Extensions of the No-Three-In-Line Problem⋆

Hanno Lefmann

Fakultät für Informatik, TU Chemnitz, D-09107 Chemnitz, Germany
phone: (+49)-(0)-371-531-31276 fax: (+49)-(0)-371-531-25439
lefmann@informatik.tu-chemnitz.de

Abstract. We investigate generalizations of the No-Three-In-Line problem in $\mathbb{Z}^d$. For several pairs (k, ℓ) of given positive integers we give algorithmic lower, and upper bounds on the largest sizes of subsets S of points from the d-dimensional $T \times \cdots \times T$-grid, where no ℓ points in S are contained in a k-dimensional affine or linear subspace, respectively.

1 Introduction

The No-Three-in-Line problem, which has been raised originally by Dudeney [7], asks for the maximum number of points, which can be chosen from the $T \times T$-grid in $\mathbb{Z}^2$, i.e., from the set $\{0, \ldots, T-1\} \times \{0, \ldots, T-1\}$, such that no three points are on a line, see [5, 14]. Erdős [9] observed that this maximum number is $\Theta(T)$. The lower bound follows by considering for primes T the grid-points $(x, x^2 \bmod T)$, $x = 0, \ldots, T-1$. The upper bound is derived from the fact that each horizontal line may contain at most two grid-points. For constructions of (near-)optimal solutions for small values of T see Flammenkamp [10, 11].
Cohen, Eades, Lin and Ruskey [6] investigated compact embeddings of graphs into $\mathbb{Z}^3$ such that distinct edges (represented by segments) do not cross each other in a point distinct from the endpoints. Compact embeddings minimize the volume of an axis-aligned bounding box in $\mathbb{Z}^3$, which contains the drawing. The endpoints of crossing edges in a drawing of a graph are coplanar. In connection with this, it was proved in [6] that there exists a set of $\Omega(T)$ points in the $T \times T \times T$-grid, which does not contain four distinct coplanar points, and up to a constant factor this lower bound is best possible. Thus, the minimum volume of an axis-aligned bounding box for a crossing-free drawing of the complete graph K_n on n vertices in $\mathbb{Z}^3$ is equal to $\Theta(n^3)$, see [6] and compare [15].
Pór and Wood [16] considered embeddings of graphs into $\mathbb{Z}^3$, where the line segments, which represent the edges, do not cross any vertex distinct from its endpoints. Then, n points in $\mathbb{Z}^3$ yield a crossing-free drawing of K_n, if no three points are on a line. They proved in [16] that there are $\Theta(T^2)$ points in the $T \times T \times T$-grid with no three collinear points, by considering the set of all triples $(x, y, (x^2 + y^2) \bmod T)$, $x, y \in \{0, \ldots, T-1\}$, for T a prime with $T \equiv 3 \bmod 4$. This gives an upper bound of $O(n^{3/2})$ on the minimum volume of a bounding box of a drawing of K_n in $\mathbb{Z}^3$.

⋆ A preliminary version of this work appeared in Proceedings AAIM 2008.

For higher dimensions, Pór and Wood [16] raised the question of determining $\mathrm{vol}(n,d,k)$, which is defined as the minimum volume of an axis-aligned bounding box for embeddings of n points in $\mathbb{Z}^d$, such that no $(k+2)$ points are contained in a k-dimensional affine space. By partitioning of such a bounding box into $\mathrm{vol}(n,d,k)^{(d-k)/d}$ many affine k-dimensional subspaces they observed $\mathrm{vol}(n,d,k) \geq (n/(k+1))^{d/(d-k)}$. Known from [6] is $\mathrm{vol}(n,3,2) = \Theta(n^3)$, and in general, vol $(n,d,d-1) = \Theta(n^d)$.
For fixed integers $d,k,\ell \geq 1$ with $k < d$ let $f_d(\ell,k,T)$ be defined as the maximum number of points in the d-dimensional $T \times \cdots \times T$-grid, such that no ℓ of these points are contained in a k-dimensional affine subspace. A lower bound of $f_d(k+2,k,T) = \Omega(T^\beta)$ yields immediately the upper bound $\mathrm{vol}(n,d,k) = O(n^{d/\beta})$ on the minimum volume of a bounding box. In the following we focus on the investigation of the growth of the function $f_d(\ell,k,T)$ rather than on $\mathrm{vol}(n,d,\ell)$. In Section 2 we investigate (constructive) lower bounds on $f_d(\ell,T) := f_d(\ell,1,T)$, i.e., no ℓ points are collinear. For fixed integers $\ell \leq d$ we prove $f_d(\ell,T) = \Omega(\max\{T^{d-2}, T^{d(\ell-2)/(\ell-1)} \cdot poly(\log T)\})$. In Section 3 we give new upper bounds on $f_d(k+2,k,T)$ for integers $k \geq 1$, in particular $f_d(k+2,k,T) = O(T^{2d/(k+2)})$ for k even. We also consider distributions of grid-points, where no ℓ points are contained in a k-dimensional linear subspace and give a counterexample for a suggested order of the corresponding function $f_d^{lin}(k+1,k,T)$, see [4, 5].
Moreover, in connection with a question of Füredi [12] for fixed integers $\ell \geq 3$ we show for any finite set $S \subset \mathbb{R}^2$, $|S| = N$ and N sufficiently large, which does not contain ℓ collinear points, a lower bound on the largest size of a subset $S' \subseteq S$, where S' does not contain k collinear points, $3 \leq k < \ell$, i.e., $|S'| = \Omega(N^{\frac{k-2}{k-1}} \cdot poly(\log N))$. All of our arguments for proving lower bounds are of a probabilistic nature, however, they easily can be made constructive in polynomial time by using derandomization arguments.

