KOLMOGOROV WIDTH OF DISCRETE LINEAR SPACES: AN APPROACH TO MATRIX RIGIDITY

Alex Samorodnitsky, Ilya Shkredov, and Sergey Yekhanin

February 12, 2016

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.

Keywords. Matrix rigidity, linear codes, Kolmogorov width.

Subject classification. F.1.2; E.4

1. Introduction

The notion of matrix rigidity was introduced by Leslie Valiant in Valiant (1977). 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 (1977) 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 R, are given in Razborov (1989); Servedio & Viola (2012). Reductions above naturally lead to the challenge of constructing rigid matrices explicitly. After more than three decades of efforts, however, this challenge remains elusive Lokam (2009).

None of the existing techniques for constructing rigid matrices Alon & Cohen (2013); Alon *et al.* (2009); Dvir (2011); Kashin & Razborov (1998); Lokam (2001); Saraf & Yekhanin (2011) surpasses the basic *untouched minor* argument of Shokrollahi *et al.* (1997) that amounts to taking a matrix where every minor has full rank, and using the bound from the Zarankiewicz problem (Jukna 2001, 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 $\log^2 n \leq R \leq n/2$. Similar parameters are known to be attainable over small finite fields Friedman (1993). Rigidity parameters above are vastly weaker the parameters of random matrices. In particular, it is not hard to show that a random matrix over any field is at least

$$\left(\frac{n}{2}, \Omega(n)\right)$$

rigid with a very high probability.

1.1. Combinatorial designs. A family $\mathcal{F}$ of w-subsets of a universe of size n is called an (n, w, λ) combinatorial 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

$$n = \frac{q^{m+1} - 1}{q - 1} \qquad w = \frac{q^m - 1}{q - 1} \qquad \lambda = \frac{q^{m-1} - 1}{q - 1}. \tag{1.1}$$

With a slight abuse of notation we write $G_{m,q}$ or just G_m to denote both geometric designs and their incidence matrices, (e.g., binary matrices whose rows / columns correspond to points / hyperplanes in $\mathbb{PG}(m + 1, q)$, and that contain a one in location (i, j) iff the point corresponding to the i-th row is contained in the hyperplane corresponding to the j-th column.)

In his original paper (Valiant 1977, 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. Let rank_p denote matrix rank over the field $\mathbb{F}_p$. By Smith (1969) rank $\mathrm{rank}_p G_{m,q}$ is given by

$$\begin{cases} n & \text{if } q \neq p^e, \quad w + (n-1)\lambda \neq 0 \bmod p; \\ n - 1 & \text{if } q \neq p^e, \quad w + (n-1)\lambda = 0 \bmod p; \\ \binom{p+m-1}{m}^e & \text{if } q = p^e. \end{cases} \tag{1.2}$$

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

$$\mathrm{rank}_2 G_{m,q} = \Theta\left(n^{\frac{\log_2(m+1)}{m}}\right). \tag{1.3}$$

Identity (1.3) 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's 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 due to Noboru Hamada (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's conjecture (Jungnickel & Tonchev 2009, 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 Jungnickel & Tonchev (2009).)

One easier natural question to ask that fits well with our approach to rigidity is whether one can prove any non-trivial 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's conjecture and identity (1.2) suggest that

$$\mathrm{rank}_p V_m \geq \Omega\left(n^{\frac{\log_p \binom{p+m-1}{m}}{m}}\right). \tag{1.4}$$

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. See formula (3.1) for a formal definition. Thus sets with large value of A_r are precisely those that are well approximated by some r-dimensional Euclidian linear space.

Now let $m = \frac{1}{\epsilon}$. We argue that for all values of $r = \omega(n^{1-\epsilon})$, the approximability measure $A_r(V_m) \approx A_r\left(\mathbb{F}_2^n\right)$. 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 $A_r(L) \geq (1+\alpha)A_r\left(\mathbb{F}_2^n\right)$, 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 $A_r(L) \geq (1+\alpha)A_r\left(\mathbb{F}_2^n\right)$, for some $r = \omega(n^{1-\epsilon})$. Currently, our main Theorem 5.12 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 5.14) already suffice to get the bound

$$\operatorname{rank}_p V_m \geq \Omega\left(n^{\frac{1}{m}} \log_2 n\right), \tag{1.5}$$

for all values of p, a result 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 bounds for $A_1(L)$ and $A_3(L)$. The proof of the latter bound is deferred to Section 7. In Section 5 we deal with high dimensional approximations and state the implications of our results for the Hamada's 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 Newman & Rabinovich (2013) (see Definition 4.1).

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 8 we discuss the relation of our approach to the natural proofs lower bounds barrier of Razborov & Rudich (1997).

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 $\operatorname{supp}(\mathbf{v})$;
- We write $f(n) \approx g(n)$, if $f(n) = g(n)(1 + o(1))$. We adopt the same agreement for $\lesssim, \gtrsim$.

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 $\mathbb{F}_2^n$ 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

$$A_r(V) = \max_{\dim W \leq r} \min_{\mathbf{v} \in V} \|\mathrm{Pr}_W(\mathbf{v})\|^2, \tag{3.1}$$

where the maximum is over all linear spaces $W \in \mathbb{R}^n, \dim W = r$, the minimum is over the non-zero elements of V, and $\mathrm{Pr}_W(\mathbf{v})$ denotes the projection of $\mathbf{v}$ onto W. 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 Temlyakov (1998); Uskov (2002).

We remark the importance of the normalization step in formula (3.1). In essence normalizing elements of V pushes all high weight vectors in V to the center of the positive orthant in $\mathbb{R}^n$, and makes $A_r(V)$ governed by the distribution of low weight vectors in V as these vectors are still pointing in different directions. This is a desirable feature, as our final goal is to argue rigidity of design matrices, whose rows have (relatively) low weight. Thus these are low weight vectors in linear spaces that interest us.

