Available online at www.sciencedirect.com
ScienceDirect
European Journal of Combinatorics 29 (2008) 946–965
European Journal of Combinatorics
www.elsevier.com/locate/ejc

Distributions of points in the unit square and large k-gons☆

Hanno Lefmann

Fakultät für Informatik, TU Chemnitz, D-09107 Chemnitz, Germany

Available online 5 March 2008

Abstract

We consider a generalization of Heilbronn's triangle problem by asking, given any integers $n \geq k$, for the supremum $\Delta_k(n)$ of the minimum area determined by the convex hull of some k of n points in the unit square $[0, 1]^2$, where the supremum is taken over all distributions of n points in $[0, 1]^2$. Improving the lower bound $\Delta_k(n) = \Omega(1/n^{(k-1)/(k-2)})$ from [C. Bertram-Kretzberg, T. Hofmeister, H. Lefmann, An algorithm for Heilbronn's problem, SIAM Journal on Computing 30 (2000) 383–390] and from [W.M. Schmidt, On a problem of Heilbronn, Journal of the London Mathematical Society (2) 4 (1972) 545–550] for $k = 4$, we show that $\Delta_k(n) = \Omega((\log n)^{1/(k-2)}/n^{(k-1)/(k-2)})$ for fixed integers $k \geq 3$ as asked for in [C. Bertram-Kretzberg, T. Hofmeister, H. Lefmann, An algorithm for Heilbronn's problem, SIAM Journal on Computing 30 (2000) 383–390]. Moreover, we provide a deterministic polynomial time algorithm which finds n points in $[0, 1]^2$, which achieve this lower bound on $\Delta_k(n)$.

1. Introduction

The problem of Heilbronn asks for a distribution of n points in the unit square $[0, 1]^2$ (or unit ball) such that the minimum area of a triangle determined by three of these n points achieves its largest value. Let $\Delta_3(n)$ denote the supremum of the minimum area of a triangle among n points, where the supremum is taken over all distributions of n points in $[0, 1]^2$. For primes n the points $1/n \cdot (i \mod n, i^2 \mod n)$, $i = 0, \ldots, n-1$, yield $\Delta_3(n) = \Omega(1/n^2)$. While for

☆ A preliminary version of a part of this paper appeared as an extended abstract in Proceedings '16th Annual ACM-SIAM Symposium on Discrete Algorithms SODA'2005, ACM and SIAM, 241–250.

E-mail address: lefmann@informatik.tu-chemnitz.de.
URL: http://www.tu-chemnitz.de/informatik/ThIS/.

0195-6698/$ - see front matter
doi:10.1016/j.ejc.2007.11.021

some time this lower bound on $\Delta_3(n)$ was believed to be also the upper bound, Komlós, Pintz and Szemerédi [12] showed that $\Delta_3(n) = \Omega(\log n/n^2)$, see [5] for a deterministic polynomial time algorithm achieving this lower bound on $\Delta_3(n)$. Upper bounds on $\Delta_3(n)$ were given by Roth [16–20] and Schmidt [21] and, improving these earlier results, the currently best upper bound $\Delta_3(n) = O(2^{c\sqrt{\log n}}/n^{8/7})$, where $c > 0$ is a constant, is due to Komlós, Pintz and Szemerédi [11]. Recently, Jiang, Li and Vitany [10] showed with methods from Kolmogorov complexity theory that if n points are distributed uniformly at random and independently of each other in the unit square $[0, 1]^2$, then the expected value of the minimum area of a triangle formed by some three of these n random points is equal to $\Theta(1/n^3)$.

Variants of Heilbronn's triangle problem in higher dimensions were investigated by Barequet [2,3], who considered the minimum volumes of simplices among n points in the d-dimensional unit cube $[0, 1]^d$, see also [14,15] and Brass [6].

Given a fixed integer $k \geq 3$, a generalization of Heilbronn's triangle problem to k points, see Schmidt [21], asks to maximize the minimum area of the convex hull of any k distinct points in a distribution of n points in the unit square $[0, 1]^2$. In particular, let $\Delta_k(n)$ be the supremum of the minimum area of the convex hull determined by some k of n points, where the supremum is taken over all distributions of n points in the unit square $[0, 1]^2$.

Some years ago, for $k = 4$, Schmidt [21] proved the lower bound $\Delta_4(n) = \Omega(1/n^{3/2})$. In [5] a deterministic polynomial time algorithm was given which, given a fixed integer $k \geq 3$, finds for any integer $n \geq k$ a configuration of n points in $[0, 1]^2$, which achieves the lower bound $\Delta_k(n) = \Omega(1/n^{(k-1)/(k-2)})$.

A closely related problem has been considered by Chazelle [7] in connection with lower bounds on the query complexity of range searching problems. In [7] he proved that for any fixed dimension $d \geq 2$ and all integers $k, n \geq 3$ with $\log n \leq k \leq n$ it is $\Delta_k(n) = \Theta(k/n)$. An improvement of the range of k might also improve his lower bound on the query complexity. Here we give an easier proof of Chazelle's bounds on $\Delta_k(n)$ for $\log n \leq k \leq n$.

In [13] the lower bound of Schmidt [21] for the case $k = 4$ has been improved to $\Delta_4(n) = \Omega((\log n)^{1/2}/n^{3/2})$. Here we extend this result to arbitrary fixed integers $k \geq 3$, and improve the lower bounds from [5] by a factor of $\Theta((\log n)^{1/(k-2)})$, as asked for in [5,21]:

Theorem 1.1. *Let $k \geq 3$ be a fixed integer. For integers $n \geq k$ it is*

$$\Delta_k(n) = \Omega\left(\frac{(\log n)^{1/(k-2)}}{n^{(k-1)/(k-2)}}\right). \tag{1}$$

Moreover, one can find deterministically in time $O(n^{2k-2+\delta})$ for any $\delta > 0$ some n points in the unit square $[0, 1]^2$ such that the minimum area of the convex hull determined by some k of these n points is $\Omega((\log n)^{1/(k-2)}/n^{(k-1)/(k-2)})$.

Concerning upper bounds, so far for fixed integers $k \geq 3$, only the bound $\Delta_k(n) = O(1/n)$ is known, compare [21], which follows easily by the pigeonhole principle by partitioning the unit square $[0, 1]^2$ into $(n-1)/(k-1)$ squares of side-lengths $\sqrt{(k-1)/(n-1)} = \Theta(1/\sqrt{n})$ each.

To prove the lower bound (1) in Theorem 1.1, in Section 2 we use probabilistic and non-discrete arguments. These arguments motivate, how we can design a deterministic algorithm for finding n points in $[0, 1]^2$, which achieve the lower bound (1), and help to understand thoroughly the algorithmic part of Theorem 1.1, which is presented in Section 3.

2. A lower bound on $\Delta_k(n)$

For distinct points $P, Q \in [0,1]^2$ let PQ denote the *line* through P and Q and let $[P, Q]$ denote the *segment* between P and Q including the endpoints. Let $\text{dist}(P, Q) := ((p_x - q_x)^2 + (p_y - q_y)^2)^{1/2}$ be the *Euclidean distance* between the points $P = (p_x, p_y)$ and $Q = (q_x, q_y)$. For points $P_1, \ldots, P_l \in [0,1]^2$ their convex hull is the set of all points $P_1 + \sum_{i=2}^{l} \lambda_i \cdot (P_i - P_1)$ with $\lambda_2, \ldots, \lambda_l \geq 0$ and $\sum_{i=2}^{l} \lambda_i = 1$. For points $P_1, \ldots, P_l \in [0,1]^2$ let $\text{area}(P_1, \ldots, P_l)$ denote the area of the convex hull of the points $P_1, \ldots, P_l$. A *strip* centered at the line PQ of width w is the set of all points in $\mathbb{R}^2$ such that their Euclidean distances from the line PQ are at most $w/2$.

First we observe the following simple facts.

Lemma 2.1. *Let* $P_1, \ldots, P_l \in [0,1]^2$ *be points. Then, it is* $\text{area}(P_1, \ldots, P_l) \geq \text{area}(P_1, \ldots, P_{l-1})$.

Proof. This follows by monotonicity, as the convex hull of $P_1, \ldots, P_{l-1}$ is contained in the convex hull of $P_1, \ldots, P_l$. □

Lemma 2.2. *Let* $P_1, \ldots, P_l \in [0,1]^2$, $l \geq 3$, *be points. If* $\text{area}(P_1, \ldots, P_l) \leq A$, *then for any distinct points* P_i, P_j *any point* P_k, $k = 1, \ldots, l$, *is contained in a strip centered at the line* $P_i P_j$ *of width* $4 \cdot A/\text{dist}(P_i, P_j)$.

Proof. Otherwise, by Lemma 2.1 it is $\text{area}(P_1, \ldots, P_l) \geq \text{area}(P_i, P_j, P_k) > (1/2 \cdot \text{dist}(P_i, P_j) \cdot (2 \cdot A))/\text{dist}(P_i, P_j) = A$, which contradicts the assumption $\text{area}(P_1, \ldots, P_l) \leq A$. □

We define a lexicographic order $\leq_{\text{lex}}$ 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

$$P \leq_{\text{lex}} Q :\Longleftrightarrow (p_x < q_x) \text{ or } (p_x = q_x \text{ and } p_y < q_y).$$

Lemma 2.3. *Let* $P, R \in [0,1]^2$ *be distinct points with* $P \leq_{\text{lex}} R$. *Then, all points* $Q \in [0,1]^2$, *such that* $P \leq_{\text{lex}} Q \leq_{\text{lex}} R$ *and* $\text{area}(P, Q, R) \leq A$, *are contained in a parallelogram of area* $4 \cdot A$.

Proof. Given the distinct points $P, R \in [0,1]^2$ with $P \leq_{\text{lex}} R$, by Lemma 2.2 all points Q with $\text{area}(P, Q, R) \leq A$ must be contained in a strip, which is centered at the line PR of width $4 \cdot A/\text{dist}(P, R)$. The condition $P \leq_{\text{lex}} Q \leq_{\text{lex}} R$ defines a parallelogram with base-length $\text{dist}(P, R)$ and height $4 \cdot A/\text{dist}(P, R)$, hence the area of this parallelogram is $4 \cdot A$. □

In the following we prove the lower bound (1) in Theorem 1.1.

Proof. Let $k \geq 3$ be a fixed integer and let $n \geq k$ be an arbitrary integer. For some constant $\beta > 0$, which will be specified later, we select uniformly at random and independently of each other $N := n^{1+\beta}$ points $P_1, \ldots, P_N \in [0,1]^2$ in $[0,1]^2$.

First, for fixed integers $i_1, \ldots, i_k$ with $1 \leq i_1 < \cdots < i_k \leq N$ we give an upper bound on the probability $\text{Prob}(\text{area}(P_{i_1}, \ldots, P_{i_k}) \leq A)$, where $A > 0$ is some number. By possibly renumbering the points, we may assume that $P_{i_1} \leq_{\text{lex}} \cdots \leq_{\text{lex}} P_{i_k}$. By Lemma 2.1, $\text{area}(P_{i_1}, \ldots, P_{i_k}) \leq A$ implies $\text{area}(P_{i_1}, P_{i_j}, P_{i_k}) \leq A$ for $j = 2, \ldots, k-1$. The points P_{i_1} and P_{i_k} with $P_{i_1} \leq_{\text{lex}} P_{i_k}$ may be anywhere in $[0,1]^2$. Given the points P_{i_1} and P_{i_k}, by Lemma 2.3 and our assumptions, i.e., $P_{i_1} \leq_{\text{lex}} \cdots \leq_{\text{lex}} P_{i_k}$ and $\text{area}(P_{i_1}, P_{i_j}, P_{i_k}) \leq A$, all points

P_{i_j}, $j = 2, \ldots, k-1$, are contained in a parallelogram of area $4 \cdot A$, which happens with probability at most $(4 \cdot A)^{k-2}$, hence

$$\text{Prob}(\text{area}(P_{i_1}, \ldots, P_{i_k}) \leq A) \leq (4 \cdot A)^{k-2}. \tag{2}$$

For convenience we use in our arguments hypergraphs.

