Improved rank bounds for design matrices and a new proof of Kelly's theorem

Zeev Dvir[*]   Shubhangi Saraf[†]   Avi Wigderson[‡]

Abstract

We study the rank of complex sparse matrices in which the supports of different columns have small intersections. The rank of these matrices, called *design* matrices, was the focus of a recent work by Barak et. al. [BDWY11] in which they were used to answer questions regarding point configurations. In this work we derive near-optimal rank bounds for these matrices and use them to obtain asymptotically tight bounds in many of the geometric applications. As a consequence of our improved analysis, we also obtain a new, linear algebraic, proof of Kelly's theorem, which is the complex analog of the Sylvester-Gallai theorem.

1 Introduction

The classical Sylvester-Gallai (SG) Theorem states the following: *Given any finite set of points in the Euclidean plane, not all on the same line, there exists a line passing through exactly two of the points.* This result was first conjectured by Sylvester in 1893 [Syl93] and then proved independently by Melchior in 1940 [Mel40] and Gallai in 1943 (in an answer to the same question independently posed by Erdos [Erd43]). The following complex variant of the Sylvester-Gallai Thoerem was proved by Kelly [Kel86] in response to a question of Serre: *Given any finite set of points in $\mathbb{C}^d$, not all on the same complex two-dimensional plane, there exists a line passing through exactly two of the points.* The above result is tight over the complex numbers, since there are two-dimensional configurations of points satisfying the condition on triples. The survey by Borwein and Moser [BM90] gives a very good overview of the SG Theorem, its proofs and the different variants that have been studied in the past. One application of our techniques, discussed later, is a new proof of Kelly's theorem, which is significantly simpler than Kelly's original proof and is very different than the recent elementary proof of [EPS06].

In recent years, variants of the SG theorem have been useful in studying certain structural questions arising in theoretical computer science. Variants of the SG theorem were useful in

---

understanding the structure of low-depth arithmetic circuits [DS06, KS09, SS10]. Quantitative versions of the SG theorem were shown to be closely linked to the structure of linear Locally Correctable Codes [BDWY11, BDSS11]. These applications join a growing number of papers in which geometric theorems regarding point/line arrangements are finding applications in theoretical computer science. We refer the reader to the recent survey [Dvi12] for an overview of some of these applications.

1.1   Rank of design matrices

Motivated by the application to point configurations, [BDWY11] studied the rank of certain matrices, called *design* matrices, and proved lower bounds on their rank. These bounds were then used to prove quantitative analogs of the SG theorem. We begin by defining design matrices formally. For a vector $v \in \mathbb{F}^n$, where $\mathbb{F}$ is a field, we denote by $\mathsf{supp}(v) = \{i \in [n] \,|\, v_i \neq 0\}$.

Definition 1.1 (Design matrix)**.** *Let A be an $m \times n$ matrix over a field $\mathbb{F}$. Let $R_1, \ldots, R_m \in \mathbb{F}^n$ be the rows of A and let $C_1, \ldots, C_n \in \mathbb{F}^m$ be the columns of A. We say that A is a (q, k, t)-design matrix if*

1. *For all $i \in [m]$, $|\mathsf{supp}(R_i)| \leq q$.*

2. *For all $j \in [n]$, $|\mathsf{supp}(C_j)| \geq k$.*

3. *For all $j_1 \neq j_2 \in [n]$, $|\mathsf{supp}(C_{j_1}) \cap \mathsf{supp}(C_{j_2})| \leq t$.*

A 'typical' setting of the parameters that often arises in applications is when q is a small constant, $m \approx n^2$, $k \approx n$ and t is a constant. The main result in [BDWY11] is the following rank bound.

Theorem 1.2 ([BDWY11])**.** *Let A by an $m \times n$ matrix. If A is a (q, k, t) design matrix then*

$$\mathsf{rank}(A) \geq n - \left(\frac{qtn}{2k}\right)^2$$

For the aforementioned 'typical' setting of the parameters, one get a lower bound of $n - O((n/k)^2)$ on the rank. By improving a key lemma in [BDWY11] using a more careful analysis, we are able to prove the following new bound.

Theorem 1.3. *Let A by an $m \times n$ matrix. If A is a (q, k, t) design matrix then*

$$\mathsf{rank}(A) \geq \frac{n}{1 + \frac{q(q-1)mt}{nk^2}} \geq n - \frac{mtq(q-1)}{k^2}$$

As a corollary, we get:

Corollary 1.4. *Let A by an $m \times n$ matrix. If A is a (q, k, t) design matrix then*

$$\mathsf{rank}(A) \geq \frac{n}{1 + \frac{q(q-1)t}{k}} \geq n - \frac{ntq(q-1)}{k}.$$

Proof. If $m \leq nk$ then we are done (substitute $m = nk$ into the bound in Theorem 1.3). Otherwise, we can remove rows of A until we are left with a new matrix A' that has exactly $m' = nk$ rows and s.t A' is also a (q, k, t)-design matrix (as long as $m > nk$ there has to be a row we can remove and maintain that each column has at least k non zeros). Now, apply Theorem 1.3 on A' and use the fact that $\mathsf{rank}(A) \geq \mathsf{rank}(A')$. $\qquad\square$

Here, for the 'typical' setting, we get a rank bound of $n - O(n/k)$, which is asymptotically better then the one obtained in Theorem 1.2.

We also prove a variant of Theorem 1.3 in which q does not appear. Note that if each column of A has support of size exactly k, then the total number of nonzero entries in A is nk. Thus the average size of the support of a row would be nk/m. In general it can be shown that replacing q with nk/m in Theorem 1.3 would give a false statement. However, we show that we can replace q with $O(nt/k)$. This is exactly the average row-support when $m = O(k^2/t)$ and, in this regime, the bound on the rank (which is now independent of q) is tight[1].

Theorem 1.5. *Let A by an $m \times n$ matrix. If A is a (q, k, t) design matrix with $k \leq nt$ then*

$$\mathsf{rank}(A) \geq n - \frac{6mn^2t^3}{k^4}.$$

As before, by replacing m with nk we get the following corollary.

Corollary 1.6. *Let A by an $m \times n$ matrix. If A is a (q, k, t) design matrix with $k \leq nt$ then*

$$\mathsf{rank}(A) \geq n - \frac{6n^3t^3}{k^3}.$$

Related work: The problem of bounding the rank of matrices with certain patterns of zeros and non-zeros is not new, and has been studied in the past in a variety of contexts. One line of research comes from Hamada's conjecture [Ham73, JT09, BIW07]. In this setting however, the notion of being a design is stricter than the notion we use in this paper. Another line of research that studies the zero-nonzero patterns of matrices and their ranks has many applications to graph theory[FH07]. Rank bounds on matrices with 'sign patterns' of positive and negative entries have also received a great deal of attention in recent years [RS08, For02], and it would be interesting to see if our techniques can say anything meaningful in this setting.

1.2 Square design matrices and monotone rigidity

Theorem 1.5, which removes the dependence on q, allows us to get meaningful lower bounds on the rank of *square* design matrices. The results of [BDWY11] did not give anything for such matrices. Let A be an $n \times n$ matrix such that every column has support of size $k \approx \sqrt{n}$ and such that for every two columns, the size of the intersection of the supports of the two columns t is $O(1)$. For instance, the zero-nonzero pattern of the projective plane incidence matrix has

---

[1]Observe that for an $m \times n$ matrix that is a (q, k, t) design, if $n = \Omega(k/t)$, then $m = \Omega(k^2/t)$. This follows by a simple inclusion-exclusion argument on the union of the supports of any $\Omega(k/t)$ columns of A.

this structure. In this case we can obtain a bound of $\Omega(n)$ on the rank of such a matrix - not by applying our rank bound directly but by applying Theorem 1.5 to the matrix after deleting a linear number of columns.

