J. Geom.

DOI 10.1007/s00022-016-0323-5

Journal of Geometry

General position subsets and independent hyperplanes in d-space

Jean Cardinal, Csaba D. Tóth, and David R. Wood

Abstract. Erdős asked what is the maximum number $\alpha(n)$ such that every set of n points in the plane with no four on a line contains $\alpha(n)$ points in general position. We consider variants of this question for d-dimensional point sets and generalize previously known bounds. In particular, we prove the following two results for fixed d:

- Every set $\mathcal{H}$ of n hyperplanes in $\mathbb{R}^d$ contains a subset $S \subseteq \mathcal{H}$ of size at least $c\,(n \log n)^{1/d}$, for some constant $c = c(d) > 0$, such that no cell of the arrangement of $\mathcal{H}$ is bounded by hyperplanes of S only.
- Every set of $cq^d \log q$ points in $\mathbb{R}^d$, for some constant $c = c(d) > 0$, contains a subset of q cohyperplanar points or q points in general position.

Two-dimensional versions of the above results were respectively proved by Ackerman et al. [*Electronic J. Combinatorics*, 2014] and by Payne and Wood [*SIAM J. Discrete Math.*, 2013].

Mathematics Subject Classification. Primary 52C35; Secondary 52C10.

1. Introduction

Points in general position A finite set of points in $\mathbb{R}^d$ is said to be in *general position* if no hyperplane contains more than d points. Given a finite set of points $P \subset \mathbb{R}^d$ in which at most $d+1$ points lie on a hyperplane, let $\alpha(P)$ be the size of a largest subset of P in general position. Let $\alpha(n,d) = \min\{\alpha(P) : |P| = n\}$.

For $d = 2$, Erdős [5] observed that $\alpha(n,2) \gtrsim \sqrt{n}$ and proposed the determination of $\alpha(n,2)$ as an open problem.[1] Füredi [6] proved $\sqrt{n \log n} \lesssim \alpha(n,2) \leq o(n)$, where the lower bound uses independent sets in Steiner triple systems,

Research of Wood is supported by the Australian Research Council.

[1] We use the shorthand notation $\lesssim$ to indicate inequality up to a constant factor for large n. Hence $f(n) \lesssim g(n)$ is equivalent to $f(n) \in O(g(n))$, and $f(n) \gtrsim g(n)$ is equivalent to $f(n) \in \Omega(g(n))$.

Published online: 11 March 2016

and the upper bound relies on the density version of the Hales–Jewett Theorem [7,8]. Füredi's argument combined with the quantitative bound for the density Hales–Jewett problem proved in the first polymath project [14] yields $\alpha(n,2) \lesssim n/\sqrt{\log^* n}$ (Theorem 2.2).

Our first goal is to derive upper and lower bounds on $\alpha(n,d)$ for fixed $d \geq 3$. We prove that the multi-dimensional Hales–Jewett theorem [8] yields $\alpha(n,3) \in o(n)$ (Theorem 2.4). But for $d \geq 4$, only the trivial upper bound $\alpha(n,d) \in O(n)$ is known. We establish lower bounds $\alpha(n,d) \gtrsim (n \log n)^{1/d}$ in a dual setting of hyperplane arrangements in $\mathbb{R}^d$ as described below.

Independent sets of hyperplanes For a finite set $\mathcal{H}$ of hyperplanes in $\mathbb{R}^d$, Bose et al. [2] defined a hypergraph $G(\mathcal{H})$ with vertex set $\mathcal{H}$ such that the set of hyperplanes containing the facets of each cell of the arrangement of $\mathcal{H}$ forms a hyperedge in $G(\mathcal{H})$. A subset $S \subseteq \mathcal{H}$ of hyperplanes is called *independent* if it is an independent set of $G(\mathcal{H})$; that is, if no cell of the arrangement of $\mathcal{H}$ is bounded by hyperplanes in S only. Denote by $\beta(\mathcal{H})$ the maximum size of an independent set of $\mathcal{H}$, and let $\beta(n,d) := \min\{\beta(\mathcal{H}) : |\mathcal{H}| = n\}$.