Definition 2.4. Let $\mathcal{G} = (V, \mathcal{E})$ be a k-uniform hypergraph, i.e., $|E| = k$ for each edge $E \in \mathcal{E}$. An unordered pair $\{E, E'\}$ of distinct edges $E, E' \in \mathcal{E}$ is called a 2-cycle if $|E \cap E'| \geq 2$. A 2-cycle $\{E, E'\}$ in $\mathcal{G}$ is called $(2, j)$-cycle if $|E \cap E'| = j$, $j = 2, \ldots, k-1$. The hypergraph $\mathcal{G}$ is called linear if it does not contain any 2-cycles. The independence number $\alpha(\mathcal{G})$ of $\mathcal{G}$ is the largest size of a subset $I \subseteq V$ which contains no edges from $\mathcal{E}$.

Set $D_0 := N^{-\gamma}$ for some constant γ with $0 < \gamma < 1$, which will be fixed later. For a number $A > 0$ we form a random hypergraph $\mathcal{G} = \mathcal{G}(D_0, A) = (V, \mathcal{E}_2 \cup \mathcal{E}_k)$ with vertex-set $V = \{1, \ldots, N\}$, where vertex $i \in V$ corresponds to the random point $P_i \in [0, 1]^2$, and with 2- and k-element edges. Let $\{i_1, i_2\} \in \mathcal{E}_2$ be a 2-element edge if and only if $\text{dist}(P_{i_1}, P_{i_2}) \leq D_0$. Moreover, let $\{i_1, \ldots, i_k\} \in \mathcal{E}_k$ be a k-element edge if and only if $\text{area}(P_{i_1}, \ldots, P_{i_k}) \leq A$ and $\{i_1, \ldots, i_k\}$ does not contain any 2-element edges from $\mathcal{E}_2$. Since there are $\binom{N}{k}$ choices for k out of N vertices, by (2), for some constant $c_k > 0$ the expected number $E[|\mathcal{E}_k|]$ of k-element edges in this random hypergraph $\mathcal{G}$ can be bounded from above as follows:

$$E[|\mathcal{E}_k|] \leq \binom{N}{k} \cdot 4^{k-2} \cdot A^{k-2} \leq c_k \cdot A^{k-2} \cdot N^k. \tag{3}$$

We want to find in the random hypergraph $\mathcal{G} = (V, \mathcal{E}_2 \cup \mathcal{E}_k)$ a large independent set $I \subseteq V$. An independent set I yields a set $P(I) = \{P_i \mid i \in I\} \subseteq \{P_1, \ldots, P_N\}$ of points in $[0, 1]^2$ of the same size $|I|$ such that for every choice of k distinct points from $P(I)$ the area of their convex hull is bigger than A.

Remark. With (3) already the lower bound $\Delta_k(n) = \Omega(k/n)$ for $\log n \leq k \leq n$ due to Chazelle [7], which has been mentioned in the introduction, follows and yields a slightly different proof of his lower bound. Namely, from (3) it follows that there exist $2 \cdot N$ points in $[0, 1]^2$ such that in the arising hypergraph $\mathcal{G} = (V, \mathcal{E}_k)$ we have $|\mathcal{E}_k| \leq \binom{2N}{k} \cdot 4^{k-2} \cdot A^{k-2}$. Then, it is $|\mathcal{E}_k| \leq N$, if

$$\begin{aligned}
&\binom{2 \cdot N}{k} \cdot 4^{k-2} \cdot A^{k-2} \leq N \\
&\Longleftarrow \left(\frac{2 \cdot e \cdot N}{k}\right)^k \cdot 4^{k-2} \cdot A^{k-2} \leq N \quad \text{as } \binom{M}{k} \leq (e \cdot M/k)^k \\
&\Longleftrightarrow A \leq \frac{k^{\frac{2}{k-2}}}{4 \cdot (2 \cdot e)^{\frac{k}{k-2}}} \cdot \frac{k}{N} \cdot \frac{1}{N^{\frac{1}{k-2}}} \\
&\Longleftarrow A \leq \frac{1}{90} \cdot \frac{k}{N} \cdot \frac{1}{N^{\frac{1}{k-2}}} \quad \text{as } \frac{k^{\frac{2}{k-2}}}{4 \cdot (2 \cdot e)^{\frac{k}{k-2}}} > 1/90.
\end{aligned} \tag{4}$$

For $k \geq \log N$, we have $N^{1/(k-2)} \leq 8$ for each integer $N \geq 8$. Then, the choice $A := (1/720) \cdot k/N$ satisfies (4) for every integer $k \geq \log N$. By removing from each edge $E \in \mathcal{E}_k$ one vertex we obtain a subset of at least N points in $[0,1]^2$ such that the area of the convex hull of each k points is at least A, i.e., for $k \geq \log N$ it is $\Delta_k(N) = \Omega(k/N)$. Concerning upper bounds on $\Delta_k(N)$, given any N points in $[0,1]^2$, we partition $[0,1]^2$ into $(N-1)/(k-1)$ squares each of side-lengths $\sqrt{(k-1)/(N-1)}$. Then, one of these little squares contains k of the N points, and the area of the convex hull of these k points certainly is at most $(k-1)/(N-1) = O(k/N)$, i.e., these arguments show:

Theorem 2.5. *For integers k, n with $3 \leq k \leq n$ it is*

$$\Delta_k(n) = \Omega\left(\frac{k}{n} \cdot \frac{1}{n^{\frac{1}{k-2}}}\right) \quad \text{and} \quad \Delta_k(n) = O\left(\frac{k}{n}\right).$$

In particular, for $\log n \leq k \leq n$ it is

$$\Delta_k(n) = \Theta\left(\frac{k}{n}\right).$$

To prove the existence of a large independent set in $\mathcal{G}$, we use the following result of Ajtai, Komlós, Pintz, Spencer and Szemerédi [1], see also [4,8,9].

Theorem 2.6. *Let $k \geq 3$ be a fixed integer. Let $\mathcal{G} = (V, \mathcal{E})$ be a k-uniform, linear hypergraph with average degree $t^{k-1} := k \cdot |\mathcal{E}|/|V|$. Then for some constant $C_k > 0$, the independence number $\alpha(\mathcal{G})$ of $\mathcal{G}$ satisfies*

$$\alpha(\mathcal{G}) \geq C_k \cdot \frac{|V|}{t} \cdot (\log t)^{\frac{1}{k-1}}. \tag{5}$$

We estimate in the random hypergraph $\mathcal{G}$ the expected numbers $E[|\mathcal{E}_2|]$ and $E[|\mathcal{E}_k|]$ of 2- and k-element edges, respectively, and $E[s_{2,j}(\mathcal{G})]$ of $(2, j)$-cycles arising from the k-element edges from $\mathcal{E}_k$, and we show that the numbers $E[|\mathcal{E}_2|]$ and $E[s_{2,j}(\mathcal{G})]$, $j = 2, \ldots, k-1$, are small compared to the number $|V| = N$ of vertices in $\mathcal{G}$. Then, by deleting some vertices from V we show the existence of a certain induced, linear k-uniform subhypergraph $\mathcal{G}^* = (V^*, \mathcal{E}_k^*)$ of the non-uniform hypergraph $\mathcal{G} = (V, \mathcal{E}_2 \cup \mathcal{E}_k)$, to which we apply Theorem 2.6.

2.1. *Upper bounds on the numbers of $(2, j)$-cycles*

In the following we use the condition that each k-element edge $E \in \mathcal{E}_k$ in the random hypergraph $\mathcal{G} = (V, \mathcal{E}_2 \cup \mathcal{E}_k)$ does not contain any 2-element edges $E \in \mathcal{E}_2$, i.e., each two distinct random points P_i and P_j, $1 \leq i < j \leq N$, which are vertices of an edge $E' \in \mathcal{E}_k$, have Euclidean distance bigger than D_0. We show next upper bounds on the expected numbers $E[s_{2,j}(\mathcal{G})]$ of $(2, j)$-cycles, $j = 2, \ldots, k-1$, in $\mathcal{G}$.

Lemma 2.7. *For $j = 2, \ldots, k-1$, there exist constants $c_{2,j} > 0$ such that for $D_0^2 \geq 2 \cdot A$ it is*

$$E[s_{2,j}(\mathcal{G})] \leq c_{2,j} \cdot A^{2k-j-2} \cdot N^{2k-j} \cdot (\log N)^3. \tag{6}$$

Proof. We prove an upper bound on the probability that $(2k - j)$ points, which are chosen uniformly at random and independently of each other in the unit square $[0,1]^2$, form two sets of k points, where the area of the convex hull of each is at most A, conditioned on the event that any two distinct of these $(2k - j)$ points have Euclidean distance bigger than $D_0 = N^{-\gamma}$, $\gamma > 0$.

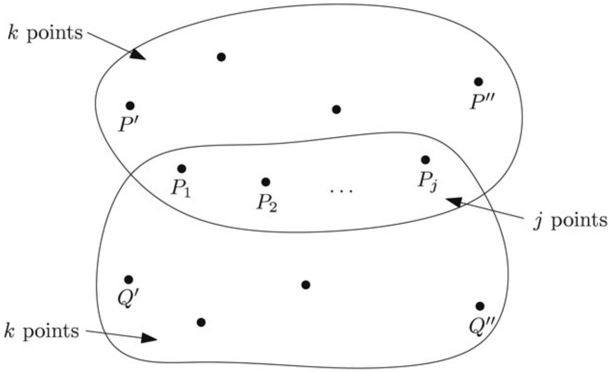

Fig. 1. Two sets of k points in $[0, 1]^2$, which have j points in common, and their extremal points P', P'' and Q', Q''.

There are $\binom{N}{2k-j}$ choices to select $(2k - j)$ out of N points. Given these $(2k - j)$ points, there are $\binom{2k-j}{j}$ possibilities to choose j points, say $P_1, \ldots, P_j$, which both k-gons have in common, and $\binom{2k-2j}{k-j}/2$ possibilities to extend $P_1, \ldots, P_j$ to two sets of k points. Let the two sets of k points be given by $P_1, \ldots, P_k$ and $P_1, \ldots, P_j, Q_{j+1}, \ldots, Q_k$ with $\text{area}(P_1, \ldots, P_k) \leq A$ and $\text{area}(P_1, \ldots, P_j, Q_{j+1}, \ldots, Q_k) \leq A$, where after renumbering $P_1 \leq_{\text{lex}} \ldots \leq_{\text{lex}} P_j$.

