On the partitioned version of the Erdős–Szekeres theorem*

Attila Pór
Rényi Institute of Mathematics, Hungarian Academy of Sciences
POB 127, Budapest, 1364 Hungary
apor@renyi.hu

Pavel Valtr
Department of Applied Mathematics, Charles University
and
Institute for Theoretical Computer Science (ITI), Charles University
Malostranské nám. 25, 118 00 Praha 1, Czech Republic
valtr@kam.mff.cuni.cz

Abstract

Let $k \geq 4$. A finite planar point set X is called a convex k-clustering, if it is a disjoint union of k sets $X_1, \ldots, X_k$ of equal sizes such that $x_1 x_2 \ldots x_k$ is a convex k-gon for each choice of $x_1 \in X_1, \ldots, x_k \in X_k$. Answering a question of Gil Kalai, we show that for every $k \geq 4$ there are two constants $c = c(k)$, $c' = c'(k)$ such that the following holds. If X is a finite set of points in general position in the plane then it has a subset X' of size at most c' such that $X \setminus X'$ can be partitioned into at most c convex k-clusterings. The special case $k = 4$ was proved earlier by Pór. Our result strengthens the so-called positive fraction Erdős–Szekeres theorem proved by Bárány and Valtr. The proof gives reasonable estimates on c and c', and it works also in higher dimensions. We also improve the previous constants for the positive fraction Erdős–Szekeres theorem obtained by Pach and Solymosi.

*The research by Attila Pór was partially supported by Hungarian National Foundation Grants T 032452 and T 038397. The research by Pavel Valtr was supported by project LN00A056 of The Ministry of Education of the Czech Republic, by Charles University grants No. 99/158 and 99/159. Part of the work on this paper was done during Attila Pór's visit at the Charles University. His visit was supported by Czech Republic Grant GAČR 201/99/0242.

1 Introduction

A set of points in the plane is said to be *in general position* if it contains no three points on a line. A finite set of points in the plane is *in convex position*, if it is the vertex set of a convex polygon. In this paper we give a generalization ("partitioned version") of the following famous result of Erdős and Szekeres [ES35]:

Theorem 1 (Erdős and Szekeres) *For any $k \geq 3$ there is a least integer $f(k)$ such that any set of at least $f(k)$ points in general position in the plane contains k points in convex position.*

For a triple (a, b, c) of points in general position in the plane, we say that its *orientation* is "+", if the clockwise order of the vertices of the triangle abc is a, b, c. Otherwise we say that the *orientation* of (a, b, c) is "−". We say that two t-tuples of points in general position $(p_1, \ldots, p_t)$ and $(p'_1, \ldots, p'_t)$ have the same *order type*, if the orientations of the two triples (p_i, p_j, p_l) and (p'_i, p'_j, p'_l) are equal for each choice of distinct indices i, j, l. A finite planar point set X is called a *k-clustering*, if it is a disjoint union of k sets $X_1, \ldots, X_k$ of equal sizes such that all k-tuples $(x_1, \ldots, x_k)$, $x_1 \in X_1, \ldots, x_k \in X_k$, have the same order type. If, moreover, $x_1, x_2, \ldots, x_k$ are in convex position, then X is called a *convex k-clustering* (see Fig. 1).

Bárány and Valtr [BV98] proved the following generalization of the Erdős–Szekeres theorem:

Theorem 2 (positive fraction Erdős–Szekeres theorem) *For any $k \geq 3$ there is an $\varepsilon_k > 0$ such that if X is a finite set of points in general position in the plane with $|X| \geq f(k)$ then it contains a convex k-clustering of size at least $\varepsilon_k |X|$. ($f(k)$ is the function from the Erdős–Szekeres theorem.)*

Repeated applications of the positive fraction Erdős–Szekeres theorem show that any set of n points in general position in the plane can be partitioned into at most $c_k \log n$ convex k-clusterings and a remaining set of size at most c'_k. Gil Kalai asked if $c_k \log n$ can be replaced by c_k, which would give a strengthening of the positive fraction Erdős–Szekeres theorem. It is easy to see for $k = 3$ and it was proved for $k = 4$ by Pór [P01]. Here we prove it for an arbitrary k:

Theorem 3 (partitioned Erdős–Szekeres theorem) *For every $k \geq 3$ there are two constants $c = c(k)$, $c' = c'(k)$ such that the following holds.*

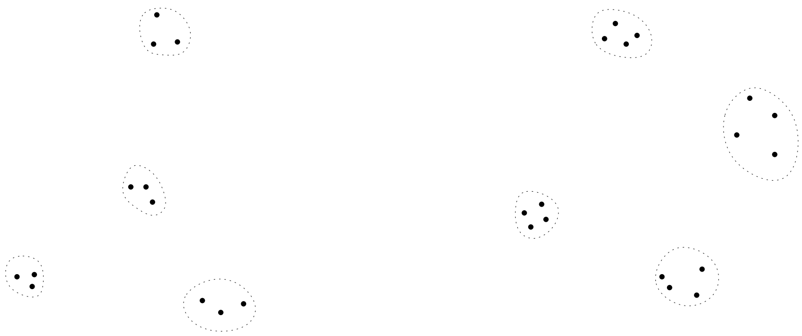

Figure 1: Two 4-clusterings.

If X is a finite set of points in general position in the plane then it has a subset X' of size at most c' such that $X \setminus X'$ can be partitioned into at most c convex k-clusterings.

Our proof gives Theorem 3 with constants $c(k) = k^{O(k^2)}$ and $c'(k) = f(k) - 1$, where $f(k) < 4^k$ is the number from the Erdős–Szekeres theorem. If $|X|$ is sufficiently large, then $c'(k)$ can be lowered to $rr(k) - 1$, where $rr(k)$ is the so-called Ramsey-remainder for convex sets (see Section 6 for details). Both bounds on $c'(k)$ are optimal, while our best lower bound on $c(k)$ is exponential in k.

We remark that Pór [P01] proved a weaker version of Theorem 3 when the k-clusterings forming the partition of $X \setminus X'$ are not required to be convex.

Our proof of Theorem 3 relies on the positive fraction Erdős–Szekeres theorem (Theorem 2). Pach and Solymosi [PS98] proved that Theorem 2 holds with $\varepsilon_k = \alpha^{-k^2}$ for some fixed $\alpha > 1$. In Section 2 we give the following better bound:

Theorem 4 *Theorem 2 holds with $\varepsilon_k = k \cdot 2^{-32k}$.*

In Section 6 we show that there is a $\beta > 1$ such that Theorem 2 does not hold with the constants $\varepsilon_k = \beta^{-k}$.

Analogues of Theorems 3 and 4 for collections of pairwise disjoint compact convex sets are given in [PV01].

Here is the organization of the paper: Section 2 contains the proof of Theorem 4. In Section 3 we derive Theorem 3 from two key lemmas, which are then proved in Sections 4 and 5, respectively. Section 6 contains concluding remarks.

2 Proof of Theorem 4

Let X be a set of n points in general position in the plane. We may suppose that $n > 2^{32k}$. It follows from the best known upper bounds [ES35, TV98] on the function $f(k)$ from the Erdős-Szekeres theorem (Theorem 1) that among any 4^{4k} points in general position it is always possible to find $4k$ points in convex position. Therefore, if we choose randomly and uniformly a 4^{4k}-point subset X_0 of X and consequently randomly and uniformly a $4k$-point subset X_1 of X_0, then X_1 is in convex position with probability at least $1/\binom{4^{4k}}{4k}$. Clearly, every $4k$-point subset of X is chosen for X_1 with the same probability. It follows that X contains at least $\frac{\binom{n}{4k}}{\binom{4^{4k}}{4k}}$ $4k$-point subsets in convex position.

If X_1 is a $4k$-point subset of X in convex position, then we say that its subset Y of size $2k$ *supports* X_1 if the points of X_1 sorted in the clockwise order alternately belong to Y and to $X_1 \setminus Y$. Clearly, X_1 is supported by two subsets.