We now derive a formula for approximability measure of the whole Boolean cube.

LEMMA 3.2. *Let $r = o(n)$ be arbitrary. We have*

$$A_r\left(\mathbb{F}_2^n\right) \approx \frac{r}{n}. \tag{3.3}$$

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

$$A_r\left(\mathbb{F}_2^n\right) \leq A_r\left(\{\mathbf{e}_1, \ldots, \mathbf{e}_n\}\right) \leq \frac{r}{n}. \tag{3.4}$$

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 (3.4) 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

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

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 3.5. *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,*

$$A_r(V) \leq \frac{1}{n} \sum_{i \leq r} \lambda_i. \tag{3.6}$$

PROOF. This is a simple corollary of a result in Ismagilov (1968). A special case of this result states that for all r

$$\frac{1}{n} \sum_{i \leq r} \lambda_i = \max_{\dim W \leq r} \mathbb{E}_{\mathbf{v} \in V} \|\mathrm{Pr}_W(\mathbf{v})\|^2, \tag{3.7}$$

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 (1.1) 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 3.5 yields

COROLLARY 3.8. *With the notation above, we have*

$$A_r\left(V_{1/\epsilon}\right) \lesssim \frac{n^{1-\epsilon} + r}{n}. \tag{3.9}$$

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[1]

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

An application of Lemma 3.5 completes the proof. ■

Combining (3.3) and (3.9), we conclude that for $r = \omega(n^{1-\epsilon})$ and $r = o(n)$,

$$A_r(V_{1/\epsilon}) \approx A_r(\mathbb{F}_2^n) \approx \frac{r}{n}. \tag{3.10}$$

Observe that identity (3.7) 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 3.11. *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,*

$$A_r(V') \leq \left(\sqrt{A_r(V)} + \sqrt{\frac{d}{w}} \right)^2. \tag{3.12}$$

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

$$(\mathbf{v}, \mathbf{v}') \geq \sqrt{1 - \frac{d}{w}}. \tag{3.13}$$

The minimum in (3.13) is attained by a vector $\mathbf{v}'$ of Hamming weight $w - d$, where the support of $\mathbf{v}'$ is a subset of the support

[1] The first eigenvalue $\lambda_1 \approx n^{1-\epsilon}$ corresponds to the obvious eigenvector $v_1 = (1, \ldots, 1)$. Since every other eigenvector $\{v_i\}_{i \geq 2}$ satisfies $v_1 \perp v_i$, we can obtain eigenvalues $\{\lambda_i\}_{i \geq 2}$ by considering the matrix $BB^t - \frac{\lambda}{w} J = \frac{w - \lambda}{w} I$. At which point it becomes clear that $\lambda_2 = \ldots = \lambda_n \approx 1$.

of $\mathbf{v}$. Let W be an r-dimensional real space in $\mathbb{R}^n$ that attains the maximum in (3.1) 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 $i \in [n]$, $(\mathbf{v}_i', \mathbf{w}_i)^2 \geq A_r\left(V'\right)$. Let $A = \sqrt{A_r(V)}$. By definition of A there exists $i \in [n]$ such that

$$|(\mathbf{v}_i, \mathbf{w}_i)| \leq A. \tag{3.14}$$

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

$$\begin{aligned}(\mathbf{v}_i', \mathbf{w}_i) = \cos\alpha &\leq \cos(\gamma - \beta) \\ &= \cos\gamma\cos\beta + \sin\gamma\sin\beta \\ &\leq |\cos\gamma| + \sin\beta \\ &\leq A + \sqrt{\tfrac{d}{w}},\end{aligned}$$

where the last inequality follows from (3.13) and (3.14). 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 (3.12) follows. ■

The above lemma and identity (3.10) yield

Proposition 3.15. *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 $r = \omega(n^{1-\epsilon})$ and $r = o(n)$. We have*

$$A_r\left(V_{1/\epsilon}'\right) \approx \frac{r}{n}. \tag{3.16}$$

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 3.17. *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* $\mathrm{rank}_{\mathbb{F}} V \geq R$.

CONJECTURE 3.18. *There exists positive constants α, δ, and $\epsilon = \frac{1}{m}$ (for an integer m) such that for all linear spaces $L \subseteq \mathbb{F}_2^n$ where* $\dim L \leq n^{2\epsilon+\delta}$, *for some $r = \omega(n^{1-\epsilon})$,*

$$A_r(L) \geq (1+\alpha)\frac{r}{n}. \tag{3.19}$$

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 3.15. Currently we can only prove the conjecture for all linear spaces L with $\dim L = o\left(n^{\epsilon}\log n\right)$. (Theorem 5.12).

REMARK 3.20. *Replacing the condition* $\dim(L) \leq n^{2\epsilon+\delta}$ *in the Conjecture above by the condition* $\dim(L) \leq O\left(n^{\epsilon\log\left(\frac{1}{\epsilon}+1\right)}\right)$ *makes the Conjecture invalid as by formula (1.3) matrices $V_{1/\epsilon}$ may have have $\mathbb{F}_2$ rank of $O\left(n^{\epsilon\log\left(\frac{1}{\epsilon}+1\right)}\right)$.*

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. We first obtain a lower bound $A_1(L)$ and then present a stronger lower bound for $A_3(L)$ deferring the proof to Section 7. 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 4.1. *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 $j \leq t$:*

$$\mathrm{supp}(\mathbf{v}_j) \not\subseteq \bigcup_{s \leq j-1} \mathrm{supp}(\mathbf{v}_s).$$

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

The term triangular rank is explained by following observation. Consider a binary matrix M, whose rows are elements of V. Triangular rank of V is exactly the size of the largest square upper-triangular minor of M (after arbitrary row / column permutations) that has ones on the diagonal.

