Point-hyperplane incidences via extremal graph theory

Aleksa Milojević[*], Benny Sudakov[*], István Tomon[†]

Abstract

The study of counting point-hyperplane incidences in the d-dimensional space was initiated in the 1990's by Chazelle and became one of the central problems in discrete geometry. It has interesting connections to many other topics, such as additive combinatorics and theoretical computer science. Assuming a standard nondegeneracy condition, i.e., that no s points are contained in the intersection of s hyperplanes, the currently best known upper bound on the number of incidences of n points and m hyperplanes in $\mathbb{R}^d$ is

$$O_{d,s}((mn)^{1-1/(d+1)}+m+n).$$

This bound by Apfelbaum and Sharir is based on geometrical space partitioning techniques.

In this paper, we propose a novel combinatorial approach to estimate the number of point-hyperplane incidences over arbitrary fields using forbidden induced patterns in incidence graphs. Perhaps surprisingly, this approach matches the best known bounds in $\mathbb{R}^d$ for many interesting values of m, n, d, e.g. when $m = n$ and d is odd. Moreover, in finite fields our bounds are sharp as a function of m and n in every dimension. We also study the size of the largest complete bipartite graph in point-hyperplane incidence graphs with a given number of edges and obtain optimal bounds as well.

1 Introduction

The Szemerédi-Trotter theorem [30] is a fundamental result in combinatorial geometry, giving a sharp upper bound on the number of incidences in point-line configurations. It states that n points and m lines on the real plane determine at most $O((mn)^{2/3}+m+n)$ incidences, and as Erdős showed (see e.g. [11]), this bound is the best possible. This deep result found numerous applications and inspired a large number of generalizations and extensions over the last decades. For example, the problem of counting point-line incidences in $\mathbb{R}^3$ under certain non-degeneracy conditions has been studied by Guth and Katz [17] in their solution to the Erdős distinct distances problem. The Szemerédi-Trotter theorem has also been applied by Elekes [12] to derive sum-product estimates over the reals. For these and other applications of incidence theorems, we refer the reader to the survey [9].

Extending incidence results to finite fields is more tricky, since standard space partitioning techniques do not apply anymore. The importance of incidence bounds in $\mathbb{F}_p^2$ stems from their close connection to sum-product estimates, as shown in the pioneering work of Bourgain, Katz and Tao [4]. Following this work, point-line incidence bounds over finite fields have been extensively studied, see e.g. [18, 22, 27].

In this paper, we consider one of the most natural incidence questions in higher dimensions, which is to find the maximum number of incidences between n points and m hyperplanes in $\mathbb{F}^d$, for an arbitrary field $\mathbb{F}$. In dimension 3 and above, it is possible that all m hyperplanes contain all n points: take n points on a line, and m hyperplanes containing this line. In order to avoid such trivialities, it is standard to impose the further condition that the *incidence graph* is $K_{s,s}$-free, where $K_{s,s}$ denotes the complete bipartite graph with vertex classes of size s. Here, we think of s as a large constant, which may depend on d, but no other parameters. Given a set of points $\mathcal{P}$ and set of hyperplanes $\mathcal{H}$, the *incidence graph* of $(\mathcal{P},\mathcal{H})$ is the bipartite graph $G(\mathcal{P},\mathcal{H})$ with vertex classes $\mathcal{P}$ and $\mathcal{H}$, and $\{x, H\} \subset \mathcal{P}\cup\mathcal{H}$ is an edge if $x \in H$. We denote by $I(\mathcal{P},\mathcal{H})$ the number of edges of the incidence graph, i.e. $I(\mathcal{P},\mathcal{H})$ is the number of incidences between $\mathcal{P}$ and $\mathcal{H}$.

[*]ETH Zurich, e-mail: **{aleksa.milojevic, benjamin.sudakov}@math.ethz.ch**. Research supported in part by SNSF grant 200021_196965.

[†]Umeå University, *e-mail*: **istvan.tomon@umu.se**.

1.1 Incidence bounds in higher dimensions

The problem of bounding the number of edges in $K_{s,s}$-free incidence graphs has a long history, with the initial motivation coming from computational geometry. How compactly can one record the incidences in a configuration of points and hyperplanes? Instead of writing down whether each of the mn pairs $(x, H) \in \mathcal{P} \times \mathcal{H}$ forms an incidence, one can compress the incidence graph by representing it as the union of few complete balanced bipartite subgraphs. Although this cannot be done very efficiently in general graphs (i.e., one needs $\Omega(\frac{n^2}{\log n})$ complete balanced bipartite subgraphs to represent a random balanced bipartite graph on $2n$ vertices with high probability), for incidence graphs Chazelle [6] developed a space partitioning technique which can be used to show that $O_{d,s}(n^{2-\frac{2}{d+1}} \log n)$ complete bipartite graphs suffice. Another motivation comes from the counting version of Hopcroft's problem, which asks, given n points and m hyperplanes in $\mathbb{R}^d$, how fast can one determine the number of incidences. The work of Chazelle [6] achieved the first subquadratic algorithm for this problem and Erickson [13] gave a lower bound for the running time of algorithms classified as partitioning algorithms. He showed that particularly hard instances for determining the number of incidences are configurations $(\mathcal{P}, \mathcal{H})$ without a compact representation, i.e. a configuration without a large complete bipartite graph in the incidence graph.

Improving the works of Chazelle [6] and Brass and Knauer [5], Apfelbaum and Sharir [2] proved that if $\mathcal{P}$ is a set of n points and $\mathcal{H}$ is a set of m hyperplanes in $\mathbb{R}^d$, whose incidence graph contains no $K_{s,s}$, then

$$I(\mathcal{P}, \mathcal{H}) \leq O_{d,s}\left((mn)^{1-\frac{1}{d+1}} + m + n\right). \tag{1}$$

This upper bound has not been improved in the past twenty years, however, matching lower bounds are only known for $d = 2$, which coincides with the Szemerédi-Trotter theorem. The currently best known lower bound for $d \geq 3$ was recently achieved by Sudakov and Tomon [29], improving on the constructions of Brass and Knauer [5] and Balko, Cibulka, and Valtr [3]: if s is sufficiently large with respect to d, then there exist a set of n points $\mathcal{P}$ and a set of m hyperplanes $\mathcal{H}$ in $\mathbb{R}^d$ whose incidence graph is $K_{s,s}$-free and

$$I(\mathcal{P}, \mathcal{H}) \geq \begin{cases} \Omega_d\left((mn)^{1-(2d+3)/(d+2)(d+3)}\right) & \text{if } d \text{ is odd,} \\ \Omega_d\left((mn)^{1-(2d^2+d-2)/(d+2)(d^2+2d-2)}\right) & \text{if } d \text{ is even.} \end{cases}$$

Generalizations of the upper bound in (1) are also established (with an extra $o(1)$ term in the exponent) for incidences of points and semi-algebraic sets in $\mathbb{R}^d$, see [14]. The common feature in all of the proofs of (1) and related results is that they rely on certain space partitioning results such as cuttings [6, 7] or polynomial partitioning [17, 26]. These techniques are highly geometric and have no analogues over finite fields.

Another natural question asks about the maximum number of incidences between points and hyperplanes in a d-dimensional space $\mathbb{F}^d$ for any field $\mathbb{F}$. If $d = 2$ and p is a prime, taking every point of $\mathbb{F}_p^2$ and $m < p^2$ arbitrary lines, we get $n = p^2$ points, and $mp = \Theta(mn^{1/2})$ incidences, beating the Szemerédi-Trotter bound as long as $m \gg p$. On the other hand, a simple application of the Kővári-Sós-Turán theorem [19] shows that one cannot have more than $O(mn^{1/2} + n)$ incidences over any field, by noting that the incidence graph of points and lines is always $K_{2,2}$-free. As we mentioned already, the work of Bourgain, Katz, and Tao [4] shows that better upper bounds can be obtained over $\mathbb{F}_p$ assuming $m, n \leq p^{2-\delta}$.

However, point-hyperplane incidence bounds in higher dimensions turn out to be more elusive. One such result in three dimensions was obtained by Rudnev [24] (see also de Zeeuw [31] for a shorter proof). He proved that in $\mathbb{F}_p^3$, if no s points are on a line, then there are at most $O(m\sqrt{n} + ms)$ incidences, assuming $n \leq p^2$. This point-plane incidence bound in $\mathbb{F}_p^3$ was later used to show various improved sum-product estimates over finite fields in [1, 22, 23]. For $d > 3$, barely anything is known about incidences of points and hyperplanes in $\mathbb{F}_p^d$.

1.2 Sparse incidence graphs in vector spaces

In this paper, we propose a novel approach to derive bounds on the maximal number of incidences between n points and m hyperplanes in a d-dimensional vector space $\mathbb{F}^d$, assuming the incidence graph is $K_{s,s}$-free. We use almost purely graph-theoretic techniques and prove bounds which are sharp for the whole range of parameters m and n with a suitable choice of field. Surprisingly, for many interesting pairs of values (m, n), our upper bound matches the upper bound in (1). This is fairly unexpected, as all proofs of (1) rely on highly geometric techniques, while we employ only combinatorial ideas. To give a snippet of our most general result, we prove the following in the special case $m = n$.

Theorem 1.1. *Let d, s be positive integers and let $\mathbb{F}$ be a field. If $\mathcal{P}$ is a set of n points and $\mathcal{H}$ is a set of n hyperplanes in $\mathbb{F}^d$ such that no s points lie on the intersection of s hyperplanes, then*

$$I(\mathcal{P}, \mathcal{H}) \leq O\left((s + d^3)n^{2-1/\lceil\frac{d+1}{2}\rceil}\right).$$

Note that this bound matches the best known bound (1) when d is odd. Before stating the general bound, which depends on the relationship of m and n, let us introduce a piece of notation. Namely, we define the quantities $\alpha_t = \frac{t}{d+2-t}$ for $t \in \{2, \dots, d\}$ and $\beta_t = \frac{t}{d+1-t}$ for $t \in \{1, \dots, d\}$. Now we can state our main theorem in its full generality.

Theorem 1.2. *Let d, s, m, n be positive integers and $\alpha > 0$ such that $m = n^\alpha$, and let $\mathbb{F}$ be a field. Let $\mathcal{P}$ be a set of n points and $\mathcal{H}$ be a set of m hyperplanes in $\mathbb{F}^d$ such that the incidence graph of $(\mathcal{P}, \mathcal{H})$ is $K_{s,s}$-free. Then*

$$I(\mathcal{P}, \mathcal{H}) \leq \begin{cases} O_{s,d}(n) & \text{if } \alpha \in [0, \beta_1], \\ O_{s,d}(m^{1-\frac{1}{t}}n) & \text{if } \alpha \in [\alpha_t, \beta_t] \text{ for some } t \in \{2, \dots, d\}, \\ O_{s,d}(mn^{1-\frac{1}{d+2-t}}) & \text{if } \alpha \in [\beta_{t-1}, \alpha_t] \text{ for some } t \in \{2, \dots, d\}, \\ O_{s,d}(m) & \text{if } \alpha \in [\beta_d, \infty). \end{cases}$$

Moreover, these bounds are tight, i.e. for every d, there exist $s = s(d)$ and $c = c(d)$ such that the following hold. For every $\alpha > 0$ and sufficiently large integer n, there exist a field $\mathbb{F} = \mathbb{F}(d, n, \alpha)$, a set of n points $\mathcal{P}$ and a set of $m = \lfloor n^\alpha \rfloor$ hyperplanes $\mathcal{H}$ in $\mathbb{F}^d$ such that the incidence graph of $(\mathcal{P}, \mathcal{H})$ is $K_{s,s}$-free and

$$I(\mathcal{P}, \mathcal{H}) \geq \begin{cases} n & \text{if } \alpha \in [0, \beta_1], \\ cm^{1-\frac{1}{t}}n & \text{if } \alpha \in [\alpha_t, \beta_t] \text{ for some } t \in \{2, \dots, d\}, \\ cmn^{1-\frac{1}{d+2-t}} & \text{if } \alpha \in [\beta_{t-1}, \alpha_t] \text{ for some } t \in \{2, \dots, d\}, \\ m & \text{if } \alpha \in [\beta_d, \infty). \end{cases}$$

Let us compare our result with (1) in the non-trivial regime $\alpha \in [1/d, d]$. Interestingly, in case $\alpha = \beta_t = \frac{t}{d+1-t}$ for some integer $t \in \{1, \dots, d\}$, Theorem 1.2 shows that the number of incidences is $O_{s,d}((mn)^{1-1/(d+1)})$, which exactly matches (1). The other extreme is when $\alpha = \alpha_t = \frac{t}{d+2-t}$ for some $t \in \{2, \dots, d\}$, in which case we get $O_{s,d}((mn)^{1-1/(d+2)})$. For other values of α, our upper bound is $O_{s,d}((mn)^{1-\gamma})$ for some $\frac{1}{d+2} \leq \gamma \leq \frac{1}{d+1}$. However, as the second half of the theorem shows, these upper bounds cannot be improved unless further assumptions on the field $\mathbb{F}$ are made.

1.3 Large complete bipartite graphs in dense incidence graphs

We also study another well-known problem of finding large complete bipartite graphs in dense incidence graphs. The motivation behind studying this problem is the intuitive understanding that a large number of incidences between points and hyperplanes is always explained by the existence of large complete bipartite subgraphs in the incidence graph. To make things more precise, given a set of points $\mathcal{P}$ and set of hyperplanes $\mathcal{H}$, we define $\mathrm{rs}(\mathcal{P}, \mathcal{H})$ as the maximum number of edges in a complete bipartite subgraph of the incidence graph of $(\mathcal{P},\mathcal{H})$, i.e. the maximum of $r\cdot s$ over all r and s such that $K_{r,s}$ is a subgraph of the incidence graph. We consider configurations of n points and m hyperplanes with εmn incidences.