A simple consequence of this result, proved in Section 4, is that, if one takes the $n \times n$ incidence matrix of the projective plane and changes a small number of 1's in the matrix to arbitrary values, then the resulting matrix has high (linear in n) rank. This can be viewed as a restricted model of the *matrix rigidity* problem of Valiant [Val77]. A matrix is rigid if changing a small number of its positions cannot decrease its rank by much. Valiant showed that a linear circuit computing a transformation given by a rigid matrix cannot have linear size and logarithmic depth. Hence, the problem of finding an explicit rigid matrix will imply circuit lower bounds that are beyond our reach at this point. Our restricted model allows one to only change positions in the matrix that are non-zero. Even though this result does not yield any interesting result on circuit lower bounds, we find it encouraging in that it gives a way to control rank under *some* type of perturbations. The full details are given in Section 4.

1.3 Configurations with many collinear triples

Given a set of points $v_1, \ldots, v_n \in \mathbb{C}^d$, we call a line that passes through exactly two of the points of the set an *ordinary* line. A line passing through at least three points is called *special*. We will use $\dim(v_1, \ldots, v_n)$ to denote the dimension of the linear span of $v_1, \ldots, v_n$ and by affine-dim$(v_1, \ldots, v_n)$ the dimension of the affine span of $v_1, \ldots, v_n$ (i.e., the minimum r such that $v_1, \ldots, v_n$ are contained in a shift of a linear subspace of dimension r).

The main geometric application studied in [BDWY11] was to extend the SG theorem to configurations of points termed δ-SG configurations.

Definition 1.7 (δ-SG configuration)**.** *We say that a set of points $v_1, v_2, \ldots, v_n \in \mathbb{C}^d$ is a δ-SG configuration if for every v_i, $i \in [n]$, at least $\delta(n-1)$ of the remaining points lie on special lines through v_i.*

Setting $\delta = 1$ we can state the original SG theorem as saying that any 1-SG configuration $v_1, \ldots, v_n$ (over $\mathbb{R}$) has affine-dim$(v_1, \ldots, v_n) \leq 1$ (i.e., is contained in a line). Kelly's theorem [Kel86] obtains the bound affine-dim$(v_1, \ldots, v_n) \leq 2$ for 1-SG configurations over $\mathbb{C}$ (the bound 2 is tight in this case). The following theorem, proved in [BDWY11] gives a generalization of Kelly's theorem to δ-SG configurations.

Theorem 1.8 ([BDWY11])**.** *Let $V = \{v_1, \ldots, v_n\} \subset \mathbb{C}^d$ be a δ-SG configuration. Then*

$$\text{affine-dim}(v_1, \ldots, v_n) \leq 13/\delta^2.$$

When $\delta = 1$ one gets a bound of 9 on the affine dimension.

There are two shortcomings of this theorem. The first is the quadratic dependence on δ. Placing the points on $1/\delta$ lines in general position one can construct a δ-SG configuration with dimension $\Omega(1/\delta)$. It was left as an open question in [BDWY11] to close this quadratic gap between the lower and upper bound on the dimension of δ-SG configurations. The second issue

4

is that one does not recover Kelly's theorem from the proof of [BDWY11], but only an inferior bound of 9 on the dimension. We are able to correct both of these issues.

Theorem 1.9. *Let $V = \{v_1, \ldots, v_n\} \subset \mathbb{C}^d$ be a δ-SG configuration. Then*

$$\text{affine-dim}(v_1, \ldots, v_n) \leq 12/\delta.$$

When $\delta = 1$ one gets a bound of 2 on the affine dimension.

There are two known proofs of Kelly's theorem. Kelly's original proof, answering a question by Serre, used deep results from algebraic geometry. An elementary proof was recently found by Elkies, Pretorius and Swanepoel [EPS06]. Our proof is conceptually very different from both of these and uses only elementary linear algebra.

Theorem 1.9 is proved, as in [BDWY11], by reduction to the rank bound for design matrices. One constructs a design matrix whose co-rank bounds the dimension of the configuration and then applies one of the bounds on the rank of these matrices.

Average-case version: A natural variant on the definition of a δ-SG configuration is to only require the presence of many pairs of points on special lines (instead of requiring each point to belong to many such pairs). In [BDWY11] it was shown that:

Theorem 1.10 ([BDWY11])**.** *Let $V = \{v_1, \ldots, v_n\} \subset \mathbb{C}^d$ be a set of n points such that at least δn^2 (unordered) pairs of them lie on special lines. Then there exists a subset $V' \subset V$ such that $|V'| \geq (\delta/6)n$ and so that V' is a δ-SG configuration.*

Combining this result with our improved bound on the dimension of δ-SG configurations we get the following improvement to a theorem from [BDWY11].

Corollary 1.11. *Let $V = \{v_1, \ldots, v_n\} \subset \mathbb{C}^d$ be a set of n distinct points. Suppose that there are at least δn^2 unordered pairs of points in V that lie on a special line. Then there exists a subset $V' \subset V$ such that $|V'| \geq (\delta/6)n$ and affine-dim$(V') \leq O(1/\delta)$.*

1.4 Flats of higher dimension

A *k-flat* is an affine subspace of dimension k. Let $\text{fl}(v_1, \ldots, v_k)$ denote the flat spanned by these k points (it can have dimension at most $k - 1$). We call $v_1, \ldots, v_k$ *independent* if their flat is of dimension $k - 1$, and say that $v_1, \ldots, v_k$ are *dependent* otherwise. Considering some fixed finite subset $V \subset \mathbb{C}^d$ of size n we call a k-flat *ordinary* if its intersection with V is contained in the union of a $(k-1)$-flat and a single point (this agrees with the definition of an ordinary line when $k = 1$). A k-flat is *elementary* if its intersection with V has exactly $k + 1$ points. Notice that for $k = 1$ (when flats are lines) the two notions of ordinary and elementary coincide.

The next definition generalizes the notion of a δ-SG configuration for higher dimensional flats in two different ways (using ordinary/elementary flats). For $k = 1$ both definitions coincide.

Definition 1.12 (δ-SG$_k$, δ-SG$_k^*$)**.** *The set V is a δ-SG$_k^*$ configuration if for every independent $v_1, \ldots, v_k \in V$ there are at least δn points $u \in V$ such that either $u \in fl(v_1, \ldots, v_k)$ or the k-flat $fl(v_1, \ldots, v_k, u)$ contains a point w outside $fl(v_1, \ldots, v_k) \cup \{u\}$. The set V is a δ-SG$_k$ configuration if for every independent $v_1, \ldots, v_k \in V$ there are at least δn points $u \in V$ such that either $u \in fl(v_1, \ldots, v_k)$ or the k-flat $fl(v_1, \ldots, v_k, u)$ is not elementary. Notice that a δ-SG$_k^*$ configuration is also a δ-SG$_k$ configuration and that for $k = 1$ both are the same.*

In [BDWY11], the following theorem was proved.

Theorem 1.13 ([BDWY11])**.** *Let V and V^* be a δ-SG$_k$ and a δ-SG$_k^*$ configurations respectively in $\mathbb{C}^d$. Then:*

1. *affine-$dim(V^*) \leq O\left((k/\delta)^2\right)$.*

2. *affine-$dim(V) \leq 2^{C^k}/\delta^2$, where $C > 0$ is a universal constant.*

Prior to this result, the only known bound for configurations with many special k-flats was a result proved by Hansen and Bonnice-Edelstein [Han65, BE67] which gives a bound of $O(k)$ on the dimension of a 1-SG$_k$ (or 1-SG$_k^*$) configuration over the reals. Since Theorem 1.13 is proved in a black-box manner using the result for δ-SG configurations we can plug-in our improvement, given in Theorem 1.9, to obtain the following.

Theorem 1.14. *Let V and V^* be a δ-SG$_k$ and a δ-SG$_k^*$ configurations respectively in $\mathbb{C}^d$. Then:*

1. *affine-$dim(V^*) \leq O\left(k/\delta\right)$.*

2. *affine-$dim(V) \leq C^k/\delta$, where $C > 0$ is a universal constant.*

Notice that, whereas the improvement for δ-SG$_k^*$ configurations is only quadratic, the improvement for $\delta - SG_k$ configurations is exponential (this is due to the way the basic bound is amplified in the induction on k). The proof of this theorem is identical to the proof of Theorem 1.13 appearing in [BDWY11], only with Theorem 1.8 replaced by Theorem 1.9. For completeness, we give the details in Section 5.2.

