defined in (18) and (19). Fix a vector $\mathbf{w} \in \mathbb{R}^n$, where for all $i \in [n]$, $\mathbf{w}_i = \sqrt{\frac{1}{\mu w_i}}$. Clearly $\|\mathbf{w}\| = 1$. Let $\mathbf{v} \in L$ be arbitrary, $|\operatorname{supp}(\mathbf{v})| = w$. Note that for all $i \in \operatorname{supp}(\mathbf{v})$, $w_i \leq w$. It remains to note that

$$(\mathbf{w}, \mathbf{v}) \geq \frac{w\sqrt{1/\mu w}}{\sqrt{w}} = \frac{1}{\sqrt{\mu}} \geq \frac{1}{\sqrt{k}},$$

where the last inequality follows from Lemma 9. ∎

Theorem 10 exhibits a vast gap between $A_1(\mathbb{F}_2^n) \approx \frac{1}{n}$ and $A_1(L) \geq \frac{1}{k}$ for $\mathbb{F}_2$-linear spaces L that have polynomially low dimension k. This Theorem alone already implies that our main Conjecture 7 holds for all linear spaces of dimension up to $o(n^\epsilon)$. In fact it shows that even one-dimensional approximations of discrete linear spaces suffice to get this result.

5   High dimensional approximations from bounded triangular rank

In this Section we deal with high dimensional approximations and state the implications of our results for the Hamada conjecture. Our main result is given by Theorem 15. As in the previous Section our arguments apply not just to $\mathbb{F}_2$-linear spaces but to all families of vectors that have bounded triangular rank. To simplify notation in this Section we do not distinguish between binary vectors $\mathbf{v}$ and their support sets $\mathrm{supp}(\mathbf{v})$.

Definition 11 *Let L be a family of subsets of some universe. Let S be a subset of the same universe. We say that S is a (c, k)-attractor for L if $\mathrm{trk}(L) \leq k$ and for all $\mathbf{v} \in L$ such that $\mathbf{v} \cap S \neq \emptyset$,*

$$|\mathbf{v} \cap S| \geq c \cdot \frac{|S|}{k}. \tag{20}$$

Below is the key Lemma of this Section.

Lemma 12 *Let L be a family of binary vectors. Let $[N]$ be the union of supports of vectors in L. Let $\mathrm{trk}(L) \leq k$. Assume that Hamming weights of all vectors in L lie in the segment $[w, 2w]$. Further assume $k \geq 2^{5c+2}$ where c is an integer. Then there exists a (c, k)-attractor for L of size at least $\frac{N}{2^{4c}}$.*

Proof: Note that if $\frac{N}{2^{4c+2}} < w$, then the set $[N]$ is a (c, k)-attractor for L. In fact, let $\mathbf{v} \in L$ be arbitrary. We have

$$|\mathrm{supp}(\mathbf{v}) \cap [N]| \geq w \geq \frac{N}{2^{4c+2}} \geq c \cdot \frac{N}{2^{5c+2}} \geq c \cdot \frac{|[N]|}{k}.$$

Thus without loss of generality we assume

$$\frac{N}{2^{4c+2}} \geq w. \tag{21}$$

We now execute the following simple greedy algorithm that constructs a tower in the family L.

1. Set the tower T to be an empty family of sets. Set $R = [N]$.

2. WHILE $R \neq \emptyset$ DO

3.   BEGIN

4.   Identify a set $\mathbf{v} \in L$ that minimizes $|\mathbf{v} \cap R|$;

5.   Add $\mathbf{v}$ to the tower T;

6.   Drop the elements in $\mathbf{v}$ from R;

7.   END

The algorithm above terminates producing a tower of height at most k. On step j the algorithm adds a new vector to T and reduces the set R by $\Delta_j = T \cap R$. We partition the steps of the algorithm into *stages*. A step falls into stage number i if in the beginning on the step

$$\frac{N}{2^i} < |R| \leq \frac{N}{2^{i-1}}. \tag{22}$$

Observe that among the first $4c$ stages there is at least one stage i, such that the height of T increases by $t \leq \frac{k}{4c}$ during that stage. Let S be the change in the set R on the stage i. We have

$$|S| \geq \left(\frac{N}{2^{i-1}} - 2w \right) - \frac{N}{2^i} \geq \frac{N}{2^{i+1}}.$$