The point P_1 is somewhere in $[0, 1]^2$. Given $P_1 \in [0, 1]^2$, with $P_1 \leq_{\text{lex}} P_j$ we have

$$\text{Prob}(r \leq \text{dist}(P_1, P_j) \leq r + \text{d}r) \leq \pi \cdot r\text{d}r. \tag{7}$$

Given the points P_1 and P_j with $\text{dist}(P_1, P_j) = r$, by using $P_1 \leq_{\text{lex}} \cdots \leq_{\text{lex}} P_j$ and by Lemma 2.3 all points $P_2, \ldots, P_{j-1}$ are contained in a parallelogram of area $4 \cdot A$, which happens with probability

$$\text{Prob}(\text{area}(P_1, \ldots, P_j) \leq A \mid P_1, P_j) \leq (4 \cdot A)^{j-2}. \tag{8}$$

Given $P_1, \ldots, P_j \in [0, 1]^2$ with $P_1 \leq_{\text{lex}} \ldots \leq_{\text{lex}} P_j$ and $\text{dist}(P_1, P_j) = r$, with $\text{area}(P_1, \ldots, P_k) \leq A$ and $\text{area}(P_1, \ldots, P_j, Q_{j+1}, \ldots, Q_k) \leq A$ and by Lemma 2.2 all points $P_{j+1}, \ldots, P_k, Q_{j+1}, \ldots, Q_k$ are contained in a strip S centered at the line $P_1 P_j$ of width $w = 4 \cdot A/r$. Let $S^* := S \cap [0, 1]^2$ and observe that the area of S^* is at most $4 \cdot \sqrt{2} \cdot A/r$.

For the convex hulls of $P_1, \ldots, P_k$ and $P_1, \ldots, P_j, Q_{j+1}, \ldots, Q_k$ denote their *(lexicographically) extremal* points by P', P'' and Q', Q', respectively, that is, $P', P'' \in \{P_1, \ldots, P_k\}$ and $Q', Q'' \in \{P_1, \ldots, P_j, Q_{j+1}, \ldots, Q_k\}$ and, say $P' \leq_{\text{lex}} P''$ and $Q' \leq_{\text{lex}} Q''$, and $P' \leq_{\text{lex}} P_1, \ldots, P_k \leq_{\text{lex}} P''$ as well as $Q' \leq_{\text{lex}} P_1, \ldots, P_j, Q_{j+1}, \ldots, Q_k \leq_{\text{lex}} Q''$, see Fig. 1.

Given the points $P_1 \leq_{\text{lex}} \ldots \leq_{\text{lex}} P_j$ with $\text{dist}(P_1, P_j) = r$, for the convex hulls of $P_1, \ldots, P_k$ and $P_1, \ldots, P_j, Q_{j+1}, \ldots, Q_k$ we distinguish three cases each:

(i) both points, P_1 and P_j, are extremal, or
(ii) exactly one point, P_1 or P_j, is extremal, or
(iii) none of the points P_1, P_j is extremal.

Given the points $P_1, \ldots, P_j \in [0, 1]^2$ with $P_1 \leq_{\text{lex}} \ldots \leq_{\text{lex}} P_j$, first we consider the convex hull of $P_1, \ldots, P_k$.

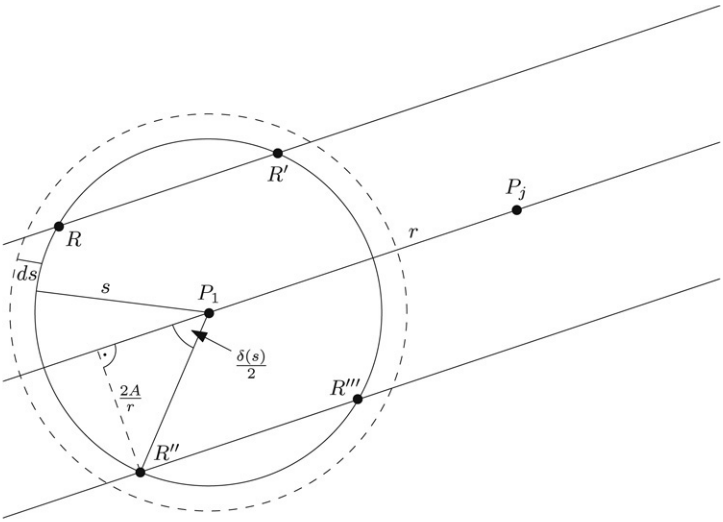

Fig. 2. The circle with radius s intersects the boundaries of the strip S in four points.

In case (i), the points P_1 and P_j are extremal for the convex hull of $P_1, \ldots, P_k$, hence $P_1 \leq_{\text{lex}} P_{j+1}, \ldots, P_k \leq_{\text{lex}} P_j$. By Lemma 2.3 all points $P_{j+1}, \ldots, P_k$ are contained in a parallelogram of area $4 \cdot A$, hence

$$\text{Prob}(\text{area}(P_1, \ldots, P_k) \leq A \mid P_1, \ldots, P_j \text{ and case (i)}) \leq (4 \cdot A)^{k-j}. \tag{9}$$

In case (ii), exactly one of the points P_1 or P_j is extremal for the convex hull of $P_1, \ldots, P_k$. By Lemma 2.2, the second extremal point, P' or P'', is contained in the set S^*, which happens with probability at most $4 \cdot \sqrt{2} \cdot A/r$. Given both extremal points P' and P'', by Lemma 2.3 all points $P_{j+1}, \ldots, P_k \neq P', P''$ are contained in a parallelogram of area $4 \cdot A$, which happens with probability at most $(4 \cdot A)^{k-j-1}$, hence

$$\begin{aligned} &\text{Prob}(\text{area}(P_1, \ldots, P_k) \leq A \mid P_1, \ldots, P_j \text{ and case (ii)}) \\ &\quad \leq \frac{4 \cdot \sqrt{2} \cdot A}{r} \cdot (4 \cdot A)^{k-j-1} = (4 \cdot A)^{k-j} \cdot \frac{\sqrt{2}}{r}. \end{aligned} \tag{10}$$

Next we consider case (iii), where neither point P_1 nor point P_j is extremal for the convex hull of $P_1, \ldots, P_k$. By Lemma 2.2, since $\text{area}(P_1, \ldots, P_k) \leq A$, both extremal points P' and P'', say $P' \leq_{\text{lex}} P_1 \leq_{\text{lex}} P_j \leq_{\text{lex}} P''$, must lie in the strip S centered at the line $P_1 P_j$ of width $4 \cdot A/r$. Since $P' \leq_{\text{lex}} P_1$, the probability that $\text{dist}(P_1, P') \in [s, s + \mathrm{d}s]$ is given by one-half of the difference of the areas of the balls with center P_1 and with radii $(s + \mathrm{d}s)$ and s, respectively, intersected with the strip S. Since we condition on the event that any two distinct points have Euclidean distance bigger than D_0, we have $r, s > D_0$. The circle with center P_1 and radius $s > D_0$ intersects both boundaries of the strip S of width $4 \cdot A/r$ in four points $R \leq_{\text{lex}} R'$ and $R'' \leq_{\text{lex}} R'''$, where $R, R'' \leq_{\text{lex}} P_1$, compare Fig. 2. To see this, we have to show that $s > 2 \cdot A/r$. Since $r, s > D_0$ it suffices to observe that $D_0 \geq 2 \cdot A/D_0$, which holds by assumption.

Let $\delta(s)$ be the angle between the lines $P_1 R$ and $P_1 R''$. Then one-half of the difference of the areas of the balls with center P_1 and with radii $(s + \mathrm{d}s)$ and s, respectively, intersected with the strip S is at most

$$\leq \frac{\delta(s)}{2 \cdot \pi} \cdot 2 \cdot \pi \cdot s\mathrm{d}s \leq 4 \cdot \sin(\delta(s)/2) \cdot s\mathrm{d}s \leq 4 \cdot \frac{2 \cdot A}{r \cdot s} \cdot s\mathrm{d}s = \frac{8 \cdot A}{r}\mathrm{d}s,$$

where we used the inequality $\delta/2 \leq \sin \delta$ for $\delta \leq 1$, since by assumption we have $\sin(\delta(s)/2) = 2 \cdot A/(r \cdot s) < 2 \cdot A/D_0^2 \leq 1$, and we infer by assuming that $P' \leq_{\text{lex}} P_1$ that

$$\text{Prob}(P' \in S \text{ and } \text{dist}(P_1, P') \in [s, s + \mathrm{d}s] \mid P_1) \leq \frac{8 \cdot A}{r}\mathrm{d}s. \tag{11}$$

Given the extremal point P' with $\text{dist}(P_1, P') = s$, the second extremal point P'' is contained in a strip centered at the line $P_1 P'$ of width $4 \cdot A/s$, which happens with probability at most

$$4 \cdot \sqrt{2} \cdot A/s. \tag{12}$$

Given both points P' and P'', by Lemma 2.3 all points $P_{j+1}, \ldots, P_k \neq P', P''$ are contained in a parallelogram of area $4 \cdot A$, which happens with probability at most

$$(4 \cdot A)^{k-j-2}. \tag{13}$$

With (11)–(13) and $s > D_0 = N^{-\gamma}$ for a constant $\gamma > 0$, we obtain

$$\begin{aligned} &\text{Prob}(\text{area}(P_1, \ldots, P_k) \leq A \mid P_1, \ldots, P_j \text{ and case (iii)}) \\ &\quad \leq (4 \cdot A)^{k-j-2} \cdot \int_{D_0}^{\sqrt{2}} \frac{4 \cdot \sqrt{2} \cdot A}{s} \cdot \frac{8 \cdot A}{r}\mathrm{d}s \\ &\quad = (4 \cdot A)^{k-j} \cdot \frac{2 \cdot \sqrt{2}}{r} \int_{D_0}^{\sqrt{2}} \frac{\mathrm{d}s}{s} \\ &\quad = 2 \cdot \sqrt{2} \cdot (4 \cdot A)^{k-j} \cdot \frac{\ln \sqrt{2} + \gamma \cdot \ln N}{r}. \end{aligned} \tag{14}$$

Summarizing (9), (10) and (14), we infer:

$$\begin{aligned} &\text{Prob}(\text{area}(P_1, \ldots, P_k) \leq A \mid P_1, \ldots, P_j) \\ &\quad \leq (4 \cdot A)^{k-j} \cdot \left(1 + \frac{\sqrt{2}}{r} + \sqrt{2} \cdot \frac{\ln 2 + 2 \cdot \gamma \cdot \ln N}{r}\right) \\ &\quad \leq (4 \cdot A)^{k-j} \cdot \left(\frac{2 \cdot \sqrt{2}}{r} + \frac{\sqrt{2} \cdot \ln 2}{r} + \frac{4 \cdot \sqrt{2} \cdot \gamma \cdot \ln N}{r}\right) \quad \text{as } r \leq \sqrt{2} \\ &\quad \leq (4 \cdot A)^{k-j} \cdot \left(\frac{10 \cdot \ln N}{r}\right) \quad \text{since } 0 < \gamma < 1. \end{aligned} \tag{15}$$

For the probability $\text{Prob}(\text{area}(P_1, \ldots, P_j, Q_{j+1}, \ldots, Q_k) \leq A \mid P_1, \ldots, P_j)$, the same upper bound as in (15) holds. Hence, for $j = 2, \ldots, k-1$, with (7), (8) and (15) we obtain for constants $c^*_{2,j} > 0$:

$$\begin{aligned} &\text{Prob}(P_1, \ldots, P_k \text{ and} P_1, \ldots, P_j, Q_{j+1}, \ldots, Q_k \text{ yield a } (2, j)\text{-cycle}) \\ &\quad \leq \int_{D_0}^{\sqrt{2}} (4 \cdot A)^{j-2} \cdot \left((4 \cdot A)^{k-j} \cdot \left(\frac{10 \cdot \ln N}{r}\right)\right)^2 \cdot \pi \cdot r\mathrm{d}r \end{aligned}$$

$$= 100 \cdot \pi \cdot 4^{2k-j-2} \cdot A^{2k-j-2} \cdot (\ln N)^2 \cdot \int_{D_0}^{\sqrt{2}} \frac{\mathrm{d}r}{r}$$
$$= 100 \cdot \pi \cdot 4^{2k-j-2} \cdot A^{2k-j-2} \cdot (\ln N)^2 \cdot (\ln \sqrt{2} - \ln D_0)$$
$$\leq c^*_{2,j} \cdot A^{2k-j-2} \cdot (\log N)^3 \quad \text{as } D_0 = N^{-\gamma}, \gamma > 0 \text{ a constant.} \tag{16}$$