1.5   A variation on Freiman's Lemma

Consider a finite set A in some abelian group. One can define the *sumset* $A + A = \{a_1 + a_2 \mid a_1, a_2 \in A\}$ in a natural way. A well-known result in additive combinatorics is the following lemma, known as Freiman's lemma, which derives structural information on A, given bounds on the size of $A + A$.

Lemma 1.15 (Freiman's lemma. See [TV06])**.** *Let A be a finite subset of $\mathbb{R}^d$ and suppose $|A + A| \leq K|A|$. Then A is contained in a linear subspace of dimension at most $\lfloor K - 1 \rfloor$.*

Clearly, the condition $|A+A| \leq K|A|$ can be replaced by $|\{(a_1+a_2)/2 \mid a_1, a_2 \in A\}| \leq K|A|$, where we replace sums with mid-points. Surprisingly enough, the original proof of this lemma

works also when we replace mid points with *any point on the line segment connecting the two points.* More formally, for two sets $A, B \subset \mathbb{R}^d$ and any function $f : A \times B \mapsto \mathbb{R}^d$ we can define $A +_f B = \{f(a, b) \mid a \in A, b \in B\}$. Then, as long as $f(a, b)$ is on the line segment connecting a, b (and is different from a, b) we get the same conclusion as in Lemma 1.15, assuming $|A +_f B| \leq K|A|$.

Intuitively, our results for δ-SG configurations are of a similar flavor since the assumption of Freiman's Lemma (in its generalized form just stated) implies the existence of many pairs of points on special lines. We are able to use our techniques to derive the following theorem (whose proof appears in Section 6).

Theorem 1.16. *Let A be a finite subset of $\mathbb{C}^d$ and let $f : A \times A \mapsto \mathbb{C}^d$ be any function such that for all $a_1 \neq a_2 \in A$ we have $f(a_1, a_2) = \alpha a_1 + (1 - \alpha)a_2$ for some $\alpha = \alpha(a_1, a_2) : A \times A \mapsto \in \mathbb{C} \setminus \{0, 1\}$ (i.e., $f(a_1, a_2)$ is on the line passing through a_1, a_2 minus the two points a_1, a_2). Suppose that*

$$|A +_f A| \leq K \cdot A.$$

Then $\dim(A) \leq O(K^2)$

This theorem relaxes the conditions of Freiman's lemma by allowing (a) points in complex space and (b) the value of $f(a_1, a_2)$ to be outside the convex hull of a_1, a_2. On the other hand, we get a worse bound of $O(K^2)$ instead of $O(K)$. We do not know if this quadratic loss is needed or not.

1.6 Organization

In Section 2 we introduce some preliminaries related to the technique of matrix scaling. In Section 3 we prove Theorem 1.3 and Theorem 1.5. The rank bound for square design matrices and the application for monotone linear circuits is given in Section 4. In Section 5 we prove our main application, Theorem 1.9. In Section 5.2 we prove the high dimensional variant. In Section 6 we prove Theorem 1.16.

2 Preliminaries – Matrix Scaling

One of the most important ingredients in the proof of the rank bound for design matrices is the notion of *matrix scaling*. Informally, the matrix scaling theorem states that if a matrix does not have any large zero sub-matrices, then one can multiply the rows and columns of the matrix by non-zero scalars so that all the row sums are equal and all the column sums are equal (assuming the entries are non-negative).

The technique of matrix scaling originated in a paper of Sinkhorn [Sin64] and has been widely studied since then (see [LSW00] for more background). It was used in [BDWY11] for the first time to study design matrices, and we build upon their work and extend it.

We first set up some notation. For a complex matrix X, we let X^* denote the matrix X conjugated and transposed. Also we let X_{ij} denote the (i, j) entry of X. For two complex

vectors $u, v \in \mathbb{C}^m$, we denote their inner product by $\langle u, v \rangle = \sum_{i=1}^{m} u_i \cdot \bar{v}_j$ and let $\|u\| = \sqrt{\langle u, u \rangle}$ denote the ℓ_2 norm of the complex vector u.

Definition 2.1. *[Matrix scaling] Let A be an $m \times n$ complex matrix. Let $\rho \in \mathbb{C}^m, \gamma \in \mathbb{C}^n$ be two complex vectors with all entries non-zero. We denote by*

$$\mathbf{SC}(A, \rho, \gamma)$$

the matrix obtained from A by multiplying the (i,j)'th element of A by $\rho_i \cdot \gamma_j$. We say that two matrices A, B of the same dimensions are a scaling of each other if there exist non-zero vectors ρ, γ such that $B = \mathbf{SC}(A, \rho, \gamma)$. It is easy to check that this is an equivalence relation. We refer to the elements of the vector ρ as the row scaling coefficients *and to the elements of γ as the* column scaling coefficients*. Notice that two matrices which are a scaling of each other have the same rank and the same pattern of zero and non-zero entries.*

Below we define a property of matrices that gives sufficient conditions for finding a scaling of a matrix which has certain row and column sums.

Definition 2.2 (Property-S)**.** *Let A be an $m \times n$ matrix over some field. We say that A satisfies* Property-S *if for every zero sub-matrix of A of size $a \times b$ it holds that*

$$\frac{a}{m} + \frac{b}{n} \leq 1. \tag{1}$$

For example, a square matrix has Property-S if is has a non-zero generalized diagonal. Also, notice that this property is maintained under concatenation (say, putting two matrices with the same number of columns one under the other). The following theorem is the main tool we will use. Its proof uses ideas from convex optimization but we will only need to use it as a black box.

Theorem 2.3 (Matrix scaling theorem, Theorem 3 in [RS89])**.** *Let A be an $m \times n$ real matrix with non-negative entries which satisfies Property-S. Then, for every $\epsilon > 0$, there exists a scaling A' of A such that the sum of each row of A' is at most $1 + \epsilon$ and the sum of each column of A' is at least $m/n - \epsilon$. Moreover, the scaling coefficients used to obtain A' are all positive real numbers.*

In our proof will use the following easy corollary of the above theorem that appeared in [BDWY11]. This corollary is obtained by applying the matrix scaling theorem to the matrix obtained by squaring all entries of the original matrix.

Corollary 2.4 (Corollary from [BDWY11])**.** *Let $A = (a_{ij})$ be an $m \times n$ complex matrix which satisfies Property-S. Then, for every $\epsilon > 0$, there exists a scaling A' of A such that for every $i \in [m]$*

$$\sum_{j \in [n]} |a_{ij}|^2 \leq 1 + \epsilon$$

and for every $j \in [n]$

$$\sum_{i \in [m]} |a_{ij}|^2 \geq m/n - \epsilon.$$

8

3 Proof of the rank bound

In this section we will present the proofs for Theorem 1.3 and Theorem 1.5. The proof follows the same general outline as the one appearing in [BDWY11]:

Step 1 – Scaling: Given the design matrix A, we construct a scaling A' of A where every column has large ℓ_2 norm and every row has small ℓ_2 norm. Since A need not satisfy Property-S, we are not be able to apply Corollary 2.4 directly. Instead we first find a matrix B whose rows are chosen from the rows of A, with repetitions, such that no row is chosen too many times. If each row of A occurs a maximum of c times in B, then we call B a c-cover of A. We can then apply Corollary 2.4 to get a scaling of B with equal row norms and equal column norms, and then use the scaling of B to derive a scaling of A with the desired properties. In [BDWY11] B is taken to be a q-cover of A.

Step 2 – Obtaining a diagonal dominant matrix: Given the scaling A' of A we consider the matrix $M = A'^* A'$. Clearly all the diagonal entries of M, which correspond to the squared ℓ_2 norms of the columns of A', are large. We use the design properties of A, as well as the properties of the scaling to show that the sum of square of the off-diagonal entries of M is small. Matrices such as M are called 'diagonal dominant' and bounding their rank can be done in various ways (see e.g., [Alo09]). In this step, our calculation gives a tighter analysis of the bounds of the entries of M and we are hence able to obtain the stronger rank bounds compared to [BDWY11].