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

$$w_i = \min_{\mathbf{v} \in L \ : \ i \in \mathrm{supp}(\mathbf{v})} |\mathrm{supp}(\mathbf{v})|. \tag{4.2}$$

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

$$\mu = \sum_{i=1}^{n} w_i^{-1}. \tag{4.3}$$

Our proof of the following lemma resembles some of the arguments in (Newman & Rabinovich 2013, Section 2).

Lemma 4.4. *Let $L \subseteq \mathbb{F}_2^n$, $\mathrm{trk}(L) \leq k$; then*

$$\mu \leq k,$$

where μ is defined by (4.3).

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 $i_j \in \mathrm{supp}(\mathbf{v}_j)$, such that $i_j \notin \mathrm{supp}(\mathbf{v}_s)$, 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 $i \in \mathrm{supp}(\mathbf{v}_j)$;
3. We set

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

   It is important to note that Δ is necessarily at most 1 since all w_s^{-1} in the above sum are at most $w_i^{-1} = 1/|\mathrm{supp}(\mathbf{v}_j)|$.

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

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

Theorem 4.5. *Let* $L \subseteq \mathbb{F}_2^n$, $\mathrm{trk}(L) \leq k$; *then*

$$A_1(L) \geq \frac{1}{k}. \tag{4.6}$$

Proof. Let $\{w_i\}$ and μ be as defined in (4.2) and (4.3). 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, $|\mathrm{supp}(\mathbf{v})| = w$. Note that for all $i \in \mathrm{supp}(\mathbf{v})$, $w_i \leq w$. 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 4.4. ■

Theorem 4.5 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 3.18 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. To see this set $k = \beta(n) \cdot n^\epsilon$, for an arbitrary nonzero $\beta(n)$ that goes to zero as n grows. Further, set $r = n^{1-\epsilon}/2\beta(n)$. By Theorem 4.5, for all L, with $\operatorname{trk}(L) \leq k$ we have

$$A_r(L) \geq A_1(L) \geq \frac{1}{k} = \frac{1}{\beta n^\epsilon} = 2 \cdot \frac{r}{n}.$$

The approximation measure $A_r(L)$ is obviously non-decreasing with r. In fact it seems natural to expect that $A_r(L)$ should grow rapidly at least when r is small. The next theorem partly confirms this intuition showing that $A_3(L)$ is indeed a constant fraction larger than the bound (4.6). (We do not have a result like that for $A_2(L)$.) We defer the proof to Section 7.

Theorem 4.7. *There exist positive constants δ and k_0 such that for all $k \geq k_0$, for all sets $L \subseteq \mathbb{F}_2^n$, $\operatorname{trk}(L) \leq k$:*

$$A_3(L) \geq \frac{1+\delta}{k}. \tag{4.8}$$

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's conjecture. Our main result is given by Theorem 5.12. 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 $\operatorname{supp}(\mathbf{v})$.

Definition 5.1. *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* $\operatorname{trk}(L) \leq k$ *and for all* $\mathbf{v} \in L$ *such that* $\mathbf{v} \cap S \neq \emptyset$,

$$|\mathbf{v} \cap S| \geq c \cdot \frac{|S|}{k}. \tag{5.2}$$

Informally, an attractor is a subset of coordinates such that every low weight vector (i.e., a vector of weight around $\frac{n}{k}$ or less) in L whose support intersects the subset, intersects it by more than one would expect. Below is the key lemma of this section.

Lemma 5.3. *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

$$\frac{N}{2^{4c+2}} \geq w. \tag{5.4}$$

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

```
Set the tower $T$ to be an empty family of sets. Set $R = [N]$.
WHILE $R \neq \emptyset$ DO
    BEGIN
        Identify $\mathbf{v} \in L$ yielding the smallest non-zero $|\mathbf{v} \cap R|$;
        Add $\mathbf{v}$ to the tower $T$;
        Drop the elements in $\mathbf{v}$ from $R$;
    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 $\Delta_j = \mathbf{v} \cap R$. We partition the steps of the algorithm into *stages*. A step falls into stage number i if in the beginning of the step

$$\frac{N}{2^i} < |R| \leq \frac{N}{2^{i-1}}. \tag{5.5}$$

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 (5.4). 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. As step j belongs to stage i, by (5.5) we have

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

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 5.6. *Let L be a family of binary vectors, $\operatorname{trk}(L) \leq k$. 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*

$$A_{c \cdot 2^{4c}}(L) \geq \Omega\left(\frac{c}{k}\right), \tag{5.7}$$

where the constant in Ω-notation is absolute.

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

