

Convex Hulls of Point-Sets and Non-Uniform Hypergraphs

Hanno Lefmann

Fakultät für Informatik, TU Chemnitz, D-09107 Chemnitz, Germany
 lefmann@informatik.tu-chemnitz.de

Abstract. For fixed integers $k \geq 3$ and hypergraphs \mathcal{G} on N vertices, which contain edges of cardinalities at most k , and are uncrowded, i.e., do not contain cycles of lengths 2, 3, or 4, and with average degree for the i -element edges bounded by $O(T^{i-1} \cdot (\ln T)^{(k-i)/(k-1)})$, $i = 3, \dots, k$, for some number $T \geq 1$, we show that the independence number $\alpha(\mathcal{G})$ satisfies $\alpha(\mathcal{G}) = \Omega((N/T) \cdot (\ln T)^{1/(k-1)})$. Moreover, an independent set I of size $|I| = \Omega((N/T) \cdot (\ln T)^{1/(k-1)})$ can be found deterministically in polynomial time. This extends a result of Ajtai, Komlós, Pintz, Spencer and Szemerédi for uncrowded uniform hypergraphs. We apply this result to a variant of Heilbronn's problem on the minimum area of the convex hull of small sets of points among n points in the unit square $[0, 1]^2$.

1 Introduction

An independent set I in a graph or hypergraph $\mathcal{G} = (V, \mathcal{E})$ with vertex-set V and edge-set \mathcal{E} is a subset of the vertex-set V , which does not contain any edges, 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 $\alpha(\mathcal{G})$. For graphs $G = (V, E)$ with average degree $t := 2 \cdot |E|/|V| \geq 1$ Turán's theorem gives $\alpha(G) \geq |V|/(2 \cdot t)$. Turán's theorem for hypergraphs says, see [20]: If $\mathcal{G} = (V, \mathcal{E}_k)$ is a k -uniform hypergraph, i.e., all edges have cardinality k , with average degree $t^{k-1} := k \cdot |\mathcal{E}_k|/|V| \geq 1$, then $\alpha(\mathcal{G}) \geq ((k-1)/k) \cdot (|V|/t)$. An independent set $I \subseteq V$ in \mathcal{G} achieving this lower bound can be found deterministically in time $O(|V| + |\mathcal{E}_k|)$. For uncrowded k -uniform hypergraphs $\mathcal{G} = (V, \mathcal{E}_k)$, i.e., \mathcal{G} contains no cycles of length 2, 3, or 4, Ajtai, Komlós, Pintz, Spencer and Szemerédi [1] improved this lower bound by a factor of $\Theta((\log t)^{1/(k-1)})$. Several applications of this result have been found, see [5]. Here we extend this result from [1] to non-uniform uncrowded hypergraphs:

Theorem 1. *Let $k \geq 3$ be a fixed integer. Let $\mathcal{G} = (V, \mathcal{E}_3 \cup \dots \cup \mathcal{E}_k)$ be an uncrowded hypergraph on $|V| = N$ vertices, where \mathcal{E}_i is the set of all i -element edges in \mathcal{G} , such that the average degrees $t_i^{i-1} := i \cdot |\mathcal{E}_i|/|V|$ for the i -element edges satisfy $t_i^{i-1} \leq c_i \cdot T^{i-1} \cdot (\ln T)^{(k-i)/(k-1)}$ for some number $T \geq 1$ with constants c_i , where $0 < c_i < 1/8 \cdot \binom{k-1}{i-1} / (10^{3(k-i)/(k-1)} \cdot k^2)$, $i = 3, \dots, k$. Then, for some constant $C_k > 0$ the independence number $\alpha(\mathcal{G})$ satisfies*

$$\alpha(\mathcal{G}) \geq C_k \cdot (N/T) \cdot (\ln T)^{1/(k-1)}. \quad (1)$$

An independent set $I \subseteq V$ with $|I| = \Omega((N/T) \cdot (\ln T)^{1/(k-1)})$ can be found deterministically in time $o(N \cdot T^{4k-4})$.

The corresponding result also holds for linear hypergraphs \mathcal{G} , which have the property that they do not contain cycles of length 2, i.e., each two distinct edges have at most one vertex in common, provided that \mathcal{G} does not contain any 2-element edges. Theorem 1 is best possible up to a constant factor for a certain range $k < T < N$, as can be seen by considering random non-uniform hypergraphs $\mathcal{G} = (V, \mathcal{E}_3 \cup \dots \cup \mathcal{E}_k)$ on $|V| = N$ vertices.

As an application we consider a variant of Heilbronn's problem for the convex hull of sets of points in the unit square $[0, 1]^2$. The original problem of Heilbronn asks for a distribution of n points in $[0, 1]^2$ such that the minimum area of a triangle determined by three of these n points achieves its largest value. For this problem, the points $1/n \cdot (i \bmod n, i^2 \bmod n)$, $i = 0, \dots, n-1$, where n is a prime, give the lower $\Omega(1/n^2)$ on the minimum area of a triangle. This lower bound has been improved in [12] by a factor $\Omega(\log n)$, see [6] for a deterministic polynomial time algorithm. Upper bounds on the minimum area of a triangle among n points in $[0, 1]^2$ were given by Roth [15–18] and Schmidt [19] and, the currently best upper bound $O(2^{c\sqrt{\log n}}/n^{8/7})$, $c > 0$ a constant, is due to Komlós, Pintz and Szemerédi [11].

Variants of Heilbronn's triangle problem in higher dimensions were investigated in [2–4, 7, 8, 13]. A generalization of Heilbronn's triangle problem to k points, see Schmidt [19], asks, given an integer $k \geq 3$, for the supremum $\Delta_k(n)$ over all distributions of n points in $[0, 1]^2$ of the minimum area of the convex hull determined by some k of n points. In [6] it has been shown that $\Delta_k(n) = \Omega(1/n^{(k-1)/(k-2)})$ for fixed $k \geq 3$, and any integers $n \geq k$; for $k = 4$ this was proved in [19]. This has been improved in [14] to $\Delta_k(n) = \Omega((\log n)^{1/(k-1)} / n^{(k-1)/(k-2)})$ for fixed $k \geq 3$. Currently, for $k \geq 4$ only the upper bound $\Delta_k(n) = O(1/n)$ is known.

Here we show for fixed integers $k \geq 3$, that one can achieve these lower bounds simultaneously for $j = 3, \dots, k$ by a single configuration of n points in $[0, 1]^2$.

