# SOME APPLICATIONS OF ALGEBRAIC METHODS IN COMBINATORIAL GEOMETRY

**BY ABDUL BASIT**

**A dissertation submitted to the**

**Graduate School—New Brunswick**

**Rutgers, The State University of New Jersey**

**in partial fulfillment of the requirements**

**for the degree of**

**Doctor of Philosophy**

**Graduate Program in Computer Science**

**Written under the direction of**

**William Steiger**

**and approved by**

**New Brunswick, New Jersey**

**May, 2017**

## ABSTRACT OF THE DISSERTATION

# Some Applications of Algebraic Methods in Combinatorial Geometry

**by ABDUL BASIT**

**Dissertation Director: William Steiger**

This dissertation explores problems in combinatorial geometry relating to incidences and to applications of incidence problems in other areas of combinatorics. In recent years, various tools from algebra have been applied to make significant progress on long-standing combinatorial geometric problems. These breakthroughs have stimulated work in developing additional algebraic tools and applying them to other problems. We study two different flavors of incidence problems using algebraic techniques.

- Given a set of points $\mathcal{P}$ and a set of *objects* $\mathcal{V}$, an incidence is defined to be a point-object pair $(p, v) \in \mathcal{P} \times \mathcal{V}$ such that $p \in v$, i.e., the point is contained in the object. In Chapter 2, we introduce a new notion of degeneracy and give bounds on the maximum number of point-plane and point-sphere incidences in $\mathbb{R}^3$ that are non-degenerate under this notion.

- Given a set of points $\mathcal{P}$, an *ordinary line* is defined to be a line incident to exactly two points. In 1893, Sylvester posed the following question: *"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."* In other words, Sylvester asked if every finite point set in $\mathbb{R}^2$, not all on a

line, determines an ordinary line. The question was resolved in the affirmative by Gallai in 1944 and many others subsequently. For point sets in complex space, Kelly's theorem states that point sets in $\mathbb{C}^3$, not all on a plane, must determine an ordinary line. In Chapter 3, we give bounds on the minimum number of ordinary lines determined by sets of points in $\mathbb{C}^3$.

Lastly, we give some applications of incidence bounds to other combinatorial problems. We study the $k$-most-frequent distances problem in $\mathbb{R}^3$. This generalizes the *unit distance* problem of Erdős, which asks for the maximum number of times a distance can be realized among the $\binom{n}{2}$ pairs of $n$ given points. In the $k$-most-frequent distances problem, we give a bound on the number of times a set of $k$ distances can be realized by a set of $n$ points in $\mathbb{R}^3$. Next, we consider the *sum product conjecture*, first stated by Erdős and Szemerédi in 1983. Informally, the conjecture states that a finite subset of the reals can not have both additive and multiplicative *structure* at the same time. We give new bounds for a more general version of this problem which considers subsets of complex numbers.

# Acknowledgements

First and foremost, I am indebted to my advisor William Steiger for his support and guidance throughout my PhD. His contributions to my research cannot be overestimated. I am extremely grateful to Jeff Kahn for his encouragement and guidance throughout the years. Most of the mathematics I know, I learned from him. Many thanks to Boris Aronov and Shubhangi Saraf for all the encouragement and always being available. Thanks to Nabil Mustafa, who started me on this path and showed me the joys of research.

I am very grateful to my coauthors Zeev Dvir, Ben Lund, Subhangi Saraf, Adam Sheffer, Charles Wolf for their contributions to this work and many fruitful discussions. I would also like to thank many others that I have had the chance to interact with mathematically. Unfortunately, the list is too long to include here. A special thanks to everyone I met while visiting the Institute of Pure and Applied Mathematics for providing a very stimulating environment.

My years at Rutgers were some of the happiest of my life. Thanks to everyone that helped make it so: Malihe Alikhani, Amey Bhangale, Erick Chastain, Fatima Betul Durak, Rashmi G, Edinah Gnang, Mrinal Kumar, Ben Lund, Chetan Tonde, Tomáš Vyskočil, and many others!

Many many thanks to Amirali Abdullah, Qasim Aslam, Imran Iftikhar and Umair Khalil for being there throughout the years. Their moral and emotional support was invaluable. Finally, my deep and sincere gratitude to my family: Baba, Mama, Athar, Tahir, Sarah, Durab and Hafsa. I owe them everything, and this journey would not have been possible without them.

# Dedication

For

Baba and Mama

# Table of Contents

# Chapter 1

# Introduction

Combinatorial Geometry is a field bringing together ideas from geometry and combinatorics. Loosely speaking, it deals with arrangements of geometric objects and involves studying the combinatorial properties of such arrangements. Problems in the area involve topics such as packing, covering, tiling, partitioning, decomposition, incidences, and much more. While such problems have been studied for centuries by mathematicians, the field gained much traction over the last century due to the work of Paul Erdős. Problems in the area (many posed by Erdős) have stimulated beautiful mathematics, and often, despite decades of work, have eluded resolution. When found, solutions to these problems involve deep mathematics that have influenced different branches of mathematics. This thesis presents results on an important subfield of combinatorial geometry, referred to as incidence geometry.

Consider a set $\mathcal{P}$ of points and a set $\mathcal{V}$ of "objects" (for example, one might consider lines, circles, cylinders, or tubes) in some vector space. An *incidence* is a pair $(p, v) \in \mathcal{P} \times \mathcal{V}$ such that the point $p$ is contained in the object $v$. The area of incidence geometry deals with questions about the set of incidences. For example, some celebrated results in the area, upon which this thesis builds, are the following:

**Theorem 1.1** (The Szemerédi-Trotter Theorem [72]). *Let $\mathcal{P}$ be a set of $m$ points and $\mathcal{L}$ a set of $n$ lines, both in $\mathbb{R}^2$. Then the number of incidences between points of $\mathcal{P}$ and lines of $\mathcal{L}$, i.e., pairs $(p, l) \in \mathcal{P} \times \mathcal{L}$ such that $p \in l$, is $O(n^{2/3}m^{2/3} + n + m)$. Furthermore, this bound is tight up to constant factors.*

**Theorem 1.2** (The Sylvester-Gallai Theorem [36]). *Let $\mathcal{P}$ be a finite set of points in $\mathbb{R}^2$ with the property that the line through any two points of $\mathcal{P}$ contains a third. Then all points of $\mathcal{P}$ must lie on a single line.*

In mathematics, incidence theorems have been shown to have connections to problems in arithmetic combinatorics (e.g., see [31, 20]), combinatorial geometry (e.g., see [39, 44, 63]) and harmonic analysis (e.g., see [76]). In theoretical computer science, incidence theorems are related to the structure of arithmetic circuits (e.g., see [47, 61]), locally correctable codes (e.g., see [12, 29, 3]) and expanders and extractors (e.g., see [19]).

In recent years, algebraic methods have been used to make progress on (and often resolve) many outstanding problems in combinatorial geometry. This thesis presents some results in incidence geometry based on these methods. In Chapter 2, we study Szemerédi-Trotter type problems, i.e., problems seeking bounds on the maximum number of incidences. Chapter 3 deals with Sylvester-Gallai type problems, i.e., problems considering the lines *determined* by a set of points. In Chapter 4, we give applications of incidence theorems to some problems in geometry and combinatorics.

## 1.1 Szemerédi-Trotter Type Problems

Let $\mathcal{P}$ be a set of points and $\mathcal{V}$ be a set of objects. A large set of problems in combinatorial geometry deal with counting the maximum number of incidences in $\mathcal{P} \times \mathcal{V}$, taken over all possible pairs of sets $\mathcal{P}, \mathcal{V}$ of a given size. Possibly the best known result of this type is the Szemerédi-Trotter theorem (Theorem 1.1). It deals with the case when the objects under consideration are lines in $\mathbb{R}^2$. In a recent work, Guth and Katz [39] almost completely resolved (up to a factor of $\sqrt{\log n}$) the long-standing planar *distinct distances problem* of Erdős [34]. A key idea in their proof was reducing the problem to a question about incidences between points and lines in $\mathbb{R}^3$. To bound the number of these incidences, they introduced novel tools from algebraic geometry. The introduction of this approach resulted in progress on several other combinatorial geometric problem (e.g., see [44, 45, 64, 69, 77, 78]). In Chapter 2, we present yet another application of these new algebraic techniques, by studying the maximum number of point-sphere and point-plane incidences in three-dimensional Euclidean space.

### 1.1.1 Non-degenerate sets of spheres

Given a set $\mathcal{P}$ of $m$ points and a set $\mathcal{S}$ of $n$ spheres, both in $\mathbb{R}^3$, we denote the number of point-sphere incidences in $\mathcal{P} \times \mathcal{S}$ as $I(\mathcal{P}, \mathcal{S})$. Consider a point set $\mathcal{P}$ with every point contained in a circle $c$, and a set of spheres $\mathcal{S}$ with every sphere containing $c$. In this case, every sphere of $\mathcal{S}$ is incident to every point of $\mathcal{P}$, and we have $I(\mathcal{P}, \mathcal{S}) = mn$. This led Agarwal et al. [1] to study *non-degenerate spheres* (based on the definition of $\eta$-degenerate hyperplanes of Elekes and Tóth [32]). In their terminology, a sphere $\sigma$ is said to be $\eta$-degenerate with respect to $\mathcal{P}$, for $0 < \eta < 1$, if there exists a circle $c$ contained in $\sigma$ with

$$|c \cap \mathcal{P}| \geq \eta \cdot |\sigma \cap \mathcal{P}|.$$

The above example, which led to $I(\mathcal{P}, \mathcal{S}) = mn$, cannot occur when the spheres of $\mathcal{S}$ are $\eta$-non-degenerate, for any $0 < \eta < 1$, which hints that the maximum value of $I(\mathcal{P}, \mathcal{S})$ should be smaller in this case. Indeed, when $\mathcal{S}$ is a set of $\eta$-non-degenerate spheres, Apfelbaum and Sharir [8] derived the bound

$$I(\mathcal{P}, \mathcal{S}) = O^*(n^{8/11} m^{9/11} + nm^{1/2}), \tag{1.1}$$

where the $O^*(\cdot)$-notation hides sub-polynomial factors, with a constant of proportionality depending on $\eta$.

We present a different notion of degeneracy, which does not depend on the point set $\mathcal{P}$ and is not defined on a single sphere, but rather on a set of spheres.

**Definition 1.3.** *We say that a set $\mathcal{S}$ of $n$ spheres is* $k$-non-degenerate*, for a constant $k$, $1 < k < n$, if there does not exist a circle that is contained in $k$ spheres of $\mathcal{S}$.*

Notice that the above construction with $I(\mathcal{P}, \mathcal{S}) = mn$ also does not apply to $k$-non-degenerate sets of spheres. This new definition is a natural notion of non-degeneracy (at least in the eyes of the authors), and is independent of the point set $\mathcal{P}$. It also has the advantage of being easy to analyze with the recent algebraic technique of *partitioning polynomials*. The algebraic techniques seem to work well in cases where the incidence graph contains no copy of some constant-sized complete bipartite graph $K_{s,t}$. Specifically, we prove the following theorem.

**Theorem 1.4.** *Let $\mathcal{P}$ be a set of $m$ points and let $\mathcal{S}$ be a $k$-non-degenerate set of $n$ spheres, both in $\mathbb{R}^3$, for some $1 < k < n$. Then for every $\varepsilon > 0$,*

$$I(\mathcal{P}, \mathcal{S}) = O(m^{3/4+\varepsilon} n^{3/4} k^{1/4} + n + mk),$$

*where the constant of proportionality depends on $\varepsilon$.*

By placing every point of $\mathcal{P}$ on a circle $c$ and taking $k-1$ spheres that contain $c$, we obtain $\Theta(mk)$ incidences. Thus, the term $mk$ (and obviously also the term $n$) cannot be removed from the bound.

The case of $k = 3$ was considered 25 years ago by Chung [23]. Specifically, Chung proved that when $\mathcal{S}$ is a 3-non-degenerate set, then $I(\mathcal{P}, \mathcal{S}) = O\left(m^{4/5} n^{4/5} + nm^{1/2} + m\right)$. Recently, Zahl [77] presented an improved bound of $O\left(m^{3/4} n^{3/4} + n + m\right)$ for this case. Notice that any set of unit spheres in $\mathbb{R}^3$ is 3-non-degenerate; this special case has also been independently considered in [44]. Theorem 1.4 implies the same bound, up to the $m^\varepsilon$ factor, for any constant $k$. As far as we know, there are no previous results for $k > 3$. We note that a direct extension of Zahl's technique to cases where $k > 3$ would imply much weaker bounds.

### 1.1.2 Non-degenerate sets of planes

We next consider incidences between points and planes in $\mathbb{R}^3$. Let $\mathcal{P}$ be a set of $m$ points and $\Pi$ a set of $n$ planes, both in $\mathbb{R}^3$. We denote the number of point-plane incidences in $\mathcal{P} \times \Pi$ as $I(\mathcal{P}, \Pi)$.

Similarly as in the case of spheres, we want to avoid the existence of a line $\ell$ such that every point in $\mathcal{P}$ is contained in $\ell$, and every plane of $\Pi$ contains $\ell$, a situation that would lead to $I(\mathcal{P}, \Pi) = mn$. This has prompted the study of several types of "non-degenerate" scenarios. Edelsbrunner, Guibas, and Sharir [30] derived the bound $O(m^{3/5-\varepsilon} n^{4/5+2\varepsilon} + m + n \log m)$, for any $\varepsilon > 0$, under the assumption that there are no three collinear points in $\mathcal{P}$. Late, Apfelbaum and Sharir [7] derived the bound of $O\left(m^{3/4} n^{3/4} + m + n\right)$ under the condition that the incidence graph does not contain a copy of $K_{r,r}$ for some fixed constant $r \in \mathbb{N}$ (this is a slight improvement on an earlier result by Brass and Knauer [21]). Elekes and Tóth [32] define a plane $h$ to be *$\eta$-degenerate*,[1] for $0 < \eta < 1$, if for any line $\ell$ contained in $h$

$$|\ell \cap \mathcal{P}| \leq \eta \cdot |h \cap \mathcal{P}|.$$

They established the bound $\Theta(m^{3/4}n^{3/4} + mn^{1/2} + n)$ for the maximum number of incidences between $m$ points and $n$ $\eta$-degenerate planes.

**Definition 1.5.** *We say that a set $\Pi$ of $n$ planes is $k$-*non-degenerate*, for a constant $k$, $1 < k < n$, if there does not exist a line that is contained in $k$ planes of $\Pi$.*

Consider a set $\mathcal{P}$ of points, a set $\Pi$ of planes, and a point $p$ that is not incident to any plane of $\Pi$, all in $\mathbb{R}^3$. The *inversion transformation around $p$* transforms every plane of $\Pi$ to a sphere, all incident to $p$, while preserving the number of incidences with $\mathcal{P}$ (e.g., see [41, Chapter 37]). Moreover, if $\Pi$ is a $k$-non-degenerate set, then so is the resulting set of spheres. (Indeed, a circle not incident to $p$ can be contained in at most one of the image spheres, and a circle incident to $p$ is the image of a line which can lie in at most $k-1$ planes of $\Pi$.) Combining this with Theorem 1.4 immediately implies the following result.

**Corollary 1.6.** *Let $\mathcal{P}$ be a set of $m$ points and let $\Pi$ be a $k$-non-degenerate set of $n$ planes, both in $\mathbb{R}^3$, for some $1 < k < n$. Then for every $\varepsilon > 0$,*

$$I(\mathcal{P}, \Pi) = O(m^{3/4+\varepsilon}n^{3/4}k^{1/4} + n + mk),$$

*where the constant of proportionality depends on $\varepsilon$.*

We prove the following improved bound in Section 2.2 (this bound is stronger than the bound of Corollary 1.6 when $k = O(n/m^{1/3})$; for larger values of $k$, both bounds are dominated by the term $mk$ and are thus equivalent).

**Theorem 1.7.** *Let $\mathcal{P}$ be a set of $m$ points and let $\Pi$ be a $k$-non-degenerate set of $n$ planes, both in $\mathbb{R}^3$, for some $1 < k < n$. Then for every $\varepsilon > 0$,*

$$I(\mathcal{P}, \Pi) = O\left(m^{4/5+\varepsilon}n^{3/5}k^{2/5} + n + mk\right),$$

---

[1] Notice the somewhat confusing change of notation. Similar to Agarwal et al. [1], we use the term *non-degenerate* configurations, while Elekes and Tóth [32] use the term *degenerate* plane.

*where the constant of proportionality depends on $\varepsilon$.*

The scenario described in Theorem 1.7 is dual to the problem of point-plane incidences in $\mathbb{R}^3$, with no $k$ points lying on a common line. The aforementioned work of Edelsbrunner, Guibas, and Sharir [30] considers this problem for the special case of $k = 3$. By using a point-plane duality argument (e.g., see [26, Chapter 8]), we obtain the following generalization.

**Corollary 1.8.** *Let $\mathcal{P}$ be a set of $m$ points and let $\Pi$ be a set of $n$ planes, both in $\mathbb{R}^3$, such that no $k$ points of $\mathcal{P}$ are collinear. Then*

$$I(\mathcal{P}, \Pi) = O\left(n^{4/5+\varepsilon} m^{3/5} k^{2/5} + m + nk\right).$$

### 1.1.3 Applications

We believe that, due to their natural definition, $k$-non-degenerate sets of spheres/planes would have various applications. As an example, we present such an application in Section 4.1, which is an extension of the three-dimensional unit distances problem, in which one considers the $k$ most frequent distances. Given a set $\mathcal{P}$ of points in $\mathbb{R}^2$ and a set $D$ of $k$ distinct distances, we let $g(\mathcal{P}, D)$ denote the number of pairs of points of $\mathcal{P}$ that span a distance in $D$. Akutsu, Tamaki, and Tokuyama [4] studied $g_k(m) = \max g(\mathcal{P}, D)$, where the maximum is taken over all sets $\mathcal{P}$ of $m$ points in $\mathbb{R}^2$ and all sets $D$ of $k$ distinct distances. They derived the bound $g_k(m) = O(m^{10/7}k^{5/7}) \approx O(m^{1.4286}k^{0.7143})$. Solymosi, Tardos, and Tóth [70] established the bound $g_k(m) = O(m^{1.4571}k^{0.6286})$, which is an improvement when $k > n^{1/3}$.

We consider the three-dimensional variant of this problem and denote its value as $f_k(m)$. That is, given a set $\mathcal{P}$ of points in $\mathbb{R}^3$ and a set $D$ of $k$ distinct distances, we denote by $f(\mathcal{P}, D)$ the number of pairs of points of $\mathcal{P} \subset \mathbb{R}^3$ that span a distance in $D$. We set $f_k(m) = \max f(\mathcal{P}, D)$, where the maximum is taken over all sets $\mathcal{P}$ of $m$ points in $\mathbb{R}^3$ and all sets $D$ of $k$ distinct distances. Currently, the best known bound for the three-dimensional unit distances problem is $f_1(m) = O(m^{3/2})$ [44, 77]. By applying this bound independently for every distance in $D$, we obtain the trivial bound $f_k(m) = O(km^{3/2})$. In Section 4.1, we establish the following bound, which is an improvement for $k = \Omega(m^{25/48+\varepsilon}) \approx \Omega(m^{0.52})$.

**Theorem 1.9.** *For any $\varepsilon > 0$, $f_k(m) = O(m^{236/149+\varepsilon} k^{125/149}) \approx O(m^{1.58} k^{0.84})$.*

## 1.2 Sylvester-Gallai Type Problems

The Sylvester-Gallai theorem (Theorem 1.2) is perhaps one of the best known result in combinatorial geometry. The statement was conjectured by Sylvester in 1893 [71], and the first published proof is by Melchior [55]. Later proofs were given by Gallai in 1944 [36] and others. There are now several different proofs of the theorem. We study and improve upon a generalization of this theorem, referred to as Kelly's theorem, to the complex numbers.

Let $\mathcal{V} = \{v_1, v_2, \ldots, v_n\}$ be a set of $n$ points in $\mathbb{C}^d$. We denote by $\mathcal{L}(\mathcal{V})$ the set of lines determined by points in $\mathcal{V}$, and by $\mathcal{L}_r(\mathcal{V})$ (resp. $\mathcal{L}_{\geq r}(\mathcal{V})$) the set of lines in $\mathcal{L}(\mathcal{V})$ that contain exactly (resp. at least) $r$ points. Let $t_r(\mathcal{V})$ denote the size of $\mathcal{L}_r(\mathcal{V})$. Throughout the write-up we omit the argument $\mathcal{V}$ when the context makes it clear. We refer to $\mathcal{L}_2$ as the set of *ordinary lines*, and $\mathcal{L}_{\geq 3}$ as the set of *special lines*.

Melchior [55] proved the following, using arguments based on the Euler characteristic of $\mathbb{R}^2$.

**Theorem 1.10** (Melchior's inequality)**.** *Let $\mathcal{V}$ be a set of $n$ points in $\mathbb{R}^2$ that are not collinear. Then*

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

Theorem 1.10 in fact proves something stronger than the Sylvester-Gallai theorem, namely that 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** (Dirac-Motzkin conjecture)**.** *For every $n \neq 7, 13$, the number of ordinary lines determined by $n$ noncollinear points in the plane is at least $\lceil \frac{n}{2} \rceil$.*

There were several results on this question (see [57, 49, 25]), before Green and Tao [37] resolved it for large enough point sets.