```
 1. $\pi = \{\pi_i\}_{i \geq -1}$ is a partition of $[N]$. Initially $\pi$ consists of two
    sets: $\pi_{-1} = \emptyset$, $\pi_0 = [N]$ and $i = 0$.
 2. WHILE  $((i < c \cdot 2^{4c})$ AND $(\pi_0 \neq \emptyset))$ DO
 3.     BEGIN
 4.       Find a $(c,k)$-attractor $\pi_i$ for $L$ of relat. size at least $\frac{1}{2^{4c};}$
 5.       Remove elements of $\pi_i$ from $\pi_0$ and from all sets $\mathbf{v} \in L$;
 6.         Add the set $\pi_i$ to $\pi$;
 7.         Drop every element $\mathbf{v}$ such that $|\mathbf{v}| < w$ from $L$;
 8.         Move elements of $\pi_0$ with no support in $L$ to $\pi_{-1}$;
 9.         Increment $i$;
10.     END
```

Our algorithm above maintains the invariant that the union of supports of elements of L is π_0 and every element of L has weight in the segment $[w, 2w]$. Thus in step 4, we can safely invoke Lemma 5.3 to obtain an the attractor. Observe that by the end of the execution of the algorithm we have

$$|\pi_0| \leq N \cdot \left(1 - \frac{1}{2^{4c}}\right)^{c \cdot 2^{4c}} \leq \frac{N}{c}. \tag{5.8}$$

Recall that $\pi = \{\pi_i\}_{i \geq -1}$ is a partition of $[N]$. Let $r = |\pi| - 2$. Let W be the real linear space spanned by binary vectors $\mathbf{p}_0, \ldots, \mathbf{p}_r$

whose supports are the corresponding elements of this partition. Clearly, $\dim W \le c \cdot 2^{4c}$. 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. Let $i \in [r]$ be the index of the step at that $\mathbf{v}$ has been removed from L; if $\mathbf{v}$ has not been removed from L by the end of the execution we set $s = r$. To approximate $\mathbf{v} \in L$ consider the set

  $$J = \{j \in [1,s] \mid \mathbf{v} \cap \pi_j \neq \emptyset\}.$$

  For all $j \in J$ by (5.2) we have

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

  Therefore

  $$|\pi_j| \le |\mathbf{v} \cap \pi_j| \cdot \frac{k}{c}. \tag{5.9}$$

  Consider the vector $\mathbf{p} = \sum_{j \in J} \mathbf{p}_j$. Summing (5.9) over all $j \in J$ we obtain

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

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

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

- ○ $|\mathbf{v} \cap \pi_0| \ge w$. By (5.8) we have $|\pi_0| \le \frac{2wk}{c}$. Consider the binary real vector $\mathbf{p}$ whose support is π_0. We have

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

■

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

THEOREM 5.10. *Let* $L \subseteq \mathbb{F}_2^n$, $\operatorname{trk}(L) \le k$. *Assume* $k \ge 2^{5c+2}$ *where* c *is an integer. We have*

$$A_{2c\cdot 2^{4c}\cdot \log n}\,(L) \ge \Omega\left(\frac{c}{k}\right), \tag{5.11}$$

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 5.6 to each L_i. Consider the joint span of $2\log n$ resulting real spaces to approximate L. ■

THEOREM 5.12. *Let* $L \subseteq \mathbb{F}_2^n$, $\operatorname{trk}(L) \le k$; *then*

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

$$A_{n^\tau}(L) \ge \Omega\left(\frac{\log k}{k}\right), \tag{5.13}$$

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 3.18 holds for all linear spaces* L, *where* $\dim L = o(n^\epsilon \log n)$.

PROOF. We start with the first bullet. Set

$$c = \left\lfloor \min\left\{\frac{\tau}{8}\log k, \frac{\log k - 2}{5}\right\}\right\rfloor,$$

which ensures that $k \ge 2^{5c+2}$. Theorem 5.10 yields

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

which immediately yields (5.13) 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 (5.13)

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

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 5.12 for the triangular rank of combinatorial designs.

COROLLARY 5.14. *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*

$$\operatorname{trk}(V_m) = \Omega\left(n^\epsilon \log n\right), \tag{5.15}$$

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

PROOF. Let $\operatorname{trk}(V_m) = k$. Set $r = n^{1-\epsilon}$. From Corollary 3.8 we have $A_r(V_m) \lesssim \frac{2}{n^\epsilon}$. However from Theorem 5.12 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 $k = \Omega\left(n^\epsilon \log n\right)$. ■

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 3.18 for a certain class a linear spaces called cut-spaces and give a simpler proof of a slightly weaker version of Theorem 5.12. 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 $\operatorname{supp}(\mathbf{v})$.

Definition 6.1. *Let L be a family of subsets of $[n]$ and α be a positive constant. We say that the r-partition $\pi = \bigsqcup_{j \le r} \pi_j$ of $[n]$ is α-attractive for L, if for every $\mathbf{v} \in L$ we have*

$$\left| \bigsqcup_{j \,:\, \mathbf{v} \cap \pi_j \neq \emptyset} \pi_j \right| \le \frac{|\mathrm{supp}(\mathbf{v})|}{(1+\alpha)} \cdot \frac{n}{r}. \tag{6.2}$$

Lemma 6.3. *Let L be a set of n-dimensional binary vectors. Suppose there exists an r-partition of $[n]$ that is α-attractive for L; then $A_r(L) \ge (1+\alpha)\frac{r}{n}$.*

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 α-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

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

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

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 3.18.

Theorem 6.4. *Let $L \subseteq \mathbb{F}_2^n$ be a cut space, $\dim L \le o\left(\frac{n}{\log n}\right)$. Then for some $r = \Theta\left(\frac{n}{\log n}\right)$ and $\alpha > 0$ we have*

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

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 $n \geq 3k$.[2] 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 $\pi$.
2. WHILE $n \geq 3k$ DO
3.     BEGIN
4.         Identify a cycle $C$ in $G$, $|C| \leq 2\log n$;
5.         Include $C$ into $\pi$ 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 α-attractive for L, for a positive α. Observe that every element $\mathbf{v} \in L$ intersects each non-singleton element of π in an even number of coordinates, since every

[2] While this fact is standard, we include a proof sketch for a seeming lack of a good reference. Consider a k-node graph G with n edges, where $n \geq 3k$. Repeatedly remove all nodes of degree at most 2 and their incident edges from G to obtain a graph G'. Note that G' is non-empty and satisfies the same relation between the number of nodes and the number of edges. Observe that each node of G' has degree three or higher. Pick an arbitrary node in G' and build a binary tree out of it to obtain a cycle of length at most $2\log n$.

cycle in a graph intersects a cut in an even number of edges. 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 using δ' in place of δ.

This concludes the proof. ■

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; the only difference being in the bound for the weight of dual codewords.

Theorem 6.5. *Let* $L \subseteq \mathbb{F}_2^n$ *be a linear space,* $\dim L = k$, $k \leq n^{\frac{1}{2}-\beta}$, $0 < \beta < 1/2$. *Then for some* $r = \Theta\left(\frac{n}{k}\log k\right)$ *and* $\alpha > 0$,

$$A_r(L) \geq (1+\alpha)\frac{r}{n}. \tag{6.6}$$

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 MacWilliams & Sloane (1977) for any linear subspace of $\mathbb{F}_2^m$ of dimension k, there is a nonzero 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 $\pi$.
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 supp($\mathbf{v}$) into $\pi$ as a new set;
6.         Drop the coordinates in supp($\mathbf{v}$) from $[n]$. Reduce $n$.
7.     END
```

Observe that every vector $\mathbf{v}$ that we select on step 4 above is, in fact, a dual vector for the space L, even though we pick it with respect to the "current universe". 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 α-attractive for L, for a positive α. 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 using δ' in place of δ.