Since X has $\binom{n}{2k}$ $2k$-point subsets, there is a $2k$-point subset Y of X which supports at least $2 \cdot \frac{\binom{n}{4k}}{\binom{4^{4k}}{4k}} / \binom{n}{2k} > \frac{(n-4k)^{2k} \cdot \frac{(2k)!}{(4k)!}}{\binom{4^{4k}}{4k}} > \frac{(n-4k)^{2k} \cdot (2k)!}{(4^{4k})^{4k}} > \frac{(n-4k)^{2k}}{2^{32k^2}}$ $4k$-point subsets of X in convex position. We fix such a set Y.

Let $y_1, \ldots, y_{2k}$ be the points of Y listed in the clockwise order. For each $i = 1, \ldots, 2k$, let T_i denote the region outside of $\operatorname{conv} Y$ bounded by the segment $y_i y_{i+1}$ and by parts of the lines $y_{i-1}y_i, y_{i+1}y_{i+2}$ (indices are counted modulo $2k$) – see Fig. 2.

Observation 1 *If $Y = \{y_1, \ldots, y_{2k}\}$ supports a $4k$-point subset X_1 of X, then $X_1 = Y \cup \{x_1, \ldots, x_{2k}\}$, where x_i lies in the interior of T_i for each $i = 1, \ldots, 2k$.* $\square$

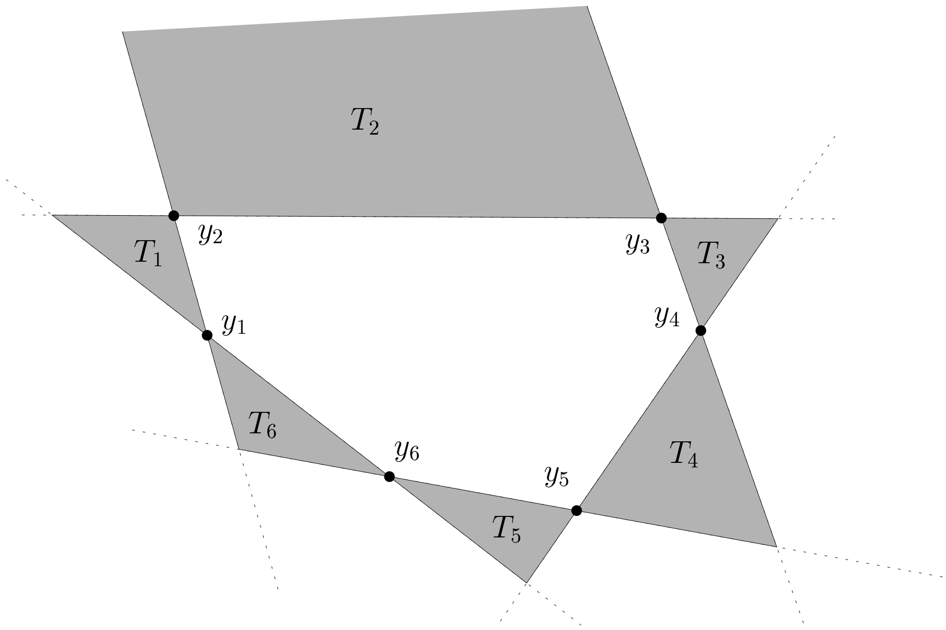

Figure 2: The regions T_i.

For $i = 1, \dots, 2k$, let t_i denote the number of points of X lying in the interior of T_i. It follows from Observation 1 that Y supports at most $\prod_{i=1}^{2k} t_i$ $4k$-point sets. Thus,

$$\frac{(n-4k)^{2k}}{2^{32k^2}} \leq \prod_{i=1}^{2k} t_i.$$

If α denotes the k-th smallest number t_i, then $\prod_{i=1}^{2k} t_i \leq \alpha^k n^k$. It follows that

$$\frac{(n-4k)^{2k}}{2^{32k^2}} \leq \alpha^k n^k$$

and thus

$$\alpha \geq \frac{(n-4k)^2}{2^{32k} n} > \frac{n-8k}{2^{32k}} > \frac{n}{2^{32k}} - 1.$$

