

Fractional Sylvester–Gallai theorems

Boaz Barak^a, Zeev Dvir^{b,1}, Avi Wigderson^c, and Amir Yehudayoff^d

^aMicrosoft Research New England, Microsoft, 1 Memorial Drive, Cambridge, MA 02142; ^bDepartment of Computer Science, Princeton University, 35 Olden Street, Princeton, NJ 08540; ^cSchool of Mathematics, Institute for Advanced Studies, Einstein Drive, Princeton, NJ 08540; and ^dDepartment of Mathematics, Technion—Israel Institute of Technology, Haifa 32000, Israel

Edited by Terence Chi-Shen Tao, University of California, Los Angeles, CA, and approved August 17, 2012 (received for review March 7, 2012)

We prove fractional analogs of the classical Sylvester–Gallai theorem. Our theorems translate local information about collinear triples in a set of points into global bounds on the dimension of the set. Specifically, we show that if for every points v in a finite set $V \subset \mathbb{C}^d$, there are at least $\delta|V|$ other points $u \in V$ for which the line through v, u contains a third point in V , then the V resides in a $(13/\delta^2)$ -dimensional affine subspace of \mathbb{C}^d . This result, which is one of several variants we study, is motivated by questions in theoretical computer science and, in particular, from the area of error correcting codes. Our proofs combine algebraic, analytic, and combinatorial arguments. A key ingredient is a new lower bound for the rank of design matrices, specified only by conditions on their zero/non-zero pattern.

discrete geometry | line arrangements

In 1893 James Joseph Sylvester posed the following, well known problem (1): “Prove that it is not possible to arrange any finite number of real points so that a right line through every two of them shall pass through a third, unless they all lie in the same right line.” This beautiful problem was solved by Eberhard Melchior in 1940 (2), and, independently, by Tibor Gallai in 1944 in response to a question of Erdos (3). This statement is commonly known as the Sylvester–Gallai theorem.

It is convenient to restate this result using the notions of special and ordinary lines. A *special* line is a line that contains at least three points from the given set. Lines that contain exactly two points from the set are called *ordinary*.

Theorem 1. (Sylvester–Gallai theorem). *If m distinct points v_1, \dots, v_m in \mathbb{R}^d are not collinear, then they define at least one ordinary line.*

In its contrapositive form, the theorem says that if for every $i \neq j$ in $[m]$ the line through v_i, v_j passes through a third point $v_k \notin \{v_i, v_j\}$, then $\dim\{v_1, \dots, v_m\} \leq 1$, where $\dim\{v_1, \dots, v_m\}$ is the dimension of the smallest affine subspace containing the points. In this formulation, the theorem can be thought of as converting local information (on collinear triples) into a global bound on the dimension of the system.

In 1966 Jean-Pierre Serre (4) asked for a complex version of this theorem. The complex version was first proved by Leroy Milton Kelly (5) using deep results from algebraic geometry. An elementary proof was later found by Elkies et al. (6).

Theorem 2. (Kelly’s theorem). *If m distinct points in \mathbb{C}^d are not coplanar, then they define at least one ordinary line.*

Both theorems are tight: In the real case, if all points are collinear (and there are at least three points) then no line passes through exactly two of them. In the complex plane, one can find noncollinear (but coplanar) configurations of points such that every line passing through two of them contains a third point.

This work studies scenarios in which the local geometric information is incomplete. We are no longer in a situation where every line is special, but are only guaranteed that, for every point, the special lines through this point cover some positive fraction of the set (later we will consider even more relaxed scenarios). To articulate this scenario we use the following definition: Call the points $v_1, \dots, v_n \in \mathbb{C}^d$ a δ -SG configuration if for every $i \in [n]$ there exists at least δn values of $j \in [n]$ such that the line through

v_i, v_j contains a third point in the set. We provide the following bound over the complex numbers. Later (see *Other Fields*) we will generalize our results to any field of characteristic zero or of sufficiently large positive characteristic.

Theorem 3. (Fractional SG theorem). *For any $\delta > 0$, the dimension of a δ -SG configuration over the complex numbers is at most $13/\delta^2$.*

The upper bound on the dimension should be compared with the trivial lower bound of $\Omega(1/\delta)$ that arises from a partition of the points into $1/\delta$ generically positioned lines. Here, and for the rest of the paper, $O(\cdot)$ and $\Omega(\cdot)$ are used to hide universal constants only.

One of the motivations for studying this problem is its connection to problems in theoretical computer science and coding theory. The problem of bounding the dimension of δ -SG configurations is closely related to locally correctable codes (LCCs). To read more about this connection, we refer the reader to ref. 7.

In section *Several Extensions*, using Theorem 3 and its proof, we derive the following additional results:

- An analog of Theorem 3 with lines replaced with higher dimensional flats (as in Hansen’s theorem).
- A fractional analog of the Motkin–Rabin theorem which is a two-color version of the Sylvester–Gallai theorem.
- A three-color “nonfractional” analog of the Motzkin–Rabin theorem (the proof, nevertheless, uses the “fractional” version in an essential way).
- Average-case versions of Theorem 3 in which we are only guaranteed that a quadratic number of pairs of points are on special lines and find a large subset of points that is low dimensional.
- Extensions of Theorem 3 to any field of characteristic zero or sufficiently large positive characteristic (as a function of n).

Methods: Rank of Design Matrices

The main ingredient in the proof of Theorem 3 is a general lower bound on the rank of matrices with certain zero/nonzero patterns. The connection between the two problems is not surprising, because both convert local combinatorial information into global algebraic information (i.e., rank/dimension bounds). The type of zero/nonzero patterns we consider are called designs:

Definition 4: (Design matrix). Let A be an $m \times n$ matrix over some field. For $i \in [m]$ let $R_i \subset [n]$ denote the set of indices of all nonzero entries in the i ’th row of A . Similarly, let $C_j \subset [m]$, $j \in [n]$, denote the set of nonzero indices in the j ’th column. We say that A is a (q, k, t) -design matrix if

1. For all $i \in [m]$, $|R_i| \leq q$.
2. For all $j \in [n]$, $|C_j| \geq k$.
3. For all $j_1 \neq j_2 \in [n]$, $|C_{j_1} \cap C_{j_2}| \leq t$.