2 No ℓ Collinear Points

For integers d,ℓ,T with $d \geq 2$ and $3 \leq \ell \leq T$ let $f_d(\ell,T)$ denote the largest size of a subset S of points in the d-dimensional $T \times \cdots \times T$-grid, such that no ℓ points of S are collinear. By monotonicity we have $f_d(\ell+1,T) \geq f_d(\ell,T)$. Well-known is the following upper bound on $f_d(\ell,T)$:

Proposition 1. *For integers d,ℓ,T with $d \geq 2$ and $3 \leq \ell \leq T$, it is*

$$f_d(\ell,T) \leq (\ell-1) \cdot T^{d-1}. \tag{1}$$

Proof. Let S be a subset of points in the d-dimensional $T \times \cdots \times T$-grid, such that no ℓ points of S are collinear. Partition the set of points in the $T \times \cdots \times T$-grid into T^{d-1} lines, where each line is of the form $(a_1,\ldots,a_i,x,a_{i+2},\ldots,a_d)$ for fixed $a_1,\ldots,a_i,a_{i+2},\ldots,a_d \in \{0,\ldots,T-1\}$. Each line contains at most $(\ell-1)$ points from S, hence $|S| \leq (\ell-1) \cdot T^{d-1}$. □

Next we give lower bounds on $f_d(\ell, T)$ for arbitrary integers $\ell \geq 3$.

Proposition 2. *For fixed integers $d \geq 2$, there exists a constant $c = c(d) > 0$ such that for all integers ℓ, T with $3 \leq \ell \leq T$ it is*

$$f_d(\ell, T) \geq \begin{cases} c \cdot \ell \cdot T^{d-1} & \text{if} \quad \ell \geq d+1 \\ \max\left\{T^{d-2}, c \cdot T^{d\frac{\ell-2}{\ell-1}}\right\} & \text{if} \quad \ell \leq d. \end{cases}$$

Notice, that we have $T^{d-2} > T^{d\frac{\ell-2}{\ell-1}}$ for $\ell < (d+2)/2$ and T sufficiently large, and for d even and for $\ell = (d+2)/2$ it is $T^{d-2} = T^{d\frac{\ell-2}{\ell-1}}$. The lower bound on $f_d(d+1, T)$ in Proposition 2, i.e., $\ell = d+1$, is by Brass and Knauer [4]. They observed that for fixed primes T and integers q the set of all integer points $(x_1, \ldots, x_d)$ with $x_1 + x_2^2 + \cdots + x_d^d \equiv q \bmod T$ in the d-dimensional $T \times \cdots \times T$-grid contains at most d collinear points, thus $f_d(d+1, T) = \Omega(T^{d-1})$. Hence, by Proposition 1 we have $f_d(d+1, T) = \Theta(T^{d-1})$.
As mentioned in the introduction, Pór and Wood [16] obtained $f_3(3, T) = \Omega(T^2)$, which is bigger than the lower bound in Theorem 2. However, Proposition 2 holds for all pairs (d, ℓ) for fixed d, and the lower bounds match up to constant factors the upper bounds (1) for every $\ell \geq d+1$.
Before proving Proposition 2, we introduce some useful notation.
For integers $a_1, \ldots, a_d$, which are not all equal to 0, let gcd $(a_1, \ldots, a_d) > 0$ denote the *greatest common divisor* of $a_1, \ldots, a_d$. Let $P = (p_1, \ldots, p_d)$ and $Q = (q_1, \ldots, q_d)$ be distinct points in the d-dimensional $T \times \cdots \times T$-grid. Let PQ denote the segment between the points P and Q, including P and Q. The segment PQ contains exactly (gcd $(p_1 - q_1, \ldots, p_d - q_d) + 1$) grid-points.
A *hypergraph* $\mathcal{G}$ is given by a pair $(V, \mathcal{E})$ with V its vertex-set and $\mathcal{E} \subseteq \mathcal{P}(V)$ its edge-set. A subset $I \subseteq V$ of the vertex-set V is called *independent*, if I does not contain any edges from $\mathcal{E}$, i.e., $E \not\subseteq I$ for each edge $E \in \mathcal{E}$. The largest size of an independent set in $\mathcal{G}$ is the *independence number* $\alpha(\mathcal{G})$. A 2-cycle in $\mathcal{G} = (V, \mathcal{E})$ is a pair $\{E, E'\}$ of distinct edges $E, E' \in \mathcal{E}$ with $|E \cap E'| \geq 2$. A 2-cycle $\{E, E'\}$ is called $(2, j)$-cycle if $|E \cap E'| = j$. A hypergraph $\mathcal{G}$ without any 2-cycles is called *linear.* A hypergraph $\mathcal{G} = (V, \mathcal{E})$ is called *ℓ-uniform,* if each edge $E \in \mathcal{E}$ contains exactly ℓ vertices.
In our arguments we use Túran's theorem for uniform hypergraphs, see [17]:

Theorem 1. *Let $\mathcal{G} = (V, \mathcal{E}_\ell)$ be an ℓ-uniform hypergraph on $|V| = N$ vertices with average-degree $t^{\ell-1} := \ell \cdot |\mathcal{E}_\ell|/N \geq 1$.*
Then, the independence number $\alpha(\mathcal{G})$ of $\mathcal{G}$ fulfills:

$$\alpha(\mathcal{G}) \geq \frac{\ell - 1}{\ell} \cdot \frac{N}{t}. \tag{2}$$

An independent set $I \subseteq V$ with $|I| \geq ((\ell-1)/\ell) \cdot (N/t)$ can be found in time $O(N + |\mathcal{E}_\ell|)$.