The following relation between $\alpha(n,d)$ and $\beta(n,d)$ was observed by Ackerman et al. [1] in the case $d = 2$.

Lemma 1.1 (Ackerman et al. [1]). *For $d \geq 2$ and $n \in \mathbb{N}$, we have $\beta(n,d) \leq \alpha(n,d)$.*

Proof. For every set P of n points in $\mathbb{R}^d$ in which at most $d+1$ points lie on a hyperplane, we construct a set $\mathcal{H}$ of n hyperplanes in $\mathbb{R}^d$ such that $\beta(\mathcal{H}) \leq \alpha(P)$. Consider the set $\mathcal{H}_0$ of hyperplanes obtained from P by duality. Since at most $d+1$ points of P lie on a hyperplane, at most $d+1$ hyperplanes in $\mathcal{H}_0$ have a common intersection point. Perturb the hyperplanes in $\mathcal{H}_0$ so that the $d+1$ hyperplanes that intersect form a simplicial cell, and denote by $\mathcal{H}$ the resulting set of hyperplanes. An independent subset of hyperplanes corresponds to a subset in general position in P. Thus $\alpha(P) \geq \beta(\mathcal{H})$. □

Ackerman et al. [1] proved that $\beta(n,2) \gtrsim \sqrt{n \log n}$, using a result by Kostochka et al. [11] on independent sets in bounded-degree hypergraphs. Lemma 1.1 implies that any improvement on this lower bound would immediately improve Füredi's lower bound for $\alpha(n,2)$. We generalize the lower bound to higher dimensions by proving that $\beta(n,d) \gtrsim (n \log n)^{1/d}$ for fixed $d \geq 2$ (Theorem 3.3).

Subsets either in General Position or in a Hyperplane We also consider a generalization of the first problem, and define $\alpha(n,d,\ell)$, with a slight abuse of notation, to be the largest integer such that every set of n points in $\mathbb{R}^d$ in which at most ℓ points lie in a hyperplane contains a subset of $\alpha(n,d,\ell)$ points in general position. Note that $\alpha(n,d) = \alpha(n,d,d+1)$ with this notation, and every set of n points in $\mathbb{R}^d$ contains $\alpha(n,d,\ell)$ points in general position or $\ell+1$ points in a hyperplane.

Motivated by a question of Gowers [9], Payne and Wood [13] studied $\alpha(n,2,\ell)$; that is, the minimum, taken over all sets of n points in the plane with at most ℓ collinear, of the maximum size a subset in general position. They combine the Szemerédi–Trotter Theorem [17] with lower bounds on maximal independent sets in bounded-degree hypergraphs to prove $\alpha(n,2,\ell) \gtrsim \sqrt{n\log n/\log\ell}$ for $\ell \lesssim n^{1/2-\varepsilon}$. We generalize some of their techniques, and show that for fixed $d \geq 2$ and all $\ell \lesssim \sqrt{n}$, we have $\alpha(n,d,\ell) \gtrsim (n/\log\ell)^{1/d}$ (Theorem 4.1). It follows that every set of at least $Cq^d\log q$ points in $\mathbb{R}^d$, where $C = C(d) > 0$ is a sufficiently large constant, contains q cohyperplanar points or q points in general position (Corollary 4.2).

2. Subsets in general position and the Hales–Jewett theorem

Let $[k] := \{1,2,\dots,k\}$ for every positive integer k. A subset $S \subseteq [k]^m$ is a *t-dimensional combinatorial subspace* of $[k]^m$ if there exists a partition of $[m]$ into sets $W_1,W_2,\dots,W_t,X$ such that $W_1,W_2,\dots,W_t$ are nonempty, and S is exactly the set of elements $x \in [k]^m$ for which $x_i = x_j$ whenever $i,j \in W_\ell$ for some $\ell \in [t]$, and x_i is constant if $i \in X$. A one-dimensional combinatorial subspace is called a *combinatorial line*.