An extended abstract from STOC 2011 called “Rank Bounds for Design Matrices with Applications to Combinatorial Geometry and Locally Correctable Codes” contains parts of this work.

Author contributions: B.B., Z.D., A.W., and A.Y. performed research.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

¹To whom correspondence should be addressed. E-mail: zeev.dvir@gmail.com.

The zero/nonzero pattern of the columns of a design matrix, C_1, \dots, C_n , form a design in that each set is large but the pairwise intersections are small. The following theorem gives a lower bound on the rank of design matrices and is proved in section *Bounds on the Rank of Design Matrices*.

Theorem 5. (Rank bound for design matrices). For every complex matrix A with n columns that is a (q, k, t) -design,

$$\text{rank}(A) \geq n - \left(\frac{q \cdot t \cdot n}{2k} \right)^2.$$

To get a feeling of the parameters, consider an $m \times n$ matrix with $O(1)$ nonzeros in each row, with $\Omega(n)$ nonzeros in each column and with $t = O(1)$ pairwise intersections of columns; Theorem 5 tells us that such a matrix has almost full rank, $n - O(1)$.

We begin, in section *Proof of the Fractional SG Theorem*, with the proof of Theorem 5. The main technical tool, Theorem 5 is proved in section *Bounds on the Rank of Design Matrices*. In section *Several Extensions* we consider various extensions to Theorem 3.

Proof of the Fractional SG Theorem

The following lemma is an easy consequence of (8) and will be used in the proof below.

Lemma 6. Let $r \geq 3$. Then 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 six triples in T which contain both i and j as elements.

Proof: By [(8), Theorem 4] there exists a Latin square $\{A_{ij}\}_{i,j \in [r]}$ with $A_{ii} = i$ for all $i \in [r]$. Taking all triples of the form (i, j, A_{ij}) with $i \neq j$ proves the Lemma.

Proof of Theorem 3: Let V be the $n \times d$ matrix whose i 'th row is the vector v_i . Assume without loss of generality (w.l.o.g.) that $v_1 = 0$. Thus $\dim\{v_1, \dots, v_n\} = \text{rank}(V)$. We will first build an $m \times n$ matrix A that will satisfy $A \cdot V = 0$. Then, we will argue that the rank of A is large because it is a design matrix. This argument will show that the rank of V is small.

Consider a special line ℓ which passes through three points v_i, v_j, v_k . This line gives a linear dependency among the three vectors v_i, v_j, v_k . In other words, this gives a vector $a = (a_1, \dots, a_n)$ which is nonzero only in the three coordinates i, j, k and such that $a \cdot V = 0$. If a is not unique, choose an arbitrary vector a with these properties. Our strategy is to pick a family of collinear triples among the points in our configuration and to build the matrix A from rows corresponding to these triples in the above manner.

Let \mathcal{L} denote the set of all special lines in the configuration (i.e., all lines containing at least three points). 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 $T_\ell \subset V_\ell^3$, given by Lemma 6 (we identify V_ℓ with $[r]$, where $r = |V_\ell|$ in some arbitrary way). We now construct the matrix A by going over all lines $\ell \in \mathcal{L}$ and for each triple in T_ℓ adding as a row of A the vector with three nonzero coefficients $a = (a_1, \dots, a_n)$ described above (so that a is the linear dependency between the three points in the triple).

Because the matrix A satisfies $A \cdot V = 0$ by construction, we only have to argue that A is a design matrix and bound its rank.

Claim 7. The matrix A is a $(3, 3k, 6)$ -design matrix, where $k = \lceil \delta n \rceil - 1$.

Proof: By construction, each row of A has exactly three nonzero entries. The number of nonzero entries in column i of A corresponds to the number of triples we used that contain the point v_i . These triples can come from all special lines containing v_i . Suppose there are s special lines containing v_i and let r_1, \dots, r_s denote the number of points on each of those lines. Then, because the lines through v_i have only the point v_i in common, we have that $\sum_{j=1}^s (r_j - 1) \geq k$. The properties of the families of triples T_ℓ guarantee that there are $3(r_j - 1)$ triples containing v_i coming from the j 'th line. Therefore there are at least $3k$ triples in total containing v_i .

The size of the intersection of columns i_1 and i_2 is equal to the number of triples containing the points v_{i_1}, v_{i_2} that were used in the construction of A . These triples can only come from one special line (the line containing these two points) and so, by Lemma 6, there can be at most six of those.

Applying Theorem 5 we get that

$$\begin{aligned} \text{rank}(A) &\geq n - \left(\frac{3 \cdot 6 \cdot n}{2 \cdot 3k} \right)^2 \geq n - \left(\frac{3 \cdot n}{\delta n - 2} \right)^2 \\ &\geq n - \left(\frac{3 \cdot n \cdot 13}{11 \cdot \delta n} \right)^2 > n - 13/\delta^2, \end{aligned}$$

where the third inequality holds as $\delta n \geq 13$ because otherwise the theorem trivially holds. This inequality implies that $\text{rank}(V) < 13/\delta^2$, which completes the proof. For $\delta = 1$, the calculation above yields $\text{rank}(V) < 11$.

Bounds on the Rank of Design Matrices

In this section we prove Theorem 5. For a set of complex vectors $V \in \mathbb{C}^n$ we denote by $\text{rank}(V)$ the dimension of the vector space spanned by elements of V . We denote the ℓ_2 -norm of a vector v by $\|v\|$. We denote by I_n the $n \times n$ identity matrix.

