On the Number of Ordinary Lines Determined by Sets in Complex Space

Abdul Basit\textsuperscript{1,5} · Zeev Dvir\textsuperscript{2,3} · Shubhangi Saraf\textsuperscript{1,4} · Charles Wolf\textsuperscript{4,6}

Received: 1 March 2017 / Revised: 23 September 2018 / Accepted: 28 September 2018 / Published online: 24 October 2018
© Springer Science+Business Media, LLC, part of Springer Nature 2018

Abstract
Kelly’s theorem states that a set of \( n \) points affinely spanning \( \mathbb{C}^3 \) must determine at least one ordinary complex line (a line incident to exactly two of the points). Our main theorem shows that such sets determine at least \( 3n/2 \) ordinary lines, unless the configuration has \( n - 1 \) points in a plane and one point outside the plane (in which case there are at least \( n - 1 \) ordinary lines). In addition, when at most \( n/2 \) points are contained in any plane, we prove stronger bounds that take advantage of the existence of lines with four or more points (in the spirit of Melchior’s and Hirzebruch’s inequalities). Furthermore, when the points span four or more dimensions, with at most \( n/2 \) points contained in any three-dimensional affine subspace, we show that there must be a quadratic number of ordinary lines.

Keywords
Ordinary lines · Sylvester–Gallai · Combinatorial geometry · Designs
1 Introduction

Let $V = \{v_1, v_2, \ldots, v_n\}$ be a set of $n$ points in $\mathbb{C}^d$. We denote by $L(V)$ the set of lines determined by points in $V$, and by $L_r(V)$ (resp. $L_{\geq r}(V)$) the set of lines in $L(V)$ that contain exactly (resp. at least) $r$ points. Let $t_r(V)$ denote the size of $L_r(V)$. Throughout the write-up we omit the argument $V$ when the context makes it clear. We refer to $L_2$ as the set of ordinary lines, and $L_{\geq 3}$ as the set of special lines.

A well-known result in combinatorial geometry is the Sylvester–Gallai theorem.

**Theorem 1.1** (Sylvester–Gallai theorem) Let $V$ be a set of $n$ points in $\mathbb{R}^2$ not all on a line. Then there exists an ordinary line determined by points of $V$.

The statement was conjectured by Sylvester in 1893 [23], and the first published proof is by Melchior [18]. Later proofs were given by Gallai in 1944 [11] and others; there are now several different proofs of the theorem. Of particular interest is the following result by Melchior [18].

**Theorem 1.2** (Melchior’s inequality) Let $V$ be a set of $n$ points in $\mathbb{R}^2$ that are not collinear. Then

$$t_2(V) \geq 3 + \sum_{r \geq 4} (r - 3)t_r(V).$$

Theorem 1.2 in fact proves something stronger than the Sylvester–Gallai theorem, i.e., there are at least three ordinary lines. A natural question to ask is how many ordinary lines must a set of $n$ points, not all on a line, determine. This led to what is known as the Dirac–Motzkin conjecture.

**Conjecture 1.3** (Dirac–Motzkin conjecture) For every $n \neq 7, 13$, the number of ordinary lines determined by $n$ non-collinear points in the plane is at least $\lceil n/2 \rceil$.

There were several results on this question (see [6,16,19]), before Green and Tao [12] resolved it for large enough point sets.

**Theorem 1.4** (Green and Tao [12]) Let $V$ be a set of $n$ points in $\mathbb{R}^2$, not all on a line. Suppose that $n \geq n_0$ for a sufficiently large absolute constant $n_0$. Then $t_2(V) \geq n/2$ for even $n$ and $t_2(V) \geq \lfloor 3n/4 \rfloor$ for odd $n$.

There are several survey articles on the topic, see for example [4], and [12] provides a nice history of the problem.

The Sylvester–Gallai theorem is not true when the field $\mathbb{R}$ is replaced by $\mathbb{C}$. The well-known Hesse configuration, realized by the nine inflection points of a non-degenerate cubic, provides a counterexample. A more general example is the following:

**Examples 1.5** (Fermat configuration) For any positive integer $k \geq 3$, let $V$ be inflection points of the Fermat Curve $X^k + Y^k + Z^k = 0$ in $\mathbb{P}^2\mathbb{C}$. Specifically, let

$$V_1 = \bigcup_{i=1}^{k} \{[0 : 1 : \omega^i]\}, \quad V_2 = \bigcup_{i=1}^{k} \{[\omega^i : 0 : 1]\}, \quad \text{and} \quad V_3 = \bigcup_{i=1}^{k} \{[1 : \omega^i : 0]\},$$
where \( \omega \) is a \( k \)th root of \(-1\). Then \( \mathcal{V} = \mathcal{V}_1 \cup \mathcal{V}_2 \cup \mathcal{V}_3 \).

The point set \( \mathcal{V} \) does not determine any ordinary lines. It is easy to check that the sets \( \mathcal{V}_1 \), \( \mathcal{V}_2 \) and \( \mathcal{V}_3 \) lie on the lines \( X = 0 \), \( Y = 0 \), and \( Z = 0 \), respectively. Every other line contains exactly three points. Indeed, consider, without loss of generality, points \( u = [0 : 1 : \omega^i] = [0 : \omega^{-i} : 1] \in \mathcal{V}_1 \) and \( v = [\omega^j : 0 : 1] \in \mathcal{V}_2 \). Then the point \( w = [-\omega^j : \omega^{-i} : 0] = [1 : -\omega^{-i-j} : 0] = [1 : \omega^{k-i-j} : 0] \in \mathcal{V}_3 \) is collinear with \( u \) and \( v \). \(^1\)

In response to a question of Serre [22], Kelly [15] showed that when the points span more than two dimensions, the point set must determine at least one ordinary line.

**Theorem 1.6** (Kelly’s theorem [15]) Let \( \mathcal{V} \) be a set of \( n \) points in \( \mathbb{C}^3 \) that are not contained in a plane. Then there exists an ordinary line determined by points of \( \mathcal{V} \).

Kelly’s proof of Theorem 1.6 used a deep result of Hirzebruch [14], referred to as Hirzebruch’s inequality, from algebraic geometry. More elementary proofs of Theorem 1.6 were given in [10] and [9]. To the best of our knowledge, no lower bound greater than one is known for the number of ordinary lines determined by point sets spanning \( \mathbb{C}^3 \). Improving on the techniques of [9], we make the first progress in this direction.

**Theorem 1.7** Let \( \mathcal{V} \) be a set of \( n \geq 24 \) points in \( \mathbb{C}^3 \) not contained in a plane. Then \( \mathcal{V} \) determines at least \( 3n/2 \) ordinary lines, unless \( n - 1 \) points are on a plane in which case there are at least \( n - 1 \) ordinary lines.

In the latter case, let \( v \) be the point not on the plane. Every line determined by \( \mathcal{V} \) containing \( v \) is ordinary, so there are at least \( n - 1 \) ordinary lines. On the other hand, it is possible that there are exactly \( n - 1 \) ordinary lines. In particular, let \( \mathcal{V} \) consist of the Fermat Configuration, for some \( k \geq 3 \), on a plane, and one point \( v \) not on the plane. Then \( \mathcal{V} \) has \( 3k + 1 \) points, and the only ordinary lines determined by \( \mathcal{V} \) are lines that contain \( v \), so there are exactly \( 3k \) ordinary lines. We are not aware of any examples that achieve the \( 3n/2 \) bound when at most \( n - 2 \) points are contained in any plane. Using a similar argument, for point sets in \( \mathbb{R}^3 \), Theorems 1.4 and 1.7 give us the following easy corollary.

**Corollary 1.8** Let \( \mathcal{V} \) be a set of \( n \) points in \( \mathbb{R}^3 \) not contained in a plane. Suppose that \( n \geq n_0 \) for a sufficiently large absolute constant \( n_0 \). Then \( \mathcal{V} \) determines at least \( 3(n - 1)/2 \) ordinary lines.

Hirzebruch’s inequality is similar in spirit to Melchior’s inequality and gives bounds on the number of ordinary lines by taking into consideration the number of lines with four or more points.

**Theorem 1.9** (Hirzebruch’s inequality [14]) Let \( \mathcal{V} \) be a set of \( n \) points in \( \mathbb{C}^2 \), such that \( t_n(\mathcal{V}) = t_{n-1}(\mathcal{V}) = t_{n-2}(\mathcal{V}) = 0 \). Then

\[
\begin{align*}
t_2(\mathcal{V}) + \frac{3}{4} t_3(\mathcal{V}) & \geq n + \sum_{r \geq 5} (2r - 9)t_r(\mathcal{V}).
\end{align*}
\]

\(^1\) We note that while the Fermat configuration as stated lives in the projective plane, it can be made affine by any projective transformation that moves a line with no points to the line at infinity.
More recently, Bojanowski [3] and Pokora [20] used a theorem of Langer [17] to prove the following result.

**Theorem 1.10** (Langer’s Inequality) Let $V$ be a set of $n$ points in $\mathbb{C}^2$, such that $t_j(V) = 0$ for $j > 2n/3$. Then

$$t_2(V) + \frac{3}{4} t_3(V) \geq n + \sum_{r \geq 5} \frac{r^2 - 4r}{4} t_r(V).$$

While imposing a stricter condition on the point set, Theorem 1.10 is strictly better than Theorem 1.9 as long as there is a line containing at least five points. For more detail on Langer’s inequality and combinatorial implications, see [8]. In $\mathbb{C}^3$, when $V$ is sufficiently non-degenerate, i.e., no plane contains too many points, we are able to give a more refined bound by taking into account the existence of lines with more than three points. In particular, we show the following (the constant $1/2$ in Theorem 1.11 is arbitrary and can be replaced by any positive constant smaller than 1):

**Theorem 1.11** There exists an absolute constant $c > 0$ and a positive integer $n_0$ such that the following holds. Let $V$ be a set of $n \geq n_0$ points in $\mathbb{C}^3$ with at most $n/2$ points contained in any plane. Then

$$t_2(V) \geq \frac{3}{2} n + c \sum_{r \geq 4} r^2 t_r(V).$$

Suppose that $V$ consists of $n - k$ points on a plane, and $k$ points not on the plane. There are at least $n - k$ lines incident to each point not on the plane, at most $k - 1$ of which could contain three or more points. So we get that there are at least $k(n - 2k)$ ordinary lines determined by $V$. Then if $k = \varepsilon n$, for $0 < \varepsilon < 1/2$, we get that $V$ has $\Omega(\varepsilon) (n^2)$ ordinary lines, where the hidden constant depends on $\varepsilon$. Therefore, the bound in Theorem 1.11 is only interesting when no plane contains too many points.

On the other hand, having at most a constant fraction of the points on any plane is necessary to obtain a bound as in Theorem 1.11. Indeed, let $V$ consist of the Fermat Configuration for some $k \geq 3$ on a plane and $o(k)$ points not on the plane. Then $V$ has $O(k)$ points and determines $o(k^2)$ ordinary lines. On the other hand, $\sum_{r \geq 4} r^2 t_r(V) = \Omega(k^2)$.

Both Hirzebruch’s and Langer’s inequalities (which also give bounds in $\mathbb{C}^3$, though without requiring that every plane contains not too many points) only give lower bounds on $t_2(V) + 3t_3(V)/4$, whereas both Theorems 1.7 and 1.11 give lower bounds on the number of ordinary lines, i.e., $t_2(V)$. We also note that lines with four points do not play any role in the previous inequalities, where the summation starts at $r = 5$. This is not the case for Theorem 1.11, which also takes into account lines with four points.

