The largest complete bipartite subgraph in point-hyperplane incidence graphs

Thao Do

Akuna Capital
Chicago, Illinois, U.S.A.

dothithuthao11191@gmail.com

Submitted: Oct 20, 2018; Accepted: Oct 21, 2019; Published: Jan 10, 2020

Abstract

Given m points and n hyperplanes in $\mathbb{R}^d$ $(d \geqslant 3)$, if there are many incidences, we expect to find a big cluster $K_{r,s}$ in their incidence graph. Apfelbaum and Sharir [1] found lower and upper bounds for the largest size of rs, which match (up to a constant) only in three dimensions. In this paper we close the gap in four and five dimensions, up to some polylogarithmic factors.

Mathematics Subject Classifications: 05C35, 52C10, 05D10

1 Introduction

Throughout this paper, let d denote an integer at least three. Given a set P of m points and a set Q of n hyperplanes in $\mathbb{R}^d$, their *incidence graph* $G(P, Q)$ is a bipartite graph with vertex set $P \cup Q$ and $(p, q) \in P \times Q$ forms an edge iff $p \in q$. It is proved in [1] that if this graph does not contain $K_{r,s}$ (a complete bipartie graph with two parts A, B, $|A| = r, |B| = s$ and (a, b) is an edge for every $a \in A, b \in B$) as a subgraph for some fixed integers $r, s \geqslant 2$, then it can have at most $O_d((mn)^{d/(d+1)} + m + n)$ edges. Here the notation $f = O_d(g)$ means there exists some constant C that depends on d such that $f \leqslant Cg$. The number of incidences between P and Q, denoted by $I(P, Q)$ is the number of edges of $G(P, Q)$.

Conversely, when the incidence graph has many edges, we expect to find a big subgraph isomorphic to $K_{r,s}$. How big can rs (the number of edges of $K_{r,s}$) be in terms of m, n and the number of edges of the incidence graph? To make the question precise, we use the following definition:

Definition 1. Given a set P of points and Q of hyperplanes in $\mathbb{R}^d$, let $rs(P, Q)$ be the maximum size of a complete bipartite subgraph of its incidence graph, and $rs_d(m, n, I)$

be the minimum of this quantity over all choices of m points and n hyperplanes in $\mathbb{R}^d$ with I incidences. To be precise:

$$rs(P,Q) := \max\{rs : K_{r,s} \subset G(P,Q)\}$$

$$rs_d(m,n,I) := \min_{|P|=m, |Q|=n, |I(P,Q)|=I} rs(P,Q).$$

We are interested in how big $rs_d(m,n,I)$ can be in terms of m, n and I. Apfelbaum and Sharir [1] gave a satisfactory answer to this question when $d=3$.

Theorem 2 (Apfelbaum and Sharir [1]). *The following statements hold true in three dimensions.*

1. *If $I = \Omega(mn^{1/2} + nm^{1/2})$, then $rs_3(m,n,I) = \frac{I^2}{mn} - \Theta(m+n)$.*
2. *If $m \leqslant n$, $I = O(nm^{1/2})$ and $I = \Omega(m^{3/4}n^{3/4})$, then $rs_3(m,n,I) = \Theta(\frac{I^4}{m^2n^3} + \frac{I}{m})$.*
3. *Symmetrically, if $n \leqslant m$, $I = O(mn^{1/2})$ and $I = \Omega(m^{3/4}n^{3/4})$, then $rs_3(m,n,I) = \Theta(\frac{I^4}{n^2m^3} + \frac{I}{n})$.*
4. *If $I = O(m^{3/4}n^{3/4} + m + n)$, then $rs_3(m,n,I) = \Theta(\frac{I}{m} + \frac{I}{n})$.*

However, much less is known in higher dimensions. One thing we know is that $rs_d(m,n,I) \geqslant \max\{\frac{I}{m}, \frac{I}{n}\}$ by looking at the star subgraphs centering at the point and the hyperplane with maximum degrees in $G(P,Q)$. As noted in [1], when the dimension d increases beyond 3, there are progressively more ranges of I (as a function of m and n) where the bounds for rs_d change qualitatively. At one extreme, when I is small enough, we expect the graph not to contain any big complete bipartite subgraph. Indeed, if $I = O_{d,\varepsilon}((mn)^{1-\frac{2}{d+2}-\varepsilon})$ for some $\varepsilon > 0$ and d is odd, Brass and Knauer [3] constructed an example of point-hyperplane incidence graph in $\mathbb{R}^d$ with I incidences and no $K_{t,t}$ for some fixed integer $t \geqslant 2$. A similar result holds if d is even, and $I = O_{d,\varepsilon}((mn)^{1-\frac{2(d+1)}{(d+2)^2}-\varepsilon})$. These bounds have been slightly improved in [2]. At another extreme, when I is very large, we expect to find a large $K_{r,s}$.