Next we prove Proposition 2:

Proof. Due to the results in [4] we only have to consider the case $\ell \neq d+1$. It is easy to see that $f_d(\ell, T) \geq T^{d-2}$. Namely, consider for integers r, $0 \leq r \leq (T-1)^2$, in the d-dimensional $T \times \cdots \times T$-grid the spheres S_r, which consist of all grid-points $P = (p_1, \ldots, p_d)$ with $\sum_{i=1}^{d} (p_i)^2 = r$. Clearly, $S_0 \cup \cdots \cup S_{(T-1)^2}$ covers the d-dimensional $T \times \cdots \times T$-grid, hence for some R we have $|S_R| \geq T^{d-2}$. Now any sphere S_R does not contain three collinear points. Indeed, for contradiction, assume that $P, P + \lambda \cdot V, P + \mu \cdot V \in S_R$, $\lambda \neq \mu$ and $\lambda, \mu \neq 0$, are collinear, where $P = (p_1, \ldots, p_d)$ and $V = (v_1, \ldots, v_d) \neq (0, \ldots, 0)$. We infer

$$\sum_{i=1}^{d} (p_i)^2 = \sum_{i=1}^{d} (p_i + \lambda \cdot v_i)^2 = \sum_{i=1}^{d} (p_i + \mu \cdot v_i)^2 = R,$$

and therefore,

$$2 \cdot \lambda \cdot \sum_{i=1}^{d} p_i \cdot v_i + \lambda^2 \cdot \sum_{i=1}^{d} (v_i)^2 = 2 \cdot \mu \cdot \sum_{i=1}^{d} p_i \cdot v_i + \mu^2 \cdot \sum_{i=1}^{d} (v_i)^2 = 0,$$

which implies $\sum_{i=1}^{d} (v_i)^2 = 0$, and this is not possible.
Next we prove the other lower bounds. Form an ℓ-uniform hypergraph $\mathcal{G} = (V, \mathcal{E}_\ell)$ with vertex-set V consisting of all T^d points in the d-dimensional $T \times \cdots \times T$-grid. For distinct grid-points $P_1, \ldots, P_\ell$ let $\{P_1, \ldots, P_\ell\} \in \mathcal{E}_\ell$ be an edge if and only if $P_1, \ldots, P_\ell$ are collinear. We want to find a large independent set $I \subseteq V$ in $\mathcal{G}$, as I yields a subset of grid-points, where no ℓ points are on a line.
We upper bound the size $|\mathcal{E}_\ell|$ of the edge-set. Let $P_1, \ldots, P_\ell$ be distinct, collinear points in the $T \times \cdots \times T$-grid, where $P_2, \ldots, P_{\ell-1}$ are contained in the segment $P_1 P_\ell$. There are T^d choices for the grid-point $P_1 = (p_{1,1}, \ldots, p_{1,d})$. Any d-tuple $(s_1, \ldots, s_d) \in \{-T+1, -T+2, \ldots, T-1\}^d$ fixes at most one point $P_\ell = (p_{1,1} + s_1, \ldots, p_{1,d} + s_d)$ in the $T \times \cdots \times T$-grid. By symmetry, which we take into account by a factor of 2^d, we may assume that $s_1, \ldots, s_d \geq 0$. Given the grid-points P_1 and P_ℓ with $P_\ell - P_1 = (s_1, \ldots, s_d) \neq (0, \ldots, 0)$, on the segment $P_1 P_\ell$ there are $\binom{\gcd\,(s_1, \ldots, s_d) - 1}{\ell - 2}$ choices for the $(\ell - 2)$ grid-points $P_2, \ldots, P_{\ell-1} \neq P_1, P_\ell$. By using $\binom{N}{k} \leq ((e \cdot N)/k)^k$ we obtain

$$\begin{aligned} |\mathcal{E}_\ell| &\leq 2^d \cdot T^d \cdot \sum_{s_1=0}^{T-1} \cdots \sum_{s_d=0}^{T-1} \binom{\gcd\,(s_1, \ldots, s_d) - 1}{\ell - 2} \\ &\leq 2^d \cdot T^d \cdot \sum_{s_1=0}^{T-1} \cdots \sum_{s_d=0}^{T-1} \left(\frac{e \cdot \gcd\,(s_1, \ldots, s_d)}{\ell - 2} \right)^{\ell-2}. \end{aligned} \tag{3}$$

For a given divisor $g \in \{1, \ldots, T-1\}$ there are at most $2 \cdot T/g$ integers $x \in \{0, \ldots, T-1\}$ which are divisible by g, hence (3) becomes

$$|\mathcal{E}_\ell| \leq (2 \cdot T)^d \cdot \sum_{g=1}^{T} \left(\frac{2 \cdot T}{g} \right)^d \cdot \left(\frac{e \cdot g}{\ell - 2} \right)^{\ell-2} \leq 4^d \cdot T^{2d} \cdot \left(\frac{9}{\ell} \right)^{\ell-2} \cdot \sum_{g=1}^{T} g^{\ell-d-2} \tag{4}$$

The sum $\sum_{g=1}^{T} g^{\ell-d-2}$ is $O(T^{\ell-d-1}/\ell)$ for $\ell \geq d+2$, and $O(\log T)$ for $\ell = d+1$, and $O(1)$ for $\ell \leq d$. Thus, by (4) for fixed $d \geq 2$ for a constant $c = c(d) > 0$ we infer

$$|\mathcal{E}_l| \leq \begin{cases} c \cdot 9^{\ell-1} \cdot \frac{T^{\ell+d-1}}{\ell^{\ell-1}} & \text{if} \quad \ell \geq d+2 \\ c \cdot T^{2d} & \text{if} \quad \ell \leq d. \end{cases}$$

Hence, the average-degree $t^{\ell-1} = \ell \cdot |\mathcal{E}_\ell|/T^d$ of $\mathcal{G}$ fulfills for a constant $c' = c'(d) > 0$:

$$t \leq \begin{cases} c' \cdot \frac{T}{\ell} & \text{if} \quad \ell \geq d+2 \\ c' \cdot T^{\frac{d}{\ell-1}} & \text{if} \quad \ell \leq d. \end{cases}$$

By Theorem 1 we find in time $O(T^d + |\mathcal{E}_\ell|)$ an independent set $I \subseteq V$ with

$$|I| \geq \begin{cases} \frac{1}{2 \cdot c'} \cdot \ell \cdot T^{d-1} & \text{if} \quad \ell \geq d+2 \\ \frac{1}{2 \cdot c'} \cdot T^{d\frac{\ell-2}{\ell-1}} & \text{if} \quad \ell \leq d. \end{cases}$$

The grid-points, which correspond to the vertices of the independent set I, satisfy that no ℓ points are collinear. $\square$

To improve the results from Theorem 2 for fixed d, ℓ with $(d+2)/2 \leq \ell \leq d$ by a logarithmic factor, we use the following result of Ajtai, Komlós, Pintz, Spencer and Szemerédi [1] in a version arising from work in [3] and [8].