Theorem 2. *Let $k \geq 3$ be a fixed integer. For integers $n \geq k$ there exists a configuration of n points in $[0, 1]^2$, such that, simultaneously for $j = 3, \dots, k$, the area of the convex hull of any j of the n points is $\Omega((\log n)^{1/(j-1)} / n^{(j-1)/(j-2)})$.*

By considering the standard $L \times L$ -grid for a suitable integer $L \geq n$ one can also give a polynomial time algorithm which achieves the lower bounds from Theorem 2 on the areas of the convex hulls. (Details are omitted.)

2 Uncrowded and Linear Hypergraphs

Definition 1. *A hypergraph is a pair $\mathcal{G} = (V, \mathcal{E})$ with vertex-set V and edge-set \mathcal{E} , where $E \subseteq V$ for each edge $E \in \mathcal{E}$. For a hypergraph \mathcal{G} the notation $\mathcal{G} = (V, \mathcal{E}_2 \cup \dots \cup \mathcal{E}_k)$ indicates that \mathcal{E}_i is the set of all i -element edges in \mathcal{G} , $i = 2, \dots, k$. For a vertex $v \in V$ let $d_i(v)$ denote the number of i -element edges $E \in \mathcal{E}_i$ which contain v , i.e., $d_i(v)$ is the degree of v for the i -element edges in \mathcal{G} . The independence number $\alpha(\mathcal{G})$ of $\mathcal{G} = (V, \mathcal{E})$ is the largest size of a subset $I \subseteq V$ which contains no edges from \mathcal{E} . A j -cycle in a hypergraph $\mathcal{G} = (V, \mathcal{E})$ is a sequence E_1, \dots, E_j of distinct edges from \mathcal{E} , such that $E_i \cap E_{i+1} \neq \emptyset$,*

$i = 1, \dots, j-1$, and $E_j \cap E_1 \neq \emptyset$, and a sequence v_1, \dots, v_j of distinct vertices with $v_{i+1} \in E_i \cap E_{i+1}$, $i = 1, \dots, j-1$, and $v_1 \in E_1 \cap E_j$. An unordered pair $\{E, E'\}$ of distinct edges $E, E' \in \mathcal{E}$ with $|E \cap E'| \geq 2$ is a 2-cycle. A 2-cycle $\{E, E'\}$ in $\mathcal{G} = (V, \mathcal{E}_2 \cup \dots \cup \mathcal{E}_k)$ with $E \in \mathcal{E}_i$ and $E' \in \mathcal{E}_j$ is called $(2; (g, i, j))$ -cycle if and only if $|E \cap E'| = g$, $2 \leq g \leq i \leq j$ but $g < j$. The hypergraph \mathcal{G} is called linear if it does not contain any 2-cycles, and \mathcal{G} is called uncrowded if it does not contain any 2-, 3-, or 4-cycles.

For uncrowded k -uniform hypergraphs with average degree t^{k-1} the Turán bound on the independence number has been improved in [1] by a factor $\Theta((\log t)^{1/(k-1)})$, see [5] and [10] for a deterministic polynomial time algorithm.

Theorem 3. *Let $k \geq 3$ be a fixed integer. Let $\mathcal{G} = (V, \mathcal{E}_k)$ be an uncrowded k -uniform hypergraph on $|V| = N$ vertices and with average degree $t^{k-1} := k \cdot |\mathcal{E}_k|/N$. Then, for some constant $C_k > 0$, the independence number $\alpha(\mathcal{G})$ satisfies $\alpha(\mathcal{G}) \geq C_k \cdot (N/t) \cdot (\log t)^{1/(k-1)}$.*

To prove Theorem 3, in [1] the following central lemma has been used to construct iteratively a large independent set in a hypergraph, which we use in our arguments too; see [10] for a deterministic polynomial time algorithm.

Lemma 1. *Let T and N be large positive integers. Let s be an integer with $0 \leq s \leq (\ln T)/10^2$. Let $w_s := (s+1)^{1/(k-1)} - s^{1/(k-1)}$ and $\varepsilon := 10^{-6}/\ln T$. Let $N/(2 \cdot e^s) \leq n \leq N/e^s$ and $T/(2 \cdot e^s) \leq t \leq T/e^s$.*

Let $\mathcal{G} = (V, \mathcal{E}_2 \cup \dots \cup \mathcal{E}_k)$ be an uncrowded hypergraph with $|V| = n$ vertices, where for each vertex $v \in V$ the degrees $d_i(v)$ for the i -element edges satisfy $d_i(v) \leq \binom{k-1}{i-1} \cdot s^{(k-i)/(k-1)} \cdot t^{i-1}$, $i = 2, \dots, k$.

Then, one can find in time $O(n \cdot t^{4(k-1)})$ an independent set $I \subseteq V$ in \mathcal{G} , a subset $V^ \subset V$ with $V^* \cap I = \emptyset$, and a hypergraph $\mathcal{G}^* = (V^*, \mathcal{E}_2^* \cup \dots \cup \mathcal{E}_k^*)$ such that*

- (i) $\alpha(\mathcal{G}) \geq |I| + \alpha(\mathcal{G}^*)$ and (ii) $|I| \geq 0.99 \cdot \frac{n \cdot w_s}{e \cdot t}$ and (iii) $|V^*| \geq \frac{n \cdot (1-\varepsilon)}{e}$
- (iv) $d_i^*(v) \leq \binom{k-1}{i-1} \cdot (s+1)^{(k-i)/(k-1)} \cdot (t \cdot (1+\varepsilon)/e)^{i-1}$ for each vertex $v \in V^*$, where $d_i^*(v)$ denotes the degree of v for the i -element edges in \mathcal{G}^* , $i = 2, \dots, k$.