Theorem 3 (Apfelbaum and Sharir [1]). *For any $d \geqslant 3$, if $I = \Omega_d(mn^{1-\frac{1}{d-1}} + nm^{1-\frac{1}{d-1}})$, then*

$$rs_d(m,n,I) = \Omega_d\left(\left(\frac{I}{mn}\right)^{d-1} mn\right). \tag{1}$$

Morever, if $I = \Omega_d((mn)^{1-\frac{1}{d-1}})$, then

$$rs_d(m,n,I) = O_d\left(\left(\frac{I}{mn}\right)^{\frac{d+1}{2}} mn\right). \tag{2}$$

These lower and upper bounds only match (up to a constant) when $d = 3$ (which is why we have a tight bound in part 1 of Theorem 2). In this paper we close the gap (up to polylogarithmic factors) in four and five dimensions for this range of I. More specifically we prove the following two results.

Theorem 4. *When $d = 4$, there exist positve constants C_4 and C'_4 such that if $I \geqslant C_4(mn^{2/3} + nm^{3/5})$, then*

$$rs_4(m, n, I) \geqslant C'_4 \left(\frac{I}{mn}\right)^{5/2} mn(\log mn)^{-4}.$$

Theorem 5. *When $d = 5$, there exist positive constants C_5 and C'_5 such that if $I \geqslant C_5(mn^{3/4} + nm^{2/3})$, then*

$$rs_5(m, n, I) \geqslant C'_5 \left(\frac{I}{mn}\right)^{3} mn(\log mn)^{-10}.$$

The main tool used to prove Theorem 4 and Theorem 5 is an incidence bound between points and *nondegenerate* hyperplanes by Elekes and Tóth [5], which is reviewed in the next section. We then present the proof of Theorem 4 and sketch the proof of Theorem 5 in the subsequent sections. At the end we explain why our method does not work in six dimensions.

Acknowledgements

The author would like to thank Larry Guth and the referees for many helpful feedback.

2 Incidences with nondegenerate hyperplanes

We use the following notation. Let A and B be two sets of geometric objects in $\mathbb{R}^d$. Their *incidence graph* $G(A, B)$ is a bipartite graph on $A \times B$, where (a, b) forms an edge iff $a \subset b$. The number of incidences between A and B, denoted by $I(A, B)$, is the number of edges of this graph. In this paper, A is either a set of points or a set of lines, and B is a set of higher dimensional *flats*. An affine d'-dimensional flat, or a d'-flat is a subset of $\mathbb{R}^d$ that is congruent to $\mathbb{R}^{d'}$ for some integer $0 \leqslant d' \leqslant d$. Points, lines, planes and hyperplanes are flats of dimensions 0, 1, 2 and $d-1$ respectively. Given a flat F_1, F_2 is a subflat of F_1 if it is a flat and a subset of F_1; it is a proper subflat if $F_2 \subsetneq F_1$.

Given a set $\mathcal{S}$ of m points in $\mathbb{R}^d$ and some $\beta \in (0, 1)$, an affine hyperplane H is *β-degenerate* with respect to (w.r.t.) $\mathcal{S}$ if there exists a proper subflat $F \subset H$ that contains more than β fraction of the number of points of $\mathcal{S}$ in H, i.e. $|F \cap \mathcal{S}| > \beta |H \cap \mathcal{S}|$. Otherwise, H is *β-nondegenerate*. Elekes and Tóth proved the following incidence bound.

Theorem 6 (Elekes-Tóth [5]). *If $\mathcal{S}$ is a set of m points and $\mathcal{H}$ is a set of n β-nondegenerate hyperplanes w.r.t. $\mathcal{S}$ (for any $0 < \beta < 1$)*[1] *in $\mathbb{R}^d$ (for any integer $d \geqslant 2$), then there exists a constant $C_{\beta,d} > 0$ such that*

$$I(\mathcal{S}, \mathcal{H}) \leqslant C_{\beta,d} \left((mn)^{\frac{d}{d+1}} + mn^{1-\frac{1}{d-1}} \right). \tag{3}$$

This implies the maximum number of β-nondegenerate, k-rich (i.e. containing at least k points of $\mathcal{S}$) hyperplanes is $O_{\beta,d}\left(\frac{m^{d+1}}{k^{d+2}} + \frac{m^{d-1}}{k^{d-1}}\right)$. Elekes-Tóth in fact proved the second statement; it is shown to be equivalent to (3) in [1]. When $d = 2$, it reduces to the well known Szemerédi-Trotter point-line incidence bound [7].