Theorem 2. *Let $\ell \geq 3$ be a fixed integer. Let $\mathcal{G} = (V, \mathcal{E}_\ell)$ be an ℓ-uniform, linear hypergraph on $|V| = N$ vertices with average-degree $t^{\ell-1} = \ell \cdot |\mathcal{E}_\ell|/N$.*
Then, the independence number $\alpha(\mathcal{G})$ of $\mathcal{G}$ satisfies for a constant $C = C(\ell) > 0$:

$$\alpha(\mathcal{G}) \geq C \cdot \frac{N}{t} \cdot (\log t)^{\frac{1}{\ell-1}}. \tag{5}$$

An independent set $I \subseteq V$ with $|I| = \Omega((N/t) \cdot (\log t)^{1/(\ell-1)})$ can be found in polynomial time.

Theorem 3. *Let $d, \ell \geq 2$ be fixed integers with $(d+2)/2 \leq \ell \leq d$. Then, there exists a constant $c = c(d) > 0$ such that for all integers $T \geq 1$:*

$$f_d(\ell, T) \geq c \cdot T^{d\frac{\ell-2}{\ell-1}} \cdot (\log T)^{\frac{1}{\ell-1}}. \tag{6}$$

Proof. We form a non-uniform hypergraph $\mathcal{G} = (V, \mathcal{E}_\ell \cup \mathcal{E}_{\ell+1})$. The vertex-set consists of all T^d points from the d-dimensional $T \times \cdots \times T$-grid and for $m = \ell, \ell+1$ and distinct grid-points $P_1, \ldots, P_m$ it is $\{P_1, \ldots, P_m\} \in \mathcal{E}_m$ if and only if $P_1, \ldots, P_m$ are collinear. By the remarks following (4), for $\ell \leq d$ we have for constants $c_1, c_2 > 0$ that

$$|\mathcal{E}_\ell| \leq c_1 \cdot T^{2d} \quad \text{and} \quad |\mathcal{E}_{\ell+1}| \leq c_2 \cdot T^{2d} \cdot \log T. \tag{7}$$

Set $\varepsilon := d/\ell^2$ and select with probability $p := T^{\varepsilon}/T^{d/(\ell-1)}$ uniformly at random and independently of each other points from the d-dimensional $T \times \cdots \times T$-grid. Let V^* be the random set of chosen grid-points, let $\mathcal{E}_m^* := \mathcal{E}_m \cap [V^*]^m$, and let $E(|V^*|)$, $E(|\mathcal{E}_m^*|)$, $m = \ell, \ell+1$, be their expected sizes. We infer with (7):

$$\begin{aligned}
E(|V^*|) &= p \cdot T^d = T^{\varepsilon + d(\ell-2)/(\ell-1)} \\
E(|\mathcal{E}_\ell^*|) &= p^\ell \cdot |\mathcal{E}_\ell| \leq p^\ell \cdot c_1 \cdot T^{2d} \leq c_1 \cdot T^{\varepsilon\ell + d(\ell-2)/(\ell-1)} \\
E(|\mathcal{E}_{\ell+1}^*|) &= p^{\ell+1} \cdot |\mathcal{E}_{\ell+1}| \leq p^{\ell+1} \cdot c_2 \cdot T^{2d} \cdot \log T \leq c_2 \cdot T^{\varepsilon(\ell+1) + d(\ell-3)/(\ell-1)} \cdot \log T.
\end{aligned}$$

By Markov's and Chernoff's inequalities (this argument can be easily derandomized in time polynomial in T by using the method of conditional probabilities) there exists a subset $V^* \subseteq V$ of grid-points such that

$$|V^*| = (1 - o(1)) \cdot T^{\varepsilon + d(\ell-2)/(\ell-1)} \tag{8}$$

$$|\mathcal{E}_\ell^*| \leq 3 \cdot c_1 \cdot T^{\varepsilon\ell + d(\ell-2)/(\ell-1)} \tag{9}$$

$$|\mathcal{E}_{\ell+1}^*| \leq 3 \cdot c_2 \cdot T^{\varepsilon(\ell+1) + d(\ell-3)/(\ell-1)} \cdot \log T. \tag{10}$$

By (8) and (10) with $\varepsilon = d/\ell^2$ we have

$$|\mathcal{E}_{\ell+1}^*| = o(|V^*|). \tag{11}$$

Let $\mathcal{G}^* = (V^*, \mathcal{E}_\ell^* \cup \mathcal{E}_{\ell+1}^*)$ be the on the vertex-set V^* induced subhypergraph of $\mathcal{G}$. We delete one vertex from each edge $E \in \mathcal{E}_{\ell+1}^*$. For distinct edges $E, E' \in \mathcal{E}$ with $|E \cap E'| \geq 2$, all points in $E \cup E'$ are collinear, as two distinct points determine a line. Thus, we have destroyed all 2-cycles in $\mathcal{G}^*$. Let $V^{**} \subseteq V^*$ be the set of remaining vertices, where $|V^{**}| = (1 - o(1)) \cdot |V^*| \geq |V^*|/2$ by (11). The on the vertex-set V^{**} induced, uniform subhypergraph $\mathcal{G}^{**} = (V^{**}, \mathcal{E}_\ell^{**})$ of $\mathcal{G}$ with $\mathcal{E}_\ell^{**} := \mathcal{E}_\ell^{**} \cap [V^{**}]^\ell$ is linear, and with (8) and (9) its average-degree $t^{\ell-1}$ satisfies

$$t^{\ell-1} = \frac{\ell \cdot |\mathcal{E}_\ell^{**}|}{|V^{**}|} \leq 6 \cdot c_1 \cdot \ell \cdot T^{\varepsilon(\ell-1)} := t_0^{\ell-1}. \tag{12}$$