In the real d-dimensional space, Apfelbaum and Sharir [2] proved that $\mathrm{rs}(\mathcal{P},\mathcal{H})=\Omega_d(\varepsilon^{d-1}mn)$ for $\varepsilon>\Omega(n^{-1/(d-1)})$. Also, if $\varepsilon>\Omega((mn)^{-\frac{1}{d-1}})$, then there exist a set of points $\mathcal{P}$ and set of hyperplanes $\mathcal{H}$ in $\mathbb{R}^d$ with $\mathrm{rs}(\mathcal{P},\mathcal{H})\leq O_d(\varepsilon^{\frac{d+1}{2}}mn)$. Note that in case $d=3$, the lower and upper bounds match for $\varepsilon>\Omega(n^{-1/2})$. Do [8] improved the lower bound in $\mathbb{R}^4$ and $\mathbb{R}^5$ for a large range of m,n,ε to match the upper bound up to logarithmic factors. One might be also interested how the minimum of $\mathrm{rs}(\mathcal{P},\mathcal{H})$ depends on the dimension d as well, which is a question motivated by certain variants of the celebrated log-rank conjecture of Lovász and Saks [20]. To this end, it follows from Fox, Pach, and Suk [15] that $\mathrm{rs}(\mathcal{P},\mathcal{H})>\varepsilon^{d+1}2^{-O(d\log d)}mn$. For constant ε, this is improved by a recent result of Singer and Sudan [25], who obtained an exponential dependence on d in the lower bound $\mathrm{rs}(\mathcal{P},\mathcal{H})=\Omega(\varepsilon^{2d}mn/d)$. We refer the interested reader to [25] about the relationship of this problem and the log-rank conjecture.

Here, we extend and improve some of these results by considering arbitrary fields, and obtain tight bounds in all cases. Following Apfelbaum and Sharir [2], we define $\mathrm{rs}_d(n,m,I)=\min\mathrm{rs}(\mathcal{P},\mathcal{H})$, where the minimum is taken over all fields $\mathbb{F}$, all sets $\mathcal{P}\subset\mathbb{F}^d$ of at most n points, all sets $\mathcal{H}$ of at most m hyperplanes in $\mathbb{F}^d$, that satisfy $I(\mathcal{P},\mathcal{H})\geq I$. Note that requiring that $\mathcal{P}$ is a set of at most n points instead of exactly n points makes no crucial difference, since one can always add points to the set $\mathcal{P}$ which are incident to no hyperplanes (if $\mathbb{F}$ is a finite field, one might need to pass to an extension of $\mathbb{F}$). A similar remark holds for hyperplanes.

Theorem 1.3. *Let d and $m\geq n$ be positive integers, $\varepsilon\in(0,1)$ and $I=\varepsilon mn$. Then*

$$\mathrm{rs}_d(m,n,I),\mathrm{rs}_d(n,m,I)=\begin{cases}\Theta_d(\varepsilon^{d-1}mn) & \text{if } \varepsilon>100\max\{m^{-\frac{1}{d}},n^{-\frac{1}{d-1}}\},\\ \Theta_d(\varepsilon m) & \text{if } \varepsilon<\frac{1}{4}\max\{m^{-\frac{1}{d}},n^{-\frac{1}{d-1}}\}.\end{cases}$$

In the regime $\varepsilon\gg\max\{m^{-\frac{1}{d}},n^{-\frac{1}{d-1}}\}$, the implied constant of the lower bound is only exponential in d. Therefore, our theorem matches the dependence on ε coming from the result of Apfelbaum and Sharir [2], as well as the exponential dependence on the dimension obtained by Sudan and Singer [25].

On the other hand, in case $\varepsilon\ll\max\{m^{-\frac{1}{d}},n^{-\frac{1}{d-1}}\}$, Theorem 1.3 tells us that the largest complete bipartite subgraph might be just a star. Finally, observe that in case $m=n$, there is a large jump around $\epsilon\approx n^{-1/d}$, where $\mathrm{rs}_d(n,n,I)$ suddenly jumps from $\approx n^{1-1/d}$ to $\approx n^{1+1/d}$.

1.4 Forbidden induced patterns in incidence graphs

The key idea used to prove our theoretical bounds is that incidence graphs of points and hyperplanes in $\mathbb{F}^d$ avoid a simple family $\mathcal{F}_d$ of graphs as induced subgraph. The family $\mathcal{F}_d$ includes all bipartite graphs on vertex classes $\{a_1,\dots,a_d\}$ and $\{b_1,\dots,b_d\}$, where a_ib_j is an edge for $i\geq j-1$, and a_ib_{i+2} is a non-edge for $i=1,\dots,d-2$.

An alternative and perhaps easier way to think about the family is through the notion of a *pattern.* We define a *pattern* to be an edge labeling of a complete bipartite graph, in which every edge receives one of the labels 0, 1 or $*$. Every such pattern Π defines a family of graphs $\mathcal{F}_\Pi$ on the same vertex set, where a graph F is in the family if all edges of the pattern labeled by 0 do not appear in F, all edges labeled by 1 appear in F, while the edges labeled by $*$ may or may not appear. To simplify the terminology, we say that a graph G contains a pattern Π if it contains an induced copy of some graph in the family $\mathcal{F}_\Pi$.

The family $\mathcal{F}_d$ discussed above can be constructed in this way from the following pattern Π_d. Namely, Π_d is a pattern on a balanced bipartite graph on vertices $a_1,\dots,a_d$, $b_1,\dots,b_d$ where a_ib_j is labeled by 1 for $i\geq j-1$ and a_ib_{i+2} is labeled by 0, while all other edges are labeled by $*$. See Figure 1 for an illustration.

The following lemma is the main geometric ingredient that we use, which allows us to reduce the incidence bounds to a problem in extremal graph theory.

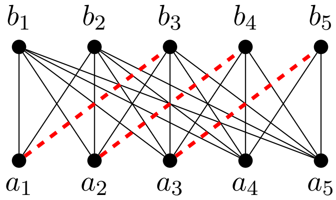

Figure 1: The pattern Π_5, where black edges are labeled 1, and the red dashed edges are labeled 0.

Lemma 1.4. *Let $d \geq 3$ be an integer and let $\mathcal{P}$ be a set of distinct points and $\mathcal{H}$ be a set of distinct hyperplanes in $\mathbb{F}^d$. Then, the incidence graph $G(\mathcal{P}, \mathcal{H})$ does not contain the pattern Π_d.*

Proof. Assume, for the sake of contradiction, that $G = G(\mathcal{P}, \mathcal{H})$ contains Π_d. Since the pattern Π_d is symmetric, we may assume that the vertices $a_1, \dots, a_d$ correspond to points $x_1, \dots, x_d \in \mathcal{P}$, and the vertices $b_1, \dots, b_d$ correspond to hyperplanes $H_1, \dots, H_d \in \mathcal{H}$. We will prove by induction that $\dim \bigcap_{i=1}^{k} H_i = d - k$ for all $2 \leq k \leq d$. This will suffice to derive the contradiction by observing that $x_{d-1}, x_d \in \bigcap_{i=1}^{d} H_i$, which is at most a 0-dimensional space, i.e. a single point. This contradicts the assumption that $x_1, \dots, x_d$ are distinct.

To show that $\dim \bigcap_{i=1}^{k} H_i = d - k$, we begin by observing that $H_1 \cap H_2$ must have dimension $d - 2$, since H_1, H_2 are distinct hyperplanes with non-empty intersection. For any $k \geq 3$, the induction hypothesis implies that $\bigcap_{i=1}^{k-1} H_i$ is an affine space of dimension $d - k + 1$, which is not a subspace of H_k since $x_{k-2} \in \bigcap_{i=1}^{k-1} H_i$ and $x_{k-2} \notin H_k$. Thus, H_k must intersect $\bigcap_{i=1}^{k-1} H_i$ in an affine subspace of dimension $\dim \bigcap_{i=1}^{k-1} H_i - 1 = d - k$. This completes the proof. □

For a linear algebraic perspective on the pattern Π_d, one might consider the $(d-2) \times (d-2)$ matrix whose rows and columns are indexed by $a_1, \dots, a_{d-2}$ and $b_3, \dots, b_d$, where the entry corresponding to a_i and b_j equals 1 if the edge $a_i b_j$ is labeled 0 and equals 0 if the edge $a_i b_j$ is labeled 1 (entries corresponding to the edges labeled by $*$ may be arbitrary). This defines an upper-triangular matrix with non-zero diagonal entries, a condition which ensures full rank. A similar family of graphs was also studied recently by Sudakov and Tomon [28] to establish Ramsey theoretic properties of algebraically defined graphs. Using the above lemma, we can just focus on the hereditary family of graphs that avoid the pattern Π_d. Then, we use the dependent random choice technique [16] and combinatorial methods to show that if G is a Π_d-free graph with sufficiently many edges, then G must contain a large complete bipartite graph.

In order to find configurations of points and hyperplanes matching our theoretical bounds, we consider optimal constructions of so called subspace evasive sets [10, 29]. We follow the ideas of [3, 5, 29] to transform large subspace evasive sets into point-hyperplane configuration with many incidences and no large bipartite subgraphs.

2 Illustration of our techniques in dimension 3

The purpose of this section is to introduce the main ideas of our paper in a relatively simple and concise way, while leaving the most general statements for later sections. We do this by focusing on the case $d = 3$, and we begin by discussing the point-plane incidence bounds in case $G(\mathcal{P}, \mathcal{H})$ does not contain $K_{s,s}$. We highlight that the forbidden pattern Π_3 is quite simple, since containing the pattern Π_3 is equivalent to containing the bipartite graph $K_{3,3}$ minus an edge as an induced subgraph.

Throughout this paper, we use standard graph theoretic notation. Given a graph G and a vertex $v \in V(G)$, we denote by $\deg v$ the degree of v, and $N(v)$ the neighbourhood of v. Also, if $v_1, \dots, v_t \in V(G)$, we denote by $N(v_1, \dots, v_t)$ the *common neighbourhood* of $v_1, \dots, v_t$, i.e. $N(v_1, \dots, v_t) = N(v_1) \cap \dots \cap N(v_t)$.

Lemma 2.1. *Let $s \geq 2$ be an integer, let $\mathcal{P}$ be a set of n points and $\mathcal{H}$ be a set of n hyperplanes in $\mathbb{F}^3$. If $G(\mathcal{P},\mathcal{H})$ does not contain $K_{s,s}$, then $I(\mathcal{P},\mathcal{H}) \leq 2sn^{3/2}$.*

Proof. Throughout the proof, we denote the bipartition of the incidence graph $G = G(\mathcal{P},\mathcal{H})$ by $A \cup B$, where A corresponds to the set of points and B to the set of hyperplanes. This will help to separate the graph-theoretic arguments from the geometric ones.

We show that if $G = G(\mathcal{P},\mathcal{H})$ satisfies $e(G) \geq 2sn^{3/2}$ and does not contain $K_{s,s}$, then it must contain the pattern Π_3. To do this, we use the method of dependent random choice. More precisely, our goal is to find four vertices $v_1, v_2 \in A, u_1, u_2 \in B$ forming a copy of $K_{2,2}$ such that the common neighbourhoods $N(v_1) \cap N(v_2)$ and $N(u_1) \cap N(u_2)$ have size at least s. Since the graph G does not contain $K_{s,s}$, there will be an edge missing between these two common neighbourhoods, which allows us to find the pattern Π_3 in G.

Let us begin by picking a random two vertex subset $\{v_1, v_2\} \subset A$ chosen from the uniform distribution on all $\binom{n}{2}$ pairs, and consider the common neighbourhood $X = N(v_1) \cap N(v_2)$ in B. The expected size of this common neighbourhood can be lower bounded by counting the number of three element sets $\{v_1, v_2, u\}$ with $v_1, v_2 \in A, u \in B$ and $v_1u, v_2u \in E(G)$ (such sets are sometimes also called as cherries). If the number of such sets is denoted by T, we have

$$\mathbb{E}\Big[|X|\Big] = \frac{T}{\binom{n}{2}} = \frac{1}{\binom{n}{2}} \sum_{u \in B} \binom{\deg u}{2} \geq \frac{2}{n-1} \binom{\frac{1}{n}\sum_{u\in B} \deg u}{2} \geq \frac{e(G)^2}{4n^3} \geq s^2,$$

where the first inequality holds by the convexity of the function $\binom{x}{2} = \frac{x(x-1)}{2}$ for $x \geq 1$. On the other hand, let us denote by Y the number of two element sets $\{u_1, u_2\} \subset X$ with common neighbourhood $N(u_1) \cap N(u_2)$ of size less than s. Then the expected value of Y can be bounded as follows. For each two element set $\{u_1, u_2\} \subset B$, we have $u_1, u_2 \in X$ if $v_1, v_2 \in N(u_1) \cap N(u_2)$. Hence, if $|N(u_1) \cap N(u_2)| < s$, then the probability of $u_1, u_2 \in X$ is less than $\binom{s}{2}/\binom{n}{2} < \left(\frac{s}{n}\right)^2$, so

$$\mathbb{E}\left[Y\right] \leq \sum_{\{u_1,u_2\}\subset B} \left(\frac{s}{n}\right)^2 \leq \binom{n}{2}\left(\frac{s}{n}\right)^2 \leq \frac{s^2}{2}.$$

Since $\mathbb{E}[|X| - Y] \geq s^2/2$, there exists a pair of vertices v_1, v_2 satisfying $|X| - Y \geq s^2/2$. This means that at least one pair $u_1, u_2 \in X$ has a common neighbourhood of size at least s. Also, since $X \geq s^2/2 \geq s$, the vertices v_1, v_2 have common neighbourhood of size at least s. As G does not contain $K_{s,s}$, there must be vertices $v_3 \in N(u_1) \cap N(u_2), u_3 \in N(v_1) \cap N(v_2)$ for which $u_3v_3 \notin E(G)$. But then $v_1, v_2, v_3, u_1, u_2, u_3$ form the pattern Π_3, which is a contradiction. This completes the proof. □

In what follows, we give the proof of Theorem 1.3 in the case $d = 3$, with an additional assumption that every point lies in many planes and every plane contains many points. This is mostly a technical assumption, which we will remove in the proof of the general case.