Finally, when a point set $V$ spans four or more dimensions in a sufficiently non-degenerate manner, i.e., no three-dimensional affine subspace contains too many points, we prove that there must be a quadratic number of ordinary lines.
Theorem 1.12 There exists an absolute constant $c' > 0$ and a positive integer $n_0$ such that the following holds. Let $V$ be a set of $n \geq n_0$ points in $\mathbb{C}^4$ with at most $n/2$ points contained in any three-dimensional affine subspace. Then

$$t_2(V) \geq c'n^2.$$ 

Here, again, the constant $1/2$ is arbitrary and can be replaced by any positive constant less than 1. However, increasing this constant will shrink the constant $c'$ in front of $n^2$.

While we state Theorems 1.7 and 1.11 over $\mathbb{C}^3$ and Theorem 1.12 over $\mathbb{C}^4$, the same bounds hold in higher dimensions as well, since projecting a point set in $\mathbb{C}^d$ onto a generic lower dimensional subspace can only increase the number of incidences. A quadratic lower bound may also be possible under the weaker hypothesis of at most $n/2$ points are contained in any two-dimensional space, but we have no proof or counterexample. A recent result of de Zeeuw [7] resolves this question for points over the reals.

Theorem 1.13 For every $\alpha < 1$ there is a $c_{\alpha} > 0$ such that, if a set $P$ of $n$ points in $\mathbb{R}^3$ has at most $\alpha n$ points on any plane, then $P$ spans at least $c_{\alpha}n^2$ ordinary lines.

Organization: In Sect. 2, we give a short overview of the new ideas in our proof (which builds upon [9]). In In Sect. 3, we develop the necessary machinery on matrix scaling and Latin squares. In Sect. 4, we prove some key lemmas that will be used in the proofs of our main results. Section 5 gives the proof of Theorems 1.7 and 1.12, which are considerably simpler than Theorem 1.11. In Sect. 6, we develop additional machinery needed for the proof of Theorem 1.11. The proof of Theorem 1.11 is presented in Sect. 7.

2 Proof Overview

The starting point for the proofs of Theorems 1.7, 1.11 and 1.12 is the method developed in [2,9] which uses rank bounds for design matrices—matrices in which the supports of different columns do not intersect in too many positions. We augment the techniques in these papers in several ways which give us more flexibility in analyzing the number of ordinary lines. We devote this short section to an overview of the general framework (starting with [9]) outlining the places where new ideas come into play.

Let $V = \{v_1, \ldots, v_n\}$ be points in $\mathbb{C}^d$ and denote by $V$ the $n \times (d + 1)$ matrix whose $i$th row is the vector $(v_i, 1) \in \mathbb{C}^{d+1}$, i.e., the vector obtained by appending a 1 to the vector $v_i$. The dimension of the (affine) space spanned by the point set can be seen to be equal to $\text{rank}(V) - 1$. We would now like to argue that too many collinearities in $V$ (or too few ordinary lines) imply that all (or almost all) points of $V$ must be contained in a low dimensional affine subspace, i.e., $\text{rank}(V)$ is small. To do this, we construct a matrix $A$, encoding the dependencies in $V$, such that $AV = 0$. Then we must have

$$\text{rank}(V) \leq n - \text{rank}(A),$$

and so it suffices to lower bound the rank of $A$. 

\(\mathbb{C}\) Springer
We construct the matrix $A$ in the following manner so that each row of $A$ corresponds to a collinear triple in $\mathcal{V}$. For any collinear triple $\{v_i, v_j, v_k\}$, there exist coefficients $a_i, a_j, a_k$ such that $a_i v_i + a_j v_j + a_k v_k = 0$. We can thus form a row of $A$ by taking these coefficients as the non-zero entries in the appropriate columns. By carefully selecting the triples using constructions of Latin squares (see Lemma 3.11), we can ensure that $A$ is a design matrix. Roughly speaking, this means that the supports of every two columns in $A$ intersect in a small number of positions. Equivalently, every pair of points appears together only in a small number of triples.

The proof in [9] now proceeds to prove a general rank lower bound on any such design matrix. To understand the new ideas in our proof, we need to “open the box” and see how the rank bound from [9] is actually proved. To get some intuition, suppose that $A$ is a matrix with 0/1 entries. To bound the rank of $A$, we can consider the matrix $M = A^*A$ and note that $\text{rank}(M) = \text{rank}(A)$. Since $A$ is a design matrix, $M$ has the property that the diagonal entries are very large (since we can show that each point is in many collinear triples) and that the off-diagonal elements are very small (since columns have small intersections). Matrices with this property are called diagonal dominant matrices, and it is easy to lower bound their rank using trace inequalities (see Lemma 3.5).

However, the matrix $A$ that we construct could have entries of arbitrary magnitude and so bounding the rank requires more work. To do this, [9] relies on matrix scaling techniques. We are allowed to multiply each row and each column of $A$ by a non-zero scalar and would like to reduce to the case where the entries of $A$ are “mostly balanced” (see Theorem 3.3 and Corollary 3.4). Once scaled, we can consider $M = A^*A$ as before and use the bound for diagonal dominant matrices.

Our proof introduces two new main ideas into this picture. The first idea has to do with the conditions needed to scale $A$. It is known (see Corollary 3.4) that a matrix $A$ has a good scaling if it does not contain a “too large” zero submatrix. This is referred to as having Property-$S$ (see Definition 3.2). The proof of [9] uses $A$ to construct a new matrix $B$, whose rows are the same as those of $A$ but with some rows repeating more than once. Then one shows that $B$ has Property-$S$ and continues to scale $B$ (which has rank at most that of $A$) instead of $A$. This loses the control on the exact number of rows in $A$ which is crucial for bounding the number of ordinary lines. We instead perform a more careful case analysis: If $A$ has Property-$S$ then we scale $A$ directly and gain more information about the number of ordinary lines. If $A$ does not have Property-$S$, then we carefully examine the large zero submatrix that violates Property-$S$. Such a zero submatrix corresponds to a set of points and a set of lines such that no line is incident to any of the points. We argue in Lemma 4.4 that such a submatrix implies the existence of many ordinary lines. In fact, the conclusion is slightly more delicate: We either get many ordinary lines (in which case we are done) or we get a point with many ordinary lines incident to it (but not enough to complete the proof). In the second case, we need to perform an iterative argument which removes the point we found and applies the same argument again to the remaining points.

The second new ingredient in our proof comes into play only in the proof of Theorem 1.11. Here, our goal is to improve on the rank bound of [9] using the existence of lines with four or more points. Recall that our goal is to give a good upper bound on the off-diagonal entries of $M = A^*A$. Consider the $(i, j)$th entry of $M$, obtained
by taking the inner product of columns $i$ and $j$ in $A$. The $i$th column of $A$ contains the coefficients of $v_i$ in a set of collinear triples containing $v_i$ (we might not use all collinear triples). In [9] this inner product is bounded using the Cauchy–Schwartz inequality, and uses the fact that we picked our triple family carefully so that $v_i$ and $v_j$ appear together in a small number of collinear triples. This does not use any information about possible cancellations that may occur in the inner product (considering different signs over the reals or angles of complex numbers). One of the key insights of our proof is to notice that having more than three points on a line, gives rise to such cancellations (which increase the more points we have on a single line).

To get a rough idea, let us focus on a set of points over the reals. Consider two points $v_1, v_2$ on a line that has two more points $v_3, v_4$ on it. Suppose that $v_3$ is ‘between’ $v_1$ and $v_2$ and that $v_4$ is outside the interval $v_1, v_2$. Then, in the collinearity equation for the triple $v_1, v_2, v_3$ the signs of the coefficients of $v_1, v_2$ will be the same. On the other hand, in the collinearity equation for $v_1, v_2, v_4$ the signs of the coefficients of $v_1, v_2$ will be different (one will be positive and the other negative). Thus, if both of these triples appear as rows of $A$, we will have non-trivial cancellations! Of course, we need to also worry about the magnitudes of the coefficients but, luckily, this is possible since, if the coefficients are of magnitudes that differ from each other too much, we can “win” in another Cauchy–Schwartz (which again translates into a better rank bound, see Lemma 6.6). To formalize the previous example, let $v_1, v_2, v_3, v_4$ be collinear points in $\mathbb{R}^d$. Then there exist coefficients such that

$$r \cdot v_1 + (1 - r) \cdot v_2 - v_3 = 0,$$

$$s \cdot v_1 + (1 - s) \cdot v_2 - v_4 = 0,$$

and

$$t \cdot v_1 + (1 - t) \cdot v_3 - v_4 = 0.$$  

Now at least one of $r(1 - r), s(1 - s)$ and $t(1 - t)$ must be negative, and at least one must be positive. Without loss of generality, say $r(1 - r)$ is positive and $s(1 - s)$ is negative. In order for the Cauchy–Schwarz inequality to be tight, we need that $r(1 - r) = s(1 - s)$, which cannot happen because one is positive and the other is negative. This phenomena is captured in Lemma 6.3 which generalizes this idea to the complex numbers. The lemma only analyzes the case of four points since we can bootstrap the lemma for lines with more points by applying it to a random four tuple (see Item 4 of Lemma 6.4).

3 Preliminaries

3.1 Matrix scaling and Rank Bounds

One of the main ingredients in our proof is rank bounds for design matrices. These techniques were first used for incidence type problems in [2] and improved upon in [9]. We first set up some notation. For a complex matrix $A$, let $A^*$ denote the matrix conjugated and transposed. Let $A_{ij}$ denote the entry in the $i$th row and $j$th column of $A$, and $R_k := R_k(A)$ and $C_i := C_i(A)$ denote the rows and column, respectively, of $A$. For
two complex vectors $u, v \in \mathbb{C}^d$, we denote their inner product by $\langle u, v \rangle = \sum_{i=1}^{d} u_i \cdot \overline{v_i}$. The support of a vector $u \in \mathbb{R}^n$ is defined to be $\text{supp}(u) = \{i \in [n] : u_i \neq 0\}$.

Central to obtaining the rank bounds for matrices is the notion of matrix scaling. We now introduce this notion and provide some definitions and lemmas.

**Definition 3.1** (Matrix scaling) Let $A$ be an $m \times n$ matrix over some field $\mathbb{F}$. For every $\rho \in \mathbb{F}^m$ and $\gamma \in \mathbb{F}^n$ with all entries non-zero, the matrix $A'$ with $A'_{ij} = A_{ij} \cdot \rho_i \cdot \gamma_j$ is referred to as a scaling of $A$. We refer to the elements of the vectors $\rho$ and $\gamma$ as the scaling coefficients. Note that two matrices that are scalings of each other have the same rank and the same pattern of zero and non-zero entries.

We will be interested in scalings of matrices that control the row and column sums. The following property provides a sufficient condition under which such scalings exist.

**Definition 3.2** (Property-S) Let $A$ be an $m \times n$ matrix over some field. We say that $A$ satisfies Property-S if for every zero submatrix of size $a \times b$, we have $a \cdot m + b \cdot n \leq 1$.

The following theorem is given in [21].

**Theorem 3.3** (Matrix Scaling Theorem) Let $A$ be an $m \times n$ real matrix with non-negative entries satisfying Property-S. Then, for every $\epsilon > 0$, there exists a scaling $A'$ of $A$ such that the sum of every row of $A'$ is at most $1 + \epsilon$, and the sum of every column of $A'$ is at least $m/n - \epsilon$. Moreover, the scaling coefficients are all positive real numbers.