**Theorem 1.11** (Green and Tao [37]). *Let $\mathcal{V}$ 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 $t_2(\mathcal{V}) \geq \frac{n}{2}$ for even $n$ and $t_2(\mathcal{V}) \geq \lfloor \frac{3n}{4} \rfloor$ for odd $n$.*

A nice history of the problem is given in [37] and there are several survey articles on the topic, see for example [18].

The Sylvester-Gallai theorem does not hold 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 counter example. A more general example is the following:

**Example 1** (Fermat configuration). *For any positive integer $k \geq 3$, let $\mathcal{V}$ be inflection points of the Fermat Curve $X^k + Y^k + Z^k = 0$ in $\mathbb{PC}^2$. Then $\mathcal{V}$ has $n = 3k$ points, in particular*

$$\mathcal{V} = \bigcup_{i=1}^{k} \{[1 : \omega^i : 0]\} \cup \{[\omega^i : 0 : 1]\} \cup \{[0 : 1 : \omega^i]\},$$

*where $\omega$ is a $k^{th}$ root of $-1$.*

*It is easy to check that $\mathcal{V}$ determines three lines containing k points each, while every other line contains exactly three points. In particular, $\mathcal{V}$ determines no ordinary lines.*[2]

In response to a question of Serre [62], Kelly [48] showed that when the points span more than two complex dimensions, the point set must determine at least one ordinary line.

**Theorem 1.12** (Kelly's theorem [48]). *Let $\mathcal{V}$ be a set of $n$ points in $\mathbb{C}^3$ that are not contained in a (complex) plane. Then there exists an ordinary line determined by points of $\mathcal{V}$.*

Kelly's proof of Theorem 1.12 used a deep result of Hirzebruch [43] from algebraic geometry. More specifically, it used the following result, known as Hirzebruch's inequality.

[2] *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.*

**Theorem 1.13** (Hirzebruch's inequality [43])**.** *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*

$$t_2(\mathcal{V}) + \frac{3}{4}t_3(\mathcal{V}) \geq n + \sum_{r\geq 5}(2r-9)t_r(\mathcal{V}).$$

Theorem 1.13 requires that no line contains more than $n-2$ points. Under a stronger assumption, i.e., no line contains more than $2n/3$ points, Bojanowski [17] obtained a better lower bound.

**Theorem 1.14** (Bojanowski [17])**.** *Let $\mathcal{V}$ be a set of $n$ points in $\mathbb{C}^2$, such that $t_r(\mathcal{V}) = 0$ for $r > 2n/3$. Then*

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

More elementary proofs of Theorem 1.12 were given in [33] and [29]. 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 [29], we make the first progress in this direction.

**Theorem 1.15.** *Let $\mathcal{V}$ be a set of $n \geq 24$ points in $\mathbb{C}^3$ not contained in a (complex) plane. Then $\mathcal{V}$ determines at least $\frac{3}{2}n$ ordinary lines, unless $n-1$ points are on a plane in which case there are at least $n-1$ ordinary lines.*

Clearly if $n-1$ points are coplanar, it is possible to have only $n-1$ 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 $\frac{3}{2}n$ bound when at most $n-2$ points are contained in any plane. Using a similar argument, for point sets in $\mathbb{R}^3$, Theorems 1.11 and 1.15 give us the following easy corollary.

**Corollary 1.16.** *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 $\frac{3}{2}n - 1$ ordinary lines.*

When $\mathcal{V}$ is sufficiently non-degenerate, i.e., no plane contains too many points, we are able to give a more refined bound in the spirit of Melchior's and Hirzebruch's inequalities, taking into account the existence of lines with more than three points. In particular, we show the following (the constant 1/2 in Theorem 1.17 is arbitrary and can be replaced by any positive constant smaller than 1):

**Theorem 1.17.** *There exists an absolute constant $c > 0$ and a positive integer $n_0$ such that the following holds. Let $\mathcal{V}$ be a set of $n \geq n_0$ points in $\mathbb{C}^3$ with at most $\frac{1}{2}n$ points contained in any plane. Then*

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

We note that having at most a constant fraction of the points on any plane is necessary in order to obtain a bound as in Theorem 1.17. Indeed, let $\mathcal{V}$ consist of the Fermat Configuration for some $k \geq 3$ on a plane and $o(k)$ points not on the plane. Then $\mathcal{V}$ has $O(k)$ points and determines $o(k^2)$ ordinary lines. On the other hand, $\sum_{r\geq 4} r^2 t_r(\mathcal{V}) = \Omega(k^2)$.

Suppose that $\mathcal{V}$ consists of $n-k$ points on a plane, and $k$ points not on the plane. There are at least $n-k$ lines through each point not on the plane, at most $k-1$ of which could contain three or more points, i.e., $\mathcal{V}$ determines at least $k(n-2k)$ ordinary lines. Then if $k = \epsilon n$, $0 < \epsilon < 1/2$, $\mathcal{V}$ will have $\Omega_\epsilon(n^2)$ ordinary lines, where the hidden constant depends on $\epsilon$. Therefore, the bound in Theorem 1.17 is only interesting when no plane contains too many points.

Theorems 1.13 and 1.14 (which also give a bound in $\mathbb{C}^3$) only give a lower bound on $t_2(\mathcal{V}) + \frac{3}{4}t_3(\mathcal{V})$, whereas both Theorems 1.15 and 1.17 give lower bounds on the number of ordinary lines, i.e., $t_2(\mathcal{V})$. We also note that lines with four points do not play any role in Theorems 1.13 and 1.14, where the summation starts at $r = 5$. This is not the case for Theorem 1.17. As a consequence, we have that if a non-planar configuration over $\mathbb{C}$ has many 4-rich lines, then it must have many ordinary lines.

Finally, when a point set $\mathcal{V}$ spans four or more dimensions in a sufficiently non-degenerate manner, i.e., no three-dimensional affine subspace contains too many points, we can show that there must be a quadratic number of ordinary lines.

**Theorem 1.18.** *There exists a positive integer $n_0$ such that the following holds. Let $\mathcal{V}$ be a set of $n \geq n_0$ points in $\mathbb{C}^4$ with at most $\frac{1}{2}n$ points contained in any three-dimensional affine subspace. Then*

$$t_2(\mathcal{V}) \geq \frac{1}{16}n^2.$$

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 in front of $n^2$. A quadratic lower bound may also be possible if at most $\frac{1}{2}n$ points are contained in any two dimensional space, but we have no proof or counterexample.

Note that while we state Theorems 1.15 and 1.17 over $\mathbb{C}^3$ and Theorem 1.18 over $\mathbb{C}^4$, the same bounds hold in higher dimensions as well since we may project a point set in $\mathbb{C}^d$, $d > 4$, onto a generic lower dimensional subspace, preserving the incidence structures. Finally, although these theorems are stated over $\mathbb{C}$, these results are also new and interesting over $\mathbb{R}$.

## 1.3 Sum-Product Estimates

Let $\mathbb{F}$ be a field, and $A \subset \mathbb{F}$ be a finite set. We define the *sum set* of $A$ to be the set

$$A + A = \{a + b : a, b \in A\}.$$

Similarly, we define the *difference set, product set* and *ratio set* respectively as

$$A - A = \{a - b : a, b \in A\}, \quad AA = \{ab : a, b \in A\}, \quad A/A = \{a/b : a, b \in A\}.$$

When $A \subset \mathbb{R}$, the *sum-product conjecture* of Erdős and Szemerédi [35] states that, for any $\varepsilon > 0$,

$$\max\{|AA|, |A + A|\} = \Omega(|A|^{2-\varepsilon}).$$

The conjecture quantifies the idea that finite subsets of the reals can not have additive and multiplicative structure simultaneously. The central theme of our result is based on ideas introduced by Solymosi [68] in the proof of the following result.

**Theorem 1.19** (Solymosi [68])**.** *Let $A \subset \mathbb{R}$ be a finite set. Then*

$$|A + A|^2|AA| = \Omega^*\left(|A|^4\right).$$

*As a consequence, we obtain*

$$\max\{|AA|, |A+A|\} = \Omega^*\left(|A|^{4/3}\right).$$

Throughout this thesis, we use the $O^*(\cdot)$ and $\Omega^*(\cdot)$ notation to hide sub-polynomial factors. By taking $A$ to be an arithmetic progression, it is easy to see that the bound in Theorem 1.19 can not be improved. Recently, Konyagin and Shkredov [51, 52] were able to obtain a better bound on the sum-product problem by studying what happens when the bound in Theorem 1.19 is close to being tight.

**Theorem 1.20** (Konyagin, Shkredov [52])**.** *Let $A \subset \mathbb{R}$ be a finite set. Then, for any $c < 5/9813$,*

$$\max\{|AA|, |A+A|\} = \Omega^*\left(|A|^{4/3+c}\right).$$

We refer the reader to the excellent exposition of the proof of Theorem 1.20 by Sheffer [65].

In a similar vein to the sum-product problem, it is expected that a set defined by a combination of additive and multiplicative operations on a given set $A$ should be large compared to $A$. One such example is the following result of Balog and Roche-Newton.

**Theorem 1.21** (Balog, Roche-Newton [11])**.** *Let $A \subset \mathbb{R}$ be a finite set. Then*

$$\left|\frac{A+A}{A+A}\right| \geq 2|A|^2 - 1.$$

Later work by Roche-Newton [59] and then Lund [53] gave the following improvement on Theorem 1.21.

**Theorem 1.22** (Lund [53])**.** *Let $A \subset \mathbb{R}$ be a finite set. Then*

$$\left|\frac{A+A}{A+A}\right| = \Omega^*\left(\frac{|A|^{2+1/4}}{|A/A|^{1/8}}\right).$$

Konyagin and Rudnev [50] used an elegant argument to show that Theorem 1.19 also holds when $A$ is a subset of $\mathbb{C}$. We extend their work and show that Theorems 1.20 and 1.22 also hold over the complex numbers. As a consequence, we obtain the following theorems.

**Theorem 1.23.** *Let $A \subset \mathbb{C}$ be a finite set. Then, for any $c < 5/9813$,*

$$\max\{|AA|, |A+A|\} = \Omega^*\left(|A|^{4/3+c}\right).$$

**Theorem 1.24.** *Let $A \subset \mathbb{C}$ be a finite set. Then, for every $\varepsilon > 0$,*

$$\left|\frac{A+A}{A+A}\right| = \Omega^*\left(\frac{|A|^{2+1/4-\varepsilon}}{|A/A|^{1/8}}\right),$$

*where the constant of proportionality depends on $\varepsilon$.*


**Acknowledgments**

The results in Chapter 2 and Section 4.1 are based on joint work with Adam Sheffer that was published in [15]. Chapter 3 is based on joint work with Zeev Dvir, Shubhangi Saraf and Charles Wolf and was published in [13]. Section 4.2 is based on ongoing work with Ben Lund and Adam Sheffer [14].

# Chapter 2

# Incidences with Non-Degenerate Sets

A common approach to bounding the number of incidences over $\mathbb{R}$ is the following. One first proves a *weak bound* using Cauchy-Schwarz or Hölder's inequality. Then, one can partition the space into *well-behaved* cells, i.e., in such a way that the objects being studied do not interact with too many cells. The weak bound is then applied to each of these cells to obtain the final bound. One way to partition the space is the use the notion of *cuttings* (e.g., see [54, Chapter 4]). In their solution of the Erdős distinct distances problem, Guth and Katz introduced a new method to partition space using polynomials. This method is now referred to as the *polynomial partitioning method*. This method appears to be better than cuttings in terms of better bounds on complexity and handing of lower-dimensional objects.

## 2.1 Algebraic Preliminaries

Consider a set $\mathcal{P}$ of $m$ points in $\mathbb{R}^d$. Given a polynomial $f \in \mathbb{R}[x_1, \ldots, x_d]$, we define the *zero set* of $f$ to be $Z(f) = \{p \in \mathbb{R}^d \mid f(p) = 0\}$. For $1 < r \leq m$, we say that $f \in \mathbb{R}[x_1, \ldots, x_d]$ is an *r-partitioning polynomial* for $\mathcal{P}$ if no connected component (referred to as cell *cell*) of $\mathbb{R}^d \setminus Z(f)$ contains more than $m/r$ points of $\mathcal{P}$. Notice that there is no restriction on the number of points of $\mathcal{P}$ that lie in $Z(f)$. The following result is due to Guth and Katz [39].

**Theorem 2.1. (Polynomial partitioning [39])** *Let $\mathcal{P}$ be a set of $m$ points in $\mathbb{R}^d$, for some fixed constant dimension $d$. Then for every $1 < r \leq m$, there exists an $r$-partitioning polynomial $f \in \mathbb{R}[x_1, \ldots, x_d]$ of degree $D = O(r^{1/d})$, where the implicit constant depend on the dimension $d$.*

To use the above theorem effectively, we need to bound the number of cells in the resulting partition. For this, we rely on the following theorem (see also [5, 45]).

**Theorem 2.2** (Warren's Theorem [75]). *For a polynomial $f \in \mathbb{R}[x_1, \ldots, x_d]$ of degree $D$, the number of connected components of $\mathbb{R}^d \setminus Z(f)$ is $O\left((2D)^d\right)$.*

As a consequence, for an $r$-partitioning polynomial as defined above, the number of connected components in the resulting partition is $O(r)$ (the degree is $O(r^{1/d})$ and the dimension $d$ is a fixed constant). In this work, we are only interested in the case $d = 3$. When using a partitioning polynomial for an incidence problem, usually the main difficulty in the analysis is bounding the number of incidences with points contained in the zero set $Z(f)$ of the partitioning polynomial. One technique for overcoming this difficulty is presented in the following theorem, which is proved in [77] and in [44]. Our formulation follows the one presented in [44].

**Theorem 2.3. (Second partitioning polynomial [44, 77])** *Given an irreducible trivariate polynomial $f$ of degree $D$, a parameter $E \geq D$, and a finite point set $\mathcal{P}$ in $\mathbb{R}^3$, there is a polynomial $g$ of degree at most $E$, co-prime with $f$, which partitions $\mathcal{P}$ into subsets $Q_0 \subset Z(g)$ and $Q_1, \ldots, Q_t$, for $t = \Theta(DE^2)$, so that each $Q_i$, $i = 1, \ldots, t$, lies in a distinct connected component of $\mathbb{R}^3 \setminus Z(g)$, and $|Q_i| \leq |Q|/t$.*

This theorem provides us with a second partitioning polynomial that can be used to partition the points contained in the zero set of the first partitioning polynomial. It is possible that some of the points will be contained in both zero sets. However, since the two polynomials have no common factors, their intersection is a one-dimensional variety, making the analysis easier (though by no means trivial).

In order to study the interaction of the spheres and planes with the partitioning polynomial, we will rely on several results from algebraic geometry. The following result gives bounds on the number of connected components of an algebraic variety.

**Theorem 2.4** (The Milnor-Thom Theorem [56, 73] (see also [5])). *Let $V$ be a real variety in $\mathbb{R}^d$, that is the solution set of real polynomial equations*

$$f_i(x_1, \ldots, x_d) = 0, \quad (i = 1, \ldots, m),$$

*and suppose that the degree of each polynomial $f_i$ is at most $D$. Then the number of connected components of $V$ is at most $D(2D-1)^{d-1}$.*

We also need the following basic properties of the zero set of polynomials in the plane.

**Theorem 2.5** (Bézout's theorem (e.g., see [45])). *Let $f, g$ be two polynomials in $\mathbb{R}[x_1, x_2]$ of degrees $D_f$ and $D_g$ respectively, with no common factors. Then $Z(f)$ and $Z(g)$ have at most $D_f D_g$ points in common, i.e., $|Z(f) \cap Z(g)| \leq D_f D_g$.*

A detailed explanation of Theorem 2.5, though focusing on the complex plane, can be found in [24, Section 8.7]. The following theorem, similar to Theorem 2.4, gives bounds on the number of connected components of a plane curve. The following version is from [45].

**Theorem 2.6** (Harnack's curve theorem [40]). *Let $f \in \mathbb{R}[x, y]$ be a bivariate polynomial of degree $D$. Then the number of (arcwise) connected components of $Z(f)$ is at most $1 + \binom{D-1}{2}$. The bound is tight in the worst case.*

## 2.2 Proving the Incidence Bounds

In this section we prove our incidence bounds, i.e., Theorems 1.4 and 1.7. Since the proofs are very similar, we present the full details of the first proof and skip several identical arguments in the second one. Several algebraic issues that are related to intersections of varieties are deferred to Section 2.3. The proofs rely on the so-called "second partitioning polynomial technique," presented in [77] and [44]. We use constant-degree partitioning polynomials, as in [64, 69]. As far as we know, this is the first combination of these two techniques (i.e., using a second partitioning polynomial where both polynomials are constant-degree).

### 2.2.1 Proof of Theorem 1.4

We start by deriving a weaker bound, which will then be used in the derivation of the sharper bound of Theorem 1.4.

**Lemma 2.7.** *Let $\mathcal{P}$ be a set of $m$ points and $\mathcal{S}$ be a $k$-non-degenerate set of $n$ spheres, both in $\mathbb{R}^3$, for some $1 < k < n$. Then*

$$I(\mathcal{P}, \mathcal{S}) \le mn^{2/3}k^{1/3} + 2n.$$

*Proof.* The proof is a variant of the standard proof of the Kővari–Sós–Túran theorem in extremal graph theory (e.g., see [54, Section 4.5]). We double-count the number $Q$ of quadruples $(a, b, c, \sigma)$, where $\sigma \in \mathcal{S}$, $a, b, c \in \mathcal{P} \cap \sigma$ and $a, b, c$ are distinct. On one hand, there are $\binom{m}{3}$ triples of points of $\mathcal{P}$, and every sphere that contains all three points must contain the unique circle that they span. Since any such circle can be contained in at most $k - 1$ spheres of $\mathcal{S}$, we have

$$Q \le \binom{m}{3}(k-1) < \frac{1}{6}m^3k. \tag{2.1}$$

On the other hand, for each $\sigma \in \mathcal{S}$, put $d_\sigma = |\mathcal{P} \cap \sigma|$, so that $I(\mathcal{P}, \mathcal{S}) = \sum_{\sigma \in \mathcal{S}} d_\sigma$. Then, by Hölder's inequality, we have

$$Q = \sum_{\sigma \in \mathcal{S}} \binom{d_\sigma}{3} \ge \frac{1}{6}\sum_{\sigma \in \mathcal{S}}(d_\sigma - 2)^3 \ge \frac{1}{6n^2}\left(\sum_{\sigma \in \mathcal{S}}(d_\sigma - 2)\right)^3 = \frac{1}{6n^2}\Big(I(\mathcal{P}, \mathcal{S}) - 2n\Big)^3. \tag{2.2}$$

By combining (2.1) and (2.2), we obtain

$$\frac{1}{6n^2}\Big(I(\mathcal{P}, \mathcal{S}) - 2n\Big)^3 < \frac{1}{6}m^3k.$$

Hence $I(\mathcal{P}, \mathcal{S}) \le mn^{2/3}k^{1/3} + 2n$, as asserted. □

We are now ready to prove our bound on the number of incidences with $k$-non-degenerate sets of spheres. For the convenience of the reader, we repeat the statement of the theorem before presenting its proof.

**Theorem 1.4.** *Let $\mathcal{P}$ be a set of $m$ points and $\mathcal{S}$ be a $k$-non-degenerate set of $n$ spheres, both in $\mathbb{R}^3$, for some $1 < k < n$. Then for every $\varepsilon > 0$,*

$$I(\mathcal{P}, \mathcal{S}) = O(m^{3/4+\varepsilon}n^{3/4}k^{1/4} + n + mk),$$

*where the constant of proportionality depends on $\varepsilon$.*

*Proof.* When $m = O(n^{1/3}/k^{1/3})$, the bound in Lemma 2.7 implies $I(\mathcal{P},\mathcal{S}) = O(n)$. Thus, in the rest of the proof we may assume that $n = O\left(m^3 k\right)$.

We prove the theorem by induction on $m+n$. Specifically, we prove by induction that

$$I(\mathcal{P},\mathcal{S}) \le \alpha_1 m^{3/4+\varepsilon} n^{3/4} k^{1/4} + \alpha_2(n+mk),$$

for any fixed $\varepsilon > 0$ and sufficiently large constants $\alpha_1, \alpha_2$ that depend on $\varepsilon$.

For the induction basis, by choosing $\alpha_2$ to be sufficiently large, we obtain that the bound holds for small values of $m, n$. We next consider the induction step.

**Partitioning the space.**

We construct an $r$-partitioning polynomial $f$ for $\mathcal{P}$ of degree $D = O(r^{1/3})$, where $r$ is a sufficiently large constant which will be determined later. For convenience we work with $D$ instead of $r$. Denote the open cells of the partitioning as $C_1, \ldots, C_t$, where $t = O(D^3)$. Let $n_i$ denote the number of spheres of $\mathcal{S}$ that intersect the cell $C_i$, and let $m_i$ denote the number of points contained in $C_i$. We have $m_i = O(m/D^3)$ for every $1 \le i \le t$. We write $m' = \sum_{i=1}^t m_i$.