This concludes the proof. ■

Similarly to Theorem 5.12, Theorem 6.5 above can be used to argue that Conjecture 3.18 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 (6.6) yields

$$A_r(L) \geq (1+\alpha)\frac{r}{n}. \tag{6.7}$$

Theorem 5.12 however presents a stronger result. Firstly, in the proof of Theorem 5.12 we use real spaces of dimension as low as n^τ to arrive at the bound (6.7) for some $r = \omega(n^{1-\epsilon})$. This leaves plenty of room for potential further improvements. Secondly, Theorem 5.12 applies to all sets of vectors of bounded triangular rank, while Theorem 6.5 only deals with $\mathbb{F}_2$-linear spaces.

7. Proof of Theorem 4.7

Our goal here is to prove the following theorem from Section 4.

THEOREM. *There exist positive constants δ and k_0 such that for all $k \geq k_0$, for all sets $L \subseteq \mathbb{F}_2^n$ with* $\mathrm{trk}(L) \leq k$:

$$A_3(L) \geq \frac{1+\delta}{k}. \tag{7.1}$$

We begin by generalizing some notation that was introduced in Section 4. Let $L \subseteq \mathbb{F}_2^n$. Assume $\{w_i\}$ are defined with respect to L as in (4.2). For $S \subseteq [n]$ we set

$$\mu(S) = \sum_{i\in S} w_i^{-1} \qquad \text{and} \qquad \mu'(S) = \sum_{i\in S} w_i^{-1/2}.$$

Let $\mu = \mu([n])$. For $\mathbf{v} \in \mathbb{F}_2^n$ we often write $\mu(\mathbf{v})$ instead of $\mu(\mathrm{supp}(\mathbf{v}))$ and $\mu'(\mathbf{v})$ instead of $\mu'(\mathrm{supp}(\mathbf{v}))$.

Definition 7.2. *Let $S \subseteq [n]$. Let $\mathbf{v}$ be a non-zero element of L. Let $\beta \leq \frac{1}{2}$ be a positive real. We say that $\mathbf{v}$ is β-balanced on S if*

$$\beta \cdot \mu'(\mathbf{v}) \leq \mu'(\mathrm{supp}(\mathbf{v}) \cap S) \leq (1-\beta) \cdot \mu'(\mathbf{v}). \tag{7.3}$$

If $\mathbf{v}$ is not β-balanced we say that $\mathbf{v}$ is β-unbalanced.

Let $S \subseteq [n], \mu(S) = \alpha\mu$ be fixed. We define a two-dimensional real space $X_S = \mathrm{Span}(\mathbf{x}_1, \mathbf{x}_2)$, where

$$\mathbf{x}_1(i) = \sqrt{\frac{1}{\mu w_i}}, \quad \text{for all } i \in [n] \tag{7.4}$$

and

$$\mathbf{x}_2(i) = \begin{cases} -\sqrt{\frac{1-\alpha}{\alpha}}\sqrt{\frac{1}{\mu w_i}}, & i \in S; \\ \sqrt{\frac{\alpha}{1-\alpha}}\sqrt{\frac{1}{\mu w_i}}, & \text{otherwise.} \end{cases} \tag{7.5}$$

In what follows we establish two preliminary lemmas and then proceed to the proof of Theorem 4.7.

Lemma 7.6. *There exist positive constants $\tau < \frac{1}{2}, \beta, \delta$ such that for all $L \subseteq \mathbb{F}_2^n$ with $\mathrm{trk}(L) \leq k$ and all sets $S \subseteq [n]$ satisfying*

$$\left|\frac{1}{2} - \frac{\mu(S)}{\mu}\right| \leq \frac{1}{2} - \tau$$

and all non-zero vectors $\mathbf{v} \in L$ that are β-unbalanced on S

$$\|\mathrm{Pr}_{X_S}(\mathbf{v})\|^2 \geq \frac{1+\delta}{k}. \tag{7.7}$$

Proof. Recall from (4.3) that $\|\mathbf{x}_1\| = 1$. Note that

$$\|\mathbf{x}_2\| = \frac{1-\alpha}{\alpha}\sum_{i \in S}\frac{1}{\mu w_i} + \frac{\alpha}{1-\alpha}\sum_{i \notin S}\frac{1}{\mu w_i} = 1.$$

Also note that

$$\begin{aligned}(\mathbf{x}_1, \mathbf{x}_2) &= -\sqrt{\tfrac{1-\alpha}{\alpha}} \sum_{i \in S} \tfrac{1}{\mu w_i} + \sqrt{\tfrac{\alpha}{1-\alpha}} \sum_{i \notin S} \tfrac{1}{\mu w_i} \\ &= -\sqrt{(1-\alpha)\alpha} + \sqrt{\alpha(1-\alpha)} \\ &= 0.\end{aligned}$$

Thus

$$\|\mathrm{Pr}_{X_S}(\mathbf{v})\|^2 = (\mathbf{v}, \mathbf{x}_1)^2 + (\mathbf{v}, \mathbf{x}_2)^2.$$