Since points and hyperplanes are dual to each other, we also have a dual version of the above result. Given a set $\mathcal{H}$ of n hyperplanes in $\mathbb{R}^d$, a point p is β-nondengenerate with respect to (w.r.t.) $\mathcal{H}$ if there does not exist a line ℓ such that $\#\{H \in \mathcal{H} : \ell \subset H\} \geqslant \beta\#\{H \in \mathcal{H} : p \in H\}$.

Corollary 7. *If $\mathcal{H}$ is a set of n hyperplanes in $\mathbb{R}^d$ and P is a set of m β-nondegenerate points w.r.t. $\mathcal{H}$ (for any $d \geqslant 2, 0 < \beta < 1$), there exists a constant $C_{\beta,d} > 0$ such that*

$$I(\mathcal{S}, \mathcal{H}) \leqslant C_{\beta,d} \left((mn)^{\frac{d}{d+1}} + nm^{1-\frac{1}{d-1}} \right). \tag{4}$$

Equivalently, given n hyperplanes in $\mathbb{R}^d$, the number of k-rich, β-nondegenerate points is $O_{\beta,d}\left(\frac{n^{d+1}}{k^{d+2}} + \frac{n^{d-1}}{k^{d-1}}\right)$.

3 Proof in four dimensions

We first outline our strategy. Let $\mathcal{S}$ be a set of m points, and $\mathcal{H}$ be a set of n hyperplanes in $\mathbb{R}^4$. There are two ways to form a big $K_{r,s}$ in the incidence graph $G(\mathcal{H}, \mathcal{S})$: either a plane contains many points of $\mathcal{S}$ and belongs to many hyperplanes of $\mathcal{H}$, or a line does. By an averaging argument, we can assume that each hyperplane is $\Omega(\frac{I}{m})$-rich (i.e. contains at least $\Omega(\frac{I}{m})$ points of $\mathcal{S}$). By Theorem 6, the contribution from β-nondegenerate hyperplanes is negligible, so we can assume that each hyperplane is β-degenerate, i.e. it contains some plane with at least β portion of the total number of points in that plane, hence the plane is $\Omega(\beta\frac{I}{m})$-rich. In this case, we say each hyperplane *degenerates* to a rich plane. Either one of those planes belongs to many hyperplanes, which would form a big $K_{r,s}$, or we can find a subset $\mathcal{P}_i$ of planes such that $I(\mathcal{S}, \mathcal{P}_i)$ is large. We then repeat our argument: using the averaging argument and Corollary 7, we can assume that each point in $\mathcal{S}$ belongs to many planes in $\mathcal{P}_i$ and degenerates to a line. Either one of those lines contains many points, which then form a big $K_{r,s}$, or we can find a subset $\mathcal{L}_j$ of lines such that $I(\mathcal{L}_j, \mathcal{P}_i)$ is large. But after some transformation, this number is the same as the number of incidences between points and lines in $\mathbb{R}^2$ and hence cannot be too large by Theorem 6 for $d = 2$, or equivalently, Szemerédi and Trotter's theorem in [7].

We now give a detailed proof.

[1] Elekes and Tóth actually proved this only for $\beta < \beta_d$ for some small β_d. It is later shown in [4] that we can take $\beta_d = \frac{1}{d-1}$ and in [6] that we can take $\beta_d = 1$.

Proof of Theorem 4. Assume for contradiction that there exist a set $\mathcal{S}$ of m points and a set $\mathcal{H}$ of n hyperplanes in $\mathbb{R}^4$ with I incidences where $I \geqslant C_4(mn^{2/3} + nm^{3/5})$ and the incidence graph $G(\mathcal{S}, \mathcal{H})$ contains no $K_{r,s}$ where $rs \geqslant C_4' \left(\frac{I}{mn}\right)^{5/2} mn(\log mn)^{-4}$. We shall choose the suitable positive constants C_4 and C_4' to derive a contradiction.

Step 1: We can assume that each hyperplane is $\frac{I}{4n}$-rich and β-degenerate with respect to $\mathcal{S}$ for some $\beta \in (0, 1)$, say $\beta = 1/2$.

Indeed, remove all the hyperplanes that contain fewer than $\frac{I}{4n}$ points and the hyperplanes that are β-nondegenerate. The number of incidences from the non-rich hyperplanes is at most $n\frac{I}{4n} = \frac{I}{4}$. By Theorem 6, the number of incidences from the β-nondegenerate hyperplanes is at most $C_{\beta,4}((mn)^{4/5}+mn^{2/3}) < \frac{C_4}{4}(mn^{2/3}+nm^{3/5})$ if we choose $C_4 > 8C_{\beta,4}$. Indeed, this only fails if $(mn)^{4/5} \geqslant mn^{2/3}$ and $(mn)^{4/5} \geqslant nm^{3/5}$, which are equivalent to $m \leqslant n^{2/3}$ and $n \leqslant m$, but those two inequalities cannot hold at the same time. Therefore, after the removal, there remain at least $\frac{I}{2}$ incidences. Assume that there remain n_1 hyperplanes for some $n_1 \leqslant n$. Throughout the proof, there are many inequalities that involve n_1, but we can always use n to upper bound n_1 in the correct direction. Therefore, without loss of generality, we can assume that $n_1 = n$.