Let $\mathcal{P}_0$ denote the subset of points of $\mathcal{P}$ contained in $Z(f)$, and let $\mathcal{P}' = \mathcal{P} \setminus \mathcal{P}_0$. Clearly

$$I(\mathcal{P},\mathcal{S}) = I(\mathcal{P}_0,\mathcal{S}) + I(\mathcal{P}',\mathcal{S}). \tag{2.3}$$

By the induction hypothesis, we have

$$\begin{aligned} I(\mathcal{P}',\mathcal{S}) &\le \sum_{i=1}^t \left(\alpha_1 m_i^{3/4+\varepsilon} n_i^{3/4} k^{1/4} + \alpha_2(n_i + m_i k)\right) \\ &\le O\left(\alpha_1 \frac{m^{3/4+\varepsilon}}{D^{9/4+3\varepsilon}} k^{1/4} \sum_{i=1}^t n_i^{3/4}\right) + \alpha_2 \sum_{i=1}^t n_i + \alpha_2 m' k. \end{aligned}$$

By Lemma 2.9 in Section 2.3, every sphere of $\mathcal{S}$ intersects $O(D^2)$ cells of the partitioning. Therefore, $\sum_{i=1}^t n_i = O\left(nD^2\right)$. Combining this with Hölder's inequality implies

$$\sum_{i=1}^t n_i^{3/4} = O\left(\left(nD^2\right)^{3/4} t^{1/4}\right) = O\left(\left(nD^2\right)^{3/4} \left(D^3\right)^{1/4}\right) = O\left(n^{3/4} D^{9/4}\right).$$

Hence

$$I(\mathcal{P}',\mathcal{S}) = O\left(\alpha_1 \frac{m^{3/4+\varepsilon} n^{3/4} k^{1/4}}{D^{3\varepsilon}} + \alpha_2 n D^2\right) + \alpha_2 m' k.$$

Since $n = O\left(m^3 k\right)$, we have $n^{1/4} = O\left(m^{3/4}k^{1/4}\right)$ or $n = O\left(m^{3/4}n^{3/4}k^{1/4}\right)$. By choosing $\alpha_1$ to be sufficiently larger than $\alpha_2$ and and choosing $D$ to be sufficiently larger than the constant in the $O()$-notation, we have

$$I(\mathcal{P}', \mathcal{S}) \leq \frac{\alpha_1}{3} m^{3/4+\varepsilon} n^{3/4} k^{1/4} + \alpha_2 m' k. \tag{2.4}$$

**Bounding the number of incidences on the partitioning polynomial.**

To bound $I(\mathcal{P}_0, \mathcal{S})$, we construct a second partitioning polynomial with respect to $\mathcal{P}_0$, as stated in Theorem 2.3. Since this theorem requires the first partitioning polynomial to be irreducible, we factor $f$ into irreducible factors $f_1, \ldots, f_s$. Then, for each $f_i$ of degree $D_i$, we apply Theorem 2.3 to construct a second partitioning polynomial $g_i$, co-prime with $f_i$ and of a large constant degree $E_i \geq D_i$, which we will specify later. We assume that every point of $\mathcal{P}_0$ is incident to exactly one of the zero sets $Z(f_i)$. If a point $p \in \mathcal{P}_0$ is incident to more than one such zero set, we arbitrarily choose one of these sets $Z(f_i)$ and treat $p$ as if it were incident only to $Z(f_i)$. Let $\mathcal{P}_0^i$ denote the set of points of $\mathcal{P}_0$ contained in $Z(f_i)$ (in the sense just defined, for points contained in more than one zero set) but not in $Z(g_i)$, and set $\mathcal{P}_{0,0}^i = \mathcal{P} \cap Z(f_i) \cap Z(g_i)$. We have

$$I(\mathcal{P}_0, \mathcal{S}) = \sum_{i=1}^{s} \Big( I(\mathcal{P}_0^i, \mathcal{S}) + I(\mathcal{P}_{0,0}^i, \mathcal{S}) \Big). \tag{2.5}$$

Let us denote the cells of the second partitioning polynomial $g_i$ that contain at least one point of $\mathcal{P}_0^i$ as $C_{i,1}, \ldots, C_{i,t_i}$, where $t_i = O(D_i E_i^2)$. Let $n_{i,j}$ denote the number of spheres that intersect the cell $C_{i,j}$, and set $m_0^i = |\mathcal{P}_0^i|$. The number of points of $\mathcal{P}_0^i \cap C_{i,j}$ is $O\left(m_0^i/(D_i E_i^2)\right)$.

When $m_0^i = O(n^{1/3}/k^{1/3})$, the bound in Lemma 2.7 implies $I(\mathcal{P}_0^i, \mathcal{S}) = O(n)$. Let $A$ be the set of indices $i$ with the property that $m_0^i = O(n^{1/3}/k^{1/3})$. By taking $\alpha_2$ to be sufficiently larger than $r$ (and thus also to $s$), we have

$$\sum_{i \in A} I(\mathcal{P}_0^i, \mathcal{S}) \leq \alpha_2 n/2. \tag{2.6}$$

We next consider the case of $i \notin A$. We have $n = O\left(\left(m_0^i\right)^3 k\right)$, which in turn implies $n = O\left(\left(m_0^i\right)^{3/4} n^{3/4} k^{1/4}\right)$. If $Z(f_i)$ is a sphere of $\mathcal{S}$, then this sphere is incident to all $m_0^i$

points of $\mathcal{P}_0^i$. By Lemma 2.11 in Section 2.3, any other sphere of $\mathcal{S}$ is incident to points of $\mathcal{P}_0^i$ in $O(D_iE_i)$ cells of the second partitioning. Therefore, $\sum_{j=1}^{t_i} n_{i,j} = O(nD_iE_i)$ (possibly ignoring a sphere that coincides with $Z(f_i)$). Using Hölder's inequality as above implies $\sum_{j=1}^{t_i} n_{i,j}^{3/4} = O\left(n^{3/4}D_iE_i^{5/4}\right)$. Thus, for $i \notin A$, we have

$$\begin{aligned}
I(\mathcal{P}_0^i,\mathcal{S}) &\le O\left(\sum_{j=1}^{t_i} \alpha_1 \frac{(m_0^i)^{3/4+\varepsilon}}{D_i^{3/4+\varepsilon}E_i^{6/4+2\varepsilon}} n_{i,j}^{3/4}k^{1/4}\right) + \alpha_2\left(\sum_{j=1}^{t_i} n_{i,j} + m_0^i k\right)\\
&= O\left(\alpha_1 (m_0^i)^{3/4+\varepsilon} n^{3/4}k^{1/4}\cdot \frac{D_i^{1/4-\varepsilon}}{E_i^{1/4+2\varepsilon}} + \alpha_2 nD_iE_i\right) + \alpha_2 m_0^i k\\
&\le \frac{\alpha_1}{3D}(m_0^i)^{3/4+\varepsilon}n^{3/4}k^{1/4} + \alpha_2 m_0^i k, \qquad (2.7)
\end{aligned}$$

when $\alpha_1$ is sufficiently larger than $\alpha_2$, and $E_i$ is sufficiently large than the constant of the $O()$-notation (recall that $n = O\left((m_0^i)^{3/4}n^{3/4}k^{1/4}\right)$). Set $m'' = \sum_{i=1}^{s} m_0^i$. Combining (2.7) with (2.6) and recalling that $s \le D$, we have

$$\sum_{i=1}^{s} I(\mathcal{P}_0^i,\mathcal{S}) \le \frac{\alpha_1}{3} m^{3/4+\varepsilon}n^{3/4}k^{1/4} + \alpha_2(n/2 + m''k). \qquad (2.8)$$

Set $m_{0,0}^i = |\mathcal{P}\cap Z(f_i)\cap Z(g_i)|$ and $m_{0,0} = \sum_{i=1}^{s} m_{0,0}^i$. Lemma 2.12 implies

$$I(\mathcal{P}_{0,0}^i,\mathcal{S}) \; = \; O(m_{0,0}^i kD_iE_i \; + \; nD_iE_i).$$

Once again, by choosing $\alpha_2$ to be sufficiently large compared to $E_i$ and $D$, we have $I(\mathcal{P}_{0,0}^i,\mathcal{S}) \le \alpha_2(m_{0,0}^i k + n/(2D))$, which in turn implies

$$\sum_{i=1}^{s} I(\mathcal{P}_{0,0}^i,\mathcal{S}) \le \alpha_2(m_{0,0}k + n/2). \qquad (2.9)$$

Finally, the assertion of the theorem follows immediately by combining (2.3), (2.4), (2.5), (2.8), and (2.9). □

### 2.2.2 Proof of Theorem 1.7

This proof is very similar to the one in the previous section. Once again, we begin by deriving a weaker bound.

**Lemma 2.8.** *Let $\mathcal{P}$ be a set of $m$ points and $\Pi$ be a $k$-non-degenerate set of $n$ planes, both in $\mathbb{R}^3$, for some $1 < k < n$. Then*

$$I(\mathcal{P},\Pi) \le m\sqrt{nk} + n.$$

The proof of Lemma 2.8 is almost identical to the proof of Lemma 2.7. The main change is that in this case we double-count a set of triples (of $\mathcal{P} \times \mathcal{P} \times \Pi$), rather than a set of quadruples.

**Theorem 1.7.** *Let $\mathcal{P}$ be a set of $m$ points and $\Pi$ be a $k$-non-degenerate set of $n$ planes, both in $\mathbb{R}^3$, for some $1 < k < n$. Then for every $\varepsilon > 0$,*

$$I(\mathcal{P}, \Pi) = O\left(m^{4/5+\varepsilon} n^{3/5} k^{2/5} + n + mk\right),$$

*where the constant of proportionality depends on $\varepsilon$.*

*Proof.* This proof is very similar to the one of Theorem 1.4. Thus, we skip various details that are identical in both proofs.

By Lemma 2.8, we may assume that $n = O\left(m^2 k\right)$. We prove the theorem by induction on $m + n$. Specifically, we prove by induction that for any fixed $\varepsilon > 0$,

$$I(\mathcal{P}, \Pi) \le \alpha_1 m^{4/5+\varepsilon} n^{3/5} k^{2/5} + \alpha_2 (n + mk),$$

for sufficiently large constants $\alpha_1, \alpha_2$ that depend on $\varepsilon$. As before, the induction basis is straight forward.

**Partitioning the space.**

For the induction step, we construct an $r$-partitioning polynomial $f$ for $\mathcal{P}$ of degree at most $D = O(r^{1/3})$, where $r$ is a sufficiently large constant. Denote the open cells of the partitioning as $C_1, \ldots, C_t$, where $t = O\left(D^3\right)$. Let $n_i$ denote the number of planes of $\Pi$ intersecting the cell $C_i$, and let $m_i$ denote the number of points contained in $C_i$. We write $m' = \sum_{i=1}^{t} m_i$.

Let $\mathcal{P}_0$ denote the subset of points of $\mathcal{P}$ contained in $Z(f)$, and write $\mathcal{P}' = \mathcal{P} \setminus \mathcal{P}_0$. Notice that

$$I(\mathcal{P}, \Pi) = I(\mathcal{P}_0, \Pi) + I(\mathcal{P}', \Pi). \tag{2.10}$$

We analyze the incidences with $\mathcal{P}'$ just as before. The only difference is how we bound $\sum_{i=1}^{t} n_i$. The number of cells intersected by a hyperplane $h$ of $\Pi$ is at most the number of connected components of $h \setminus Z(f)$. According to Theorem 2.2, this number is $O(D^2)$. Therefore, $\sum_{i=1}^{t} n_i = O\left(nD^2\right)$. Completing the analysis as before yields

$$I(\mathcal{P}', \Pi) \le \frac{\alpha_1}{3} m^{4/5} n^{3/5} k^{2/5} + \alpha_2 m' k. \tag{2.11}$$

**Bounding the number of incidences on the partitioning polynomial.**

To bound $I(\mathcal{P}_0, \Pi)$, we construct a second partitioning polynomial with respect to $\mathcal{P}_0$, as stated in Theorem 2.3. Since this theorem requires the first partitioning polynomial to be irreducible, we factor $f$ into irreducible factors $f_1, \ldots, f_s$. Then, for each $f_i$ of degree $D_i$, we apply Theorem 2.3 to construct a second partitioning polynomial $g_i$, co-prime with $f_i$ and of a large constant degree $E_i \geq D_i$. We assume that every point of $\mathcal{P}_0$ is incident to exactly one of the zero sets $Z(f_i)$. If a point $p \in \mathcal{P}_0$ is incident to more than one such zero set, we arbitrarily choose one of these sets $Z(f_i)$ and treat $p$ as if it were incident only to $Z(f_i)$. Let $\mathcal{P}_0^i$ denote the set of points of $\mathcal{P}_0$ contained in $Z(f_i)$ but not in $Z(g_i)$, and set $\mathcal{P}_{0,0}^i = \mathcal{P} \cap Z(f_i) \cap Z(g_i)$. We have

$$I(\mathcal{P}_0, \Pi) = \sum_{i=1}^{s} \Big( I(\mathcal{P}_0^i, \Pi) + I(\mathcal{P}_{0,0}^i, \Pi) \Big). \tag{2.12}$$

Set $m'' = \sum_{i=1}^{s} |\mathcal{P}_0^i|$ and $m_{0,0} = \sum_{i=1}^{s} |\mathcal{P}_{0,0}^i|$. We again repeat the analysis from the proof of Theorem 1.4. By replacing Lemma 2.11 with Lemma 2.10, we obtain

$$\sum_{i=1}^{s} I(\mathcal{P}_0^i, \Pi) \leq \frac{\alpha_1}{3} m^{4/5+\varepsilon} n^{3/5} k^{2/5} + \alpha_2 (nk/2 + m''k). \tag{2.13}$$

Similarly, by relying on the proof of Theorem 1.4 and replacing Lemma 2.12 with Corollary 2.13, we obtain

$$\sum_{i=1}^{s} I(\mathcal{P}_{0,0}^i, \Pi) \leq \alpha_2 (m_0 k + n/2). \tag{2.14}$$

Finally, the assertion of the theorem is immediately obtained by combining (2.10), (2.11), (2.12), (2.13), and (2.14). □

## 2.3 Intersections with Partitioning Polynomials

In this section we establish several claims that were made in the proofs of Theorems 1.4 and 1.7 concerning intersections of spheres and planes with zero sets of partitioning polynomials. Some of these proofs are variants of proofs from [44].

Let $\psi : \mathbb{R}^2 \to \mathbb{R}^3$ be the *inverse stereographic projection* given by

$$\psi(u, v) = \left( x_0 + \frac{2ur}{u^2 + v^2 + 1}, \quad y_0 + \frac{2vr}{u^2 + v^2 + 1}, \quad z_0 + \frac{r(u^2 + v^2 - 1)}{u^2 + v^2 + 1} \right),$$

for a given point $(x_0, y_0, z_0)$ and constant $r$. That is, $\psi$ is a mapping from $\mathbb{R}^2$ to the sphere in $\mathbb{R}^3$ whose center is $(x_0, y_0, z_0)$ and whose radius is $r$, excluding the top point of the sphere. Since the following lemmas present asymptotic bounds, ignoring this single point will not affect them.

**Lemma 2.9.** *Let $f \in \mathbb{R}[x, y, z]$ be a polynomial of degree $D \geq 1$ and let $\sigma$ be a sphere in $\mathbb{R}^3$. Then $\sigma$ intersects at most $c'D^2$ cells of $\mathbb{R}^3 \setminus Z(f)$, for some absolute constant $c'$.*

*Proof.* If $\sigma$ is contained in $Z(f)$, then $\sigma$ does not intersect any cells. We may thus assume that $\sigma$ is not contained in $Z(f)$.

Let $\tilde{f} = (u^2 + v^2 + 1)^D \cdot (f \circ \psi)$, where $\psi$ maps $\mathbb{R}^2$ to the sphere $\sigma$. It is not difficult to verify that $\tilde{f}$ is a polynomial of degree at most $2D$. The zero set $Z(\tilde{f}) \subset \mathbb{R}^2$ is the image of a mapping of $Z(f) \cap \sigma$ to the plane. Every cell of $\mathbb{R}^3 \setminus Z(f)$ that is intersected by $\sigma$ corresponds to at least one cell of $\mathbb{R}^2 \setminus Z(\tilde{f})$ (i.e., the image of the planar cell under $\psi$ is contained in the three dimensional cell). Thus, the number of cells of $\mathbb{R}^3 \setminus Z(f)$ that $\sigma$ intersects is upper bounded by the number of cells in $\mathbb{R}^2 \setminus Z(\tilde{f})$. According to Theorem 2.2, this number is $O(D^2)$. □

**Lemma 2.10.** *Let $f, g \in \mathbb{R}[x, y, z]$ be co-prime polynomials of degrees $D$ and $E$, respectively, with $D \leq E$. Let $h$ be a plane in $\mathbb{R}^3$, such that $h \not\subseteq Z(f)$. Then the number of cells of $\mathbb{R}^3 \setminus Z(g)$ intersected by $h \cap Z(f)$ is at most $c'DE$ for some absolute constant $c'$.*

*Proof.* We may assume that $h$ is not contained in $Z(g)$, since otherwise no cell of $\mathbb{R}^3 \setminus Z(g)$ is intersected by $h$. Set $\gamma_f = h \cap Z(f)$ and $\gamma_g = h \cap Z(g)$. Each three-dimensional cell of $\mathbb{R}^3 \setminus Z(g)$ that intersects $h$ corresponds to a set of cells in $h$. Every cell of $\mathbb{R}^3 \setminus Z(g)$ that is intersected by $\gamma_f$ corresponds to at least one connected component of $\gamma_f \setminus \gamma_g \subset h$. Thus, the number of cells intersected by $h$ is upper bounded by the number of connected components of $\gamma_f \setminus \gamma_g$. Each such connected component is either a full connected component of $\gamma_f$ or an open portion of $\gamma_f$ whose closure meets $\gamma_g$. According to Theorem 2.6, $\gamma_f$ has $O(D^2) = O(DE)$ connected components. For every irreducible component of $\gamma_f$ that is also contained in $\gamma_g$, we can remove the component from $\gamma_f$. We denote by $\gamma_f' \subset h$ the curve that is obtained by removing these irreducible components from $\gamma_f$. Notice that $\gamma_g$ and $\gamma_f'$ do not have common components. By Theorem 2.5, the number of intersection points between $\gamma_f'$ and $\gamma_g$ is $O(DE)$. This implies that the number of connected components of $Z(\tilde{f}) \setminus Z(\tilde{g})$ is $O(DE)$, concluding the proof of the lemma. □

**Lemma 2.11.** *Let $f, g \in \mathbb{R}[x, y, z]$ be co-prime polynomials of degrees $D$ and $E$, respectively, with $D \leq E$. Let $\sigma$ be a sphere in $\mathbb{R}^3$, such that $\sigma \not\subseteq Z(f)$. Then the number of cells of $\mathbb{R}^3 \setminus Z(g)$ intersected by $\sigma \cap Z(f)$ is at most $c'DE$ for some absolute constant $c'$.*

*Proof.* As in the proof of Lemma 2.9, we set

$$\tilde{f} = (u^2 + v^2 + 1)^D \cdot (f \circ \psi) \quad \text{and} \quad \tilde{g} = (u^2 + v^2 + 1)^E \cdot (g \circ \psi),$$

where $\psi$ maps $\mathbb{R}^2$ to the sphere $\sigma$. As before, $\tilde{f}$ is of a polynomial of degree at most $2D$ and $\tilde{g}$ is a polynomial of degree at most $2E$. The zero set $Z(\tilde{g}) \subset \mathbb{R}^2$ is the image of a mapping of $Z(g) \cap \sigma$ to the plane (i.e., each three-dimensional cell of $\mathbb{R}^3 \setminus Z(g)$ that intersects $\sigma$ corresponds to a set of cells in $\mathbb{R}^2$). Every cell of $\mathbb{R}^3 \setminus Z(g)$ that is intersected by $\sigma \cap Z(f)$ corresponds to at least one connected component of $\big(\sigma \cap Z(f)\big) \setminus Z(g)$. Thus, the number of cells intersected by $\sigma$ is upper bounded by the number of connected components of $Z(\tilde{f}) \setminus Z(\tilde{g})$. By repeating the corresponding analysis in the proof of Lemma 2.10, we obtain that the number of such components is $O(DE)$. □

**Lemma 2.12.** *Let $f, g \in \mathbb{R}[x, y, z]$ be co-prime polynomials of degrees $D$ and $E$, respectively, with $D \leq E$. Let $\mathcal{S}$ be a $k$-non-degenerate set of $n$ spheres and $\mathcal{P}$ be a set of $m$ points, both in $\mathbb{R}^3$, such that $\mathcal{P} \subset Z(f) \cap Z(g)$. Then*

$$I(\mathcal{P}, \mathcal{S}) = O(nDE + mkDE).$$

*Proof.* The number of spheres that are contained in $Z(f)$ or in $Z(g)$ is $O(D + E)$, and thus such spheres participate in $O(m(D + E)) = O(mE)$ incidences. In the rest of the analysis we consider only spheres that are contained neither in $Z(f)$ nor in $Z(g)$.

As in Lemma 2.11, given a sphere $\sigma \in \mathcal{S}$, we set

$$\tilde{f}_\sigma = (u^2 + v^2 + 1)^D \cdot (f \circ \psi) \quad \text{and} \quad \tilde{g}_\sigma = (u^2 + v^2 + 1)^E \cdot (g \circ \psi),$$

where $\psi$ maps $\mathbb{R}^2$ to the sphere $\sigma$. As before, $\tilde{f}_\sigma$ is of degree at most $2D$ and $\tilde{g}_\sigma$ is of degree at most $2E$.