Since $\mathcal{G}^{**}$ is linear, we may apply Thereom 2 and we infer with (12) for the independence number $\alpha(\mathcal{G})$ for constants $C_\ell, C'_\ell > 0$:

$$\begin{aligned}
\alpha(\mathcal{G}) \geq \alpha(\mathcal{G}^{**}) &\geq C_\ell \cdot \frac{|V^{**}|}{t} \cdot (\log t)^{1/(\ell-1)} \geq C_\ell \cdot \frac{|V^{**}|}{t_0} \cdot (\log t_0)^{1/(\ell-1)} \geq \\
&\geq C_\ell \cdot \frac{(1/2) \cdot T^{\varepsilon + d(\ell-2)/(\ell-1)}}{(6 \cdot c_1 \cdot \ell)^{1/(\ell-1)} \cdot T^{\varepsilon}} \cdot \left(\log\left((6 \cdot c_1 \cdot \ell)^{1/(\ell-1)} \cdot T^{\varepsilon}\right)\right)^{1/(\ell-1)} \\
&\geq C'_\ell \cdot T^{d(\ell-2)/(\ell-1)} \cdot (\log T)^{1/(\ell-1)},
\end{aligned}$$

and by Theorem 2 such an independent set can be constructed in time polynomial in T. This shows $f_d(\ell, T) = \Omega(T^{d(\ell-2)/(\ell-1)} \cdot (\log T)^{1/(\ell-1)})$. □

Related here is a problem, which has been investigated by Füredi [12]. He considered finite sets $S \subset \mathbb{R}^2$ of points, which for fixed $\ell \geq 3$ do not contain ℓ

collinear points. He investigated the largest size $\alpha_k^\ell(S)$ of a subset of S, which does not contain any k points on a line, where $k < \ell$, and proved in [12] that $\alpha_k^\ell(S) = \Omega(|S|^{(k-2)/(k-1)})$. Moreover, for $\ell = 4$ and $k = 3$ Füredi showed the lower bound $\alpha_k^\ell(S) = \Omega(\sqrt{|S| \cdot \log |S|})$, while on the other hand by using the density-result of Hales-Jewett's theorem from [13] he obtained $\alpha_k^\ell(S) = o(|S|)$. As asked for in [12], the lower bound on $\alpha_k^\ell(S)$ given above can be improved by a polylogarithmic factor as the following considerations show.

Theorem 4. *Let $d, k, \ell \geq 2$ be fixed integers with $3 \leq k < \ell$. Let $S \subset \mathbb{R}^d$ be a finite set with $|S| = N$, where S does not contain ℓ collinear points.*
Then, one can find in time polynomial in N a subset $S' \subseteq S$, such that S' does not contain k collinear points, with

$$|S'| = \Omega\left(N^{\frac{k-2}{k-1}} \cdot (\log N)^{\frac{1}{k-1}}\right). \tag{13}$$

Proof. We construct a k-uniform hypergraph $\mathcal{G} = (S, \mathcal{E}_k)$ with vertex-set S. For any k distinct points $P_1, \ldots, P_k \in S$ let $\{P_1, \ldots, P_k\} \in \mathcal{E}_k$ if and only if $P_1, \ldots, P_k$ are collinear. We want to find a large independent set in $\mathcal{G}$. The set S with $|S| = N$ generates at most $\binom{N}{2}$ lines. Each line contains at most $(\ell - 1)$ points from S, hence on each line the number of k-element sets of collinear points is at most $\binom{\ell-1}{k}$, and we infer for a constant $c = c(\ell) > 0$:

$$|\mathcal{E}_k| \leq \binom{N}{2} \cdot \binom{\ell-1}{k} \leq c \cdot N^2. \tag{14}$$

Next we give upper bounds on the numbers $s_{2,j}(\mathcal{G})$ of $(2,j)$-cycles in $\mathcal{G}$, $j = 2, \ldots, k-1$. For a $(2,j)$-cycle $\{E, E'\}$ in $\mathcal{G}$ all points in $E \cup E'$ are collinear. Thus we have $s_{2,j}(\mathcal{G}) = 0$ for $j \leq 2 \cdot k - \ell$, as the set S does not contain ℓ collinear points. For $j > 2 \cdot k - \ell$ we obtain as in (14) for some constant $c' = c'(\ell) > 0$:

$$s_{2,j}(\mathcal{G}) \leq \binom{N}{2} \cdot \binom{\ell-1}{2 \cdot k - j} \leq c' \cdot N^2. \tag{15}$$