Lemma 2.2. *Let $\mathcal{P}$ be a set of n points and $\mathcal{H}$ be a set of $m \geq n$ hyperplanes in $\mathbb{F}^3$. Assume that $I(\mathcal{P},\mathcal{H}) \geq \varepsilon mn$ for some ε satisfying $\varepsilon \geq 6\max\{n^{-1/2}, m^{-1/3}\}$. Furthermore, assume that every point $x \in \mathcal{P}$ lies on at least $\varepsilon m/3$ hyperplanes of $\mathcal{H}$, and every hyperplane $H \in \mathcal{H}$ contains at least $\varepsilon n/3$ points of $\mathcal{P}$. Then $\mathrm{rs}(\mathcal{P},\mathcal{H}) \geq \left(\frac{\varepsilon}{6}\right)^2 mn$.*

Proof. Let A and B be the vertex classes of $G(\mathcal{P},\mathcal{H})$ with $|A| = n$ and $|B| = m$. Our proof relies on a simple observation that if $X = N(v_1) \cap N(v_2)$ for some two distinct vertices $v_1, v_2 \in A$, then any vertex $v_3 \in A$ either satisfies $|N(v_3) \cap X| \leq 1$ or $X \subseteq N(v_3)$. Suppose this is not the case and there exists a vertex $v_3 \in A$ with a non-neighbour $u_1 \in X$ and two neighbours $u_2, u_3 \in X$. Then, the vertices $v_1, v_2, v_3, u_1, u_2, u_3$ form the pattern Π_3, which is not possible due to Lemma 1.4.

Let us take vertices $v_1, v_2 \in A$ with the largest possible common neighbourhood $X = N(v_1) \cap N(v_2)$. By a similar computation as in the proof of Lemma 2.1 we get that the average size of the common neighbourhood of two vertices of A is

$$\frac{1}{\binom{n}{2}} \sum_{u\in B} \binom{\deg u}{2} \geq \frac{m}{\binom{n}{2}} \binom{\frac{1}{m}\sum_{u\in B} \deg u}{2} \geq \frac{\varepsilon^2 m}{2}, \tag{2}$$

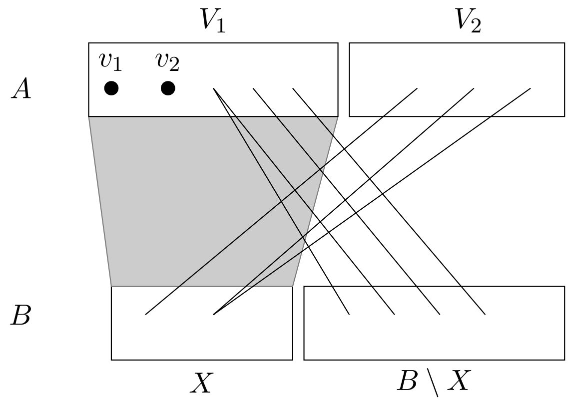

Figure 2: An illustration for the proof of Lemma 2.2

thus we have $|X| \geq \varepsilon^2 m/2$.

Let V_1 be the set of vertices $v_3 \in A$ for which $X \subseteq N(v_3)$ and let V_2 be the set of vertices $v_3 \in A$ for which $|N(v_3) \cap X| \leq 1$. See Figure 2 for an illustration of these sets. By our previous observation, $A = V_1 \cup V_2$. Let us now count the edges incident to X. Since every vertex of X has degree at least $\varepsilon n/3$ by the conditions of our theorem, X is incident to at least $\varepsilon n|X|/3$ edges. On the other hand, there are $|V_1||X|$ edges between X and V_1, and at most $|V_2|$ edges between X and V_2. Thus,

$$\frac{\varepsilon}{3}n|X| \leq e(A, X) = e(V_1, X) + e(V_2, X) \leq |V_1||X| + |V_2|.$$

Since $\varepsilon > 6m^{-1/3}$, we have $\varepsilon n|X|/6 \geq \varepsilon^3 nm/12 > n \geq |V_2|$, and so $|V_1||X| \geq \varepsilon n|X|/6$. Therefore, $|V_1| \geq \varepsilon n/6$. From this, we get the complete bipartite graph on the vertex set $V_1 \cup X$ with $|X| \cdot |V_1| \geq \varepsilon^3 mn/6$ edges. This is weaker than our promised bound, so let us improve it as follows.

As X is the largest common neighbourhood of two vertices in A and any two vertices of V_1 contain X in their common neighbourhood, this shows that no two vertices of V_1 have a common neighbour outside X. In particular, every vertex outside X sends at most 1 edge to V_1, and so $e(V_1, B) \leq |X||V_1| + m$. On the other hand, every vertex of V_1 has degree at least $\varepsilon m/3$, and therefore we have

$$|V_1| \cdot \frac{\varepsilon}{3}m \leq e(V_1, B) \leq |X||V_1| + m.$$

Since $\varepsilon > 6n^{-1/2}$, we have $|V_1| \cdot \varepsilon m/6 \geq \frac{\varepsilon^2 n}{36}m \geq m$ and so $|X||V_1| \geq |V_1| \cdot \varepsilon m/6$. Thus, we conclude that $|X| \geq \varepsilon m/6$ and so G contains a complete bipartite subgraph between V_1 and X, with at least $|V_1||X| \geq \varepsilon^2 mn/36$ edges. $\square$

3 Point-hyperplane incidences with no $K_{s,s}$

In this section, we prove the upper bounds in Theorem 1.2. We prepare the proof with the following lemma about finding the pattern Π_d under some simple conditions. We recall that Π_d is the pattern on a balanced bipartite graph on vertices $a_1, \ldots, a_d, b_1, \ldots, b_d$ where $a_i b_j$ is labeled by 1 for $i \geq j-1$ and $a_i b_{i+2}$ is labeled by 0, while all other edges are labeled by $*$.

Lemma 3.1. *Let s and $2 \leq t \leq d$ be positive integers, and let $G = (A, B; E)$ be a $K_{s,s}$-free bipartite graph. Assume that $v_1, \ldots, v_t \in A$ satisfy the following conditions:*

- $N(v_1, v_2) \supsetneq N(v_1, v_2, v_3) \supsetneq \cdots \supsetneq N(v_1, \ldots, v_t)$,
- *there exists a set $X \subseteq N(v_1, \ldots, v_t)$ of size at least $d^{d+1}s$ such that every $(d+1-t)$-tuple of elements in X has a common neighbourhood of size at least s.*

Then G contains the pattern Π_d.

Proof. The first condition implies that there exist vertices $u_1, \ldots, u_{t-2}$ which satisfy

$$u_i \in N(v_1, \ldots, v_{i+1}) \backslash N(v_{i+2}).$$

To find vertices $u_{t-1}, \ldots, u_d$ and $v_{t+1}, \ldots, v_d$, we define a $(d+2-t)$-uniform hypergraph F on the set of vertices X. Let $\prec$ be an arbitrary total ordering on X, then the edges of the hypergraph F will be those sets $\{u_{t-1}, \ldots, u_d\}$ with $u_{t-1} \prec \cdots \prec u_d$ for which there exists an index $i \in \{t-1, \ldots, d-1\}$ such that $N(u_i, \ldots, u_d) = N(u_{i+1}, \ldots, u_d)$. We claim that F has at most

$$s(d+1-t)|X|^{d+1-t}$$

edges. To this end, note that there are $d+1-t$ choices for the index i at which the equality $N(u_i, \ldots, u_d) = N(u_{i+1}, \ldots, u_d)$ can occur. Furthermore, for each of these choices, there are at most $|X|^{d-i}$ ways to choose $\{u_{i+1}, \ldots, u_d\}$. Having fixed these vertices, which by assumption satisfy $|N(u_{i+1}, \ldots, u_d)| \geq s$, we have less than s vertices u_i for which $N(u_i) \supseteq N(u_{i+1}, \ldots, u_d)$, otherwise we have a $K_{s,s}$. Finally, there are at most $|X|^{i-(t-1)}$ ways to choose vertices $u_{t-1}, \ldots, u_{i-1}$. Thus, the total number of edges of the hypergraph F is at most

$$|E(F)| < (d+1-t)|X|^{d-i} \cdot s \cdot |X|^{i-(t-1)} = s(d+1-t)|X|^{d+1-t},$$

as claimed.

Observe that the inequality $\binom{|X|}{d+2-t} \geq \left(\frac{|X|}{d+2-t}\right)^{d+2-t} \geq s(d+1-t)|X|^{d+1-t}$ is satisfied by our assumption that $|X| \geq sd^{d+1}$. Hence, there exist $u_{t-1} \prec \cdots \prec u_d$ in X such that $\{u_{t-1}, \ldots, u_d\}$ is not an edge of F, or equivalently, $N(u_i, \ldots, u_d) \subsetneq N(u_{i+1}, \ldots, u_d)$ for all indices $t-1 \leq i \leq d-1$. Picking vertices

$$v_{i+2} \in N(u_{i+1}, \ldots, u_d) \backslash N(u_i, \ldots, u_d)$$

for all $t-1 \leq i \leq d-1$, the vertices $v_1, \ldots, v_d, u_1, \ldots, u_d$ span a subgraph with pattern Π_d. $\square$

The following is the main graph-theoretical lemma which allows us to prove the upper bounds on the number of incidences of points and hyperplanes. Let us recall that we have defined $\alpha_t = \frac{t}{d+2-t}$ for $t \in \{2, \ldots, d\}$ and $\beta_t = \frac{t}{d+1-t}$ for $t \in \{1, \ldots, d\}$.

Lemma 3.2. *For integers $d, s \geq 2$, there exists $C = C(d,s) > 0$ such that the following holds for every sufficiently large n. Let $G = (A, B; E)$ be a bipartite graph, $|A| = n, |B| = m$ and $m = n^\alpha$. If G contains no $K_{s,s}$ as a subgraph and does not contain the pattern Π_d, then*

$$e(G) \leq \begin{cases} Cn & \text{if } \alpha \in [0, \beta_1], \\ Cm^{1-\frac{1}{t}}n & \text{if } \alpha \in [\alpha_t, \beta_t] \text{ for some } t \in \{2, \ldots, d\}, \\ Cmn^{1-\frac{1}{d+2-t}} & \text{if } \alpha \in [\beta_{t-1}, \alpha_t] \text{ for some } t \in \{2, \ldots, d\}, \\ Cm & \text{if } \alpha \in [\beta_d, \infty). \end{cases}$$

Moreover, if $\alpha = 1$, i.e. when $m = n$, the constant $C = 8(s + d^3)$ suffices.

Proof. Let us start by addressing the cases when $\alpha \in [\alpha_t, \beta_t]$ for some $t \in \{2, \ldots, d\}$ or $\alpha \in [\beta_d, \infty)$. In the latter case, we set $t := d$. Suppose that $e(G) \geq C(m^{1-\frac{1}{t}}n + m)$ for some large constant C, and observe that $m^{1-\frac{1}{t}}n \geq m$ if $\alpha \in [\alpha_t, \beta_t]$, and $m \geq m^{1-\frac{1}{t}}n$ if $\alpha \in [\beta_d, \infty)$. We show that there exists a t-tuple of vertices $v_1, \ldots, v_t \in A$ which satisfies the assumptions of Lemma 3.1, which then implies that G contains the pattern Π_d, contradiction. To do this, we use a variant of the dependent random choice method.

We call an ordered t-tuple of vertices $(v_1, \ldots, v_t)$ *bad* if $N(v_1, \ldots, v_t) \geq s$ and $N(v_1, \ldots, v_i) = N(v_1, \ldots, v_{i+1})$ for some $i \in \{1, \ldots, t-1\}$, and *good* otherwise. We choose a random good t-tuple $(v_1, \ldots, v_t)$ with the following sampling procedure. We sample the vertices v_i in order, starting from v_1 which is a uniformly random element of A. At step i, having chosen vertices $v_1, \ldots, v_{i-1}$, we have two cases.

Case 1. $|N(v_1,\dots,v_{i-1})| < s$. In this case, any t-tuple containing $v_1,\dots,v_{i-1}$ is good, and then all subsequent vertices v_j for $j \ge i$ can be sampled uniformly at random from $A \backslash \{v_1,\dots,v_{j-1}\}$.

Case 2. $|N(v_1,\dots,v_{i-1})| \ge s$. Define the set $S_i = \{v \in A : N(v) \supseteq N(v_1,\dots,v_{i-1})\}$. Since $G(P,H)$ does not contain a copy of $K_{s,s}$, there are less than s vertices in the set S_i. Sample v_i uniformly at random from $A \backslash (S_i \cup \{v_1,\dots,v_{i-1}\})$.

Let us denote by X the size of the common neighbourhood of $v_1,\dots,v_t$, and let Y be the number of $(d+1-t)$-tuples of vertices $u_t,\dots,u_d \in N(v_1,\dots,v_t)$ which has less than s common neighbours. If we show that $\mathbb{E}[X-Y] \ge d^{d+1}s$, then one can pick a good t-tuple $v_1,\dots,v_t$ satisfying $X-Y \ge d^{d+1}s$ and delete one vertex from every $(d+1-t)$-tuple in $N(v_1,\cdots,v_t)$ which has less than s common neighbours. Then, one is left with the t-tuple $v_1,\dots,v_t$ and a subset of their common neighbourhood which satisfy all conditions of Lemma 3.1.