We first consider incidences between a sphere $\sigma \in \mathcal{S}$ and points of $\mathcal{P}$ that are not on a one dimensional component of $Z(f) \cap Z(g) \cap \sigma$. The number of zero-dimensional components of $Z(\tilde{f}_\sigma) \cap Z(\tilde{g}_\sigma)$ is an upper bound on the number of such incidences with $\sigma$. By removing from $\tilde{f}_\sigma$ and $\tilde{g}_\sigma$ factors that are common to both of these polynomials, Theorem 2.5 implies that the number of such points is $O(DE)$. Now by Theorem 2.4, the removed common components contain $O(D^2) = O(DE)$ connected components. Thus, the overall number of incidences that are not on a one dimensional component of $Z(f) \cap Z(g) \cap \sigma$ is $O(nDE)$.

We next consider incidences between a sphere $\sigma \in \mathcal{S}$ and points of $\mathcal{P}$ that are on a one dimensional component of $Z(f) \cap Z(g) \cap \sigma$. We consider a point $p \in \mathcal{P}$, and bound the number of such incidences that $p$ can participate in. A curve that is contained in more than one sphere of $\mathcal{S}$ is a circle. Since $\mathcal{S}$ is $k$-non-degenerate, at most $k-1$ spheres of $\mathcal{S}$ can contain such a circle. In the following paragraph, we prove that $Z(f) \cap Z(g)$ contains $O(DE)$ irreducible components. Since each such curve is contained in at most $k-1$ spheres of $\mathcal{S}$, then $p$ is incident to $O(kDE)$ spheres. Thus, there are $O(mkDE)$ incidences between a sphere $\sigma \in \mathcal{S}$ and points of $\mathcal{P}$ that are on a one dimensional component of $Z(f) \cap Z(g) \cap \sigma$.

It remains to prove that $Z(f) \cap Z(g)$ contains $O(DE)$ irreducible components. We project $Z(f) \cap Z(g)$ on a generic plane by using a *resultant* (for some basic details about resultants, see [38] and [24, Sections 3.5-3.6]). The projection is a two-dimensional curve of degree $O(DE)$, and thus contains $O(DE)$ irreducible components. This in turns implies that the original curve contains $O(DE)$ irreducible components. $\square$

**Corollary 2.13.** *Let $f, g \in \mathbb{R}[x, y, z]$ be co-prime polynomials of degrees $D$ and $E$, respectively, such that $D \leq E$. Let $\Pi$ be a $k$-non-degenerate set of $n$ planes and let $\mathcal{P}$*

*be a set of $m$ points, both in $\mathbb{R}^3$, such that $\mathcal{P} \subset Z(f) \cap Z(g)$. Then*

$$I(\mathcal{P}, \Pi) = O(nDE + mkDE).$$

*Proof.* This follows immediately by applying a generic inversion transformation, transforming the planes into spheres, and then applying Lemma 2.12 (such a transformation at most doubles the degrees of $f$ and $g$). A different approach would be to repeat the proof of Lemma 2.12, but without using the inverse stereographic projection, and with lines instead of circles. □

## 2.4 Lower Bounds

The following lower bound is a simple variant of the lower bound constructions for the planar Szemerédi-Trotter theorem (e.g., see [58]).

**Theorem 2.14.** *(i) For any $n, m$, and $1 < k < n$, there exist a set $\mathcal{P}$ of $m$ points and a $k$-non-degenerate set $\Pi$ of $n$ planes, both in $\mathbb{R}^3$, such that $I(\mathcal{P}, \Pi) = \Theta(m^{2/3}n^{2/3}k^{1/3} + n + mk)$.*
*(ii) For any $n, m$, and $1 < k < n$, there exist a set $\mathcal{P}$ of $m$ points and a $k$-non-degenerate set $\mathcal{S}$ of $n$ spheres, both in $\mathbb{R}^3$, such that $I(\mathcal{P}, \mathcal{S}) = \Theta(m^{2/3}n^{2/3}k^{1/3} + n + mk))$.*

*Proof.* We consider a planar set of $m$ points and $\ell = n/(k-1)$ lines with $\Theta(\ell^{2/3}m^{2/3} + \ell + m)$ point-line incidences (e.g., see [58, Section 2]). We embed this plane as the $xy$-plane in $\mathbb{R}^3$, and pass $k-1$ distinct (and otherwise generic) planes through each line of the planar configuration. Thus, the total number of planes is $n = (k-1)\ell$, and the number of point-plane incidences is

$$\begin{aligned}(k-1) \cdot \Theta(\ell^{2/3}m^{2/3} + \ell + n) &= \Theta\left(k\left((n/k)^{2/3}m^{2/3} + n/k + m\right)\right) \\ &= \Theta(m^{2/3}n^{2/3}k^{1/3} + n + mk).\end{aligned}$$

This concludes the proof of part (i). Part (ii) is immediately obtained by applying an inversion transformation. □

Theorem 2.14 implies that Theorem 1.7 is tight when $n = O(\sqrt{m}k^2)$ and when $n = \Omega(m^2k)$.

# Chapter 3

# Ordinary Lines in Complex Space

In this section we establish bounds on the number of ordinary lines determined by points in complex space, i.e., we prove Theorems 1.15, 1.17 and 1.18. The starting point for the proofs of these theorems is the method developed in [12, 29] 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 Section 3.1 to an overview of the general framework (starting with [29]), outlining the places where new ideas come into play. In Section 3.2 we develop the necessary machinery for matrix scaling and Latin squares. In Section 3.3, we prove some key lemmas that will be used in the proofs of our main results. Section 3.4 gives the proof of Theorems 1.15 and 1.18, which are considerably simpler than that of Theorem 1.17. Section 3.5 is devoted to developing additional machinery. We present the proof of Theorem 1.17 in Section 3.6.

## 3.1 Proof Overview

Let $\mathcal{V} = \{v_1, \ldots, v_n\}$ 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 (complex affine) space spanned by the point set in $\mathbb{C}^d$ can be seen to be equal to $\text{rank}(V) - 1$. We would now like to argue that too many collinearities in $\mathcal{V}$ (or too few ordinary lines) imply that all (or almost all) points of $\mathcal{V}$ must be contained in a low-dimensional (complex) affine subspace, i.e., $\text{rank}(V)$ is small. To do this, we construct a matrix $A$, encoding the dependencies in $\mathcal{V}$, such that

$AV = 0$. Then we must have

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

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

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 nonzero 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 [29] now proceeds to establish 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 [29] is actually proved. To provide some intuition, suppose that $A$ is a matrix with 0/1 entries. To bound the rank of $A$, we can consider the matrix $M = A^*A$ (where $A^*$ is the matrix $A$ conjugated and transposed) and note that $\mathrm{rank}(M) = \mathrm{rank}(A)$. 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 this end, we rely on *matrix scaling* techniques. We are allowed to multiply each row and each column of $A$ by a nonzero 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 [29] 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 passes through any of the points. We argue in Lemma 3.15 that such a submatrix implies the existence of many ordinary lines. In fact, the conclusion is slightly more delicate: Either there exist many ordinary lines (in which case we are done) or there exists a point incident to many ordinary lines (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 1.17. Here, our goal is to improve on the rank bound of [29] 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 [29] 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. Furthermore, the number of such cancellations increase the more points we have on a single line.

To get a rough idea, we focus on a set of points in real space. 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 both 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 opposite. 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 still obtain a better bound. This again translates into a better rank bound, see Lemma 3.23. To formalize the previous example, let $v_1, v_2, v_3, v_4$ be collinear points in $\mathbb{R}^d$. Then there exist $r, s, t \in \mathbb{R}$ such that

$$
\begin{aligned}
& r \cdot v_1 + (1-r) \cdot v_2 - v_3 = 0, \\
& s \cdot v_1 + (1-s) \cdot v_2 - v_4 = 0, \\
\text{and} \quad & t \cdot v_1 + (1-t) \cdot v_3 - v_4 = 0.
\end{aligned}
$$

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 $r(1-r) = s(1-s)$, which cannot happen because they have opposite signs. This phenomena is captured in Lemma 3.20, 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 quadruple (see Item 4 of Lemma 3.21).

## 3.2 Preliminaries

### 3.2.1 Matrix scaling and rank bounds

One of the main ingredients in our proof is a technique for bounding the rank of design matrices. These ideas were first used for incidence-type problems in [12] and improved upon in [29].

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 $A$. For two complex vectors $u, v \in \mathbb{C}^d$, we denote their inner product by $\langle u, v\rangle = \sum_{i=1}^{d} u_i \cdot \overline{v_i}$.

We now introduce the notion of matrix scaling, which is central to obtaining the rank bounds.

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

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*

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

We are interested in scalings of a matrix that normalize the row and column sums. The following theorem (see [60]) shows that satisfying Property $S$ is a sufficient condition for such a scaling to exist.

**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.*

We may assume that the sum of every row of the scaling $A'$ is exactly $1 + \epsilon$. Otherwise, we may scale each rows to make the sum $1 + \epsilon$, and note that the column sums can only increase.

The following Corollary to Theorem 3.3 appeared in [12].

**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 $i \in [m]$*

$$\sum_{j\in[n]} \left|A'_{ij}\right|^2 \leq 1 + \epsilon,$$

*and for every $j \in [n]$*

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

*Moreover, the scaling coefficients are all positive real numbers.*

Corollary 3.4 follows from 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 $\sum_{j\in[n]} \left|A'_{ij}\right|^2 = 1+\epsilon$.

To bound the rank of a matrix $A$, we will bound the rank of $M = A'^*A'$, where $A'$ is a scaling of $A$. Clearly $\text{rank}(A) = \text{rank}(A') = \text{rank}(M)$. 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 [6]) and appeared in this form in [29].

**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*

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

The matrix scaling theorem allows us to control the $\ell_2$ norms of the columns and rows of $A$, which in turn allow us to bound the sums of squares of entries of $M$. For this, we use a variation of a lemma from [29]. While the proof idea is the same, our proof requires a somewhat more careful analysis. First, we need some definitions.

**Definition 3.6.** *For an $m \times n$ matrix $A$ with complex entries, define:*

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

*and*

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

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

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

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

*Proof.* Note that

$$\begin{aligned}\sum_{i\neq j}|M_{ij}|^2 &= \sum_{i\neq j}|\langle C_i, C_j\rangle|^2\\ &= \sum_{i\neq j}\left|\sum_{k=1}^{m} A_{ki}\overline{A_{kj}}\right|^2.\end{aligned}$$

Since the supports of any two columns of $A$ intersect in exactly $t$ locations, the Cauchy-Schwarz inequality implies that $\left|\sum_{k=1}^{m} A_{ki}\overline{A_{kj}}\right|^2 \leq t\sum_{k=1}^{m}|A_{ki}|^2|A_{kj}|^2$. Our approach requires somewhat more careful analysis, so we use the following equality:

$$\begin{aligned}\sum_{i\neq j}\left|\sum_{k=1}^{m} A_{ki}\overline{A_{kj}}\right|^2 &= \sum_{i\neq j}\left(t\sum_{k=1}^{m}|A_{ki}|^2|A_{kj}|^2 - \sum_{k<k'}\left|A_{ki}\overline{A_{kj}} - A_{k'i}\overline{A_{k'j}}\right|^2\right)\\ &= t\sum_{i\neq j}\sum_{k=1}^{m}|A_{ki}|^2|A_{kj}|^2 - D(A)\\ &= t\sum_{k=1}^{m}\left(\sum_{i=1}^{n}|A_{ki}|^2\right)^2 - t\sum_{k=1}^{m}\left(\sum_{i=1}^{n}|A_{ki}|^4\right) - D(A).\end{aligned}$$

Since there are $q$ nonzero entries for every row of $A$, the Cauchy-Schwarz inequality implies that $\sum_{i=1}^{n}|A_{ki}|^4 \geq \frac{1}{q}\left(\sum_{i=1}^{n}|A_{ki}|^2\right)^2$. Again, this turns out to be insufficient for our purpose and we consider the equality $\sum_{i=1}^{n}|A_{ki}|^4 = \frac{1}{q}\left(\sum_{i=1}^{n}|A_{ki}|^2\right)^2 + \sum_{i<j}\left(|A_{ki}|^2 - |A_{kj}|^2\right)^2$, which gives:

$$\begin{aligned}\sum_{i\neq j}|M_{ij}|^2 &= \left(1-\frac{1}{q}\right)t\sum_{k=1}^{m}\left(\sum_{i=1}^{n}|A_{ki}|^2\right)^2 - \frac{t}{q}\sum_{k=1}^{m}\sum_{i<j}\left(|A_{ki}|^2 - |A_{kj}|^2\right)^2 - D(A)\\ &= \left(1-\frac{1}{q}\right)t\sum_{k=1}^{m}\left(\sum_{i=1}^{n}|A_{ki}|^2\right)^2 - \frac{t}{q}E(A) - D(A)\\ &= \left(1-\frac{1}{q}\right)tm\alpha^4 - \left(D(A) + \frac{t}{q}E(A)\right).\end{aligned}$$

□

From Lemma 3.7, we obtain the following easy corollary.

**Corollary 3.8.** *Let $A$ be an $m\times n$ matrix over $\mathbb{C}$ and $M = A^*A$. 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 $q$. Then*

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

### 3.2.2 Latin squares

Latin squares were used in the construction of design-matrices in both [29] and [12]. In our proof they play a more central role and we use 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 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 [42]. 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 [22] 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 is a strengthening of a lemma from [12].

**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, and $T$ 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. That is, $L_{ij} \notin \{i, j\}$ for $i \neq j \in [r]$, giving Property 1.

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, $i$ 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 $(*, j, i), (j, *, i), (*, i, j)$ and $(i, *, j)$ appears exactly once, where $*$ is some other element of $[r]$. This shows that every pair appears in exactly six triples.

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

## 3.3 The Dependency Matrix

Let $\mathcal{V} = \{v_1, \ldots, v_n\}$ be a set of $n$ points in $\mathbb{C}^d$. We will use $\dim(\mathcal{V})$ to denote the dimension of the linear span of $\mathcal{V}$ and by affine-dim$(\mathcal{V})$ the dimension of the affine span of $\mathcal{V}$ (i.e., the minimum $r$ such that points of $\mathcal{V}$ are contained in a shift of a linear subspace of dimension $r$). We projectivize $\mathbb{C}^d$ and consider the set of vectors $\mathcal{V}' = \{v'_1, \ldots, v'_n\}$, 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

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

We now construct a matrix $A$, which we refer to as the *dependency matrix* of $\mathcal{V}$. Note that the construction we give here is preliminary, but suffices to prove Theorems 1.15 and 1.18. A refined construction is given in Section 3.5, where we select the triples more carefully. The rows of the matrix will consist of linear dependency coefficients, which we define below.

**Definition 3.12** (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 $\mathbb{C}^{d+1}$. There exist nonzero 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 $v_1, v_2, v_3$. Note that the coefficients are determined up to scaling by a complex number. Throughout our proof, the specific choice of coefficients does not matter, so we fix a canonical choice by setting $a_3 = 1$.*

**Definition 3.13** (Dependency Matrix)**.** *For every line $l \in \mathcal{L}_{\geq 3}(\mathcal{V})$, let $\mathcal{V}_l$ denote the points lying on $l$. Then $|\mathcal{V}_l| \geq 3$ and we assign each line a triple system $T_l \subseteq \mathcal{V}_l^3$, the existence of which is guaranteed by Lemma 3.11 (with $r = |\mathcal{V}_l|$). Let $A$ be the $m \times n$ matrix obtained by going over every line $l \in \mathcal{L}_{\geq 3}$ and for each triple $(i, j, k) \in T_l$, adding as a row of $A$ the vector with three nonzero 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 $\mathcal{V}$.*

Note that we have $AV = 0$. Every row of $A$ has exactly three nonzero entries. By Property 2 of Lemma 3.11, the supports of any two distinct columns intersect in exactly six entries when the two corresponding points lie on a special line[1], and 0 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 nonzero entries in columns $i$ and $j$. The number of times a pair appears is the number of rows with nonzero 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, $t_2(\mathcal{V})$ of which determine ordinary lines, i.e., the number of pairs appearing in $A$ is $\binom{n}{2} - t_2$. Since each pair appears six times, the total number of times these pairs appear is $6\left(\binom{n}{2} - t_2\right)$. Note that each row gives three distinct pairs of points. It follows that the number of rows of $A$ is

$$m = 6\left(\binom{n}{2} - t_2\right)/3 = n^2 - n - 2t_2(\mathcal{V}). \tag{3.1}$$

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: first when $A$ satisfies Property $S$ and second when it does not. We now prove some facts dealing with the two cases. The following lemma applies to the first case.

**Lemma 3.14.** *Let $\mathcal{V}$ be a set of $n$ points affinely spanning $\mathbb{C}^d$, $d \geq 3$, and let $A$ be the dependency matrix for $\mathcal{V}$. If $A$ satisfies Property $S$, then*

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

*Proof.* Fix $\epsilon > 0$. Since $A$ satisfies Property $S$, Corollary 3.4 shows that there is a scaling $A'$ such that the $\ell_2$ norm of each row is at most $\sqrt{1+\epsilon}$ and the $\ell_2$ norm of each

[1] Note that while the triple system $T_l$ consists of ordered triples, the supports of the rows of $A$ are unordered.

column is at least $\sqrt{\frac{m}{n}-\epsilon}$. Let $M := A'^*A'$. Then $M_{ii} \geq \frac{m}{n}-\epsilon$ 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 implies that $\sum_{i\neq j} |M_{ij}|^2 \leq 4m(1+\epsilon)^2$. Lemma 3.5 implies that,

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

Taking the limit as $\epsilon$ approaches 0, and combining with equation (3.1) gives

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

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

$$\begin{aligned} d+1 &\leq \frac{4n^2}{n^2+3n-2t_2(\mathcal{V})} \\ \text{i.e.,}\quad t_2(\mathcal{V}) &\geq \frac{(d-3)}{2(d+1)}n^2 + \frac{3}{2}n.\end{aligned}$$

□

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

**Lemma 3.15.** *Let $\mathcal{V}$ be a set of $n$ points in $\mathbb{C}^d$, and let $A$ be the dependency matrix for $\mathcal{V}$. Suppose that $A$ does not satisfy Property $S$. Then, for every integer $b^*$, $1 < b^* < 2n/3$, one of the following must hold:*

1. *There exists a point $v \in \mathcal{V}$ contained in at least $\frac{2}{3}(n+1) - b^*$ ordinary lines;*

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

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

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

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 $\mathcal{V}_1 = \{v_1, \ldots, v_b\} \subset \mathcal{V}$ and $\mathcal{V}_2 = \mathcal{V} \setminus \mathcal{V}_1$. We may assume that $U$ is maximal, so every row in the submatrix $X \times W$ has at least one nonzero entry. We consider two cases: when $b < b^*$, and when $b \geq b^*$.

**Case 1** ($b < b^*$). Partition the rows of $X$ into three sets: Let $X_1, X_2$ and $X_3$ be rows with one, two and three nonzero entries in columns of $W$ respectively. We will obtain a lower bound on the number of ordinary lines containing exactly one point in $\mathcal{V}_1$ and one point in $\mathcal{V}_2$ by bounding the number of pairs $\{v, w\} \in \mathcal{V}_1 \times \mathcal{V}_2$ 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 $X_1$ contributes two pairs of points $\{v, u\}, \{v, w\} \in \mathcal{V}_1 \times \mathcal{V}_2$ that lie on a special line. Each row of $X_2$ contributes two pairs of points $\{v, w\}, \{u, w\} \in \mathcal{V}_1 \times \mathcal{V}_2$ that lie on a special line. Rows of $X_3$ have all zero entries in the submatrix supported on $X \times Y$ and do 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 $\mathcal{V}_1$ and one point in $\mathcal{V} \setminus \mathcal{V}_1$ is $\frac{2|X_1|+2|X_2|}{6} \leq \frac{2|X|}{6}$. Hence, the number of ordinary lines containing exactly one of $v_1, \ldots, v_b$ must be at least $b(n - b) - \frac{|X|}{3}$.

Now recall that

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

Substituting $m \leq n^2 - n$, from equation (3.1). gives

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

This shows that the number of ordinary lines containing exactly one point in $\mathcal{V}_1$ is at least

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

It follows that there exists $v \in \mathcal{V}_1$ 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 $t_2(\mathcal{V})$ by counting the number of nonzero 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 $\mathcal{V}_2$, 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 nonzero entries, i.e., the number of pairs of entries equals $3a$. It follows that

$$3a \leq 6\binom{n-b}{2}. \tag{3.2}$$

Recall equation (3.1) and that $\frac{a}{m} + \frac{b}{n} > 1$, which implies

$$a > m\left(1 - \frac{b}{n}\right) = \left(n^2 - n - 2t_2(\mathcal{V})\right)\left(1 - \frac{b}{n}\right). \tag{3.3}$$

Combining (3.2) and (3.3), we obtain

$$\left(n^2 - n - 2t_2(\mathcal{V})\right)\left(1 - \frac{b}{n}\right) < 2\binom{n-b}{2}.$$

Finally, solving for $t_2(\mathcal{V})$ yields

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

□

## 3.4 Proofs of Theorems 1.15 and 1.18

The proofs of both Theorems 1.15 and 1.18 rely on Lemmas 3.14 and 3.15. 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 where a point incident to many ordinary lines is removed, and then the same argument is applied to the remaining points.

### Proof of Theorem 1.15

The following corollary follows easily from Lemma 3.14 and Lemma 3.15.