Step 2: For each $\frac{I}{4n}$-rich β-degerenate hyperplane H, we can find a proper subflat $P \subset H$ so that $|P| \geqslant \beta|H| \geqslant \frac{\beta I}{4n}$. Since $\dim(H) = 3$, we can assume that P is a plane. Let $\mathcal{P}$ denote the set of these planes. We claim that no plane in $\mathcal{P}$ belongs to more than s_0 hyperplanes in $\mathcal{H}$, where

$$s_0 := \frac{c_1 I^{3/2}}{m^{3/2}n^{1/2}(\log mn)^4} \tag{5}$$

where c_1 is a sufficiently small constant to be specified later. Indeed, assume that there are at least s_0 hyperplanes that degenerate to a same plane, then we have a configuration of $K_{r,s}$ with $r = \frac{\beta I}{4n}$ and $s = s_0$. This leads to a contradiction if we choose $C_4' < \frac{\beta c_1}{4}$:

$$rs \geqslant \frac{\beta I}{4n} \cdot \frac{c_1 I^{3/2}}{m^{3/2}n^{1/2}(\log mn)^4} \geqslant C_4' \left(\frac{I}{mn}\right)^{5/2} mn(\log mn)^{-4}.$$

Step 3: We use a dyadic decomposition to find a subset of planes with many incidences with $\mathcal{S}$. Let $\mathcal{P}_j$ denote the set of all planes that are assigned to at least 2^j and fewer than 2^{j+1} hyperplanes, for $1 \leqslant j < \log s_0 < \log n$ (here the logarithm is in base 2). The contribution to incidences from the planes must be at least β fraction of the number of incidences from the β-degenerate hyperplanes, which implies $\sum_{j=0}^{\log s_0} 2^{j+1} I(\mathcal{S}, \mathcal{P}_j) \geqslant \frac{\beta}{4} I$. Hence there must exist some i such that

$$I' := I(\mathcal{S}, \mathcal{P}_i) \geqslant \frac{\beta I}{4 \log s_0 2^{i+1}} \geqslant \frac{c_2 I}{2^i \log n}, \tag{6}$$

where $c_2 = \beta/8$. We claim that the following holds where β is the same as before, and $C_{\beta,3}$ is defined in Theorem 6:

$$I' = I(\mathcal{S}, \mathcal{P}_i) > 4C_{\beta,3} \left[(|\mathcal{P}_i||\mathcal{S}|)^{\frac{3}{4}} + |\mathcal{P}_i||\mathcal{S}|^{\frac{1}{2}}\right]. \tag{7}$$

Indeed, assume otherwise. By (6), $I \leqslant c_2^{-1} 2^i I' \log n \leqslant c_3 2^i \log n \left[(|\mathcal{P}_i||\mathcal{S}|)^{\frac{3}{4}} + |\mathcal{P}_i||\mathcal{S}|^{\frac{1}{2}}\right]$, where $c_3 = c_2^{-1} 4 C_{\beta,3}$. We shall now derive a contradiction using two facts: $|\mathcal{P}_i| \leqslant \frac{n}{2^i}$ and $2^i \leqslant s_0$. The first fact follows from $\sum_{j=0}^{\log s_0} 2^j |\mathcal{P}_j| \leqslant n$, since each hyperplane is assigned to exactly one plane. Using the formula for s_0 in (5), we have:

$$\begin{aligned} I &\leqslant c_3 2^i \log n \left[(|\mathcal{P}_i||\mathcal{S}|)^{\frac{3}{4}} + |\mathcal{P}_i||\mathcal{S}|^{\frac{1}{2}}\right] \\ &\leqslant c_3 \log n 2^i \left(\left(\frac{nm}{2^i}\right)^{3/4} + \frac{n}{2^i} m^{1/2}\right) \\ &\leqslant c_3 \log n (mn)^{3/4} s_0^{1/4} + c_3 (\log n) n m^{1/2} \\ &\leqslant c_3 \log n (mn)^{3/4} \left(\frac{c_1 I^{3/2}}{m^{3/2} n^{1/2} (\log mn)^4}\right)^{1/4} + c_3 (\log n) n m^{1/2}. \end{aligned} \tag{8}$$