For $\varepsilon := 1/(2 \cdot k^2)$, we select uniformly at random and independently of each other points from S with probability $p := N^\varepsilon / N^{1/(k-1)}$. Let $S^* \subseteq S$ be the random set of chosen points, and let $\mathcal{G}^* = (S^*, \mathcal{E}_k^*)$ with $\mathcal{E}_k^* := \mathcal{E}_k \cap [S^*]^k$ be the on the vertex-set S^* induced subhypergraph of $\mathcal{G}$. The expected numbers satisfy $E[|S^*|] = p \cdot |S| = p \cdot N$, and $E[|\mathcal{E}_k^*|] = p^k \cdot |\mathcal{E}_k|$, and $E[s_{2,j}(\mathcal{G}^*)] = p^{2k-j} \cdot s_{2,j}(\mathcal{G})$, $j = 2, \ldots, k-1$. By Markov's and Chernoff's inequality with (14) and (15) there exists an induced subhypergraph $\mathcal{G}^* = (S^*, \mathcal{E}_k^*)$ of $\mathcal{G}$ such that

$$|S^*| \geq p \cdot N/2 = N^{\frac{k-2}{k-1} + \varepsilon}/2 \tag{16}$$

$$|\mathcal{E}_k^*| \leq 3 \cdot p^k \cdot |\mathcal{E}_k| \leq 3 \cdot c \cdot N^{\frac{k-2}{k-1} + \varepsilon k} \tag{17}$$

$$s_{2,j}(\mathcal{G}^*) \leq 3 \cdot p^{2k-j} \cdot s_{2,j}(\mathcal{G}) \leq 3 \cdot c' \cdot N^{\frac{j-2}{k-1} + \varepsilon(2k-j)}. \tag{18}$$

For $j = 2, \ldots, k-1$, and $0 < \varepsilon \leq 1/(2 \cdot k^2)$, by (16) and (18) we have

$$s_{2,j}(\mathcal{G}^*) = o(|S^*|). \tag{19}$$

Discard one vertex from each $(2, j)$-cycle in $\mathcal{G}^*$, $j = 2, \ldots, k-1$. The set $S^{**} \subseteq S^*$ of all remaining vertices satisfies by (19) that $|S^{**}| = (1 - o(1)) \cdot |S^*| \geq |S^*|/2$. The on the vertex-set S^{**} induced subhypergraph $\mathcal{G}^{**} = (S^{**}, \mathcal{E}_k^{**})$ of $\mathcal{G}^*$ with $\mathcal{E}_k^{**} := \mathcal{E}_k^* \cap [S^{**}]^k$ is linear. With $|\mathcal{E}_k^{**}| \leq |\mathcal{E}_k^*|$ and (17) we obtain for the average-degree $(t^{**})^{k-1}$ of $\mathcal{G}^{**}$:

$$(t^{**})^{k-1} := \frac{k \cdot |\mathcal{E}_k^{**}|}{|S^{**}|} \leq 12 \cdot k \cdot c \cdot N^{\varepsilon(k-1)} =: (t_0^{**})^{k-1}. \tag{20}$$

By Theorem 2 with (20) one can find in time polynomial in N an independent set $I \subseteq S^{**}$, such that, as $\varepsilon > 0$ is fixed, for constants $C_k, C'_k > 0$ we have

$$\begin{aligned} |I| &\geq C_k \cdot \frac{|S^{**}|}{t^{**}} \cdot (\log t^{**})^{\frac{1}{k-1}} \geq C_k \cdot \frac{|S^{**}|}{t_0^{**}} \cdot (\log t_0^{**})^{\frac{1}{k-1}} \geq \\ &\geq C_k \cdot \frac{(1/2) \cdot N^{\frac{k-2}{k-1}+\varepsilon}}{(12 \cdot k \cdot c)^{\frac{1}{k-1}} \cdot N^{\varepsilon}} \cdot \left(\log\left((6 \cdot k \cdot c)^{\frac{1}{k-1}} \cdot N^{\varepsilon}\right)\right)^{\frac{1}{k-1}} \\ &\geq C'_k \cdot N^{\frac{k-2}{k-1}} \cdot (\log N)^{\frac{1}{k-1}}. \end{aligned}$$

The set I does not contain k distinct collinear points. □

3 No $(k+2)$ Points in Affine k-Space or Linear $(k+1)$-Space

Here we consider higher dimensional versions of Theorem 2. For fixed positive integers k, ℓ with $\ell \geq k+2$, let $f_d(\ell, k, T)$ denote the maximum number of points in the d-dimensional $T \times \cdots \times T$-grid, such that no ℓ points are contained in a k-dimensional affine subspace of $\mathbb{R}^d$. We have by monotonicity $f_d(\ell+1, k, T) \geq f_d(\ell, k, T)$.

The d-dimensional $T \times \cdots \times T$-grid can be partitioned into T^{d-k} many k-dimensional affine spaces, namely for fixed $a_1, \ldots, a_{d-k} \in \{0, \ldots, T-1\}$, into the k-dimensional affine spaces given by all points $(a_1, \ldots, a_{d-k}, x_{d-k+1}, \ldots, x_d)$, hence it follows

$$f_d(\ell, k, T) \leq (\ell - 1) \cdot T^{d-k}. \tag{21}$$

For $k = d-1$ and fixed $\ell \geq d+1$ the upper bound (21) is asymptotically sharp, namely for primes T the set of points $(x \bmod T, x^2 \bmod T, \ldots, x^d \bmod T)$, $x = 0, \ldots, T-1$, on the modular moment-curve meets every $(d-1)$-dimensional affine space in at most d points, compare [5, 18], thus

$$f_d(\ell, d-1, T) = \Theta(T). \tag{22}$$

We can improve on the upper bound (21) for pairs $(\ell = k+2, k)$ as follows.

Lemma 1. *Let $d, k \geq 1$ with $k \leq d-1$ be fixed integers. Then, for some constant $c = c(k) > 0$ it is:*

$$f_d(k+2, k, T) \leq c \cdot T^{\frac{d}{\lceil (k+1)/2 \rceil}} \tag{23}$$

For even $k \geq 2$, the upper bound (23) on $f_d(k+2,k,T)$ is smaller than (21) for $k < d-2$, and for $k = d-2$ in both bounds the exponents of T are equal, and (21) is less than (23) only for $k = d-1$. For odd $k \geq 1$, the upper bound (23) is smaller than (21) for the range $(d-1)/2 - \sqrt{(d-1)^2/4 - d} \leq k \leq (d-1)/2 + \sqrt{(d-1)^2/4 - d}$.