To obtain a quantitative upper bound for $\alpha(n,2)$, we combine Füredi's argument with the quantitative version of the density Hales–Jewett theorem for $k=3$ obtained in the first polymath project (Note that the latter also consider Moser numbers, involving geometric lines and not only combinatorial lines, but this is not needed here.)

Theorem 2.1 (Polymath [14]). *The size of the largest subset of $[3]^m$ without a combinatorial line is $O(3^m/\sqrt{\log^* m})$.*

Theorem 2.2. *$\alpha(n,2) \lesssim n/\sqrt{\log^* n}$.*

Proof. Consider the m-dimensional grid $[3]^m$ in $\mathbb{R}^m$ and project it onto $\mathbb{R}^2$ using a generic projection; that is, so that three points in the projection are collinear if and only if their preimages in $[3]^m$ are collinear. Denote by P the resulting planar point set and let $n = 3^m$. Since the projection is generic, the only collinear subsets of P are projections of collinear points in the original m-dimensional grid, and $[3]^m$ contains at most three collinear points. From Theorem 2.1, the largest subset of P with no three collinear points has size at most the indicated upper bound. □

To bound $\alpha(n,3)$, we use the multidimensional version of the density Hales–Jewett Theorem.

Theorem 2.3 (See [7,14]). *For every $\delta > 0$ and every pair of positive integers k and t, there exists a positive integer $M := M(k,\delta,t)$ such that for every $m > M$, every subset of $[k]^m$ of density at least δ contains a t-dimensional subspace.*

Theorem 2.4. *$\alpha(n,3) \in o(n)$.*

Proof. Consider the m-dimensional hypercube $[2]^m$ in $\mathbb{R}^m$ and project it onto $\mathbb{R}^3$ using a generic projection. Let P be the resulting point set in $\mathbb{R}^3$ and let $n := 2^m$. Since the projection is generic, the only coplanar subsets of P are projections of points of the m-dimensional grid $[2]^m$ lying in a two-dimensional subspace. Therefore P does not contain more than four coplanar points. From Theorem 2.3 with $k = t = 2$, for every $\delta > 0$ and sufficiently large m, every subset of P with at least δn elements contains $k^t = 4$ coplanar points. Hence every independent subset of P has $o(n)$ elements. □

We would like to prove $\alpha(n, d) \in o(n)$ for fixed d. However, we cannot apply the same technique, because an m-cube has too many co-hyperplanar points, which remain co-hyperpanar in projection. By the multidimensional Hales–Jewett theorem, every constant fraction of vertices of a high-dimensional hypercube has this property. It is a coincidence that a projection of a hypercube to $\mathbb{R}^d$ works for $d = 3$, because $2^{d-1} = d + 1$ in that case.

3. Lower bounds for independent hyperplanes

We also give a lower bound on $\beta(n, d)$ for $d \geq 2$. By a simple charging argument (see Cardinal and Felsner [3]), one can establish that $\beta(n, d) \gtrsim n^{1/d}$. Inspired by the recent result of Ackerman et al. [1], we improve this bound by a factor of $(\log n)^{1/d}$.

Lemma 3.1. *Let $\mathcal{H}$ be a finite set of hyperplanes in $\mathbb{R}^d$. For every subset of d hyperplanes in $\mathcal{H}$, there are at most 2^d simplicial cells in the arrangement of $\mathcal{H}$ such that all d hyperplanes contain some facets of the cell.*

Proof. A simplicial cell σ in the arrangement of $\mathcal{H}$ has exactly $d + 1$ vertices, and exactly $d + 1$ facets. Any d hyperplanes along the facets of σ intersect in a single point, namely at a vertex of σ. Every set of d hyperplanes in $\mathcal{H}$ that intersect in a single point can contain d facets of at most 2^d simplicial cells (since no two such cells can lie on the same side of all d hyperplanes). □

The following is a reformulation of a result of Kostochka et al. [11], that is similar to the reformulation of Ackerman et al. [1] in the case $d = 2$. In what follows, $d = O(1)$, and asymptotic notations refer to $n \to \infty$.