Thus, the main objective is to show that $\mathbb{E}[X-Y] \ge d^{d+1}s$. By linearity of expectation, we have $\mathbb{E}[X] = \sum_{u \in B} \mathbb{P}[u \in N(v_1,\dots,v_t)]$. Let us fix a vertex $u \in B$ with $\deg u \ge d+s$, and estimate $\mathbb{P}[u \in N(v_1,\dots,v_t)]$. We have

$$\begin{aligned}\mathbb{P}[u \in N(v_1,\dots,v_t)] = \mathbb{P}[v_1,\dots,v_t \in N(u)] &= \prod_{i=1}^{t} \mathbb{P}\Big[v_i \in N(u) \Big| v_1,\dots,v_{i-1} \in N(u)\Big] \\ &\ge \prod_{i=1}^{t} \frac{\deg u - s - i + 1}{n - i + 1} > \left(\frac{\deg u - s - d}{n}\right)^t.\end{aligned}$$

Let $B_0 \subseteq B$ be the set of vertices u satisfying $\deg u \ge 2(d+s)$. Then, the number of edges incident to B_0 is $\sum_{u \in B_0} \deg u \ge C(m^{1-1/t}n + m) - 2(d+s)m \ge e(G)/2$ for $C \ge 4(d+s)$. Therefore,

$$\begin{aligned}\mathbb{E}[X] &\ge \sum_{u \in B_0} \left(\frac{\deg u - s - d}{n}\right)^t \ge \sum_{u \in B_0} \left(\frac{\deg u}{2n}\right)^t \\ &\ge |B_0| \left(\frac{\sum_{u \in B_0} \deg u}{2|B_0|n}\right)^t \ge m \left(\frac{e(G)}{4mn}\right)^t \ge \left(\frac{e(G)}{4m^{1-1/t}n}\right)^t,\end{aligned}$$

where the third inequality holds by convexity. On the other hand, the probability that a given $(d+1-t)$-tuple of vertices $u_t,\dots,u_d \in B$ which has less than s common neighbours lies in $N(v_1,\dots,v_t)$ can be bounded by

$$\mathbb{P}[u_t,\dots,u_d \in N(v_1,\dots,v_t)] = \mathbb{P}[v_1,\dots,v_t \in N(u_t,\dots,u_d)] \le \left(\frac{|N(u_t,\dots,u_d)|}{n-s-d}\right)^t \le (2s)^t n^{-t},$$

assuming n is sufficiently large with respect to s and d. Thus, the expectation of Y can be bounded as

$$\mathbb{E}[Y] \le \sum_{\substack{u_t,\dots,u_d \in B \\ |N(u_t,\dots,u_d)| < s}} \left(\frac{2s}{n}\right)^t \le m^{d+1-t} \left(\frac{2s}{n}\right)^t.$$

When $\alpha \le d = \beta_d$, we have $\alpha \le \beta_t = \frac{t}{d+1-t}$. Hence, for a sufficiently large value of C,

$$\mathbb{E}[X-Y] \ge \left(\frac{e(G)}{4m^{1-1/t}n}\right)^t - m^{d+1-t}\left(\frac{2s}{n}\right)^t \ge (C/4)^t - (2s)^t n^{\alpha(d+1-t)-t} \ge (C/4)^t - (2s)^t \ge d^{d+1}s.$$

On the other hand, when $\alpha \ge d = \beta_d$, we have

$$\mathbb{E}[X-Y] \ge \Big(\frac{Cm}{4m^{1-1/t}n}\Big)^t - m\Big(\frac{2s}{n}\Big)^t = \big((C/4)^d - (2s)^d\big)\frac{m}{n^d} \ge (C/4)^d - (2s)^d \ge d^{d+1}s,$$

for a sufficiently large value of C. Thus, we indeed have $\mathbb{E}[X-Y] \ge (C/4)^t - (2s)^t \ge d^{d+1}s$ in all cases and therefore this completes the first case of the proof.

Before we consider the rest of the cases, let us justify the assertion that when $\alpha = 1$ one can take $C = 8(s + d^3)$. Namely, if $\alpha = 1$ we have $\alpha_t \le \alpha \le \beta_t$ for $t = \lceil \frac{d+1}{2} \rceil$. Then, for the proof to go through, one needs to verify that the inequality $\mathbb{E}[X - Y] \ge (C/4)^t - (2s)^t \ge sd^{d+1}$ holds with the proposed value of C, which is a consequence of a simple calculation.

Observe that the rest of the cases already follow by symmetry. Since $1/\alpha_t = \alpha_{d+2-t}$ and $1/\beta_{t-1} = \beta_{d+2-t}$, we have that if $\alpha \in [\beta_{t-1}, \alpha_t]$ for some $t \in \{2, \dots, d\}$, then $1/\alpha \in [\alpha_{d+2-t}, \beta_{d+2-t}]$. Also, if $\alpha \in [0, \beta_1]$, then $1/\alpha \in [\beta_d, \infty)$. Thus, if we reverse the roles of m and n, together with the observation $n = m^{1/\alpha}$, we find ourselves in the setting of the cases already proved. This completes the proof of the theorem. □

Now we are ready to prove the upper bounds of Theorem 1.1 and Theorem 1.2, which we restate as follows.

Theorem 3.3. *Let $d, s \ge 2$ be integers, then there exists $C = C(d, s) > 0$ such that the following holds for every $\alpha > 0$ and field $\mathbb{F}$. Let $\mathcal{P}$ be a set of n points and $\mathcal{H}$ be a set of $m = n^\alpha$ hyperplanes in $\mathbb{F}^d$ such that the incidence graph of $(\mathcal{P}, \mathcal{H})$ is $K_{s,s}$-free. Then*

$$I(\mathcal{P}, \mathcal{H}) \le \begin{cases} Cn & \text{if } \alpha \in [0, \beta_1], \\ Cm^{1-\frac{1}{t}}n & \text{if } \alpha \in [\alpha_t, \beta_t] \text{ for some } t \in \{2, \dots, d\}, \\ Cmn^{1-\frac{1}{d+2-t}} & \text{if } \alpha \in [\beta_{t-1}, \alpha_t] \text{ for some } t \in \{2, \dots, d\}, \\ Cm & \text{if } \alpha \in [\beta_d, \infty). \end{cases}$$

Moreover, if $\alpha = 1$, the implied constant can be taken to be $C = 8(s + d^3)$.

Proof. By Lemma 1.4, the graph $G(\mathcal{P}, \mathcal{H})$ does not contain the pattern Π_d. Since the conclusion of Lemma 3.2 does not hold, one of its assumptions must fail. Thus, we conclude that $G(\mathcal{P}, \mathcal{H})$ either contains $K_{s,s}$ as a subgraph or $I(\mathcal{P}, \mathcal{H})$ satisfies the required upper bound. □

4 Large complete bipartite subgraphs in incidence graph

The main goal of this section is to prove our lower bounds on $\mathrm{rs}_d(m, n, I)$ and $\mathrm{rs}_d(n, m, I)$ for $m \ge n$, that is, to prove Theorem 1.3. As observed before, the lower bounds are only interesting in the regime $\varepsilon > 100 \max\{n^{-\frac{1}{d-1}}, m^{-\frac{1}{d}}\}$, otherwise taking the maximum degree vertex with its neighborhood gives a complete bipartite graph of the required size. Let us restate our theorem in this case.

Theorem 4.1. *Let $m \ge n$ be integers, $\varepsilon > 0$, and $I = \varepsilon mn$. If $\varepsilon > 100 \max\{m^{-\frac{1}{d}}, n^{-\frac{1}{d-1}}\}$, then*

$$\mathrm{rs}_d(m, n, I), \mathrm{rs}_d(n, m, I) \ge \left(\frac{\varepsilon}{100}\right)^{d-1} mn.$$

Let us prepare the proof with some definitions. For a bipartite graph G, we denote by $\mathrm{rs}(G)$ the number of edges in the largest complete bipartite subgraph of $G = (A, B; E)$. Also, we say that a t-tuple of vertices $v_1, \dots, v_t \in A$ is a *good t-tuple* if $N(v_1, v_2) \supsetneq N(v_1, v_2, v_3) \supsetneq \dots \supsetneq N(v_1, \dots, v_t)$ and $|N(v_1, \dots, v_t)| \ge 2$. The reason this definition is useful is the following lemma, which allows us to embed the pattern Π_d using good d-tuples. One can think of Lemma 4.2 as a simpler version of Lemma 3.1 from Section 3.

Lemma 4.2. *Suppose that G contains a good d-tuple $v_1, \dots, v_d$. Then G contains the pattern Π_d.*

Proof. Picking $u_i \in N(v_1, \dots, v_{i+1}) \backslash N(v_1, \dots, v_{i+2})$ for $i \le d - 2$ and $u_{d-1}, u_d \in N(v_1, \dots, v_d)$, the vertices $v_1, \dots, v_d, u_1, \dots, u_d$ induce a subgraph with pattern Π_d. □

Let us reformulate Theorem 4.1 in terms of good d-tuples.

Lemma 4.3. *Let $d \geq 3$ be a positive integer and let $G = (A, B; E)$ be a bipartite graph with at least εmn edges, where $|A| = n$, $|B| = m$ and $n \leq m$. If $\varepsilon \geq 100 \max\{n^{-\frac{1}{d-1}}, m^{-\frac{1}{d}}\}$, then either A contains a good d-tuple or*

$$\mathrm{rs}(G) \geq \left(\frac{\varepsilon}{100}\right)^{d-1} mn.$$

We prove this lemma after some preparations. The following lemma shows how to build good t-tuples under the assumption that G does not contain large complete bipartite subgraphs.

Lemma 4.4. *Let $\gamma, \varepsilon > 0$, and let $G = (A, B; E)$ be a bipartite graph such that $|A| = n$, $|B| = m$, and every vertex $u \in B$ has degree at least $\deg u \geq \frac{\varepsilon}{3} n$. If $\mathrm{rs}(G) < \frac{\varepsilon}{6}\gamma \cdot mn$, then for any set $X \subseteq B$ of size at least γm, there exists a vertex $v \in A$ for which $|N(v) \cap X| \geq \frac{\varepsilon}{6}|X|$ and $N(v) \cap X \neq X$.*

Proof. Let $X \subseteq B$ be a subset of size at least γm. The idea of this proof is to double count the edges incident to X. Let V_1 denote the set of vertices $v \in A$ for which $X \subseteq N(v)$ and let V_2 denote the set of vertices $v \in A$ with $|N(v) \cap X| < \frac{\varepsilon}{6}|X|$. To show the lemma, it suffices to find a vertex $v \in A$ belonging to neither V_1 nor V_2.

Suppose, for the sake of contradiction, that $V_1 \cup V_2 = A$. Then, $e(A, X) = e(V_1, X) + e(V_2, X)$. Since the degree of every vertex in B is at least $\frac{\varepsilon n}{3}$, we have $e(A, X) \geq \frac{\varepsilon}{3} n|X|$. On the other hand, since the graph induced on $V_1 \cup X$ is complete bipartite, we have $e(V_1, X) \leq \mathrm{rs}(G) \leq \frac{\varepsilon}{6}\gamma \cdot mn$. Also, by the definition of V_2, we have $e(V_2, X) = \sum_{v \in V_2} |N(v) \cap X| < |V_2| \cdot \frac{\varepsilon}{6}|X| \leq \frac{\varepsilon}{6} n|X|$. Hence,

$$\frac{\varepsilon}{6}\gamma \cdot mn + \frac{\varepsilon}{6} n|X| > \frac{\varepsilon}{3} n|X|.$$

Rearranging this inequality gives $|X| < \gamma m$, which is a contradiction to our initial assumption. $\square$

Now, we explain how to build good $(d-1)$-tuples in the graph G.

Lemma 4.5. *Let $d \geq 3$ be a positive integer and let $G = (A, B; E)$ be a bipartite graph with εmn edges, where $|A| = n$, $|B| = m$. If $\mathrm{rs}(G) \leq (\frac{\varepsilon}{6})^{d-1} mn$, then G contains a good $(d-1)$-tuple $v_1, \ldots, v_{d-1} \in A$ with common neighbourhood of size at least $|N(v_1, \ldots, v_{d-1})| \geq 2\left(\frac{\varepsilon}{6}\right)^{d-1} m$.*

Proof. For $t = 2, \ldots, d-1$, we prove that there is a good t-tuple $v_1, \ldots, v_t \in A$ with $|N(v_1, \ldots, v_t)| \geq 2\left(\frac{\varepsilon}{6}\right)^t m$. We prove this by induction on t. For $t = 2$, we simply need to find a pair of vertices $v_1, v_2 \in A$ whose common neighbourhood has size at least $2\left(\frac{\varepsilon}{6}\right)^2 m$. In fact, the average size of the common neighbourhood of two vertices of A is $\frac{\varepsilon^2}{2} m$, as shown by equation (2) in Section 2, and therefore we are done.

For $t > 2$, the induction hypothesis implies that G contains a good $(t-1)$-tuple $v_1, \ldots, v_{t-1} \in A$ with the common neighbourhood of size $|N(v_1, \ldots, v_{t-1})| \geq 2\left(\frac{\varepsilon}{6}\right)^{t-1} m$. Let $\gamma = (\frac{\varepsilon}{6})^{t-1}$, then $\mathrm{rs}(G) \leq (\frac{\varepsilon}{6})^{d-1} mn \leq \frac{\varepsilon}{6}\gamma \cdot mn$ and $|X| \geq \gamma m$ are satisfied, so we can apply Lemma 4.4 to $X = N(v_1, \ldots, v_{t-1})$. We get that there exists a vertex $v_t \in A$ for which $|N(v_1, \ldots, v_t)| \geq \frac{\varepsilon}{6}|X| \geq (\frac{\varepsilon}{6})^t m$ and $N(v_1, \ldots, v_{t-1}) \supsetneq N(v_1, \ldots, v_t)$. We conclude that $v_1, \ldots, v_t$ is a good t-tuple with sufficiently large common neighbourhood, completing the proof. $\square$

Now, we are ready to prove Lemma 4.3.

Proof of the Lemma 4.3. We begin the proof by showing that we can assume large minimum degree, more precisely that $\min_{v \in A} \deg v \geq \left(1 - \frac{2}{d}\right) \varepsilon |B|$ and $\min_{v \in B} \deg v \geq \left(1 - \frac{2}{d}\right) \varepsilon |A|$. Then, we will present the main argument, in which we strengthen Lemma 4.5 and show there exists a good $(d-1)$-tuple with common neighbourhood of size $\left(\frac{\varepsilon}{6}\right)^{d-2} m$, which allows us to apply Lemma 4.4 to find a good d-tuple and complete the proof.

We will consider a proof by double induction, first on d and then on $m + n$. Note that if $\min\{m, n\} = 1$, then the statement of the lemma is trivial. Our first step will be, as mentioned above, to show that if G has a vertex of low degree, then one can remove this vertex and obtain a graph G' with $\mathrm{rs}(G') \geq (\frac{\varepsilon}{100})^{d-1} mn$.

Since A and B are symmetric, we will only discuss the case when G has a vertex $v \in A$ with degree $\deg v \leq \left(1 - \frac{2}{d}\right) \varepsilon|B|$. Define the graph $G' = G - \{v\}$, which has parts of size $n' = n - 1$ and $m' = m$. Furthermore, G' has $\varepsilon' m' n'$ edges, where $\varepsilon' m' n' \geq \varepsilon mn - \left(1 - \frac{2}{d}\right) \varepsilon m$. Thus, if $G - \{v\}$ does not contain the pattern Π_d, we conclude

$$\begin{aligned}\mathrm{rs}(G) \geq \mathrm{rs}(G') &\geq \left(\frac{\varepsilon'}{100}\right)^{d-1} m'n' \geq \left(\frac{\varepsilon mn - \left(1 - \frac{2}{d}\right)\varepsilon m}{100(n-1)m}\right)^{d-1} (n-1)m \\ &= \left(\left(\frac{\varepsilon}{100}\right)^{d-1} mn\right) \cdot \frac{n-1}{n} \left(\frac{n - (1 - \frac{2}{d})}{n-1}\right)^{d-1} \geq \left(\frac{\varepsilon}{100}\right)^{d-1} mn.\end{aligned}$$

The last inequality is true due to simple calculations. We should also verify that G' satisfies the conditions of Lemma 4.3 by showing that $\varepsilon'^{d-1} n' \geq 100^{d-1}$ and $\varepsilon'^{d} m' \geq 100^{d}$. But this also follows, since we have already shown that $\varepsilon'^{d-1} n' m' \geq \varepsilon^{d-1} mn$, which implies both $\varepsilon'^{d-1} n' \geq \varepsilon^{d-1} n \geq 100^{d-1}$ and $\varepsilon'^{d} m' \geq \varepsilon^{d} m \geq 100^{d}$ (since $\varepsilon'^{d-1} n' m' \geq \varepsilon^{d-1} mn$ also implies $\varepsilon' \geq \varepsilon$). We conclude that $\mathrm{rs}(G) \geq (\varepsilon/100)^{d-1} mn$, as promised.