Thus, for some constants $c^*_{2,j}, c_{2,j} > 0$, $j = 2, \ldots, k-1$, we obtain with (16) for the expected numbers $E[s_{2,j}(\mathcal{G})]$ of $(2, j)$-cycles in $\mathcal{G}$:

$$E[s_{2,j}(\mathcal{G})] \leq \binom{N}{2k-j} \cdot \binom{2k-j}{j} \cdot \binom{2k-2j}{k-j} \cdot c^*_{2,j} \cdot A^{2k-j-2} \cdot (\log N)^3$$
$$\leq c_{2,j} \cdot A^{2k-j-2} \cdot N^{2k-j} \cdot (\log N)^3,$$

which finishes the proof. □

2.2. *Choosing a subhypergraph*

Concerning edges $E \in \mathcal{E}_2$, for two points P, Q, which are chosen uniformly at random and independently of each other in $[0, 1]^2$, we have

$$\text{Prob}(\text{dist}(P, Q) \leq D_0) \leq \pi \cdot D_0^2,$$

since the point P can be anywhere in $[0, 1]^2$ and, if $\text{dist}(P, Q) \leq D_0$, the point Q is contained in the ball with center P and radius D_0. Thus, the expected number $E[|\mathcal{E}_2|]$ of unordered pairs of distinct points with Euclidean distance at most D_0 among the N random points $P_1, \ldots, P_N \in [0, 1]^2$ satisfies with $D_0 = N^{-\gamma}$ for some constant $c_2 > 0$:

$$E[|\mathcal{E}_2|] \leq \binom{N}{2} \cdot \pi \cdot D_0^2 \leq c_2 \cdot N^{2-2\gamma}. \tag{17}$$

By Markov's inequality, i.e., $\text{Prob}(X > k \cdot E[X]) < 1/k$ for every non-negative random variable X and any number $k \geq 1$, by using the estimates (3), (6) and (17) there exist N points $P_1, \ldots, P_N \in [0, 1]^2$ such that for $D_0^2 \geq 2 \cdot A$ the resulting hypergraph $\mathcal{G} = (V, \mathcal{E}_2 \cup \mathcal{E}_k)$ satisfies for $j = 2, \ldots, k-1$:

$$|V| = N \tag{18}$$
$$|\mathcal{E}_k| \leq k \cdot c_k \cdot A^{k-2} \cdot N^k \tag{19}$$
$$s_{2,j}(\mathcal{G}) \leq k \cdot c_{2,j} \cdot A^{2k-j-2} \cdot N^{2k-j} \cdot (\log N)^3 \tag{20}$$
$$|\mathcal{E}_2| \leq k \cdot c_2 \cdot N^{2-2\gamma}. \tag{21}$$

By (18) and (19), the *average degree* t^{k-1} for the k-element edges of $\mathcal{G}$ fulfills

$$t^{k-1} = \frac{k \cdot |\mathcal{E}_k|}{|V|} \leq \frac{k^2 \cdot c_k \cdot A^{k-2} \cdot N^k}{N} = k^2 \cdot c_k \cdot A^{k-2} \cdot N^{k-1} =: t_0^{k-1}.$$

For a suitable constant $c > 0$, which will be fixed later, we set

$$A := c \cdot \frac{(\log n)^{1/(k-2)}}{n^{(k-1)/(k-2)}}. \tag{22}$$

We show next that the numbers $|\mathcal{E}_2|$ and $s_{2,j}(\mathcal{G})$ of 2-element edges and $(2, j)$-cycles in $\mathcal{G}$, $j = 2, \ldots, k-1$, in $\mathcal{G}$, respectively, are very small compared to the number $|V|$ of vertices.

Lemma 2.8. *For every fixed* $\gamma > 1/2$ *it is*

$$|\mathcal{E}_2| = o(|V|). \tag{23}$$

Proof. By (18) and (21) we infer

$$\begin{aligned} &|\mathcal{E}_2| = o(|V|) \\ \Longleftarrow\ &N^{2-2\gamma} = o(N) \\ \Longleftrightarrow\ &N^{1-2\gamma} = o(1), \end{aligned}$$

which holds for fixed $\gamma > 1/2$. □

Lemma 2.9. *For* $D_0^2 \geq 2 \cdot A$ *and for* $j = 2, \ldots, k-1$, *and every fixed* β *with* $0 < \beta < (k-j)/((k-2)\cdot(2k-j-1))$ *it is*

$$s_{2,j}(\mathcal{G}) = o(|V|). \tag{24}$$

Proof. By (18), (20) and (22) and $N = n^{1+\beta}$ with fixed $\beta > 0$ we obtain for $j = 2, \ldots, k-1$:

$$\begin{aligned} &s_{2,j}(\mathcal{G}) = o(|V|) \\ \Longleftarrow\ &A^{2k-j-2} \cdot N^{2k-j} \cdot (\log N)^3 = o(N) \\ \Longleftrightarrow\ &A^{2k-j-2} \cdot N^{2k-j-1} \cdot (\log N)^3 = o(1) \\ \Longleftrightarrow\ &(\log n)^{3+\frac{2k-j-2}{k-2}} \cdot n^{(1+\beta)(2k-j-1)-\frac{(k-1)(2k-j-2)}{k-2}} = o(1) \\ \Longleftrightarrow\ &(1+\beta)\cdot(2k-j-1) < \frac{(k-1)\cdot(2k-j-2)}{k-2}, \end{aligned}$$

which holds for $\beta < (k-j)/((k-2)\cdot(2k-j-1))$. □

We fix $\beta := 1/k^2$ and $\gamma := k/(2\cdot(k-1))$. Then all assumptions in Lemmas 2.8 and 2.9 are fulfilled. Also the assumption $D_0^2 \geq 2 \cdot A$ in Lemma 2.7 is satisfied, namely, by choice of $\beta, \gamma > 0$ with $D_0 = N^{-\gamma}$ and $N = n^{1+\beta}$ and (22) we have

$$\begin{aligned} &D_0^2 \geq 2 \cdot A \\ \Longleftrightarrow\ &N^{-2\gamma} \geq 2 \cdot c \cdot \frac{(\log n)^{\frac{1}{k-2}}}{n^{\frac{k-1}{k-2}}} \\ \Longleftrightarrow\ &n^{\frac{k-1}{k-2}-2(1+\beta)\gamma} \geq 2 \cdot c \cdot (\log n)^{\frac{1}{k-2}} \\ \Longleftrightarrow\ &n^{\frac{2}{k(k-1)(k-2)}} \geq 2 \cdot c \cdot (\log n)^{\frac{1}{k-2}}. \end{aligned}$$

We delete from the hypergraph $\mathcal{G} = (V, \mathcal{E}_2 \cup \mathcal{E}_k)$ one vertex from each 2-element edge $E \in \mathcal{E}_2$ and from each $(2, j)$-cycle, $j = 2, \ldots, k-1$. Let $V^* \subseteq V$ be the set of all remaining vertices, where $|V^*| = (1-o(1))\cdot N \geq N/2$ by Lemmas 2.8 and 2.9. The resulting induced subhypergraph $\mathcal{G}^* = (V^*, \mathcal{E}_k^*)$ of $\mathcal{G}$ is k-uniform and does not contain any 2-cycles anymore, i.e., is linear, and by (19) satisfies $|V^*| \geq N/2$ and $|\mathcal{E}_k^*| \leq k \cdot c_k \cdot A^{k-2} \cdot N^k$, hence $\mathcal{G}^* = (V^*, \mathcal{E}_k^*)$ has average degree

$$t^{k-1} = k \cdot |\mathcal{E}_k^*|/|V^*| \leq 2 \cdot k^2 \cdot c_k \cdot A^{k-2} \cdot N^{k-1} =: t_1^{k-1}. \tag{25}$$

With (25) and $A = c \cdot (\log n)^{1/(k-2)}/n^{(k-1)/(k-2)}$ from (22), and $N = n^{1+\beta}$ for $\beta = 1/k^2$, and by Theorem 2.6 the independence number $\alpha(\mathcal{G}^*)$ of $\mathcal{G}^*$ satisfies for some sufficiently small constant $c > 0$ in (22) for some constants $C_k, C'_k > 0$:

$$\begin{aligned}\alpha(\mathcal{G}) &\ge \alpha(\mathcal{G}^*) \ge C_k \cdot \frac{|V^*|}{t} \cdot (\log t)^{\frac{1}{k-1}} \ge C_k \cdot \frac{|V^*|}{t_1} \cdot (\log t_1)^{\frac{1}{k-1}} \\ &\ge \frac{C_k \cdot N/2}{(2 \cdot k^2 \cdot c_k \cdot A^{k-2})^{\frac{1}{k-1}} \cdot N} \cdot \left(\log((2 \cdot k^2 \cdot c_k \cdot A^{k-2})^{\frac{1}{k-1}} \cdot N)\right)^{\frac{1}{k-1}} \\ &\ge \frac{C_k \cdot n}{2 \cdot (2 \cdot k^2 \cdot c_k)^{\frac{1}{k-1}} \cdot c^{\frac{k-2}{k-1}} \cdot (\log n)^{\frac{1}{k-1}}} \cdot \left(C'_k + \frac{(k-2) \cdot \log c}{k-1} + \frac{\log n}{k^2}\right)^{\frac{1}{k-1}} \\ &\ge n.\end{aligned}$$

The vertices of an independent set I of size $|I| = n$ yield a set $P(I) \subset [0, 1]^2$ of n points among the N points $P_1, \ldots, P_N \in [0, 1]^2$ such that the area of the convex hull of any k distinct points from $P(I)$ is $\Omega((\log n)^{1/(k-2)}/n^{(k-1)/(k-2)})$. □

3. A deterministic algorithm

Here we prove the algorithmic part of Theorem 1.1. To provide a deterministic polynomial time algorithm, which for fixed integer $k \ge 3$ and any integers $n \ge k$ finds n points in $[0, 1]^2$ that achieve the lower bound $\Delta_k(n) = \Omega((\log n)^{1/(k-2)}/n^{(k-1)/(k-2)})$, we discretize the unit square $[0, 1]^2$ by considering the standard $T \times T$-grid, where $T = n^{1+\beta}$ for some constant $\beta > 0$, which will be specified later. With this discretization we have to take care of collinear triples of grid-points in the $T \times T$-grid, as the area of the convex hull of k collinear grid-points is equal to zero.

To some extent, we proceed as in Section 2, but with some crucial differences due to the occurring collinear triples of grid-points.