Proof. Let $k \geq 2$ be even and set $g := k/2$. Let S be a subset of points from the d-dimensional $T \times \cdots \times T$-grid, where no $(k+2)$ points in S are contained in a k-dimensional affine subspace, w.l.o.g. $|S| \geq k+2$. Consider the set S_{g+1} of all $(g+1)$-term sums of pairwise distinct elements from S with addition componentwise:

$$S_{g+1} := \{s_1 + \cdots + s_{g+1} \mid s_1, \ldots, s_{g+1} \in S \text{ are pairwise distinct}\}.$$

We claim that for distinct points $s_1, \ldots, s_{g+1} \in S$ and distinct $t_1, \ldots, t_{g+1} \in S$ with $\{s_1, \ldots, s_{g+1}\} \neq \{t_1, \ldots, t_{g+1}\}$ it is

$$s_1 + \cdots + s_{g+1} \neq t_1 + \cdots + t_{g+1}. \tag{24}$$

Otherwise, we have $s_1 + \cdots + s_{g+1} = t_1 + \cdots + t_{g+1}$ for some distinct points $s_1, \ldots, s_{g+1} \in S$ and distinct $t_1, \ldots, t_{g+1} \in S$. Assume that for some integer $j \geq 1$ it is $s_i = t_i$, $i = 0, \ldots, j-1$, and that $s_j, \ldots, s_{g+1}, t_j, \ldots, t_{g+1}$ are pairwise distinct points. Then, it is $s_j + \cdots + s_{g+1} = t_j + \cdots + t_{g+1}$, hence we have found $2 \cdot (g+2-j) = k+4-2 \cdot j$ distinct points in S, which are contained in a $(k+2-2 \cdot j)$-dimensional affine space. Adding $2 \cdot j$ further distinct grid-points from S to $s_j, \ldots, s_{g+1}, t_j, \ldots, t_{g+1}$ yields $(k+2)$ grid-points in the set S, which are contained in a k-dimensional affine space, a contradiction.
By (24) we infer $|S_{g+1}| = \binom{|S|}{g+1}$, and all points in S_{g+1} are contained in a $((g+1) \cdot T) \times \cdots \times ((g+1) \cdot T)$-grid, thus we obtain

$$\binom{|S|}{g+1} = |S_{g+1}| \leq ((g+1) \cdot T)^d,$$

and with $k = 2 \cdot g$ for a constant $c = c(k) > 0$ we have $|S_{g+1}| \leq c \cdot T^{2d/(k+2)}$, hence $f_d(k+2,k,T) \leq c \cdot T^{2d/(k+2)}$.
Let $k \geq 1$ be odd. If a subset S of points from the d-dimensional $T \times \cdots \times T$-grid does not contain $(k+2)$ points, which are contained in a k-dimensional affine subspace, then S also does not contain $(k+1)$ points, which are contained in a $(k-1)$-dimensional affine subspace. With the already proved upper bound for even values we infer for $k \geq 1$ odd that $f_d(k+2,k,T) \leq f_d(k+1,k-1,T) \leq c \cdot T^{2d/(k+1)}$. □

Concerning lower bounds, Brass and Knauer proved in [4] for fixed integers $d, k, \ell \geq 2$ by a random selection of points from the $T \times \cdots \times T$-grid that

$$f_d(\ell,k,T) = \Omega(T^{d-k-(d(k+1)/(\ell-1))}). \tag{25}$$

Then (25) guarantees $f_d(\ell,k,T) = \Omega(T)$ for $\ell - 1 \geq d(k+1)/(d-k-1)$ and $k \leq d-2$. One can improve (25) a little by using a (slightly) different argument:

Lemma 2. *For fixed integers $d, k, \ell \geq 1$ with $k \leq d-1$ and $\ell \geq k+2$ and integers $T \geq 1$ it is:*

$$f_d(\ell, k, T) = \Omega(T^{d-k-(k(d+1)/(\ell-1))}). \tag{26}$$

Notice that (26) is bigger than (25) for $k < d$. However, (26) as well as (25) are close to the lower bound (21) for ℓ large.

Proof. Form a ℓ-uniform hypergraph $\mathcal{G} = (V, \mathcal{E}_\ell)$ with vertex-set V consisting of all T^d points from the d-dimensional $T \times \cdots \times T$-grid. For grid-points $P_1, \ldots, P_\ell \in V$ let $\{P_1, \ldots, P_\ell\} \in \mathcal{E}_\ell$ if and only if $P_1, \ldots, P_\ell$ are contained in a k-dimensional affine subspace. We want to guarantee a large independent set in $\mathcal{G}$. Each k-dimensional affine subspace contains at most T^k points from the d-dimensional $T \times \cdots \times T$-grid. The number of k-dimensional affine subspaces, which intersect the d-dimensional $T \times \cdots \times T$-grid in at least $(k+1)$ points, is at most $\binom{T^d}{k+1}$. We infer for a constant $c > 0$

$$|\mathcal{E}_\ell| \leq \binom{T^k}{\ell} \cdot \binom{T^d}{k+1} \leq c \cdot T^{k\ell + d(k+1)},$$

hence the average-degree $t^{\ell-1}$ of $\mathcal{G}$ fulfills for some constant $c' > 0$:

$$t^{\ell-1} = \frac{\ell \cdot |\mathcal{E}_\ell|}{|V|} \leq \frac{\ell \cdot c \cdot T^{k\ell + d(k+1)}}{T^d} \leq c' \cdot T^{k(d+\ell)}. \tag{27}$$

By Theorem 1 and (27) we can find in time polynomial in T an independent set $I \subseteq V$ in $\mathcal{G}$, such that for a constant $c'' > 0$ it is

$$|I| \geq \frac{\ell - 1}{\ell} \cdot \frac{T^d}{c'^{1/(\ell-1)} \cdot T^{k(d+\ell)/(\ell-1)}} \geq c'' \cdot T^{d-k-(k(d+1)/(\ell-1))}.$$