In fact, in the above theorem, we may assume that the sum of every row of $A'$ is exactly $1 + \epsilon$, since scaling the rows to make the row sums $1 + \epsilon$ can only increase the column sums.

The following corollary to Theorem 3.3 appeared in [2].

**Corollary 3.4** ($\ell_2$ Scaling) Let $A$ be an $m \times n$ complex matrix satisfying 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 $j \in [n]$

$$\sum_{i \in [m]} |A'_{ij}|^2 \geq \frac{m}{n} - \epsilon.$$

Moreover, the scaling coefficients are all positive real numbers.

Corollary 3.4 is obtained by applying Theorem 3.3 to the matrix obtained by squaring the absolute values of the entries of the matrix $A$. Once again, we may assume that $\sum_{j \in [n]} |A'_{ij}|^2 = 1 + \epsilon$. 
To bound the rank of a matrix $A$, we will bound the rank of the matrix $M = A^*A'$, where $A'$ is some scaling of $A$. Then we have that $\text{rank}(A) = \text{rank}(A') = \text{rank}(M)$. We use Corollary 3.4, along with rank bounds for diagonal dominant matrices. The following lemma is a variant of a folklore lemma on the rank of diagonal dominant matrices (see [1]) and appeared in this form in [9].

**Lemma 3.5** Let $A$ be an $n \times n$ complex hermitian matrix, such that $|A_{ii}| \geq L$ for all $i \in n$. Then

$$\text{rank}(A) \geq \frac{n^2 L^2}{nL^2 + \sum_{i \neq j} |A_{ij}|^2}.$$

The matrix scaling theorem allows us to control the $\ell_2$ norms of the columns and rows of $A'$, which in turn allows us to bound the sums of squares of entries of $M$. For this, we use a variation of a lemma from [9]. While the proof idea is the same, our proof requires a somewhat more careful analysis.

For any $x_1, x_2, \ldots, x_n \in \mathbb{R}$, a standard application of the Cauchy–Schwarz inequality implies

$$\left( \sum_{i=1}^{n} x_i \right)^2 \leq n \sum_{i=1}^{n} x_i^2,$$

with equality if and only if $x_1 = x_2 = \cdots = x_n$. As mentioned in Sect. 2, a key ingredient in our proof is to use information about cancellations in the inner product. To this end, we make use of the following:

$$\left( \sum_{i=1}^{n} x_i \right)^2 + \sum_{i<j} (x_i - x_j)^2 = n \sum_{i=1}^{n} x_i^2.$$

The latter term on the LHS in (2) quantifies the difference between the terms on two sides of (1). The proof for Lemma 3.7 will use (2) twice. To simplify the presentation, we need the following definitions:

**Definition 3.6** Let $A$ be an $m \times n$ matrix over $\mathbb{C}$, and let $\text{supp}(i, j) := \text{supp}(C_i) \cap \text{supp}(C_j)$. Then we define:

$$D(A) := \sum_{i \neq j} \sum_{k < k'}^{\text{supp}(i, j)} |A_{ki} \overline{A_{kj}} - A_{k'i} \overline{A_{k'j}}|^2,$$

and

$$E(A) := \sum_{k=1}^{m} \sum_{i < j}^{\text{supp}(R_k)} \left( |A_{ki}|^2 - |A_{kj}|^2 \right)^2.$$

Note that both $D(A)$ and $E(A)$ are non-negative real numbers.

**Discussion on $D(A)$ and $E(A)$:** To give some insight into the proof, we include a brief discussion about $D(A)$ and $E(A)$. Consider the matrix

$$\text{ Springer}$$
and let $M = A^* A$. Observe that $M_{ij}^2 = (C_i, C_j)^2 = \left(\sum_{k=1}^{4} A_{ki} A_{kj}\right)^2$. Equation (1) implies that $\left(\sum_{k=1}^{4} A_{ki} A_{kj}\right)^2 \leq 4 \sum_{k=1}^{4} A_{ki}^2 A_{kj}^2$. An obvious improvement here comes from the fact the supports of any two columns intersect in only two entries, giving $\left(\sum_{k=1}^{4} A_{ki} A_{kj}\right)^2 \leq 2 \sum_{k=1}^{4} A_{ki}^2 A_{kj}^2$. For $i = 1, j = 2$, this gives $(1 + 1)^2 = 4 \leq 2(1^2 + 1^2) = 4$ and is tight because $A_{11} A_{12} = A_{21} A_{22} = 1$. For $i = 1, j = 3$, it implies $(2 + 1)^2 = 9 \leq 2(2^2 + 2^2) = 10$. The difference is precisely captured by $(A_{11} A_{13} - A_{21} A_{23})^2 = (2 - 1)^2 = D(A)$.

$E(A)$ comes into the picture in bounding $\sum_{i=1}^{4} A_{ki}^4$, for fixed $k$. Equation (1) implies $\sum_{i=1}^{4} A_{ki}^4 \geq \frac{1}{3} \left(\sum_{i=1}^{4} A_{ki}^2\right)^2$. Equality holds if and only if all non-zero entries in row $k$ are equal, e.g., when $k \in \{2, 3, 4\}$. Otherwise, e.g., when $k = 1$, $E(A)$ quantifies the difference.

In Sect. 6, we will show that for the dependency matrix, each row will either have all non-zero entries far from each other, i.e., it will contribute to $E(A)$, or it will, together with other rows, contribute to $D(A)$.

We now give the main result for this section.

**Lemma 3.7** Let $A$ be an $m \times n$ matrix over $\mathbb{C}$. Suppose that each row of $A$ has $\ell_2$ norm $< \alpha$, the supports of every two columns of $A$ intersect in at most $t$ locations, and the size of the support of every row is at most $q$. Let $M = A^* A$. Then

$$
\sum_{i \neq j} |M_{ij}|^2 \leq \left(1 - \frac{1}{q}\right) t m q^4 - \left(D(A) + \frac{t}{q} E(A)\right).
$$

**Proof** Note that

$$
\sum_{i \neq j} |M_{ij}|^2 = \sum_{i \neq j} |(C_i, C_j)|^2 = \sum_{i \neq j} \left| \sum_{k=1}^{m} A_{ki} A_{kj} \right|^2 = \sum_{i \neq j} \sum_{k \in \text{supp}(i, j)} \left| A_{ki} A_{kj} \right|^2.
$$

Here the last equality uses the fact that the only terms contributing to the sum in the inner summation come from $k \in \text{supp}(i, j)$. Since the supports of any two columns of $A$ intersect in at most $t$ locations, $|\text{supp}(i, j)| \leq t$ for all $i \neq j$. Combining this with (2) gives:

$$
\sum_{i \neq j} \sum_{k \in \text{supp}(i, j)} \left| A_{ki} A_{kj} \right|^2 \leq \sum_{i \neq j} \left( t \sum_{k \in \text{supp}(i, j)} |A_{ki}|^2 |A_{kj}|^2 - \sum_{k < k'} \left| A_{ki} A_{kj} - A_{k'i} A_{k'j} \right|^2 \right)
$$

\(\square\) Springer
\[
\sum_{i \neq j} \left( t \sum_{k=1}^{m} |A_{ki}|^2 |A_{kj}|^2 - \sum_{k<k'}_{k,k' \in \text{supp}(i,j)} |A_{ki} A_{kj} - A_{k'i} A_{k'j}|^2 \right)
\]
\[
= t \sum_{i \neq j} \sum_{k=1}^{m} |A_{ki}|^2 |A_{kj}|^2 - D(A)
\]
\[
= t \sum_{k=1}^{m} \left( \sum_{i=1}^{n} |A_{ki}|^2 \right)^2 - t \sum_{k=1}^{m} \left( \sum_{i=1}^{n} |A_{ki}|^4 \right) - D(A).
\]

We now use (2) again to bound the second term in the expression above. This time we
may restrict our attention to the at most \( q \) non-zero entries in each row. This gives
\[
\sum_{i \neq j} |M_{ij}|^2 \leq t \sum_{k=1}^{m} \left( \sum_{i=1}^{n} |A_{ki}|^2 \right)^2 - t \sum_{k=1}^{m} \frac{1}{q} \left( \sum_{i=1}^{n} |A_{ki}|^2 \right)^2
\]
\[
+ \sum_{i < j}_{i,j \in \text{supp}(R_k)} \left( |A_{ki}|^2 - |A_{kj}|^2 \right)^2 - D(A)
\]
\[
= \left( 1 - \frac{1}{q} \right) t \sum_{k=1}^{m} \left( \sum_{i=1}^{n} |A_{ki}|^2 \right)^2
\]
\[
- \frac{t}{q} \sum_{k=1}^{m} \sum_{i < j}_{i,j \in \text{supp}(R_k)} \left( |A_{ki}|^2 - |A_{kj}|^2 \right)^2 - D(A)
\]
\[
= \left( 1 - \frac{1}{q} \right) t \sum_{k=1}^{m} \left( \sum_{i=1}^{n} |A_{ki}|^2 \right)^2 - \frac{t}{q} E(A) - D(A)
\]
\[
= \left( 1 - \frac{1}{q} \right) t m \alpha^4 - \left( D(A) + \frac{t}{q} E(A) \right). \tag*{\square}
\]

From this, we get the following easy corollary.

**Corollary 3.8** Let \( A \) be an \( m \times n \) matrix over \( \mathbb{C} \). Suppose that each row of \( A \) has \( \ell_2 \) norm \(<\alpha\), the supports of every two columns of \( A \) intersect in at most \( t \) locations, and the size of the support of every row is at most \( q \). Let \( M = A^* A \). Then
\[
\sum_{i \neq j} |M_{ij}|^2 \leq \left( 1 - \frac{1}{q} \right) t m \alpha^4.
\]

### 3.2 Latin Squares

Latin squares play a central role in our proof. While Latin squares play a role in both [9] and [2], our proof exploits their design properties more strongly.
Definition 3.9 (Latin square) An $r \times r$ Latin square is an $r \times r$ matrix $L$ such that $L_{ij} \in [r]$ for all $i, j$ and every number in $[r]$ appears exactly once in each row and exactly once in each column.

If $L$ is a Latin square and $L_{ii} = i$ for all $i \in [r]$, we call it a diagonal Latin square.

Lemma 3.10 For every $r \geq 3$, there exists an $r \times r$ diagonal Latin square. For $r \geq 4$, there exist diagonal Latin squares with the additional property that, for every $i \neq j$, $L_{ij} \neq L_{ji}$.

Proof For $r \geq 3$, the existence of $r \times r$ diagonal Latin squares was proved by Hilton [13]. Therefore, we need only show the second part of the theorem. For this we rely on self-orthogonal Latin squares.

Two Latin squares $L$ and $L'$ are called orthogonal if every ordered pair $(k, l) \in [r]^2$ occurs uniquely as $(L_{ij}, L'_{ij})$ for some $i, j \in [r]$. A Latin square is called self-orthogonal if it is orthogonal to its transpose, denoted by $L^T$. A theorem of Brayton, Coppersmith, and Hoffman [5] proves the existence of $r \times r$ self-orthogonal Latin squares for $r \in \mathbb{N}$, $r \neq 2, 3, 6$. Let $L$ be a self-orthogonal Latin square. Since $L_{ii} = L'^T_{ii}$, the diagonal entries give all pairs of the form $(i, i)$ for every $i \in [r]$, i.e., the diagonal entries must be a permutation of $[r]$. Without loss of generality, we may assume that $L_{ii} = i$ and so $L$ is also a diagonal Latin square. Clearly a self-orthogonal Latin square satisfies the property that $L_{ij} \neq L_{ji}$ if $i \neq j$.