The first inequality above follows from the fact that in the beginning of stage i the size of R is at least $\frac{N}{2^{i-1}}$ minus the size of the last step of stage $(i-1)$ and the fact that every step size is bounded by $2w$. The second inequality follows from (21). Let a be the index of the first step of stage i. We have

$$\sum_{j=a}^{a+t-1} |\Delta_j| = |S| \geq \frac{N}{2^{i+1}}.$$

Thus at stage i there exists a step j such that

$$|\Delta_j| \geq \frac{N}{2^{i+1}} \cdot \frac{4c}{k}.$$

Let A be the set R at the beginning of step j. Clearly,

$$\frac{N}{2^i} < |A| \leq \frac{N}{2^{i-1}}.$$

Combining the last two inequalities we get

$$\frac{|A|}{|\Delta_j|} \leq \frac{k}{c}.$$

Therefore by the greedy property of our algorithm for every set $\mathbf{v} \in L$ that intersects A we have

$$\frac{|A|}{|A \cap \mathbf{v}|} \leq \frac{k}{c},$$

or equivalently

$$|A \cap \mathbf{v}| \geq c \cdot \frac{|A|}{k}.$$

Thus A is a (c, k)-attractor for L of size at least $\frac{N}{2^{4c}}$. ∎

Lemma 13 *Let L be a family of binary vectors, $\mathrm{trk}(L) \leq k$. Assume that Hamming weights of all vectors in L lie in some segment $[\sqrt{2}w, 2w]$. Further assume $k \geq 2^{5c+2}$ where c is an integer. We have*

$$A_{c \cdot 2^{4c}}(L) \geq \Omega \left(\frac{c}{k} \right), \tag{23}$$

where the constant in Ω-notation is absolute.

Proof: Set $[N]$ be the union of supports of vectors in L. Clearly, $N \leq 2wk$. We now construct a basis for the approximating real space.

1. $\pi = \{\pi_i\}_{i \geq 0}$ is a partition of $[N]$. Initially π consists of a single set $\pi_0 = [N]$ and $i = 0$.

2. WHILE $((i < c \cdot 2^{4c})$ AND $(\pi_0 \neq \emptyset))$ DO

3.   BEGIN

4.   Identify a (c, k)-attractor π_i for L of relative size at least $\frac{1}{2^{4c}}$;

5.   Remove elements of π_i from π_0 and from all sets $\mathbf{v} \in L$;

6.   Add the set π_i to π;

7.   Drop every element $\mathbf{v}$ such that $|\mathbf{v}| < w$ from L;

8.   Increment i;

9.   END

Lemma 12 ensures that the attractor constructed on step 4 above always exists. Observe that by the end of the execution of the algorithm we have

$$|\pi_0| \leq N \cdot \left(1 - \frac{1}{2^{4c}}\right)^{c \cdot 2^{4c}} \leq \frac{N}{c}. \tag{24}$$

Recall that $\pi = \{\pi_i\}_{i \geq 0}$ is a partition of $[N]$. Let W be the real linear space spanned by binary vectors $\mathbf{p}_0, \ldots, \mathbf{p}_r$ whose supports are elements of this partition. Clearly, $\dim W \leq c \cdot 2^{4c}$. We claim that W approximates all $\mathbf{v} \in L$ well. Consider two cases:

- $|\mathbf{v} \cap \pi_0| < w$. At least $(\sqrt{2} - 1)w$ elements of $\mathrm{supp}(\mathbf{v})$ fall onto (c, k)-attractors in π. To approximate $\mathbf{v} \in L$ consider the set $J = \{j \mid \mathbf{v} \cap \pi_j \neq \emptyset \text{ and } j \neq 0\}$. For all $j \in J$ by (20) we have

$$|\mathbf{v} \cap \pi_j| \geq c \cdot \frac{|\pi_j|}{k}.$$

  Therefore

$$|\pi_j| \leq |\mathbf{v} \cap \pi_j| \cdot \frac{k}{c}. \tag{25}$$

  Consider the vector $\mathbf{p} = \sum_{j \in J} \mathbf{p}_j$. Summing (25) over all $j \in J$ we obtain

$$\mathrm{wt}(\mathbf{p}) \leq 2w \cdot \frac{k}{c},$$

  where $\mathrm{wt}(\mathbf{p})$ denotes the Hamming weight. Thus

$$\left(\frac{\mathbf{p}}{\|\mathbf{p}\|}, \frac{\mathbf{v}}{\|\mathbf{v}\|}\right)^2 \geq \frac{|\mathbf{v} \cap \mathrm{supp}(\mathbf{p})|^2}{\mathrm{wt}(\mathbf{v})} \cdot \frac{c}{2wk} \geq \frac{(\sqrt{2} - 1)^2}{4} \cdot \frac{c}{k}.$$