Theorem 3.2 (Kostochka et al. [11]). *Consider an n-vertex $(d + 1)$-uniform hypergraph H such that every d-tuple of vertices is contained in at most $t = O(1)$ edges, and apply the following procedure:*

1. *let X be the subset of vertices obtained by choosing each vertex independently at random with probability p, such that $pn = (n/(t \log\log\log n))^{3/(3d-1)}$,*
2. *remove the minimum number of vertices of X so that the resulting subset Y induces a triangle-free linear[2] hypergraph $H[Y]$.*

[2] A hypergraph is *linear* if it has no pair of distinct edges sharing two or more vertices.

Then with high probability $H[Y]$ has an independent set of size at least $\left(\frac{n}{t}\log\frac{n}{t}\right)^{\frac{1}{d}}$.

Theorem 3.3. *For fixed $d \geq 2$, we have $\beta(n,d) \gtrsim (n\log n)^{1/d}$.*

Proof. Let $\mathcal{H}$ be a set of n hyperplanes in $\mathbb{R}^d$ and consider the $(d+1)$-uniform hypergraph H having one vertex for each hyperplane in $\mathcal{H}$, and a hyperedge of size $d+1$ for each set of $d+1$ hyperplanes forming a simplicial cell in the arrangement of $\mathcal{H}$. From Lemma 3.1, every d-tuple of vertices of H is contained in at most $t := 2^d$ edges. We can apply Theorem 3.2 and obtain a subset S of hyperplanes of size $\Omega\left(((\frac{n}{2^d})\log(\frac{n}{2^d}))^{1/d}\right)$ such that no simplicial cell is bounded by hyperplanes of S only. However, there might be nonsimplicial cells of the arrangement that are bounded by hyperplanes of S only.

Let p be the probability used to define X in Theorem 3.2. It is known [10] that the total number of cells in an arrangement of d-dimensional hyperplanes is less than dn^d. Hence for an integer $c \geq d+1$, the expected number of cells of size c that are bounded by hyperplanes of X only is at most

$$p^c dn^d \leq \frac{n^{(4-3d)c/(3d-1)}}{(2^d \log\log\log n)^{3/(3d-1)}} \cdot dn^d \lesssim dn^{(4-3d)c/(3d-1)+d}.$$

Note that for $c \geq d+2$, the exponent of n satisfies

$$\frac{(4-3d)c}{3d-1} + d < 0.$$

Therefore the expected number of such cells of size at least $d+2$ is vanishing.

On the other hand we can bound the expected number of cells that are of size at most d, and that are bounded by hyperplanes of X only, where the expectation is again with respect to the choice of X. Note that cells of size d are necessarily unbounded, and in a simple arrangement, no cell has size less than d. The number of unbounded cells in a d-dimensional arrangement is $O(dn^{d-1})$ [10]. Therefore, the number we need to bound is at most

$$p^d O(dn^{d-1}) \lesssim n^{(4-3d)d/(3d-1)+d-1} \lesssim n^{1/(3d-1)} = o(n^{1/d}).$$

Consider now a maximum independent set S in the hypergraph $H[Y]$, where Y is defined as in Theorem 3.2, and for each cell that is bounded by hyperplanes of S only, remove from S one of the hyperplanes bounding the cell. Since $S \subseteq X$, the expected number of such cells is $o(n^{1/d})$, hence there exists an X for which the number of remaining hyperplanes in $S \subseteq X$ is still $\Omega\left((n\log n)^{1/d}\right)$, and they now form an independent set. $\square$

We have the following coloring variant of Theorem 3.3.