Lemma 2. *Let $k \geq 3$ be a fixed integer. Let $\mathcal{G} = (V, \mathcal{E}_2 \cup \dots \cup \mathcal{E}_k)$ be a hypergraph with $|V| = N$ and $N \geq 65 \cdot (\ln k)^{1000/998}$, where the average degrees $t_i^{i-1} := i \cdot |\mathcal{E}_i|/N$ for the i -element edges in \mathcal{E}_i fulfill $t_i^{i-1} \leq c_i \cdot T^{i-1} \cdot (\ln T)^{(k-i)/(k-1)}$ for some number $T \geq 1$ and for some constants $c_i > 0$ with $c_i < 1/8 \cdot \binom{k-1}{i-1} / (10^{(3(k-i))/(k-1)} \cdot k^2)$, $i = 2, \dots, k$.*

Then, for $s := 10^{-3} \cdot \ln T$, one can find in time $O(|V| + \sum_{i=2}^k |\mathcal{E}_i|)$ an induced subhypergraph $\mathcal{G}^ = (V^*, \mathcal{E}_2^* \cup \dots \cup \mathcal{E}_k^*)$ on $|V^*| = n$ vertices with $\mathcal{E}_i^* := \mathcal{E}_i \cap [V^*]^i$, $i = 2, \dots, k$, such that $(3/4) \cdot N/e^s \leq n \leq N/e^s$ and for each vertex $v \in V^*$ the degrees $d_i^*(v)$ for the i -element edges in \mathcal{G}^* satisfy*

$$d_i^*(v) \leq \binom{k-1}{i-1} \cdot s^{\frac{k-i}{k-1}} \cdot (T/e^s)^{i-1}. \quad (2)$$

Proof. We pick vertices with probability $p := 1/e^s$ uniformly at random and independently of each other from the vertex-set V in \mathcal{G} . Let V^* be the random set of chosen vertices of expected size $E[|V^*|] = p \cdot N$. With $s = 10^{-3} \cdot \ln T$ and $T = O(N)$, we have by Chernoff's inequality for $N \geq 65 \cdot (\ln k)^{1000/998}$:

$$\text{Prob} (E[|V^*|] - |V^*| > N/(8 \cdot e^s)) \leq e^{-\frac{N^2/(64 \cdot e^{2s})}{N}} = e^{-N/(64 \cdot e^{2s})} < 1/k. \quad (3)$$

Let $\mathcal{E}_i^* := \mathcal{E}_i \cap [V^*]^i$, $i = 2, \dots, k$, and let $\mathcal{G}^* = (V^*, \mathcal{E}_2^* \cup \dots \cup \mathcal{E}_k^*)$ be the on V^* induced random subhypergraph of \mathcal{G} . For $i = 2, \dots, k$, we have for the expected numbers $E[|\mathcal{E}_i^*|] = p^i \cdot |\mathcal{E}_i| = p^i \cdot N \cdot t_i^{i-1}/i \leq p^i \cdot c_i \cdot T^{i-1} \cdot (\ln T)^{(k-i)/(k-1)} \cdot N/i$. By Markov's inequality it is $\text{Prob} (|\mathcal{E}_i^*| > k \cdot E[|\mathcal{E}_i^*|]) \leq 1/k$, hence with (3) there exists a subhypergraph $\mathcal{G}^* = (V^*, \mathcal{E}_2^* \cup \dots \cup \mathcal{E}_k^*)$ of \mathcal{G} such that for $i = 2, \dots, k$:

$$|V^*| \geq (7/8) \cdot N/e^s \quad \text{and} \quad |\mathcal{E}_i^*| \leq k \cdot p^i \cdot c_i \cdot T^{i-1} \cdot (\ln T)^{(k-i)/(k-1)} \cdot N/i. \quad (4)$$

Let n_i be the number of vertices $v \in V^*$ with degree $d_i^*(v) \geq 8 \cdot e^s \cdot k^2 \cdot p^i \cdot c_i \cdot T^{i-1} \cdot (\ln T)^{(k-i)/(k-1)}$ for the i -element-edges in \mathcal{G}^* , $i = 2, \dots, k$. By (4) we infer $n_i \leq N/(8 \cdot k \cdot e^s) \leq |V^*|/(7 \cdot k)$, thus $\sum_{i=2}^k n_i < |V^*|/7$. We delete these vertices from V^* and obtain a subset $V^{**} \subseteq V^*$ with $|V^{**}| \geq (6/7) \cdot |V^*|$. For the induced subhypergraph $\mathcal{G}^{**} = (V^{**}, \mathcal{E}_2^{**} \cup \dots \cup \mathcal{E}_k^{**})$ of \mathcal{G}^* with $\mathcal{E}_i^{**} := \mathcal{E}_i \cap [V^{**}]^i$, $i = 2, \dots, k$, we infer with (4) for each vertex $v \in V^{**}$:

$$|V^{**}| \geq (3/4) \cdot N/e^s \quad \text{and} \quad d_i^{**}(v) \leq 8 \cdot k^2 \cdot c_i \cdot (T/e^s)^{i-1} \cdot (\ln T)^{(k-i)/(k-1)},$$

where $d_i^{**}(v)$ is the degree of v for the i -element edges in \mathcal{G}^{**} . For $s := 10^{-3} \cdot \ln T$ and $c_i < 1/8 \cdot \binom{k-1}{i-1} / (10^{(3(k-i))/(k-1)} \cdot k^2)$, $i = 2, \dots, k$, we have

$$d_i^{**}(v) \leq 8 \cdot k^2 \cdot c_i \cdot (T/e^s)^{i-1} \cdot (\ln T)^{\frac{k-i}{k-1}} \leq \binom{k-1}{i-1} \cdot s^{\frac{k-i}{k-1}} \cdot (T/e^s)^{i-1},$$

which proves (2). By possibly deleting some more vertices and all incident edges we obtain $(3/4) \cdot N/e^s \leq |V^{**}| \leq N/e^s$. This probabilistic argument can be derandomized by using the method of conditional probabilities and yields a deterministic algorithm with running time $O(|V| + \sum_{i=2}^k |\mathcal{E}_i|)$. \square

We prove Theorem 1 with an approach similar to that in [1]. The difference between their arguments and ours is, that we do not apply Lemma 1 step by step from the beginning, but use first Lemma 2 to jump to a suitable subhypergraph:

Proof. Apply Lemma 2 with $s := 10^{-3} \cdot \ln T$ to the hypergraph $\mathcal{G} = (V, \mathcal{E}_2 \cup \dots \cup \mathcal{E}_k)$ on N vertices and obtain an induced subhypergraph $\mathcal{G}_{s-1} := (V_{s-1}, \mathcal{E}_{2;s-1} \cup \dots \cup \mathcal{E}_{k;s-1})$ on n vertices with $\mathcal{E}_{i;s-1} := \mathcal{E}_i \cap [V_{s-1}]^i$, $i = 2, \dots, k$, and with $(3/4) \cdot N/e^s \leq n \leq N/e^s$, and for each vertex $v \in V_{s-1}$ its degree $d_{i;s-1}(v)$ in \mathcal{G}_{s-1} for the i -element edges in $\mathcal{E}_{i;s-1}$ satisfies $d_{i;s-1}(v) \leq \binom{k-1}{i-1} \cdot s^{(k-i)/(k-1)} \cdot (T/e^s)^{i-1}$. Set $n_{s-1} := n$ and $t_{s-1} := T/e^s$. By iteratively applying Lemma 1 as in [1] with $\varepsilon := 10^{-6}/\ln T$ to the hypergraphs \mathcal{G}_{r-1} , we obtain for $r = s, \dots, 10^{-2} \cdot \ln T$ independent sets $I_r \subseteq V_{r-1}$ and hypergraphs $\mathcal{G}_r = (V_r, \mathcal{E}_{2;r} \cup \dots \cup \mathcal{E}_{k;r})$ with

$|V_r| = n_r$, where $(3/4) \cdot N \cdot (1 - \varepsilon)^{r+1-s} / e^{r+1} \leq n_r \leq N / e^{r+1}$ with numbers $t_r \leq T \cdot (1 + \varepsilon)^{r+1-s} / e^{r+1}$, such that

$$\begin{aligned} \alpha(\mathcal{G}_r) &\geq |I_r| + \alpha(\mathcal{G}_{r+1}) \quad \text{and} \quad |I_r| \geq (0.99 \cdot n_{r-1} \cdot w_r) / (e \cdot t_{r-1}) \\ |V_r| &\geq (n_{r-1} \cdot (1 - \varepsilon)) / e \\ d_{i,r}(v) &\leq \binom{k-1}{i-1} \cdot (r+1)^{\frac{k-i}{k-1}} \cdot (t_r)^{i-1} \end{aligned}$$

for each $v \in V_r$, where $d_{i,r}(v)$ is the degree for the i -element edges in \mathcal{G}_r of v . With $(1 + \varepsilon)^n > 1 + \varepsilon \cdot n$, $1 + \varepsilon \leq e^\varepsilon$, $r \leq 10^{-2} \cdot \ln T$ and $\varepsilon = 10^{-6} / \ln T$ we have

$$\frac{n_r}{t_r} \geq \frac{(3/4) \cdot N \cdot (1 - \varepsilon)^{r+1-s} / e^{r+1}}{T \cdot (1 + \varepsilon)^{r+1-s} / e^{r+1}} \geq \frac{(3/4) \cdot N}{T} \cdot \frac{(1 - \varepsilon)^r}{(1 + \varepsilon)^r} \geq 0.74 \cdot \frac{N}{T}.$$

Hence, with $w_r = (r+1)^{1/(k-1)} - r^{1/(k-1)}$ and $s = 10^{-3} \cdot \ln T$, we obtain for some constant $C_k > 0$ an independent set $I = I_s \cup \dots \cup I_{(\ln T)/10^2}$ in \mathcal{G} with

$$\begin{aligned} \alpha(\mathcal{G}) \geq |I| &= \sum_{r=s}^{(\ln T)/10^2} |I_r| \geq 0.99 \cdot \frac{0.74}{e} \cdot \frac{N}{T} \cdot \sum_{r=s}^{(\ln T)/10^2} w_r \geq \\ &\geq \frac{0.73}{e} \cdot \frac{N}{T} \cdot \sum_{r=s}^{(\ln T)/10^2} ((r+1)^{\frac{1}{k-1}} - r^{\frac{1}{k-1}}) \geq C_k \cdot \frac{N}{T} \cdot (\ln T)^{\frac{1}{k-1}}, \end{aligned}$$

which gives the lower bound (1) in Theorem 1. The time bound for the corresponding deterministic algorithm can be estimated as follows: Lemma 2 is applied in time $O(|V| + \sum_{i=2}^k |\mathcal{E}_i|)$ and all applications of Lemma 1 are done in time $O(\sum_{r=(\ln T)/10^3}^{(\ln T)/10^2} ((N/e^r) \cdot (T \cdot (1 + \varepsilon)^{r+1-s} / e^{r+1})^{4(k-1)})) = o(N \cdot T^{4(k-1)})$, compare Lemma 1, hence we have the time bound $o(N \cdot T^{4(k-1)})$. \square

In [9] it has been shown that one may relax in Theorem 3 the assumptions: it suffices to have a linear hypergraph. Similarly, one can show:

Theorem 4. *Let $k \geq 3$ be a fixed integer. Let $\mathcal{G} = (V, \mathcal{E}_3 \cup \dots \cup \mathcal{E}_k)$ be a linear hypergraph with $|V| = N$ such that the average degrees $t_i^{i-1} := i \cdot |\mathcal{E}_i| / |V|$ for the i -element edges satisfy $t_i^{i-1} \leq c_i \cdot T^{i-1} \cdot (\ln T)^{(k-i)/(k-1)}$ for some number $T \geq 1$, where $c_i > 0$ are constants with $c_i < 1/32 \cdot \binom{k-1}{i-1} / (10^{3(k-i)/(k-1)} \cdot k^6)$, $i = 3, \dots, k$.*

Then, for some constant $C_k > 0$, one can find deterministically in time $O(N \cdot T^{4k-2})$ an independent set $I \subseteq V$ such that $|I| = \Omega((N/T) \cdot (\ln T)^{1/(k-1)})$.

3 Areas of the Convex Hull of j Points

For distinct points $P, Q \in [0, 1]^2$ let PQ denote the *line* through P and Q , and let $[P, Q]$ be the *segment* between P and Q . Let $\text{dist}(P, Q)$ denote the *Euclidean distance* between the points P and Q . For points $P_1, \dots, P_l \in [0, 1]^2$

let $\text{area}(P_1, \dots, P_l)$ be the area of the convex hull of P_1, \dots, P_l . A *strip* centered at the line PQ of width w is the set of all points in \mathbb{R}^2 , which are at Euclidean distance at most $w/2$ from the line PQ . We define a lexicographic order \leq_{lex} on the unit square $[0, 1]^2$: for points $P = (p_x, p_y) \in [0, 1]^2$ and $Q = (q_x, q_y) \in [0, 1]^2$ let $P \leq_{lex} Q : \iff (p_x < q_x) \text{ or } (p_x = q_x \text{ and } p_y < q_y)$.