3.1 Covering lemmas

The following lemma is implicit in [BDWY11] and shows how one can find a cover of a matrix A that satisfies Property-S. Recall that a matrix B is a c-cover of A if each row of B is a row of A and each row of A appears at most c times in B.

Lemma 3.1. *Let A by an $m \times n$ matrix over $\mathbb{C}$ that is a (q, k, t) design matrix. Then there exists an $nk \times n$ matrix B that is a q-cover of A, and such that B satisfies Property-S.*

Proof sketch. B is constructed as follows: for each $i \in [n]$, we let B_i be a $k \times n$ submatrix of A which has no zeros in the i'th column. Let B be the $nk \times n$ matrix which is composed of the concatenation all matrices $B_i, i \in [n]$. □

We now prove a variant of Lemma 3.1 where every row of A appears at most $6nt/k$ times in B. In some settings $6nt/k$ might be smaller than q and then this variant might give a potentially stronger rank bound (as stated in Theorem 1.5). Observe that this rank bound is independent of the parameter q.

Lemma 3.2. *Let A by an $m \times n$ matrix over $\mathbb{C}$ that is a (q, k, t) design. Suppose $k \leq nt$, then there exists an $nk \times n$ matrix B that is a $6nt/k$-cover of A, and that satisfies Property-S.*

Proof. We split the set of n columns of A into ℓ sets, where $\ell \leq \lceil 2nt/k \rceil$, each of size at most $k/2t$. Call these sets $S_1, S_2, \ldots, S_\ell$. For each S_i, we will first construct a matrix B_i which is an $(|S_i| \cdot k) \times n$ matrix that will be composed of the rows of A, where each row appears at most 2 times. The matrix B will be an $nk \times n$ matrix which is composed of all the matrices B_i, $1 \leq i \leq \ell$. In other words the set of rows of B is the multi-set obtained by taking all the rows of all the B_i.

The matrix B_i is constructed as follows. For each column in S_i, there are at least $k/2$ rows such that none of the other columns in S_i has a nonzero entry in that row. This is because the intersection of the support of any two columns has size at most t, and there are at most $k/2t$ columns in S_i. Thus the support of each column can intersect the union of the support of all other columns in S_i in at most $k/2$ locations. For each column of S_i, pick some $k/2$ rows such that none of the other columns in S_i has support which intersects that row, and add 2 copies of each of those rows to the matrix B_i. Do this for each column in S_i. It follows immediately from construction that each row of A appears at most two times in each B_i.

Since each row of A appears at most two times in each B_i, and B is composed of $\leq \lceil 2nt/k \rceil$ such matrices B_i, thus each row of A appears at most $2\lceil 2nt/k \rceil \leq 4nt/k + 2 \leq 6nt/k$ times in B (using the bound $k \leq nt$).

To see that B satisfies property-S, observe that B can be written as the union of k square $n \times n$ matrices each with nonzero entries on the main diagonal. For each of the $n \times n$ matrices, we would take $|S_i|$ rows per set S_i, where each row corresponds to one of the columns in S_i, such that none of the other columns in S_i have a nonzero entry in that row. $\qquad\square$

The next lemma shows the one can use a cover of A to find a good scaling of A. This lemma is also implicit in [BDWY11] and we give the proof sketch only for completeness.

Lemma 3.3. *Let A by an $m \times n$ matrix over $\mathbb{C}$, and let B be an $nk \times n$ matrix that is a c-cover of A, and such that B satisfies Property-S. Then, for every $\epsilon > 0$, there exists a scaling A' of A in which each row of A' has ℓ_2 norm at most $\sqrt{1+\epsilon}$ and each column of A' has ℓ_2 norm at least $\sqrt{(k-\epsilon)/c}$.*

Proof sketch. Fix $\epsilon > 0$ and apply Corollary 2.4 on B to obtain a scaling B' of B such that the ℓ_2 norm of each column is at least $\sqrt{k-\epsilon}$, and the ℓ_2 norm of each row is at most $\sqrt{1+\epsilon}$. We now use this scaling B' of B to obtain the scaling A' of A. The scaling of the columns used to get A' is the same as the scaling coefficients for the columns of B'. We pick the scaling coefficients of the rows of A' as follows: for each row R that appears in A, we look at the occurrences $R_1, R_2, \ldots, R_i$ of the same row in B and look at the scaling coefficients for those rows in B'. Say the coefficients are $s_1, s_2, \ldots, s_i$ ($i \leq c$), then we take $\max\{s_1, \ldots s_i\}$ to be the scaling coefficient of row R. If the row R does not appear in B at all, then we pick the scaling coefficient to be such that the final ℓ_2 norm of the row is 1. One can easily verify that A' is a scaling of A with each row of A' having ℓ_2 norm at most $\sqrt{1+\epsilon}$ and each column having ℓ_2 norm at least $\sqrt{(k-\epsilon)/c}$. $\qquad\square$

3.2 Proof of Theorem 1.3 and Theorem 1.5

Before proving the theorems we prove two more lemmas. The first lemma is the main new ingredient in our proof enables us to get a tighter bound on the entries of the diagonal dominant matrix M (see proof outline above).

Lemma 3.4. *Let A be an $m \times n$ matrix over $\mathbb{C}$. Suppose that each row of A has ℓ_2 norm $< \alpha$ and suppose that the supports of every two columns of A intersect in at most t locations. Let $M = A^*A$. Then*

$$\sum_{i \neq j} |M_{ij}|^2 \leq tm\alpha^4.$$

Moreover if we know that the size of the support of every row in A is at most q, then

$$\sum_{i \neq j} |M_{ij}|^2 \leq \left(1 - \frac{1}{q}\right) tm\alpha^4.$$

Proof. For $1 \leq i \leq n$, Let C_i denote the ith column of A. Then

$$\sum_{i \neq j} |M_{ij}|^2 = \sum_{i \neq j} |\langle C_i, C_j \rangle|^2$$

$$= \sum_{i \neq j} \left| \sum_{k=1}^{m} A_{ki} \overline{A}_{kj} \right|^2$$

$$\leq \sum_{i \neq j} t \sum_{k=1}^{m} |A_{ki}|^2 |A_{kj}|^2$$

$$\leq t \sum_{k=1}^{m} \left(\sum_{i=1}^{n} |A_{ki}|^2 \right)^2$$

$$\leq tm\alpha^4.$$

When there are at most q nonzero entries per row, we have

$$t \sum_{k=1}^{m} \sum_{i \neq j} |A_{ki}|^2 |A_{kj}|^2 = t \sum_{k=1}^{m} \left(\sum_{i=1}^{n} |A_{ki}|^2 \right)^2 - t \sum_{k=1}^{m} \left(\sum_{i=1}^{n} |A_{ki}|^4 \right)$$

$$\leq t \sum_{k=1}^{m} \left(\sum_{i=1}^{n} |A_{ki}|^2 \right)^2 - t \sum_{k=1}^{m} \frac{1}{q} \left(\sum_{i=1}^{n} |A_{ki}|^2 \right)^2$$

$$= \left(1 - \frac{1}{q}\right) t \sum_{k=1}^{m} \left(\sum_{i=1}^{n} |A_{ki}|^2 \right)^2$$

$$\leq \left(1 - \frac{1}{q}\right) tm\alpha^4.$$

$\square$

11

The second lemma is a variant of a folklore lemma on the rank of diagonal dominant matrices (see, e.g., [Alo09]).