This leaves us only with the case $r = 6$, which requires separate treatment. It is known that $6 \times 6$ self-orthogonal Latin squares do not exist. Fortunately, the property we require is weaker and we are able to give an explicit construction of a matrix that is sufficient for our needs. Let $L$ be the following matrix:

$$
\begin{bmatrix}
1 & 4 & 5 & 3 & 6 & 2 \\
3 & 2 & 6 & 5 & 1 & 4 \\
2 & 5 & 3 & 6 & 4 & 1 \\
6 & 1 & 2 & 4 & 3 & 5 \\
4 & 6 & 1 & 2 & 5 & 3 \\
5 & 3 & 4 & 1 & 2 & 6
\end{bmatrix}
$$

It is straightforward to verify that $L$ has the required properties. □

The following lemma is a strengthening of a lemma from [2].

Lemma 3.11 Let $r \geq 3$. Then there exists a set $T \subseteq [r]^3$, referred to as a triple system, of $r^2 - r$ triples that satisfies the following properties:

1. Each triple consists of three distinct elements;
2. For every pair $i, j \in [r], i \neq j$, there are exactly six triples containing both $i$ and $j$;
3. If $r \geq 4$, for every $i, j \in [r], i \neq j$, there are at least two triples containing $i$ and $j$ such that the remaining elements are distinct.
Proof Let $L$ be a Latin square as in Lemma 3.10. Let $T$ be the set of triples $(i, j, k) \subseteq [r]^3$ with $i \neq j$ and $k = L_{ij}$. Clearly the number of such triples is $r^2 - r$. We verify that the properties mentioned hold.

Recall that we have $L_{ii} = i$ for all $i \in [r]$, and every value appears once in each row and column. So for $i \neq j \in [r]$, it cannot happen that $L_{ij} = i$ or $L_{ij} = j$ and we get Property 1, i.e., all elements of a triple must be distinct.

For Property 2, note that a pair $i, j$ appears once as $(i, j, L_{ij})$ and once as $(j, i, L_{ji})$. And since every element appears exactly once in every row and column, we have that $i$ must appear once in the $j$th row, $j$ must appear once in the $i$th row and the same for the columns. It follows that each of $(\ast, j, i), (j, \ast, i), (\ast, i, j)$ and $(i, \ast, j)$ appears exactly once, where $\ast$ is some other element of $[r]$. This gives us that every pair appears in exactly six triples.

For $r \geq 4$ and $i \neq j$, since $L_{ij} \neq L_{ji}$, the triples $(i, j, L_{ij})$ and $(j, i, L_{ji})$ are sufficient to satisfy Property 3. \qed

4 The Dependency Matrix

Let $V = \{v_1, \ldots, v_n\}$ be a set of $n$ points in $\mathbb{C}^d$. We will use $\dim(V)$ to denote the dimension of the linear span of $V$, and by affine-$\dim(V)$ the dimension of the affine span of $V$ (i.e., the minimum $r$ such that points of $V$ are contained in a shift of a linear subspace of dimension $r$). We projectivize $\mathbb{C}^d$ and consider the set of vectors $V' = \{v'_1, \ldots, v'_n\}$, where $v'_i = (v_i, 1)$ is the vector in $\mathbb{C}^{d+1}$ obtained by appending a 1 to the vector $v_i$. Let $V$ be the $n \times (d + 1)$ matrix whose $i$th row is the vector $v'_i$. Now note that

$$\text{affine-$\dim(V)$} = \dim(V') - 1 = \text{rank}(V) - 1.$$

We now construct a matrix $A$, which we refer to as the dependency matrix of $V$. The construction we give here is preliminary, but suffices to prove Theorems 1.7 and 1.12. A refined construction is given in Sect. 6, where we select the triples more carefully. The rows of the matrix will consist of linear dependency coefficients, which we define below.

**Definition 4.1 (Linear dependency coefficients)** Let $v_1, v_2$ and $v_3$ be three distinct collinear points in $\mathbb{C}^d$, and let $v'_i = (v_i, 1), i \in \{1, 2, 3\}$, be vectors in $\mathbb{C}^{d+1}$. Recall that $v_1, v_2, v_3$ are collinear if and only if there exist non-zero coefficients $a_1, a_2, a_3 \in \mathbb{C}$ such that

$$a_1 v'_1 + a_2 v'_2 + a_3 v'_3 = 0.$$ 

We refer to the $a_1, a_2$ and $a_3$ as the linear dependency coefficients between $v_1, v_2, v_3$. Since these coefficients are determined up to scaling by a complex number, we fix a canonical choice by setting $a_3 = 1$.

**Definition 4.2 (Dependency matrix)** For every line $l \in L_{\geq 3}(V)$, let $V_l$ denote the points lying on $l$. Then $|V_l| \geq 3$ and we assign each line a triple system $T_l \subseteq [|V_l|]^3$. 

\[\text{Springer}\]
the existence of which is guaranteed by Lemma 3.11. Let $A$ be the $m \times n$ matrix obtained by going over every line $l \in L_{\geq 3}$ and for each triple $(i, j, k) \in T_l$, adding as a row of $A$ a vector with three non-zero coefficients in positions $i, j, k$ corresponding to the linear dependency coefficients among the points $v_i, v_j, v_k$. We refer to $A$ as the dependency matrix for $V$.

Note that we have $AV = 0$. Every row of $A$ has exactly three non-zero entries. By Property 2 of Lemma 3.11, the supports of any distinct two columns intersect in exactly six entries when the two corresponding points lie on a special line, and are disjoint otherwise. That is, the supports of any two distinct columns intersect in at most six entries.

We say a pair of points $v_i, v_j, i \neq j$, appears in the dependency matrix $A$ if there exists a row with non-zero entries in columns $i$ and $j$. The number of times a pair appears is the number of rows with non-zero entries in both columns $i$ and $j$.

Every pair of points that lies on a special line appears exactly six times. The only pairs not appearing in the matrix are pairs of points that determine ordinary lines. There are $\binom{n}{2}$ pairs of points, $t_2(V)$ of which determine ordinary lines. So the number of pairs appearing in $A$ is $\binom{n}{2} - t_2$. The total number of times these pairs appear is then $6 \left( \binom{n}{2} - t_2 \right)$. Every row gives three distinct pairs of points, so it follows that the number of rows of $A$ is

$$m = 6 \left( \binom{n}{2} - t_2 \right) / 3 = n^2 - n - 2t_2(V). \quad (4)$$

Note that $m > 0$, unless $t_2 = \binom{n}{2}$, i.e., all lines are ordinary.

As mentioned in the proof overview, we will consider two cases: when $A$ satisfies Property-$S$ and when it does not. We now prove lemmas dealing with the two cases. The following lemma deals with the former case.

**Lemma 4.3** Let $V$ be a set of $n$ points affinely spanning $\mathbb{C}^d$, $d \geq 3$, and let $A$ be the dependency matrix for $V$. Suppose that $A$ satisfies Property-$S$. Then

$$t_2(V) \geq \frac{(d-3)}{2(d+1)} n^2 + \frac{3}{2} n.$$  

**Proof** Fix $\varepsilon > 0$. Since $A$ satisfies Property-$S$, by Corollary 3.4 there is a scaling $A'$ such that the $\ell_2$ norm of each row is at most $\sqrt{1 + \varepsilon}$ and the $\ell_2$ norm of each column is at least $\sqrt{m/n - \varepsilon}$. Let $M := A'^* A'$. Then $M_{ii} \geq m/n - \varepsilon$ for all $i$. Since every row in $A$ has support of size three, and the supports of any two columns intersect in at most six locations, Corollary 3.8 gives us that $\sum_{i \neq j} |M_{ij}|^2 \leq 4m(1 + \varepsilon)^2$. Applying Lemma 3.5 to $M$, we have

$$\text{rank}(M) \geq \frac{n^2(m/n - \varepsilon)^2}{n(m/n - \varepsilon)^2 + 4m(1 + \varepsilon)^2}.$$  

$\blacksquare$ Springer
Taking $\varepsilon$ to 0, and combining with (4) gives

$$\text{rank}(A) = \text{rank}(A') = \text{rank}(M) \geq \frac{n^2 m^2 / n^2}{n m^2 / n^2 + 4m} = \frac{mn}{m + 4n} = n - \frac{4n^2}{m + 4n} = n - \frac{4n^2}{n^2 - n - 2t_2(V) + 4n} = n - \frac{4n^2}{n^2 + 3n - 2t_2(V)}.$$

Recall that $\text{affine-dim}(V) = d = \text{rank}(V) - 1$. Since $AV = 0$, we have $\text{rank}(V) \leq n - \text{rank}(A)$. It follows that

$$d + 1 \leq \frac{4n^2}{n^2 + 3n - 2t_2(V)},$$

i.e., $t_2(V) \geq \frac{(d - 3)}{2(d + 1)} n^2 + \frac{3}{2} n$. □

We now consider the case when Property-S is not satisfied.

**Lemma 4.4** Let $V$ be a set of $n$ points in $\mathbb{C}^d$, and let $A$ be the dependency matrix for $V$. Suppose that $A$ does not satisfy Property-S. Then, for every integer $b^*$, $1 < b^* < 2n/3$, at least one of the following holds:

1. There exists a point $v \in V$ contained in at least $2(n + 1)/3 - b^*$ ordinary lines;
2. $t_2(V) \geq nb^*/2$.

**Proof** Since $A$ violates Property-S, there exists a zero submatrix supported on rows $U \subseteq [m]$ and columns $W \subseteq [n]$ of the matrix $A$, where $|U| = a$ and $|W| = b$, such that

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

Let $X = [m]\setminus U$ and $Y = [n]\setminus W$ and note that $|X| = m - a$ and $|Y| = n - b$. Let the violating columns correspond to the set $V_1 = \{v_1, \ldots, v_b\} \subset V$. We consider two cases: when $b < b^*$, and when $b \geq b^*$.

**Case 1** ($b < b^*$). We may assume that $U$ is maximal, so every row in the submatrix $X \times W$ has at least one non-zero entry. Partition the rows of $X$ into three parts: Let $X_1$, $X_2$ and $X_3$ be rows with one, two and three non-zero entries in columns of $W$ respectively. We will get a lower bound on the number of ordinary lines containing exactly one point in $V_1$ and one point in $V \setminus V_1$ by bounding the number of pairs $(v_i, w)$, with $v_i \in V_1$ and $w \in V \setminus V_1$, that lie on special lines. Note that there are at most $b(n - b)$ such pairs, and each pair that does not lie on a special line determines an ordinary line.

Each row of $X_1$ gives two pairs of points $(v_i, w_1)$ and $(v_i, w_2)$ that lie on a special line, where $v_i \in V_1$ and $w_1, w_2 \in V \setminus V_1$. Each row of $X_2$ gives two pairs of points $(v_i, w)$ and $(v_j, w)$, where $v_i, v_j \in V_1$ and $w \in V \setminus V_1$ that lie on special lines. Each
row of $X_3$ has all zero entries in the submatrix supported on $X \times Y$, so does not contribute any pairs. Recall that each pair of points on a special line appears exactly six times in the matrix. This implies that the number of pairs that lie on special lines with at least one point in $V_1$ and one point in $V \setminus V_1$ is $(2|X_1| + 2|X_2|)/6 \leq |X|/6$. Hence, the number of ordinary lines containing exactly one of $v_1, \ldots, v_b$ is at least $b(n - b) - |X|/3$.