The starting point of the proof is the observation that, if the matrix entries are all in the set $\{0, 1\}$ then the proof is quite easy: simply consider the product $A^t \cdot A$ and observe that its diagonal elements are much larger than its off-diagonal elements. Such matrices, called *diagonal dominant matrices*, are known to have high rank and so A must have high rank as well (cf. Lemma 12 below). The choice of the set $\{0, 1\}$ is not important—as long as the ratios between different nonzero entries are bounded, the same proof strategy will work. We reduce to this case using a technique called *matrix scaling*.

Definition 8: [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 nonzero. We denote by $\text{SC}(A, \rho, \gamma)$ 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 nonzero vectors ρ, γ such that $B = \text{SC}(A, \rho, \gamma)$. 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 nonzero entries.

Matrix scaling originated in a paper of Sinkhorn R (9) and has been widely studied since (see ref. 10 for more background). The following is a special case of a theorem from (11) that gives sufficient conditions for finding a scaling of a matrix which has certain row and column sums.

Definition 9: (Property-S). Let A be an $m \times n$ matrix over some field. We say that A satisfies Property-S if for every zero submatrix of A of size $a \times b$ it holds that

$$\frac{a}{m} + \frac{b}{n} \leq 1. \quad [1]$$

Theorem 10. (Matrix scaling theorem, Theorem 3 in ref. 11). Let A be an $m \times n$ real matrix with nonnegative 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.

We will use the following easy corollary of this theorem obtained by applying it to the matrix whose elements are the squares of absolute values of A .

Corollary 11. 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 $j \in [n]$

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

We will also use a variant of a well known lemma (see for example ref. 12) which provides a bound on the rank of matrices whose diagonal entries are much larger than the off-diagonal ones.

Lemma 12. Let $A = (a_{ij})$ be an $n \times n$ complex Hermitian matrix and let $0 < \ell < L$ be integers. Suppose that $a_{ii} \geq L$ for all $i \in [n]$ (the diagonal elements of a Hermitian matrix are real) and that $|a_{ij}| \leq \ell$ for all $i \neq j$. Then

$$\text{rank}(A) \geq \frac{n}{1 + n \cdot (\ell/L)^2} \geq n - (n\ell/L)^2.$$

Proof: We can assume w.l.o.g. that $a_{ii} = L$ for all i . If not, then we can make the inequality into an equality by multiplying the i 'th row and column by $(L/a_{ii})^{1/2} < 1$ without changing the rank or breaking the symmetry. Let $r = \text{rank}(A)$ and let $\lambda_1, \dots, \lambda_r$ denote the nonzero eigenvalues of A (counting multiplicities). Because A is Hermitian we have that the λ_i 's are real. We have

$$\begin{aligned} n^2 \cdot L^2 = \text{tr}(A)^2 &= \left(\sum_{i=1}^r \lambda_i \right)^2 \leq r \cdot \sum_{i=1}^r \lambda_i^2 \\ &= r \cdot \sum_{i,j=1}^n |a_{ij}|^2 \leq r \cdot (n \cdot L^2 + n^2 \cdot \ell^2). \end{aligned}$$

Rearranging we get the required bound. The second inequality in the statement of the lemma follows from the fact that $1/(1+x) \geq 1-x$ for all x .

Proof of Theorem 5: To prove the theorem we will first find a scaling of A so that the norms (squared) of the columns are large and such that each entry is small. Our first step is to find an $nk \times n$ matrix B that will satisfy Property-S and will be composed from rows of A such that (s.t.) each row is repeated with multiplicity between 0 and q . Such a matrix can be constructed from A as follows: for each $i \in [n]$ let B_i be a $k \times n$ submatrix of A with no zeros in the i 'th column. Take B to be the $nk \times n$ matrix composed of all matrices B_i , $i \in [n]$. It is easy to check that B satisfies property-S and that each row of A appears in B at most q times.

Our next step is to obtain a scaling of B and, from it, a scaling of A . Fix some $\epsilon > 0$ (which will later tend to zero). Applying Corollary 11 we get a scaling B' of B such that the ℓ_2 -norm of each row is at most $\sqrt{1+\epsilon}$ and the ℓ_2 -norm of each column is at least $\sqrt{nk/n-\epsilon} = \sqrt{k-\epsilon}$. We now obtain a scaling A' of A as follows: The scaling of the columns are the same as for B' . For the rows of A appearing in B we take the maximal scaling coefficient used for these rows in B' , that is, if row i in A appears as rows $i_1, i_2, \dots, i_{q'}$ in B , then the scaling coefficient of row i in A' is the maximal scaling coefficient of rows $i_1, i_2, \dots, i_{q'}$ in B' . For rows not in B , we pick scaling coefficients so that their ℓ_2 norm (in the final scaling) is equal to 1. It is easy to verify that the matrix A' is a scaling of A such that each row has ℓ_2 -norm at most $\sqrt{1+\epsilon}$ and each column has ℓ_2 -norm at least $\sqrt{(k-\epsilon)/q}$.

Next, consider the matrix $M = (A')^* \cdot A'$, where $(A')^*$ is A' transposed conjugate. Then $M = (m_{ij})$ is an $n \times n$ Hermitian matrix. The diagonal entries of M are exactly the squares of the ℓ_2 -norm of the columns of A' . Therefore, $m_{ii} \geq (k-\epsilon)/q$ for all $i \in [n]$. The off-diagonal entries of M are the inner products of different columns of A' . The intersection of the support of each pair of different columns is at most t . The norm of each row is at most $\sqrt{1+\epsilon}$. For every two real numbers α, β so that $\alpha^2 + \beta^2 \leq 1 + \epsilon$ we have $|\alpha \cdot \beta| \leq 1/2 + \epsilon'$, where ϵ' tends to zero as ϵ tends to zero. Therefore $|m_{ij}| \leq t \cdot (1/2 + \epsilon')$ for all $i \neq j \in [n]$. Applying Lemma 12 we get that

$$\text{rank}(A) = \text{rank}(A') \geq n - \left(\frac{q \cdot t(1/2 + \epsilon') \cdot n}{k - \epsilon} \right)^2.$$

Because this inequality holds for all $\epsilon > 0$ it holds also for $\epsilon = 0$, which gives the required bound.