Proof. For some number $A \ge 1$, which will be specified later, we form a hypergraph $\mathcal{G} = \mathcal{G}(A) = (V, \mathcal{E}_3^0 \cup \mathcal{E}_k)$, which contains 3-element and k-element edges. The vertex-set V consists of the T^2 grid-points from the $T \times T$-grid. The edge-sets $\mathcal{E}_3^0$ and $\mathcal{E}_k$ are defined as follows. For distinct grid-points $P, Q, R \in V$ in the $T \times T$-grid let $\{P, Q, R\} \in \mathcal{E}_3^0$ if and only if the grid-points P, Q, R are collinear. Moreover, for distinct grid-points $P_1, \ldots, P_k \in V$ in the $T \times T$-grid let $\{P_1, \ldots, P_k\} \in \mathcal{E}_k$ if and only if area$(P_1, \ldots, P_k) \le A$ and no three of the grid-points $P_1, \ldots, P_k$ are collinear. Notice, that for $k = 3$ there are two types of 3-element edges in the hypergraph $\mathcal{G}$, namely those edges describing collinear triples of points and those edges describing triples of points, which form triangles of area at most A. We are looking for a large independent set in this hypergraph $\mathcal{G} = (V, \mathcal{E}_3^0 \cup \mathcal{E}_k)$. An independent set $I \subseteq V$ corresponds to $|I|$ many grid-points in the $T \times T$-grid, such that the area of the convex hull of each k of these $|I|$ points is bigger than A.

We use the following algorithmic version of Theorem 2.6 of Bertram–Kretzberg and this author [4], compare also Fundia [9].

Theorem 3.1. *Let $k \ge 3$ be a fixed integer. Let $\mathcal{G} = (V, \mathcal{E})$ be a k-uniform linear hypergraph with average degree $t^{k-1} := k \cdot |\mathcal{E}|/|V|$. Then one can find for any $\delta > 0$ in time $O(|V| + |\mathcal{E}| + |V|^3/t^{3-\delta})$ an independent set $I \subseteq V$ with*

$$|I| = \Omega\left(\frac{|V|}{t} \cdot (\log t)^{1/(k-1)}\right).$$

The difficulty in our arguments is to find a suitable induced subhypergraph of $\mathcal{G} = (V, \mathcal{E}_3^0 \cup \mathcal{E}_k)$ to which Theorem 3.1 may be applied, and yields a solution with the desired quality. To do so, first we give upper bounds on the numbers $|\mathcal{E}_3^0|$ and $|\mathcal{E}_k|$ of 3- and k-element edges, respectively, and the numbers of 2-cycles arising from the k-element edges $E \in \mathcal{E}_k$ in the hypergraph $\mathcal{G}$. Then in a certain induced subhypergraph $\mathcal{G}^*$ of $\mathcal{G}$ we delete some vertices to destroy all 3-element edges from $\mathcal{E}_3^0$ and all 2-cycles. The resulting induced subhypergraph $\mathcal{G}^{**}$ is k-uniform and linear, and then we may apply to $\mathcal{G}^{**}$ the algorithm from Theorem 3.1.

For integers h and s let $\gcd(h, s) \geq 1$ denote the *greatest common divisor* of h and s. For distinct grid-points $P = (p_x, p_y)$ and $Q = (q_x, q_y)$ there are exactly $\gcd(q_x - p_x, q_y - p_y) - 1$ grid-points on the segment $[P, Q]$ excluding P and Q.

We use a *lexicographic order* $\leq_{\text{lex}}$ on the $T \times T$-grid: for grid-points $P = (p_x, p_y)$ and $Q = (q_x, q_y)$ let

$$P \leq_{\text{lex}} Q :\Longleftrightarrow (p_x < q_x) \text{ or } (p_x = q_x \text{ and } p_y < q_y).$$

Lemma 3.2. *The number $|\mathcal{E}_3^0|$ of 3-element edges in the hypergraph $\mathcal{G} = (V, \mathcal{E}_3^0 \cup \mathcal{E}_k)$ satisfies for some constant $c_3 > 0$:*

$$|\mathcal{E}_3^0| \leq c_3 \cdot T^4 \cdot \log T. \tag{26}$$

We remark that in [5] an upper bound of $O(T^{4+\varepsilon})$, for any $\varepsilon > 0$, on the number of collinear triples of grid-points in the $T \times T$-grid has been proved.

Proof. For distinct grid-points $P, Q, R \in V$ we have $\{P, Q, R\} \in \mathcal{E}_3^0$ if and only if P, Q, R are collinear. Let $P \leq_{\text{lex}} Q \leq_{\text{lex}} R$ with $P = (p_x, p_y)$ and $R = (r_x, r_y)$. If $p_x = r_x$ or $p_y = r_y$, the number of unordered collinear triples P, Q, R of grid-points is less than T^4, since there are $2 \cdot T$ horizontal or vertical lines and on each of these lines we can choose $\binom{T}{3}$ unordered triples of grid-points, which yield $2 \cdot T \cdot \binom{T}{3} < T^4$ unordered collinear triples in the $T \times T$-grid.

Let $h, s \neq 0$. A grid-point P may be chosen in at most T^2 ways. Given P, any other grid-point R with $P \leq_{\text{lex}} R$ in the $T \times T$-grid is determined by a pair (s, h) of integers with $s := r_x - p_x > 0$ and $h := r_y - p_y$. Without loss of generality let $1 \leq h, s \leq T$, as those pairs (s, h) of integers with $1 \leq s \leq T$ and $-T \leq h \leq -1$ are taken into account by an additional constant factor of 2.

Having fixed the grid-points P and R, on the segment $[P, R]$ there are less than $\gcd(h, s)$ grid-points Q excluding P and R, hence with $P \leq_{\text{lex}} Q \leq_{\text{lex}} R$ there are less than $\gcd(h, s)$ choices for the grid-point Q. Thus, the number of unordered collinear triples in the $T \times T$-grid is bounded from above as follows:

$$|\mathcal{E}_3^0| \leq T^4 + 2 \cdot T^2 \cdot \sum_{s=1}^{T} \sum_{h=1}^{T} \gcd(h, s).$$

Each divisor $d \in \{1, \ldots, T\}$ divides at most T/d integers from the set $\{1, \ldots, T\}$, hence, with $\sum_{d=1}^{T} 1/d \leq 1 + \int_1^T (1/x) \mathrm{d}x \leq 1 + \ln T$, we infer for a constant $c_3 > 0$:

$$\begin{aligned} |\mathcal{E}_3^0| &\leq T^4 + 2 \cdot T^2 \cdot \sum_{s=1}^{T} \sum_{h=1}^{T} \gcd(h, s) \leq T^4 + 2 \cdot T^2 \cdot \sum_{d=1}^{T} d \cdot \left(\frac{T}{d}\right)^2 \\ &\leq c_3 \cdot T^4 \cdot \log T, \end{aligned}$$

as was claimed. □

The next result from [5] is the discrete analogue of Lemmas 2.2 and 2.3 for the $T \times T$-grid.

Lemma 3.3. *For distinct grid-points $P = (p_x, p_y)$ and $R = (r_x, r_y)$ with $P \leq_{\text{lex}} R$ from the $T \times T$-grid, where $s := r_x - p_x \geq 0$ and $h := r_y - p_y$, the following hold:*

(a) *There are at most $4 \cdot A$ grid-points Q in the $T \times T$-grid such that*
  (i) *$P \leq_{\text{lex}} Q \leq_{\text{lex}} R$, and*
  (ii) *P, Q, R are not collinear, and* $\text{area}(P, Q, R) \leq A$.

(b) *The number of grid-points Q in the $T \times T$-grid which fulfills only* (ii) *from* (a) *is at most $12 \cdot A \cdot T/s$ for $s > 0$, and at most $12 \cdot A \cdot T/|h|$ for $|h| > s$.*

Lemma 3.4. *For fixed integers $k \geq 3$, the number $|\mathcal{E}_k|$ of unordered k-tuples $P_1, \ldots, P_k$ of distinct grid-points in the $T \times T$-grid with area $(P_1, \ldots, P_k) \leq A$, where no three of $P_1, \ldots, P_k$ are collinear, satisfies for some constant $c_k > 0$:*

$$|\mathcal{E}_k| \leq c_k \cdot A^{k-2} \cdot T^4. \tag{27}$$

Proof. Let $P_1, \ldots, P_k$ be distinct grid-points, no three on a line, in the $T \times T$-grid with $\text{area}(P_1, \ldots, P_k) \leq A$. We may assume after renumbering that $P_1 \leq_{\text{lex}} \ldots \leq_{\text{lex}} P_k$. For $P_1 = (p_{1,x}, p_{1,y})$ and $P_k = (p_{k,x}, p_{k,y})$ let $s := p_{k,x} - p_{1,x} \geq 0$ and $h := p_{k,y} - p_{1,y}$. If $s = 0$, then for $k \geq 3$ the grid-points $P_1, \ldots, P_k$ are collinear, hence we have $s > 0$.

There are T^2 choices for the grid-point P_1. Given P_1, any other grid-point P_k with $P_1 \leq_{\text{lex}} P_k$ is determined by a pair (s, h) of integers with $1 \leq s \leq T$ and $-T \leq h \leq T$. With $\text{area}(P_1, \ldots, P_k) \leq A$, by Lemma 2.1 it is $\text{area}(P_1, P_j, P_k) \leq A$ for $j = 2, \ldots, k-1$. Then, given the grid-points P_1 and P_k, since $P_1 \leq_{\text{lex}} P_j \leq_{\text{lex}} P_k$, $j = 2, \ldots, k-1$, and no three of the grid-points $P_1, \ldots, P_k$ are collinear, by Lemma 3.3(a) there are at most $4 \cdot A$ choices for each grid-point P_j, hence $(4 \cdot A)^{k-2}$ choices for the grid-points $P_2, \ldots, P_{k-1}$ altogether, thus for a constant $c_k > 0$:

$$|\mathcal{E}_k| \leq T^2 \cdot \sum_{s=1}^{T} \sum_{h=-T}^{T} (4 \cdot A)^{k-2} \leq c_k \cdot A^{k-2} \cdot T^4,$$

as desired. □

By (27) we infer that the average degree t^{k-1} of the hypergraph $\mathcal{G} = (V, \mathcal{E}_3^0 \cup \mathcal{E}_k)$ for the k-element edges $E \in \mathcal{E}_k$ satisfies

$$t^{k-1} = \frac{k \cdot |\mathcal{E}_k|}{|V|} \leq \frac{k \cdot c_k \cdot A^{k-2} \cdot T^4}{T^2} = k \cdot c_k \cdot A^{k-2} \cdot T^2 =: t_0^{k-1}. \tag{28}$$

3.1. Upper bounds on the number of $(2, j)$-cycles

Let $s_{2,j}(\mathcal{G})$ denote the number of $(2, j)$-cycles, $j = 2, \ldots, k-1$, which arise from the k-element edges $E \in \mathcal{E}_k$ in the hypergraph $\mathcal{G} = (V, \mathcal{E}_3^0 \cup \mathcal{E}_k)$.