Recall that

$$1 < \frac{a}{m} + \frac{b}{n} = \left(1 - \frac{|X|}{m}\right) + \frac{b}{n}. $$

Substituting $m \leq n^2 - n$ (a consequence of (4)) gives

$$|X| < \frac{bn}{n} \leq b(n - 1).$$

This gives that the number of ordinary lines containing exactly one point in $V_1$ is at least

$$b(n - b) - \frac{|X|}{3} > \frac{2b}{3}n - \frac{3b^2 - b}{3}. $$

We now have that there exists $v \in V_1$ such that the number of ordinary lines containing $v$ is at least

$$\left\lfloor \frac{2}{3}n - \frac{3b - 1}{3} \right\rfloor \geq \left\lfloor \frac{2}{3}n - b^* + \frac{4}{3} \right\rfloor \geq \frac{2}{3}(n + 1) - b^*. $$

**Case 2** ($b \geq b^*$). We will determine a lower bound for $t_2(V)$ by counting the number of non-zero pairs of entries $A_{ij}, A_{ij'}$ with $j \neq j'$, that appear in the submatrix $U \times Y$. There are $\binom{n - b}{2}$ pairs of points in $V \setminus V_1$, each of which appears at most six times, therefore the number of pairs of such entries is at most $6\binom{n - b}{2}$. Each row of $U$ has three pairs of non-zero entries, i.e., the number of pairs of entries equals $3a$. It follows that

$$3a \leq 6\binom{n - b}{2}. $$

Combining (4) with the fact that $a/m + b/n > 1$ implies

$$a > m\left(1 - \frac{b}{n}\right) = \left(n^2 - n - 2t_2(V)\right)\left(1 - \frac{b}{n}\right). $$

Equations (5) and (6) together give

$$\left(n^2 - n - 2t_2(V)\right)\left(1 - \frac{b}{n}\right) < 2\binom{n - b}{2}. $$
Finally, solving for $t_2(\mathcal{V})$, we have

$$t_2(\mathcal{V}) > \frac{nb}{2} \geq \frac{nb^*}{2}.$$ 

\[\Box\]

5 Proofs of Theorems 1.7 and 1.12

The proofs of both Theorems 1.7 and 1.12 rely on Lemmas 4.3 and 4.4. Together, these lemmas imply that there must be a point with many ordinary lines containing it, or there are many ordinary lines in total. As mentioned in the proof overview, the theorems are then obtained by using an iterative argument removing a point contained in many ordinary lines, and then applying the same argument to the remaining points.

5.1 Proof of Theorem 1.7

We get the following easy corollary from Lemmas 4.3 and 4.4.

Corollary 5.1 Let $\mathcal{V}$ be a set of $n \geq 5$ points in $\mathbb{C}^d$ not contained in a plane. Then at least one of the following holds:

1. There exists a point $v \in \mathcal{V}$ contained in at least $\frac{2n}{3} - \frac{7}{3}$ ordinary lines;
2. $t_2(\mathcal{V}) \geq \frac{3n}{2}$.

Proof Let $A$ be the dependency matrix for $\mathcal{V}$. If $A$ satisfies Property-S, then we are done by Lemma 4.3. Otherwise, let $b^* = 3$, and note that Lemma 4.4 gives us the statement of the corollary when $n \geq 5$. \[\Box\]

We are now ready to prove Theorem 1.7. For convenience, we state the theorem again.

Theorem 1.7 Let $\mathcal{V}$ be a set of $n \geq 24$ points in $\mathbb{C}^3$ not contained in a plane. Then $\mathcal{V}$ determines at least $3n/2$ ordinary lines, unless $n - 1$ points are on a plane in which case there are at least $n - 1$ ordinary lines.

Proof If $t_2(\mathcal{V}) \geq 3n/2$ then we are done. Else, by Corollary 5.1, there exists a point $v_1$ incident to at least $(2n - 7)/3$ ordinary lines and hence, at most $(n + 4)/6$ special lines. Let $\mathcal{V}_1 = \mathcal{V}\setminus\{v_1\}$. If $\mathcal{V}_1$ is planar, then there are exactly $n - 1$ ordinary lines incident to $v_1$. This is the only case where there exists fewer than $3n/2$ ordinary lines.

Suppose now that $\mathcal{V}_1$ is not planar. Again, by Corollary 5.1, there are either $3(n - 1)/2$ ordinary lines determined by $\mathcal{V}_1$, or there exists a point $v_2 \in \mathcal{V}_1$ incident to at least $2(n - 1)/3 - 7/3 = (2n - 9)/3$ ordinary lines (determined by $\mathcal{V}_1$). In the former case, there are $3(n - 1)/2$ ordinary lines determined by $\mathcal{V}_1$, at most $(n + 4)/6$ of which could contain $v_1$. Then, the total number of ordinary lines in $\mathcal{V}$ is

$$t_2(\mathcal{V}) \geq \frac{3}{2} (n - 1) - \frac{1}{6} (n + 4) + \frac{1}{3} (2n - 7) = \frac{1}{2} (4n - 9).$$

When $n \geq 9$, this implies $t_2(\mathcal{V}) \geq 3n/2$. \[\Box\]
In the latter case there exists a point \( v_2 \in V_1 \) with at least \( (2n - 9)/3 \) ordinary lines determined by \( V_1 \) incident to it. At most one of these could contain \( v_1 \), so there are at least \( (2n - 7)/3 + (2n - 9)/3 - 1 = (4n - 19)/3 \) ordinary lines incident to one of \( v_1 \) or \( v_2 \). Note also that the number of special lines containing one of \( v_1 \) or \( v_2 \) is at most \( (n + 4)/6 + (n + 3)/6 = (2n + 7)/6 \).

Let \( V_2 = V_1 \setminus \{ v_2 \} \). If \( V_2 \) is contained in a plane, there are at least \( n - 3 \) ordinary lines incident to each of \( v_1 \) and \( v_2 \) giving a total of \( 2n - 6 \) ordinary lines determined by \( V \). It follows that when \( n \geq 12 \), \( t_2(V) \geq 3n/2 \).

Otherwise, \( V_2 \) is not contained in a plane, and again Corollary 5.1 gives two cases. If there are \( 3(n - 2)/2 \) ordinary lines determined by \( V_2 \), then the total number of ordinary lines is

\[
t_2(V) \geq \frac{1}{2} (2n - 2) - \frac{1}{6} (2n + 7) + \frac{1}{3} (4n - 19) = \frac{1}{2} (5n - 21).
\]

When \( n \geq 11 \), this implies \( t_2(V) \geq 3n/2 \).

Otherwise, there exists a point \( v_3 \) contained in at least \( 2(n - 2)/3 - 7/3 \) ordinary lines. At most two of these could contain \( v_1 \) or \( v_2 \), so there are \( 2(n - 2)/3 - 7/3 - 2 = (2n - 17)/3 \) ordinary lines incident to \( v_3 \) determined by \( V \). Summing up the number of lines containing one of \( v_1, v_2 \) and \( v_3 \), we have

\[
t_2(V) \geq \frac{1}{3} (2n - 17) + \frac{1}{3} (4n - 19) = 2n - 12.
\]

When \( n \geq 24 \), this implies \( t_2(V) \geq 3n/2 \).

\[\square\]

5.2 Proof of Theorem 1.12

We get the following easy corollary from Lemmas 4.3 and 4.4.

**Corollary 5.2** There exists a positive integer \( n_0 \) such that the following holds. Let \( V \) be a set of \( n \geq n_0 \) points in \( \mathbb{C}^d \) not contained in a three-dimensional affine subspace. Then at least one of the following holds:

1. There exists a point contained in at least \( n/2 \) ordinary lines;
2. \( t_2(V) \geq n^2/12 \).

**Proof** Let \( A \) be the dependency matrix of \( V \). If \( A \) satisfies Property-S, then we are done by Lemma 4.3. Otherwise, let \( b^* = n/6 \). Now by Lemma 4.4, either the number of ordinary lines

\[
t_2(V) \geq \frac{n}{2} b^* \geq \frac{1}{12} n^2,
\]

or, there exists a point \( v \in V \), such that the number of ordinary lines containing \( v \) is at least

\[
\frac{2}{3} (n + 1) - b^* > \frac{1}{2} n.
\]

\[\square\]
We are now ready to prove Theorem 1.12. For convenience, we state the theorem again.

**Theorem 1.12** There exists an absolute constant $c' > 0$ and a positive integer $n_0$ such that the following holds. Let $V$ be a set of $n \geq n_0$ points in $C^4$ with at most $n/2$ points contained in any three-dimensional affine subspace. Then

$$t_2(V) \geq c'n^2.$$ 

**Proof** The basic idea of the proof uses the following algorithm: We use Corollary 5.2 to find a point incident to a large number of ordinary lines, “prune” this point, and then repeat this on the smaller set of points. We stop when either we cannot find such a point, in which case Corollary 5.2 guarantees a large number of ordinary lines, or when we have accumulated enough ordinary lines.

Consider the following algorithm: Let $V_0 := V$ and $j = 0$.

1. If $V_j$ satisfies Case 2 of Corollary 5.2, then stop.
2. Otherwise, there must exist a point $v_{j+1} \in V_j$ incident to at least $(n-j)/2$ ordinary lines determined by $V_j$. Let $V_{j+1} = V_j \setminus \{v_{j+1}\}$.
3. Set $j = j + 1$. If $j \geq n/2$, then stop. Otherwise, go to Step 1.

Since no three-dimensional subspace contains more than $n/2$ points, at no point will the algorithm stop because the configuration becomes three-dimensional. That is, we can use Corollary 5.2 at every step of the algorithm.

We now analyze the two stopping conditions for the algorithm, and show that we can always find enough ordinary lines by the time the algorithm stops.

Suppose that we stop because $V_j$ satisfies Case 2 of Corollary 5.2 for some $1 \leq j < n/2$. Case 2 of Corollary 5.2 implies

$$t_2(V_j) \geq \frac{(n-j)^2}{12}.$$  \hspace{1cm} (7)

On the other hand, each pruned point $v_i$, $1 \leq i \leq j$, is incident to at least $(n - i + 1)/2 > (n - i)/2$ ordinary lines determined by $V_{i-1}$, and hence, at most $(n - i - (n - i + 1)/2)/2 < (n - i)/4$ special lines. An ordinary line in $V_i$ might not be ordinary in $V_{i-1}$ if it contains $v_i$. Thus, in order to lower bound the total number of ordinary lines in $V$, we sum over the number of ordinary lines contributed by each of the pruned points $v_i$, $1 \leq i \leq j$, and subtract from the count the number of potential lines that could contain $v_i$. Then the number of ordinary lines in $V$ contributed by the pruned points is at least

$$\sum_{i=1}^{j} \left( \frac{n - i}{2} - \frac{n - i}{4} \right) = \frac{1}{4} \sum_{i=1}^{j} (n - i).$$  \hspace{1cm} (8)
Combining (7) and (8) gives

\[ t_2(\mathcal{V}) \geq \frac{1}{12} (n - j)^2 + \frac{1}{4} \sum_{i=1}^{j} (n - i) = \frac{1}{24} \left( -j^2 + j(2n - 3) + 2n^2 \right). \]

This is an increasing function for \( j < n - 3/2 \), i.e., \( j \leq n - 2 \) implying that

\[ t_2(\mathcal{V}) \geq \frac{n^2}{12}. \]

We now consider the case when the algorithm stops because \( j \geq n/2 \). Note that at this point, we will have pruned exactly \( j \) points. Each pruned point \( v_i, 1 \leq i \leq j \), is incident to at least \( (n - i + 1)/2 > (n - i)/2 \) ordinary lines determined by \( \mathcal{V}_{i-1} \). The only way such an ordinary line is not ordinary in \( \mathcal{V} \) is if it contains one of the previously pruned points. At most \( i - 1 < i \) of the ordinary lines incident to \( v_i \) contain other pruned points \( v_k, k < i \). Therefore the total number of ordinary lines determined by \( \mathcal{V} \) satisfies

\[ t_2(\mathcal{V}) \geq \frac{1}{2} \sum_{i=1}^{j} \frac{n - i}{2} - \frac{1}{2} \sum_{i=1}^{j} i = \frac{1}{2} \sum_{i=1}^{j} (n - 3i) \geq \frac{n^2 - 8n - 9}{16}. \]

This gives us that for some absolute constant \( c' > 0 \) and \( n \) large enough,

\[ t_2(\mathcal{V}) \geq c'n^2. \]

\[ \square \]

6 A Dependency Matrix for a More Refined Bound

In this section we give a refined construction for the dependency matrix of a point set \( \mathcal{V} \). Recall that we defined the dependency matrix in Definition 4.2 to contain a row for each collinear triple from a triple system constructed on each special line. The goal was to not have too many triples containing the same pair (as can happen when there are many points on a single line). At the end of this section (Definition 6.7) we will give a construction of a dependency matrix that will have an additional property (captured in Item 4 of Lemma 6.4) which is used to obtain cancellation in the diagonal dominant argument, as outlined in the proof overview.