In what follows, we assume that G contains no vertex $v \in A$ of degree less than $\left(1 - \frac{2}{d}\right) \varepsilon m$ and no vertex $v \in B$ of degree less than $\left(1 - \frac{2}{d}\right) \varepsilon n$. Let us now focus on the main argument. Recall that we proceed by induction on d, where the case $d = 3$ was discussed in Section 2 as Lemma 2.2.

We choose a good $(d-1)$-tuple $(v_1, \ldots, v_{d-1})$ with the maximal size of the common neighbourhood, which we denote by $X = N(v_1, \ldots, v_{d-1})$. By Lemma 4.5, there must be a good $(d-1)$-tuple with $|X| \geq \left(\frac{\varepsilon}{6}\right)^{d-1} |B|$. If A does not contain a good d-tuple in G, then any vertex $v \in A$ either has $X \subseteq N(v)$ or $|N(v) \cap X| = 1$. Lemma 4.4 applied with $\gamma = (\frac{\varepsilon}{6})^{d-2}$ shows that $|X| \leq \left(\frac{\varepsilon}{6}\right)^{d-2} |B|$. We conclude that one can partition A into two sets, V_1 and V_2, where $V_1 = \{v \in A : X \subseteq N(v)\}$ and $V_2 = \{v \in A : |X \cap N(v)| \leq 1\}$.

Let us consider the edges incident to X. By our assumption on the minimum degree of vertices of X, we have $e(X, A) \geq |X| \cdot \left(1 - \frac{2}{d}\right) \varepsilon n \geq \frac{\varepsilon}{3} n |X|$. On the other hand,

$$\frac{\varepsilon}{3} n|X| \leq e(X, A) = e(X, V_1) + e(X, V_2) \leq |X||V_1| + |V_2| \leq |X||V_1| + n.$$

Since $|X| \geq \left(\frac{\varepsilon}{6}\right)^{d-1} m$ and $(\frac{\varepsilon}{100})^{d} m \geq 1$, we have $|X| \geq \frac{12}{\varepsilon}$. Thus, $\frac{\varepsilon}{12} n|X| \geq n$ and so $\frac{\varepsilon}{4} n|X| \leq |X||V_1|$, showing that $|V_1| \geq \frac{\varepsilon}{4} n$.

Let us now consider the bipartite graph $G' = G[V_1 \cup (B \backslash X)]$. The number of edges of this graph is at least

$$\begin{aligned} e(G') = e(V_1, B) - e(V_1, X) &\geq |V_1| \cdot \left(1 - \frac{2}{d}\right) \varepsilon |B| - |V_1||X| \\ &\geq |V_1| \left(\frac{d-2}{d} \varepsilon |B| - \left(\frac{\varepsilon}{6}\right)^{d-2} |B|\right) \geq \left(1 - \frac{5/2}{d}\right) \varepsilon |B||V_1|, \end{aligned}$$

since $6^{d-2} \geq 2d$ for all $d \geq 3$. Thus, the sizes of the vertex classes of G' are $n' = |V_1| \geq \frac{\varepsilon n}{4}$ and $m' = |B \backslash X| \geq (1 - \frac{5/2}{d}) m$. Finally, the above computation shows that $e(G') = \varepsilon' m' n' \geq (1 - \frac{5/2}{d}) \varepsilon m' n'$.

Next, our plan is to apply the induction hypothesis for $d-1$ on the graph G' to find a large complete bipartite subgraph. Let us first argue that G' does not have a good $(d-1)$-tuple. Suppose for contradiction that G' contains a good $(d-1)$-tuple $u_1, \ldots, u_{d-1}$. The vertices $u_1, \ldots, u_{d-1}$ form a good $(d-1)$-tuple in the graph G as well, and their common neighbourhood contains both X and at least two vertices from $B \backslash X$ (since this is a good $(d-1)$-tuple in G'). But this implies $u_1, \ldots, u_{d-1}$ is a $(d-1)$-tuple with a larger common neighbourhood in G than $v_1, \ldots v_{d-1}$, contradicting the maximality. Thus, we conclude that G' cannot contain a good $(d-1)$-tuple.

We also need to verify that ε' is large enough so that we are able to apply the induction hypothesis on d, i.e. we need to check that $\varepsilon'^{d-2} n' \geq 100^{d-2}$ and $\varepsilon'^{d-1} m' \geq 100^{d-1}$. This is a consequence of an easy computation since

$$\varepsilon'^{d-2} n' \geq \left(1 - \frac{5/2}{d}\right)^{d-2} \varepsilon^{d-2} \frac{\varepsilon n}{4} \geq \frac{e^{-5/2}}{4} 100^{d-1} \geq 100^{d-2}.$$

In the above calculation, we have used that $\varepsilon^{d-1}n \geq 100^{d-1}$ and that $\left(1 - \frac{5/2}{d}\right)^{d-2} \geq e^{-5/2}$, which holds since $\left(1 - \frac{5/2}{d}\right)^{d-2}$ is a decreasing function for $d \geq 3$ with the limit $e^{-5/2}$ when $d \to \infty$. A similar computation shows that $\varepsilon'^{d-1}m' \geq \left(1 - \frac{5/2}{d}\right)^{d}\varepsilon^{d-1}m \geq 100^{d}\varepsilon^{-1}\left(1 - \frac{5/2}{4}\right)^{4} \geq 100^{d-1}$, where we have used that $\left(1 - \frac{5/2}{d}\right)^{d}$ is an increasing function of d, when $d \geq 3$. Having verified all the assumptions, we apply the induction hypothesis to the graph G' and conclude that

$$\mathrm{rs}(G') \geq \left(1 - \frac{5/2}{d}\right)^{d-2} \left(\frac{\varepsilon}{100}\right)^{d-2} |V_1||B| \geq e^{-5/2}\frac{\varepsilon^{d-1}}{4 \cdot 100^{d-2}}mn \geq \frac{\varepsilon^{d-1}}{100^{d-1}}mn.$$

Hence, we have $\mathrm{rs}(G) \geq \mathrm{rs}(G') \geq (\frac{\varepsilon}{100})^{d-1}mn$, which completes the proof. □

Proof of the lower bounds in Theorem 4.1. Proving the lower bounds on $\mathrm{rs}_d(m, n, I)$ and $\mathrm{rs}_d(n, m, I)$ is completely symmetric, so we just discuss $\mathrm{rs}_d(m, n, I)$. If $\varepsilon > 100 \max\{m^{-\frac{1}{d}}, n^{-\frac{1}{d-1}}\}$, let us consider any sets of n points and m hyperplanes, say $\mathcal{P}$ and $\mathcal{H}$. The incidence graph $G = G(\mathcal{P}, \mathcal{H})$ avoids the pattern Π_d by Lemma 1.4 and so G contains no good d-tuples. Since G has at least εmn edges, Lemma 4.3 implies $\mathrm{rs}(G) \geq \left(\frac{\varepsilon}{100}\right)^{d-1} mn$. □

5 Constructions and examples

The main goal of this section is to provide examples which show that the bounds given in Theorem 1.2 and Theorem 1.3 are tight. The main building block of all constructions in this section are *subspace evasive sets*. Throughout this section, we consider only finite fields $\mathbb{F}$.

Definition 5.1. A set of points $S \subset \mathbb{F}^d$ is (k, s)-*subspace evasive* if every k-dimensional affine subspace of $\mathbb{F}^d$ contains less than s elements of S.

It is clear that the size of a (k, s)-subspace evasive set is at most $O_s(|\mathbb{F}|^{d-k})$, as one can partition $\mathbb{F}^d$ into $|\mathbb{F}|^{d-k}$ affine subspaces of dimension k. It was proved by Dvir and Lovett [10] that for sufficiently large s, this trivial bound is optimal. Sudakov and Tomon [29] gave an alternative proof using the random algebraic method.

Lemma 5.2 ([10])**.** *For integers $1 \leq k \leq d$, there exists $s \leq d^d$ such that every finite space $\mathbb{F}^d$ contains a (k, s)-subspace evasive set of size at least $|\mathbb{F}|^{d-k}$.*

To be more precise, Dvir and Lovett [10] construct a (k, s)-subspace evasive set of size at least $\frac{1}{3}|\mathbb{F}|^{d-k}$. However, taking the union of four random translates of such a set results in a $(k, 4s)$-subspace evasive set of size at least $|\mathbb{F}|^{d-k}$ with high probability. Removing the factor $1/3$ makes our calculations a bit nicer, but other than that has no real effect.

5.1 Constructions without $K_{s,s}$

In this section, we prove the following equivalent formulation of the upper bounds in Theorem 1.2.

Theorem 5.3. *For every dimension d, there exist $s = s(d)$ and $c = c(d) > 0$ such that the following holds. For every $\alpha > 0$ and sufficiently large integer n, there exists a field $\mathbb{F}$, a set of n points $\mathcal{P}$ and a set of $m = \lfloor n^\alpha \rfloor$ hyperplanes $\mathcal{H}$ in $\mathbb{F}^d$ such that the incidence graph of $(\mathcal{P}, \mathcal{H})$ is $K_{s,s}$-free and*

$$I(\mathcal{P}, \mathcal{H}) \geq \begin{cases} n & \text{if } \alpha \in [0, \beta_1], \\ cm^{1-\frac{1}{t}}n & \text{if } \alpha \in [\alpha_t, \beta_t] \text{ for some } t \in \{2, \ldots, d\}, \\ cmn^{1-\frac{1}{d+2-t}} & \text{if } \alpha \in [\beta_{t-1}, \alpha_t] \text{ for some } t \in \{2, \ldots, d\}, \\ m & \text{if } \alpha \in [\beta_d, \infty). \end{cases}$$

The main building block of the proof of this theorem is the following lemma.

Lemma 5.4. *Let $2 \le t \le d$ be integers, then there exists an integer $s = s(d) > 0$ such that for every prime p the following holds. There exists a set of points $\mathcal{P}$ and a set of hyperplanes $\mathcal{H}$ in $\mathbb{F}_p^d$ such that $G(\mathcal{P},\mathcal{H})$ does not contain $K_{s,s}$, $|\mathcal{P}| \ge p^t$, $|\mathcal{H}| \ge p^{d-t+2}/s$, and*

$$I(\mathcal{P},\mathcal{H}) = \frac{|\mathcal{P}||\mathcal{H}|}{p} \ge \Omega_d\left((|\mathcal{P}||\mathcal{H}|)^{1-\frac{1}{d+2}}\right).$$

Proof. By Lemma 5.2, there exist a constant $s = s(d)$ and sets $\mathcal{P}, \mathcal{N}_0 \subset \mathbb{F}_p^d$ such that $|\mathcal{P}| \ge p^t$, $|\mathcal{N}_0| \ge p^{d-t+1}$, $\mathcal{P}$ is $(d-t,s)$-subspace evasive and $\mathcal{N}_0$ is $(t-1,s)$-subspace evasive. In other words, any s points of $\mathcal{P}$ span an affine subspace of dimension at least $d-t+1$, and any s points of $\mathcal{N}_0$ span an affine subspace of dimension at least t.

In particular, no line through the origin contains more than s points of $\mathcal{N}_0$. Thus, by keeping at most one point in each line through the origin, we get a subset $\mathcal{N}$ of $\mathcal{N}_0$ of size at least $\ge \frac{1}{s}|\mathcal{N}_0|$.

Let $\mathcal{H}$ be the set of all hyperplanes whose normal vectors are from $\mathcal{N}$, i.e. all hyperplanes of the form $\langle w, x\rangle = b$ for $w \in \mathcal{N}, b \in \mathbb{F}_p$. Since no line through the origin contains more than one point of $\mathcal{N}$, we observe that every pair $(w,b) \in \mathcal{N} \times \mathbb{F}_p$ yields a different hyperplane. Thus, the size of $\mathcal{H}$ is $|\mathcal{H}| = p|\mathcal{N}| \ge p^{d-t+2}/s$.

To count the incidences between $\mathcal{P}$ and $\mathcal{H}$, we observe that for every $w \in \mathcal{N}$ and every $x \in \mathcal{P}$, there is a unique $b \in \mathbb{F}_p$ such that $\langle w,x\rangle = b$. This means that out of p hyperplanes of $\mathcal{H}$ with normal vector $w \in \mathcal{N}$, exactly one is incident to x. Hence, the number of incidences between $\mathcal{P}$ and $\mathcal{H}$ is $I(\mathcal{P},\mathcal{H}) = |\mathcal{P}||\mathcal{N}| = \frac{1}{p}|\mathcal{P}||\mathcal{H}|$ and so $I(\mathcal{P},\mathcal{H}) \ge \Omega_d\left((|P||\mathcal{H}|)^{1-\frac{1}{d+2}}\right)$.

Finally, we argue that the incidence graph $G(\mathcal{P},\mathcal{H})$ is $K_{s,s}$-free. Indeed, any s points of $\mathcal{P}$ span an affine subspace of dimension at least $d-t+1$. On the other hand, the intersection of s hyperplanes is nonempty only when no two of them are parallel, meaning that their normal vectors are distinct. If these s hyperplanes intersect in a $(d-t+1)$-dimensional affine space, this means that the s normal vectors lie in a $(t-1)$-dimensional linear space orthogonal to it. But this is not possible since $\mathcal{N}$ is $(t-1,s)$-subspace evasive. Hence, the intersection of s hyperplanes can be at most $(d-t)$-dimensional. As $d-t < d-t+1$, this shows that $G(\mathcal{P},\mathcal{H})$ is $K_{s,s}$-free. □

Observe that Lemma 5.4 implies the existence of $\mathcal{P},\mathcal{H}$ with $|\mathcal{P}| = n_0 = \Omega_d(p^t), |\mathcal{H}| = m_0 = \Omega_d(p^{d+2-t})$ and $I(P,\mathcal{H}) = \Omega_d((m_0n_0)^{1-\frac{1}{d+2}})$. This only covers a sparse set of potential values of m and n. However, by using an approximation and a subsampling trick, one can cover all possible values of m and n.