Lemma 3.5. *Let $k \geq 3$ be a fixed integer. For $j = 2, \ldots, k-1$, there exist constants $c_{2,j} > 0$ such that the numbers $s_{2,j}(\mathcal{G})$ of $(2, j)$-cycles in the hypergraph $\mathcal{G} = (V, \mathcal{E}_3^0 \cup \mathcal{E}_k)$ fulfill*

$$s_{2,j}(\mathcal{G}) \leq c_{2,j} \cdot A^{2k-j-2} \cdot T^4 \cdot (\log T)^3. \tag{29}$$

Proof. For $j = 2, \ldots, k-1$, let us denote the grid-points corresponding to the vertices of two distinct k-element edges $E, E' \in \mathcal{E}_k$, which yield a $(2, j)$-cycle in $\mathcal{G}$, i.e., $|E \cap E'| = j$, by $P_1, \ldots, P_k$ and $P_1, \ldots, P_j, Q_{j+1}, \ldots, Q_k$, where $P_1 \leq_{\text{lex}} \ldots \leq_{\text{lex}} P_j$ and no three points of $P_1, \ldots, P_k$ and of $P_1, \ldots, P_j, Q_{j+1}, \ldots, Q_k$ are collinear. By assumption we have

$$\text{area}(P_1, \ldots, P_k) \leq A \quad \text{and} \quad \text{area}(P_1, \ldots, P_j, Q_{j+1}, \ldots, Q_k) \leq A. \tag{30}$$

There are T^2 choices for the grid-point P_1. Given $P_1 = (p_{1,x}, p_{1,y})$, any other grid-point $P_j = (p_{j,x}, p_{j,y})$ with $P_1 \leq_{\text{lex}} P_j$ is determined by a pair $(s, h) \neq (0, 0)$ of integers with $s = p_{j,x} - p_{1,x} \geq 0$ and $h = p_{j,y} - p_{1,y}$. By symmetry we may assume that $s > 0$ and $0 \leq h \leq s \leq T$, otherwise we interchange the role of h and s. This is taken into account by the additional constant factor $c' = 2$. Given P_1 and P_j, by Lemma 2.1 and (30) we have area $(P_1, P_i, P_j) \leq A$ for $i = 2, \ldots, j-1$ and since $P_1 \leq_{\text{lex}} P_i \leq_{\text{lex}} P_j$, where P_1, P_i, P_j are not on a line, by Lemma 3.3(a) there are at most $4 \cdot A$ choices for each grid-point P_i, hence there are at most $(4 \cdot A)^{j-2}$ choices for the grid-points $P_2, \ldots, P_{j-1}$. Thus, for fixed h, s the number of choices for the grid-points $P_1, \ldots P_j$ is at most

$$c' \cdot (4 \cdot A)^{j-2} \cdot T^2. \tag{31}$$

As in the arguments in Section 2, for the convex hulls of the points $P_1, \ldots, P_k$ and $P_1, \ldots, P_j, Q_{j+1}, \ldots, Q_k$ we denote their *(lexicographically) extremal* points by $P', P'' \in \{P_1, \ldots, P_k\}$ and $Q', Q'' \in \{P_1, \ldots, P_j, Q_{j+1}, \ldots, Q_k\}$, respectively, i.e., if say $P' \leq_{\text{lex}} P''$ and $Q' \leq_{\text{lex}} Q''$, then we have $P' \leq_{\text{lex}} P_1, \ldots, P_k \leq_{\text{lex}} P''$ and $Q' \leq_{\text{lex}} P_1, \ldots, P_j, Q_{j+1}, \ldots, Q_k \leq_{\text{lex}} Q''$, compare Fig. 1.

Given the points $P_1 \leq_{\text{lex}} \ldots \leq_{\text{lex}} P_j$, there are three possibilities for each of the convex hulls of $P_1, \ldots, P_k$ and $P_1, \ldots, P_j, Q_{j+1}, \ldots, Q_k$:

(i) both points, P_1 and P_j, are extremal, or
(ii) exactly one point, P_1 or P_j, is extremal, or
(iii) none of the points P_1, P_j is extremal.

We consider the convex hull of $P_1, \ldots, P_k$ as the considerations for the convex hull of $P_1, \ldots, P_j, Q_{j+1}, \ldots, Q_k$ are similar.

In case (i) the grid-points P_1 and P_j are extremal for the convex hull of $P_1, \ldots, P_k$, hence we have $P_1 \leq_{\text{lex}} P_{j+1}, \ldots, P_k \leq_{\text{lex}} P_j$. By (30) and Lemma 3.3(a), since area$(P_1, P_l, P_j) \leq A$, $l = j+1, \ldots, k$, and no three of $P_1, \ldots, P_k$ are collinear, there are at most $4 \cdot A$ choices for each grid-point P_l, hence in case (i), given $P_1, \ldots, P_j$, the number of choices for the grid-points $P_{j+1}, \ldots, P_k$ is at most

$$(4 \cdot A)^{k-j}. \tag{32}$$

In case (ii) exactly one of the grid-points P_1 or P_j is extremal for the convex hull of $P_1, \ldots, P_k$. By Lemma 3.3(b) there are at most $12 \cdot A \cdot T/s$ choices for the second extremal grid-point P' or P''. Having fixed this second extremal grid-point, for each grid-point $P_{j+1}, \ldots, P_k \neq P', P''$ there are by Lemma 3.3(a) at most $4 \cdot A$ choices, hence in case (ii), given $P_1, \ldots, P_j$, the number of choices for the grid-points $P_{j+1}, \ldots, P_k$ is at most

$$(4 \cdot A)^{k-j-1} \cdot \frac{12 \cdot A \cdot T}{s} = (4 \cdot A)^{k-j} \cdot \frac{3 \cdot T}{s}. \tag{33}$$

In case (iii) none of the grid-points P_1, P_j is extremal for the convex hull of $P_1, \ldots, P_k$. Given the grid-points $P_1, \ldots, P_j$, by (30) and Lemma 2.2 all grid-points $P_{j+1}, \ldots, P_k$ and

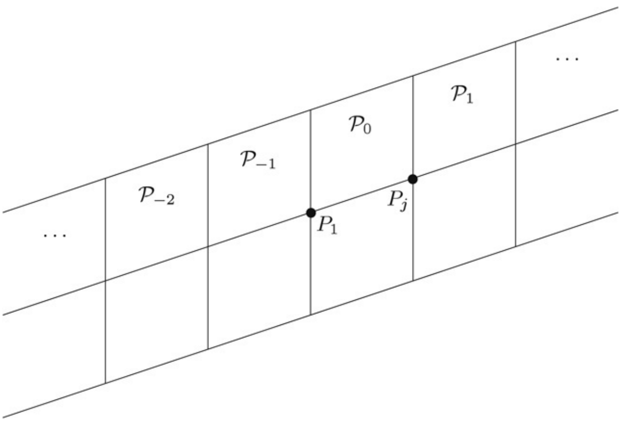

Fig. 3. The parallelograms $\mathcal{P}_{-i}, \mathcal{P}_i$, $i = 1, 2, \ldots,$ are copies of $\mathcal{P}_0$.

$Q_{j+1}, \ldots, Q_k$ are contained in a strip S centered at the line $P_1 P_j$ of width $4 \cdot A/\sqrt{h^2+s^2}$. Consider the parallelogram $\mathcal{P}_0 = \{(p_x, p_y) \in S \mid p_{1,x} \le p_x \le p_{j,x}\}$ within the strip S of width $4 \cdot A/\sqrt{h^2+s^2}$, where $P_1 = (p_{1,x}, p_{1,y})$ and $P_j = (p_{j,x}, p_{j,y})$ and $s = p_{j,x} - p_{1,x} \ge 0$. By Lemma 3.3(a) this parallelogram $\mathcal{P}_0$ contains at most $4 \cdot A$ grid-points P, such that P_1, P_j, P are not collinear. We divide the strip S within the $T \times T$-grid into congruent parallelograms $\mathcal{P}_0, \mathcal{P}_g$, $g = -l, -l+1, \ldots, m$ with $1 \le l, m \le \lfloor T/s \rfloor + 2$, each of side-lengths $4 \cdot A/s$ and $\sqrt{h^2+s^2}$ and area $4 \cdot A$, where all parallelograms $\mathcal{P}_{-g}$, $g \ge 1$, are on the left of the parallelogram $\mathcal{P}_0$, and all parallelograms $\mathcal{P}_h$, $h \ge 1$, are on the right of $\mathcal{P}_0$, i.e., $\mathcal{P}_{-g} := \{(p_x, p_y) \in S \mid p_{1,x} - g \cdot s \le p_x \le p_{1,x} - (g-1) \cdot s\}$ and $\mathcal{P}_h := \{(p_x, p_y) \in S \mid p_{j,x} + (h-1) \cdot s \le p_x \le p_{j,x} + h \cdot s\}$, compare Fig. 3. Each grid-point $P = (p_x, p_y) \in \mathcal{P}_{-g} \cup \mathcal{P}_g$, $g \ge 1$, satisfies $|p_x - p_{1,x}| \ge g \cdot s$ or $|p_x - p_{j,x}| \ge g \cdot s$. By Lemma 3.3(a) each parallelogram $\mathcal{P}_{-g}$ or $\mathcal{P}_g$, $g \ge 1$, contains at most $4 \cdot A$ grid-points P, such that P_1, P_j, P are not collinear. Each for the convex hull of $P_1, \ldots, P_k$ extremal grid-point is contained in a parallelogram $\mathcal{P}_{-g}$ or $\mathcal{P}_g$, since by our assumption neither $P_1 \in \mathcal{P}_0$ nor $P_j \in \mathcal{P}_0$ are extremal. If $P' \in \mathcal{P}_{-g} \cup \mathcal{P}_g$ or $P'' \in \mathcal{P}_{-g} \cup \mathcal{P}_g$, $g \ge 1$, then by Lemma 3.3(b) there are at most $12 \cdot A \cdot T/(g \cdot s)$ choices for the second extremal grid-point. Having chosen both extremal grid-points P' and P'', for the other grid-points $P_{j+1}, \ldots, P_k \ne P', P''$, by (30) and Lemma 3.3(a) there are at most $(4 \cdot A)^{k-j-2}$ choices.

Hence, with $\sum_{i=1}^{l} 1/i \le 1 + \ln l$ we obtain in case (iii), given $P_1, \ldots, P_j$, the following upper bound on the number of choices for $P_{j+1}, \ldots, P_k$:

$$2 \cdot (4 \cdot A)^{k-j-2} \cdot \sum_{g=1}^{\lfloor T/s \rfloor + 2} 4 \cdot A \cdot \frac{12 \cdot A \cdot T}{g \cdot s} = (4 \cdot A)^{k-j} \cdot \frac{6 \cdot T}{s} \cdot \sum_{g=1}^{\lfloor T/s \rfloor + 2} \frac{1}{g}$$

$$\le (4 \cdot A)^{k-j} \cdot \frac{12 \cdot T}{s} \cdot \ln T. \tag{34}$$

By (32)–(34) using $s \le T$, altogether the number of choices for the grid-points $P_{j+1}, \ldots, P_k$ is at most

$$(4 \cdot A)^{k-j} \cdot \left(1 + \frac{3 \cdot T}{s} + \frac{12 \cdot T \cdot \ln T}{s}\right) \le (4 \cdot A)^{k-j} \cdot \frac{16 \cdot T \cdot \ln T}{s}. \tag{35}$$