- $|\mathbf{v} \cap \pi_0| \geq w$. By (24) we have $|\pi_0| \leq \frac{2wk}{c}$. Consider the binary real vector $\mathbf{p}$ whose support is π_0. We have

$$\left(\frac{\mathbf{p}}{\|\mathbf{p}\|}, \frac{\mathbf{v}}{\|\mathbf{v}\|}\right)^2 \geq \frac{w^2}{2w} \cdot \frac{c}{2wk} = \frac{c}{4k}.$$

$\blacksquare$

In what follows all $\log$'s are base 2 unless otherwise specified.

Theorem 14 *Let $L \subseteq \mathbb{F}_2^n$, $\mathrm{trk}(L) \leq k$. Assume $k \geq 2^{5c+2}$ where c is an integer. We have*

$$A_{2c \cdot 2^{4c} \cdot \log n}(L) \geq \Omega\left(\frac{c}{k}\right), \tag{26}$$

where the constant in Ω-notation is absolute.

Proof: Partition the set L into $2 \log n$ subsets $L_1, \ldots, L_{2 \log n}$ where every set L_i contains elements of L whose Hamming weight is between $2^{(i-1)/2}$ an $2^{i/2}$. Apply Lemma 13 to each L_i. Consider the joint span of $2 \log n$ resulting real spaces to approximate L. ∎

Theorem 15 *Let $L \subseteq \mathbb{F}_2^n$, $\mathrm{trk}(L) \leq k$; then*

- *For all $\tau > 0$ and sufficiently large k and n,*

$$A_{n^\tau}(L) \geq \Omega\left(\frac{\log k}{k}\right),\tag{27}$$

   where the constant in the Ω-notation depends only on τ.

- *The bullet above implies that for all α and all $\epsilon > 0$ our main Conjecture 7 holds for all linear spaces L, where $\dim L = o(n^\epsilon \log n)$.*

Proof: We start with the first bullet. Set $c = \left\lfloor \min\left\{\frac{\tau}{8} \log k, \frac{\log k - 2}{5}\right\} \right\rfloor$. Theorem 14 yields

$$A_{\frac{\tau}{4} \log k \log n k^{\tau/2}}(L) \geq \Omega\left(\frac{\log k}{k}\right),$$

which immediately yields (27) for large enough n.

We proceed to the second bullet. Let $k = \dim L = \beta(n) n^\epsilon \log n$, where $\beta(n) \to 0$ but $\beta \log n$ grows. Fix an arbitrary $\tau < 1 - \epsilon$. By (27)

$$A_{n^\tau}(L) \geq \frac{c \log k}{k},$$

for some constant c. Set $r(n) = \frac{c\epsilon n^{1-\epsilon}}{\beta(n)(1+\alpha)}$. Observe that for sufficiently large n,

$$A_r(L) \geq A_{n^\tau}(L) \geq \frac{c \log k}{k} \geq \frac{c\epsilon \log n}{\beta n^\epsilon \log n} = \frac{c\epsilon}{\beta n^\epsilon} = (1+\alpha)\frac{r}{n}.$$

This concludes the proof. ∎

The following Corollary gives the implication of Theorem 15 for the triangular rank of combinatorial designs.

Corollary 16 *Let $m = \frac{1}{\epsilon}$ and let V_m be the $n \times n$ incidence matrix of a combinatorial design that has the parameters of the geometric design $G_{m,q}$. We have*

$$\mathrm{trk}(V_m) = \Omega\left(n^\epsilon \log_2 n\right),\tag{28}$$

where $\mathrm{trk}(V_m)$ denotes the triangular rank of the collection of rows of V_m.

Proof: Let $\mathrm{trk}(V_m) = k$. Set $r = n^{1-\epsilon}$. From Corollary 3 we have $A_r(V_m) \lesssim \frac{2}{n^\epsilon}$. However from Theorem 15 we have $A_r(V_m) \geq \Omega\left(\frac{\log k}{k}\right)$. Therefore

$$\frac{2}{n^\epsilon} \geq \Omega\left(\frac{\log k}{k}\right).$$

Thus $k = \Omega\left(n^\epsilon \log_2 n\right)$. ∎

Note that triangular rank of V_m gives a lower bound for the rank of V_m over any field.