Let $I \subset \{1, \ldots, 2k\}$ be a set of k indices i with $t_i \geq \alpha > \frac{n}{2^{32k}} - 1$. For each $i \in I$, we fix a set C_i of $\lceil \frac{n}{2^{32k}} - 1 \rceil$ points of X in the interior of T_i. Then the sets $C_i \cup \{y_i\}, i \in I$, form a convex k-clustering of size $k \cdot \lceil \frac{n}{2^{32k}} \rceil$. $\square$

In the proof of Theorem 4 we used that the best known bounds on $f(k)$ give $f(k) \leq 4^k$. If the conjecture of Erdős and Szekeres [ES60] that $f(k) = 2^{k-2} + 1$ is true then the above proof gives still a better constant $\varepsilon_k = k \cdot 2^{-16k}$.

3 Proof of the partitioned version of the Erdős-Szekeres theorem

In the whole proof, $k \geq 3$ is fixed and $K := 2^{32(k+1)+1}$.

For $t \geq 3$, we say that a family $\{X_1, X_2, \ldots, X_t\}$ of t planar point sets has *the same-type property*, if all t-tuples $(x_1, \ldots, x_t)$, $x_1 \in X_1, \ldots, x_t \in X_t$, have the same order type (and we say that the pair $\{X_1, X_2\}$ of two planar point sets has *the same-type property*, if $\operatorname{conv} X_1 \cap \operatorname{conv} X_2 = \emptyset$). If, moreover, $x_1, x_2, \ldots, x_t$ are in convex position, then we say that the sets $X_1, X_2, \ldots, X_t$ are *in convex position*.

We say that a sequence $(X_1, X_2, \ldots, X_t)$ of planar point sets has *the sequential same-type property*, if the family $\{X_1, X_2, \ldots, X_i, X_{i+1} \cup X_{i+2} \cup \ldots \cup X_t\}$ has the same-type property for each $i = 1, \ldots, t-1$. A crucial notion in our proof is the notion of so-called s-configurations. An *s-configuration* is a finite set X of points in general position in the plane which is partitioned into $s+1$ sets $\mathcal{C}_1, \ldots, \mathcal{C}_s, Y$ such that:

(i) $\mathcal{C}_i$ is a convex $(k+1)$-clustering for each $i = 1, \ldots, s$,

(ii) for each $i = 1, \ldots, s$, one of the $k+1$ clusters of $\mathcal{C}_i$ is denoted C_i so that the sequence $(C_1, \ldots, C_s, Y)$ has the sequential same-type property, and

(iii) $|C_1| = |C_2| = \ldots = |C_s| = \frac{|X|}{K} = \frac{|Y|}{K - s(k+1)}$, where $K = 2^{32(k+1)+1}$.

Sometimes we mean by an *s-configuration* not the set X but the corresponding $(2s+1)$-tuple $(\mathcal{C}_1, \ldots, \mathcal{C}_s; C_1, \ldots, C_s; Y)$ of sets. The definition of an s-configuration is somewhat technical. At this point we give two remarks to it:

Remark 1: The convex $(k+1)$-clustering $\mathcal{C}_i$ can be partitioned almost completely into few $(k+1)$ convex k-clusterings (Observation 2 in Section 5).

Moreover, if we use the points of the cluster $C_i \subset \mathcal{C}_i$ elsewhere in the final partition, then the rest of $\mathcal{C}_i$ is a convex k-clustering. It is the reason why we require $\mathcal{C}_i$ to be a convex $(k+1)$-clustering.

Remark 2: s-configurations are defined so that the following two conditions hold. An s-configuration can be partitioned almost completely into a bounded number of $(s+1)$-configurations (Lemma 1 below), and if s is large enough ($s = 2k^2$ suffices) then an s-configuration can be partitioned almost completely into a bounded number of convex k-clusterings (Lemma 2 below).