Given the grid-points $P_1, \ldots, P_j$, the same upper bound (35) holds for the number of choices of the grid-points $Q_{j+1}, \ldots, Q_k$. With (31) and (35), for $j = 2, \ldots, k-1$, we obtain for some constants $c', c_{2,j} > 0$:

$$s_{2,j}(\mathcal{G}) \leq c' \cdot (4 \cdot A)^{j-2} \cdot T^2 \cdot \sum_{s=1}^{T} \sum_{h=0}^{s} \left((4 \cdot A)^{k-j} \cdot \frac{16 \cdot T \cdot \ln T}{s} \right)^2$$
$$\leq c' \cdot 4^{2k-j+2} \cdot A^{2k-j-2} \cdot T^4 \cdot (\ln T)^2 \cdot \sum_{s=1}^{T} \sum_{h=0}^{s} \frac{1}{s^2}$$
$$\leq c_{2,j} \cdot A^{2k-j-2} \cdot T^4 \cdot (\log T)^3, \tag{36}$$

which finishes the proof of the lemma. □

3.2. Selecting a subhypergraph

For a suitable constant $c > 0$ we set

$$A := \frac{c \cdot T^2 \cdot (\log n)^{1/(k-2)}}{n^{(k-1)/(k-2)}}. \tag{37}$$

Towards our estimate of the running times we observe that $A \geq 1$ for n large enough. For the moment we use a probabilistic argument, which will be derandomized shortly. With probability $p := T^{\varepsilon}/t_0$, thus $p = O(T^{\varepsilon}/(A^{(k-2)/(k-1)} \cdot T^{2/(k-1)})) = o(1)$ by (28), where $\varepsilon > 0$ is a small constant, we pick uniformly at random and independently of each other vertices from the vertex-set V. Let $V^* \subseteq V$ be the resulting random subset of the picked vertices and let $\mathcal{G}^* = (V^*, \mathcal{E}_3^{0*} \cup \mathcal{E}_k^*)$ with $\mathcal{E}_3^{0*} := \mathcal{E}_3^0 \cap [V^*]^3$ and $\mathcal{E}_k^* := \mathcal{E}_k \cap [V^*]^k$ be the resulting random induced subhypergraph of $\mathcal{G}$. Let $E[|V^*|]$, $E[|\mathcal{E}_3^{0*}|]$, $E[|\mathcal{E}_k^*|]$, $E[s_{2,j}(\mathcal{G}^*)]$ denote the expected numbers of vertices, 3-element edges, k-element edges, and $(2, j)$-cycles arising from the k-element edges $E \in \mathcal{E}_k^*$, respectively, in $\mathcal{G}^* = (V^*, \mathcal{E}_3^{0*} \cup \mathcal{E}_k^*)$. By (26), (27) and (29) we infer for $j = 2, \ldots, k-1$ and constants $c_1', c_3', c_k', c_{2,j}' > 0$:

$$E[|V^*|] = p \cdot T^2 \geq c_1' \cdot T^2 \cdot T^{\varepsilon}/(A^{\frac{k-2}{k-1}} \cdot T^{\frac{2}{k-1}}) = c_1' \cdot T^{\frac{2k-4}{k-1}+\varepsilon}/A^{\frac{k-2}{k-1}} \tag{38}$$

$$E[|\mathcal{E}_3^{0*}|] = p^3 \cdot |\mathcal{E}_3^0| \leq c_3' \cdot (T^4 \cdot \log T) \cdot T^{3\varepsilon}/(A^{\frac{k-2}{k-1}} \cdot T^{\frac{2}{k-1}})^3$$
$$\leq c_3' \cdot T^{\frac{4k-10}{k-1}+3\varepsilon} \cdot \log T/A^{\frac{3k-6}{k-1}} \tag{39}$$

$$E[|\mathcal{E}_k^*|] = p^k \cdot |\mathcal{E}_k| \leq c_k' \cdot (A^{k-2} \cdot T^4) \cdot T^{k\varepsilon}/(A^{\frac{k-2}{k-1}} \cdot T^{\frac{2}{k-1}})^k$$
$$\leq c_k' \cdot T^{\frac{2k-4}{k-1}+k\varepsilon}/A^{\frac{k-2}{k-1}} \tag{40}$$

$$E[s_{2,j}(\mathcal{G}^*)] = p^{2k-j} \cdot s_{2,j}(\mathcal{G})$$
$$\leq c_{2,j}' \cdot (A^{2k-j-2} \cdot T^4 \cdot (\log T)^3) \cdot T^{(2k-j)\varepsilon}/(A^{\frac{k-2}{k-1}} \cdot T^{\frac{2}{k-1}})^{2k-j}$$
$$\leq c_{2,j}' \cdot T^{\frac{2j-4}{k-1}+(2k-j)\varepsilon} \cdot (\log T)^3/A^{\frac{j-2}{k-1}}. \tag{41}$$

By Chernoff's inequality, for n binomially distributed random variables X_i, $i = 1, \ldots, n$, with values in $\{0, 1\}$ and with sum $X := X_1 + \cdots + X_n$ having expected value $E[X]$, it is $\text{Prob}(E[X] - X \geq u) \leq e^{-u^2/n}$. With this, (38)–(41), and Markov's inequality we infer that there exists an induced subhypergraph $\mathcal{G}^* = (V^*, \mathcal{E}_3^{0*} \cup \mathcal{E}_k^*)$ of $\mathcal{G}$ such that

$$|V^*| \geq (c_1'/2) \cdot T^{\frac{2k-4}{k-1}+\varepsilon} / A^{\frac{k-2}{k-1}} \tag{42}$$

$$|\mathcal{E}_3^{0*}| \leq (k+1) \cdot c_3' \cdot T^{\frac{4k-10}{k-1}+3\varepsilon} \cdot \log T / A^{\frac{3k-6}{k-1}} \tag{43}$$

$$|\mathcal{E}_k^*| \leq (k+1) \cdot c_k' \cdot T^{\frac{2k-4}{k-1}+k\varepsilon} / A^{\frac{k-2}{k-1}} \tag{44}$$

$$s_{2,j}(\mathcal{G}^*) \leq (k+1) \cdot c_{2,j}' \cdot T^{\frac{2j-4}{k-1}+(2k-j)\varepsilon} \cdot (\log T)^3 / A^{\frac{j-2}{k-1}}. \tag{45}$$

This probabilistic argument can be turned into a deterministic polynomial time algorithm by using the method of conditional probabilities. Namely, for $j = 2, \ldots, k-1$, let $\mathcal{C}_j$ be the (multi-)set of all $(2k-j)$-element subsets $E \cup E'$ of V such that the pair $\{E, E'\}$ of k-element edges $E, E' \in \mathcal{E}_k$ yields a $(2, j)$-cycle in $\mathcal{G}$, i.e., $|E \cap E'| = j$. We enumerate the vertices of the $T \times T$-grid as $P_1, \ldots, P_{T^2}$. To each vertex P_i we associate a parameter $p_i \in [0, 1]$, $i = 1, \ldots, T^2$, and we define a potential function $F(p_1, \ldots, p_{T^2})$ by

$$F(p_1, \ldots, p_{T^2}) := 2^{p \cdot T^2/2} \cdot \prod_{i=1}^{T^2} \left(1 - \frac{p_i}{2}\right) + \frac{\sum_{\{i,j,k\} \in \mathcal{E}_3^0} p_i \cdot p_j \cdot p_k}{(k+1) \cdot c_3' \cdot T^{\frac{4k-10}{k-1}+3\varepsilon} \cdot \log T / A^{\frac{3k-6}{k-1}}}$$
$$+ \frac{\sum_{\{i_1,\ldots,i_k\} \in \mathcal{E}_k} \prod_{l=1}^{k} p_{i_l}}{(k+1) \cdot c_k' \cdot T^{\frac{2k-4}{k-1}+k\varepsilon} / A^{\frac{k-2}{k-1}}}$$
$$+ \sum_{j=2}^{k-1} \frac{\sum_{\{i_1,\ldots,i_{2k-j}\} \in \mathcal{C}_j} \prod_{l=1}^{2k-j} p_{i_l}}{(k+1) \cdot c_{2,j}' \cdot T^{\frac{2j-4}{k-1}+(2k-j)\varepsilon} \cdot (\log T)^3 / A^{\frac{j-2}{k-1}}}.$$

We initialize $p_1 := \cdots := p_{T^2} := p := T^\varepsilon / t_0$. Using $1 - x \leq e^{-x}$, with (39)–(41) we infer $F(p, \ldots, p) < (2/e)^{pT^2/2} + k/(k+1)$. Hence, in the beginning we have $F(p, \ldots, p) < 1$, if $p \cdot T^2 \geq 7 \cdot \ln(k+1)$. This is fulfilled since $p = T^\varepsilon / t_0 \geq (T^\varepsilon \cdot n)/T^2$ by (28) and (37), and $\varepsilon < 1$, and $T = n^{1+\beta}$ with $\beta > 0$. Using the linearity of $F(p_1, \ldots, p_{T^2})$ in each p_i, we minimize $F(p_1, \ldots, p_{T^2})$ step by step by fixing one after the other $p_i := 0$ or $p_i := 1$ for $i = 1, \ldots, T^2$. Finally, we obtain $F(p_1, \ldots, p_{T^2}) \leq F(p, \ldots, p) < 1$. With $V^* = \{P_i \in V \mid p_i = 1\}$ this yields an induced subhypergraph $\mathcal{G}^* = (V^*, \mathcal{E}_3^{0*} \cup \mathcal{E}_k^*)$ of $\mathcal{G}$ with $\mathcal{E}_3^{0*} := \mathcal{E}_3^0 \cap [V^*]^3$ and $\mathcal{E}_k^* := \mathcal{E}_k \cap [V^*]^k$.

We now have $|V^*| \geq p \cdot T^2/2$, as otherwise $F(p_1, \ldots, p_{T^2}) \geq 2^{pT^2/2} \cdot \prod_{i=1}^{T^2} (1 - p_i/2) > 2^{pT^2/2} \cdot (1/2)^{pT^2/2} = 1$, which is a contradiction. Moreover, it is $|\mathcal{E}_3^{0*}| \leq (k+1) \cdot c_3' \cdot T^{(4k-10)/(k-1)+3\varepsilon} \cdot \log T / A^{(3k-6)/(k-1)}$, else we have $F(p_1, \ldots, p_n) > 1$, a contradiction. Similarly we infer $|\mathcal{E}_k^*| \leq (k+1) \cdot c_k' \cdot T^{(2k-4)/(k-1)+k\varepsilon} / A^{(k-2)/(k-1)}$ and $s_{2,j}(\mathcal{G}^*) \leq (k+1) \cdot c_{2,j}' \cdot T^{(2j-4)/(k-1)+(2k-j)\varepsilon} \cdot (\log T)^3 / A^{(j-2)/(k-1)}$.

Hence the induced subhypergraph $\mathcal{G}^* = (V^*, \mathcal{E}_3^{0*} \cup \mathcal{E}_k^*)$ satisfies (42)–(45). When fixing $p_i := 0$ or $p_i := 1$, $i = 1, \ldots, T^2$, during the algorithm, we consider only those edges and 2-cycles, which are incident to vertex P_i, hence for fixed $k \geq 3$ the running time is linear in $(|V| + |\mathcal{E}_3^0| + |\mathcal{E}_k| + \sum_{j=2}^{k-1} |\mathcal{C}_j|)$. By (26), (27), (29) and (37), and since $T = n^{1+\beta}$ with $\beta > 0$, the time for this derandomization is

$$O\left(\left(|V| + |\mathcal{E}_3^0| + |\mathcal{E}_k| + \sum_{j=2}^{k-1} |\mathcal{C}_j|\right)\right) = O(|\mathcal{C}_2|) = O(A^{2k-4} \cdot T^4 \cdot (\log T)^3)$$

$$= O\left(\frac{T^{4k-4} \cdot (\log n)^5}{n^{2k-2}}\right) = O\left(n^{2k-2+4\beta(k-1)} \cdot (\log n)^5\right). \quad (46)$$

We show next that, for a certain choice of the parameters $\beta, \varepsilon > 0$, the numbers $|\mathcal{E}_3^{0*}|$ and $s_{2,j}(\mathcal{G}^*)$ of 3-element edges and of $(2, j)$-cycles, $j = 2, \ldots, k-1$, in $\mathcal{G}^*$, respectively, are small in comparison to the number $|V^*|$ of vertices in $\mathcal{G}^*$.