Proof of Theorem 5.3. Let us first discuss the boundary cases, i.e. when $\alpha \in [0,\beta_1]$ or $\alpha \in [\beta_d,\infty)$. In this case, we show that there exist sets $\mathcal{P}$ and $\mathcal{H}$ with $|\mathcal{P}| = n, |\mathcal{H}| = m = n^\alpha$, $I(\mathcal{P},\mathcal{H}) \ge \max\{m,n\}$ and such that $G(\mathcal{P},\mathcal{H})$ does not contain $K_{2,2}$. In the case $\alpha \in [0,\beta_1]$, it suffices to choose $p \ge \max\{m,n\}$ and take $\mathcal{H}$ to be m parallel hyperplanes in $\mathbb{F}_p^d$ and $\mathcal{P}$ to be n points on one of the hyperplanes. On the other hand, when $\alpha \in [\beta_d,\infty)$, let $\mathcal{P}$ be n collinear points and $\mathcal{H}$ be a set of m hyperplanes containing exactly one of these points.

Let us now present the main argument. The idea of the proof is to pick a configuration $(\mathcal{P}_0,\mathcal{H}_0)$ given by Lemma 5.4 and randomly sample a proportion of it to obtain the required values of m,n.

First, consider the case when $\alpha_t \le \alpha \le \beta_t$ for some $t \in \{2,\dots,d\}$. Let p be the smallest prime larger than $2(sm)^{1/t}$. By Bertrand's postulate, we have $p \le 4(sm)^{1/t}$. By applying Lemma 5.4 with $d+2-t$ instead of t, we obtain a set of points $\mathcal{P}_0$ and a set of hyperplanes $\mathcal{H}_0$ in $\mathbb{F}_p^d$ with $|\mathcal{P}_0| = n_0 \ge p^{d+2-t}, |\mathcal{H}_0| = m_0 \ge p^t/s$ and $I(\mathcal{P}_0,\mathcal{H}_0) = \frac{1}{p}|\mathcal{P}_0||\mathcal{H}_0|$. Moreover, $G(\mathcal{P}_0,\mathcal{H}_0)$ does not contain $K_{s,s}$ as a subgraph.

Let $\mathcal{P}$ and $\mathcal{H}$ be random subsets of $P_0,\mathcal{H}_0$ containing n and m elements, respectively. Observe that $m_0 \ge \frac{1}{s}p^t \ge m$ and $n_0 \ge p^{d+2-t} \ge m^{\frac{d+2-t}{t}} \ge m^{1/\alpha} = n$, so it indeed makes sense to consider such subsets of $\mathcal{P}_0,\mathcal{H}_0$. The expected number of incidences between $\mathcal{P}$ and $\mathcal{H}$ is simply $\mathbb{E}[I(P,\mathcal{H})] = \frac{m}{m_0}\cdot\frac{n}{n_0}I(P_0,\mathcal{H}_0) \ge \frac{1}{p}mn$. Since $p \le 4(sm)^{-1/t}$, picking a pair $(\mathcal{P},\mathcal{H})$ with at least as many incidences as the expectation, we get

$$I(\mathcal{P},\mathcal{H}) \ge \frac{1}{p}mn \ge \frac{1}{4s^{1/t}}m^{-\frac{1}{t}}\cdot mn.$$

In the case $\beta_{t-1} \le \alpha \le \alpha_t$, we proceed similarly. We pick p to be the smallest prime larger than $2s^{\frac{1}{t}}n^{\frac{1}{d+2-t}}$, then $p < 4s^{\frac{1}{t}}n^{\frac{1}{d+2-t}}$. We apply Lemma 5.4 with $d+2-t$ instead of t to find a set of points $\mathcal{P}_0$ and a set of hyperplanes $\mathcal{H}_0$ in $\mathbb{F}_p^d$ with $|\mathcal{P}_0| = n_0 \ge p^{d+2-t}$, $|\mathcal{H}_0| = m_0 \ge p^t/s$ and $I(\mathcal{P}_0, \mathcal{H}_0) = \frac{1}{p}|\mathcal{P}_0||\mathcal{H}_0|$. By an almost identical computation as before one can check that $n_0 \ge n$ and $m_0 \ge n^{\alpha} \ge m$. Hence, if $\mathcal{P} \subset \mathcal{P}_0$ is a random n-element subset and $\mathcal{H} \subset \mathcal{H}_0$ is a random m-element subset, they satisfy $\mathbb{E}(I(\mathcal{P}, \mathcal{H})) \ge \frac{1}{p}mn \ge \frac{1}{4s^{1/t}}n^{-\frac{1}{d+2-t}} \cdot mn$. Therefore, we can pick a pair $(\mathcal{P}, \mathcal{H})$ with at least as many incidences as the expectation, completing the proof. $\square$

5.2 Constructions with no large complete bipartite subgraphs

In this section, we prove the upper bounds of Theorem 1.3. However, instead of assuming $m \ge n$ and giving bounds on both $\mathrm{rs}_d(m, n, I)$ and $\mathrm{rs}_d(n, m, I)$, as originally stated in Theorem 1.3, we prefer to keep the notation consistent by always denoting the number of hyperplanes by m and the number of points by n. Thus, we will give bounds on $\mathrm{rs}_d(m, n, I)$ which depend on $\max\{m, n\}$ and $\min\{m, n\}$. With this notational change, Theorem 1.3 can be equivalently restated as follows.

Theorem 5.5. *Let d, m, n be positive integers, $\varepsilon \in (0,1)$ and $I = \varepsilon mn$.*

- *If $\varepsilon \ge \frac{1}{4}\max\{m, n\}^{-1/d}$ and $\varepsilon \ge \frac{1}{4}\min\{m, n\}^{-1/(d-1)}$, then $\mathrm{rs}_d(m, n, I) \le O_d(\varepsilon^{d-1}mn)$.*
- *If $\varepsilon < \frac{1}{4}\max\{m, n\}^{-1/d}$ or $\varepsilon < \frac{1}{4}\min\{m, n\}^{-1/(d-1)}$ then $\mathrm{rs}_d(m, n, I) \le O_d(\varepsilon \max\{m, n\})$.*

The implied constant depending on the dimension in the above theorem is at most $2^{O(d \log d)}$. The main geometric construction is contained in the following lemma, which can be thought of as a strengthening of Lemma 5.4.

Lemma 5.6. *Let p be a prime and let d, t, k, ℓ be positive integers satisfying $t \le d-1$ and $k, \ell \le p^{d-t}$. Then, there exist an integer $s = O(d^d)$, a set of points $\mathcal{P}$ and a set of hyperplanes $\mathcal{H}$ in $\mathbb{F}_p^d$ satisfying the following properties:*

(i) $kp^t/2 \le |\mathcal{P}| \le kp^t$, $\frac{1}{4}\ell p^t \le |\mathcal{H}| \le \ell sp^t$, and every point of $\mathcal{P}$ is contained in exactly $|\mathcal{H}|/p$ hyperplanes of $\mathcal{H}$. In particular, $I(\mathcal{P}, \mathcal{H}) = \frac{|\mathcal{P}|\cdot|\mathcal{H}|}{p}$.

(ii) For $a \in \{1, \dots, t-1\}$ and any set $A \subseteq \mathcal{P}$ of size $|A| \ge skp^a$, A is not contained in a $(d-t+a)$-dimensional affine subspace.

(iii) For $b \in \{1, \dots, t-2\}$ and any set $B \subseteq \mathcal{H}$ of size $|B| \ge s^2\ell p^b$, the intersection $\bigcap_{H_i \in B} H_i$ does not contain a $(t-1-b)$-dimensional affine subspace.

Proof. By Lemma 5.2, there exists $s = O(d^d)$ such that for every prime p there exists a $(d-t, s)$-subspace evasive set $\mathcal{P}_0$ containing p^t points of $\mathbb{F}_p^d$. In other words, any s points of $\mathcal{P}_0$ span an affine subspace of dimension at least $d-t+1$. Furthermore, there exists a set of points $\mathcal{Q}_0$ containing p^{t-1} points of $\mathbb{F}_p^d$ such that $\mathcal{Q}_0$ is $(d-t+1, s)$-subspace evasive, which means that any s points of $\mathcal{Q}_0$ span an at least $(d-t+2)$-dimensional affine subspace.

The set $\mathcal{P}$ is constructed as the union of k random translates of $\mathcal{P}_0$. More precisely, we choose k random vectors $u_1, \dots, u_k \in \mathbb{F}_p^d$ (with repetition) and set $\mathcal{P} = \bigcup_{i=1}^k (\mathcal{P}_0 + u_i)$. The set $\mathcal{P}$ constructed in this way is $(d-t, ks)$-subspace evasive, since any affine space of dimension $d-t$ containing ks elements of $\mathcal{P}$ must contain at least s elements of one of the translates $\mathcal{P}_0 + u_i$. But this is impossible, since $\mathcal{P}_0$ is a $(d-t, s)$-subspace evasive set.

Next, we compute the expected size of $\mathcal{P}$ and argue that $u_1, \dots, u_k$ can be chosen such that $|\mathcal{P}| \ge \frac{1}{2}kp^t$. For a given point $x \in \mathbb{F}_p^d$, the probability that x lies in $\mathcal{P}_0 + u_i$ is

$$\mathbb{P}[x \in (\mathcal{P}_0 + u_i)] = \mathbb{P}[u_i \in (x - \mathcal{P}_0)] = \frac{1}{p^d}|\mathcal{P}_0|.$$

Thus, the expected size of $\mathcal{P}$ is

$$\mathbb{E}\big[|\mathcal{P}|\big] = \sum_{x\in\mathbb{F}_p^d} \mathbb{P}[x\in\mathcal{P}] = \sum_{x\in\mathbb{F}_p^d}\Big(1-\prod_{i=1}^{k}\mathbb{P}[x\notin(\mathcal{P}_0+u_i)]\Big) = \sum_{x\in\mathbb{F}_p^d}\left(1-\Big(1-\frac{|\mathcal{P}_0|}{p^d}\Big)^k\right)$$
$$\geq p^d\Big(1-\exp\Big(-\frac{k|\mathcal{P}_0|}{p^d}\Big)\Big) \geq p^d\frac{k|\mathcal{P}_0|}{2p^d} = \frac{1}{2}kp^t.$$

In the above calculation, we have used the inequalities $1-\frac{x}{2}\geq e^{-x}\geq 1-x$ for $x\in(0,1)$. Therefore, we conclude that $u_1,\dots,u_k$ can be chosen such that $\mathcal{P}$ has size at least $\frac{1}{2}kp^t$. On the other hand, we always have $|\mathcal{P}|\leq k|\mathcal{P}_0| = kp^t$.

Let us now explain how to process the set $\mathcal{Q}_0$ and how to construct from it the set of hyperplanes $\mathcal{H}$. Let $U=\{x\in\mathbb{F}_p^d : x(d)=0\}$. Since we may translate the set $\mathcal{Q}_0$ without altering its properties, we may assume that at most $\frac{|\mathcal{Q}_0|}{p}$ elements of $\mathcal{Q}_0$ lie in U. Then, we define the set $\mathcal{Q}_1=\mathcal{Q}_0\backslash U$, which is still $(d-t+1,s)$-subspace evasive and has size $|\mathcal{Q}_1|\geq|\mathcal{Q}_0|/2$.

Pick $s\ell$ uniformly random vectors $v_1,\dots,v_{s\ell}\in U$ and define $\mathcal{Q}_2=\bigcup_{i=1}^{s\ell}(\mathcal{Q}_1+v_i)$. In a similar manner as above, we observe that $\mathcal{Q}_2$ is a $(d-t+1,\ell s^2)$-subspace evasive set. Our goal now is to show that $v_1,\dots,v_{s\ell}$ can be chosen in a way such that at least $\frac{1}{4}\ell p^{t-1}$ lines through the origin contain a point of $\mathcal{Q}_2$. If $L\subset\mathbb{F}_p^d$ is a fixed line through the origin, note that L intersects $\mathcal{Q}_1+v_i$ if and only if the translate of L given by $L-v_i$ intersects $\mathcal{Q}_1$. Note that union of all the translates of L covers $\mathcal{Q}_1$. Since $\mathcal{Q}_1$ is a $(d-t+1,s)$-subspace evasive set, each translate of L contains at most s points of $\mathcal{Q}_1$ and so at least $|\mathcal{Q}_1|/s$ translates of L intersect $\mathcal{Q}_1$. Thus, we conclude $\mathbb{P}[L\cap(\mathcal{Q}_1+v_i)\neq\emptyset]\geq\frac{|\mathcal{Q}_1|/s}{p^{d-1}}$. If we denote the number of lines through the origin intersecting $\mathcal{Q}_2$ by Y, by a similar calculation as above we have

$$\mathbb{E}[Y] = \sum_{L\text{ line through }0} 1-\mathbb{P}[L\cap\mathcal{Q}_2=\emptyset] \geq \sum_{L\text{ line through }0} 1-\left(1-\frac{|\mathcal{Q}_1|}{sp^{d-1}}\right)^{s\ell} \geq p^{d-1}\frac{\ell|\mathcal{Q}_1|}{2p^{d-1}} \geq \frac{1}{4}\ell p^{t-1}.$$

We form the subset $\mathcal{Q}\subseteq\mathcal{Q}_2$ by retaining at most one point on each line through the origin. The above calculation implies that $v_1,\dots,v_{s\ell}$ can be chosen such that $|\mathcal{Q}|\geq\frac{1}{4}\ell p^{t-1}$. On the other hand, the size of $\mathcal{Q}$ is bounded by $|\mathcal{Q}|\leq|\mathcal{Q}_2|\leq s\ell|\mathcal{Q}_0|\leq s\ell p^{t-1}$ and $\mathcal{Q}$ is a $(d-t+1,\ell s^2)$-subspace evasive set.