**Corollary 3.16.** *Let $\mathcal{V}$ be a set of $n$ points in $\mathbb{C}^d$ not contained in a plane. Then one of the following holds:*

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

2. $t_2(\mathcal{V}) \geq \frac{3}{2}n$*.*

*Proof.* Let $A$ be the dependency matrix for $\mathcal{V}$. If $A$ satisfies Property $S$, then we are done by Lemma 3.14. Otherwise, let $b^* = 3$, and note that Lemma 3.15 gives us the statement of the corollary when $n \geq 5$. The statement holds trivially when $n < 5$. □

We are now ready to prove Theorem 1.15. For convenience, we state the theorem again.

**Theorem 1.15.** *Let $\mathcal{V}$ be a set of $n \geq 24$ points in $\mathbb{C}^3$ not contained in a (complex) plane. Then $\mathcal{V}$ determines at least $\frac{3}{2}n$ ordinary lines, unless $n-1$ points are on a plane in which case there are at least $n-1$ ordinary lines.*

*Proof.* We may assume, by Corollary 3.16, that there exists a point $v_1$ incident to at least $\frac{1}{3}(2n-7)$ ordinary lines and hence contained in at most $\frac{1}{6}(n+4)$ 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 containing $v_1$. We note here that this is the only case where there are fewer then $\frac{3}{2}n$ ordinary lines.

Suppose now that $\mathcal{V}_1$ is not planar. Again, by Corollary 3.16, there are either $\frac{3}{2}(n-1)$ ordinary lines in $\mathcal{V}_1$ or there exists a point $v_2 \in \mathcal{V}_1$ incident to at least $\frac{2}{3}(n-1) - \frac{7}{3} = \frac{1}{3}(2n-9)$ ordinary lines. In the former case, there exist $\frac{3}{2}(n-1)$ ordinary lines in $\mathcal{V}_1$, at most $\frac{1}{6}(n+4)$ of which could contain $v_1$. This shows that the total number of ordinary lines in $\mathcal{V}$ satisfies

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

If $n \geq 9$, then $t_2(\mathcal{V}) \geq \frac{3}{2}n$.

In the latter case there exists a point $v_2 \in \mathcal{V}_1$ incident to at least $\frac{1}{3}(2n-9)$ ordinary lines in $\mathcal{V}_1$. Note that at most one of these could contain $v_1$, so at least $\frac{1}{3}(2n-7) + \frac{1}{3}(2n-9) - 1 = \frac{1}{3}(4n-19)$ ordinary lines containing either $v_1$ or $v_2$. The number of special lines containing either $v_1$ or $v_2$ is at most $\frac{1}{6}(n+4) + \frac{1}{6}(n+3) = \frac{1}{6}(2n+7)$.

Let $\mathcal{V}_2 = \mathcal{V}_1 \setminus \{v_2\}$. If $\mathcal{V}_2$ 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 in $\mathcal{V}$. It follows that when $n \geq 12$, $t_2(\mathcal{V}) \geq \frac{3}{2}n$.

Otherwise, $\mathcal{V}_2$ is not contained in a plane, and again, Corollary 3.16 gives two cases. If there are $\frac{3}{2}(n-2)$ ordinary lines in $\mathcal{V}_2$, then the total number of ordinary lines is

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

If $n \geq 11$, then $t_2(\mathcal{V}) \geq \frac{3}{2}n$.

Finally, if none of the above hold, there exists a point $v_3$ incident to at least $\frac{2}{3}(n-2) - \frac{7}{3}$ ordinary lines. At most two of these could contain one of $v_1$ or $v_2$, so there are $\frac{2}{3}(n-2) - \frac{7}{3} - 2 = \frac{1}{3}(2n-17)$ ordinary lines through $v_3$ in $\mathcal{V}$. Summing up the number of lines containing one of $v_1, v_2$ and $v_3$, we obtain

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

If $n \geq 24$, $t_2(\mathcal{V}) \geq \frac{3}{2}n$ completing the proof. □

## Proof of Theorem 1.18

The following corollary follows from Lemma 3.14 and Lemma 3.15.

**Corollary 3.17.** *There exists a positive integer $n_0$ such that the following holds. Let $\mathcal{V}$ be a set of $n \geq n_0$ points in $\mathbb{C}^d$ not contained in a three-dimensional affine subspace. Then one of the following must hold:*

1. *There exists a point incident to at least $\frac{n}{2}$ ordinary lines,*

2. *$t_2(\mathcal{V}) \geq \frac{1}{12}n^2$.*

*Proof.* Let $A$ be the dependency matrix of $\mathcal{V}$. If $A$ satisfies Property $S$, then we are done by Lemma 3.14. Otherwise, let $b^* = n/6$. Now, by Lemma 3.15, either the number of ordinary lines

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

or there exists a point $v \in \mathcal{V}$, such that the number of ordinary lines containing $v$ is at least

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

□

We are now ready to prove Theorem 1.18. For convenience, we state the theorem again.

**Theorem 1.18.** *There exists a positive integer $n_0$ such that the following holds. Let $\mathcal{V}$ be a set of $n \geq n_0$ points in $\mathbb{C}^4$ with at most $\frac{1}{2}n$ points contained in any three-dimensional affine subspace. Then*

$$t_2(\mathcal{V}) \geq \frac{1}{16}n^2.$$

*Proof.* The basic idea of the proof uses the following algorithm: We apply Corollary 3.17 and 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 are unable to find such a point, in which case Corollary 3.17 guarantees a large number of ordinary lines, or when we have accumulated enough ordinary lines. Consider the following algorithm:

Let $\mathcal{V}_0 := \mathcal{V}$ and $j = 0$.

1. If $\mathcal{V}_j$ satisfies case (2) of Corollary 3.17, then stop.
2. Otherwise, there must exist a point $v_{j+1} \in \mathcal{V}_j$ incident to at least $\frac{n-j}{2}$ ordinary lines. Let $\mathcal{V}_{j+1} = \mathcal{V}_j \setminus \{v_{j+1}\}$.
3. Set $j = j + 1$. If $j = n/2$, then stop. Otherwise go to Step 1.

Note that since no three-dimensional subspace contains more than $n/2$ points, the algorithm will never stop because the configuration becomes three-dimensional; that is, we can use Corollary 3.17 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 the algorithm stops because $\mathcal{V}_j$ satisfies case (2) of Corollary 3.17 for some $1 \leq j < n/2$. From case (2) of Corollary 3.17, we have

$$t_2(\mathcal{V}_j) \geq \frac{(n-j)^2}{12}. \tag{3.4}$$

On the other hand, each pruned point $v_i$, $1 \leq i \leq j$, incident to at least $\frac{n-i+1}{2} > \frac{n-i}{2}$ ordinary lines determined by $\mathcal{V}_{i-1}$, and hence contained in at most $(n - i - \frac{n-i+1}{2})/2 <$

$\frac{n-i}{4}$ special lines. Note that an ordinary line in $\mathcal{V}_i$ might not be ordinary in $\mathcal{V}_{i-1}$ if it contains $v_i$. Thus, in order to lower bound the total number of ordinary lines in $\mathcal{V}$, we can sum over the number of ordinary lines contributed by each of the pruned points $v_i$, $1 \le i \le j$, and subtract from the count the number of potential lines that could contain $v_i$. The number of ordinary lines in $\mathcal{V}$ contributed by the pruned points is at least

$$\sum_{i=1}^{j}\left(\frac{n-i}{2}-\frac{n-i}{4}\right)=\frac{1}{4}\sum_{i=1}^{j}(n-i)\,. \tag{3.5}$$

Combining (3.4) and (3.5), we obtain

$$\begin{aligned} t_2(\mathcal{V}) &\ge \frac{1}{12}(n-j)^2+\frac{1}{4}\sum_{i=1}^{j}(n-i) \\ &= \frac{1}{24}\left(-j^2+j(2n-3)+2n^2\right). \end{aligned}$$

This is an increasing function for $j < n-1$, implying that

$$t_2(\mathcal{V}) \ge \frac{n^2}{12}.$$

We now consider the case when the algorithm stops because $j = n/2$. Note that at this point, we will have pruned exactly $j$ points. Each pruned point $v_i$, $1 \le i \le j$, is incident to $\frac{n-i+1}{2} > \frac{n-i}{2}$ ordinary lines determined by $\mathcal{V}_{i-1}$. The only way such an ordinary line is not ordinary in $\mathcal{V}$ is that 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 $\mathcal{V}$ must satisfy

$$t_2(\mathcal{V}) \ge \sum_{i=1}^{j}\frac{n-i}{2}-\sum_{i=1}^{j} i=\frac{1}{2}\sum_{i=1}^{j}(n-3i)=\frac{n^2-6n}{16}.$$

It follows that for $n$ large enough,

$$t_2(\mathcal{V}) \ge \frac{1}{16}n^2.$$

□

## 3.5 A Refined Dependency Matrix Construction

In this section we give a more careful construction for the dependency matrix of a point set $\mathcal{V}$. Recall from Definition 3.13 that we defined the dependency matrix to contain a row for each collinear triple from a triple system constructed on each special line. The goal was to avoid having 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 3.24) we will give a construction of a dependency matrix that will have an additional property (captured in Item 4 of Lemma 3.21) which is used to obtain cancellations in the diagonal dominant argument (as was outlined in the proof overview).

We denote the argument of a complex number $z$ by $\arg(z)$. We use the convention that for every complex number $z$, $\arg(z) \in (-\pi, \pi]$.

**Definition 3.18** (angle between two complex numbers)**.** *We define the* angle between two complex numbers $a$ and $b$ *to be the the absolute value of the argument of* $a\overline{b}$*, denoted by* $\left|arg\left(a\overline{b}\right)\right|$*. Note that the angle between* $a$ *and* $b$ *equals the angle between* $b$ *and* $a$*.*

**Definition 3.19** (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 quadruples on the line.

**Lemma 3.20.** *Let* $v_1, v_2, v_3, v_4$ *be four distinct collinear points in* $\mathbb{C}^d$*. Then at least one of the following must hold:*

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

$$a_1 v_1' + a_2 v_2' + a_3 v_3' = 0 \tag{3.6}$$

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

$$b_1 v_1' + b_2 v_2' + b_4 v_4' = 0. \tag{3.7}$$

We may assume, without loss of generality, that $a_3 = b_4 = 1$. Equations (3.6) and (3.7) show 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 equations (3.6) and (3.7), we obtain:

$$(b_2 a_1 - b_1 a_2) v_1' + b_2 v_3' - a_2 v_4' = 0. \tag{3.8}$$

Equation (3.8) implies that $C_{(1,3)}(4) = \frac{(b_2 a_1 - b_1 a_2)\overline{b_2}}{|b_2 a_1 - b_1 a_2||b_2|}$ and $C_{(1,4)}(3) = -\frac{(b_2 a_1 - b_1 a_2)\overline{a_2}}{|b_2 a_1 - b_1 a_2||a_2|}$. It follows that the angle between $C_{(1,2)}(3)$ and $C_{(1,2)}(4)$ is

$$\left| \arg \left( \frac{a_1 \overline{a_2}}{|a_1||a_2|} \frac{\overline{b_1} b_2}{|b_1||b_2|} \right) \right| = \left| \arg \left( a_1 \overline{a_2} \overline{b_1} b_2 \right) \right|. \tag{3.9}$$

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

$$\left| \arg \left( \frac{(b_2 a_1 - b_1 a_2)\overline{b_2}}{|b_2 a_1 - b_1 a_2||b_2|} \frac{\overline{a_1}}{|a_1|} \right) \right| = \left| \arg \left( \overline{a_1} \overline{b_2} (b_2 a_1 - b_1 a_2) \right) \right|, \tag{3.10}$$

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

$$\left| \arg \left( -\frac{b_1}{|b_1|} \frac{\overline{(b_2 a_1 - b_1 a_2)} a_2}{|b_2 a_1 - b_1 a_2||a_2|} \right) \right| = \left| \arg \left( -b_1 a_2 \overline{(b_2 a_1 - b_1 a_2)} \right) \right|. \tag{3.11}$$

Note that the product of expressions inside the arg functions in (3.9), (3.10) and (3.11) is a negative real number, and so the sum of (3.9), (3.10) and (3.11) must be $\pi$. It follows that one of the angles must be at least $\pi/3$. □

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

**Lemma 3.21.** *Let $l$ be a line in $\mathbb{C}^d$ and $\mathcal{V}_l = \{v_1, \ldots v_r\}$ denote points on $l$ with $r \geq 3$ Then there exists an $m \times r$ matrix $A = A(l)$, with $m = r^2 - r$, such that the following hold:*

1. *$AV_l = 0$, where $V_l$ is the $r \times (d+1)$ matrix whose $i^{th}$ row is the vector $(v_i, 1)$.*

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. *Let $R_k$ denote the $k^{th}$ row of $A$ and suppose $supp(R_k) = \{i, j, s\}$. If $r \geq 4$ then for at least $1/3$ of choices of $k \in [m]$, there exists $k' \in [m]$ such that $supp(R_{k'}) = \{i, j, t\}$ $(t \neq s)$ 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 $[r]^3$. Let $\Sigma$ be the set of all bijective maps from $[r]$ to the points $\mathcal{V}_l$. For every bijective map $\sigma \in \Sigma$, construct a matrix $A_\sigma$ in the following manner: Let $T_l$ be the triple system on $\mathcal{V}_l^3$ induced by composing $\sigma$ 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 $A_\sigma V_l = 0$, and Property 1 is satisfied. Properties 2 and 3 follow from properties of the triple system from Lemma 3.11.

We use a probabilistic argument to show that there exists a matrix $A$ that has Property 4. Let $\sigma \in \Sigma$ be a uniformly random element, and 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 $\{v_i, v_j, v_s, v_t\}$ 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 3.20, 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

$$\Pr(R_k \text{ satisfies Property 4}) \geq 1/3.$$

Define the random variable $X$ to be the number of rows satisfying Property 4, and note that we have

$$\mathbb{E}[X] \geq (r^2 - r)\frac{1}{3}.$$

It follows that there exists a matrix $A$ in which at least 1/3 of the rows satisfy Property 4.

□

To bound the sum the off diagonal entries of $M$, we use the following notion of balanced rows. The main idea here is that if there are many rows that are balanced then we can obtain an improved bound from cancellations that show up via the different angles. On the other hand, if many rows are unbalanced then we obtain a better bound by showing that the sum of entries squares must be far from the bound implied by Cauchy-Schwarz inequality.

**Definition 3.22** ($\eta$-balanced row)**.** *Given an $m \times n$ matrix $A$, we say a row $R_k$ is* $\eta$-balanced *for some constant $\eta$ if $\left||A_{ki}|^2 - |A_{kj}|^2\right| \leq \eta$, for every $i, j \in supp(R_k)$. 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 3.23.** *There exists an absolute constant $c_0 > 0$ such that the following holds. Let $l$ be a line in $\mathbb{C}^d$ and $\mathcal{V}_l = \{v_1, \ldots v_r\}$ be points on $l$ with $r \geq 4$. Let $A = A(l)$ be the dependency matrix for $l$, defined in Lemma 3.21, and $A'$ be a scaling of $A$ such that the $\ell_2$ norm of every row is $\alpha$. Let $M = A'^* A'$. Then*

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

*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 $A'$ we have

$$\sum_{i \neq j} |M_{ij}|^2 = 4(r^2 - r)\alpha^4 - \left(D(A') + 2E(A')\right). \tag{3.12}$$

We can give a lower bound on $D(A') + 2E(A')$ using Property 4 of Lemma 3.21. So 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 are $\eta$-unbalanced. We will show that each such row contributes to either $D(A')$ or $E(A')$.

If a row $R_k$ is $\eta$-imbalanced, note that

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

Alternatively suppose that $R_k$ is $\eta$-balanced. Recall that $\sum_{i=1}^{n} |A'_{ki}|^2 = \alpha$ and note that we must have $|A'_{ki}|^2 \in [\frac{\alpha}{3} - \frac{2\eta}{3}, \frac{\alpha}{3} + \frac{2\eta}{3}]$ 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 nonzero entry in column $s$ and $R_{k'}$ has a third nonzero entry in column $t$, $s \neq t$. 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., $\cos\theta \leq 1/2$. This implies that

$$\begin{aligned}&\left|A'_{ki}\overline{A'_{kj}} - A'_{k'i}\overline{A'_{k'j}}\right|^2 = |A'_{ki}\overline{A'_{kj}}|^2 + |A'_{k'i}\overline{A'_{k'j}}|^2 - 2|A'_{ki}\overline{A'_{kj}}||A'_{k'i}\overline{A'_{k'j}}|\cos\theta \\ &\geq |A'_{ki}\overline{A'_{kj}}|^2 + |A'_{k'i}\overline{A'_{k'j}}|^2 - |A'_{ki}\overline{A'_{kj}}||A'_{k'i}\overline{A'_{k'j}}|.\end{aligned}$$

For any positive real numbers $a, b$

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

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

$$\begin{aligned}&|A'_{ki}\overline{A'_{kj}}|^2 + |A'_{k'i}\overline{A'_{k'j}}|^2 - |A'_{ki}\overline{A'_{kj}}||A'_{k'i}\overline{A'_{k'j}}| \geq \frac{3}{4}|A'_{ki}\overline{A'_{kj}}|^2 \\ &\geq \frac{3}{4}\left(\frac{\alpha}{3} - \frac{2\eta}{3}\right)^2 = \frac{1}{12}\left(\alpha - 2\eta\right)^2.\end{aligned}$$

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

$$E(A') \geq \beta \frac{(r^2 - r)}{3} \eta^2. \tag{3.13}$$

Summing over all the $\eta$-balanced rows, we get

$$\begin{aligned}D(A') &= \sum_{i\neq j}\sum_{k<k'} \left|A'_{ki}\overline{A'_{kj}} - A'_{k'i}\overline{A'_{k'j}}\right|^2 = \frac{1}{2}\sum_{k\neq k'}\sum_{i\neq j}\left|A'_{ki}\overline{A'_{kj}} - A'_{k'i}\overline{A'_{k'j}}\right|^2 \\ &\geq \frac{1}{2}\cdot(1-\beta)\frac{(r^2 - r)}{3}\cdot\frac{1}{12}\left(\alpha - 2\eta\right)^2 = (1-\beta)\frac{(r^2 - r)}{72}\left(\alpha - 2\eta\right)^2.\end{aligned} \tag{3.14}$$

Combining (3.14) and (3.13), and setting $\eta = \alpha/10$ yields

$$\begin{aligned} D(A') + 2E(A') &\geq (1-\beta)\frac{(r^2-r)}{72}(\alpha - 2\eta)^2 + 2\beta\frac{(r^2-r)}{3}\eta^2 \\ &= (r^2-r)\left((1-\beta)\frac{1}{72}\left(\frac{4}{5}\alpha\right)^2 + \beta\frac{2}{3}\left(\frac{1}{10}\alpha\right)^2\right) \\ &\geq c_0(r^2-r)\alpha^2, \end{aligned} \tag{3.15}$$

for some absolute constant $c_0$. Finally, the assertion of lemma now follows by combining (3.12) and (3.15)

□

We are now ready to define the dependency matrix that will be used in the proof of Theorem 1.17.

**Definition 3.24** (Dependency Matrix, second construction)**.** *Let $\mathcal{V} = \{v_1, \dots v_n\}$ 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 $A(l)$, $l \in \mathcal{L}_{\geq 3}(\mathcal{V})$, adjoin $n - r$ column vectors consisting of all zeroes and having length $r^2 - r$ in the column locations corresponding to points not in $l$. This gives an $(r^2 - r) \times n$ matrix. Let $A$ be the matrix obtained by taking the union of rows of these matrices for every $l \in \mathcal{L}_{\geq 3}(\mathcal{V})$. We refer to $A$ as the* dependency matrix *of $\mathcal{V}$.*

Note that this construction is a special case of the one given in Definition 3.13 and so it satisfies all the properties mentioned there. In particular, $AV = 0$ and the number of rows in $A$ is $n^2 - n - 2t_2(\mathcal{V})$.

## 3.6 Proof of Theorem 1.17

We begin with some key lemmas. As before, there are two cases: first when the dependency matrix $A$ satisfies Property $S$ and second when it does not. In the second case, we rely on Lemma 3.15. The following lemma deals with the first case.

**Lemma 3.25.** *There exists an absolute constant $c_1 > 0$ such that the following holds. Let $\mathcal{V} = \{v_1, v_2, \dots, v_n\}$ be a set of points in $\mathbb{C}^d$ not contained in a plane. Let $A$ be the*

*$m \times n$ dependency matrix for $\mathcal{V}$, and suppose that $A$ satisfies Property $S$. Then*

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

*Proof.* By Corollary 3.4, for every $\epsilon > 0$ there exists a scaling $A'$ of $A$ such that

$$\forall i \in [m] \quad \sum_{j \in [n]} \left| A'_{ij} \right|^2 = 1 + \epsilon,$$

and

$$\forall j \in [n] \quad \sum_{i \in [m]} \left| A'_{ij} \right|^2 \geq \frac{m}{n} - \epsilon. \tag{3.16}$$

Let $C_i$ denote the $i^{th}$ column of $A'$, and let $M = A'^* A'$. From (3.16) it follows that $|M_{ii}| = \langle C_i, C_i \rangle \geq \left( \frac{m}{n} - \epsilon \right)$. 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 $l \in \mathcal{L}_{\geq 3}$. Thus, $A'$ 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 3.23 and writing $\alpha = 1 + \epsilon$ gives

$$\begin{aligned}
\sum_{i \neq j} |M_{ij}|^2 &\leq \sum_{l \in \mathcal{L}_3} 4(r^2 - r)\alpha^4 + \sum_{l \in \mathcal{L}_{\geq 4}} \left( 4(r^2 - r)\alpha^4 - c_0 (r^2 - r)\alpha^2 \right) \\
&= \sum_{l \in \mathcal{L}_{\geq 3}} 4(r^2 - r)\alpha^4 - \sum_{l \in \mathcal{L}_{\geq 4}} c_0 (r^2 - r)\alpha^2 \\
&= 4m(1 + \epsilon)^4 - (1 + \epsilon)^2 c_0 \sum_{r \geq 4} (r^2 - r) t_r.
\end{aligned}$$

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

$$\operatorname{rank}(M) \geq \frac{n^2 L^2}{nL^2 + \sum_{i \neq j} |M_{ij}|^2} \geq \frac{n^2 \left( \frac{m}{n} - \epsilon \right)^2}{n \left( \frac{m}{n} - \epsilon \right)^2 + 4m(1 + \epsilon)^4 - (1 + \epsilon)^2 F}.$$

Taking the limit as $\epsilon$ approaches 0, we obtain

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

Note that

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

It follows that if

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

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 yields

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

This holds when

$$t_2(\mathcal{V}) < \frac{3n}{2} + \frac{F}{8} = \frac{3n}{2} + \frac{c_0}{8}\sum_{r=4}^{n}(r^2 - r)t_r(\mathcal{V}),$$

which completes the proof. □

We now have the following easy corollary.