6 High dimensional approximations from short dual vectors

In this Section we establish Conjecture 7 for a certain class a linear spaces called cut-spaces and give a simpler proof of a slightly weaker version of Theorem 15. In both of these results we use low weight vectors in the dual space of an $\mathbb{F}_2$-linear space L to construct the approximating real space for L. As in the previous Section, we often write $\mathbf{v}$ to denote both a vector $\mathbf{v}$ and its support set $\mathrm{supp}(\mathbf{v})$.

Definition 17 *Let L be a family of subsets of $[n]$. We say that the r-partition $\pi = \bigsqcup_{j \leq r} \pi_j$ of $[n]$ is attractive for L, if for some constant $\alpha > 0$ for every $\mathbf{v} \in L$ we have*

$$\left| \bigsqcup_{j \,:\, \mathbf{v} \cap \pi_j \neq \emptyset} \pi_j \right| \leq \frac{|\mathrm{supp}(\mathbf{v})|}{(1+\alpha)} \cdot \frac{n}{r}. \tag{29}$$

Lemma 18 *Let L be a set of n-dimensional binary vectors. Suppose there exists an r-partition of $[n]$ that is attractive for L; then $A_r(L) \geq (1+\alpha)\frac{r}{n}$.*

Proof: Let W be the real linear space spanned by binary vectors $\mathbf{p}_1, \ldots, \mathbf{p}_r$ whose supports are elements of the attractive r-partition. To approximate $\mathbf{v} \in L$ consider the vector $\mathbf{p} = \sum_{j \,:\, \mathbf{v} \cap \pi_j \neq \emptyset} \mathbf{p}_j$. We have

$$\mathrm{wt}(\mathbf{p}) \leq \frac{\mathrm{wt}(\mathbf{v})}{(1+\alpha)} \cdot \frac{n}{r}.$$

Thus $(\mathbf{p}/\|\mathbf{p}\|, \mathbf{v})^2 \geq (1+\alpha)\frac{r}{n}$. ∎

6.1 Approximating cut spaces

A cut space is a subspace of $\mathbb{F}_2^n$ that has a $k \times n$ generator matrix where every column has weight two. Equivalently, a cut space is defined by a k-node graph G with n edges. Elements of the cut space are incidence vectors of cuts in the graph. Elements of the dual space are incidence vectors of even degree subgraphs of G. In what follows we restrict our attention to connected graphs G. For such graphs the dimension of the corresponding cut space is $k-1$. We now argue that cut spaces satisfy Conjecture 7.

Theorem 19 *Let $L \subseteq \mathbb{F}_2^n$ be a cut space, $\dim L \leq o\left(\frac{n}{\log n}\right)$. Then for some $r = \Theta\left(\frac{n}{\log n}\right)$ and $\alpha > 0$ we have*

$$A_r(L) \geq (1+\alpha)\frac{r}{n}.$$

Proof: We rely on the fact that any graph with k nodes and n edges contains a cycle of length at most $2\log n$ provided $n \geq 3k$. We consider the graph G corresponding to L. We construct a family π of disjoint subsets of edges of G (coordinates of L). We start by executing the following simple algorithm:

1. Start with an empty family of sets π.

2. WHILE $n \geq 3k$ DO

3.   BEGIN

4.   Identify a cycle C in G, $|C| \leq 2\log n$;

5.   Include C into π as a new set;

6.   Drop edges in C from G;

7. END

Our next goal is to make sure that most sets in π have approximately the same size. Firstly, we repeatedly join together any two sets in π, if the sum of their sizes is below $2\log n$. Secondly, we drop the smallest set from π. Now every set in π has size in the range $[\log n, 2\log n]$. We fix a small $\delta > 0$ and consider two alternatives:

- *The average size of a set in π is larger than* $(1 + \delta)\log n$. We extend π to become a partition of the set $[n]$ by including all remaining coordinates as singleton sets. Let $r = |\pi|$. We have

$$r \leq \frac{n}{(1 + \delta)\log n} + 3k + \log n \lesssim \frac{n}{(1 + \delta)\log n}.$$

  We claim that π is attractive for L. Observe that every element $\mathbf{v} \in L$ intersects each non-singleton element of π in an even number of coordinates. Therefore we have

$$\left| \bigsqcup_{j \,:\, \mathbf{v} \cap \pi_j \neq \emptyset} \pi_j \right| \leq \frac{\mathrm{wt}(\mathbf{v})}{2} \cdot 2\log n = \mathrm{wt}(\mathbf{v}) \cdot \log n.$$