We denote the argument of a non-zero complex number \( z \) by \( \arg z \), and use the convention that \( \arg z \in (-\pi, \pi] \).

**Definition 6.1** (Angle between two complex numbers) We define the angle between two non-zero complex numbers \( a \) and \( b \) to be the absolute value of the argument of \( a \overline{b} \), denoted by \( |\arg ab| \). Note that the angle between \( a \) and \( b \) equals the angle between \( b \) and \( a \).

**Definition 6.2** (Co-factor) Let \( v_1, v_2 \) and \( v_3 \) be three distinct collinear points in \( \mathbb{C}^d \), and let \( a_1, a_2 \) and \( a_3 \) be the linear dependency coefficients among the three points.
Define the co-factor of $v_3$ with respect to $(v_1, v_2)$, denoted by $C_{1,2}(3)$, to be $\frac{a_1 \overline{a_2}}{|a_1||a_2|}$. Notice that this is well defined with respect to the points, and does not depend on the choice of coefficients.

The next lemma will be used to show that “cancellations” must arise in a line containing four points (as mentioned earlier in the proof overview). We will later use this lemma as a black box to quantify the cancellations in lines with more than four points by applying it to random four tuples inside the line.

**Lemma 6.3** Let $v_1, v_2, v_3, v_4$ be four collinear points in $\mathbb{C}^d$. Then at least one of the following holds:

1. The angle between $C_{1,2}(3)$ and $C_{1,2}(4)$ is at least $\pi/3$;
2. The angle between $C_{1,3}(4)$ and $C_{1,3}(2)$ is at least $\pi/3$;
3. The angle between $C_{1,4}(2)$ and $C_{1,4}(3)$ is at least $\pi/3$.

**Proof** For $i \in \{1, 2, 3, 4\}$, let $v_i' = (v_i, 1)$, i.e., the vector obtained by appending 1 to $v_i$. Since $v_1, v_2, v_3, v_4$ are collinear, there exist $a_1, a_2, a_3 \in \mathbb{C}$ such that

$$a_1 v_1' + a_2 v_2' + a_3 v_3' = 0,$$

and $b_1, b_2, b_4 \in \mathbb{C}$ such that

$$b_1 v_1' + b_2 v_2' + b_4 v_4' = 0. \tag{10}$$

We may assume, without loss of generality, that $a_3 = b_4 = 1$. Now (9) and (10) give us that $C_{1,2}(3) = \frac{a_1 \overline{a_2}}{|a_1||a_2|}$, $C_{1,2}(4) = \frac{b_1 \overline{b_2}}{|b_1||b_2|}$, $C_{1,3}(2) = \frac{a_1}{|a_1|}$ and $C_{1,4}(2) = \frac{b_1}{|b_1|}$.

Combining (9) and (10), we get the following linear equation:

$$(b_2 a_1 - b_1 a_2) v_1' + b_2 v_3' - a_2 v_4' = 0. \tag{11}$$

From (11), we get $C_{1,3}(4) = \frac{(b_2 a_1 - b_1 a_2) \overline{b_2}}{|b_2 a_1 - b_1 a_2||b_2|}$ and $C_{1,4}(3) = -\frac{(b_2 a_1 - b_1 a_2) \overline{a_2}}{|b_2 a_1 - b_1 a_2||a_2|}$.

Then the angle between $C_{1,2}(3)$ and $C_{1,2}(4)$ is

$$\left|\arg \frac{a_1 \overline{a_2}}{|a_1||a_2|} \frac{b_1 b_2}{|b_1||b_2|}\right| = |\arg a_1 \overline{a_2} b_1 b_2|. \tag{12}$$

The angle between $C_{1,3}(4)$ and $C_{1,3}(2)$ is

$$\left|\arg \frac{(b_2 a_1 - b_1 a_2) \overline{a_1}}{|b_2 a_1 - b_1 a_2||a_1|}\right| = |\arg a_1 \overline{b_2} (b_2 a_1 - b_1 a_2)|. \tag{13}$$

The angle between $C_{1,4}(2)$ and $C_{1,4}(3)$ is

$$\left|\arg \frac{b_1 (b_2 a_1 - b_1 a_2) a_2}{|b_1||b_2 a_1 - b_1 a_2||a_2|}\right| = |\arg -b_1 a_2 (b_2 a_1 - b_1 a_2)|. \tag{14}$$
Note that the product of expressions inside the arg functions in (12), (13) and (14) is a negative real number, and so the sum of (12), (13) and (14) must be π. It follows that one of the angles must be at least π/3. □

Our final dependency matrix will be composed of blocks, each given by the following lemma. Roughly speaking, we construct a block of rows $A(l)$ for each special line $l$. The rows in $A(l)$ will be chosen carefully and will correspond to triples that will eventually give nontrivial cancellations.

**Lemma 6.4** Let $l$ be a line in $\mathbb{C}^d$ and $V_l = \{v_1, \ldots , v_r\}$ be points on $l$ with $r \geq 3$. Let $V_l$ be the $r \times (d + 1)$ matrix whose $i^{th}$ row is the vector $(v_i, 1)$. Then there exists an $(r^2 - r) \times r$ matrix $A = A(l)$, which we refer to as the dependency matrix of $l$, such that the following hold:

1. $AV_l = 0$;
2. Every row of $A$ has support of size three;
3. The support of every two columns of $A$ intersects in exactly six locations;
4. If $r \geq 4$ then for at least $1/3$ of choices of $k \in [r^2 - r]$, there exists $k' \in [r^2 - r]$ such that following holds: For $k \in [r^2 - r]$, let $R_k$ denote the $k^{th}$ row of $A$. Suppose $\text{supp}(R_k) = \{i, j, s\}$. Then $\text{supp}(R_k') = \{i, j, t\}$ (for some $t \neq s$) and the angle between the co-factors $C_{i, j}(s)$ and $C_{i, j}(t)$ is at least $\pi/3$.

**Proof** Recall that Lemma 3.11 gives us a family of triples $T_r$ on the set $[r]^3$. For every bijective map $\sigma : V_l \rightarrow [r]$, construct a matrix $A_\sigma$ in the following manner: Let $T_l$ be the triple system on $V_l^3$ induced by $\sigma$ and $T_r$. For each triple $(v_i, v_j, v_k) \in T_l$, add a row with three non-zero entries in positions $i, j, k$ corresponding to the linear dependency coefficients between $v_i, v_j$ and $v_k$.

Note that for every $\sigma$, $A_\sigma$ has $r^2 - r$ rows and $r$ columns. Since the rows correspond to linear dependency coefficients, clearly we have $A_\sigma V_l = 0$ satisfying Property 1. Properties 2 and 3 follow from properties of the triple system from Lemma 3.11.

We will use a probabilistic argument to show that there exists a matrix $A$ that has Property 4. Let $\Sigma$ be the collection of all bijective maps from $[r]$ to the points $V_l$, and let $\sigma \in \Sigma$ be a uniformly random element. Consider $A_\sigma$. Since every pair of points occurs in at least two distinct triples, for every row $R_k$ of $A_\sigma$, there exists a row $R_k'$ such that the supports of $R_k$ and $R_k'$ intersect in two entries. Suppose that $R_k$ and $R_k'$ have supports contained in $\{i, j, s, t\}$. Suppose that $\sigma$ maps $\{v_j, v_j, v_k, v_i\}$ to $\{1, 2, 3, 4\}$ and that $(1, 2, 3)$ and $(1, 2, 4)$ are triples in $T_r$. Without loss of generality, assume $v_i$ maps to 1. Then by Lemma 6.3, the angle between at least one of the pairs $\{C_{i, j}(s), C_{i, j}(t)\}, \{C_{i, s}(j), C_{i, s}(t)\}, \{C_{i, t}(j), C_{i, t}(s)\}$ must be at least $\pi/3$. That is, given that $v_i$ maps to 1, we have that the probability that $R_k$ satisfies Property 4 is at least $1/3$. Then it is easy to see that

$$\text{Pr}(R_k \text{satisfies Property 4}) \geq \frac{1}{3}.$$
\[ \mathbb{E}[X] \geq (r^2 - r) \frac{1}{3}. \]

It follows that there exists a matrix \( A \) in which at least \( \frac{1}{3} \) of the rows satisfy Property 4.

To argue about the off-diagonal entries of \( M \), we will use the following notion of balanced rows. The main idea here is that, if there are many rows that are not balanced then we win in one of the Cauchy–Schwartz applications and, if many rows are balanced then we win from cancellations that show up via the different angles.

**Definition 6.5** (\( \eta \)-Balanced row) Given an \( m \times n \) matrix \( A \), we say a row \( R_k \) is \( \eta \)-balanced for some constant \( \eta \) if \( \vert |A_{ki}|^2 - |A_{kj}|^2 \vert \leq \eta \), for every \( i, j \in \text{supp}(R_k) \). Otherwise, we say that \( R_k \) is \( \eta \)-unbalanced. When \( \eta \) is clear from the context, we say that the row is balanced/unbalanced.

**Lemma 6.6** There exists an absolute constant \( c_0 > 0 \) such that the following holds. Let \( l \) be a line in \( \mathbb{C}^d \) and \( V_1 = \{v_1, \ldots v_r\} \) be points on \( l \) with \( r \geq 4 \). Let \( A = A(l) \) be the dependency matrix for \( l \), defined in Lemma 6.4, and \( A' \) a scaling of \( A \) such that the \( \ell_2 \) norm of every row is \( \alpha \). Let \( M = A'^*A' \).

\[
\sum_{i \neq j} |M_{ij}|^2 \leq 4(r^2 - r)\alpha^4 - c_0(r^2 - r)\alpha^2.
\]

**Proof** Recall that \( A \) is an \( (r^2 - r) \times r \) matrix, that the support of every row has size exactly three, and that the supports of any two distinct columns of \( A \) intersect in six locations. Clearly, any scaling \( A' \) of \( A \) will also satisfy these properties. Applying Lemma 3.7 to \( A' \) gives

\[
\sum_{i \neq j} |M_{ij}|^2 = 4(r^2 - r)\alpha^4 - (D(A') + 2E(A')).
\]

We are able to give a lower bound on \( D(A') + 2E(A') \) using Property 4 of Lemma 6.4. From here on, we focus on the rows mentioned in Property 4. Recall that there are at least \((r^2 - r)/3\) such rows. For some \( \eta \) to be determined later, suppose that \( \beta \) fraction of these rows is \( \eta \)-unbalanced. We will show each such row contributes to either \( D(A') \) or \( E(A') \).