Theorem 4.5 yields $(\mathbf{v}, \mathbf{x}_1)^2 \geq \frac{1}{k}$. Thus to establish (7.7) it suffices to have $(\mathbf{v}, \mathbf{x}_2)^2 \geq \frac{\delta}{k}$. We set $w = \mathrm{wt}(\mathbf{v})$ and

$$T = (\mathbf{v}, \mathbf{x}_1) = \sqrt{\frac{1}{\mu w}} \cdot \mu'(\mathbf{v}).$$

Let

$$T_1 = \sqrt{\frac{1}{\mu w}} \cdot \mu'(\mathrm{supp}(\mathbf{v}) \cap S)$$

and let

$$T_2 = \sqrt{\frac{1}{\mu w}} \cdot \mu'(\mathrm{supp}(\mathbf{v}) \setminus S).$$

It is not hard to see that

$$(\mathbf{v}, \mathbf{x}_2) = -\sqrt{\frac{1-\alpha}{\alpha}}\, T_1 + \sqrt{\frac{\alpha}{1-\alpha}}\, T_2. \tag{7.8}$$

Since the vector $\mathbf{v}$ is β-unbalanced on S we either have $T_1 \leq \beta T$ and $T_2 \geq (1-\beta)T$ which implies

$$|(\mathbf{v}, \mathbf{x}_2)| \geq \left((1-\beta)\sqrt{\frac{\alpha}{1-\alpha}} - \beta\sqrt{\frac{1-\alpha}{\alpha}}\right) T \tag{7.9}$$

or $T_1 \geq (1-\beta)T$ and $T_2 \leq \beta T$ which implies

$$|(\mathbf{v}, \mathbf{x}_2)| \geq \left((1-\beta)\sqrt{\frac{1-\alpha}{\alpha}} - \beta\sqrt{\frac{\alpha}{1-\alpha}}\right) T. \tag{7.10}$$

It remains to note that the coefficients in front of T in the right hand sides of (7.9) and (7.10) can be uniformly bounded from below by a constant $\delta > 0$ provided β is somewhat small and α is sufficiently close to $\frac{1}{2}$. Combining this observation with $T \geq \frac{1}{\sqrt{k}}$ concludes the proof. ■

We now proceed to the next lemma. For $B \subseteq [n]$ let $\bar{B} = [n] \setminus B$.

Lemma 7.11. *Let $L \subseteq \mathbb{F}_2^n$. Let $B \subseteq [n]$ be such that for all $i \in B$ and $j \in \bar{B}$, $w_i \leq w_j$. Let $\mathbf{v}$ be a non-zero element of L and let $w = \mathrm{wt}(\mathbf{v})$. Suppose $\mathbf{v}$ is β-balanced on B and*

$$\mu'(\mathbf{v}) \leq (1+\epsilon)\sqrt{w}; \tag{7.12}$$

then for all $i \in \bar{B}$,

$$w_i \geq \left(1 - \frac{\epsilon}{(1+\epsilon)\beta}\right)^2 w. \tag{7.13}$$

Proof. Let

$$\sigma = \min_{i \in \bar{B}} \frac{w_i}{w}.$$

Clearly, for all $i \in B$ we have $w_i \leq \sigma w$. Let $t = |\mathrm{supp}(\mathbf{v}) \cap B|$. Inequality (7.12) yields

$$\frac{t}{\sqrt{\sigma w}} + \frac{w-t}{\sqrt{w}} \leq (1+\epsilon)\sqrt{w}.$$

This implies

$$t \leq \frac{\epsilon w}{\left(\frac{1}{\sqrt{\sigma}} - 1\right)}. \tag{7.14}$$

Inequality above tells us that t has to be small if σ is small. Recall that $\mathbf{v}$ is β-balanced on B. Thus

$$\mu'(\mathrm{supp}(\mathbf{v}) \cap \bar{B}) \leq (1-\beta)\mu'(\mathbf{v}).$$

Combining the above inequality with (7.12) and with

$$\frac{w-t}{\sqrt{w}} \leq \mu'(\mathrm{supp}(\mathbf{v}) \cap \bar{B})$$

we conclude that

$$w - t \leq (1+\epsilon)(1-\beta)w, \tag{7.15}$$

thus t is somewhat large, implying that σ cannot be too small. Formally, combining (7.14) with (7.15) we obtain

$$w(\beta + \epsilon\beta - \epsilon) \leq \frac{\epsilon w}{\left(\frac{1}{\sqrt{\sigma}} - 1\right)}.$$

The latter inequality yields

$$\sigma \geq \left(1 - \frac{\epsilon}{(1+\epsilon)\beta}\right)^2$$

and concludes the proof. ■

We now proceed to the main proof of this section.

Proof of Theorem 4.7: Fix an arbitrary set $L \subseteq \mathbb{F}_2^n$, $\mathrm{trk}(L) \leq k$. Let γ and ϵ be two small positive universal constants that we fix later. Without a loss of generality we assume that $\mu = \mu([n]) \geq (1-\gamma)k$, since otherwise we have

$$A_3(L) \geq A_1(L) \geq \frac{1}{\mu} \geq \frac{1+\delta}{k},$$

for a positive δ, where the second inequality follows from the proof of Theorem 4.5.

Let $\tau < \frac{1}{2}$ and β be positive constants from the statement of Lemma 7.6. As in the proof of Lemma 4.4 we consider an n-node hypergraph H, where the edges are the supports of non-zero elements of L. We perform a sequence of steps maintaining a set $B \subseteq [n]$ of black nodes in H, and a sequence $V \subseteq L$ of vectors that we have already picked. Initially, $B = \emptyset$ and $V = \emptyset$. On each step we apply some rule to pick a vector $\mathbf{v} \in L$ such that $\mathrm{supp}(\mathbf{v}) \not\subseteq B$. We add $\mathbf{v}$ to V and add all white elements of $\mathrm{supp}(\mathbf{v})$ to B. Our process has three stages, on which we apply different rules to pick the next vector $\mathbf{v}$ to add.