□

Next we consider *linear* subspaces. Let $f_d^{lin}(\ell, k, T)$ denote the maximum number of points in the d-dimensional $T \times \cdots \times T$-grid, such that no ℓ points are contained in a k-dimensional *linear* subspace. From number theory it is known [5] that for fixed $d \geq 2$ it is $f_d^{lin}(2, 1, T) = \Theta(T^d)$. Bárány, Harcos, Pach and Tardos proved in [2] that $f_d^{lin}(d, d-1, T) = \Theta(T^{d/(d-1)})$ for fixed $d \geq 2$. Based on this, Brass and Krauer [4] conjectured (stated as a problem in [5]) that

$$f_d^{lin}(k+1, k, T) = \Theta(T^{\frac{(d-k)d}{d-1}}). \quad (?) \tag{28}$$

However, we can show the following:

Lemma 3. *For fixed integers d, k with $1 \leq k \leq d-1$ there exists a constant $c > 0$, such that for every integer $T \geq 1$ it is*

$$f_d^{lin}(k+1, k, T) \leq c \cdot T^{\frac{d}{\lceil k/2 \rceil}}. \tag{29}$$

For odd k the upper bound (29) is asymptotically smaller than the suggested growth of $f_d^{lin}(k+1, k, T)$ in (28) for $1 < k < d-2$ with equality for $k = d-2$. Similarly, for even k the upper bound (29) is smaller than in (28) for the range $d/2 - \sqrt{d^2/4 - 2d + 2} < k < d/2 + \sqrt{d^2/4 - 2d + 2}$. Hence, (28) does not hold for several values of k, d.

Proof. The proof is similar to that of Lemma 1, therefore we only sketch it. Let $k \geq 1$ be an odd integer and set $g := (k+1)/2$. Let S be a subset of points from the d-dimensional $T \times \cdots \times T$-grid, where no $(k+1)$ distinct points from S are contained in a k-dimensional linear subspace, w.l.o.g. $|S| \geq k+1$. Let

$$S_g := \{s_1 + \cdots + s_g \mid s_1, \ldots, s_g \in S \text{ are pairwise distinct}\}.$$

As in the proof of Lemma 1, for distinct grid-points $s_1, \ldots, s_g \in S$ and distinct $t_1, \ldots, t_g \in S$ with $\{s_1, \ldots, s_g\} \neq \{t_1, \ldots, t_g\}$ it is $s_1 + \cdots + s_g \neq t_1 + \cdots + t_g$, as otherwise we can find $(k+1)$ distinct grid-points in S, which are contained in a k-dimensional linear subspace, a contradiction, hence $|S_g| = \binom{|S|}{g} \leq (g \cdot T)^d$, and we infer $f_d^{lin}(k+1, k, T) = O(T^{2d/(k+1)})$ for odd $k \geq 1$.
For even $k \geq 2$ we conclude as in the proof of Lemma 1 that for a constant $c > 0$ it is $f_d^{lin}(k+1, k, T) \leq f_d^{lin}(k, k-1, T) \leq c \cdot T^{2d/k}$. □

References

1. M. Ajtai, J. Komlós, J. Pintz, J. Spencer and E. Szemerédi, *Extremal Uncrowded Hypergraphs*, J. Comb. Theory Ser. A, 32, 1982, 321–335.
2. I. Bárány, G. Harcos, J. Pach and G. Tardos, Covering Lattice Points by Subspaces, Per. Math. Hung. 43, 2001, 93–103.
3. C. Bertram–Kretzberg and H. Lefmann, *The Algorithmic Aspects of Uncrowded Hypergraphs,* SIAM J. on Comput. 29, 1999, 201–230.
4. P. Brass and C. Knauer, *On Counting Point-Hyperplane Incidences,* Comp. Geo. 25, 2003, 13–20.
5. P. Brass, W. Moser and J. Pach, *Research Problems in Discrete Geometry,* Springer, 2005, 417–421.
6. R.–F. Cohen, P. Eades, T. Lin and F. Ruskey, *Three-Dimensional Graph Drawing,* Algorithmica 17, 1996, 199–208.
7. H. E. Dudeney, *Amusements in Mathematics,* Nelson, London, 1917, p. 94 and 222.
8. R. A. Duke, H. Lefmann and V. Rödl, *On Uncrowded Hypergraphs,* Rand. Struct. & Algorithms 6, 1995, 209–212.
9. P. Erdős, in K.–F. Roth, *On a Problem of Heilbronn,* J. London Math. Soc. 26, 1951, 198–204.
10. A. Flammenkamp, *Progress in the No-three-in-Line-Problem,* J. Comb. Theory, Ser. A 60, 1992, 305–311.
11. A. Flammenkamp, *Progress in the No-three-in-Line-Problem. II,* J. Comb. Theory, Ser. A 81, 1998, 108–113.
12. Z. Füredi, *Maximal Independent Subsets in Steiner Systems and in Planar Sets,* SIAM J. Disc. Math. 4, 1991, 196–199.

13. H. Fürstenberg and Y. Katznelson, *A Density Version of the Hales–Jewett Theorem,* Discrete Mathematics 25, 1989, 227-241.
14. R. Guy, *Unsolved Problems in Number Theory,* 2nd ed., Springer, 1994.
15. J. Pach, T. Thiele and G. Tóth, *Three Dimensional Grid Drawings of Graphs,* Advances in Discrete and Computational Geometry, eds. B. Chazelle, J. E. Goodman and R. Pollack, Contemporary Math. 223, AMS, 1999, 251–255.
16. A. Pór and D. W. Wood, *No-Three-in-Line-in-3D,* Algorithmica 47, 2007, 481–488.
17. J. Spencer, *Turán's Theorem for k-Graphs,* Disc. Math. 2, 1972, 183–186.
18. T. Thiele, *Geometric Selection Problems and Hypergraphs,* PhD thesis, FU Berlin, 1995.