By our assumption, $I \geqslant C_4 n m^{3/5}$, thus the second term on the right hand side of (8), $c_3 (\log n) n m^{1/2}$, is less than $\frac{I}{2}$ for large values of m, n. This implies the first term of (8) must be at least $\frac{I}{2}$. Rearranging we get

$$\begin{aligned} I^{5/8} &\leqslant 2 c_3 c_1^{1/4} m^{3/8} n^{5/8} \frac{\log n}{\log mn} \\ &\leqslant (c_4 n m^{3/5})^{5/8} \end{aligned}$$

where $c_4 = (2 c_3 c_1^{1/4})^{8/5}$. However, we can choose c_1 small enough and C_4 big enough so that $c_4 < C_4$ and hence this contradicts $I \geqslant C_4 n m^{3/5}$. So (7) must hold.

Step 4: Let us consider $G(S, \mathcal{P}_i)$, the incidence graph between the points S and the set of planes $\mathcal{P}_i$ from Step 3. We can use an averaging argument similar to Step 1 to assume that each point in S is $\frac{I'}{4m}$-rich (i.e. belongs to at least $\frac{I'}{4m}$ planes in $\mathcal{P}_i$). Since the bound in (7) is the same as that in Corollary 7, we can also assume each point in S is β-degenerate w.r.t. $\mathcal{P}_i$ (in the sense defined before Corollary 7). Each such point degenerates to a line that is $\beta\frac{I'}{4m}$-rich. Let $\mathcal{L}$ denote the set of all these lines. We claim that no line in $\mathcal{L}$ contains more than r_0 points, where

$$r_0 := \frac{c_5 I^{3/2}}{m^{1/2} n^{3/2} (\log mn)^3} \tag{9}$$

for some small enough positive constant c_5 to be chosen later. Indeed, since each plane in $\mathcal{P}_i$ belongs to at least 2^i hyperplanes in $\mathcal{H}$ by definition, each line in $\mathcal{L}$ belongs to at least $\frac{\beta I'}{4m}$ planes in $\mathcal{P}_i$, and thus belongs to at least $\frac{\beta I' 2^i}{4m} \geqslant \frac{\beta c_2 I}{4m \log n}$ hyperplanes in $\mathcal{H}$ by (6). If there are at least r_0 points that degenerate (or belong) to a same line, then we have a configuration of $K_{r,s}$ where

$$rs \geqslant r_0 \cdot \frac{\beta c_2 I}{4m \log n}$$

$$= \frac{c_5 I^{3/2}}{m^{1/2}n^{3/2}(\log mn)^3} \cdot \frac{\beta c_2 I}{4m\log n}$$
$$\geqslant \frac{c_2c_5\beta}{4}\left(\frac{I}{mn}\right)^{5/2} mn(\log mn)^{-4}.$$

This contradicts our assumption if we choose $C_4' < c_2c_5\beta/4$.

Step 5: Similar to Step 3, we use a dyadic decomposition to find a subset of lines in $\mathcal{L}$ that form many incidences with $\mathcal{P}_i$. Here we say a line ℓ is incident to a plane P if $\ell \subset P$. Let $\mathcal{L}_k$ denote the set of all lines that contain at least 2^k and fewer than 2^{k+1} points for $1 \leqslant k < \log r_0 < \log m$. The contribution to I' from the lines must be at least β fraction, which implies $\sum_{k=0}^{\log r_0} 2^{k+1} I(\mathcal{L}_k, \mathcal{P}_i) \geqslant \frac{\beta}{4} I'$. Hence there must exist some j such that

$$I'' := I(\mathcal{L}_j, \mathcal{P}_i) \geqslant \frac{\beta I'}{8\log r_0 2^j} \geqslant \frac{\beta^2 I}{64\log m \log n 2^{i+j}} \geqslant \frac{c_6 I}{2^{i+j}\log m \log n}, \tag{10}$$

where $c_6 = \beta^2/64$. We claim that the following holds where C_{ST} is the constant in Szemerédi and Trotter's theorem [7]:

$$I'' = I(\mathcal{L}_j, \mathcal{P}_i) \geqslant C_{ST}\left(|\mathcal{P}_i|^{2/3}|\mathcal{L}_j|^{2/3} + |\mathcal{P}_i| + |\mathcal{L}_j|\right). \tag{11}$$