Lemma 3.5. *Let M be an $n \times n$ Hermitian matrix such that for each $i \in [n]$, $M_{ii} \geq L$, where L is some positive real number. Then,*

$$\mathsf{rank}(M) \geq \frac{n^2 L^2}{nL^2 + \sum_{i \neq j} |M_{ij}|^2}.$$

Proof. First, note that, w.l.o.g, we can assume $M_{ii} = L$ for all $i \in [n]$. If not, we can replace M with a scaling M' of M defined as $M'_{ij} = \frac{L}{\sqrt{M_{ii}M_{jj}}} \cdot M_{ij}$. Since all scaling coefficients are at most 1 we have

$$\sum_{i \neq j} |M'_{ij}|^2 \leq \sum_{i \neq j} |M_{ij}|^2$$

and both matrices M and M' have the same rank. To bound the rank of M (assuming all diagonal elements are equal to L) we denote its (real) non-zero eigenvalues by $\lambda_1, \lambda_2, \ldots, \lambda_r$, where $r = \mathsf{rank}(M)$. Then

$$n^2 L^2 = \mathsf{tr}(M)^2 = \left(\sum_{i=1}^{r} \lambda_i \right)^2$$

$$\leq r \sum_{i=1}^{r} \lambda_i^2 = r \sum_{i,j=1}^{n} |M_{ij}|^2$$

$$= r \left(nL^2 + \sum_{i \neq j} |M_{ij}|^2 \right).$$

Rearranging, we get the required bound on r. $\qquad\square$

Proof of Theorem 1.3. Let A be an (q, k, t) design matrix and fix some $\epsilon > 0$. Using Lemmas 3.1 and 3.3, we obtain a scaling A' of A where each row of A' has ℓ_2 norm at most $\sqrt{1 + \epsilon}$ and each column has ℓ_2 norm at least $\sqrt{(k - \epsilon)/q}$. Let $M = A'^* A'$. Then $M_{ii} \geq (k - \epsilon)/q$ and, by Lemma 3.4, $\sum_{i \neq j} |M_{ij}|^2 \leq \left(1 - \frac{1}{q} \right) tm(1 + \epsilon)^2$. Applying Lemma 3.5 to M we get that

$$\mathsf{rank}(M) \geq \frac{n^2 \left(\frac{k-\epsilon}{q} \right)^2}{n \left(\frac{k-\epsilon}{q} \right)^2 + \left(1 - \frac{1}{q} \right) tm(1 + \epsilon)^2}.$$

Taking ϵ to zero and simplifying, we get

$$\mathsf{rank}(M) \geq \frac{n}{1 + \frac{q(q-1)mt}{nk^2}} \geq n - mtq(q - 1)/k^2$$

where the second inequality follows from the fact that $1/(1 + x) \geq 1 - x$ for all x. Since $\mathsf{rank}(A) = \mathsf{rank}(A') \geq \mathsf{rank}(M)$, Theorem 1.3 follows. $\qquad\square$

Proof of Theorem 1.5. The only change in the proof of Theorem 1.5 is that instead of Lemma 3.1 we use Lemma 3.2. By Lemmas 3.1 and 3.3, we get a scaling A' of A where each row of A' has ℓ_2 norm at most $\sqrt{1+\epsilon}$ and each column has ℓ_2 norm at least $\sqrt{(k-\epsilon)k/6nt}$. Letting $M = A'^* A'$ as before, we have $M_{ii} \geq (k-\epsilon)k/6nt$ and, by Lemma 3.4, $\sum_{i \neq j} |M_{ij}|^2 \leq \left(1 - \frac{1}{q}\right) tm(1+\epsilon)^2$. Applying Lemma 3.5 as before we get

$$r \geq \frac{n}{1 + \frac{6mnt^3}{k^4}} \geq n - 6mn^2t^3/k^4.$$

$\square$

Tight examples for Theorem 1.3: One might hope that the bound of tm in the Theorem 1.3 (where t is the maximum intersection of any two columns) can be replaced by $\bar{t}m$, where $\bar{t}$ is some kind of average of intersections of all pairs of columns. One attempt towards showing such a statement would be by introducing another Cauchy-Schwarz after the one used in the proof of Lemma 3.4. However, this does not seem to help. The resulting bound after the second Cauchy-Schwarz seems to give a bound of $\sqrt{\sum_{i \neq j} t_{ij}^2} m$, which is worse than tm. To see why *max t* should not be replaceable with *average t* in the final statement of the rank bound is by the example of a square $n \times n$ matrix A with n/s blocks of size $s \times s$ arranged along the diagonal. These blocks have all entries equal to 1, and 0 elsewhere. Then the rank of this matrix is n/s. Now, max t equals s. If n is much larger than s, then average t is much smaller than 1. Now take this matrix A and randomly pick $n/100$ columns of A and call this new matrix A'. Then we can still have that average $t < 1$, $k = s$ and $q = s/100$. Plugging it into our rank bound for design matrices (with max t replaced by average t) will give us a lower bound of $\Omega(n)$ on the rank of A', whereas we know that rank of A' is at most n/s.

4   Rank bound for square design matrices

We start by deriving an easy corollary that bounds the rank of square design matrices.

Theorem 4.1. *Let A be an $n \times n$ matrix that is a (q, k, t) design matrix. Then* $\mathsf{rank}(A) = \Omega(k^4/nt^3)$.

Proof. Delete any $n - k^4/10nt^3$ columns of A to get a new matrix A'. Then A' is a $n \times k^4/10nt^3$ matrix that is also a (q, k, t) design. Applying Theorem 1.5 to A', and doing the calculation, we get that $\mathsf{rank}(A') \geq k^4/10nt^3 - k^4/20nt^3)$. Thus $\mathsf{rank}(A) = \Omega(k^4/nt^3)$. $\square$

Note that the incidence matrix of the projective plane of order p is a $(p^2+p+1) \times (p^2+p+1)$ matrix which is a $(p, p, 1)$ design. Thus any square matrix that has the same zero-nonzero pattern as the incidence matrix of the projective plane must have linear rank. Such a result was known for the $0-1$ valued incidence matrix of the projective plane matrix using the argument by Alon [Alo09]. However the result above allows us to get a bound on the rank with just the information of the zero-nonzero pattern. In contrast, over finite fields, the rank of the projective plane incidence matrix is sub-linear in n.

4.1 Monotone rigidity

One of the motivation for proving rank bounds for matrices, using only information on their support, comes from a longstanding open problem in complexity theory, known as *matrix rigidity*. Informally, a matrix is *rigid* if one cannot reduce its rank by much by changing a small number of its entries in each column. More formally, we have:

Definition 4.2 (Matrix rigidity [Val77]). *Let A be an $n \times n$ matrix over some field. We say that A is (r, s)-rigid if A cannot be written as $A = L + S$ with*

1. *L a matrix of rank at most r and*

2. *S a matrix with at most s non zeros per column.*

In [Val77], Valiant showed that if A is $(n^\alpha, \alpha n)$-rigid (for some constant $\alpha > 0$) than a linear circuit (a circuit with fan-in 2 gates, each computing a linear combination of previously computed gates) computing the mapping $x \mapsto x^t A$ cannot have both $O(n)$ size and $O(\log(n))$ depth. Since proving such lower bounds is beyond the reach of current techniques, constructing *explicit* rigid matrix (over any field) has become a much sought after goal. We refer the reader to the survey [Lok09] for more background on this longstanding open problem.

Using our results on design matrices, we can construct an explicit matrix that is highly rigid, as long as one only changes its non-zero entries.

Definition 4.3 (Monotone rigidity). *Let A be an $n \times n$ matrix over some field. We say that A is (r, s)-monotonically rigid if A cannot be written as $A = L + S$ with*

1. *L a matrix of rank at most r,*

2. *S a matrix with at most s non zeros per column and*

3. *The support of S is contained in that of A (that is, $S_{ij} \neq 0$ implies $A_{ij} \neq 0$).*

The following is an immediate corollary of Theorem 4.1.

Corollary 4.4. *Let A be an $n \times n$ matrix with non-negative real entries that is a (q, k, t)-design matrix with $k \geq \Omega(\sqrt{n})$ and $t \leq O(1)$. Then A is $(\alpha\sqrt{(n)}, \alpha n)$-monotonically rigid for some $\alpha > 0$. For example, one can take A to be the projective plane incidence matrix.*

We observe that this result can be used to derive super linear lower bounds (via Valiant's argument) for *monotone* circuits, which are circuits that can use linear combinations with non negative coefficients only. Such lower bounds, however, can be achieved using much simpler arguments (in fact, much stronger lower bounds of the form $\approx n^{1.5}$). We are not aware, however, of a simple way to construct monotonically rigid matrices.