Finally, we choose $\mathcal{H}$ to be the set of hyperplanes given by the equations $\langle q,x\rangle=b$, for $q\in\mathcal{Q}$ and $b\in\mathbb{F}_p$. We ensured that no two points $q_1,q_2\in\mathcal{Q}$ lie on the same line through the origin, meaning that the hyperplanes given by $\langle q_1,x\rangle=b_1$, $\langle q_2,x\rangle=b_2$ are distinct for all $q_1,q_2\in\mathcal{Q}$ and $b_1,b_2\in\mathbb{F}_p$.

All that is left is to show that the conditions of the theorem hold for the sets $\mathcal{P},\mathcal{H}$. We have already shown that $|\mathcal{P}|\in[kp^t/2,kp^t]$. Furthermore, since every pair $(q,b)\in\mathcal{Q}\times\mathbb{F}_p$ yields a different hyperplane of $\mathcal{H}$, we conclude $|\mathcal{H}|=p|\mathcal{Q}|\in\big[\frac{1}{4}\ell p^t, s\ell p^t\big]$. Finally, for every $x\in\mathcal{P}$ and every $q\in\mathcal{Q}$, there is a unique b for which $\langle x,q\rangle=b$, and therefore the number of hyperplanes containing x is precisely $|\mathcal{Q}|=|\mathcal{H}|/p$. This verifies (i).

Now, we prove that for any $a\in\{1,\dots,t-1\}$, no $(d-t+a)$-dimensional affine subspace contains skp^a points of A. Suppose this is not the case and there was a set $A\subseteq P$ of size $|A|\geq skp^a$ which was contained in a $(d-t+a)$-dimensional affine space V. Note that V can be partitioned into p^a translates of some $(d-t)$-dimensional affine subspace $V'<V$. But then there exists a translate of V' containing at least ks points of $\mathcal{P}$.This is a contradiction since $\mathcal{P}$ is a $(d-t,ks)$-subspace evasive set, thus establishing (ii).

We proceed by a similar argument to prove (iii). Let B be a set of $s^2\ell p^b$ hyperplanes, denoted by $H_1,\dots,H_{s^2\ell p^b}$. Furthermore, we assume that the equation of H_i is given by $\langle q_i,x\rangle=b_i$. Observe that if the intersection of the hyperplanes H_i is nonempty, one must have $q_i\neq q_j$ for all $i,j\in[s^2\ell p^b]$, since otherwise two parallel hyperplanes belong to B. Let us denote the set of all points q_i for $i\in[s^2\ell p^b]$ by $B'\subseteq\mathcal{Q}$. Assume, for contradiction, that $\bigcap_{i=1}^{s^2\ell p^b}H_i$ contains a $(t-1-b)$-dimensional affine subspace. This affine subspace can be written as $V+q$ for some $q\in\mathbb{F}_p^d$ and some $(t-1-b)$-dimensional linear subspace V. Then all points q_i are contained in the orthogonal complement of

V, which we denote by $V^\perp$. But $V^\perp$ is a $(d-t+1+b)$-dimensional linear space, which contains $s^2\ell p^b$ points of $\mathcal{Q}$. By partitioning this space into p^b translates of a $(d-t+1)$-dimensional affine space, and using that $\mathcal{Q}$ is a $(d-t+1, \ell s^2)$-subspace evasive set, one arrives at a contradiction as above. $\square$

In what follows, we prove Theorem 5.5 in a series of lemmas, addressing different regimes.

Lemma 5.7. *Let $\varepsilon \in (0,1)$, and let d and n, m be integers which satisfy*

$$\varepsilon \leq \frac{1}{4}\max\{m,n\}^{-\frac{1}{d}}.$$

Then there exists a prime p, a set $\mathcal{P}$ of n points and a set $\mathcal{H}$ of m hyperplanes in $\mathbb{F}_p^d$ such that $I(\mathcal{P},\mathcal{H}) \geq \varepsilon mn$ and $\operatorname{rs}(\mathcal{P},\mathcal{H}) \leq (2s)^4\varepsilon\max\{m,n\}$, where s is given by Lemma 5.6.

Proof. For notational convenience, we assume that $\max\{m,n\} = m$, noting that the same argument applies if $\max\{m,n\} = n$. Let p be a prime satisfying $\varepsilon^{-1}/2 < p \leq \varepsilon^{-1}$, and set $t = \lfloor \log_p 4m \rfloor$ and $\ell = \lceil \frac{4m}{p^t} \rceil \leq p$. Note that $4m \leq \ell p^t \leq 8m$. Since $p > \varepsilon^{-1}/2 \geq 2m^{1/d}$, we have $p^d > 4m$ and so $t \leq d-1$. Thus, Lemma 5.6 can be applied with $k = \ell$ to get a set of points $\mathcal{P}_0$ and a set of hyperplanes $\mathcal{H}_0$ with $n_0 = |\mathcal{P}_0| \geq \ell p^t/2$, $m_0 = |\mathcal{H}_0| \geq \ell p^t/4$ and $I(\mathcal{P}_0,\mathcal{H}_0) = \frac{1}{p}m_0n_0$, satisfying (i), (ii) and (iii).

Let us now argue that $\operatorname{rs}(\mathcal{P}_0,\mathcal{H}_0) \leq (2s)^4\varepsilon m$. Let $A \subseteq \mathcal{P}_0, B \subseteq \mathcal{H}_0$ such that every point of A lies on every hyperplane in B. We consider three cases: either $|A| \geq s^2\ell p^{t-2}$, $|B| \geq s^2\ell p^{t-2}$ or $\max\{|A|,|B|\} < s^2\ell p^{t-2}$. In all cases, we show that

$$|A|\cdot|B| \leq s^4\ell p^{t-1} \leq 2s^4\varepsilon\ell p^t \leq 16s^4\varepsilon m.$$

If $|A| \geq s^2\ell p^{t-2}$, then by (ii) of Lemma 5.6, the points of A span an affine subspace of dimension at least $d-t+t-1 = d-1$, which means that at most one of the hyperplanes contains them all. On the other hand, again by (ii) of Lemma 5.6, any hyperplane can contain at most $s\ell p^{t-1}$ points of $\mathcal{P}_0$ and thus $|A| \leq s\ell p^{t-1}$. In conclusion, we have $|A| \leq s\ell p^{t-1}$ and $|B| = 1$, implying $|A|\cdot|B| \leq s\ell p^{t-1}$ as claimed.

The case $|B| \geq s^2\ell p^{t-2}$ is very similar. By (iii) of Lemma 5.6, the intersection of hyperplanes of B does not contain a line, meaning that $\bigcap_{H\in B} H$ is a single point. Furthermore, there are exactly $|\mathcal{H}_0|/p$ hyperplanes of $\mathcal{H}_0$ containing any point of $\mathcal{P}_0$. Since $|\mathcal{H}_0| \leq \ell s p^t$, we have $|B| \leq |\mathcal{H}_0|/p \leq s\ell p^{t-1}$, and therefore $|A|\cdot|B| \leq s\ell p^{t-1}$.

Finally, if $\max\{|A|,|B|\} < s^2\ell p^{t-2}$, we let a and b be the smallest integers for which $s^2\ell p^a \leq |A| < s^2\ell p^{a+1}$ and $s^2\ell p^b \leq |B| < s^2\ell p^{b+1}$. By property (ii) of Lemma 5.6, the points of A span an at least $(d-t+a+1)$-dimensional affine space. Similarly, by property (iii) of Lemma 5.6, the hyperplanes of B intersect in at most a $(t-2-b)$-dimensional space. Since all points of A belong to all hyperplanes of B, it must be that $d-t+a+1 \leq t-2-b$, i.e. $a+b+2 \leq 2t-d-1$. Thus,

$$|A|\cdot|B| \leq s^4\ell^2p^{a+b+2} \leq s^4\ell p\cdot p^{2t-d-1} \leq s^4\ell p^{t-1},$$

where the last inequality holds by noting that $t \leq d-1$. This completes the third case.

To finish the proof, let $\mathcal{P}$ be a random subset of $\mathcal{P}_0$ containing n points and $\mathcal{H}$ be a random subset of $\mathcal{H}_0$ containing m hyperplanes. One should verify that $n \leq |\mathcal{P}_0|$ and $m \leq |\mathcal{H}_0|$, but this is clear since $|\mathcal{P}_0|, |\mathcal{H}_0| \geq \frac{1}{4}\ell p^t \geq m$ and $m \geq n$. We also have $\operatorname{rs}(\mathcal{P},\mathcal{H}) \leq \operatorname{rs}(\mathcal{P}_0,\mathcal{H}_0) \leq (2s)^4\varepsilon m$. Furthermore, the expected number of incidences between $\mathcal{P}$ and $\mathcal{H}$ is

$$\mathbb{E}\big[I(P,\mathcal{H})\big] = \frac{n}{n_0}\frac{m}{m_0}I(P_0,\mathcal{H}_0) = \frac{1}{p}mn \geq \varepsilon mn.$$

Thus, there exist subsets $\mathcal{P},\mathcal{H}$ of $\mathcal{P}_0,\mathcal{H}_0$ with at least εmn incidences and $\operatorname{rs}(P,\mathcal{H}) \leq (2s)^4\varepsilon m$. $\square$

Lemma 5.8. *Let $\varepsilon \in (0,1)$, and let d and $n \leq m$ be positive integers which satisfy*

$$\varepsilon \leq \frac{1}{4} n^{-\frac{1}{d-1}}.$$

Then there exists a prime power q, a set $\mathcal{P}$ of n points and a set $\mathcal{H}$ of m hyperplanes in $\mathbb{F}_q^d$ such that $I(\mathcal{P},\mathcal{H}) \geq \varepsilon mn$ and $\operatorname{rs}(\mathcal{P},\mathcal{H}) \leq (4s)^4 \varepsilon m$, where s is given by Lemma 5.6.

Proof. The main idea of the proof is to take a symmetric $(d-1)$-dimensional construction described in Lemma 5.7 and "blow it up" to obtain our desired configuration.

More precisely, we apply Lemma 5.7 to find a set of n points $\mathcal{P}_0$ and a set of n hyperplanes $\mathcal{H}_0$ in $\mathbb{F}_p^{d-1}$ for some prime p for which $I(\mathcal{P}_0,\mathcal{H}_0) \geq \varepsilon n^2$ and $\operatorname{rs}(\mathcal{P}_0,\mathcal{H}_0) \leq (2s)^4\varepsilon n$. The hyperplanes and points may be considered in the larger ambient space $\mathbb{F}_q^{d-1}$, for any prime power $q = p^k$. Passing to $\mathbb{F}_q^{d-1}$ does not change the incidence graph. In the rest of the proof, we fix q to be a power of p larger than m/n.

Now, for each $x \in \mathbb{F}_q^{d-1}$, let $x' \in \mathbb{F}_q^d$ be the point whose first $d-1$ coordinates are the same as x, and the last coordinate is 0. Set $\mathcal{P} = \{x' | x \in \mathcal{P}_0\}$. Furthermore, we fix a set $S = \{s_1, \dots, s_\ell\} \subseteq \mathbb{F}_q$ of size $\ell = \lceil m/n \rceil$. Then, for every hyperplane $H \in \mathcal{H}_0$, if H is given by the equation $\langle a, x\rangle = a_1x_1 + \dots + a_{d-1}x_{d-1} = b$, we define hyperplanes $H^{(1)}, \dots, H^{(\ell)}$ in $\mathbb{F}_q^d$, where $H^{(i)}$ is given by the equations $a_1x_1 + \dots + a_{d-1}x_{d-1} + s_i x_d = b$. Then, we let $\mathcal{H}_1 = \{H^{(i)} | i \in [\ell], H \in \mathcal{H}_0\}$.

The sizes of the newly constructed sets are $|\mathcal{P}| = n, |\mathcal{H}_1| = n\lceil \frac{m}{n} \rceil > m$. Furthermore, if $x \in \mathcal{P}_0$ is incident to $H \in \mathcal{H}_0$, then x' is incident to all of the hyperplanes $H^{(1)}, \dots, H^{(\ell)}$. Thus, the number of incidences between $\mathcal{P}$ and $\mathcal{H}_1$ is $I(\mathcal{P},\mathcal{H}_1) \geq \ell I(\mathcal{P}_0,\mathcal{H}_0) \geq \varepsilon \ell |\mathcal{P}_0||\mathcal{H}_0| = \varepsilon |\mathcal{P}||\mathcal{H}_1|$. Finally, the size of the largest complete bipartite graph is bounded by

$$\operatorname{rs}(\mathcal{P},\mathcal{H}_1) \leq \ell \cdot \operatorname{rs}(\mathcal{P}_0,\mathcal{H}_0) \leq \ell \cdot (2s)^4 \varepsilon n \leq (4s)^4 \varepsilon m.$$

Taking a random m element subset $\mathcal{H}$ of $\mathcal{H}_1$, the expectation of $I(\mathcal{P},\mathcal{H})$ is at least εmn, so there is a choice for $\mathcal{H}$ such that $(\mathcal{P},\mathcal{H})$ satisfies our desired conditions. $\square$

Lemma 5.9. *Let $\varepsilon \in (0,1)$, and let d and $n \geq m$ be positive integers which satisfy*

$$\varepsilon \leq \frac{1}{4} m^{-\frac{1}{d-1}}.$$

Then there exists a prime power q, a set $\mathcal{P}$ of n points and a set $\mathcal{H}$ of m hyperplanes in $\mathbb{F}_q^d$ such that $I(\mathcal{P},\mathcal{H}) \geq \varepsilon mn$ and $\operatorname{rs}(\mathcal{P},\mathcal{H}) \leq (4s)^4 \varepsilon n$, where s is given by Lemma 5.6.

Proof. The proof is verbatim the same as the proof of Lemma 5.8, the only difference being that we "blow up" points instead of hyperplanes. More precisely, we start from a set of m points $\mathcal{P}_0$ and a set of m hyperplanes $\mathcal{H}_0$ in $\mathbb{F}_p^{d-1}$ for which $I(\mathcal{P}_0,\mathcal{H}_0) \geq \varepsilon m^2$ and $\operatorname{rs}(\mathcal{P}_0,\mathcal{H}_0) \leq (2s)^4 \varepsilon m$ and consider them in $\mathbb{F}_q^{d-1}$, where $q > n/m$.