Indeed, assume otherwise. By (10):

$$I \leqslant c_6^{-1} 2^{i+j} I'' \log m \log n \leqslant c_7 2^{i+j}\left(|\mathcal{P}_i|^{2/3}|\mathcal{L}_j|^{2/3} + |\mathcal{P}_i| + |\mathcal{L}_j|\right)\log m \log n$$

where $c_7 = c_6^{-1} C_{S-T} = 64C_{ST}/\beta^2$. We make use of the following four facts: $|\mathcal{P}_i| \leqslant \frac{n}{2^i}$, $|\mathcal{S}_j| \leqslant \frac{m}{2^j}$, $2^i \leqslant s_0$ and $2^j \leqslant r_0$. We already showed $|\mathcal{P}_i| \leqslant \frac{n}{2^i}$. The second fact holds for a similar reason: since each point is assigned to exactly one line we have $\sum_{k=0}^{\log r_0} 2^k |\mathcal{L}_k| \leqslant m$, and thus $|\mathcal{L}_j| \leqslant \frac{m}{2^j}$. We now write:

$$I \leqslant c_7(\log m \log n) 2^{i+j}\left[\left(\frac{n}{2^i}\frac{m}{2^j}\right)^{2/3} + \frac{n}{2^i} + \frac{m}{2^j}\right]$$
$$\leqslant c_7(\log m \log n)\left((2^i 2^j)^{1/3}(mn)^{2/3} + n2^j + m2^i\right)$$
$$\leqslant c_7(\log m \log n)\left((s_0 r_0)^{1/3}(mn)^{2/3} + nr_0 + ms_0\right)$$

Using the formula for s_0 and r_0 in (5) and (9), we have:

$$I \leqslant c_7(\log m \log n)\left[(mn)^{2/3}\left(\frac{c_1 I^{3/2}}{m^{3/2}n^{1/2}(\log mn)^4}\frac{c_5 I^{3/2}}{m^{1/2}n^{3/2}(\log mn)^3}\right)^{1/3}\right.$$
$$\left. + m\frac{c_1 I^{3/2}}{m^{3/2}n^{1/2}(\log mn)^4} + n\frac{c_5 I^{3/2}}{m^{1/2}n^{3/2}(\log mn)^3}\right]$$
$$\leqslant I\left[\frac{(c_1c_5)^{1/3}c_7 \log m \log n}{(\log mn)^{7/3}} + \left(\frac{I}{mn}\right)^{1/2}\left(\frac{c_1 \log m \log n}{(\log mn)^4} + \frac{c_5 \log m \log n}{(\log mn)^3}\right)\right].$$

We can choose c_1 and c_5 small enough so that

$$\frac{(c_1c_5)^{1/3}c_7\log m\log n}{(\log mn)^{7/3}} < \frac{1}{2}$$

and

$$\left(\frac{I}{mn}\right)^{1/2}\left(\frac{c_1\log m\log n}{(\log mn)^4}+\frac{c_5\log m\log n}{(\log mn)^3}\right) < \frac{1}{2}$$

for large values of m, n. In this case, the right hand side of the last inequality is strictly less than I, a contradiction. So (11) must hold.

Step 6: Project the set of planes $\mathcal{P}_i$ and the set of lines $\mathcal{L}_j$ to a generic three dimensional subspace, then intersect them with a generic plane $\mathcal{P}i$ within this subspace. After this transformation, $\mathcal{P}_i$ becomes a set of lines P^* and $\mathcal{L}_j$ becomes a set of points L^* in $\mathcal{P}i$. By (11), $I(P^*, L^*) = I(\mathcal{L}_j, \mathcal{P}_i) > C_{ST}(|P^*|^{2/3}|L^*|^{2/3} + |P^*| + |L^*|)$, which violates the Szemerédi-Trotter theorem [7]. This gives the desired contradiction and finishes our proof. □

4 Sketch of the proof in five dimensions

The proof method is the same as that in four dimensions, but the exponents are different and the method is repeated one more time. In particular, in the previous section, we *unwrap* a point-hyperplane configuration in $\mathbb{R}^4$ with many incidences in two layers: hyperplanes degenerate to planes and points degenerate to lines. At each layer, either we can find a big $K_{r,s}$ subgraph, or the number of incidences remain larger than the nondegenerate bound in Theorem 6, and we can keep unwrapping. In $\mathbb{R}^5$, we unwrap in three layers: hyperplanes degenerate to 3-flats, points degenerate to lines, and 3-flats degenerate to planes. The detailed proof is quite similar to that in the four dimensions case, so we only give an outline here. For simplicity, we ignore the polylogarithmic factors and write $f \gtrsim g$ (or $f \lesssim g$) to indicate there exists some constants $a, b > 0$ such that $f \geqslant a(\log mn)^b g$ (or $f \leqslant a(\log mn)^b g$.

Proof sketch of Theorem 5. Prove by contradiction. Let $\mathcal{S}$ denote the set of m points and $\mathcal{H}$ denote the set of n hyperplanes in $\mathbb{R}^5$. Assume that $I(\mathcal{S}, \mathcal{H}) \gtrsim (mn^{3/4} + nm^{2/3})$ but their incidence graph does not contain any $K_{r,s}$ subgraph where $rs \gtrsim \left(\frac{I}{mn}\right)^3 mn$.