Lemma 3. (a) Let $P_1, \dots, P_l \in [0, 1]^2$, $l \geq 3$, be points. If $\text{area}(P_1, \dots, P_l) \leq A$, then for any distinct points P_i, P_j every other point P_k , $k \neq i, j$, is contained in a strip centered at the line $P_i P_j$ of width $4 \cdot A / \text{dist}(P_i, P_j)$.

(b) Let $P, R \in [0, 1]^2$ be distinct points with $P \leq_{lex} R$. Then all points $Q \in [0, 1]^2$ with $P \leq_{lex} Q \leq_{lex} R$ and $\text{area}(P, Q, R) \leq A$ are contained in a parallelogram of area $4 \cdot A$.

In the following we prove Theorem 2.

Proof. Let $k \geq 3$ be a fixed and let $n \geq k$ be any integer. For a constant $\beta > 0$, which will be specified later, we select uniformly at random and independently of each other $N := n^{1+\beta}$ points P_1, \dots, P_N in $[0, 1]^2$. Set $D_0 := N^{-\gamma}$ for a constant γ with $0 < \gamma < 1$ and let $A_3, \dots, A_k > 0$ be numbers, which will be fixed later. We form a random hypergraph $\mathcal{G} = (V, \mathcal{E}_2 \cup \dots \cup \mathcal{E}_k)$ with vertex-set $V = \{1, \dots, N\}$, where vertex $i \in V$ corresponds to the random point $P_i \in [0, 1]^2$, and with edges of cardinality at most k . Let $\{i_1, i_2\} \in \mathcal{E}_2$ if and only if $\text{dist}(P_{i_1}, P_{i_2}) \leq D_0$. Moreover, for $j = 3, \dots, k$, let $\{i_1, \dots, i_j\} \in \mathcal{E}_j$ if and only if $\text{area}(P_{i_1}, \dots, P_{i_j}) \leq A_j$ and $\{i_1, \dots, i_j\}$ does not contain any edges from \mathcal{E}_2 . We want to find a large independent set $I \subseteq V$ in \mathcal{G} , as I yields a subset $P(I) \subseteq [0, 1]^2$ of size $|I|$ such that the area of the convex hull of each j distinct points, $j = 3, \dots, k$, from $P(I)$ is bigger than A_j . To do so, first we estimate the expected numbers $E[|\mathcal{E}_j|]$ of j -element edges and $E[s_{2;(g,i,j)}(\mathcal{G})]$ of $(2; (g, i, j))$ -cycles in \mathcal{G} , and we prove that these numbers are not too big. Then we show the existence of a certain induced, linear subhypergraph $\mathcal{G}^* = (V, \mathcal{E}_3^* \cup \dots \cup \mathcal{E}_k^*)$ (no 2-element edges anymore) of \mathcal{G} , which satisfies the assumptions of Theorem 4, and then we obtain a large independent set.

Lemma 4. For $j = 3, \dots, k$, there exist constants $c_j > 0$ such that

$$E[|\mathcal{E}_j|] \leq c_j \cdot A_j^{j-2} \cdot N^j. \quad (5)$$

Proof. For integers i_1, \dots, i_j with $1 \leq i_1 < \dots < i_j \leq N$ we estimate the probability $\text{Prob}(\text{area}(P_{i_1}, \dots, P_{i_j}) \leq A_j)$. We may assume that $P_{i_1} \leq_{lex} \dots \leq_{lex} P_{i_j}$. Then $\text{area}(P_{i_1}, \dots, P_{i_j}) \leq A_j$ implies $\text{area}(P_{i_1}, P_{i_g}, P_{i_j}) \leq A_j$ for $g = 2, \dots, j-1$. The points P_{i_1} and P_{i_j} with $P_{i_1} \leq_{lex} P_{i_j}$ may be anywhere in $[0, 1]^2$. Given $P_{i_1}, P_{i_j} \in [0, 1]^2$, by Lemma 3(b) all points P_{i_g} , $g = 2, \dots, j-1$, are contained in a parallelogram of area $4 \cdot A_j$, which happens with probability at most $(4 \cdot A_j)^{j-2}$. As there are $\binom{N}{j}$ choices for j out of N points, for some constants $c_j > 0$, $j = 3, \dots, k$, we obtain $E[|\mathcal{E}_j|] \leq c_j \cdot A_j^{j-2} \cdot N^j$. \square

Next we estimate the expected numbers $E[s_{2;(g,i,j)}(\mathcal{G})]$ of $(2; (g, i, j))$ -cycles, $2 \leq g \leq i \leq j \leq k$ but $g < j$, in $\mathcal{G} = (V, \mathcal{E}_2 \cup \dots \cup \mathcal{E}_k)$.

Lemma 5. *Let $2 \leq g \leq i \leq j \leq k$ with $i \geq 3$ and $g < j$, and let $0 < A_3 \leq \dots \leq A_k$. Then, there exist constants $c_{2;(g,i,j)} > 0$ such that for $D_0^2 \geq 2 \cdot A_j$ it is*

$$E[s_{2;(g,i,j)}(\mathcal{G})] \leq c_{2;(g,i,j)} \cdot A_i^{i-2} \cdot A_j^{j-g} \cdot N^{i+j-g} \cdot (\log N)^3. \quad (6)$$

Proof. We estimate the probability that $(i + j - g)$ points, which are chosen uniformly at random and independently of each other in $[0, 1]^2$, form sets of i and j points with areas of the convex hulls at most A_i and A_j , respectively, conditioned on the event that distinct points have Euclidean distance bigger than D_0 . Both sets have g points in common, say P_1, \dots, P_g , where $P_1 \leq_{lex} \dots \leq_{lex} P_g$. Let the sets of i and j points be P_1, \dots, P_i and $P_1, \dots, P_g, Q_{g+1}, \dots, Q_j$ with area $(P_1, \dots, P_i) \leq A_i$ and area $(P_1, \dots, P_g, Q_{g+1}, \dots, Q_j) \leq A_j$, respectively. The point P_1 may be anywhere in $[0, 1]^2$. Given $P_1 \in [0, 1]^2$, we have $\text{Prob}(r \leq \text{dist}(P_1, P_g) \leq r + dr) \leq \pi \cdot r \, dr$. Given $P_1, P_g \in [0, 1]^2$ with $\text{dist}(P_1, P_g) = r$, by Lemma 3(b) all points P_2, \dots, P_{g-1} are contained in a parallelogram with area $4 \cdot A_i$, which happens with probability at most $(4 \cdot A_i)^{g-2}$.