**Corollary 3.26.** *There exists a positive integer $n_0$ such that the following holds. Let $c_1$ be the constant from Lemma 3.25 and $\mathcal{V}$ be a set of $n \geq n_0$ points in $\mathbb{C}^d$ not contained in a plane. Then one of the following must hold:*

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

*Proof.* If $A$ satisfies Property $S$, then we are done by Lemma 3.25. Otherwise, let $b^*$ be an integer such that

$$\frac{n}{2}(b^* - 1) < \frac{3n}{2} + c_1 \sum_{r \geq 4}(r^2 - r)t_r(\mathcal{V}) \leq \frac{n}{2}b^*. \tag{3.17}$$

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

$$b^* < 4 + \frac{2c_1}{n}\sum_{r \geq 4}(r^2 - r)t_r(\mathcal{V}) < \frac{1}{6}n. \tag{3.18}$$

Now by Lemma 3.15 and (3.17), 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 (3.18), there exists a point $v \in \mathcal{V}$, such that the number of ordinary lines containing $v$ is at least

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

□

The following lemma will be crucial for the proof of Theorem 1.17.

**Lemma 3.27.** *Let $\mathcal{V}$ be a set of $n$ points in $\mathbb{C}^d$, and $\mathcal{V}' = \mathcal{V} \setminus \{v\}$ for some $v \in \mathcal{V}$. Then*

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

*Proof.* Note that when we remove $v$ from the set $\mathcal{V}$, we only affect lines incident to $v$. In particular, ordinary lines through $v$ are removed and the number of points on every special line through $v$ decreases by 1. All other lines remain unaffected and so it suffices to consider only lines that contain the point $v$.

Consider the difference

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

We now determine the contribution of a line $l$ determined by $\mathcal{V}$ to $K$.

Each line $l \in \mathcal{L}_{\geq 5}(\mathcal{V})$ (i.e., a line that has $r \geq 5$ points) that contains $v$ contributes $r^2 - r$ to the summation $\sum_{r\geq 4}(r^2 - r)t_r(\mathcal{V})$. In $\mathcal{V}'$, $l$ has $r-1$ points, and contributes $(r-1)^2 - (r-1)$ to the summation $\sum_{r\geq 4}(r^2 - r)t_r(\mathcal{V}')$. Therefore, $l$ contributes $2(r-1)$ to the difference K. We may charge this contribution to the points on $l$ that are different from $v$. There are $r-1$ other points on $l$, so each point contributes 2 to $K$.

Each line $l \in \mathcal{L}_4(\mathcal{V})$ that contains $v$ contributes $r^2 - r = 12$ to the summation $\sum_{r\geq 4}(r^2 - r)t_r(\mathcal{V})$. These lines contain three points in $\mathcal{V}'$, and so do not contribute anything in the $\sum_{r\geq 4}(r^2 - r)t_r(\mathcal{V}')$ term. Once again, we charge this contribution to the points lying on $l$ that are different from $v$. 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 through $v$ and any other point, and each point either contributes 0, 2 or 4 to $K$. This shows that

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

Rearranging completes the proof. □

We are now ready to prove our main result. For convenience, we repeat the statement of the theorem again.

**Theorem 1.17.** *There exists an absolute constant $c > 0$ and a positive integer $n_0$ such that the following holds. Let $\mathcal{V}$ be a set of $n \geq n_0$ points in $\mathbb{C}^3$ with at most $\frac{1}{2}n$ points contained in any (complex) plane. Then*

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

*Proof.* The remainder of the proof is similar to the proof of Theorem 1.18, in that we use Corollary 3.26 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 are unable to find such a point, in which case Corollary 3.26 guarantees a large number of ordinary lines, or when we have accumulated enough ordinary lines. As before, consider the following algorithm:

Let $\mathcal{V}_0 := \mathcal{V}$ and $j = 0$.

1. If $\mathcal{V}_j$ satisfies case (2) of Lemma 3.26, then stop.

2. Otherwise, there must exist a point $v_{j+1}$ incident to at least $\frac{n-j}{2}$ ordinary lines. Let $\mathcal{V}_{j+1} = \mathcal{V}_j \setminus \{v_{j+1}\}$.

3. Set $j = j + 1$. If $j = n/2$, then stop. Otherwise go to Step 1.

Note that since no plane contains more than $n/2$ points, the algorithm will never stop because the configuration becomes planar; that is, we can use Corollary 3.26 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 the algorithm stop because $\mathcal{V}_j$ satisfies case (2) of Corollary 3.26 for some $1 \le j < n/2$. From case (2) of Lemma 3.26 and Lemma 3.27, we have

$$\begin{aligned} t_2(\mathcal{V}_j) &\ge \frac{3(n-j)}{2} + c_1 \sum_{r\ge 4}(r^2-r)t_r(\mathcal{V}_j) \\ &\ge \frac{3(n-j)}{2} + c_1 \left( \sum_{r\ge 4}(r^2-r)t_r(\mathcal{V}) - 4\sum_{i=1}^{j}(n-i) \right). \end{aligned} \tag{3.19}$$

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

$$\sum_{i=1}^{j} \left( \frac{n-i}{2} - \frac{n-i}{4} \right) = \frac{1}{4} \sum_{i=1}^{j} (n-i)\,. \tag{3.20}$$

By combining (3.19) and (3.20), we obtain

$$\begin{aligned} t_2(\mathcal{V}) &\ge \frac{3}{2}(n-j) + c_1 \left( \sum_{r\ge 4}(r^2-r)t_r(\mathcal{V}) - 4\sum_{i=1}^{j}(n-i) \right) + \frac{1}{4}\sum_{i=1}^{j}(n-i) \\ &= \frac{3}{2}n + c_1 \sum_{r\ge 4}(r^2-r)t_r(\mathcal{V}) + \left(\frac{1}{4} - 4c_1\right)\sum_{i=1}^{j}(n-i) - \frac{3}{2}j. \end{aligned}$$

For $c_1$ small enough and $n$ large, the term $\left(\frac{1}{4} - 4c_1\right)\sum_{i=1}^{j}(n-i) - \frac{3}{2}j$ is positive. Thus, there must exist an absolute constant $c > 0$ such that

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

We now consider the case when the algorithm stops because $j = n/2$. Note that at this point, we will have pruned exactly $j$ points. Each pruned point $v_i$, $1 \le i \le j$, is incident to at least $\frac{n-i+1}{2} > \frac{n-i}{2}$ ordinary lines determined by $\mathcal{V}_{i-1}$. However, as many as $i-1 < i$ ordinary lines through $v_i$ 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

$$t_2(\mathcal{V}) \geq \sum_{i=1}^{j} \frac{n-i}{2} - \sum_{i=1}^{j} i = \frac{1}{2} \sum_{i=1}^{j} (n-3i) = \frac{n^2-6n}{16}.$$

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

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

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

□

# Chapter 4

# Applications of Incidence Bounds

## 4.1 The $k$-Most-Frequent Distances

In this section we study the $k$ most frequent distances problem that was described in the introduction. For convenience, we reiterate the definition of the problem. Given a set $\mathcal{P}$ of points in $\mathbb{R}^3$ and a set $D$ of $k$ distinct distances, we denote by $f(\mathcal{P}, D)$ the number of pairs of points of $\mathcal{P} \subset \mathbb{R}^3$ that span a distance in $D$. We set $f_k(m) = \max f(\mathcal{P}, D)$, where the maximum is taken over all sets $\mathcal{P}$ of $m$ points in $\mathbb{R}^3$ and all sets $D$ of $k$ distinct distances. Notice that $D$ can be any set of distinct distances, though the extremal case occurs when $D$ is the set of the most frequent distances.

**Definition 4.1.** *Given a point set $\mathcal{P}$, a set $\mathcal{S}$ of spheres, both in $\mathbb{R}^3$, and a positive constant $\ell$, we say that an incidence between a point $p \in \mathcal{P}$ and a sphere $\sigma \in \mathcal{S}$ is* $\ell$-proper *if there is no circle $c$ such that*

1. *$c$ is incident to $p$,*
2. *$c$ is contained in $\sigma$,*
3. *$c$ is contained in at least $\ell$ spheres of $\mathcal{S}$.*

To derive our bound for the $k$ most frequent distances problem, we require the following generalization of Theorem 1.4.

**Theorem 4.2.** *Let $\mathcal{P}$ be a set of $m$ points, let $\mathcal{S}$ be a set of $n$ spheres, both in $\mathbb{R}^3$. Let $\ell \in (1, n)$ be an integer. Then for every $\varepsilon > 0$, the number of $\ell$-proper incidences between $\mathcal{P}$ and $\mathcal{S}$ is $O(m^{3/4+\varepsilon} n^{3/4} \ell^{1/4} + n + m\ell)$, where the constant of proportionality depends on $\varepsilon$.*

*Proof.* It suffices to perform one minor change in the proof of Theorem 1.4 (after switching $k$ with $\ell$ in the proof). The proof of Theorem 1.4 relied on the non-degeneracy of $\mathcal{S}$ in two places, and we now verify that both remain valid:

- In obtaining the weaker bound in Lemma 2.7, we count the number of quadruples $(a, b, c, \sigma)$, where $\sigma \in \mathcal{S}$ and $a, b, c \in \mathcal{P} \cap \sigma$. In the current proof we only count quadruples where each of the three points forms an $\ell$-proper incidence with the sphere. Notice that each triple of points still gets counted at most $\ell - 1$ times, since if more than $\ell - 1$ spheres contain a triple, none of these spheres form an $\ell$-proper incidence with a point of the triple. The rest of the proof of Lemma 2.7 remains unchanged.
- In Lemma 2.12 we rely on the property that every circle is contained in at most $k$ spheres and use it to bound incidences on "one-dimensional intersections" with the partitioning polynomials. Replacing the non-degeneracy with $\ell$-proper incidences does not affect the analysis.

□

**Theorem 1.9.** *For any* $\varepsilon > 0$, $f_k(m) = O(m^{236/149+\varepsilon}k^{125/149}) \approx O(m^{1.58}k^{0.84})$.

*Proof.* Let $D$ be a set of $k$ distances and $\mathcal{P}$ be a set of $m$ points in $\mathbb{R}^3$ such that $f_k(m) = f(\mathcal{P}, D)$. For each point $p \in \mathcal{P}$ we generate $k$ spheres with center $p$, each with a distinct radius from $D$. Let $\mathcal{S}$ denote the resulting set of $mk$ spheres, and notice that $f_k(m) = f(\mathcal{P}, D) = I(\mathcal{P}, \mathcal{S})/2$ (a pair of points with distance $d$ corresponds to a pair of incidences between $\mathcal{P}$ and $\mathcal{S}$). Since the spheres of $\mathcal{S}$ have only $k$ different radii, every circle in $\mathbb{R}^3$ is contained in at most $2k$ spheres of $\mathcal{S}$. However, applying Theorem 1.4 only implies the worse-than-trivial bound $I(\mathcal{P}, \mathcal{S}) = O(m^{3/4+\varepsilon}(mk)^{3/4}k^{1/4} + mk) = O(m^{3/2+\varepsilon}k)$.

To obtain a better bound we set $\ell = m^{50/149}k^{53/149}$ and separately bound the number of $\ell$-proper incidences and the number of non-$\ell$-proper incidences. According to Theorem 4.2, the number of $\ell$-proper incidences between $\mathcal{P}$ and $\mathcal{S}$ is

$$O(m^{3/4+\varepsilon}(mk)^{3/4}\ell^{1/4} + mk + m\ell) = O(m^{236/149+\varepsilon}k^{125/149}). \tag{4.1}$$

Now we bound the number of non-$\ell$-proper incidences. Consider a circle $c$ that is contained in a set of spheres, and notice that the centers of all of these spheres must be collinear. Specifically, these centers are all on the line $L_c$ that is incident to the center of $c$ and perpendicular to the plane that contains $c$; we will say that $L_c$ is *the line corresponding to* $c$. We rely on a bound for the number of lines that can contain many points of $\mathcal{P}$ to bound the number of circles that are contained in many spheres, and thus the number non-$\ell$-proper incidences. Specifically, we use the Szemerédi-Trotter theorem (e.g., see [58, 72]) to bound the number of "heavy" lines.

Partition the circles that are contained in at least $\ell$ spheres of $\mathcal{S}$ into $\log_2(m/\ell)$ classes $C_1, C_2, \ldots$, where $C_i$ consists of the circles that are contained in at least $2^{i-1}\ell$ spheres of $\mathcal{S}$, and fewer than $2^i\ell$ such spheres.

We bound the maximum possible of circles $C_i$ can have. For every $1 \leq j \leq \log_2(m/\ell) - i + 1$, let $\mathcal{L}_j^{(i)}$ denote the set of lines that are incident to at least $2^{i+j-2}\ell$ points of $\mathcal{P}$, and incident to at most $2^{i+j-1}\ell$ such points. According to the Szemerédi-Trotter theorem, the number of lines that contain at least $2^{i+j-2}\ell$ points of $\mathcal{P}$ is $O(m^2/(2^{i+j-2}\ell)^3)$, i.e., $|\mathcal{L}_j^{(i)}| = O(m^2/(2^{i+j-2}\ell)^3)$.

Consider a line $L \in \mathcal{L}_j^{(i)}$ and translate and rotate the space so that $L$ becomes the $x$-axis. Notice that every circle of $C_i$ to which $L$ corresponds must intersect the $xy$-plane in a pair of points that are symmetric around $L$. Let $\mathcal{P}_{xy}^L$ denote the set of these points that have a positive $y$-coordinate (i.e., one point out of every pair). Similarly, every sphere whose center is incident to $L$ intersects the $xy$-plane in a circle whose center is incident to $L$. Let $C_{xy}^L$ denote the set of these circles. There is a bijection between circles of $C_i$ to which $L$ corresponds and points of $\mathcal{P}_{xy}^L$ that are incident to at least $2^{i-1}\ell$ circles of $C_{xy}^L$. Since every point of $\mathcal{P}$ that is incident to $L$ corresponds to $k$ spheres of $\mathcal{S}$, we have $|C_{xy}^L| \leq 2^{i+j-1}\ell k$. According to [2] and [10], the number of incidences between a set of $M$ points and a set of $N$ circles, both in $\mathbb{R}^2$, is[1]

$$O^*\left(M^{2/3}N^{2/3} + M^{6/11}N^{9/11} + M + N\right). \tag{4.2}$$

This implies that the number of points of $\mathcal{P}_{xy}^L$ that are incident to at least $2^{i-1}\ell$

[1]Recall that the $O^*(\cdot)$-notation hides sub-polynomial factors.

circles of $C_{xy}^L$ is $O^*\left(2^{2j-i}k^2/\ell + 2^{9j/5-2i/5}k^{9/5}/\ell^{2/5}\right)$. We thus have

$$|C_i| = \sum_{j=1}^{\log_2(m/\ell)-i} O^*\left(|\mathcal{L}_j^{(i)}| \cdot \left(\frac{2^{2j-i}k^2}{\ell} + \frac{2^{9j/5-2i/5}k^{9/5}}{\ell^{2/5}}\right)\right)$$

$$= \sum_{j=1}^{\log_2(m/\ell)-i} O^*\left(\frac{m^2k^2}{2^{4i+j}\ell^4} + \frac{m^2k^{9/5}}{\ell^{17/5}2^{17i/5+6j/5}}\right) = O^*\left(\frac{m^2k^2}{2^{4i}\ell^4} + \frac{m^2k^{9/5}}{\ell^{17/5}2^{17i/5}}\right)$$

$$= O^*\left(\frac{m^{98/149}k^{86/149}}{2^{4i}} + \frac{m^{128/149}k^{88/149}}{2^{17i/5}}\right) = O^*\left(\frac{m^{128/149}k^{88/149}}{2^{17i/5}}\right).$$

Now that we have established an upper bound on the cardinality of $C_i$ we can bound the number of incidences with the circles of this class. By [9], the point-circle incidence bound (4.2) holds also in $\mathbb{R}^3$, and we obtain

$$I(\mathcal{P}, C_i) = O^*\left(m^{2/3}\left(\frac{m^{128/149}k^{88/149}}{2^{17i/5}}\right)^{2/3}\right.$$

$$\left. + m^{6/11}\left(\frac{m^{128/149}k^{88/149}}{2^{17i/5}}\right)^{9/11} + \frac{m^{128/149}k^{88/149}}{2^{17i/5}}\right)$$

$$= O^*\left(\frac{m^{554/447}k^{176/447}}{2^{34i/15}} + \frac{m^{186/149}k^{72/149}}{2^{153i/55}} + \frac{m^{128/149}k^{88/149}}{2^{17i/5}}\right).$$

Therefore, the number of incidences that are not $\ell$-proper due to circles of $C_i$ is at most

$$2^i\ell \cdot I(\mathcal{P}, C_i) = O^*\left(\frac{m^{704/447}k^{335/447}}{2^{19i/15}} + \frac{m^{236/149}k^{125/149}}{2^{98i/55}} + \frac{m^{178/149}k^{141/149}}{2^{12i/5}}\right).$$

This implies that the number of non-$\ell$-proper incidences between $\mathcal{P}$ and $\mathcal{S}$ is

$$\sum_{i=1}^{\log_2(m/\ell)} O^*\left(\frac{m^{704/447}k^{335/447}}{2^{19i/15}} + \frac{m^{236/149}k^{125/149}}{2^{98i/55}} + \frac{m^{178/149}k^{141/149}}{2^{12i/5}}\right)$$

$$= O^*\left(m^{704/447}k^{335/447} + m^{236/149}k^{125/149} + m^{178/149}k^{141/149}\right)$$

$$= O^*\left(m^{236/149}k^{125/149}\right), \tag{4.3}$$

where the final step holds since $k \le m$. Combining (4.1) and (4.3) yields the assertion of the theorem. A somewhat more tedious analysis can show that the chosen value of $\ell$ is optimal. □

## 4.2 Sum-Product Estimates

In this section we consider the sum-product estimates described in the introduction. Incidence bounds and other geometric ideas appear frequently when proving such bounds.

Our contribution is in showing that some of these bounds extend from the real to the complex setting.

A common theme in the proofs of Theorems 1.20 and 1.22 was the following. Given a finite set $A \subset \mathbb{R}$ with all elements positive, consider the point set $\mathcal{P} = A \times A \subset \mathbb{R}^2$. Let $\mathcal{L}$ be the set of lines through the origin covering $\mathcal{P}$. Now to prove a lower bound on the size of the set in question, it suffices to consider the set $\mathcal{P} + \mathcal{P}$. A key observation here is that if $a_1 = (x_1, y_1) \in l_1$ and $a_2 = (x_2, y_2) \in l_2$, the point $a_1 + a_2 = (x_1 + x_2, y_1 + y_2) \in (\mathcal{P}+\mathcal{P}) \times (\mathcal{P}+\mathcal{P})$ lies inside the *wedge* between $l_1$ and $l_2$ centered at the origin. For lines $l_1, l_2, l_3, l_4$, suppose that the wedge defined by $l_1, l_2$ does not overlap the wedge defined by $l_3, l_4$. Then for all $a_1 \in l_1, b_1 \in l_2, a_2 \in l_3, b_2 \in l_4$, $a_1 + b_1 \neq a_2 + b_2$, and it suffices to consider the points contained in the respective pairs independently. On the other hand, if the wedges intersect, *overlapping sums* need to be considered. In this case, incidence bounds are useful to show that there can not be too many overlaps.

When incidence bounds are needed, we rely on the following results that generalize classical incidence theorems to the complex plane.

**Theorem 4.3** (The Complex Szemerédi-Trotter Theorem [74, 78])**.** *Let $\mathcal{P}$ be a set of $m$ points and $\mathcal{L}$ a set of $n$ lines, both in $\mathbb{C}^2$. Then the number of incidences between points of $\mathcal{P}$ and lines of $\mathcal{L}$ is $O(m^{2/3}n^{2/3} + m + n)$.*

The following theorem can be considered a generalization of Theorem 4.3, where lines are replaced with more general curves.