If a row \( R_k \) is \( \eta \)-imbalanced, we have

\[
\sum_{i < j} \left( |A'_{ki}|^2 - |A'_{kj}|^2 \right)^2 > \eta^2.
\]

Alternatively suppose that \( R_k \) is \( \eta \)-balanced. Recall that \( \sum_{i=1}^n |A'_{ki}|^2 = \alpha \), and note that we must have that \( |A'_{ki}|^2 \in [\alpha/3 - 2\eta/3, \alpha/3 + 2\eta/3] \) for all \( i \in \text{supp}(R_k) \). Suppose that both \( R_k \) and \( R_k' \) have non-zero entries in columns \( i \) and \( j \), but \( R_k \) has a third non-zero entry in column \( s \) and \( R_k' \) has a third non-zero entry in column \( t \), where
s \neq t$. Suppose further that the angle $\theta$ between the co-factors $C_{i,j}(s)$ and $C_{i,j}(t)$ is at least $\pi/3$, i.e., $\cos \theta \leq 1/2$. Then

$$
|A'_{ki} \overline{A'_{kj}} - A'_{k'i} \overline{A'_{k'j}}|^2 \\
= |A'_{ki} \overline{A'_{kj}}|^2 + |A'_{k'i} \overline{A'_{k'j}}|^2 - 2 |A'_{ki} \overline{A'_{kj}}| |A'_{k'i} \overline{A'_{k'j}}| \cos \theta \\
\geq |A'_{ki} \overline{A'_{kj}}|^2 + |A'_{k'i} \overline{A'_{k'j}}|^2 - |A'_{ki} \overline{A'_{kj}}| |A'_{k'i} \overline{A'_{k'j}}|.$$

Any positive real numbers $a, b$ satisfy

$$
a^2 + b^2 - ab = \left(\frac{a}{2} - b\right)^2 + \frac{3}{4} a^2 \geq \frac{3}{4} a^2.
$$

Substituting $a = |A'_{ki} \overline{A'_{kj}}|$ and $b = |A'_{k'i} \overline{A'_{k'j}}|$ gives

$$
\geq \frac{3}{4} |A'_{ki} \overline{A'_{kj}}|^2 \geq \frac{3}{4} \left(\frac{\alpha}{3} - \frac{2\eta}{3}\right)^2 = \frac{1}{12} (\alpha - 2\eta)^2.
$$

Summing over the $\eta$-unbalanced rows, we have

$$
E(A') \geq \beta \left(\frac{r^2 - r}{3}\right) \eta^2.
$$

Summing over all the $\eta$-balanced rows gives

$$
D(A') = \sum_{i \neq j} \sum_{k < k'} |A'_{ki} \overline{A'_{kj}} |^2 \\
= \frac{1}{2} \sum_{k < k'} \sum_{i} |A'_{ki} \overline{A'_{kj}} |^2 \\
\geq \frac{1}{2} \cdot (1 - \beta) \left(\frac{r^2 - r}{3}\right) \cdot \frac{1}{12} (\alpha - 2\eta)^2. \\
= (1 - \beta) \left(\frac{r^2 - r}{72}\right) (\alpha - 2\eta)^2.
$$

Combining the lower bounds for $D(A)$ and $E(A)$, and setting $\eta = \alpha/10$ gives

$$
D(A') + 2E(A') \geq (1 - \beta) \left(\frac{r^2 - r}{72}\right) (\alpha - 2\eta)^2 + 2\beta \left(\frac{r^2 - r}{3}\right) \eta^2 \\
= (r^2 - r) \left(1 - \beta\right) \frac{1}{72} \left(\frac{4}{5} \alpha\right)^2 + \beta \frac{2}{3} \left(\frac{1}{10} \alpha\right)^2 \\
\geq c_0 (r^2 - r) \alpha^2
$$
for some absolute constant $c_0$. Combining the above with (15), we have

$$\sum_{i \neq j} |M_{ij}|^2 \leq 4(r^2 - r)\alpha^4 - c_0(r^2 - r)\alpha^2. \quad \square$$

We are now ready to define the full dependency matrix that we will use in the proof of Theorem 1.11.

**Definition 6.7 (Dependency matrix, second construction)** Let $V = \{v_1, \ldots, v_n\}$ be a set of $n$ points in $\mathbb{C}^d$ and let $V$ be the $n \times (d + 1)$ matrix whose $i$th row is the vector $(v_i, 1)$. For each matrix $A(l)$, where $l \in \mathcal{L}_{\geq 3}(V)$, add $n - r$ column vectors of all zeroes, with length $r^2 - r$, in the column locations corresponding to points not in $l$, giving an $(r^2 - r) \times n$ matrix. Let $A$ be the matrix obtained by taking the union of rows of these matrices for every $l \in \mathcal{L}_{\geq 3}(V)$. We refer to $A$ as the *dependency matrix* of $V$.

Note that this construction is a special case of the one given in Definition 4.2 and so satisfies all the properties mentioned there. In particular, we have $AV = 0$ and the number of rows in $A$ equals $n^2 - n - 2t_2(V)$.

### 7 Proof of Theorem 1.11

Before we prove the theorem, we give some key lemmas. As before, we consider two cases: When the dependency matrix $A$ satisfies Property-S and when it does not. In the latter case, we rely on Lemma 4.4. The following lemma deals with the former case.

**Lemma 7.1** There exists an absolute constant $c_1 > 0$ such that the following holds. Let $V = \{v_1, v_2, \ldots, v_n\}$ be a set of points in $\mathbb{C}^d$ not contained in a plane. Let $A$ be the $m \times n$ dependency matrix for $V$, and suppose that $A$ satisfies Property-S. Then

$$t_2(V) \geq \frac{3}{2} n + c_1 \sum_{r \geq 4} (r^2 - r)t_r(V).$$

**Proof** Since $A$ satisfies Property-S, by Corollary 3.4 for every $\varepsilon > 0$, there exists a scaling $A'$ of $A$ such that for every $i \in [m]$

$$\sum_{j \in [n]} |A'_{ij}|^2 = 1 + \varepsilon,$$

and for every $j \in [n]$

$$\sum_{i \in [m]} |A'_{ij}|^2 \geq \frac{m}{n} - \varepsilon. \quad (16)$$

Let $C_i$ be denote the $i$th column of $A'$, and let $M = A'^* A'$. From (16), we have $|M_{ii}| = \langle C_i, C_i \rangle \geq \left( \frac{m}{n} - \varepsilon \right)$. 

\[ Springer \]
To bound the sum of squares of the off-diagonal entries, we go back to the construction of the dependency matrix. Recall that the matrix $A$ was obtained by taking the union of rows of matrices $A(l)$, for each $l \in \mathcal{L}_{\geq 3}$. Then we have that $A'$ is the union of scalings of the rows of the matrices $A(l)$, for each $l \in \mathcal{L}_{\geq 3}$. Note that $|M_{ij}| = \langle C_i, C_j \rangle$ and that the intersection of the supports of any two distinct columns in contained within a scaling of $A(l)$, for some $l \in \mathcal{L}_{\geq 3}$. Therefore, to get a bound on $\sum_{i \neq j} |M_{ij}|^2$, it suffices to consider these component matrices. Combining the bounds obtained from Lemma 6.6, with $\alpha = 1 + \varepsilon$, we have

$$
\sum_{i \neq j} |M_{ij}|^2 \leq \sum_{l \in \mathcal{L}_{3}} 4(r^2 - r)\alpha^4 + \sum_{l \in \mathcal{L}_{\geq 4}} \left( 4(r^2 - r)\alpha^4 - c_0(r^2 - r)\alpha^2 \right)
$$

$$
= \sum_{l \in \mathcal{L}_{\geq 3}} 4(r^2 - r)\alpha^4 - \sum_{l \in \mathcal{L}_{\geq 4}} c_0(r^2 - r)\alpha^2
$$

$$
= 4m(1 + \varepsilon)^4 - (1 + \varepsilon)^2c_0 \sum_{r \geq 4}(r^2 - r)t_r.
$$

Let $F = c_0 \sum_{r \geq 4}(r^2 - r)t_r$. Lemma 3.5 implies

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

$$
\geq \frac{n^2 (m/n - \varepsilon)^2}{n (m/n - \varepsilon)^2 + 4m(1 + \varepsilon)^4 - (1 + \varepsilon)^2 F}.
$$

Taking $\varepsilon$ to 0, we get

$$
\text{rank}(M) \geq \frac{n^2 (m/n)^2}{n (m/n)^2 + 4m - F} = n - \frac{4n^2 m - n^2 F}{m^2 + 4mn - nF}.
$$

Since

$$
\text{affine-dim}(\mathcal{V}) = \text{rank}(\mathcal{V}) - 1 \leq \frac{4n^2 m - n^2 F}{m^2 + 4mn - nF} - 1,
$$

if

$$
\frac{4n^2 m - n^2 F}{m^2 + 4mn - nF} < 4,
$$

then $\mathcal{V}$ must be contained in a plane, contradicting the assumption of the theorem. Substituting $m = n^2 - n - 2t_2(\mathcal{V})$ and simplifying gives

$$
4t_2^2 - (2n^2 + 4n)t_2 + 3n^3 - 3n^2 + \frac{n^2 F}{4} - nF > 0.
$$
This holds when

\[ t_2(\mathcal{V}) \leq \frac{3n}{2} + \frac{F}{8} = \frac{3n}{2} + \frac{c_0}{8} \sum_{r=4}^{n} (r^2 - r) t_r(\mathcal{V}). \]

which completes the proof. \qed

We now have the following easy corollary.

**Corollary 7.2** There exists a positive integer \( n_0 \) such that the following holds. Let \( c_1 \) be the constant from Lemma 7.1 and let \( \mathcal{V} \) be a set of \( n \geq n_0 \) points in \( \mathbb{C}^d \) not contained in a plane. Then at least one of the following holds:

1. There exists a point \( v \in \mathcal{V} \) contained in at least \( n/2 \) ordinary lines;
2. \( t_2(\mathcal{V}) \geq \frac{3n}{2} + c_1 \sum_{r \geq 4} (r^2 - r) t_r(\mathcal{V}). \)

**Proof** If \( A \) satisfies Property-S, then we are done by Lemma 7.1. Otherwise, let \( b^* \) be an integer such that

\[ \frac{n}{2} (b^* - 1) < \frac{3n}{2} + c_1 \sum_{r \geq 4} (r^2 - r) t_r(\mathcal{V}) \leq \frac{n}{2} b^*. \]  

(17)

Clearly we have \( b^* > 1 \). Recall that \( \sum_{r \geq 4} (r^2 - r) t_r(\mathcal{V}) < n^2 \), implying that for \( c_1 \) small enough and \( n \) large enough,

\[ b^* < 4 + \frac{2c_1}{n} \sum_{r \geq 4} (r^2 - r) t_r(\mathcal{V}) < \frac{1}{6} n. \]  

(18)

Now by Lemma 4.4 and (17), either the number of ordinary lines

\[ t_2(\mathcal{V}) \geq \frac{n}{2} b^* \geq \frac{3n}{2} + c_1 \sum_{r \geq 4} (r^2 - r) t_r(\mathcal{V}), \]

or, using (18), there exists a point \( v \in \mathcal{V} \), such that the number of ordinary lines containing \( v \) is at least
The following lemma will be crucially used in the proof of Theorem 1.11.