Given $P_1, \dots, P_g \in [0, 1]^2$ with $\text{dist}(P_1, P_g) = r$, by Lemma 3(a) all points P_{g+1}, \dots, P_i are contained in a strip S_i of width $w = 4 \cdot A_i/r$, and all points Q_{g+1}, \dots, Q_j are contained in a strip S_j of width $w = 4 \cdot A_j/r$, where both strips are centered at the line P_1P_g . Set $S_i^* := S_i \cap [0, 1]^2$ and $S_j^* := S_j \cap [0, 1]^2$, which have areas at most $4 \cdot \sqrt{2} \cdot A_i/r$ and $4 \cdot \sqrt{2} \cdot A_j/r$, respectively.

For the convex hulls of P_1, \dots, P_i and $P_1, \dots, P_g, Q_{g+1}, \dots, Q_j$ we denote their extremal points by P', P'' and Q', Q'' , respectively, i.e., $P', P'' \in \{P_1, \dots, P_i\}$ and $Q', Q'' \in \{P_1, \dots, P_g, Q_{g+1}, \dots, Q_j\}$ and, say $P' \leq_{lex} P''$ and $Q' \leq_{lex} Q''$, it is $P' \leq_{lex} P_1, \dots, P_i \leq_{lex} P''$ and $Q' \leq_{lex} P_1, \dots, P_g, Q_{g+1}, \dots, Q_j \leq_{lex} Q''$. Given $P_1 \leq_{lex} \dots \leq_{lex} P_g$, there are three possibilities each for the convex hulls of P_1, \dots, P_i and $P_1, \dots, P_g, Q_{g+1}, \dots, Q_j$: extremal are (i) P_1 and P_g , or (ii) only one point, P_1 or P_g , or (iii) none of P_1 and P_g .

Consider the convex hull of the points P_1, \dots, P_i . In case (i), given $P_1, \dots, P_g \in [0, 1]^2$ with $\text{dist}(P_1, P_g) = r$, as in the proof of Lemma 4 we infer

$$\text{Prob}(\text{area}(P_1, \dots, P_i) \leq A_i \mid P_1, \dots, P_g; \text{case (i)}) \leq (4 \cdot A_i)^{i-g}. \quad (7)$$

In case (ii), either P_1 or P_g is extremal for the convex hull of P_1, \dots, P_i . By Lemma 3(a), the second extremal point is contained in the set S_i^* , which happens with probability at most $4 \cdot \sqrt{2} \cdot A_i/r$. Given both extremal points $P', P'' \in [0, 1]^2$, by Lemma 3(b) all points $P_{g+1}, \dots, P_i \neq P', P''$ are contained in a parallelogram of area $4 \cdot A_i$, hence, with $\text{dist}(P_1, P_g) = r$ we infer

$$\text{Prob}(\text{area}(P_1, \dots, P_i) \leq A_i \mid P_1, \dots, P_g; \text{case (ii)}) \leq ((4 \cdot A_i)^{i-g} \cdot \sqrt{2})/r. \quad (8)$$

In case (iii) neither point P_1 nor P_g is extremal for the convex hull of P_1, \dots, P_i . With area $(P_1, \dots, P_i) \leq A_i$, by Lemma 3(a) both extremal points P' and P'' , say $P' \leq_{lex} P_1 \leq_{lex} P_g \leq_{lex} P''$, are contained in the strip S_i of width $4 \cdot A_i/r$, which is centered at the line P_1P_g . Given $P_1 \in [0, 1]^2$, the probability that $\text{dist}(P_1, P') \in [s, s + ds]$ is at most the difference of the areas of the balls with center P_1 and with radii $(s + ds)$ and s , respectively, intersected with the strip S_i .

Since distinct points have Euclidean distance bigger than D_0 , we have $r, s > D_0$. A circle with center P_1 and radius $s > D_0$ intersects both boundaries of the strip S_i of width $4 \cdot A_i/r$ in four points $R \leq_{lex} R'$ and $R'' \leq_{lex} R'''$, where R, R' are on one boundary of the strip S_i and R'', R''' are on the other boundary. To see this, notice that $s > 2 \cdot A_i/r$ follows from $r, s > D_0$ and $D_0^2 > 2 \cdot A_j \geq 2 \cdot A_i$. Let $\varepsilon(s)$ be the angle between the lines P_1R and P_1R'' . Then, by using $\varepsilon/2 \leq \sin \varepsilon$ for $\varepsilon \leq \pi/2$ and $\sin(\varepsilon(s)/2) = 2 \cdot A_i/(r \cdot s) < 2 \cdot A_i/D_0^2 \leq 1$, we infer

$$\begin{aligned} \text{Prob}(\text{dist}(P_1, P') \in [s, s + ds] \mid P_1) &\leq ((2 \cdot \varepsilon(s)))/(2 \cdot \pi) \cdot 2 \cdot \pi \cdot s \, ds \leq \\ &\leq 8 \cdot \sin(\varepsilon(s)/2) \cdot s \, ds = (16 \cdot A_i/r) \, ds. \end{aligned}$$

Given $P' \in [0, 1]^2$ with $\text{dist}(P_1, P') = s$, the second extremal point $P'' \in [0, 1]^2$ is contained in a strip centered at the line P_1P' of width $4 \cdot A_i/s$, which occurs with probability at most $4 \cdot \sqrt{2} \cdot A_i/s$. Given both points P', P'' , by Lemma 3(b) all points $P_{g+1}, \dots, P_i \neq P', P''$ are contained in a parallelogram of area $4 \cdot A_i$. Hence, given $P_1, \dots, P_g \in [0, 1]^2$, with $s > D_0 = N^{-\gamma}$ and $\gamma > 0$, we infer:

$$\begin{aligned} &\text{Prob}(\text{area}(P_1, \dots, P_i) \leq A_i \mid P_1, \dots, P_g; \text{case (iii)}) \\ &\leq (4 \cdot A_i)^{i-g} \cdot \int_{D_0}^{\sqrt{2}} \frac{4 \cdot \sqrt{2}}{r \cdot s} \, ds = \sqrt{32} \cdot (4 \cdot A_i)^{i-g} \cdot \frac{\ln \sqrt{2} + \gamma \cdot \ln N}{r}. \quad (9) \end{aligned}$$