Step 1: We can assume that every hyperplane is $\frac{I}{n}$-rich, and β-degenerate with respect to $\mathcal{S}$ for some $\beta \in (0, 1)$. The choice of β is quite flexible, so we can assume $\beta = 1/2$.

Step 2: For each such hyperplane H, we can find a 3-dimensional flat (or a 3-flat) F such that $F \subset H$ and $|F \cap \mathcal{S}| \geqslant \beta|H \cap \mathcal{S}| \geqslant \frac{\beta I}{n}$. Let $\mathcal{F}$ denote the set of these 3-flats. Using our assumption on rs, no flat in $\mathcal{F}$ belongs to more than s_0 hyperplanes where $s_0 \leqslant c_1 \frac{I^2}{m^2 n(\log mn)^{10}}$ for some sufficiently small positive constant c_1 to be chosen later.

Step 3: Let $\mathcal{F}_j$ denote the set of all 3-flats in $\mathcal{F}$ that are assigned to at least 2^j and fewer than 2^{j+1} hyperplanes where $j \leqslant \log s_0 < \log n$. Then there exists an i such that $I(\mathcal{F}_i, \mathcal{S}) \gtrsim \frac{I}{2^i}$. We show that

$$I' := I(\mathcal{F}_i, \mathcal{S}) \gtrsim (|\mathcal{F}_i||\mathcal{S}|)^{4/5} + |\mathcal{F}_i||\mathcal{S}|^{2/3}.$$

Indeed, assume otherwise. Using $I' \gtrsim 2^i I$, $|\mathcal{F}_i| \leqslant \frac{n}{2^i}$ and $2^i \leqslant s_0 \lesssim \frac{I^2}{m^2 n}$, we have

$$I \lesssim 2^i I' \lesssim 2^i \left[\left(\frac{nm}{2^i}\right)^{4/5} + \frac{n}{2^i} m^{2/3}\right] \lesssim (mn)^{4/5} \left(\frac{I^2}{m^2 n}\right)^{1/5} + nm^{2/3},$$

which cannot happen given our condition $I \gtrsim mn^{3/4} + nm^{2/3}$.

Step 4: Since I' is large, using Corollary 7, we can assume that each point in $\mathcal{S}$ is $\frac{I'}{m}$-rich (i.e. belongs to at least $\frac{I'}{m}$ flats in $\mathcal{F}_i$, and is β-degenerate w.r.t. $\mathcal{F}_i$. Each such point degenerates to a $\frac{\beta I'}{m}$-rich line. Let $\mathcal{L}$ denote that set of these lines. Then no line in $\mathcal{L}$ can contain more than r_0 points where $r_0 \leqslant c_2 \frac{I^2}{mn^2(\log mn)^{10}}$ for some sufficiently small positive constant c_2 to be chosen later

Step 5: We use a dyadic decomposition to find a subset of lines with many incidences with $\mathcal{F}_i$. Let $\mathcal{L}_k$ denote the set of all lines in $\mathcal{L}$ that contain more than 2^k and fewer than 2^{k+1} points. Then there exists a j such that $I(\mathcal{F}_i, \mathcal{L}_j) \gtrsim \frac{I'}{2^j} \gtrsim \frac{I}{2^{i+j}}$. We show that

$$I'' := I(\mathcal{F}_i, \mathcal{L}_j) \gtrsim |\mathcal{F}_i|^{3/4}|\mathcal{L}_j|^{3/4} + |\mathcal{F}_i||\mathcal{L}_j|^{1/2}.$$

Indeed, assume otherwise. Using $I'' \gtrsim I/2^{i+j}$, $|\mathcal{F}_i| \leqslant \frac{n}{2^i}$, $|\mathcal{L}_j| \leqslant \frac{m}{2^j}$, $2^i \leqslant s_0 \lesssim \frac{I^2}{m^2 n}$ and $2^j \leqslant r_0 \lesssim \frac{I^2}{mn^2}$, we have

$$\begin{aligned} I &\lesssim 2^{i+j} I'' \\ &\lesssim 2^{i+j} \left[\left(\frac{mn}{2^{i+j}}\right)^{3/4} + \frac{n}{2^i}\left(\frac{m}{2^j}\right)^{1/2}\right] \\ &\lesssim (mn)^{3/4} \left(\frac{I^2}{m^2 n} \frac{I^2}{mn^2}\right)^{1/4} + nm^{1/2} \left(\frac{I^2}{mn^2}\right)^{1/2} \\ &= 2I \end{aligned}$$

This cannot happen with an appropriate choice of c_1, c_2 and logarithmic factors.