  It remains to note that

$$\mathrm{wt}(\mathbf{v}) \cdot \frac{n}{r} \gtrsim (1 + \delta) \cdot \mathrm{wt}(\mathbf{v}) \cdot \log n.$$

- *The average size of a set in π is below* $(1+\delta)\log n$. The fraction of sets of size above $1.5\log n$ is at most 2δ. We pair up sets of size less than $1.5\log n$ arbitrarily (possibly dropping one set). We replace pairs by their unions. This leads us to a new family of sets, where the size of each set is in the range $[1.5\log n, 3\log n]$ and the average size is above $1.5(1 + \delta')\log n$. Here we apply the argument from the previous bullet.

This concludes the proof. $\blacksquare$

6.2   Approximating general $\mathbb{F}_2$-linear spaces

We now apply the method used in the previous Section to approximate cut spaces to generic linear spaces.

Theorem 20 *Let $L \subseteq \mathbb{F}_2^n$ be a linear space,* $\dim L = k$, $k \leq n^{\frac{1}{2} - \beta}$. *Then for some $r = \Theta\left(\frac{n}{k}\log k\right)$ and $\alpha > 0$,*

$$A_r(L) \geq (1 + \alpha)\frac{r}{n}. \tag{30}$$

Proof: Fix $\epsilon > 0$ such that $k^{1+\epsilon} = o\left(\frac{n}{k}\log k\right)$. We rely on the fact that by the Hamming bound for any linear subspace of $\mathbb{F}_2^m$ of dimension k, there is a dual vector of weight at most $ck/\log k$ provided $k^{1+\epsilon} \leq m$, for a universal constant c. We construct a family π of disjoint subsets of $[n]$. We start by executing the following simple algorithm:

1. Start with an empty family of sets π.

2. WHILE $n \geq k^{1+\epsilon}$ DO

3.   BEGIN

4.   Identify a light dual vector $\mathbf{v}$, $\mathrm{wt}(\mathbf{v}) \leq \frac{ck}{\log k}$;

5.   Include $\mathrm{supp}(\mathbf{v})$ into π as a new set;

6. Drop the coordinates in $\mathrm{supp}(\mathbf{v})$ from $[n]$. Reduce n appropriately.

7. END

Our next goal is to make sure that most sets in π have approximately the same size. Firstly, we repeatedly join together any two sets in π, if the sum of their sizes is below $2ck/\log k$. Secondly, we drop the smallest set from π. Now every set in π has size in the range $[ck/\log k, 2ck/\log k]$. We fix a small $\delta > 0$ and consider two alternatives:

- *The average size of a set in π is larger than* $(1+\delta)ck/\log k$. We extend π to become a partition of the set $[n]$ by including all remaining coordinates as singleton sets. Let $r = |\pi|$. We have

$$r \leq \frac{n\log k}{(1+\delta)ck} + k^{1+\epsilon} + \frac{ck}{\log k} \lesssim \frac{n\log k}{(1+\delta)ck}.$$

  We claim that π is attractive for L. Observe that every element $\mathbf{v} \in L$ intersects each non-singleton element of π in an even number of coordinates. Therefore we have

$$\left| \bigsqcup_{j\,:\,\mathbf{v}\cap\pi_j\neq\emptyset} \pi_j \right| \leq \frac{\mathrm{wt}(\mathbf{v})}{2} \cdot \frac{2ck}{\log k} = \mathrm{wt}(\mathbf{v}) \cdot \frac{ck}{\log k}.$$

  It remains to note that

$$\mathrm{wt}(\mathbf{v}) \cdot \frac{n}{r} \gtrsim \mathrm{wt}(\mathbf{v}) \cdot (1+\delta) \cdot \frac{ck}{\log k}.$$

- *The average size of a set in π is below* $(1+\delta)ck/\log k$. The fraction of sets of size above $1.5ck/\log k$ is at most 2δ. We pair up sets of size less than $1.5ck/\log k$ arbitrarily (possibly dropping one set). We replace pairs by their unions. This leads us to a new family of sets, where the size of each set is in the range $[1.5ck/\log k, 3ck/\log k]$ and the average size is above $1.5(1+\delta')ck/\log k$. Here we apply the argument from the previous bullet.