Here are the two key lemmas:

Lemma 1 *For $s = 1, \ldots, 2k^2$, every s-configuration is a disjoint union of at most $8s^2$ $(s+1)$-configurations and a remaining set of size at most $25s^2K$.*

Lemma 2 *Every $(2k^2)$-configuration is a disjoint union of at most $2K \log K$ convex k-clusterings and a remaining set of size at most 4^k.*

Lemma 1 will be obtained using the separation method of Pór [P01], Lemma 2 relies on Theorem 4 and on the Erdős-Szekeres theorem on monotone subsequences (Theorem 5 in Section 5). We now derive Theorem 3 from Lemmas 1 and 2.

Proof of Theorem 3. Let X be a finite set of points in general position in the plane. We may suppose that $|X|$ is divisible by $2K$ (this can be achieved by removing at most $2K-1$ points from X). By Theorem 4, X contains a convex $(k+1)$-clustering $\mathcal{C}_1$ with each cluster of size $\frac{|X|}{K}$. We denote one of the clusters by C_1.

Let l be a line ("ham-sandwich cut") cutting each of the sets $C_1, X \setminus \mathcal{C}_1$ into two equal parts. Let $C_1 = C_1^1 \cup C_1^2, X \setminus \mathcal{C}_1 = Y^1 \cup Y^2$ be the corresponding partitions such that $|C_1^1| = |C_1^2|, |Y^1| = |Y^2| = (K-(k+1))|C_1^1|$, and that l separates $C_1^1 \cup Y^2$ from $C_1^2 \cup Y^1$.

We partition the convex $(k+1)$-clustering $\mathcal{C}_1$ into two convex $(k+1)$-clusterings $\mathcal{C}_1^1, \mathcal{C}_1^2$ of equal sizes by cutting each cluster of $\mathcal{C}_1$ into two equal parts (the cluster C_1 is cut into the above subsets C_1^1, C_1^2).

We obtain a partition of the set X into two 1-configurations $X^j = (\mathcal{C}_1^j; C_1^j; Y^j), j = 1, 2$. Applying first Lemma 1 ($2k^2$-times) and then Lemma 2 on each part, we get a partition of each of the two 1-configurations X^j into at most $\left(\prod_{s=1}^{2k^2-1} 8s^2\right) \cdot 2K \log K = k^{O(k^2)}$ convex k-clusterings and a remaining set of size at most $2 \cdot k^{O(k^2)} \cdot 25(2k^2)^2 K + k^{O(k^2)} \cdot 4^k = k^{O(k^2)}$. Theorem 3 follows. $\square$

4 Proof of Lemma 1

Claim 1 *If a family $\{C_1, \ldots, C_s, Y\}$ of $s+1$ finite point sets has the same-type property, then there is a closed strip S bounded by two parallel lines such that $Y \subset S$ and $C_i \cap S = \emptyset$ for $i = 1, \ldots, s$.*

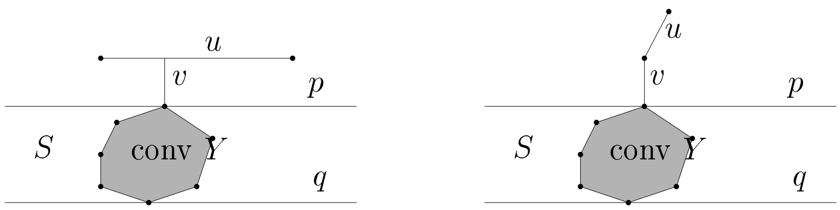

Figure 3: The strip S (two cases).

Proof. Let u be a segment connecting a pair of points in $C_1 \cup \ldots \cup C_s$ such that its distance to $\operatorname{conv} Y$ is smallest possible. Let v be a segment of length $\operatorname{dist}(u, \operatorname{conv} Y)$ connecting a point of u with a point of $\operatorname{conv} Y$ (see Fig. 3). Further, let p be the line perpendicular to v going through the point $v \cap \operatorname{conv} Y$, and let q be the other line of the same direction tangent to $\operatorname{conv} Y$.