Several Extensions

We now describe several extensions of the fractional Sylvester–Gallai theorem: two average-case versions, two high-dimensional fractional versions, a two-color fractional version, a three-color nonfractional version, and statements on other fields (than real and complex numbers).

Average-Case Versions. In this section we argue about the case where we only know that there are many collinear triples in a configuration. We start with the following average-case version in which there is *some* large family of collinear triples that satisfies some regularity condition (no pair is in too many triples). Below we will use this theorem to derive another average-case version in which we only assume that many pairs are on special lines. In both theorems the conclusion is only that there exists a large subset that lies in small dimension (which is clearly the strongest qualitative statement one can hope for in this setting).

Theorem 13. (Average-case SG theorem Let). $V = \{v_1, \dots, v_m\} \subset \mathbb{C}^d$ be a set of m distinct points. Let T be some set of (unordered) collinear triples in V . Suppose $|T| \geq \alpha m^2$ and that every two points v, v' in V appear in at most c triples in T , then there exists a subset $V' \subset V$ such that $|V'| \geq \alpha m/(2c)$ and $\dim(V') \leq O(1/\alpha^2)$.

Notice that the bound on the number of triples containing a fixed pair of points is necessary for the theorem to hold. If we remove this assumption than we could create a counter example by arranging the points so that $m^{2/3}$ of them are on a line and the rest are placed so that every large subset of them spans the entire space (e.g., in general position). The proof will use the following hypergraph lemma.

Lemma 14. Let H be a 3-regular hypergraph with vertex set $[m]$ and αm^2 edges of codegree at most c (i.e., for every $i \neq j$ in $[m]$, the set

$\{i, j\}$ is contained in at most c edges). Then there is a subset $M \subseteq [m]$ of size $|M| \geq \alpha m / (2c)$ so that the minimal degree of the subgraph of H induced by M is at least $\alpha m / 2$.

Proof: We describe an iterative process to find M . We start with $M = [m]$. While there exists a vertex of degree less than $\alpha m / 2$, remove this vertex from M and remove all edges containing this vertex from H . Continuing in this fashion we conclude with a set M such that every point in M has degree at least $\alpha m / 2$. This process removed in total at most $m \cdot \alpha m / 2$ edges and thus the new H still contains at least $\alpha m^2 / 2$ edges. As the codegree is at most c , every vertex appears in at most cm edges. Thus, the size of M is of size at least $\alpha m / (2c)$.

Proof: We describe an iterative process to find M . We start with $M = [m]$. While there exists a vertex of degree less than $\alpha m / 2$, remove this vertex from M and remove all edges containing this vertex from H . Continuing in this fashion we conclude with a set M such that every point in M has degree at least $\alpha m / 2$. This process removed in total at most $m \cdot \alpha m / 2$ edges and thus the new H still contains at least $\alpha m^2 / 2$ edges. As the codegree is at most c , every vertex appears in at most cm edges. Thus, the size of M is of size at least $\alpha m / (2c)$.

Proof of Theorem 13: The family of triples T defines a 3-regular hypergraph on V of codegree at most c . Lemma 14 thus implies that there is a subset $V' \subseteq V$ of size $|V'| \geq \alpha m / (2c)$ that is an $(\alpha/2)$ -SG configuration. By Theorem 3, V' has dimension at most $O(1/\alpha^2)$.

We now state the second average-case variant of Sylvester–Gallai in which we assume that there are many pairs on special lines. Here there is no need for further assumption and the proof is by an easy reduction to Theorem 13.

Theorem 15. (Average-case SG theorem—2nd variant). Let $V = \{v_1, \dots, v_m\} \subset \mathbb{C}^d$ be a set of m distinct points. Suppose that there are at least $\beta \cdot m^2$ unordered pairs of points in V that lie on a special line (i.e., there is a third point collinear with them). Then there exists a subset $V' \subset V$ such that $|V'| \geq (\beta/6) \cdot m$ and $\dim(V') \leq O(1/\beta^2)$.

Proof: Let ℓ_1, \dots, ℓ_t denote the special lines in the configuration V . Let r_1, \dots, r_t be integers such that ℓ_i contains exactly $r_i \geq 3$ points from V . The assumption of the theorem implies that

$$\sum_{i=1}^t (r_i^2 - r_i) \geq 2\beta \cdot m^2.$$

We now apply Lemma 6 on all t lines to find t sets of triples T_1, \dots, T_t such that each T_i contains triples of points from the line ℓ_i . We now have that each T_i contains exactly $r_i^2 - r_i$ triples and that, in each T_i , two points appear in at most six triples. This last condition translates also to the union of the T_i 's because two lines intersect in at most one point. We thus see that the set $T = \cup_{i=1}^t T_i$ satisfies the conditions of Theorem 13 with $\alpha = 2\beta$ and $c = 6$. We can thus conclude that there exists a subset $V' \subset V$ such that $|V'| \geq (\beta/6) \cdot m$ and $\dim(V') \leq O(1/\beta^2)$ as was required.

Fractional Hansen Theorems. Hansen's theorem (13) is a high-dimensional version of the SG theorem in which lines are replaced with higher dimensional flats. Let $\text{fl}(v_1, \dots, v_k)$ denote the affine span of k points, i.e., the points that can be written as linear combinations with coefficients that sum to one (fl for "flat"). We call v_1, \dots, v_k independent if their flat is of dimension $k - 1$ (dimension means affine dimension), and say that v_1, \dots, v_k are dependent otherwise. A k -flat is an affine subspace of dimension k . In the

following V is a set of n distinct points in complex space \mathbb{C}^d . A k -flat is called ordinary if its intersection with V is contained in the union of a $(k - 1)$ -flat and a single point. A k -flat is elementary if its intersection with V has exactly $k + 1$ points. Notice that for $k = 1$ —the case of lines—the two notions of ordinary and elementary coincide.