Summarizing cases (i-iii) with (7)-(9), and $r \leq \sqrt{2}$ and $0 < \gamma < 1$ we obtain:

$$\begin{aligned} &\text{Prob}(\text{area}(P_1, \dots, P_i) \leq A_i \mid P_1, \dots, P_g) \\ &\leq (4 \cdot A_i)^{i-g} \cdot \frac{\sqrt{8} + \sqrt{8} \cdot (\ln 2 + 2 \cdot \gamma \cdot \ln N)}{r} \leq (4 \cdot A_i)^{i-g} \cdot \frac{11 \cdot \ln N}{r}. \quad (10) \end{aligned}$$

Similarly, it follows $\text{Prob}(\text{area}(P_1, \dots, P_g, Q_{g+1}, \dots, Q_j) \leq A_j \mid P_1, \dots, P_g) \leq ((4 \cdot A_j)^{j-g} \cdot 11 \cdot \ln N)/r$ holds. Hence, we obtain for constants $c_{2;(g,i,j)}^* > 0$:

$$\begin{aligned} &\text{Prob}(P_1, \dots, P_i \text{ and } P_1, \dots, P_g, Q_{g+1}, \dots, Q_j \text{ is a } (2; (g, i, j)\text{-cycle}) \leq \\ &\leq \int_{D_0}^{\sqrt{2}} (4 \cdot A_i)^{g-2} \cdot \left((4 \cdot A_i)^{i-g} \cdot \frac{11 \cdot \ln N}{r} \right) \cdot \left((4 \cdot A_j)^{j-g} \cdot \frac{11 \cdot \ln N}{r} \right) \cdot \pi \cdot r \, dr \\ &\leq c_{2;(g,i,j)}^* \cdot A_i^{i-2} \cdot A_j^{j-g} \cdot (\log N)^3 \quad \text{as } D_0 = N^{-\gamma}, \gamma > 0 \text{ is constant.} \quad (11) \end{aligned}$$

There are $\binom{N}{i+j-g}$ choices for $i+j-g$ out of N points, hence for constants $c_{2;(g,i,j)} > 0$, $j = 2, \dots, k-1$, we get with (11) the upper bound:

$$E[s_{2;(g,i,j)}(\mathcal{G})] \leq c_{2;(g,i,j)} \cdot A_i^{i-2} \cdot A_j^{j-g} \cdot N^{i+j-g} \cdot (\log N)^3. \quad \square$$

For distinct points $P, Q \in [0, 1]^2$, it is $\text{Prob}(\text{dist}(P, Q) \leq D_0) \leq \pi \cdot D_0^2$. With $D_0 = N^{-\gamma}$ we infer $E[|\mathcal{E}_2|] \leq \binom{N}{2} \cdot \pi \cdot D_0^2 \leq c_2 \cdot N^{2-2\gamma}$ for some constant $c_2 > 0$. By Markov's inequality, using this and the estimates (5) and (6) there exist N points $P_1, \dots, P_N \in [0, 1]^2$ such that the resulting hypergraph $\mathcal{G} = (V, \mathcal{E}_2 \cup \dots \cup \mathcal{E}_k)$ with $|V| = N$ satisfies for $2 \leq g \leq i \leq j \leq k$ but $g < j$:

$$|\mathcal{E}_2| \leq k^3 \cdot c_2 \cdot N^{2-2\gamma} \quad \text{and} \quad |\mathcal{E}_j| \leq k^3 \cdot c_j \cdot A_j^{j-2} \cdot N^j \quad (12)$$

$$s_{2;(g,i,j)}(\mathcal{G}) \leq k^3 \cdot c_{2;(g,i,j)} \cdot A_i^{i-2} \cdot A_j^{j-g} \cdot N^{i+j-g} \cdot (\log N)^3. \quad (13)$$

For suitable constants $c'_j > 0$, $j = 3, \dots, k$, which will be fixed later, we set

$$A_j := (c'_j \cdot (\log n)^{1/(j-2)})/n^{(j-1)/(j-2)}. \quad (14)$$

Lemma 6. *For fixed $\gamma > 1/2$ it is $|\mathcal{E}_2| = o(|V|)$.*

Proof. Using (12) and $|V| = N$, we have $|\mathcal{E}_2| = o(|V|)$ provided that $N^{2-2\gamma} = o(N) \iff N^{1-2\gamma} = o(1)$, which holds for $\gamma > 1/2$. \square

Lemma 7. *For fixed $2 \leq g \leq i \leq j \leq k$ but $g < j$ and for fixed constant β with $0 < \beta < (j-g)/((j-2) \cdot (i+j-g-1))$ it is $s_{2;(g,i,j)}(\mathcal{G}) = o(|V|)$.*

Proof. By using (13) and (14) and $|V| = N = n^{1+\beta}$ with fixed $\beta > 0$ we have $s_{2;(g,i,j)}(\mathcal{G}) = o(|V|)$ for $j = 2, \dots, k-1$, provided that

$$\begin{aligned} A_i^{i-2} \cdot A_j^{j-g} \cdot N^{i+j-g} \cdot (\log N)^3 &= o(N) \\ \iff (\log n)^{4+\frac{i-g}{j-2}} \cdot n^{(1+\beta)(i+j-g-1)-(i-1)-\frac{(j-g)(j-1)}{j-2}} &= o(1) \\ \iff (1+\beta) \cdot (i+j-g-1) < i-1 + ((j-g) \cdot (j-1))/(j-2), \end{aligned}$$

which holds for $\beta < (j-g)/((j-2) \cdot (i+j-g-1))$. \square

Fix $\beta := 1/(2 \cdot k^2)$ and $\gamma := k/(2 \cdot (k-1))$. Then, with (14) and $D_0 = N^{-\gamma}$ and $N = n^{1+\beta}$ all assumptions in Lemmas 5–7 are fulfilled. We delete one vertex from each 2-element edge $E \in \mathcal{E}_2$ and each $(2; (g, i, j))$ -cycle, $2 \leq g \leq i \leq j \leq k$ but $g < j$, in \mathcal{G} . Let $V^* \subseteq V$ be the set of remaining vertices. The induced subhypergraph $\mathcal{G}^* = (V^*, \mathcal{E}_3^* \cup \dots \cup \mathcal{E}_k^*)$ of \mathcal{G} with $\mathcal{E}_j^* := \mathcal{E}_j \cap [V^*]^j$, $j = 3, \dots, k$, is linear, and by (12), and Lemmas 6 and 7 fulfills $|V^*| \geq N/2$ and $|\mathcal{E}_j^*| \leq k^3 \cdot c_j \cdot A_j^{j-2} \cdot N^j$. By (14), the hypergraph \mathcal{G}^* has average degree