5   Proof of Theorem 1.9

We first prove the general bound for $\delta > 0$ (later we will analyze the $\delta = 1$ case). Suppose $v_1, \ldots, v_n \in \mathbb{C}^d$ are a δ-SG configuration. Let V be the $n \times d$ matrix whose ith row is the

14

vector v_i. By shifting the points so that $v_i \neq 0$ for all $i \in [n]$ we have

$$\text{affine-dim}\{v_1, \ldots, v_n\} = \dim\{v_1, \ldots, v_n\} - 1 = \mathsf{rank}(V) - 1.$$

(The difference of 1 between affine and linear dimension will only matter in the $\delta = 1$ case.)

Thus we want to upper-bound the rank of V. To do so, we will construct an $m \times n$ design-matrix A so that $AV = 0$. Then we will use the design properties of A to argue that the rank of A must be high, which in turn implies that the rank of V must be small. The following lemma is implicit in [BDWY11] and we include its proof sketch here for completeness.

Lemma 5.1. *Let $v_1, v_2, \ldots, v_n$ be a δ-SG configuration. Let V be the $n \times d$ matrix whose i'th row is the vector v_i. Then there exists an $m \times n$ matrix A such that A is a $(3, 3k, 6)$-design matrix with $k = \lceil \delta(n-1) \rceil$, every row of A has support of size exactly 3, and such that $AV = 0$.*

Proof sketch. A result of [Hil73] on the existence of diagonal Latin squares implies that for all $r \geq 3$ there exists a set $T \subset [r]^3$ of $r^2 - r$ triples that satisfies the following properties:

1. Each triple $(t_1, t_2, t_3) \in T$ is of three distinct elements.

2. For each $i \in [r]$ there are exactly $3(r-1)$ triples in T containing i as an element.

3. For every pair $i, j \in [r]$ of distinct elements there are at most 6 triples in T which contain both i and j as elements.

Let $\mathcal{L}$ denote the set of all special lines in the configuration. Let L_i be a subset of $\mathcal{L}$ containing lines passing through v_i. For each $\ell \in \mathcal{L}$ let V_ℓ denote the set of points in the configuration which lie on the line ℓ. Then $|V_\ell| \geq 3$ and we can assign to it a family of triples $T_\ell \subset V_\ell^3$ satisfying the three properties above. We now construct the matrix A by going over all special lines $\ell \in \mathcal{L}$ and for each triple $t = (i, j, k) \in T_\ell$ adding as a row of A a vector with three non-zero coefficients in positions i, j, k, corresponding to the linear dependency among the collinear vectors v_i, v_j, v_k so that we have $AV = 0$. We now argue that the matrix A is a $(3, 3k, 6)$-design matrix as follows. The number of non-zeros in each row is exactly 3 by construction. For each v_i, there are at least $k = \lceil \delta(n-1) \rceil$ points (other than v_i) on special lines through v_i. Summing over all of these lines we get that v_i appears in at least $3k$ triples and so the i'th column of A will contain at least $3k$ non-zeros. Every distinct pair of points v_i, v_j determine a unique line and so, by construction, can appear in at most 6 triples together. $\square$

Given this lemma and our rank bounds from the previous section, the proof of Theorem 1.9 follows quite easily.

Proof of Theorem 1.9. By Corollary 1.4,

$$\mathsf{rank}(A) \geq \frac{n}{1 + \frac{q(q-1)t}{k}} \geq \frac{n}{1 + \frac{3 \cdot 2 \cdot 6}{3 \lceil \delta(n-1) \rceil}} \geq \frac{n}{1 + \frac{12}{\delta n - 1}}.$$

Now,

$$\frac{n}{1 + \frac{12}{\delta n - 1}} = n - \frac{12n}{\delta n + 9} > n - \frac{12}{\delta}.$$

15

Hence,

$$\text{affine-dim}\{v_1, v_2, \ldots, v_n\} = \mathsf{rank}(V) - 1 \leq n - \mathsf{rank}(A) < \frac{12}{\delta}.$$

$\square$

5.1 The case of $\delta = 1$: Kelly's Theorem

We now describe how to obtain the tight bound of 2 on the affine-dimension when $\delta = 1$. In this scenario, every pair of points is on a special line. We start with the following simple claim.

Claim 5.2. *Let A be an $m \times n$ matrix so that A is a (q, k, t) design matrix and such that the support of each row in A is exactly q. Then $m\binom{q}{2} \leq \binom{n}{2}t$*

Proof. We count the number of pairs of locations in the matrix which are in the same row and are both nonzero. Counting once by rows we get that this quantity is equal to $m\binom{q}{2}$. On the other hand, counting by columns (going over all pairs of columns) we get an upper bound of $\binom{n}{2}t$ since two columns intersect in at most t places. $\square$

Applying Lemma 5.1 we get an $m \times n$ matrix A which is a $(3, 3(n-1), 6)$-design matrix. By the above claim we have that $m \leq n(n-1)$. Using Theorem 1.3 we get that

$$\mathsf{rank}(A) \geq \frac{n}{1 + \frac{3 \cdot 2 \cdot n(n-1) \cdot 6}{n(3(n-1))^2}} = \frac{n}{1 + \frac{4}{n-1}} = \frac{n(n-1)}{n+3} > n - 4.$$

Hence,

$$\mathsf{rank}(V) \leq n - \mathsf{rank}(A) < 4.$$

Thus

$$\dim\{v_1, v_2, \ldots, v_n\} = \mathsf{rank}(V) \leq 3,$$

and

$$\text{affine-dim}\{v_1, v_2, \ldots, v_n\} \leq 2.$$

5.2 Proof of high dimensional variant

Fix some point $v_0 \in V$. By a *normalization w.r.t.* v_0 we mean an affine transformation $N : \mathbb{C}^d \mapsto \mathbb{C}^d$ which first moves v_0 to zero, then picks a hyperplane H s.t. no point in V (after the shift) is parallel to H (i.e., has inner product zero with the orthogonal vector to H) and finally multiplies each point (other than zero) by a constant s.t. it is in H. It is easy to see (see [BDWY11]) that, for such a mapping N, we have that $v_0, v_1, \ldots, v_k$ are dependent iff $N(v_1), \ldots, N(v_k)$ are dependent.

We now prove Theorem 1.14 in two parts (corresponding to the two cases of V and V^*). We denote by $f(\delta, k)$ the maximum d such that there exists a δ-SG_k^* configuration of dimensions d. We denote by $g(\delta, k)$ the maximum d such that there exists a δ-SG_k configuration of dimensions d.

Proof for δ-SG_k^* configurations: The proof is by induction on k. For $k = 1$ we know $f(\delta, 1) \leq c/\delta$ with $c > 1$ a universal constant. Suppose $k > 1$. We separate into two cases. The first case is when V^* is a $(\delta/(2k))$-SG_1 configuration and we are done using the bound on $k = 1$. In the other case there is some point $v_0 \in V^*$ s.t. the size of the set of points on special lines through v_0 is at most $\delta/(2k)$. Let S denote the set of points on special lines through v_0. Thus $|S| < \delta n/(2k)$. Let $N : \mathbb{C}^d \mapsto \mathbb{C}^d$ be a normalization w.r.t. v_0. Notice that for points $v \notin S$ the image $N(v)$ determines v. Similarly, all points on some special line map to the same point via N.

Our goal is to show that $V' = N(V^* \setminus \{v_0\})$ is a $((1 - 1/(2k))\delta)$-SG_{k-1}^* configuration (after eliminating multiplicities from V'). This will complete the proof since $\dim(V^*) \leq \dim(V') + 1$. Indeed, if this is the case we have

$$f(\delta, k) \leq \max\{2c(k/\delta), f((1 - 1/(2k))\delta, k - 1) + 1\}.$$

and, by induction, we have $f(\delta, k) \leq 4c(k/\delta)$.