Corollary 3.4. *Hyperplanes of a simple arrangement of size n in $\mathbb{R}^d$ for fixed $d \geq 2$ can be colored with $O\left(n^{1-1/d}/(\log n)^{1/d}\right)$ colors so that no cell is bounded by hyperplanes of a single color.*

Proof. From Theorem 3.3, there always exists an independent set of hyperplanes of size at least $c\,(n\log n)^{1/d}$ for some constant c. We assume here that all logarithms are base 2. We define a new constant c' such that

$$c' = \left(\frac{1}{c} + c'\right) 2^{2/d-1} \Leftrightarrow c' = \frac{2^{2/d-1}}{c(1-2^{2/d-1})}.$$

We now prove that n hyperplanes forming a simple arrangement in $\mathbb{R}^d$ can be colored with $c'\left(n^{1-1/d}/(\log n)^{1/d}\right)$ colors so that no cell is bounded by hyperplanes of a single color. We proceed by induction and suppose this holds for $n/2$ hyperplanes. We apply the greedy algorithm and iteratively pick a maximum independent set until there are at most $n/2$ hyperplanes left. We assign a new color to each independent set, then use the induction hypothesis for the remaining hyperplanes. This clearly yields a proper coloring.

Since every independent set has size at least $c\left(\frac{n}{2}\log\frac{n}{2}\right)^{1/d}$, the number of iterations before we are left with at most $n/2$ hyperplanes is at most

$$t \le \frac{\frac{n}{2}}{c\left(\frac{n}{2}\log\frac{n}{2}\right)^{1/d}}.$$

The number of colors is therefore at most

$$\begin{aligned} t + c'\left(\frac{\left(\frac{n}{2}\right)^{1-1/d}}{\left(\log\frac{n}{2}\right)^{1/d}}\right) &\le \frac{\frac{n}{2}}{c\left(\frac{n}{2}\log\frac{n}{2}\right)^{1/d}} + c'\left(\frac{\left(\frac{n}{2}\right)^{1-1/d}}{\left(\log\frac{n}{2}\right)^{1/d}}\right) \\ &= \left(\frac{1}{c} + c'\right)\left(\frac{\left(\frac{n}{2}\right)^{1-1/d}}{\left(\log\frac{n}{2}\right)^{1/d}}\right) \\ &\le \left(\frac{1}{c} + c'\right)\left(2^{2/d-1}\frac{n^{1-1/d}}{(\log n)^{1/d}}\right) \\ &= c'\left(\frac{n^{1-1/d}}{(\log n)^{1/d}}\right), \end{aligned}$$

as claimed. In the penultimate line, we used the fact that $\log\frac{n}{2} > \frac{1}{2}\log n$ for $n > 4$. □

4. Large subsets in general position or in a hyperplane

We wish to prove the following.

Theorem 4.1. *Fix $d \ge 2$. Every set of n points in $\mathbb{R}^d$ with at most ℓ cohyperplanar points, where $\ell \lesssim n^{1/2}$, contains a subset of $\Omega\left((n/\log\ell)^{1/d}\right)$ points in general position. That is,*

$$\alpha(n,d,\ell) \gtrsim (n/\log\ell)^{1/d} \quad \textit{for } \ell \lesssim \sqrt{n}.$$

This is a higher-dimensional version of the result by Payne and Wood [13]. The following Ramsey-type statement is an immediate corollary.

Corollary 4.2. *For fixed $d \geq 2$ there is a constant c such that every set of at least $cq^d \log q$ points in $\mathbb{R}^d$ contains q cohyperplanar points or q points in general position.*

In order to give some intuition about Corollary 4.2, it is worth mentioning an easy proof when $cq^d \log q$ is replaced by $q \cdot \binom{q}{d}$. Consider a set of $n = q \cdot \binom{q}{d}$ points in $\mathbb{R}^d$, and let S be a maximal subset in general position. Either $|S| \geq q$ and we are done, or S spans $\binom{|S|}{d} \leq \binom{q}{d}$ hyperplanes, and, by maximality, every point lies on at least one of these hyperplanes. Hence by the pigeonhole principle, one of the hyperplanes in S must contain at least $n/\binom{q}{d} = q$ points.