**Lemma 7.3** Let $V$ be a set of $n$ points in $\mathbb{C}^d$, and $V' = V \setminus \{v\}$ for some $v \in V$. Then

\[
\sum_{r \geq 4} (r^2 - r)t_r(V') \geq \sum_{r \geq 4} (r^2 - r)t_r(V) - 4(n - 1).
\]

**Proof** Note that when we remove $v$ from the set $V$, we only affect lines that contain $v$. In particular, ordinary lines containing $v$ are removed and the number of points on every special line containing $v$ goes down by 1. Every other line remains unchanged and so it suffices to consider only lines that contain the point $v$.

We consider the difference

\[
K = \sum_{r \geq 4} (r^2 - r)t_r(V) - \sum_{r \geq 4} (r^2 - r)t_r(V').
\]

We will consider the contribution of a line $l$ determined by $V$ to the difference $K$.

Each line $l \in \mathcal{L}_{\geq 5}(V)$, i.e., a line that has $r \geq 5$ points, that contains $v$ contributes $r^2 - r$ to the summation $\sum_{r \geq 4} (r^2 - r)t_r(V)$. In $V'$, $l$ has $r - 1$ points, and contributes $(r - 1)^2 - (r - 1)$ to the summation $\sum_{r \geq 4} (r^2 - r)t_r(V')$. Therefore, $l$ contributes $2(r - 1)$ to the difference $K$. We may charge this contribution to the points on $l$ that are not $v$. There are $r - 1$ other points on $l$, so each point contributes 2 to $K$.

Each line $l \in \mathcal{L}_4(V)$ that contains $v$ contributes $r^2 - r = 12$ to the summation $\sum_{r \geq 4} (r^2 - r)t_r(V)$. These lines contain three points in $V'$, and so do not contribute anything in the $\sum_{r \geq 4} (r^2 - r)t_r(V')$ term. Once again, we charge this contribution to the points lying on $l$ that are not $v$. Each such line has three points on it other than $v$, so each point contributes $12/3 = 4$ to $K$.

There is a unique line containing $v$ and any other point, and each point either contributes 0, 2 or 4 to $K$. This gives

\[
\sum_{r \geq 4} (r^2 - r)t_r(V) - \sum_{r \geq 4} (r^2 - r)t_r(V') \leq 4(n - 1).
\]

Rearranging completes the proof. \(\Box\)

We are now ready to prove the main theorem. For convenience, we restate the theorem.

**Theorem 1.11** There exists an absolute constant $c > 0$ and a positive integer $n_0$ such that the following holds. Let $V$ be a set of $n \geq n_0$ points in $\mathbb{C}^3$ with at most $n/2$ points contained in any plane. Then

\[
t_2(V) \geq \frac{3}{2}n + c\sum_{r \geq 4} r^2 t_r(V).
\]
Proof The remainder of the proof is similar to the proof of Theorem 1.12, i.e., we use Corollary 7.2 to find a point incident to a large number of ordinary lines, “prune” this point, and then repeat this on the smaller set of points. We stop when either we cannot find such a point, in which case Corollary 7.2 guarantees a large number of ordinary lines, or when we have accumulated enough ordinary lines.

As before, consider the following algorithm: Let \( V_0 := V \) and \( j = 0 \).

1. If \( V_j \) satisfies Case 2 of Lemma 7.2, then stop.
2. Otherwise, there must exist a point \( v_{j+1} \) incident to at least \( (n - j)/2 \) ordinary lines determined by \( V_j \). Let \( V_{j+1} = V_j \setminus \{v_{j+1}\} \).
3. Set \( j = j + 1 \). If \( j \geq n/2 \), then stop. Otherwise, go to Step 1.

Note that since no plane contains more than \( n/2 \) points, at no point will the algorithm stop because the configuration becomes planar. That is, we can use Corollary 7.2 at every step of the algorithm. We now analyze the two stopping conditions for the algorithm, and show that we can always find enough ordinary lines by the time the algorithm stops.

Suppose that we stop because \( V_j \) satisfies Case 2 of Corollary 7.2 for some \( 1 \leq j < n/2 \). From Case 2 of Lemma 7.2 and Lemma 7.3, we have

\[
\begin{align*}
t_2(V_j) & \geq \frac{3(n-j)}{2} + c_1 \sum_{r \geq 4} (r^2 - r)t_r(V_j) \\
& \geq \frac{3(n-j)}{2} + c_1 \left( \sum_{r \geq 4} (r^2 - r)t_r(V) - 4 \sum_{i=1}^{j} (n-i) \right) .
\end{align*}
\]  

(19)

On the other hand, each pruned point \( v_i, 1 \leq i \leq j \), is incident to at least \( (n - i + 1)/2 > (n - i)/2 \) ordinary lines determined by \( V_{i-1} \), and hence, at most \( (n - i - (n - i + 1)/2)/2 < (n - i)/4 \) special lines. Note that an ordinary line in \( V_i \) might not be ordinary in \( V_{i-1} \) if contains \( v_i \). Thus, in order to lower bound the total number of ordinary lines in \( V \), we sum over the number of ordinary lines contributed by each of the pruned points \( v_i, 1 \leq i \leq j \), and subtract from the count the number of potential lines that could contain \( v_i \). Then the number of ordinary lines contributed by the pruned points is at least

\[
\sum_{i=1}^{j} \left( \frac{n-i}{2} - \frac{n-i}{4} \right) = \frac{1}{4} \sum_{i=1}^{j} (n-i) .
\]  

(20)

Combining (19) and (20), we have

\[
\begin{align*}
t_2(V) & \geq \frac{3}{2} (n-j) + c_1 \left( \sum_{r \geq 4} (r^2 - r)t_r(V) - 4 \sum_{i=1}^{j} (n-i) \right) + \frac{1}{4} \sum_{i=1}^{j} (n-i) \\
& = \frac{3}{2} n + c_1 \sum_{r \geq 4} (r^2 - r)t_r(V) + \left( \frac{1}{4} - 4c_1 \right) \sum_{i=1}^{j} (n-i) - \frac{3}{2} j .
\end{align*}
\]
For $c_1$ small enough and $n$ large, the term $(1/4 - 4c_1) \sum_{i=1}^{j}(n-i) - 3j/2$ is positive. Therefore, there exists some absolute constant $c > 0$ such that

$$t_2(\mathcal{V}) \geq \frac{3}{2} n + c \sum_{r \geq 4} r^2 t_r(\mathcal{V}).$$

We now consider the case when the algorithm stops because $j \geq n/2$. Note that at this point, we will have pruned exactly $j$ points. Each pruned point $v_i, 1 \leq i \leq j$, is incident to $(n-i+1)/2 > (n-i)/2$ ordinary lines determined by $\mathcal{V}_{i-1}$. However, as many as $i-1 < i$ of these lines could contain other pruned points $v_k, k < i$, i.e., lines that could be special in $\mathcal{V}$. Therefore the total number of ordinary lines determined by $\mathcal{V}$ is at least

$$t_2(\mathcal{V}) \geq \sum_{i=1}^{j} \frac{n-i}{2} - \sum_{i=1}^{j} i = \frac{1}{2} \sum_{i=1}^{j} (n-3i) \geq \frac{n^2 - 8n - 9}{16}.$$

Note that $n^2 > \sum_{r \geq 4} (r^2 - r)t_r(\mathcal{V})$, which gives

$$t_2(\mathcal{V}) \geq \frac{3}{2} n + c \sum_{r \geq 4} r^2 t_r(\mathcal{V})$$

for some absolute constant $c > 0$ and $n$ large enough.

\[\Box\]

Acknowledgements  
Zeev Dvir: Research was supported by NSF CAREER award DMS-1451191 and NSF Grant CCF-1523816. Shubhangi Saraf: Research was supported in part by NSF Grants CCF-1350572 and CCF-1540634. Charles Wolf: Research was supported in part by NSF Grant CCF-1350572.

References

1. Alon, N.: Perturbed identity matrices have high rank: proof and applications. Combin. Probab. Comput. 18(1–2), 3–15 (2009)
2. Barak, B., Dvir, Z., Wigderson, A., Yehudayoff, A.: Fractional Sylvester–Gallai theorems. Proc. Natl. Acad. Sci. USA 110(48), 19213–19219 (2013)
3. Bojanowski, R.: Zastosowania uogólnionej nierówności Bogomolova–Miyaoki–Yau. Master’s thesis, University of Warsaw (2013)
4. Borwein, P., Moser, W.O.J.: A survey of Sylvester’s problem and its generalizations. Aequationes Math. 40(1), 111–135 (1990)
5. Brayton, R.K., Coppersmith, D., Hoffman, A.J.: Self-orthogonal latin squares of all orders $n \neq 2, 3, 6$. Bull. Am. Math. Soc. 80(1), 116–118 (1974)
6. Csima, J., Sawyer, E.T.: There exist $6n/13$ ordinary points. Discrete Comput. Geom. 9(2), 187–202 (1993)
7. de Zeeuw, F.: Ordinary lines in space (2018). arXiv:1803.09524
8. de Zeeuw, F.: Spanned lines and Langer’s inequality (2018). arXiv:1802.08015
9. Dvir, Z., Saraf, S., Wigderson, A.: Improved rank bounds for design matrices and a new proof of Kelly’s theorem. Forum Math. Sigma 2, e4 (2014)
10. Elkies, N., Pretorius, L.M., Swanepoel, K.J.: Sylvester–Gallai theorems for complex numbers and quaternions. Discrete Comput. Geom. 35(3), 361–373 (2006)
11. Gallai, T.: Solution of problem 4065. Am. Math. Monthly 51, 169–171 (1944)
12. Green, B., Tao, T.: On sets defining few ordinary lines. Discrete Comput. Geom. 50(2), 409–468 (2013)
13. Hilton, A.J.W.: On double diagonal and cross latin squares. J. London Math. Soc. 2(4), 679–689 (1973)
14. Hirzebruch, F.: Arrangements of lines and algebraic surfaces. In: Artin, M., Tate, J. (eds.) Arithmetic and Geometry, Vol. II. Progress in Mathematics, vol. 36, pp. 113–140. Birkhäuser, Boston (1983)
15. Kelly, L.: A resolution of the Sylvester–Gallai problem of J.-P. Serre. Discrete Comput. Geom. 1(1), 101–104 (1986)
16. Kelly, L., Moser, W.: On the number of ordinary lines determined by $n$ points. Canad. J. Math. 10, 210–219 (1958)
17. Langer, A.: Logarithmic orbifold Euler numbers of surfaces with applications. Proc. London Math. Soc. 86(2), 358–396 (2003)
18. Melchior, E.: Über vielseite der projektiven ebene. Deutsch. Math. 5, 461–475 (1940)
19. Motzkin, T.: The lines and planes connecting the points of a finite set. Trans. Am. Math. Soc. 70, 451–464 (1951)
20. Pokora, P.: The orbifold Langer–Miyaoka–Yau inequality and Hirzebruch-type inequalities (2016). arXiv:1612.05141
21. Rothblum, U., Schneider, H.: Scalings of matrices which have prespecified row sums and column sums via optimization. Linear Algebra Appl. 114, 737–764 (1989)
22. Serre, J.-P.: Advanced problem 5359. Am. Math. Monthly 73(1), 89 (1966)
23. Sylvester, J.J.: Mathematical question 11851. Educ. Times 59(98), 256 (1893)