During the first stage on each step we choose a white node with the smallest w_i and set $\mathbf{v}$ to be a weight-w_i vector such that

$i \in \mathrm{supp}(\mathbf{v})$. Each such step increases $\mu(B)$ by at most 1. The first stage terminates once $\mu(B) \geq \tau\mu$. We now refer to the set B as S_0. Observe that for all $i \in S_0$ and $j \in \bar{S_0}$, $w_i \leq w_j$. We consider two possibilities:

1. There exists a set $S \subseteq [n]$, $\left|\frac{1}{2} - \frac{\mu(S)}{\mu}\right| \leq \frac{1}{2} - \tau$ such that: if $\mathbf{v} \in L$ is a non-zero vector that satisfies

$$\mu'(\mathbf{v}) \leq (1+\epsilon)\sqrt{\mathrm{wt}(\mathbf{v})} \tag{7.16}$$

   and is β-balanced on S_0, then $\mathbf{v}$ is β-unbalanced on S.

   We argue that in this case (7.1) holds. To see this consider the spaces

$$X_{S_0} = \mathrm{Span}(\mathbf{x}_1, \mathbf{x}_2) \quad \text{and} \quad X_S = \mathrm{Span}(\mathbf{x}_1, \mathbf{x}_3)$$

   from Lemma 7.6. Let $\mathbf{v}$ be an arbitrary non-zero vector from L. If $\mathbf{v}$ violates (7.16); then

$$\|\mathrm{Pr}_{X_{S_0}}(\mathbf{v})\|^2 \geq (\mathbf{v}, \mathbf{x}_1)^2 = \left(\frac{\mu'(\mathbf{v})}{\sqrt{\mu \cdot \mathrm{wt}(\mathbf{v})}}\right)^2 \geq \frac{(1+\epsilon)^2}{k}.$$

   Else, if $\mathbf{v}$ is β-unbalanced on S_0, $\|\mathrm{Pr}_{X_{S_0}}(\mathbf{v})\|^2 \geq \frac{1+\delta}{k}$ by Lemma 7.6. Finally, by our assumption, if $\mathbf{v}$ satisfies (7.16) as is β-balanced on S_0, it is β-unbalanced on S. Thus we have $\|\mathrm{Pr}_{X_S}(\mathbf{v})\|^2 \geq \frac{1+\delta}{k}$. Therefore every non-zero vector in L has a squared projection of size at least $\frac{1+\delta}{k}$ on the three-dimensional linear space $\mathrm{Span}(\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3)$, for a suitably chosen positive δ.

2. For every set $S \subseteq [n]$, such that $\left|\frac{1}{2} - \frac{\mu(S)}{\mu}\right| \leq \frac{1}{2} - \tau$ there exists a non-zero vector $\mathbf{v} \in L$ that satisfies (7.16) and is simultaneously β-balanced on both S_0 and S. Our goal now is to arrive at a contradiction. We proceed to the second stage of building sets B and V. This stage terminates once $\mu(B) > (1-\tau)\mu$, i.e., once we have $\left|\frac{1}{2} - \frac{\mu(B)}{\mu}\right| > \frac{1}{2} - \tau$.

On each step we choose $\mathbf{v}$ to be a vector that satisfies (7.16) and is β-balanced on both S_0 and B. Let w be the weight of such a $\mathbf{v}$. By Lemma 7.11 for all $i \in \bar{S}_0$ we have

$$w_i \geq \sigma w \quad \text{where} \quad \sigma = \left(1 - \frac{\epsilon}{(1+\epsilon)\beta}\right)^2. \tag{7.17}$$

Let $E = \operatorname{supp}(\mathbf{v}) \cap \bar{B}$. As $\mathbf{v}$ is β-balanced on B, (7.16) yields

$$\mu'(E) = \sum_{i \in E} \frac{1}{\sqrt{w_i}} \leq (1-\beta)\mu'(\mathbf{v}) \leq (1-\beta)(1+\epsilon)\sqrt{w}.$$

Therefore

$$\sum_{i \in E} \frac{1}{\sqrt{w_i}} \frac{1}{\sqrt{\sigma w}} \leq \frac{(1-\beta)(1+\epsilon)}{\sqrt{\sigma}}.$$

Thus by (7.17)

$$\mu(E) = \sum_{i \in E} \frac{1}{w_i} \leq \frac{(1-\beta)(1+\epsilon)}{\sqrt{\sigma}}.$$

Therefore on each step of the second stage we increase $\mu(B)$ by at most $(1+\epsilon)(1-\beta)/\sqrt{\sigma}$.

Finally, when $\mu(B)$ reaches $(1-\tau)\mu$ we proceed to the last stage. On this stage on each step we again choose a white node with the smallest w_i and set $\mathbf{v}$ to be a weight-w_i vector such that $i \in \operatorname{supp}(\mathbf{v})$. Each such step increases $\mu(B)$ by at most 1.

It is not hard to see that the steps above generate a tower a height at least

$$\begin{aligned} \lfloor \tau\mu \rfloor + \left\lfloor \tfrac{((1-2\tau)\mu-1)\sqrt{\sigma}}{(1+\epsilon)(1-\beta)} \right\rfloor + \lfloor \tau\mu - 1 \rfloor \geq \\ \left(2\tau(1-\gamma) + \tfrac{(1-2\tau)(1-\gamma)\sqrt{\sigma}}{(1+\epsilon)(1-\beta)}\right) k - c, \end{aligned} \tag{7.18}$$

for a constant c, where σ is given by (7.17). Note that fixing γ and ϵ to be sufficiently small positive constants and assuming that k is large enough we can ensure that the right hand side of (7.18) exceeds k.