For dimensions higher than one, there are two different definitions that generalize that of SG configuration. The first definition is based on ordinary k -flats. The second definition, which is less restricted than the first one, uses elementary k -flats (like in Hansen's theorem).

Definition 16: The set V is a δ -SG $_k^*$ configuration if for every independent $v_1, \dots, v_k \in V$ there are at least δn points $u \in V$ s.t. either $u \in \text{fl}(v_1, \dots, v_k)$ or the k -flat $\text{fl}(v_1, \dots, v_k, u)$ contains a point $w \in V$ outside $\text{fl}(v_1, \dots, v_k) \cup \{u\}$.

Definition 17: The set V is a δ -SG $_k$ configuration if for every independent $v_1, \dots, v_k \in V$ there are at least δn points $u \in V$ s.t. either $u \in \text{fl}(v_1, \dots, v_k)$ or the k -flat $\text{fl}(v_1, \dots, v_k, u)$ is not elementary.

Both definitions coincide with that of SG configuration when $k = 1$: Indeed, $\text{fl}(v_1) = v_1$ and $\text{fl}(v_1, u)$ is the line through v_1, u . Therefore, u is never in $\text{fl}(v_1)$ and the line $\text{fl}(v_1, u)$ is not elementary if it contains at least one point $w \notin \{v_1, u\}$.

We prove two high-dimensional versions of the SG theorem, each corresponding to one of the definitions above. Both theorems hold over the complex numbers (as well as over the real numbers). To the best of our knowledge, no complex-numbers version of Hansen's theorem was previously known.

Theorem 18. Let V be a δ -SG $_k^*$ configuration. Then $\dim(V) \leq f(\delta, k)$ with $f(\delta, k) = O((k/\delta)^2)$.

Theorem 19. Let V be a δ -SG $_k$ configuration. Then $\dim(V) \leq g(\delta, k)$ with $g(\delta, k) = 2^{Ck} / \delta^2$, with $C > 1$ a universal constant.

The proofs of the two theorems are below. Theorem 18 follows by an appropriate induction on the dimension, using the (one-dimensional) robust SG theorem. Theorem 19 follows by reduction to Theorem 18.

Before proving the theorems we set some notation. Fix some point $v_0 \in V$. By a normalization with respect to (w.r.t.) v_0 we mean a mapping $N : V \mapsto \mathbb{C}^d$ which first shifts all points by $-v_0$ (so that v_0 goes 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 . Note that all points on a single line through v_0 map to the same image under N . Observe also that for any such mapping N we have $\dim(N(V)) \geq \dim(V) - 1$ because the shifting can decrease the dimension by at most one and the scaling part maintains the dimension. Another property which holds for N is:

Claim 20. For such a mapping N we have that v_0, v_1, \dots, v_k are dependent iff $N(v_1), \dots, N(v_k)$ are dependent.

Proof: Because translation and scaling does not affect dependence, w.l.o.g. we assume that $v_0 = 0$ and that the distance of the hyperplane H from zero is one. Let h be the unit vector orthogonal to H . For all $i \in [k]$ we have $N(v_i) = v_i / \langle v_i, h \rangle$. Assume that v_0, v_1, \dots, v_k are dependent, that is, w.l.o.g. $v_k = \sum_{i \in [k-1]} a_i v_i$ for some a_1, \dots, a_{k-1} . For all $i \in [k - 1]$ define $b_i = a_i \langle v_i, h \rangle / \langle v_k, h \rangle$. Thus $N(v_k) = \sum_{i \in [k-1]} a_i v_i / \langle v_k, h \rangle = \sum_{i \in [k-1]} b_i N(v_i)$ where $\sum_{i \in [k-1]} b_i = 1$, which means that $N(v_1), \dots, N(v_k)$ are dependent. Because the map $a_i \mapsto b_i$ is invertible, the other direction of the claim holds as well.

We first prove the theorem for δ -SG $_k^*$ configurations.

Proof of Theorem 18: The proof is by induction on k . For $k = 1$ we know $f(\delta, 1) \leq c\delta^{-2}$ with $c > 1$ a universal constant. Suppose $k > 1$. We separate into two cases. The first case is when V is an $(\delta/(2k))$ -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 $\delta n/(2k)$ (a line is special if it contains at least three points). 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 assertion will complete the proof because $\dim(V) \leq \dim(V') + 1$. Indeed, if this assertion holds we have

$$f(\delta, k) \leq \max\{4c(k/\delta)^2, f((1 - 1/(2k))\delta, k - 1) + 1\}$$

and by induction we have $f(\delta, k) \leq 4c(k/\delta)^2$.

Fix $v'_1, \dots, v'_{k-1} \in V'$ to be $k - 1$ independent points (if no such tuple exists then V' is trivially a configuration). Let $v_1, \dots, v_{k-1} \in V$ be points s.t. $N(v_i) = v'_i$ for $i \in [k - 1]$. Claim 20 implies that v_0, v_1, \dots, v_{k-1} are independent. Thus, there is a set $U \subset V$ of size at least δn s.t. for every $u \in U$ either $u \in \text{fl}(v_0, v_1, \dots, v_{k-1})$ or the k -flat $\text{fl}(v_0, v_1, \dots, v_{k-1}, u)$ contains a point w outside $\text{fl}(v_0, v_1, \dots, v_{k-1}) \cup \{u\}$.