Then, we fix a set $S = \{s_1, \dots, s_\ell\} \subseteq \mathbb{F}_q$ of size $\ell = \lceil n/m \rceil$ and for each point $x \in \mathbb{F}_q^{d-1}$ we define $x^{(i)} \in \mathbb{F}_q^d$ be the point whose first $d-1$ coordinates are the same as x, and the last coordinate is s_i. Set $\mathcal{P}_1 = \{x^{(i)} | x \in \mathcal{P}_0, i \in [\ell]\}$. Then, for every hyperplane $H \in \mathcal{H}_0$, if H is given by the equation $\langle a, x\rangle = a_1x_1 + \dots + a_{d-1}x_{d-1} = b$, we define the hyperplane H' in $\mathbb{F}_q^d$, given by the equation $a_1x_1 + \dots + a_{d-1}x_{d-1} + 0 \cdot x_d = b$. Then, we let $\mathcal{H} = \{H' | H \in \mathcal{H}_0\}$.

The sizes of the newly constructed sets are $|\mathcal{H}| = m, |\mathcal{P}_1| = m\lceil \frac{n}{m} \rceil > n$. Thus, by choosing a random subset $\mathcal{P} \subseteq \mathcal{P}_1$ and performing the same calculation as in the proof of Lemma 5.8 one arrives at the desired conclusion. $\square$

Lemma 5.10. *Let $\varepsilon \in (0,1)$, and let d and n, m be positive integers which satisfy*

$$4\varepsilon \geq m^{-\frac{1}{d-2}} \text{ and } 4\varepsilon \geq n^{-\frac{1}{d-2}}.$$

Then there exists a prime power q, a set $\mathcal{P}$ of n points and a set $\mathcal{H}$ of m hyperplanes in $\mathbb{F}_q^d$ such that $I(\mathcal{P},\mathcal{H}) \geq \varepsilon mn$ and $\operatorname{rs}(\mathcal{P},\mathcal{H}) \leq 2^6 s^4 (4\varepsilon)^{d-1} mn$, where s is given by Lemma 5.6.

Proof. The proof of this lemma is somewhat similar to the proofs of Lemmas 5.8 and 5.9, since we use the idea of "blowing-up" a lower-dimensional construction. However, in this case we use a $(d-2)$-dimensional configuration as a starting point and we "blow-up" both points and hyperplanes.

Let p be a prime between $\varepsilon^{-1}/2$ and ε^{-1}, and set $N = (4\varepsilon)^{-(d-2)}$. We remark that the inequalities $N \leq m, n$ are satisfied. Note that the condition of Lemma 5.7 is also satisfied with $d-2$ instead of d, and N instead of m and n. Therefore, we can find sets $\mathcal{P}_0, \mathcal{H}_0$ of points and hyperplanes in $\mathbb{F}_p^{d-2}$, both of size N, and satisfying $I(\mathcal{P}_0, \mathcal{H}_0) \geq \varepsilon N^2$ and $\mathrm{rs}(\mathcal{P}_0, \mathcal{H}_0) \leq (2s)^4\varepsilon N$.

As in the proof of Lemma 5.8, we consider $\mathcal{P}_0, \mathcal{H}_0$ in $\mathbb{F}_q^{d-2}$ instead of $\mathbb{F}_p^{d-2}$, where q is the power of p satisfying $q > \max\{\frac{m}{N}, \frac{n}{N}\}$. Furthermore, we fix two sets $S, T \subseteq \mathbb{F}_q$, such that $S = \{s_1, \dots, s_k\}$ and $T = \{t_1, \dots, t_\ell\}$, where $k = \lceil \frac{n}{N} \rceil, \ell = \lceil \frac{m}{N} \rceil$. To every point $x \in \mathcal{P}_0$, we associate k points of $\mathbb{F}_q^d$ as follows. For $i \in [k]$, let $x^{(i)} \in \mathbb{F}_q^d$ be equal to x on the first $d-2$ coordinates, equal to s_i on the coordinate $d-1$, and equal to 0 on the last coordinate. Similarly, for a hyperplane $H \in \mathcal{H}_0$ given by the equation $a_1x_1 + \dots + a_{d-2}x_{d-2} = b$, we associate the hyperplanes $H^{(j)}$ in $\mathbb{F}_q^d$ for $j \in [\ell]$ defined by the equations $a_1x_1 + \dots + a_{d-2}x_{d-2} + 0 \cdot x_{d-1} + t_jx_d = b$. Finally, we set $\mathcal{P}_1 = \{x^{(i)} | i \in [k], x \in \mathcal{P}_0\}$ and $\mathcal{H}_1 = \{H^{(j)} | H \in \mathcal{H}_0, j \in [\ell]\}$.

If $x \in H$, then it is easy to check that $x^{(i)} \in H^{(j)}$ for all $i \in [k], j \in [\ell]$, and no other type of incidences emerge in $(\mathcal{P}_1, \mathcal{H}_1)$. Thus, $I(\mathcal{P}_1, \mathcal{H}_1) \geq k\ell I(\mathcal{P}_0, \mathcal{H}_0) \geq \varepsilon k|\mathcal{P}_0|\ell|\mathcal{H}_0| = \varepsilon|\mathcal{P}_1||\mathcal{H}_1|$. Furthermore,

$$\mathrm{rs}(\mathcal{P}_1, \mathcal{H}_1) \leq k\ell\, \mathrm{rs}(\mathcal{P}_0, \mathcal{H}_0) \leq \frac{4mn}{N^2} \cdot (2s)^4\varepsilon N = 2^6 s^4 \varepsilon \frac{mn}{N}.$$

Recalling that $N = (4\varepsilon)^{-(d-2)}$, we obtain $\mathrm{rs}(\mathcal{P}_1, \mathcal{H}_1) \leq 2^6 s^4 (4\varepsilon)^{d-1} mn$. Subsampling a random n element subset $\mathcal{P}$ of $\mathcal{P}_1$, and a random m element subset $\mathcal{H}$ of $\mathcal{H}_1$, completes the proof. □

Proof of Theorem 5.5. We claim that it suffices to take $C = 4^{d+4}s^4 = O((4d)^{4d})$. Consider three cases.

Case 1. $\varepsilon \geq \frac{1}{4}\max\{m,n\}^{-1/d}$ and $\varepsilon \geq \frac{1}{4}\min\{m,n\}^{-1/(d-1)}$.

Then $4\varepsilon > m^{-1/(d-2)}$ and $4\varepsilon > n^{-1/(d-2)}$ are also satisfied, so Lemma 5.10 applies. Therefore, we have $\mathrm{rs}_d(m, n, \varepsilon mn) \leq C\varepsilon^{d-1}mn$.

Case 2. $\varepsilon \leq \frac{1}{4}\max\{m,n\}^{-1/d}$.

In this case, Lemma 5.7 shows that $\mathrm{rs}_d(m, n, \varepsilon mn) \leq (2s)^4\varepsilon \max\{m,n\} \leq C\varepsilon \max\{m,n\}$.

Case 3. $\varepsilon \leq \frac{1}{4}\min\{m,n\}^{-1/(d-1)}$.

We apply either Lemma 5.8 or Lemma 5.9, depending on whether $m \geq n$ or $n \geq m$. This shows that $\mathrm{rs}_d(m, n, \varepsilon mn) \leq (4s)^4\varepsilon \max\{m,n\} \leq C\varepsilon \max\{m,n\}$.

□

6 Concluding remarks

In this paper we proved tight bounds on the maximum number of incidences between n points and m hyperplanes in a d-dimensional space $\mathbb{F}^d$, assuming their incidence graph is $K_{s,s}$-free. More precisely, for every (m, n), we can find a field $\mathbb{F} = \mathbb{F}(d, m, n)$ for which our upper bounds can be realized up to a constant factor depending only on d and s, for sufficiently large $s \geq s_0(d)$. It would be interesting to improve our upper bounds for a fixed finite field $\mathbb{F}_p$ and arbitrary m and n. In two dimensions such estimates were obtained by Bourgain, Katz, and Tao [4].

Another interesting problem is to understand the size of the largest bipartite graph in dense incidence graphs. Given a set of n points $\mathcal{P}$ and a set of m hyperplanes $\mathcal{H}$ in $\mathbb{R}^d$ with $I(\mathcal{P}, \mathcal{H}) \geq \frac{mn}{2}$, what are the optimal bounds on the minimum of $\mathrm{rs}(\mathcal{P}, \mathcal{H})$? Singer and Sudan [25] (based on unpublished works of Lovett and Pálvölgyi and Fox, Wigderson) proved that there is a configuration of n points $\mathcal{P}$ and n linear hyperplanes $\mathcal{H}$ such that the points and the normal vectors of the hyperplanes have only 0-1 coordinates, $I(\mathcal{P}, \mathcal{H}) \geq n^2/2$ and $\mathrm{rs}(\mathcal{P}, \mathcal{H}) = n^2 2^{-\Theta(\sqrt{d})}$. Lovett [21] conjectured that this bound is the best possible, using the equivalent terminology of low-rank matrices. This problem might be also interesting over fields other than $\mathbb{R}$. For every field $\mathbb{F}$ the lower bound $\mathrm{rs}(\mathcal{P}, \mathcal{H}) = \Omega(\frac{mn}{2^{2d}d})$ follows from a recent result of Singer and Sudan [25]. On the other hand, by taking all points and all hyperplanes in $\mathbb{F}_2^d$, we have $n = 2^d$, $m = 2(2^d - 1)$, $I = \frac{mn}{2}$, and it is not difficult to show that every complete bipartite graph in the incidence graph has at most $2^d \approx \frac{mn}{2^{d+1}}$ edges.

References

[1] E. Aksoy Yazici, B. Murphy, M. Rudnev and I. Shkredov. *Growth estimates in positive characteristic via collisions*, Int. Math. Res. Not. 23 (2017): 7148–7189.

[2] R. Apfelbaum, and M. Sharir. *Large Complete Bipartite Subgraphs In Incidence Graphs Of Points And Hyperplanes.* SIAM Journal on Discrete Mathematics 21.3 (2007): 707–725.

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

[4] J. Bourgain, N. Katz, and T. Tao. *A sum-product estimate in finite fields, and applications.* Geom. funct. anal. 14 (2004): 27–57.

[5] P. Brass, and C. Knauer. *On counting point-hyperplane incidences.* Comput. Geom. Theory Appls. 25 (2003): 13–20.

[6] B. Chazelle. *Cutting hyperplanes for divide-and-conquer.* Discrete Comput. Geom. 9 (1993): 145–158.

[7] K. L. Clarkson. *New applications of random sampling in computational geometry.* Discrete Comput. Geom. 2 (1987): 195–222.

[8] T. Do. *The Largest Complete Bipartite Subgraph in Point-Hyperplane Incidence Graphs.* The Electronic Journal of Combinatorics (2020): 27.1.

[9] Z. Dvir. *Incidence Theorems and Their Applications.* Foundations and Trends in Theoretical Computer Science 6 (2012): 257–393.

[10] Z. Dvir, and S. Lovett. *Subspace evasive sets.* in Proceedings of the forty-fourth annual ACM symposium on Theory of computing (2012): 351–358.

[11] H. Edelsbrunner. *Algorithms in Combinatorial Geometry.* volume 10 of EATCS Monographs on Theoretical Computer Science. Springer-Verlag, 1987.

[12] G. Elekes. *On the number of sums and products.* Acta Arithmetica 81.4 (1997): 365–367.

[13] J. Erickson. *New lower bounds for Hopcroft's problem.* Discret. Comput. Geom., 16(4) (1996): 389–418.

[14] J. Fox, J. Pach, A. Sheffer, A. Suk, and J. Zahl. *A semi-algebraic version of Zarankiewicz's problem.* J. Eur. Math. Soc. 19 (2017): 1785–1810.

[15] J. Fox, J. Pach, and A. Suk. *A Polynomial Regularity Lemma for Semialgebraic Hypergraphs and Its Applications in Geometry and Property Testing.* SIAM Journal on Computing 45.6 (2016): 2199–2223.

[16] J. Fox, and B. Sudakov. *Dependent random choice.* Random Structures and Algorithms 38 (2011): 1–32.

[17] L. Guth, and N. H. Katz. *On the Erdős distinct distances problem in the plane.* Annals of Math. 181 N1 (2015): 155–190.

[18] J. Kollár. *Szemerédi–Trotter-type theorems in dimension 3.* Advances in Mathematics 271, (2015): 30–61.

[19] T. Kővári, V. Sós, and P. Turán. *On a problem of K. Zarankiewicz.* Colloq. Math. 3 (1954): 50–57.

[20] L. Lovász, and M. Saks. *Lattices, Möbius functions and communication complexity.* Annual Symposium on Foundations of Computer Science (1988): 81–90.

[21] S. Lovett. *Communication Is Bounded by Root of Rank.* In: Journal of the ACM (2016) 63.1. Conference version in STOC (2014): 1–9.

[22] A. Mohammadi, and S. Stevens. *Attaining the Exponent 5/4 for the Sum-Product Problem in Finite Fields*, Int. Math. Res. Not. (2023): 3516–3532.

[23] O. Roche-Newton, M. Rudnev, and I. D. Shkredov. *New sum-product type estimates over finite fields.* Advances in Mathematics 293 (2016): 589–605.

[24] M. Rudnev. *On the number of incidences between points and planes in three dimensions.* Combinatorica 38 (2018): 219–254.

[25] N. Singer, and M. Sudan. *Point-Hyperplane Incidence Geometry and the Log-Rank Conjecture.* preprint, arXiv:2101.09592.

[26] J. Solymosi, and T. Tao. *An Incidence Theorem in Higher Dimensions*, Discrete Comput. Geom. 48 (2012): 255-–280.

[27] S. Stevens, and F. de Zeeuw. *An improved point-line incidence bound over arbitrary fields*, Bull. London Math. Soc., 49 (2017): 842–858.

[28] B. Sudakov, and I. Tomon. *Ramsey properties of algebraic graphs and hypergraphs.* Forum of Mathematics, Sigma 10 (2022): e95.

[29] B. Sudakov, and I. Tomon. *Evasive sets, covering by subspaces, and point-hyperplane incidences.* Discrete Comput. Geom. (2023). doi:10.1007/s00454-023-00601-1

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

[31] F. de Zeeuw. *A short proof of Rudnev's point-plane incidence bound.* preprint, arXiv:1612.02719.