Step 6: Project the set of 3-dim flats $\mathcal{F}_i$ and the set of lines $\mathcal{L}_j$ into a generic four dimensional subspace, then intersect them with a generic 3-dimensional flat in this subspace. After this transformation, $\mathcal{F}_i$ becomes a set of planes and $\mathcal{L}_j$ becomes a set of points in $\mathbb{R}^3$. Because of the inequality in the previous step, we can assume that each 3-flats in $\mathcal{F}_i$ degenerate to a plane. Let $\mathcal{P}$ denote the set of all such planes. Then no plane belongs to more than t_0 flats in $\mathcal{F}_i$ where $t_0 \leqslant c_3 \frac{I^2}{m^2 n(\log mn)^{10}}$ for some sufficiently small positive constant c_3 to be chosen later.

Step 7: Using a dyadic decomposition, there exists some subset $\mathcal{P}_k$ of planes, each belongs to at least 2^k and fewer than 2^{k+1} 3-flats in $\mathcal{F}_i$ such that $I''' := I(\mathcal{L}_j, \mathcal{P}_k) \gtrsim |\mathcal{P}_k|^{2/3}|\mathcal{L}_j|^{2/3} + |\mathcal{P}_k| + |\mathcal{L}_j|$.

Step 8: Project the set of planes $\mathcal{P}_k$ and the set of lines $\mathcal{L}_j$ into a generic three dimensional subspace, then intersect them with a generic plane in this subspace. After this transformation, $\mathcal{P}_k$ becomes a set of lines and $\mathcal{L}_j$ becomes a set of points in $\mathbb{R}^2$. Hence, the inequality on $I(\mathcal{L}_j, \mathcal{P}_k)$ in Step 7 violates Szemerédi and Trotter's theorem for appropriate choices of c_1, c_2, c_3 and logarithmic factors. This gives the desired contradiction and finishes our proof. □

5 Discussion

We first compare our approach with that of Apfelbaum and Sharir in [1]. Their proof of (1) relies on the incidence bound (3) to *unwrap* the point-hyperplane configuration with many incidences: for each $k = d-1, d-2, \ldots, 2$, either we can find a big $K_{r,s}$ subgraph involving a $(k-1)$-flat (i.e. a flat that contains at least r points and belongs to at least s hyperplanes) or the rich k-flats degenerate to rich $(k-1)$-flats. In this paper we obtain a stronger result by combining (3) with its dual incidence bound (4) to unwrap the point-hyperplane configuration from both directions.

Our approach can be used to obtain an improved lower bound in six and higher dimensions, but is not good enough to match the upper bound (2). To understand why that is the case, let us revisit the construction that attains this upper bound. We start with a $(d-2)$-dimensional rectangular integer grid, which we will denote by $\mathcal{G}$, with many rich 'hyperplanes' in the first $d-2$ coordinates of $\mathbb{R}^d$; note that these rich hyperplanes are of dimension $d-3$. Extend the configuration to the $(d-1)^{st}$ coordinate of $\mathbb{R}^d$ so that the points of $\mathcal{G}$ become parallel lines $\mathcal{L}$ and the hyperplanes of $\mathcal{G}$ become parallel $(d-2)$-dimensional flats $\mathcal{F}$. Our point set is obtained by putting an equal number of points in each line in $\mathcal{L}$ and our hyperplane set is obtained by extending each flat in $\mathcal{F}$ to an equal number of hyperplanes in an arbitrary direction.

In this case, every hyperplane is 1-degenerate (since it has a $(d-2)$-dimensional subflat which contains the same number of points), and so is every point. Therefore, our argument in the first two layers is not wasteful. However, the further layers are no longer β-degenerate for any fixed constant $\beta \in (0,1)$. As shown in [5], the second term of (3) – which dominates the estimate in dimensions at least 4 – is no longer tight when β is not fixed. One potential fix is to obtain an incidence bound with explicit dependence of β that is stronger than (6) when $\beta = o(1)$.

References

[1] 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.

[2] M. Balko, J. Cibulka and P. Valtr. Covering Lattice Points by Subspaces and Counting Point–Hyperplane Incidences. *Discrete & Computational Geometry*, 61(2): 325–354, 2019.

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

[4] T. Do. Extending Erdős -Beck's theorem to higher dimensions, arXiv:1607.00048.

[5] G. Elekes and C. Tóth. Incidences of not-too-degenerate hyperplanes, *Proceedings of the twenty-first annual symposium on Computational geometry*, pages 16–21, 2005.

[6] B. Lund. Two theorems on point-flat incidences, arXiv:1708.00039, 2017.

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