**Theorem 4.4** (The Complex Pach-Sharir Theorem [66])**.** *Let $\mathcal{P}$ be a set of $m$ points and $\Gamma$ a set of $n$ algebraic curves such that*

1. *any two distinct curves from $\Gamma$ intersect in at most two points of $\mathcal{P}$;*
2. *for any two points $p, q \in \mathcal{P}$, there exist at most two curves in $\Gamma$ which pass through both $p$ and $q$.*

*Then, for any $\varepsilon > 0$, the number of incidences between points of $\mathcal{P}$ and curves of $\Gamma$ is $O(m^{2/3+\varepsilon}n^{2/3} + m + n)$, where the constant of proportionality depends on $\varepsilon$.*

Note that Theorem 4.4 gives a slightly weaker bound, due to the addition of $\varepsilon$ in the exponent of $m$. This results in the loss of an $\varepsilon$ in the exponent for the result stated in Theorem 1.24.

In Section 4.2.1, we describe the main tools that we need for our theorems. In Section 4.2.2, we present the proof of Theorem 1.23. We do not present the proof of Theorem 1.24, but simply note that the proof in [53] can be extended similarly using the ideas that are presented below.

### 4.2.1 Rhombi and graphs

**Definition 4.5.** *For a fixed $\varepsilon > 0$ and points $u, v \in \mathbb{R}^2$, let $R_{u,v}(\varepsilon)$ be the open rhombus whose major diagonal is the segment $uv$, and whose minor diagonal has length $\varepsilon|u-v|$. When $\varepsilon$ is clear from the context, we use the abbreviated notation $R_{u,v}$.*

We will require the following lemma of Konyagin and Rudnev [50].

**Lemma 4.6.** *Let $A$ be a subset of $\mathbb{C} \setminus \{0\}$ contained inside an angular sector $S := \{z \in \mathbb{C} : |\tan(2arg(z))| < \varepsilon\}$ for some fixed $\varepsilon > 0$. Let $u, v$ be two distinct elements of the ratio set $A/A \subset \mathbb{C} \cong \mathbb{R}^2$ with realizations $u = \frac{y_1}{x_1}$ and $v = \frac{y_2}{x_2}$. Then the point $w = \frac{y_1+y_2}{x_1+x_2} \in \mathbb{C} \cong \mathbb{R}^2$ lies in the set $R_{u,v}(\varepsilon)$.*

*Proof.* Let $t = x_2/x_1$ and note that we may write

$$w = \frac{y_1 + y_2}{x_1 + x_2} = u\frac{1}{1+t} + v\frac{t}{1+t} = v + (u-v)\frac{1}{1+t}.$$

Since $|\tan(2\arg(x_1))| < \varepsilon$ and $|\tan(2\arg(x_2))| < \varepsilon$, $t$ lies in the angular sector $S_\varepsilon := \{z \in \mathbb{C} : |\tan\arg(z)| < \varepsilon\}$. Then $\frac{1}{1+t}$ lies in the image of $S_\varepsilon$ under the map $f : z \to \frac{1}{1+z}$, which we denote as $M_\varepsilon$. The map $f$ can be viewed as a translation by $(1,0)$, followed by an inversion around the unit disk centered at the origin. It is not too hard to check that $f$ maps the lines $\{z \in \mathbb{C} : \tan\arg(z) = \varepsilon\}$ and $\{z \in \mathbb{C} : \tan\arg(z) = -\varepsilon\}$ to circles going through the origin and $(1,0)$, such that the tangents to the circles at the origin form an angle of $\varepsilon$ and $-\varepsilon$ respectively with the real axis. Then $M_\varepsilon$ is the *meniscus* formed by the intersection of the two open disks bounded by these circles. Note that $M_\varepsilon$ is contained within the open rhombus $R_\varepsilon$ whose major diagonal is the real line interval $(0, 1)$ and whose minor diagonal is of length $\varepsilon$. Let the set $M_{u,v}$ be the image of $M_\varepsilon$ under a dilation by $(u - v)$, and a translation by $v$, i.e.,

$$M_{u,v} = v + (u - v)M_\varepsilon.$$

Clearly, $w$ must lie within $M_{u,v}$. Finally, it is easily seen that $M_{u,v}$ is contained in the open rhombus $R_{u,v}$. □

**Definition 4.7.** *Let $B$ be a finite set of points in $\mathbb{R}^2$. We say that $v \in B$ is a $K$-nearest neighbor of $w \in B$ if there are at most $K - 1$ points in $B$ that are closer to $w$ than $v$ is. A $K$-nearest neighbor graph on $B$ is a directed graph with out-degree exactly $K$, such that each point $v \in B$ is connected to its $K$-nearest neighbors and ties are broken arbitrarily. An undirected $K$-nearest neighbor graph on $B$ is a $K$-nearest neighbor graph on $B$ that ignores the direction of each edge.*

A vertex may have arbitrarily many vertices at the same distance from it, but will only have out-degree $K$ in a $K$-nearest neighbors graph. Therefore, a $K$-nearest-neighbors graph is not unique. For our result, we will need an upper bound on the in-degree of a $K$-nearest-neighbor graph.

**Lemma 4.8.** *The in-degree of a vertex in a $K$-nearest neighbor graph is at most $6K$.*

*Proof.* Let $v$ be a vertex in the graph, and let $W$ be a cone with vertex $v$ and angle $\pi/3$, so that the boundary of $W$ contains no vertex other than $v$. Let $u$ be the most distant vertex in $W$ such that $v$ is a $K$-nearest neighbor of $u$. If $w$ is a vertex in $W$ and $v$ is a $K$-nearest neighbor of $w$, then $w$ is closer to $u$ than $v$ is, and so must be a $K$-nearest neighbor of $u$. Hence, $W$ contains at most $K$ vertices that have $v$ as a $K$-nearest neighbor. The lemma follows since we can cover the plane with six wedges with vertex $v$ and angle $\pi/3$. □

The next lemma shows that if we consider the $K$-nearest-neighbor graph on a point set where the edges are rhombi (instead of segments), then there are few intersections between the edges.

**Lemma 4.9.** *Let $\varepsilon > 0$ be a sufficiently small constant, $B$ be a set of points in $\mathbb{R}^2$, $K$ an integer in $[1, B)$, and $\mathcal{G}$ be an undirected $K$-nearest neighbor graph on $B$. Then, for each edge $\{u, v\} \in \mathcal{G}$, there are at most $112K^2$ edges $\{p, q\} \in \mathcal{G}$ such that $R_{u,v}(\varepsilon) \cap R_{p,q}(\varepsilon) \neq \emptyset$.*

*Proof.* Let $\{u, v\}, \{p, q\} \in \mathcal{G}$ such that $R_{u,v} \cap R_{p,q} \neq \emptyset$. We will show that either $\{u, v\}, \{p, q\}$ share a vertex, or are connected by an edge. Then the maximum number of edges $\{p', q'\}$ where $R_{u,v} \cap R_{p',q'} \neq \emptyset$ is less than the number of paths of length 1 plus the number of paths of length 2 from $u$ and $v$. Using Lemma 4.8, this quantity is easily seen to be at most $2(7K + (7K)^2) \leq 112K^2$.

We may assume that $u, v, p, q$ are all distinct (i.e., $\{u, v\}$ and $\{p, q\}$ do not share a vertex). Without loss of generality, suppose that $v$ is a $K$-nearest neighbor of $u$ and $q$ is a $K$-nearest neighbor of $p$. We show that at least one of $\{u, p\}, \{u, q\}, \{v, p\}$ is an edge in $\mathcal{G}$, and this will complete the proof of the lemma.

Since $R_{u,v}$ and $R_{p,q}$ are open and have non-empty intersection, either one is contained in the other or their boundaries must intersect at least two times. Suppose one of the following holds: (i) one of the edges is contained in the other; (ii) one of the intersections lies at $u$ or at $v$; (iii) there are intersections between the boundaries of $R_{u,v}$ and $R_{p,q}$ on both sides of the line connecting $\{u, v\}$. Then, either $u$ or $v$ is contained in the closure of $R_{p,q}$, or the segments $uv$ and $pq$ cross. Therefore, it suffices to consider three cases: 1) one of the vertices $u, v$ is contained in the closure of $R_{p,q}$, or vice-versa; 2) the segments cross properly; 3) all intersection points lie strictly *between the two segments*, i.e., the intersection points lie on the same side of the line connecting $u, v$ and on the same side of the line connecting $p, q$.

For $\varepsilon < \sqrt{3}$, if $u \in R_{p,q}$, then $\|u - p\| < \|p - q\|$, and since $q$ is a $K$-nearest neighbor of $p$, it follows that $\{u, p\} \in \mathcal{G}$. The same argument applies when $v \in R_{p,q}$, and this finishes the first case.

If the segments $\{u, v\}$ and $\{p, q\}$ cross properly, then $\{u, v, p, q\}$ lie in convex position, and the point where $\{u, v\}$ and $\{p, q\}$ intersect is interior to their convex hull. The triangle inequality implies that

$$\|u - v\| + \|p - q\| > \|u - q\| + \|v - p\|,$$

so either $\|u-q\| < \|u-v\|$ and $\{u,q\} \in \mathcal{G}$, or else $\|v-p\| < \|p-q\|$ and $\{v,p\} \in \mathcal{G}$.

Now, suppose that all intersections between $R_{u,v}$ and $R_{p,q}$ lie between the edges $\{u,v\}$ and $\{p,q\}$. Without loss of generality, we suppose that $\|u-v\| = 1$ and $\|p-q\| \le 1$. Let $a$ and $b$ be the points where the boundaries of $R_{u,v}$ and $R_{p,q}$ intersect. Suppose, again without loss of generality, that $a$ is closer than $b$ to both $u$ and $q$. Clearly $\|u-a\| < 1$ and $\|a-q\| < 1$. Write $\sigma = \tan^{-1}\varepsilon$, and note that the angles $\angle uav$ and $\angle vaq$ are at least $\pi - 2\sigma$, so the angle $\angle qau$ is at most $4\sigma$. Now, for $\varepsilon$ small enough,

$$\|u-q\|^2 = \|u-a\|^2 + \|a-q\|^2 - 2\|u-a\|\|a-q\|\cos(\angle qau) < 1.$$

That is, $\|u-q\| < \|u-v\|$ which implies that $\{u,q\} \in \mathcal{G}$ and completes the proof. $\square$

### 4.2.2 The sum-product problem

We now give an overview of the proof of Theorem 1.23. This closely follows the proof of Theorem 1.20 (a statement established by Konyagin and Shkredov [52] in the context of the reals) and we point out the changes where necessary. We begin by stating some definitions and lemmas.

#### Preliminaries

The *multiplicative energy* of $A$ is defined to be

$$E^{\times}(A) = \left|\left\{(a_1,a_2,a_3,a_4) \in A^4 : a_1a_2 = a_3a_4\right\}\right|.$$

For any $x \in AA$, let $r_A^{\times}(x) = \left|\left\{(a_1,a_2) \in A^2 : a_1a_2 = x\right\}\right|$. Since every pair in $A \times A$ contributes once, it follows that $\sum_{x\in AA} r_A^{\times}(x) = |A|^2$. The Cauchy-Schwarz inequality implies that

$$E^{\times}(A) = \sum_{x\in AA} r_A^{\times}(x)^2 \ge \frac{\left(\sum_x r_A^{\times}(x)\right)^2}{|AA|} = \frac{|A|^4}{|AA|}. \tag{4.4}$$

For $\alpha \in \mathbb{C}$, define $\alpha A = \{\alpha \cdot a : a \in A\}$. For $A_1, A_2, A_3 \subset \mathbb{C}$, define

$$\sigma(A_1,A_2,A_3) = \left|\{(a_1,a_2,a_3) \in A_1 \times A_2 \times A_3 : a_1+a_2+a_3 = 0\}\right|.$$

Given a finite set $A$ and $\lambda \in A/A$, we denote $A_\lambda = A \cap (\lambda A)$. For $\tau \in \mathbb{R}$, let

$$S_\tau = \{\lambda \in A/A : \tau < |A_\lambda| \le 2\tau\} \subseteq A/A.$$

Our main contribution is the following lemma, which generalizes Lemma 10 in [51] to the complex setting. While the statement of the lemma is essentially the same, the proof relies on the additional machinery that we developed in Section 4.2.1.

**Lemma 4.10.** *Consider a finite subset $A$ of $\mathbb{C} \setminus \{0\}$ contained inside an angular sector $\{z \in \mathbb{C} : |\tan(2arg(z))| < \varepsilon\}$ for a small enough constant $\varepsilon > 0$. Let $\tau$ and $\sigma$ be positive real numbers such that $32\sigma \le \tau^2 \le |A+A|\sqrt{\sigma}$. Moreover, let $S'_\tau \subseteq S_\tau$ with the property that for any nonzero $\alpha_1, \alpha_2, \alpha_3 \in \mathbb{C}$ and $\lambda_1, \lambda_2, \lambda_3 \in S'_\tau$ we have*

$$\sigma(\alpha_1 A_{\lambda_1}, \alpha_2 A_{\lambda_2}, \alpha_3 A_{\lambda_3}) \le \sigma. \tag{4.5}$$

*Then*

$$|A+A|^2 \ge |S'_\tau| \frac{\tau^3}{16\sqrt{14\sigma}}.$$

*Proof.* We consider the point set $\mathcal{P} = A \times A$ and double count $|\mathcal{P}+\mathcal{P}|$. On one hand, we have $|\mathcal{P}+\mathcal{P}| = |(A+A) \times (A+A)| = |A+A|^2$. We now derive a lower bound for $|\mathcal{P}+\mathcal{P}|$, relying on several observations from the proof of Theorem 1.19. Specifically, we consider lines that are incident to the origin and have a slope from $A/A$. Let $l_\lambda$ be the line $y = \lambda x$, where $\lambda \in A/A$. Clearly, a line $l_\lambda$ intersects $\mathcal{P}$ in the set $\{(x, \lambda x) : x \in A_\lambda\}$, which we denote by $\mathcal{A}_\lambda$.

Let $K \in [1, |S'_\tau|)$ be an integer which we will set below. From Lemma 4.9, we obtain a set $\mathcal{G} \subset S'_\tau \times S'_\tau$ of size at least $|S'_\tau|K/2$. We consider each pair of slopes in $\mathcal{G}$ and sum up the points that lie on the lines with those slopes. It is easy to check that any two sets $\mathcal{A}_{\lambda_1}$ and $A_{\lambda_2}$ contribute $|\mathcal{A}_{\lambda_1}| \cdot |\mathcal{A}_{\lambda_2}|$ distinct vector sums. Note that if $p \in \mathcal{A}_{\lambda_1}, q \in \mathcal{A}_{\lambda_2}, p' \in \mathcal{A}_{\lambda_3}, q' \in \mathcal{A}_{\lambda_4}$ satisfy $p+q = p'+q'$, then $R_{\lambda_1,\lambda_2} \cap R_{\lambda_3,\lambda_4} \ne \emptyset$. Equivalently, if $R_{\lambda_1,\lambda_2} \cap R_{\lambda_3,\lambda_4} = \emptyset$, then $p+q \ne p'+q'$.

Therefore

$$|\mathcal{P}+\mathcal{P}| \ge \sum_{(\lambda_1,\lambda_2)\in\mathcal{G}} \left( \tau^2 - \sum_{\substack{(\lambda_3,\lambda_4)\in\mathcal{G},\, R_{\lambda_1,\lambda_2}\cap R_{\lambda_3,\lambda_4}\ne\emptyset, \\ (\lambda_1,\lambda_2)\ne(\lambda_3,\lambda_4)}} |\{z : z \in (\mathcal{A}_{\lambda_1}+\mathcal{A}_{\lambda_2}) \cap (\mathcal{A}_{\lambda_3}+\mathcal{A}_{\lambda_4})\}| \right). \tag{4.6}$$

We will now bound the value of the inner summation. Consider one element of the sum, fixing $\lambda_1, \ldots, \lambda_4$. At least one of $\lambda_1, \lambda_2, \lambda_3, \lambda_4$ differs from the others. Say,

$\lambda_4 \notin \{\lambda_1, \lambda_2, \lambda_3\}$. Consider

$$z = (z_1, z_2) \in (\mathcal{A}_{\lambda_1} + \mathcal{A}_{\lambda_2}) \cap (\mathcal{A}_{\lambda_3} + \mathcal{A}_{\lambda_4}).$$

Then for some $a_1 \in A_1, \ldots, a_4 \in A_4$ we have $z_1 = a_1 + a_2 = a_3 + a_4$ and $z_2 = \lambda_1 a_1 + \lambda_2 a_2 = \lambda_3 a_3 + \lambda_4 a_4$. This implies that

$$\begin{aligned} 0 &= \lambda_1 a_1 + \lambda_2 a_2 - \lambda_3 a_3 - \lambda_4 a_4 - \lambda_4(a_1 + a_2 - a_3 - a_4) \\ &= (\lambda_1 - \lambda_4)a_1 + (\lambda_2 - \lambda_4)a_2 - (\lambda_3 - \lambda_4)a_3. \end{aligned} \tag{4.7}$$

Since $\lambda_4 \notin \{\lambda_1, \lambda_2, \lambda_3\}$, none of the coefficients of $a_i$, $i \in \{1, 2, 3\}$, is zero, so by (4.5) there are at most $\sigma$ solutions to (4.7). That is, $|\{z : z \in (\mathcal{A}_{\lambda_1} + \mathcal{A}_{\lambda_2}) \cap (\mathcal{A}_{\lambda_3} + \mathcal{A}_{\lambda_4})\}| \leq \sigma$. Combining this with (4.6), and recalling that every pair $(\lambda_1, \lambda_2)$ in $\mathcal{G}$ intersects at most $112K^2$ other pairs in $\mathcal{G}$, it follows that

$$|\mathcal{P} + \mathcal{P}| \geq \sum_{(\lambda_1,\lambda_2)\in\mathcal{G}} \left(\tau^2 - 112K^2\sigma\right) \geq \frac{|S'_\tau|K}{2}(\tau^2 - 112K^2\sigma) = |S'_\tau| \left(\frac{\tau^2 K}{2} - 56\sigma K^3\right).$$

Set $K = \lceil \tau/\sqrt{224\sigma} \rceil$. Since we assume that $32\sigma \leq \tau^2$, $K \geq 1$. By combining this with the trivial $|\mathcal{P} + \mathcal{P}| = |A + A|^2$, we obtain

$$|A + A|^2 \geq |S'_\tau| \left(\frac{\tau^2 K}{2} - 56\sigma K^3\right) \geq |S'_\tau| \frac{\tau^3}{16\sqrt{14\sigma}}.$$

Notice that if $K \geq |S'_\tau|$, we are not able to use Lemma 4.9. In this case, we assume for contradiction that the assertion of the lemma fails. If true, we would have

$$|A + A|^2 < |S'_\tau| \frac{\tau^3}{16\sqrt{14\sigma}} \leq K \frac{\tau^3}{16\sqrt{14\sigma}} \leq \frac{\tau^4}{896\sigma},$$

but this contradicts the assumption $\tau^2 \leq |A + A|\sqrt{\sigma}$ and so completes the proof. □

The next lemma generalizes Lemma 12 from [51] to the complex numbers. Aside from the use of Lemma 4.10, the proof from [51] remains unchanged, so we omit the proof here.

**Lemma 4.11.** *Consider a finite subset $A$ of $\mathbb{C} \setminus \{0\}$ contained inside an angular sector $\{z \in \mathbb{C} : |\tan(2arg(z))| < \varepsilon\}$ for a small enough constant $\varepsilon > 0$. Let $L \geq 1$ be a real number such that $|A + A|^2 |A/A| \leq L|A|^4$. Then there exist $\tau > E^\times(A)/(2|A|^2)$ and* *sets* $S'_\tau \subset S_\tau \subset A/A$ *such that* $|S_\tau|\tau^2 = \Omega(E^\times(A)/\lg|A|)$, $|S'_\tau| \ge |S_\tau|/2$, *and for any* $\lambda \in S'_\tau$ *we have*

$$|A_\lambda A_\lambda| = \Omega(\tau^2/L^{16}\lg^{17}|A_\lambda|).$$

We now introduce some parameters of a finite set $A \subset \mathbb{C}$ that measure additive or multiplicative properties. We say that $A$ is of Szemerédi-Trotter type (abbreviated as *SzT-type*) with parameter $D := D(A)$ if for every $B \subset \mathbb{C}$ and integer $\tau \ge 1$

$$|\{s \in A - B : |A \cap (B + s)| \ge \tau\}| \le \frac{D|A||B|^2}{\tau^3}.$$

For $Q, R \subset \mathbb{C}$ and integer $t \ge 1$, let

$$\mathbf{Sym}_t^\times(Q, R) = \{x : |Q \cap xR^{-1}| \ge t\},$$

where $R^{-1} = \{1/a : a \in R\}$. We are also interested in the following characteristic of $A$. Let

$$d_*(A) = \min_{t \ge 1} \min_{\emptyset \ne Q, R \subset \mathbb{C}\setminus\{0\}} \frac{|Q|^2|R|^2}{|A|t^3}, \tag{4.8}$$

where the second minimum in (4.8) is taken over $Q, R$ such that $A \subset \mathbf{Sym}_t^\times(Q, R)$ and $|A| \le \max\{|Q|, |R|\}$.

The following appears as Lemma 13 in [52], where it was proved for finite subsets of the real numbers. As before, we note that the proof works as is for subsets of complex numbers (aside from the use of Theorem 4.3) and leave out the proof in this thesis.

**Lemma 4.12.** *Let* $A \subset \mathbb{C}$ *be a finite set. Then* $A$ *is SzT-type with parameter* $O(d_*(A))$.

Finally, we will require the following result of Shkredov [67]. Again, while the result in [67] is stated for subsets of real numbers, the statement holds for subsets of complex numbers without any changes in the proof.