Fix $v_1', \ldots, v_{k-1}' \in V'$ to be $k - 1$ independent points (if no such tuple exists then V' is trivially 1-SG_{k-1}^* configuration). Let $v_1, \ldots, v_{k-1} \in V^*$ be (necessarily independent) points s.t. $N(v_i) = v_i'$ for $i \in [k - 1]$. Thus, there is a set $U \subset V^*$ of size at least δn s.t. for every $u \in U$ either $u \in \mathrm{fl}(v_0, v_1, \ldots, v_{k-1})$ or the k-flat $\mathrm{fl}(v_0, v_1, \ldots, v_{k-1}, u)$ contains a point w outside $\mathrm{fl}(v_0, v_1, \ldots, v_{k-1}) \cup \{u\}$.

Let $\tilde{U} = U \setminus S$ so that N is invertible on $\tilde{U}$ and

$$|\tilde{U}| \geq |U| - |S| \geq (1 - 1/(2k))\delta n.$$

Suppose $u \in \tilde{U}$ and let $u' = N(u)$. If $u \in \mathrm{fl}(v_0, v_1, \ldots, v_{k-1})$ then u' is in $\mathrm{fl}(v_1', \ldots, v_{k-1}')$. Otherwise, $\mathrm{fl}(v_0, v_1, \ldots, v_{k-1}, u)$ contains a point w outside $\mathrm{fl}(v_0, v_1, \ldots, v_{k-1}) \cup \{u\}$. Let $w' = N(w)$. We will show that w' is (a) contained in the $(k - 1)$-flat $\mathrm{fl}(v_1', \ldots, v_{k-1}', u')$ and (b) is outside $\mathrm{fl}(v_1', \ldots, v_{k-1}') \cup \{u'\}$. Property (a) follows since $v_0, v_1, \ldots, v_{k-1}, u, w$ are dependent and so $v_1', \ldots, v_{k-1}', u', w'$ are also dependent. To show (b) observe first that the points $v_1', \ldots, v_{k-1}', u'$ are independent (since $v_0, v_1, \ldots, v_{k-1}, u$ are independent) and so u' is not in $\mathrm{fl}(v_1', \ldots, v_{k-1}')$. We also need to show that $w' \neq u'$ but this follows from the fact that $u \neq w$ and so $w' = N(w) \neq N(u) = u'$ since N is invertible on $\tilde{U}$ and $u \in \tilde{U}$. Since

$$|N(\tilde{U})| = |\tilde{U}| \geq (1 - 1/(2k))\delta n \geq (1 - 1/(2k))\delta|V'|$$

the proof is complete. $\qquad\square$

We now prove the second part of Theorem 1.14.

Proof for δ-SG_k configurations. The proof follows by induction on k (the case $k = 1$ is given by Theorem 1.9). Suppose $k > 1$. Suppose that $\dim(V) > g(\delta, k)$. We want to show that there exist k independent points $v_1, \ldots, v_k$ s.t. for at least $1 - \delta$ fraction of the points $w \in V$ we have that w is not in $\mathrm{fl}(v_1, \ldots, v_k)$ **and** the flat $\mathrm{fl}(v_1, \ldots, v_k, w)$ is elementary (i.e., does not contain any other point).

Let $k' = g(1, k-1)$. Since we are trying to show by induction that $g(\delta, k) \leq C^k/\delta$ for some absolute constant C, we can pick C so that $g(\delta, k) > f(\delta, k'+1)$. Therefore, we can find $k'+1$ independent points $v_1, \ldots, v_{k'+1}$ s.t. there is a set $U \subset V$ of size at least $(1-\delta)n$ s.t. for every $u \in U$ we have that u is not in $\mathrm{fl}(v_1, \ldots, v_{k'+1})$ **and** the $(k'+1)$-flat $\mathrm{fl}(v_1, \ldots, v_{k'+1}, u)$ contains only one point, namely u, outside $\mathrm{fl}(v_1, \ldots, v_{k'+1})$.

We now apply the inductive hypothesis on the set $V \cap \mathrm{fl}(v_1, \ldots, v_{k'+1})$ which has dimension at least $k' = g(1, k-1)$. This gives us k independent points $v'_1, \ldots, v'_k$ that define an elementary $(k-1)$-flat $\mathrm{fl}(v'_1, \ldots, v'_k)$. (Saying that V is not 1-SG_{k-1} is the same as saying that it contains an elementary $(k-1)$-flat). Joining any of the points $u \in U$ to $v'_1, \ldots, v'_k$ gives us an elementary k-flat and so the theorem is proved. $\qquad\square$

6   Proof of the variation on Freiman's Lemma

In this section we prove Theorem 1.16. Let A and $f : A \times A \mapsto \mathbb{C}^d$ be as in the statement of the theorem. The proof is divided into two claims.

Claim 6.1. *There exists a subset $A' \subset A$ with $|A| \geq \Omega(|A|/K)$ and $\dim(A') \leq O(K^2)$.*

Proof. Let $B = A +_f A$ so that for every pair $(a, a') \in A \times A$ there exists some point $b \in B$ such that $f(a, a') = b$. Thus, on average, a point $b \in B$ has $|A|^2/|B|$ pairs of $A \times A$ mapping to it via the function f. Let B_1 be the set of all points in B that have at least $|A|^2/10|B|$ pairs that map to it. Let S' be the set of pairs of $A \times A$ that map to some element of B_1. Then $|S'| > |A|^2/2$. Consider $B_2 = B_1 \cup A$ and observe that $|B_2| \leq O(K|A|)$. Now, each pair in S' is on a special line determined by B_2. Thus, by Corollary 1.11, for $\alpha = |S'|/|B_2|^2$, we get that there is a subset $B_{LD} \subset B_2$ of dimension at most $O(1/\alpha) = O(K^2)$ and size bounded by

$$|B_{LD}| \geq \Omega(\alpha |B_2|) \geq \Omega(|A|^2/|B_2|) \geq \Omega(|A|/K).$$

If $|B_{LD} \cap A| \geq |B_{LD}|/2$, then take $A' = B_{LD} \cap A$ and the claim is proved. Otherwise, consider the set $B'_{LD} = B_{LD} \setminus A \subset B_1$ so that $|B'_{LD}| \geq |B_{LD}|/2$. Each point of B'_{LD} has at least $|A|^2/10|B|$ pairs of $A \times A$ that map to it via the function f. For $a \in A$ we denote

$$M(a) = \{a' \in A \mid f(a, a') \in B'_{LD}\}.$$

Then the average of $|M(a)|$ (taken over all $a \in A$) is at least

$$\frac{1}{|A|} \sum_{a \in A} |M(a)| \geq \frac{|B'_{LD}| \cdot |A|^2}{10|B| \cdot |A|} = \Omega(|A|/K^2).$$

Call a point $a \in A$ a *heavy* point if

$$|M(a)| \geq \frac{|B'_{LD}| \cdot |A|}{100|B|} \geq \Omega(|A|/K^2).$$

Case (1): Some point $a^* \in A$ has

$$|M(a')| \geq \frac{K|B'_{LD}| \cdot |A|}{100|B|} = \Omega(|A|/K).$$

In this case, consider the set $B''_{LD} = B'_{LD} \cup \{a^*\}$. Clearly this set has dimension at most $\dim(B'_{LD}) + 1 = O(1/\alpha) + 1 = O(K^2)$. Also, the span of B''_{LD} contains the set $M(a^*)$. Thus, there are $\Omega(A/K)$ points of A that are contained in a set of dimension $O(K^2)$. This completes the proof of this case.