We now use known incidence bounds to estimate the maximum number of cohyperplanar $(d+1)$-tuples in a point set. In what follows we consider a finite set P of n points in $\mathbb{R}^d$ such that at most ℓ points of P are cohyperplanar, where $\ell := \ell(n) \lesssim n^{1/2}$ is a fixed function of n. For $d \geq 3$, a hyperplane h is said to be γ-*degenerate* if at most $\gamma \cdot |P \cap h|$ points in $P \cap h$ lie on a $(d-2)$-flat. A flat is said to be k-rich whenever it contains *at least* k points of P. The following is a standard reformulation of the classic Szemerédi–Trotter theorem on point-line incidences in the plane [17].

Theorem 4.3 (Szemerédi and Trotter [17]). *For every set of n points in $\mathbb{R}^2$, the number of k-rich lines is at most*

$$O\left(\frac{n^2}{k^3} + \frac{n}{k}\right).$$

This bound is the best possible.

Elekes and Tóth proved the following higher-dimensional version, involving an additional non-degeneracy condition.

Theorem 4.4 (Elekes and Tóth [4]). *For every integer $d \geq 3$, there exist constants $C_d > 0$ and $\gamma_d > 0$ such that for every set of n points in $\mathbb{R}^d$, the number of k-rich γ_d-degenerate planes is at most*

$$C_d\left(\frac{n^d}{k^{d+1}} + \frac{n^{d-1}}{k^{d-1}}\right).$$

This bound is the best possible apart from constant factors.

We prove the following upper bound on the number of cohyperplanar $(d+1)$-tuples in a point set.

Lemma 4.5. *Fix $d \geq 2$. Let P be a set of n points in $\mathbb{R}^d$ with at most ℓ cohyperplanar points, where $\ell \lesssim n^{1/2}$. Then the number of cohyperplanar $(d+1)$-tuples in P is $O\left(n^d \log \ell\right)$.*

Proof. We proceed by induction on $d \geq 2$. The base case $d = 2$ was established by Payne and Wood [13], using the Szemerédi–Trotter bound (Theorem 4.3). We reproduce it here for completeness. We wish to bound the number of collinear triples in a set P of n points in the plane. Let h_k be the number of lines containing exactly k points of P. The number of collinear 3-tuples is

$$\begin{aligned}\sum_{k=3}^{\ell} h_k \binom{k}{3} &\leq \sum_{k=3}^{\ell} k^2 \sum_{i=k}^{\ell} h_i \\ &\lesssim \sum_{k=3}^{\ell} k^2 \left(\frac{n^2}{k^3} + \frac{n}{k}\right) \\ &\lesssim n^2 \log \ell + \ell^2 n \lesssim n^2 \log \ell.\end{aligned}$$

We now consider the general case $d \geq 3$. Let P be a set of n points in $\mathbb{R}^d$, no ℓ in a hyperplane, where $n \geq d+2$ and $\ell \lesssim \sqrt{n}$. let $\gamma := \gamma_d > 0$ be a constant specified in Theorem 4.4. We distinguish the following three types of $(d+1)$-tuples:

Type 1: $(d+1)$-tuples contained in some $(d-2)$-flat spanned by P Denote by s_k the number of $(d-2)$-flats spanned by P that contain exactly k points of P. Project P onto a $(d-1)$-flat in a generic direction to obtain a set of points P' in $\mathbb{R}^{d-1}$. Now s_k is the number of hyperplanes of P' containing exactly k points of P'. By applying the induction hypothesis on P', the number of cohyperplanar d-tuples is

$$\sum_{k=d}^{\ell} s_k \binom{k}{d} \lesssim n^{d-1/2} \log \ell.$$

Hence the number of $(d+1)$-tuples of P lying in a $(d-2)$-flat spanned by P satisfies

$$\sum_{k=d+1}^{\ell} s_k \binom{k}{d+1} \lesssim \ell n^{d-1} \log \ell \leq n^d \log \ell.$$