This concludes the proof. ∎

Similarly to Theorem 15, Theorem 20 above can be used to argue that Conjecture 7 holds for all linear spaces of dimension up to $o(n^\epsilon \log n)$, i.e., if we set $k = \beta(n) \cdot n^\epsilon \log n$, where $\beta(n) \to 0$, and $r = \Theta\left(\frac{n}{k}\log n\right) = \Theta\left(\frac{n^{1-\epsilon}}{\beta}\right)$; then (30) yields

$$A_r(L) \geq (1+\alpha)\frac{r}{n}. \tag{31}$$

Theorem 15 however presents a stronger result. Firstly, in the proof of Theorem 15 we use real spaces of dimension as low as n^τ to arrive at the bound (31) for some $r = \omega(n^{1-\epsilon})$. This leaves plenty of room for potential further improvements. Secondly, Theorem 15 applies to all sets of vectors of bounded triangular rank, while Theorem 20 only deals with $\mathbb{F}_2$-linear spaces.

7 Relation to natural proofs

In [16] Razborov and Rudich introduced the natural proofs barrier for proving lower bounds for computational complexity of Boolean functions. Stated informally their results say that if a certain property of Boolean functions is shown to imply hardness then; either typical (random) functions do not have this property, or the property should be hard to recognize, or some well accepted hardness conjectures are invalid. While the theory developed in [16] deals with properties of Boolean functions one can make an analogy in the linear setting by defining a natural

property of a matrix to be a property that holds for most matrices and can be verified efficiently [1]. In this case however the respective "hardness conjectures" are not as standard.

In light of the above it is interesting to ask if establishing our main Conjecture 7 for some ϵ would necessarily certify rigidity of random binary matrices where each element is set to 1 independently with probability $n^{-\epsilon}$. The answer to this seems to depend on the value of α for that one proves the Conjecture. If α is sufficiently small; then rigidity of random matrices would likely not be implied. The following Theorem shows that unlike the rows of a combinatorial design $V_{1/\epsilon}$, the rows of a random binary matrix of density $n^{-\epsilon}$ with high probability admit a non-trivial approximation on average even in the regime of a fairly large dimension of the approximating real space.

Theorem 21 *Let V be a random $n \times n$ matrix of zeros and ones where every entry is set to 1 independently with probability $p = n^{-\epsilon}$, $\epsilon < \frac{1}{4}$. Let B be the real matrix, where the rows of B are the normalized elements of V. Let $\lambda_1 \geq \ldots \geq \lambda_n \geq 0$ be the eigenvalues of BB^t. There exists $\alpha > 0$ such that with high probability over the choice of V for all $r = o(n)$ we have,*

$$\max_{\dim W \leq r} \mathbb{E}_{\mathbf{v} \in V} \|\mathrm{Pr}_W(\mathbf{v})\|^2 = \frac{1}{n} \sum_{i \leq r} \lambda_i \geq (1+\alpha)\frac{r}{n}. \tag{32}$$

Proof: The equality above is given by (10). We need to prove the inequality, which is a straightforward corollary of the Marchenko-Pastur law [4] determining the limiting behavior of the spectral distribution of large inner product matrices. We state a special case of this law that will suffice for our purposes (see Theorem 3.10 in [4]).

Let $\{M_n\}$ be a sequence of $n \times n$ random matrices, such that the entries of M_n are i.i.d. random variables with expectation μ_n and variance 1. Let $\lambda_1, \ldots, \lambda_n$ be the eigenvalues of $\frac{1}{n}M_n M_n^t$. Then, for any $0 \leq a \leq 4$ holds, with probability 1, that

$$\frac{\#\{i : \lambda_i \geq a\}}{n} \longrightarrow \frac{1}{2\pi} \int_a^4 \sqrt{\frac{4-x}{x}} \, dx \tag{33}$$

Fix $a = 2$. Let $c = \frac{1}{2\pi} \int_2^4 \sqrt{\frac{4-x}{x}} dx = \frac{\pi - 2}{2\pi} \approx 0.18$. Taking $M_n = \frac{1}{\sqrt{p(1-p)}} \cdot V$, we observe that, with probability tending to one with n, at least $cn - o(n)$ eigenvalues of $\frac{1}{p(1-p)n}VV^t$ are greater or equal 2.