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 $u' = N(u)$. By Claim 20 if $u \in \text{fl}(v_0, v_1, \dots, v_{k-1})$ then u' is in $\text{fl}(v'_1, \dots, v'_{k-1})$. Otherwise, $\text{fl}(v_0, v_1, \dots, v_{k-1}, u)$ contains a point w outside $\text{fl}(v_0, v_1, \dots, v_{k-1}) \cup \{u\}$. Let $w' = N(w)$. We will show that w' is (a) contained in the $(k - 1)$ -flat $\text{fl}(v'_1, \dots, v'_{k-1}, u')$ and (b) is outside $\text{fl}(v'_1, \dots, v'_{k-1}) \cup \{u'\}$. Property (a) follows from Claim 20 because $v_0, v_1, \dots, v_{k-1}, u, w$ are dependent and so $v'_1, \dots, v'_{k-1}, u', w'$ are also dependent. To show (b) observe first that by Claim 20 the points $v'_1, \dots, v'_{k-1}, u'$ are independent (because $v_0, v_1, \dots, v_{k-1}, u$ are independent) and so u' is not in $\text{fl}(v'_1, \dots, v'_{k-1})$. We also need to show that $w' \neq u'$ but this follows from the fact that $u \neq w$ and so $w' = N(w) \neq N(u) = u'$ because 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.

We can now prove the theorem for δ -SG $_k$ configurations.

Proof of Theorem 19: The proof follows by induction on k (the case $k = 1$ is given by Theorem 3). Suppose $k > 1$. Suppose that $\dim(V) > g(\delta, k)$. We want to show that there exist k independent points v_1, \dots, v_k s.t. for at least $1 - \delta$ fraction of the points $w \in V$ we have that w is not in $\text{fl}(v_1, \dots, v_k)$ and the flat $\text{fl}(v_1, \dots, v_k, w)$ is elementary (i.e., does not contain any other point).

Let $k' = g(1, k - 1)$. By choice of g we have $g(\delta, k) > f(\delta, k' + 1)$ with f from Theorem 18. Thus, by Theorem 18, we can find $k' + 1$ independent points $v_1, \dots, v_{k'+1}$ 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 $\text{fl}(v_1, \dots, v_{k'+1})$ and the $(k' + 1)$ -flat $\text{fl}(v_1, \dots, v_{k'+1}, u)$ contains only one point, namely u , outside $\text{fl}(v_1, \dots, v_{k'+1})$.

We now apply the inductive hypothesis on the set $V \cap \text{fl}(v_1, \dots, v_{k'+1})$ which has dimension at least $k' = g(1, k - 1)$. This application gives us k independent points

v'_1, \dots, v'_k that define an elementary $(k - 1)$ -flat $\text{fl}(v'_1, \dots, v'_k)$. 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, \dots, v'_k gives us an elementary k -flat and so the theorem is proved.

A Fractional Motzkin–Rabin Theorem. The Motzkin–Rabin theorem is a two-color variant of the Sylvester–Gallai theorem. Here we prove a fractional version of it.

Definition 21: (δ -MR configuration). Let V_1, V_2 be two disjoint finite subsets of \mathbb{C}^d . Points in V_1 are of color 1 and points in V_2 are of color 2. A line is called *bichromatic* if it contains at least one point from each of the two colors. We say that V_1, V_2 are a δ -MR configuration if for every $i \in [2]$ and for every point $p \in V_i$, the bichromatic lines through p contain at least $\delta|V_i|$ points from V_i .

Theorem 22. Let $V_1, V_2 \subset \mathbb{C}^d$ be a δ -MR configuration. Then

$$\dim(V_1 \cup V_2) \leq O(1/\delta^4).$$

Proof: We will call a line passing through exactly two points in V_1 (respectively V_2) a V_1 -ordinary (respectively V_2 -ordinary) line. W.l.o.g. assume $|V_1| \leq |V_2|$. We separate the proof into two cases:

Case I is when V_2 is a $(\delta/2)$ -SG configuration. Then, by Theorem 3, $\dim(V_2) \leq O(1/\delta^2)$. If in addition $\dim(V_1) \leq 13/(\delta/2)^2$ then we are done. Otherwise, by Theorem 3, there exists a point $a_0 \in V_1$ such that there are at least $(1 - \delta/2)|V_1|$ V_1 -ordinary lines through a_0 . Let a_1, \dots, a_k denote the points in V_1 that belong to these lines with $k \geq (1 - \delta/2)|V_1|$. We now claim that $V_2 \cup \{a_0\}$ spans all the points in V_1 . This claim will suffice because, in this case, $\dim(V_2) \leq O(1/\delta^2)$. Let $a \in V_1$. Then, because V_1, V_2 is a δ -MR configuration, there are at least $\delta|V_1|$ points in V_1 such that the line through them and a contains a point in V_2 . One of these points must be among a_1, \dots, a_k , say it is a_1 . Because a is in the span of V_2 and a_1 and because a_1 is in the span of V_2 and a_0 we are done.

Case II is when V_2 is not a $(\delta/2)$ -SG configuration. In this case, there is a point $b \in V_2$ such that there are at least $(1 - \delta/2)|V_2|$ V_2 -ordinary lines through b . From this fact and from the δ -MR property, we get that $|V_1| \geq (\delta/2)|V_2|$ (there are at least $(\delta/2)|V_2|$ V_2 -ordinary lines through b that have an additional point from V_1 on them). This fact implies that the union $V_1 \cup V_2$ is a $(\delta^2/4)$ -SG configuration because $\delta|V_i| \geq (\delta^2/4)|V_1 \cup V_2|$, and the result follows by applying Theorem 3.

A Three-Color Variant. Having the flexibility of arguing about δ -SG configurations is also handy in proving theorems where there is no partial information. We demonstrate this flexibility by proving a three-color analog of the Motzkin–Rabin theorem.

Definition 23: (3MR configuration). Let V_1, V_2, V_3 be three pairwise disjoint finite subsets of \mathbb{C}^d , each of distinct points. We say that V_1, V_2, V_3 is a 3MR-configuration if every line ℓ so that $\ell \cap (V_1 \cup V_2 \cup V_3)$ has more than one point intersects at least two of the sets V_1, V_2, V_3 .