Type 2: $(d+1)$-tuples of P that span a γ-degenerate hyperplane Let h_k be the number of γ-degenerate hyperplanes containing exactly k points of P. By Theorem 4.4,

$$\begin{aligned}\sum_{k=d+1}^{\ell} h_k \binom{k}{d+1} &\leq \sum_{k=d+1}^{\ell} k^d \sum_{i=k}^{\ell} h_i \\ &\lesssim \sum_{k=d+1}^{\ell} k^d \left(\frac{n^d}{k^{d+1}} + \frac{n^{d-1}}{k^{d-1}}\right) \\ &\lesssim n^d \log \ell + \ell^2 n^{d-1} \lesssim n^d \log \ell.\end{aligned}$$

Type 3: $(d+1)$-tuples of P that span a hyperplane that is not γ-degenerate Recall that if a hyperplane H panned by P is not γ-degenerate, then more than a γ fraction of its points lie in a $(d-2)$-flat $L(H)$. We may assume that $L(H)$ is also spanned by P. Consider a $(d-2)$-flat L spanned by P and containing exactly k points of P. The hyperplanes spanned by P that contain L

partition $P\backslash L$. Let n_r be the number of hyperplanes containing L and exactly r points of $P\backslash L$. We have $\sum_{r=1}^{\ell} n_r r \leq n$.

If a hyperplane H is not γ-degenerate, contains a $(d-2)$-flat $L = L(H)$ with exactly k points, and r other points of P, then $k > \gamma(r+k)$, hence $r < (\frac{1}{\gamma}-1)k$. Furthermore, all $(d+1)$-tuples that span H must contain at least one point that is not in L. Hence the number of $(d+1)$-tuples that span H is at most $O(rk^d)$. The total number of $(d+1)$-tuples of type 3 that span a hyperplane H with a common $(d-2)$-flat $L = L(H)$ is therefore at most

$$\sum_{r=1}^{\ell} n_r r k^d \leq n k^d.$$

Recall that s_k denotes the number of $(d-2)$-flats containing exactly k points. Summing over all such $(d-2)$-flats and applying the induction hypothesis yields the following upper bound on the total number of $(d+1)$-tuples spanning hyperplanes that are not γ-degenerate:

$$\sum_{k=d+1}^{\ell} s_k n k^d \lesssim n^d \log \ell.$$

Summing over all three cases, the total number of cohyperplanar $(d+1)$-tuples is $O(n^d \log \ell)$ as claimed. □

In the plane, Lemma 4.5 gives an $O(n^2 \log \ell)$ bound for the number of collinear triples in an n-element point set with no ℓ on a line, where $\ell \in O(\sqrt{n})$. This bound is tight for $\ell = \Theta(\sqrt{n})$ for a $\lfloor \sqrt{n} \rfloor \times \lfloor \sqrt{n} \rfloor$ section of the integer lattice. It is almost tight for $\ell \in \Theta(1)$, Solymosi and Stojaković [15] recently constructed n-element point sets for every constant ℓ and $\varepsilon > 0$ that contains at most ℓ points on a line and $\Omega(n^{2-\varepsilon})$ collinear ℓ-tuples, hence $\Omega(n^{2-\varepsilon}\binom{\ell}{3}) \subset \Omega(n^{2-\varepsilon})$ collinear triples.

Armed with Lemma 4.5, we now apply the following standard result from hypergraph theory due to Spencer [16].