To complete the proof, we will argue that the spectral distribution of $\frac{1}{p(1-p)n}VV^t$ is close to that of BB^t. In fact, a special case of the perturbation inequality A.41 in [4] states that for any two symmetric $n \times n$ matrices X and Y, with eigenvalues $\lambda_1, \ldots, \lambda_n$ and $\mu_1, \ldots, \mu_n$, and for any real number a holds that

$$\left| \frac{\#\{i : \lambda_i \geq a\}}{n} - \frac{\#\{i : \mu_i \geq a\}}{n} \right| \leq \left(\frac{1}{n} \cdot Tr\left((X-Y)^2 \right) \right)^{1/3} \tag{34}$$

We will apply this inequality to slightly perturbed versions of the matrices $\frac{1}{p(1-p)n}VV^t$ and BB^t. The goal of this modification would be to cancel the problematic effect of the maximal eigenvalues of the two matrices. With this in mind, we set $U = V - pJ$, where J is the all-1 matrix. Let $W = (w_{ij}) = UU^t$. We take $X = \frac{1}{p(1-p)n}W$.

Next, we consider the norms of the rows of V. Since V is a 0-1 matrix, the norm of its ith row is $\sqrt{r_i}$, where r_i is the row sum. We take $Y = \left(\frac{w_{ij}}{\sqrt{r_i r_j}} \right)$.

Note that the matrices X and Y are rank-1 perturbations of $\frac{1}{p(1-p)n}VV^t$ and BB^t. It is well-known that if two matrices differ by a matrix of rank 1, their eigenvalues interlace. Hence, using the perturbed matrices changes the LHS of (34) by at most an additive factor of $O(1/n)$, which we may ignore.

In the following analysis we may and will assume all r_i to lie in the interval $np \pm t\sqrt{np \log n}$, for a sufficiently large absolute constant t, since this holds with probability tending to one with n, by the Chernoff bound.

With this assumption, we can bound the distance between the entries of X and Y as follows.

$$(x_{ij} - y_{ij})^2 = \left(\frac{w_{ij}}{p(1-p)n} - \frac{w_{ij}}{\sqrt{r_i r_j}} \right)^2 \leq O\left(\frac{1}{n^2} \right) \cdot w_{ij}^2$$

In the last inequality we have used the fact that $np^2 \gg \sqrt{n \log n}$, which we may do since $p = n^{-\epsilon}$ and, by assumption, $\epsilon < 1/4$.

To complete the argument about proximity of the spectral distributions of X and Y, we need to estimate from above the ℓ_2 norm of W. Note first that by (33) we may assume all the eigenvalues of X to lie between 0 and 4. Hence $\sum_{i,j=1}^n x_{ij}^2 = O(n)$. Since $W = p(1-p)n \cdot X$, we deduce $\sum_{i,j=1}^n w_{ij}^2 \leq n^2 p^2 \cdot \sum_{i,j=1}^n x_{ij}^2 \leq O\left(n^3 p^2 \right)$. Finally, we can estimate the RHS of (34) from above as follows:

$$\left(\frac{1}{n} \cdot Tr\left((X-Y)^2 \right) \right)^{1/3} \leq O\left(\frac{1}{n} \right) \cdot \left(\sum_{i,j=1}^n w_{ij}^2 \right)^{1/3} \leq O\left(p^{2/3} \right) = o(n)$$

We deduce that at least $cn - o(n)$ eigenvalues of the matrices Y and (hence) BB^t are greater or equal 2, for an absolute constant $c \approx 0.18$. The claim of the theorem follows. ∎

8 Conclusions

In this paper we suggested a new path to establishing rigidity of design matrices over the field $\mathbb{F}_2$. Our approach is centered around the conjecture that says that after the natural "normalizing" embedding of the Boolean cube into $\mathbb{R}^n$, low dimensional $\mathbb{F}_2$-linear spaces exhibit some tiny amount of resemblance to real linear spaces. In particular it is easier to approximate them by Euclidian linear spaces than to approximate all of the Boolean cube. We showed that the conjecture is indeed true (by a huge margin) when approximating real spaces are of low dimension. However our approximability results for high-dimensional real spaces are not strong enough.

Currently it feels that the weakness of our results stems from the fact in that we use relatively little combinatorial structure of $\mathbb{F}_2$-linearity. In particular our strongest result (Theorem 15) applies to all sets of bounded triangular rank. Note that while it is plausible that one can make further progress based just on triangular rank; one cannot establish Conjecture 7 is such generality.

Remark 22 Replacing the condition $\dim(L) \leq n^{2\epsilon+\delta}$ in the Conjecture 7 by the condition $\mathrm{trk}(L) \leq n^{2\epsilon+\delta}$ makes the Conjecture invalid.