This concludes the proof. ■

8. Relation to natural proofs

In Razborov & Rudich (1997) 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 Razborov & Rudich (1997) 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 Alekhnovich (2003). 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 3.18 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 8.1. *Let $0 < \epsilon < 1/4$ be a small positive constant. Let V be a random $n \times n$ matrix of zero and ones where every entry is set to 1 independently with probability $p = n^{-\epsilon}$. 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 for all $r \leq o(n)$,*

$$\frac{1}{n}\sum_{i\leq r} \lambda_i \geq (1+\alpha)\frac{r}{n}. \qquad (8.2)$$

Proof. The claim of the theorem is a simple corollary of the Marchenko-Pastur law Bai & Silverstein (2010) 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 Bai & Silverstein (2010)).

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 $0 \le a \le 4$ holds, with probability 1, that

$$\frac{\#\{i : \lambda_i \ge a\}}{n} \longrightarrow_{n \to \infty} \frac{1}{2\pi} \int_a^4 \sqrt{\frac{4-x}{x}}\, dx \tag{8.3}$$

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 Bai & Silverstein (2010) 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
(8.4)

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

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 undesired 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 i^{th} 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 matrices $\frac{1}{p(1-p)n} VV^t$ and BB^t respectively. By the Courant minimax

principle, if two matrices differ by a matrix of rank 1, their eigenvalues interlace. Hence, using the perturbed matrices changes the LHS of ((8.4)) by at most an additive factor of $O\left(1/n\right)$, which we may ignore. (For a more general statement, see Theorem A.44 in Bai & Silverstein (2010).)

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.

For this we note that U is a random matrix whose entries are centered i.i.d. Bernoulli random variables. Hence, by Theorem 5.8 of Bai & Silverstein (2010), the maximal eigenvalue of $X = \left(UU^t\right)/(p(1-p)n)$ tends to 4 with probability one as n goes to infinity. Consequently, we may (and will) assume that $\sum_{i,j=1}^n x_{ij}^2 = O(n)$.

Since $W = p(1-p)n \cdot X$, we deduce

$$\sum_{i,j=1}^{n} w_{ij}^2 \leq n^2p^2 \cdot \sum_{i,j=1}^{n} x_{ij}^2 \leq O\left(n^3p^2\right).$$

Finally, we can estimate the RHS of ((8.4)) from above as follows:

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

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. ■

9. 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 the 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 that we use relatively little combinatorial structure of $\mathbb{F}_2$-linearity. In particular our strongest result (Theorem 5.12) 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 3.18 in such generality.

REMARK 9.1. *Replacing the condition* $\dim(L) \leq n^{2\epsilon+\delta}$ *in the Conjecture 3.18 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 3.15 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. ■

Acknowledgements

Alex Samorodnitsky was supported by BSF and ISF grants. Ilya Shkredov was supported by Russian Scientific Foundation grant RSF 14-11-00433. We would like to thank Noga Alon, Swastik

Kopparty, and Mark Rudelson for many helpful discussions regarding this work. We would also like to thank anonymous reviewers for their very detailed and helpful comments on the manuscript.

References

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

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

Noga Alon, Rina Panigrahy & Sergey Yekhanin (2009). 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, 339–351.

Zhidong Bai & Jack Silverstein (2010). *Spectral Analysis of Large Dimensional Random Matrices.* Springer, New York.

Zeev Dvir (2011). On matrix rigidity and locally self-correctable codes. *Computational Complexity* **20**, 367–388.

Joel Friedman (1993). A note on matrix rigidity. *Combinatorica* **13**, 235–239.

Noboru Hamada (1973). 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.

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

Stasys Jukna (2001). *Extremal combinatorics.* Springer, Berlin, Heidelberg, New York.

Dieter Jungnickel & Vladimir Tonchev (2009). Polarities, quasi-symmetric designs, and Hamada conjecture. *Designs, Codes, and Cryptography* **51**, 131–140.

Boris Kashin & Alexander Razborov (1998). Improved lower bounds on the rigidity of Hadamard matrices. *Mathematical Notes* **63**, 471–475.

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

Satyanarayana Lokam (2009). Complexity lower bounds using linear algebra. *Foundations and Trends in Theoretical Computer Science* **4**, 1–155.

F. J. MacWilliams & N. J. A. Sloane (1977). *The Theory of Error Correcting Codes*. North Holland, Amsterdam, New York.

Ilan Newman & Yuri Rabinovich (2013). On multiplicative λ-approximations and some geometric applications. *SIAM Journal on Computing* **42**, 885–883.

Alexander Razborov (1989). On rigid matrices. Manuscript. In Russian.

Alexander Razborov & Steven Rudich (1997). Natural proofs. *Journal of Computer and System Sciences* **55**, 24–35.

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

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

Amin Shokrollahi, Daniel Speilman & Voelker Stemann (1997). A remark on matrix rigidity. *Information Processing Letters* **64**, 283–285.

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

Vladimir Temlyakov (1998). Nonlinear Kolmogorov widths. *Mathematical Notes* **63**, 785–795.

KIRILL USKOV (2002). *Kolmogorov width of geometric configurations and functional compacts in Hilbert spaces.* Ph.D. thesis, Moscow State Technical University. In Russian.

LESLIE VALIANT (1977). Graph-theoretic arguments in low level complexity. In *6th Symposium on Mathematical Foundations of Computer Science (MFCS)*, 162–176.

Manuscript received 21 September 2015

ALEX SAMORODNITSKY
Hebrew University, Israel
salex@cs.huji.ac.il

ILYA SHKREDOV
Steklov Mathematical Institute
and IITP RAS, Russia
ilya.shkredov@gmail.com

SERGEY YEKHANIN
Microsoft Research, United States
yekhanin@microsoft.com