Theorem 24. Let V_1, V_2, V_3 be a 3MR configuration and denote $V = V_1 \cup V_2 \cup V_3$. Then

$$\dim(V) \leq O(1).$$

Proof: Assume w.l.o.g. that V_1 is not smaller than V_2, V_3 . Let $\alpha = 1/16$. There are several cases to consider:

V_1 is an α -SG configuration. By Theorem 3, the dimension of V_1 is at most

$$d_1 = O(1).$$

Consider the two sets

$$V'_2 = V_2 \setminus \text{span}(V_1) \quad \text{and} \quad V'_3 = V_3 \setminus \text{span}(V_1)$$

each is a set of distinct points in \mathbb{C}^d . Assume w.l.o.g. that $|V'_2| \geq |V'_3|$.

V'_2 is an α -SG configuration. By Theorem 3, the dimension of V'_2 is at most

$$d_2 = O(1).$$

Fix a point v_3 in V'_3 . For every point $v \neq v_3$ in V'_3 the line through v_3, v contains a point from $\text{span}(V_1) \cup V'_2$. Therefore,

$$\dim(V) \leq d_1 + d_2 + 1 \leq O(1).$$

V'_2 is not an α -SG configuration. There is a point v_2 in V'_2 so that for $k \geq |V'_2|/2$ of the points $v \neq v_2$ in V'_2 the line through v_2, v does not contain any other point from V'_2 . If $V'_2 = \text{span}(V_1, v_2)$ then the dimension of $V_1 \cup V_2$ is at most $d_1 + 1$ and we are done, as in the previous case. Otherwise, there is a point v'_2 in $V'_2 \setminus \text{span}(V_1, v_2)$.

We claim that in this case $|V'_3| \geq k/2$. Denote by P_2 the k points $v \neq v_2$ in V'_2 so that the line through v_2, v does not contain any other point from V'_2 . For every $v \in P_2$ there is a point $V_{1,3}(v)$ in $V_1 \cup V_3$ that is on the line through v, v_2 (the point v_2 is fixed). There are two cases to consider. (i) The first case is that for at least $k/2$ of the points v in P_2 we have $V_{1,3}(v) \in V'_3$. In this case clearly $|V'_3| \geq k/2$. (ii) The second case is that for at least $k/2$ of the points v in P_2 we have $V_{1,3}(v) \notin V'_3$. This fact means that these points $V_{1,3}(v)$ are in $\text{span}(V_1)$. Fix such a point $v \in P_2$ [which is in $\text{span}(V_1, v_2)$]. The line through v'_2, v contains a point v' from $V_1 \cup V_3$. The point v' is not in $\text{span}(V_1)$, as if it was then v'_2 would be in $\text{span}(v, v') \subseteq \text{span}(V_1, v_2)$. Therefore v' is in V'_3 . This fact also implies that $|V'_3| \geq k/2$.

Denote $V' = V'_2 \cup V'_3$. So we can conclude that for every v' in V' the special lines through v' contain at least $|V'|/8$ of the points in $V_1 \cup V_2 \cup V_3$. As in the proof of Theorem 3, we can thus define a family of triples T , each triple of three distinct collinear points in V , so that each v' in V' belongs to at least $|V'|/8$ triples in T and each two distinct v', v'' in V' belong to at most six triples.

By a slight abuse of notation, we also denote by V the matrix with rows defined by the points in V . Let V_1 be the submatrix of V with row defined by points in $\text{span}(V_1) \cap V$ and V' be the submatrix of V with row defined by points in V' . Use the triples in T to construct a matrix A so that $A \cdot V = 0$. Let A_1 be the submatrix of A consisting of the columns that correspond to $\text{span}(V_1) \cap V$ and A' be the submatrix of A consisting of the columns that correspond to V' . Therefore, $A' \cdot V' = -A_1 \cdot V_1$ which implies

$$\text{rank}(A' \cdot V') \leq \text{rank}(A_1 \cdot V_1) \leq d_1.$$

By the above discussion A' is a $(3, |V'|/8, 6)$ -design matrix and thus, by Theorem 5, has rank at least

$$|V'| - O(1)$$

and so

$$\dim(V') \leq O(1) + d_1 \leq O(1).$$

We can finally conclude that

$$\dim(V) \leq d_1 + \dim(V') \leq O(1).$$

V_1 is not an α -SG configuration. There is a point v_1 in V_1 so that for at least $|V_1|/2$ of the points $v \neq v_1$ in V_1 the line through

v_1, v does not contain any other point from V_1 . Assume w.l.o.g. that $|V_2| \geq |V_3|$. This fact implies that

$$|V_2| \geq |V_1|/4.$$

$|V_3| < |V_2|/16$. In this case the configuration defined by $V_1 \cup V_2$ is an α -SG configuration: Fix a point $v_1 \in V_1$ (a similar argument works for every $v_2 \in V_2$). We need to show that there are many special lines through v_1 . There are two options: Either (i) there are $|V_1|/2$ points $v'_1 \neq v_1$ in V_1 with a third point from V_1 on the line through v_1, v'_1 , or (ii) there are $|V_1|/2$ points $v'_1 \neq v_1$ in V_1 so that the line v_1, v'_1 contains a third point from $V_2 \cup V_3$. If case (i) holds, the point v_1 satisfies the required property. If case (ii) holds, because $|V_3| < |V_2|/16$, for at least $|V_1|/4$ points $v'_1 \in V_1$, the line through v_1, v'_1 contains a third point from V_2 . By Theorem 3, the dimension of $V_1 \cup V_2$ is at most

$$d_{1,2} = O(1).$$

Fix a point v_3 in V_3 . For every point $v \neq v_3$ in V_3 the line through v_3, v contains a point from $V_1 \cup V_2$. Therefore,

$$\dim(V) \leq d_{1,2} + 1.$$

$|V_3| \geq |V_2|/16$. In this case V is an α -SG configuration because $|V_i| \geq \alpha|V|$ for each $i \in \{1, 2, 3\}$. By Theorem 3, the dimension of V is thus at most $O(1)$.