Lemma 3.6. *For every fixed* ε *with* $0 < \varepsilon < \beta/(1+\beta)$ *it is*

$$|\mathcal{E}_3^{0*}| = o(|V^*|). \quad (47)$$

Proof. By (37), (42) and (43), and using $T = n^{1+\beta}$ for fixed $\beta, \varepsilon > 0$, we have

$$\begin{aligned}
&|\mathcal{E}_3^{0*}| = o(|V^*|)\\
&\Longleftarrow T^{\frac{4k-10}{k-1}+3\varepsilon} \cdot \log T / A^{\frac{3k-6}{k-1}} = o(T^{\frac{2k-4}{k-1}+\varepsilon} / A^{\frac{k-2}{k-1}})\\
&\Longleftrightarrow T^{\frac{2k-6}{k-1}+2\varepsilon} \cdot \log T / A^{\frac{2k-4}{k-1}} = o(1)\\
&\Longleftrightarrow n^{2-(1+\beta)(2-2\varepsilon)} \cdot (\log n)^{\frac{k-3}{k-1}} = o(1)\\
&\Longleftrightarrow (1+\beta) \cdot (2 - 2 \cdot \varepsilon) > 2,
\end{aligned}$$

which holds for fixed $\varepsilon < \beta/(1+\beta)$. □

Lemma 3.7. *For every fixed* ε *with* $0 < \varepsilon < \frac{k-j}{(2k-j-1)(k-2)(1+\beta)}$, $j = 2, \ldots, k-1$, *it is*

$$s_{2,j}(\mathcal{G}^*) = o(|V^*|). \quad (48)$$

Proof. For $j = 2, \ldots, k-1$, by (37), (42) and (45), and using $T = n^{1+\beta}$ for fixed $\beta, \varepsilon > 0$, we infer

$$\begin{aligned}
&s_{2,j}(\mathcal{G}^*) = o(|V^*|)\\
&\Longleftarrow T^{\frac{2j-4}{k-1}+(2k-j)\varepsilon} \cdot (\log T)^3 / A^{\frac{j-2}{k-1}} = o(T^{\frac{2k-4}{k-1}+\varepsilon} / A^{\frac{k-2}{k-1}})\\
&\Longleftrightarrow A^{\frac{k-j}{k-1}} \cdot (\log T)^3 / T^{\frac{2k-2j}{k-1}-(2k-j-1)\varepsilon} = o(1)\\
&\Longleftrightarrow n^{(1+\beta)(2k-j-1)\varepsilon - \frac{k-j}{k-2}} \cdot (\log n)^{3+\frac{k-j}{(k-1)(k-2)}} = o(1)\\
&\Longleftrightarrow (1+\beta) \cdot (2 \cdot k - j - 1) \cdot \varepsilon < \frac{k-j}{k-2},
\end{aligned}$$

which holds for fixed $\varepsilon < \frac{k-j}{(k-2)(2k-j-1)(1+\beta)}$. □

To satisfy $p = T^\varepsilon / t_0 \leq 1$, with (28) we need $T^\varepsilon / ((k \cdot c_k)^{1/(k-1)} \cdot A^{(k-2)/(k-1)} \cdot T^{2/(k-1)}) \leq 1$. This holds with (37) for $0 < \varepsilon \leq 2 - 1/(1+\beta)$. For $\varepsilon := 1/(C \cdot (1+\beta))$ for fixed $C \geq k^2$ and $\beta := 1/(C-1)$ this and the assumptions in Lemmas 3.6 and 3.7 are fulfilled.

From each 3-element edge $E \in \mathcal{E}_3^{0*}$, and each $(2, j)$-cycle in $\mathcal{G}^*$ we delete one vertex in time

$$O\left(|V^*| + \sum_{j=2}^{k-1} s_{2,j}(\mathcal{G}^*)\right) = O(|V^*|). \tag{49}$$

By Lemmas 3.6 and 3.7 the resulting induced subhypergraph $\mathcal{G}^{**} = (V^{**}, \mathcal{E}_k^{**})$ of $\mathcal{G}^*$ with $\mathcal{E}_k^{**} := \mathcal{E}_k^* \cap [V^{**}]^k$ satisfies $|V^{**}| = (1 - o(1)) \cdot |V^*| \geq |V^*|/2$ and does not contain any 3-element edges from $\mathcal{E}_3^{0*}$ or $(2, j)$-cycles arising from $\mathcal{E}_k^{**}$, i.e., $\mathcal{G}^{**}$ is a linear, k-uniform hypergraph. By (42) we have $|V^{**}| \geq (c_1'/4) \cdot T^{(2k-4)/(k-1)+\varepsilon}/A^{(k-2)/(k-1)}$, and using $|\mathcal{E}_k^{**}| \leq |\mathcal{E}_k^*|$, by (44) the average degree t^{k-1} of the k-uniform subhypergraph $\mathcal{G}^{**} = (V^{**}, \mathcal{E}_k^{**})$ of $\mathcal{G}$ satisfies

$$\begin{aligned} t^{k-1} &= \frac{k \cdot |\mathcal{E}_k^{**}|}{|V^{**}|} \leq \frac{k \cdot (k+1) \cdot c_k' \cdot T^{\frac{2k-4}{k-1}+k\varepsilon}/A^{\frac{k-2}{k-1}}}{(c_1'/4) \cdot T^{\frac{2k-4}{k-1}+\varepsilon}/A^{\frac{k-2}{k-1}}} \\ &= \frac{4 \cdot k \cdot (k+1) \cdot c_k'}{c_1'} \cdot T^{(k-1)\varepsilon} =: t_1^{k-1}. \end{aligned} \tag{50}$$

Since $\mathcal{G}^{**}$ is linear, the assumptions in Theorem 3.1 are fulfilled, and, using (37), (50), and $T = n^{1+\beta}$ and $\varepsilon = 1/(k^2 \cdot (1+\beta))$ we find for any $\delta > 0$ in time

$$\begin{aligned} O\left(|\mathcal{E}_k^{**}| + \frac{|V^{**}|^3}{t_1^{3-\delta}}\right) &= O\left(\frac{T^{\frac{2k-4}{k-1}+k\varepsilon}}{A^{\frac{k-2}{k-1}}} + \frac{T^{\frac{6k-12}{k-1}+\varepsilon\delta}}{A^{\frac{3k-6}{k-1}}}\right) \\ &= O\left(\frac{n^3 \cdot T^{\varepsilon\delta}}{(\log n)^{\frac{3}{k-1}}}\right) = O\left(\frac{n^{3+\delta/k^2}}{(\log n)^{\frac{3}{k-1}}}\right) \end{aligned} \tag{51}$$

an independent set I of size

$$\begin{aligned} |I| &= \Omega\left(\frac{|V^{**}|}{t} \cdot (\log t)^{\frac{1}{(k-1)}}\right) = \Omega\left(\frac{|V^{**}|}{t_1} \cdot (\log t_1)^{\frac{1}{(k-1)}}\right) \\ &= \Omega\left(\frac{T^{\frac{2k-4}{k-1}+\varepsilon}/A^{\frac{k-2}{k-1}}}{T^{\varepsilon}} \cdot (\log T^{\varepsilon})^{\frac{1}{(k-1)}}\right) = \Omega\left(\frac{n}{(\log n)^{\frac{1}{k-1}}} \cdot (\log T)^{\frac{1}{(k-1)}}\right) \\ &= \Omega(n) \quad \text{since } T = n^{1+\beta} \text{ and } \beta, \varepsilon > 0 \text{ are constants.} \end{aligned}$$

By choosing the constant $c > 0$ in (37) sufficiently small, we obtain an independent set of size n, which yields a desired set of n points in $[0, 1]^2$ such that, after rescaling, the area of the convex hull of any k distinct of these n points is at least $\Omega((\log n)^{1/(k-2)}/n^{(k-1)/(k-2)})$. Adding the running times in (46), (49) and (51) we get for $\beta = 1/(C-1)$ and $\delta < 1$ the time bound $O(n^{2k-2+(4(k-1))/(C-1)} \cdot (\log n)^5)$. Thus, we may achieve the time bound $O(n^{2k-2+\delta'})$ for any fixed $\delta' > 0$ by choosing $\varepsilon := 1/(C \cdot (1+\beta))$ and $\beta := 1/(C-1)$, where $C \geq k^2$ is a constant with $C > 1 + (4k-4)/\delta'$. □

References

[1] M. Ajtai, J. Komlós, J. Pintz, J. Spencer, E. Szemerédi, Extremal uncrowded hypergraphs, Journal of Combinatorial Theory Series 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, The on-line Heilbronn's triangle problem, Discrete Mathematics 283 (2004) 7–14.
[4] C. Bertram–Kretzberg, H. Lefmann, The algorithmic aspects of uncrowded hypergraphs, SIAM Journal on Computing 29 (1999) 201–230.
[5] C. Bertram-Kretzberg, T. Hofmeister, H. Lefmann, An algorithm for Heilbronn's problem, SIAM Journal on Computing 30 (2000) 383–390.
[6] P. Brass, An upper bound for the d-dimensional analogue of Heilbronn's triangle problem, SIAM Journal on Discrete Mathematics 19 (2005) 192–195.
[7] B. Chazelle, Lower bounds on the complexity of polytope range searching, Journal of the American Mathematical Society 2 (1989) 637–666.
[8] R.A. Duke, H. Lefmann, V. Rödl, On uncrowded hypergraphs, Random Structures and Algorithms 6 (1995) 209–212.
[9] A. Fundia, Derandomizing Chebychev's inequality to find independent sets in uncrowded hypergraphs, Random Structures and Algorithms 8 (1996) 131–147.
[10] T. Jiang, M. Li, P. Vitany, The average case area of Heilbronn-type triangles, Random Structures and Algorithms 20 (2002) 206–219.
[11] J. Komlós, J. Pintz, E. Szemerédi, On Heilbronn's triangle problem, Journal of the London Mathematical Society 24 (1981) 385–396.
[12] J. Komlós, J. Pintz, 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 and large quadrangles (extended abstract), in: Proceedings 15th Annual International Symposium on Algorithms and Computation, ISAAC'2004, in: LNCS, vol. 3341, Springer, 2004, pp. 657–668.
[14] H. Lefmann, On Heilbronn's problem in higher dimension, Combinatorica 23 (2003) 669–680.
[15] H. Lefmann, N. Schmitt, A deterministic polynomial time algorithm for Heilbronn's problem in three dimensions, SIAM Journal on Computing 31 (2002) 1926–1947.
[16] K.F. Roth, On a problem of Heilbronn, Journal of the London Mathematical Society 26 (1951) 198–204.
[17] K.F. Roth, On a problem of Heilbronn, II, Proceedings of the London Mathematical Society 25 (3) (1972) 193–212.
[18] K.F. Roth, On a problem of Heilbronn, III, Proceedings of the London Mathematical Society 25 (3) (1972) 543–549.
[19] 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, in: Proc. of Symposia in Pure Mathematics, vol. 24, AMS, Providence, 1973, pp. 251–262.
[20] K.F. Roth, Developments in Heilbronn's triangle problem, Advances in Mathematics 22 (1976) 364–385.
[21] W.M. Schmidt, On a problem of Heilbronn, Journal of the London Mathematical Society (2) 4 (1972) 545–550.