Theorem 4.6 (Spencer [16]). *Every r-uniform hypergraph with n vertices and m edges contains an independent set of size at least*

$$\frac{r-1}{r^{r/(r-1)}} \frac{n}{\left(\frac{m}{n}\right)^{1/(r-1)}}. \tag{1}$$

Proof of Theorem 4.1. We apply Theorem 4.6 to the hypergraph formed by considering all cohyperplanar $(d+1)$-tuples in a given set of n points in $\mathbb{R}^d$, with no ℓ cohyperplanar. Substituting $m \lesssim n^d \log \ell$ and $r = d+1$ in (1), we get a lower bound

$$\frac{n}{(n^{d-1} \log \ell)^{1/d}} = \left(\frac{n}{\log \ell}\right)^{1/d},$$

for the maximum size of a subset in general position, as desired. □

Note added in proof

Recall that in Sect. 2 we asked whether $\alpha(n, d) \in o(n)$ for every $d \geq 3$. Subsequently, Milićević [12] announced an answer to this question in the affirmative.

References

[1] Ackerman, E., Pach, J., Pinchasi, R., Radoičić, R., Tóth, G.:*A note on coloring line arrangements*. Electron. J. Comb. **21**(2), #P2.23 (2014)

[2] Bose, P., Cardinal, J., Collette, S., Hurtado, F., Korman, M., Langerman, S., Taslakian, P.: *Coloring and guarding arrangements*. Discret. Math. Theor. Comput. Sci. **15**(3), 139–154 (2013)

[3] Cardinal, J., Felsner, S.: Covering partial cubes with zones. In: Proceedings of the 16th Japan conference on discrete and computational geometry and graphs (JCDCG^2 2013), Lecture notes in computer science. Springer, Berlin (2014)

[4] Elekes, G., Tóth, C.D.: Incidences of not-too-degenerate hyperplanes. In: Proceedings of the ACM symposium on computational geometry (SoCG), pp. 16–21 (2005)

[5] Erdős, P.: *On some metric and combinatorial geometric problems*. Discret. Math. **60**, 147–153 (1986)

[6] Füredi, Z.: *Maximal independent subsets in Steiner systems and in planar sets*. SIAM J. Discret. Math. **4**(2), 196–199 (1991)

[7] Furstenberg, H., Katznelson, Y.: *A density version of the Hales–Jewett theorem for $k = 3$*. Discret. Math. **75**(1-3), 227–241 (1989)

[8] Furstenberg, H., Katznelson, Y.: *A density version of the Hales–Jewett theorem*. J. d'analyse Math. **57**, 64–119 (1991)

[9] Gowers, T.: A geometric Ramsey problem. http://mathoverflow.net/questions/50928/a-geometric-ramsey-problem (2012)

[10] Halperin, D.: Arrangements. In: Goodman, J.E., O'Rourke, J. (eds) Handbook of Discrete and Computational Geometry, chapter 24., 2nd edn., CRC Press, Boca Raton (2004)

[11] Kostochka, A.V., Mubayi, D., Verstraëte, J.: *On independent sets in hypergraphs*. Random Struct. Algorithms **44**(2), 224–239 (2014)

[12] Milićević, L.: Sets in almost general position (2015). arXiv:1601.07206

[13] Payne, M.S., Wood, D.R.: *On the general position subset selection problem*. SIAM J. Discret. Math. **27**(4), 1727–1733 (2013)

[14] Polymath, D.H.J.: *A new proof of the density Hales–Jewett theorem*. Ann. Math. **175**(2), 1283–1327 (2012)

[15] Solymosi, J., Stojaković, M.: *Many collinear k-tuples with no $k + 1$ collinear points*. Discret. Comput. Geom. **50**, 811–820 (2013)

[16] Spencer, J.: *Turán's theorem for k-graphs*. Discret. Math. **2**(2), 183–186 (1972)

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

General positions subsets and independent hyperplanes

Jean Cardinal
Université libre de Bruxelles (ULB)
Brussels
Belgium
e-mail: jcardin@ulb.ac.be

Csaba D. Tóth
California State University Northridge
Northridge, CA
USA
e-mail: csaba.toth@csun.edu

David R. Wood
Monash University
Melbourne
Australia
e-mail: david.wood@monash.edu

Received: October 17, 2015.