Other Fields. In this section we show that our results can be extended from the complex field to fields of characteristic zero, and even to fields with very large positive characteristic. The argument is quite generic and relies on Hilbert's Nullstellensatz. We only discuss Theorem 5 because all other theorems follow from it over any field.

Definition 25: (T -matrix) Let m, n be integers and let $T \subset [m] \times [n]$. We call an $m \times n$ matrix A a T -matrix if all entries of A with indices in T are nonzero and all entries with indices outside T are zero.

Theorem 26. (Effective Hilbert's Nullstellensatz ref. 14). Let $g_1, \dots, g_s \in \mathbb{Z}[y_1, \dots, y_t]$ be degree d polynomials with coefficients in $\{0, 1\}$ and let

$$Z \triangleq \{y \in \mathbb{C}^t \mid g_i(y) = 0 \forall i \in [s]\}.$$

Suppose $h \in \mathbb{Z}[z_1, \dots, z_t]$ is another polynomial with coefficients in $\{0, 1\}$ which vanishes on Z . Then there exist positive integers p, q and polynomials $f_1, \dots, f_s \in \mathbb{Z}[y_1, \dots, y_t]$ such that

$$\sum_{i=1}^s f_i \cdot g_i \equiv p \cdot h^q.$$

Furthermore, one can bound p and the maximal absolute value of the coefficients of the f_i 's by an explicit function $H_0(d, t, s)$.

Theorem 27. Let m, n, r be integers and let $T \subset [m] \times [n]$. Suppose that all complex T -matrices have rank at least r . Let \mathbb{F} be a field of either characteristic zero or of finite large enough characteristic $p > P_0(n, m)$, where P_0 is some explicit function of n and m . Then, the rank of all T -matrices over \mathbb{F} is at least r .

Proof: Let $g_1, \dots, g_s \in \mathbb{C}[\{x_{ij} \mid i \in [m], j \in [n]\}]$ be the determinants of all $r \times r$ submatrices of an $m \times n$ matrix of variables $X = (x_{ij})$. The statement "all T -matrices have rank at least r " can be phrased as "if $x_{ij} = 0$ for all $(i, j) \notin T$ and $g_k(X) = 0$ for all $k \in [s]$ then $\prod_{(i,j) \in T} x_{ij} = 0$." That is, if all entries outside T are zero and X has rank smaller than r then it must have at least one zero entry also inside T . From Nullstellensatz we know that

there are integers $\alpha, \lambda > 0$ and polynomials f_1, \dots, f_s and h_{ij} , $(i, j) \notin T$, with integer coefficients such that

$$\alpha \cdot \left(\prod_{(i,j) \in T} x_{ij} \right)^\lambda \equiv \sum_{(i,j) \notin T} x_{ij} \cdot h_{ij}(X) + \sum_{k=1}^s f_k(X) \cdot g_k(X). \quad [2]$$

This identity implies the high rank of T -matrices also over any field F in which $\alpha \neq 0$. Because we have a bound on α in terms of n and m the result follows.

Corollary 28. *Theorem 3 holds over any field of characteristic zero or of sufficiently large (as a function of the number of points) positive characteristic.*

1. Sylvester JJ (1893) Mathematical question 11851. *Educ Times* 59:98.
2. Melchior E (1940) Über vielseitige der projektive ebene. *Deutsche Math* 5:461–475.
3. Erdos P (1943) Problems for solution: 4065–4069. *Am Math Monthly*, 50 pp:65–66.
4. Serre SP (1966) Problem 5359. *Am Math Monthly* 73:89.
5. Kelly LM (1986) A resolution of the Sylvester–Gallai problem of J. P. Serre. *Discrete Comput Geom* 1:101–104.
6. Pretorius Elkies LM, Swanepoel ND, Swanepoel KJ (2006) Sylvester–Gallai theorems for complex numbers and quaternions. *Discrete Comput Geom* 35:361–373.
7. Barak B, Dvir Z, Yehudayoff A, Wigderson A (2011) Rank bounds for design matrices with applications to combinatorial geometry and locally correctable codes. *Proceedings of the 43rd Annual ACM Symposium on Theory of Computing, STOC '11* (ACM, New York), pp 519–528.
8. Hilton AJW (1973) On double diagonal and cross Latin squares. *J Lond Math Soc* s2-6:679–689.
9. Sinkhorn R (1964) A relationship between arbitrary positive matrices and doubly stochastic matrices. *Ann Math Statist* 35:876–879.
10. Linial N, Samorodnitsky A, Wigderson A (2000) A deterministic strongly polynomial algorithm for matrix scaling and approximate permanents. *Combinatorica* 20:545–568.
11. Rothblum U, Schneider H (1989) Scaling of matrices which have prespecified row sums and column sums via optimization. *Lin Algebra Appl* 114–115:737–764.
12. Alon N (2009) Perturbed identity matrices have high rank: Proof and applications. *Combinator Probab Comput* 18:3–15.
13. Hansen S (1965) A generalization of a theorem of Sylvester on the lines determined by a finite point set. *Math Scand* 16:175–180.
14. Kollár J (1988) Sharp effective Nullstellensatz. *J Am Math Soc* 1:963–975.

The meaning of “sufficiently large” in the corollary does not depend on n , the underlying dimension, via a random linear projection of the m points to an m -dimensional space (if $n > m$).

ACKNOWLEDGMENTS. We thank Moritz Hardt for many helpful conversations. We thank Jozsef Solymosi for helpful comments. B.B was at Princeton University during most of the research presented in this paper and was supported by NSF grants CNS-0627526, CCF-0426582 and CCF-0832797, and the Packard and Sloan fellowships. Z.D was partially supported by NSF grant CCF-0832797 and by the Packard fellowship. A.W was partially supported by NSF grants CCF-0832797 and DMS-0835373. A.Y was at the Institute for Advanced Study, Princeton, during most of the research presented in this paper and was partially supported by NSF grants CCF-0832797 and DMS-0835373.