Case (2): In this case we have that

$$|M(a)| \leq \frac{K|B'_{LD}| \cdot |A|}{100|B|}$$

for all $a \in A$ and, in particular, for all heavy points. Therefore, there must be at least $\Omega(|A|/K)$ heavy points (otherwise, the average of $|M(a)|$ would be too small). Call the set of heavy points H, so that $|H| \geq \Omega(|A|/K)$. Pick $a_1 \in H$ and consider $R_1 = \mathsf{span}(B'_{LD} \cup \{a_1\}) \cap A$. Then R_1 contains $M(a_1)$, which has size $\Omega(|A|/K^2)$, and has dimension at most $\dim(B'_{LD}) + 1 = O(K^2)$. If $H \not\subset R_1$, we can pick some $a_2 \in H \setminus R_1$ and define $R_2 = \mathsf{span}(B'_{LD} \cup \{a_1, a_2\}) \cap A$. The dimension of R_2 is at most $\dim(B'_{LD}) + 2$, and its size is at least $|R_1| + \Omega(|A|/K^2)$, since $M(a_2) \cap M(a_1) = \emptyset$, or else a_2 would be in the span of $B'_{LD} \cup \{a_1\}$. Continuing in this manner (i.e., picking $a_3, a_4, \ldots$) for at most K steps or until we run out of elements of H (which has size $\Omega(|A|/K)$) we obtain a subset $A' \subset A$ of dimension at most $O(K^2) + K = O(K^2)$ containing at least $\Omega(A/K)$ elements. This completes the proof of the claim. $\square$

Claim 6.2. *We have* $\dim(A) \leq O(K^2)$.

Proof. Let $A' \subset A$ be a subset of size $\Omega(|A|/K)$ and dimension $O(K^2)$ given by the previous claim. Let T be a minimal set in $A \setminus A'$ for which $\mathsf{span}(T \cup A')$ contains A. Notice that this implies that the points in T are linearly independent and that

$$\dim(A' \cup T) = \dim(A') + |T|.$$

Observe that for every $a_1 \neq a_2 \in T$, there do not exist $a'_1 \neq a'_2 \in A'$ and $b \in B$ such that

$$f(a_1, a'_1) = f(a_2, a'_2) = b.$$

This is because otherwise $\{a_1, a_2\} \subseteq \mathsf{span}(A' \cup \{b\})$, which means that $\dim(\{a_1, a_2\} \cup A') \leq \dim(A') + 1$, which violates the properties of T. Therefore,

$$|T| \cdot |A'| \leq |B|.$$

This gives $|T| \leq O(K^2)$ and so $\dim(A) \leq \dim(A') + T = O(K^2)$. $\square$

Corollary 6.3. *If a set $A \subset \mathbb{R}^n$ defines at most $K|A|$ directions then* $\dim(A) \leq O(K^2)$

Proof. Notice that the proof of the above theorem works also if the function f is allowed to take values in projective space. Since the point at infinity on the line through a, b is the direction $a - b$ we get the required consequence. $\square$

References

[Alo09]   Noga Alon. Perturbed identity matrices have high rank: Proof and applications. *Comb. Probab. Comput.*, 18(1-2):3–15, 2009.

[BDSS11]   A. Bhattacharyya, Z. Dvir, A. Shpilka, and S. Saraf. Tight lower bounds for 2-query lccs over finite fields. In *Proc. of FOCS 2011*, pages 638–647, 2011.

[BDWY11]   B. Barak, Z. Dvir, A. Wigderson, and A. Yehudayoff. Rank bounds for design matrices with applications to combinatorial geometry and locally correctable codes. In *Proceedings of the 43rd annual ACM symposium on Theory of computing*, STOC '11, pages 519–528, New York, NY, USA, 2011. ACM.

[BE67]   W. Bonnice and M. Edelstein. Flats associated with finite sets in $\mathbb{P}^d$. *Niew. Arch. Wisk.*, 15:11–14, 1967.

[BIW07]   Omer Barkol, Yuval Ishai, and Enav Weinreb. On locally decodable codes, self-correctable codes, and t-private pir. In *APPROX '07/RANDOM '07: Proceedings of the 10th International Workshop on Approximation and the 11th International Workshop on Randomization, and Combinatorial Optimization. Algorithms and Techniques*, pages 311–325, Berlin, Heidelberg, 2007. Springer-Verlag.

[BM90]   P. Borwein and W. O. J. Moser. A survey of sylvester's problem and its generalizations. *Aequationes Mathematicae*, 40(1), 1990.

[DS06]   Zeev Dvir and Amir Shpilka. Locally decodable codes with 2 queries and polynomial identity testing for depth 3 circuits. *SIAM Journal on Computing*, 36(5):1404–1434, 2006.

[Dvi12]   Z. Dvir. Incidence theorems and their applications., 2012. Preprint available at http://arxiv.org/abs/1208.5073.

[EPS06]   Noam D. Elkies, Lou M. Pretorius, and Konrad J. Swanepoel. Sylvester-gallai theorems for complex numbers and quaternions,. *Discrete and Computational Geometry*, 35(3):361–373, 2006.

[Erd43]   P. Erdos. Problems for solution: 40654069, 1943.

[FH07]   Shaun M. Fallat and Leslie Hogben. The minimum rank of symmetric matrices described by a graph: A survey. *Linear Algebra and its Applications*, 426(2-3):558 – 582, 2007.

[For02]   Jürgen Forster. A linear lower bound on the unbounded error probabilistic communication complexity. *J. Comput. Syst. Sci.*, 65(4):612–625, 2002.

[Ham73]   N. Hamada. On the p-rank of the incidence matrix of a balanced or partially balanced incomplete block design and its application to error correcting codes. *Hiroshima Math. J.*, 3:154–226, 1973.

[Han65]   S. Hansen. A generalization of a theorem of sylvester on the lines determined by a finite point set. *Mathematica Scandinavia*, 16:175–180, 1965.

[Hil73]   A. J. W. Hilton. On double diagonal and cross latin squares. *J. London Math. Soc.*, s2-6(4):679–689, 1973.

[JT09]   Dieter Jungnickel and Vladimir D. Tonchev. Polarities, quasi-symmetric designs, and hamada's conjecture. *Des. Codes Cryptography*, 51(2):131–140, 2009.

[Kel86]   L. M. Kelly. A resolution of the sylvester - gallai problem of j. -p. serre. *Discrete & Computational Geometry*, 1:101–104, 1986.

[KS09]   Neeraj Kayal and Shubhangi Saraf. Blackbox polynomial identity testing for depth 3 circuits. In *FOCS '09: Proceedings of the 2009 50th Annual IEEE Symposium on Foundations of Computer Science*, pages 198–207, Washington, DC, USA, 2009. IEEE Computer Society.

[Lok09]   Satyanarayana V. Lokam. Complexity lower bounds using linear algebra. *Foundations and Trends in Theoretical Computer Science*, 4(1-2):1–155, 2009.

[LSW00]   N. Linial, A. Samorodnitsky, and A. Wigderson. A deterministic strongly polynomial algorithm for matrix scaling and approximate permanents. *Combinatorica*, 20(4):545–568, 2000.

[Mel40]   E. Melchior. Uber vielseite der projektive ebene. *Deutsche Math.*, 5:461–475, 1940.

[RS89]   U. Rothblum and H. Schneider. Scaling of matrices which have prespecified row sums and column sums via optimization. *Linear Algebra Appl*, 114-115:737–764, 1989.

[RS08]   Alexander A. Razborov and Alexander A. Sherstov. The sign-rank of ac^o. In *FOCS '08: Proceedings of the 2008 49th Annual IEEE Symposium on Foundations of Computer Science*, pages 57–66, Washington, DC, USA, 2008. IEEE Computer Society.

[Sin64]   R. Sinkhorn. A relationship between arbitrary positive matrices and doubly stochastic matrices. *Ann. Math. Statist.*, 35:876–879, 1964.

[SS10]   Nitin Saxena and C. Seshadhri. From sylvester-gallai configurations to rank bounds: Improved black-box identity test for depth-3 circuits. *Foundations of Computer Science, Annual IEEE Symposium on*, 0:21–29, 2010.

[Syl93]   J. J. Sylvester. Mathematical question 11851. *Educational Times*, 59:98, 1893.

[TV06]   T. Tao and V.H. Vu. *Additive Combinatorics*. Number v. 13 in Cambridge Studies in Advanced Mathematics. Cambridge University Press, 2006.

[Val77]   Leslie G. Valiant. Graph-theoretic arguments in low-level complexity. In *MFCS*, pages 162–176, 1977.