$$t_j^{j-1} = j \cdot |\mathcal{E}_j^*|/|V^*| \leq 2 \cdot k^3 \cdot j \cdot c_j \cdot (c'_j)^{j-2} \cdot N^{j-1} \cdot \log n/n^{j-1} =: (t_j(1))^{j-1}$$

for the j -element edges. Fix a constant $c' > 0$ such that $C_k/(2 \cdot c') \cdot \beta^{1/(k-1)} > 1$ and set $T := c' \cdot (N/n) \cdot (\log n)^{1/(k-1)}$. Then fix constants $c'_j > 0$, $j = 3, \dots, k$, in (14) such that

$$\begin{aligned} (t_j(1))^{j-1} &= (2 \cdot k^3 \cdot j \cdot c_j \cdot (c'_j)^{j-2} \cdot N^{j-1} \cdot \log n)/n^{j-1} \leq \\ &\leq 1/32 \cdot \binom{k-1}{j-1} / (10^{3(k-j)/(k-1)} \cdot k^6) \cdot T^{j-1} \cdot (\log T)^{(k-j)/(k-1)}. \end{aligned}$$

Then, the assumptions in Theorem 4 are satisfied for \mathcal{G}^* , and its independence number $\alpha(\mathcal{G}^*)$ satisfies for some constant $C_k > 0$:

$$\begin{aligned} \alpha(\mathcal{G}) &\geq \alpha(\mathcal{G}^*) \geq C_k \cdot (|V^{**}|/T) \cdot (\log T)^{\frac{1}{k-1}} \geq C_k \cdot (N/(2 \cdot T)) \cdot (\log T)^{\frac{1}{k-1}} \geq \\ &\geq \frac{C_k \cdot n}{2 \cdot c' \cdot (\log n)^{\frac{1}{k-1}}} \cdot (\log(n^\beta))^{\frac{1}{k-1}} \geq n. \end{aligned}$$

The vertices of an independent set I with $|I| = n$ yield n points among the N points $P_1, \dots, P_N \in [0, 1]^2$, such that for $j = 3, \dots, k$ the area of the convex hull of any j distinct points of these n points is $\Omega((\log n)^{1/(j-2)}/n^{(j-1)/(j-2)})$ as desired. \square

References

1. M. Ajtai, J. Komlós, J. Pintz, J. Spencer and E. Szemerédi, *Extremal Uncrowded Hypergraphs*, Journal of Combinatorial Theory Ser. A, 32, 1982, 321–335.
2. G. Barequet, *A Lower Bound for Heilbronn's Triangle Problem in d Dimensions*, SIAM Journal on Discrete Mathematics 14, 2001, 230–236.
3. G. Barequet and J. Naor, *Large $k - D$ Simplices in the D -Dimensional Unit Cube*, Proceedings '17th Canadian Conference on Computational Geometry', 2005, 30–33.
4. G. Barequet and A. Shaikhet, *The On-Line Heilbronn's Triangle Problem in d Dimensions*, Proceedings '12th Annual Computing and Combinatorics Conference COCOON'06', Springer Verlag, LNCS 4112, 2006, 408–417.
5. C. Bertram-Kretzberg and H. Lefmann, *The Algorithmic Aspects of Uncrowded Hypergraphs*, SIAM Journal on Computing 29, 1999, 201–230.
6. C. Bertram-Kretzberg, T. Hofmeister and H. Lefmann, *An Algorithm for Heilbronn's Problem*, SIAM Journal on Computing 30, 2000, 383–390.
7. P. Brass, *An Upper Bound for the d -Dimensional Analogue of Heilbronn's Triangle Problem*, SIAM Journal on Discrete Mathematics 19, 2005, 192–195.
8. B. Chazelle, *Lower Bounds on the Complexity of Polytope Range Searching*, Journal of the American Mathematical Society 2, 1989, 637–666.
9. R. A. Duke, H. Lefmann and V. Rödl, *On Uncrowded Hypergraphs*, Random Structures & Algorithms 6, 1995, 209–212.
10. A. Fundia, *Derandomizing Chebychev's Inequality to find Independent Sets in Uncrowded Hypergraphs*, Random Structures & Algorithms, 8, 1996, 131–147.
11. J. Komlós, J. Pintz and E. Szemerédi, *On Heilbronn's Triangle Problem*, Journal of the London Mathematical Society, 24, 1981, 385–396.
12. J. Komlós, J. Pintz and E. Szemerédi, *A Lower Bound for Heilbronn's Problem*, Journal of the London Mathematical Society, 25, 1982, 13–24.
13. H. Lefmann, *Distributions of Points in the d Dimensions and Large k -Points Simplices*, Proceedings '11th Annual Computing and Combinatorics Conference COCOON'05', Springer Verlag, LNCS 3595, eds. Lusheng Wang, 2005, 514–523.
14. H. Lefmann, *Distributions of Points in the Unit-Square and Large k -Gons*, Proceedings '16th ACM-SIAM Symposium on Discrete Algorithms SODA', ACM and SIAM, 2005, 241–250.
15. K. F. Roth, *On a Problem of Heilbronn*, Journal of the London Mathematical Society 26, 1951, 198–204.
16. K. F. Roth, *On a Problem of Heilbronn, II, and III*, Proc. of the London Mathematical Society (3), 25, 1972, 193–212, and 543–549.
17. K. F. Roth, *Estimation of the Area of the Smallest Triangle Obtained by Selecting Three out of n Points in a Disc of Unit Area*, Proc. of Symposia in Pure Mathematics, 24, 1973, AMS, Providence, 251–262.
18. K. F. Roth, *Developments in Heilbronn's Triangle Problem*, Advances in Mathematics, 22, 1976, 364–385.
19. W. M. Schmidt, *On a Problem of Heilbronn*, Journal of the London Mathematical Society (2), 4, 1972, 545–550.
20. J. Spencer, *Turán's Theorem for k -Graphs*, Discrete Mathematics, 2, 1972, 183–186.