**Theorem 4.13.** *Let* $A$ *be a finite set of complex numbers. Then*

$$|A + A| = \Omega\left(|A|^{58/37} D(A)^{-21/37}\right). \tag{4.9}$$

### Proof of Theorem 1.23

We are now ready to prove Theorem 1.23. For convenience, we state it again.

**Theorem 1.23.** *Let $A \subset \mathbb{C} \setminus \{0\}$ be a finite set. Then for any $c < 5/9813$ we have*[2]

$$\max\{|AA|, |A+A|\} = \Omega^*\left(|A|^{4/3+c}\right).$$

*Proof.* We may assume that the elements of $A$ are contained in an angular sector around the origin. More specifically, for a sufficiently small absolute constant $\sigma > 0$ (small enough to satisfy the conditions of Lemma 4.10), $|\arg(a)| < \sigma$ for every $a \in A$. Otherwise, by the pigeonhole principle, there exists a subset $A' \subseteq A$ such for for every $a_1, a_2 \in A'$, $|\arg(a_1) - \arg(a_2)| < \sigma$. There exists a complex number $z$ such that elements of $zA'$ have the required property. Note that it suffices to prove the statement of the theorem for the set $zA'$, since $|A'| \geq |A|/\sigma$, $A'A' \subset AA$, $A' + A' \subset A + A$, and the statement of the theorem is invariant under dilation. For the rest of the proof, to simplify notation, we assume that $A$ satisfies the required property.

We may assume that $|A+A| = O(|A|^{3/4+c})$. Suppose that $|A+A|^2|AA| \leq L|A|^4$ and that $|AA|^3 \leq L'|A|^4$ for some parameters $L, L' \in \mathbb{R}$. We now show that $\max\{L, L'\} = \Omega(|A|^{3c})$, which will complete the proof.

From Lemma 4.11, there exists a real number $\tau > E^\times(A)/(2|A|^2)$ and a set $S'_\tau \subset S_\tau \subset A/A$ such that $|S_\tau|\tau^2 = \Omega(E^\times(A)/\lg|A|)$, $|S'_\tau| \geq |S_\tau|/2$. Furthermore, for every $\lambda \in S'_\tau$,

$$t := |A_\lambda A_\lambda| = \Omega^*\left(\tau^2/L^{16}\right). \tag{4.10}$$

The Katz-Koester inclusion (see [46]) states that for any $\lambda \in A/A$,

$$A_\lambda A_\lambda \subset (AA) \cap (\lambda AA).$$

It follows that for any $\lambda \in S'_\tau$, $|(AA) \cap (\lambda AA)| \geq |A_\lambda A_\lambda| \geq t$. In particular, $S'_\tau \subseteq \mathbf{Sym}_t^\times(AA, AA)$. Since $S'_\tau \subset S_\tau$, we have

$$\sum_{a \in A} |A \cap aS'_\tau| = \sum_{\lambda \in S'_\tau} |A \cap \lambda A| \geq \tau |S'_\tau|.$$

By the pigeonhole principle, there exists $a \in A$ such that the set $A' = A \cap aS'_\tau$ satisfies $|A'| \geq \tau|S'_\tau|/|A|$. From $S'_\tau \subset \mathbf{Sym}_t^\times(AA, AA)$, we observe that $A' \subset \mathbf{Sym}_t^\times(AA, AA)$.

[2] *Recall that the $O^*(\cdot)$-notation hides sub-polynomial factors.*

Setting $Q = aAA$ and $R = AA$ in equation (4.8) gives

$$d_*(A') \leq \frac{|AA|^4}{|A'|t^3} = O\left(\frac{|AA|^4 L^{48}}{|A'|\tau^6}\right) = O\left(\frac{|AA|^4|A|L^{48}}{|S'_\tau|\tau^7}\right). \tag{4.11}$$

We are now set up to finish the proof, the remainder of which uses the lemmas and observations stated so far. From Theorem 4.13 and equation (4.11) we obtain

$$\begin{aligned}|A+A| \geq |A'+A'| &= \Omega\left(|A'|^{58/37} d_*(A')^{-21/37}\right) \\ &= \Omega\left(\left(\frac{\tau|S'_\tau|}{|A|}\right)^{58/37}\left(\frac{|S'_\tau|\tau^7}{L^{48}|A||AA|^4}\right)^{21/37}\right) = \Omega\left(\frac{|S'_\tau|^{79/37}\tau^{205/37}}{|A|^{79/37}L^{1008/37}|AA|^{84/37}}\right).\end{aligned}$$

Recall that $|S'_\tau| \geq S_\tau/2$, that $|S_\tau| = \Omega^*(E^\times(A)/\tau^2)$ and that $\tau > E^\times(A)/(2|A|^2)$. It follows that

$$|A+A| = \Omega^*\left(\frac{E^\times(A)^{79/37}\tau^{47/37}}{|A|^{79/37}L^{1008/37}|AA|^{84/37}}\right) = \Omega^*\left(\frac{E^\times(A)^{126/37}}{|A|^{173/37}L^{1008/37}|AA|^{84/37}}\right).$$

Combining this with (4.4) and the assumption that $|AA|^3 \leq L'|A|^4$ yields

$$|A+A| = \Omega^*\left(\frac{|A|^{331/37}}{r^{1008/37}|AA|^{210/37}}\right) = \Omega^*\left(\frac{|A|^{51/37}}{L^{1008/37}L'^{70/37}}\right). \tag{4.12}$$

It is now straightforward to check that $\max\{L, L'\} = \Omega(|A|^{3c})$. $\square$

# Chapter 5

# Conclusion

In Chapter 2 of this thesis, we presented upper bounds on the number of point-plane and point-sphere incidences in $\mathbb{R}^3$. These results were based on the polynomial method introduced by Guth and Katz [39] in their solution of the Erdős distinct distances problem. When dealing with incidences in $\mathbb{R}^3$, certain complications arise. Specifically, all (or a large fraction) of the points and spheres could be contained in the zero set of the partitioning polynomial. In general, the surface could be arbitrarily complicated, and thus, hard to analyze. Our results are based on the following two techniques:

- The use of a constant-degree partitioning polynomial as introduced by Solymosi and Tao [69]. This reduced the complexity of the partitioning surface and simplified the analysis at the cost of an $\varepsilon$ factor in the exponent.

- Finding a second partitioning polynomial co-prime with the first polynomial, an idea introduced independently by Zahl [77] and Kaplan, Matoušek, Safernová and Sharir [44]. The intersection of the two partitioning polynomials is now one-dimensional, simplifying the analysis considerably.

Much is still unknown about these techniques. One direction of future research is to seek incidence bounds in higher dimensions. Perhaps the biggest challenge here is to prove the existence of a sequence of partitioning polynomials, each of bounded degree, and such that the polynomials do not share any common factors. Some progress was made in this direction by Basu and Sombra [16] who proved the existence of a third partitioning polynomial. It is likely that in $\mathbb{R}^d$, one should be able to find $d-1$ partitioning polynomials such that the intersection of these polynomials is one dimensional. The existence of such polynomials has so far proven to be difficult to establish.

In most cases, there are large gaps between the known upper and lower bounds for incidence problems. Some examples are: (i) point-circle incidences in $\mathbb{R}^2$; (ii) point-sphere incidences in $\mathbb{R}^3$; (iii) point-plane incidences in $\mathbb{R}^3$. Finding better bounds for incidences is a problem of considerable interest and they would have many interesting implications. However, it is likely that this would require the introduction of new techniques.

In Chapter 3, we used matrix scaling techniques that were introduced by Barak, Dvir, Wigderson and Yehudayoff [12] in order to establish bounds on the number of ordinary lines determined by point sets in complex space. These results helped make progress towards understanding the structure of lines determined by point sets, but there is still much to learn. For example, as far as we are aware, no non-trivial upper bounds for the number of ordinary lines are known in $\mathbb{R}^3$. More specifically, it is not known whether a sufficiently non-degenerate point set in $\mathbb{C}^3$ or $\mathbb{R}^3$ (i.e., a set with at most a fraction of the points contained in a plane) can determine $o(n^2)$ ordinary lines. We are currently exploring this question and hope to make progress on it.

The matrix scaling technique has been shown to have other applications in combinatorial geometry. Dvir and Gopi [28], for example, established connections to incidence problems, Dvir, Garg, Oliveira and Solymosi [27] showed applications to structural (or graph) rigidity, and there are many other examples as well. The matrix scaling technique is still in its infancy, and we believe there are many as yet undiscovered connections. It is a promising direction of research to explore and further develop these techniques.

# References

[1] P. Agarwal, R. Apfelbaum, G. Purdy, and M. Sharir. Similar simplices in a $d$-dimensional point set. In *Proceedings of the 23rd Annual Symposium on Computational Geometry*, pages 232–238. ACM, 2007.

[2] P. Agarwal, E. Nevo, J. Pach, R. Pinchasi, M. Sharir, and S. Smorodinsky. Lenses in arrangements of pseudo-circles and their applications. *Journal of the ACM (JACM)*, 51(2):139–186, 2004.

[3] A. Ai, Z. Dvir, S. Saraf, and A. Wigderson. Sylvester-Gallai type theorems for approximate collinearity. In *Forum of Mathematics, Sigma*, volume 2. Cambridge Univ Press, 2014.

[4] T. Akutsu, H. Tamaki, and T. Tokuyama. Distribution of distances and triangles in a point set and algorithms for computing the largest common point sets. *Discrete & Computational Geometry*, 20(3):307–331, 1998.

[5] N. Alon. Tools from higher algebra. *Handbook of Combinatorics*, 2:1749–1783, 1995.

[6] N. Alon. Perturbed identity matrices have high rank: Proof and applications. *Combinatorics, Probability and Computing*, 18(1–2):3–15, 2009.

[7] R. Apfelbaum and M. Sharir. Large complete bipartite subgraphs in incidence graphs of points and hyperplanes. *SIAM Journal on Discrete Mathematics*, 21(3):707–725, 2007.

[8] R. Apfelbaum and M. Sharir. Non-degenerate spheres in three dimensions. *Combinatorics, Probability and Computing*, 20(04):503–512, 2011.

[9] B. Aronov, V. Koltun, and M. Sharir. Incidences between points and circles in three and higher dimensions. *Discrete & Computational Geometry*, 33(2):185–206, 2005.

[10] B. Aronov and M. Sharir. Cutting circles into pseudo-segments and improved bounds for incidences. *Discrete & Computational Geometry*, 28(4):475–490, 2002.

[11] A. Balog and O. Roche-Newton. New sum-product estimates for real and complex numbers. *Discrete & Computational Geometry*, 53(4):825–846, 2015.

[12] B. Barak, Z. Dvir, A. Wigderson, and A. Yehudayoff. Fractional Sylvester-Gallai theorems. *Proceedings of the National Academy of Sciences*, 110(48):19213–19219, 2013.

[13] A. Basit, Z. Dvir, S. Saraf, and C. Wolf. On the number of ordinary lines determined by sets in complex space. *Proceedings of the 33rd Annual Symposium on Computational Geometry*, 2017.

[14] A. Basit, B. Lund, and A. Sheffer. Improved sum-product bounds in $\mathbb{C}$. *In preparation*, 2016.

[15] A. Basit and A. Sheffer. Incidences with $k$-non-degenerate sets and their applications. *Journal of Computational Geometry*, 5(1):284–302, 2014.

[16] S. Basu and M. Sombra. Polynomial partitioning on varieties of codimension two and point-hypersurface incidences in four dimensions. *Discrete & Computational Geometry*, 55(1):158–184, 2016.

[17] R. Bojanowski. Zastosowania uogólnionej nierówności Bogomolova-Miyaoka-Yau. Master Thesis (in Polish). `http://www.mimuw.edu.pl/~alan/postscript/bojanowski.ps`, 2003.

[18] P. Borwein and W. Moser. A survey of Sylvester's problem and its generalizations. *Aequationes Mathematicae*, 40(1):111–135, 1990.

[19] J. Bourgain. More on the sum-product phenomenon in prime fields and its applications. *International Journal of Number Theory*, 1(01):1–32, 2005.

[20] J. Bourgain, N. Katz, and T. Tao. A sum-product estimate in finite fields, and applications. *Geometric & Functional Analysis GAFA*, 14(1):27–57, 2004.

[21] P. Brass and C. Knauer. On counting point-hyperplane incidences. *Computational Geometry*, 25(1-2):13–20, 2003.

[22] R. K. Brayton, D. Coppersmith, and A. J. Hoffman. Self-orthogonal latin squares of all orders $n \neq 2, 3, 6$. *Bulletin of the American Mathematical Society*, 80, 1974.

[23] F. Chung. Sphere-and-point incidence relations in high dimensions with applications to unit distances and furthest-neighbor pairs. *Discrete & Computational Geometry*, 4(2):183–190, 1989.

[24] D. Cox, J. Little, and D. O'Shea. Ideals, varieties, and algorithms: An introduction to computational algebraic geometry and commutative algebra, 1992.

[25] J. Csima and E. Sawyer. There exist $6n/13$ ordinary points. *Discrete & Computational Geometry*, 9(2):187–202, 1993.

[26] M. De Berg, M. Van Kreveld, M. Overmars, and O. Schwarzkopf. *Computational Geometry*. Springer, 2000.

[27] Z. Dvir, A. Garg, R. Oliveira, and J. Solymosi. Rank bounds for design matrices with block entries and geometric applications. *arXiv preprint arXiv:1610.08923*, 2016.

[28] Z. Dvir and S. Gopi. On the number of rich lines in truly high dimensional sets. In *Proceedings of the 31st Annual Symposium on Computational Geometry*, pages 584–598, 2014.

[29] Z. Dvir, S. Saraf, and A. Wigderson. Improved rank bounds for design matrices and a new proof of Kelly's theorem. In *Forum of Mathematics, Sigma*, volume 2. Cambridge Univ Press, 2014.

[30] H. Edelsbrunner, L. Guibas, and M. Sharir. The complexity of many cells in arrangements of planes and related problems. *Discrete & Computational Geometry*, 5(2):197–216, 1990.

[31] G. Elekes. On the number of sums and products. *Acta Arithmetica*, 81(4):365–367, 1997.

[32] G. Elekes and C. Tóth. Incidences of not-too-degenerate hyperplanes. In *Proceedings of the 21st Annual Symposium on Computational Geometry*, pages 16–21. ACM, 2005.

[33] N. Elkies, L. Pretorius, and K. Swanepoel. Sylvester-Gallai theorems for complex numbers and quaternions. *Discrete & Computational Geometry*, 35(3):361–373, 2006.

[34] P. Erdős. On sets of distances of $n$ points. *The American Mathematical Monthly*, 53(5):248–250, 1946.

[35] P. Erdős and E. Szemerédi. On sums and products of integers. In *Studies in Pure Mathematics*, pages 213–218. Springer, 1983.

[36] T. Gallai. Solution of problem 4065. *American Mathematical Monthly*, 51:169–171, 1944.

[37] B. Green and T. Tao. On sets defining few ordinary lines. *Discrete & Computational Geometry*, 50(2):409–468, 2013.

[38] L. Guth and N. Katz. Algebraic methods in discrete analogs of the Kakeya problem. *Advances in Mathematics*, 225(5):2828–2839, 2010.

[39] L. Guth and N. Katz. On the Erdős distinct distances problem in the plane. *Annals of Mathematics*, 181(1):155–190, 2015.

[40] A. Harnack. Über Vieltheiligkeit der ebenen algebraischen Curven. *Mathematische Annalen*, 10(2):189–198, 1876.

[41] R. Hartshorne. *Geometry: Euclid and beyond.* Springer Science & Business Media, 2000.

[42] A. Hilton. On double diagonal and cross Latin squares. *Journal of the London Mathematical Society*, 2(4):679–689, 1973.

[43] F. Hirzebruch. Arrangements of lines and algebraic surfaces. In *Arithmetic and geometry*, pages 113–140. Springer, 1983.

[44] H. Kaplan, J. Matoušek, Z. Safernová, and M. Sharir. Unit distances in three dimensions. *Combinatorics, Probability and Computing*, 21(04):597–610, 2012.

[45] H. Kaplan, J. Matoušek, and M. Sharir. Simple proofs of classical theorems in discrete geometry via the Guth-Katz polynomial partitioning technique. *Discrete & Computational Geometry*, 48(3):499–517, 2012.

[46] N. Katz and P. Koester. On additive doubling and energy. *SIAM Journal on Discrete Mathematics*, 24(4):1684–1693, 2010.

[47] N. Kayal and S. Saraf. Blackbox polynomial identity testing for depth 3 circuits. In *Proceedings of the 50th Annual Symposium on Foundations of Computer Science*, pages 198–207. IEEE, 2009.

[48] L. Kelly. A resolution of the Sylvester-Gallai problem of J.-P. Serre. *Discrete & Computational Geometry*, 1(1):101–104, 1986.

[49] L. Kelly and W. Moser. On the number of ordinary lines determined by $n$ points. *Canadian Journal of Mathematics*, 10:210–219, 1958.

[50] S. Konyagin and M. Rudnev. On new sum-product-type estimates. *SIAM Journal on Discrete Mathematics*, 27(2):973–990, 2013.

[51] S. Konyagin and I. Shkredov. On sum sets of sets having small product set. *Proceedings of the Steklov Institute of Mathematics*, 290(1):288–299, 2015.

[52] S. Konyagin and I. Shkredov. New results on sum-products in $\mathbb{R}$. *arXiv preprint arXiv:1602.03473*, 2016.

[53] B. Lund. An improved bound on (A plus A)/(A plus A). *Electronic Journal of Combinatorics*, 23(3), 2016.

[54] J. Matoušek. *Lectures on Discrete Geometry*, volume 212. Springer Science & Business Media, 2002.

[55] E. Melchior. Über Vielseite der projektiven Ebene. *Deutsche Math*, 5:461–475, 1940.

[56] J Milnor. On the Betti numbers of real varieties. *Proceedings of the American Mathematical Society*, 15(2):275–280, 1964.

[57] Th. Motzkin. The lines and planes connecting the points of a finite set. *Transactions of the American Mathematical Society*, pages 451–464, 1951.

[58] J. Pach and M. Sharir. Geometric incidences. *Contemporary Mathematics*, 342:185–224, 2004.

[59] O. Roche-Newton. If $(A+A)/(A+A)$ is small, then the ratio set is large. *Journal of the London Mathematical Society*, 93(1):83–100, 2016.

[60] U. Rothblum and H. Schneider. Scalings of matrices which have prespecified row sums and column sums via optimization. *Linear Algebra and its Applications*, 114:737–764, 1989.

[61] N. Saxena and C. Seshadhri. From Sylvester-Gallai configurations to rank bounds: Improved blackbox identity test for depth–3 circuits. *Journal of the ACM (JACM)*, 60(5):33, 2013.

[62] J.-P. Serre. Advanced problem 5359. *American Mathematical Monthly*, 73(1):89, 1966.

[63] M. Sharir, A. Sheffer, and J. Solymosi. Distinct distances on two lines. *Journal of Combinatorial Theory, Series A*, 120(7):1732–1736, 2013.

[64] M. Sharir, A. Sheffer, and J. Zahl. Improved bounds for incidences between points and circles. *Combinatorics, Probability and Computing*, 24(03):490–520, 2015.

[65] A. Sheffer. The Konyagin-Shkredov sum-product bound. `https://adamsheffer.files.wordpress.com/2016/07/ks-sp.pdf`, 2016.

[66] A. Sheffer, E. Szabó, and J. Zahl. Point-curve incidences in the complex plane. *arXiv preprint arXiv:1502.07003*, 2015.

[67] I. Shkredov. On sums of Szemerédi-Trotter sets. *Proceedings of the Steklov Institute of Mathematics*, 289(1):300–309, 2015.

[68] J. Solymosi. Bounding multiplicative energy by the sumset. *Advances in mathematics*, 222(2):402–408, 2009.

[69] J. Solymosi and T. Tao. An incidence theorem in higher dimensions. *Discrete & Computational Geometry*, 48(2):255–280, 2012.

[70] J. Solymosi, G. Tardos, and C. Tóth. The $k$ most frequent distances in the plane. *Discrete & Computational Geometry*, 28(4):639–648, 2002.

[71] J. J. Sylvester. Mathematical question 11851. *Educational Times*, 59(98):256, 1893.

[72] E. Szemerédi and W. Trotter. Extremal problems in discrete geometry. *Combinatorica*, 3(3–4):381–392, 1983.

[73] R. Thom. Sur l'homologie des variétés algebriques réelles. *Differential and Combinatorial Toplogy, SS Cairns, ed.*, pages 255–265, 1965.

[74] C. Tóth. The Szemerédi-Trotter theorem in the complex plane. *Combinatorica*, 35(1):95–126, 2015.

[75] H. Warren. Lower bounds for approximation by nonlinear manifolds. *Transactions of the American Mathematical Society*, 133(1):167–178, 1968.

[76] T. Wolff. Recent work connected with the Kakeya problem. *Prospects in Mathematics (Princeton, NJ, 1996)*, 2:129–162, 1999.

[77] J. Zahl. An improved bound on the number of point-surface incidences in three dimensions. *Contributions to Discrete Mathematics*, 8(1), 2013.

[78] J. Zahl. A Szemerédi-Trotter type theorem in $\mathbb{R}^4$. *Discrete & Computational Geometry*, 54(3):513–572, 2015.