Proof: Let $m = \frac{1}{\epsilon}$ be an integer. Let V_m' be a matrix that is obtained from V_m by independently flipping every zero entry to one with probability $n^{-2\epsilon}$. It is not hard to see that with overwhelming probability V_m' does not contain an all-zeros minor of size $\Omega(n^{2\epsilon} \log n)$. Thus $\mathrm{trk}(V_m') = O(n^{2\epsilon} \log n)$. However Proposition 5 implies that for any $r = \omega(n^{1-\epsilon})$, we have $A_r(V_m') \approx \frac{r}{n}$. Thus rows of V_m' give a counterexample to this stronger version of the Conjecture. ∎

Acknowledgement

References

[1] Michael Alekhnovich. More on average case vs. approximation complexity. In *44th IEEE Symposium on Foundations of Computer Science (FOCS)*, pages 298–307, 2003.

[2] Noga Alon and Gil Cohen. On rigid matrices and U-polynomials. In *28th IEEE Computational Complexity Conference (CCC)*, pages 197–206, 2013.

[3] Noga Alon, Rina Panigrahy, and Sergey Yekhanin. Deterministic approximation algorithms for the nearest codeword problem. In *13th International Workshop on Randomization and Computation (RANDOM)*, volume 5687 of Lecture Notes in Computer Science, pages 339–351, 2009.

[4] Zhidong Bai and Jack Silverstein. *Spectral Analysis of Large Dimensional Random Matrices*. Springer, New York, 2010.

[5] Zeev Dvir. On matrix rigidity and locally self-correctable codes. *Computational Complexity*, 20:367–388, 2011.

[6] Joel Friedman. A note on matrix rigidity. *Combinatorica*, 13:235–239, 1993.

[7] Noboru Hamada. On the p-rank of the incidence matrix of a balanced or partially balanced incomplete block design and its applications to error-correcting codes. *Hiroshima Mathematical Journal*, 3:154–226, 1973.

[8] R. S. Ismagilov. n-dimensional width of compact in Hilbert space. *Journal of Functional Analysis and Applications*, 2:32–39, 1968.

[9] Stasys Jukna. *Extremal combinatorics*. Springer, Berlin, Heidelberg, New York, 2001.

[10] Dieter Jungnickel and Vladimir Tonchev. Polarities, quasi-symmetric designs, and Hamada conjecture. *Designs, Codes, and Cryptography*, 51:131–140, 2009.

[11] Boris Kashin and Alexander Razborov. Improved lower bounds on the rigidity of Hadamard matrices. *Mathematical Notes*, 63:471–475, 1998.

[12] Satyanarayana Lokam. Spectral methods for matrix rigidity with applications to size-depth trade-offs and communication complexity. *Journal of Computer and System Sciences*, 63:449–473, 2001.

[13] Satyanarayana Lokam. Complexity lower bounds using linear algebra. *Foundations and Trends in Theoretical Computer Science*, 4:1–155, 2009.

[14] Ilan Newman and Yuri Rabinovich. On multiplicative λ-approximations and some geometric applications. *SIAM Journal on Computing*, 42:885–883, 2013.

[15] Alexander Razborov. On rigid matrices, 1989. Manuscript. In Russian.

[16] Alexander Razborov and Steven Rudich. Natural proofs. *Journal of Computer and System Sciences*, 55:24–35, 1997.

[17] Shubhangi Saraf and Sergey Yekhanin. Noisy interpolation of sparse polynomials, and applications. In *26th IEEE Computational Complexity Conference (CCC)*, pages 86–92, 2011.

[18] Rocco Servedio and Emanuele Viola. On a special case of rigidity. In *Electronic Colloquium on Computational Complexity (ECCC)*, TR12-144, 2012.

[19] Amin Shokrollahi, Daniel Speilman, and Voelker Stemann. A remark on matrix rigidity. *Information Processing Letters*, 64:283–285, 1997.

[20] Kempton Smith. On the p-rank of the incidence matrix of points and hyperplanes in a finite projective geometry. *Journal of Combinatorial Theory*, 7:122–129, 1969.

[21] Vladimir Temlyakov. Nonlinear Kolmogorov widths. *Mathematical Notes*, 63:785–795, 1998.

[22] Kirill Uskov. *Kolmogorov width of geometric configurations and functional compacts in Hilbert spaces*. PhD thesis, Moscow State Technical University, 2002. in Russian.

[23] Leslie Valiant. Graph-theoretic arguments in low level complexity. In *6th Symposium on Mathematical Foundations of Computer Science (MFCS)*, pages 162–176, 1977.