Then the strip S bounded by the lines p, q contains $\operatorname{conv} Y$ and no point of $C_1 \cup \ldots \cup C_s$ (otherwise a segment connecting a point of $S \cap (C_1 \cup \ldots \cup C_s)$ with one of the end-points of u would be closer to $\operatorname{conv} Y$ than u – a contradiction). $\square$

Claim 2 *(i) Let $N \geq 1$ be an integer, and let C, Y be two disjoint finite sets of points in the plane such that $|Y| = N \cdot |C|$ and $C \cup Y$ contains no pair of points on a horizontal line. Then there are partitions $Y = Y^1 \cup Y^2 \cup Y^{\mathrm{rem}}, C = C^1 \cup C^2 \cup C^{\mathrm{rem}}$ such that $Y^1 \cup C^2$ is separated from $Y^2 \cup C^1$ by a horizontal line, $|Y^x| = N \cdot |C^x|$ for $x \in \{1, 2, \mathrm{rem}\}$, and $|C^{\mathrm{rem}}| \leq 1$.*

(ii) Let C, D, Y be three pairwise disjoint finite sets of points with $|Y| = N \cdot |C| = N \cdot |D|$, where $N \geq 1$ is an integer. Suppose further that $C \cup D \cup Y$ is in general position and any two of the sets C, D, Y are separated from each other by a horizontal line. Then there are partitions $Y = Y^1 \cup Y^2 \cup Y^{\mathrm{rem}}$, $C = C^1 \cup C^2 \cup C^{\mathrm{rem}}$, $D = D^1 \cup D^2 \cup D^{\mathrm{rem}}$ such that:

1. $|Y^x| = N \cdot |C^x| = N \cdot |D^x|$ for each $x \in \{1, 2, \mathrm{rem}\}$,

2. $|C^{\mathrm{rem}}| = |D^{\mathrm{rem}}| \leq 1$, and

3. the triple (C^j, D^j, Y^j) has the same-type property, for $j = 1, 2$.

Proof. **(i)** Let NC be the multiset of points of C, each taken with multiplicity N. Consider the multiset $Y \cup NC$ of size $2N|C|$, and choose a horizontal line p such that at most $N|C|$ points of $Y \cup NC$ lie strictly above p and also at most $N|C|$ points of $Y \cup NC$ lie strictly below p.

If no point of C lies on p, we may suppose that also no point of Y lies on p. The line p determines suitable partitions of Y, C with $Y^{\mathrm{rem}} = C^{\mathrm{rem}} = \emptyset$ in this case.

If a point $c \in C$ lies on p, set $C^{\mathrm{rem}} := \{c\}$ and put N suitable points of Y to Y^{rem}. Then the line p will again determine suitable partitions of $Y \setminus Y^{\mathrm{rem}}, C \setminus C^{\mathrm{rem}}$.

(ii) We distinguish two cases (other cases are obtained by symmetry).

Case 1: the sets C, D, Y appear in the order C, D, Y from top to bottom. (see Fig. 4)

We use a "sweep-line" argument. We take a horizontal line p separating C from D. We rotate p in the clockwise order. During the rotation we are changing the center of rotation so that at each moment p cuts $C \cup D$ into two equal parts (the partition determined by p changes at finitely many moments when two points of $C \cup D$ lie on p) – see Fig. 4 (several positions of p during the rotation are shown by dotted lines). Moreover, we can proceed so that p always contains at most one point of Y. Let p^+ be the open halfplane of points to the right of p, and let p^- be the open halfplane opposite to p^+. We rotate p until we achieve $|p^+ \cap Y| \geq N \cdot (|p^+ \cap C| + \frac{1}{2}|p \cap C|)$. Then we stop and denote the final position of p by p_0.

If p passes through a point of Y immediately before it gets to its final position p_0, then p_0 contains no point of $C \cup D \cup Y$ and it gives us the required partitions

$C^1 = C \cap p_0^+$, $C^2 = C \cap p_0^-$,
$D^1 = D \cap p_0^-$, $D^2 = D \cap p_0^+$,
$Y^1 = Y \cap p_0^+$, $Y^2 = Y \cap p_0^-$,

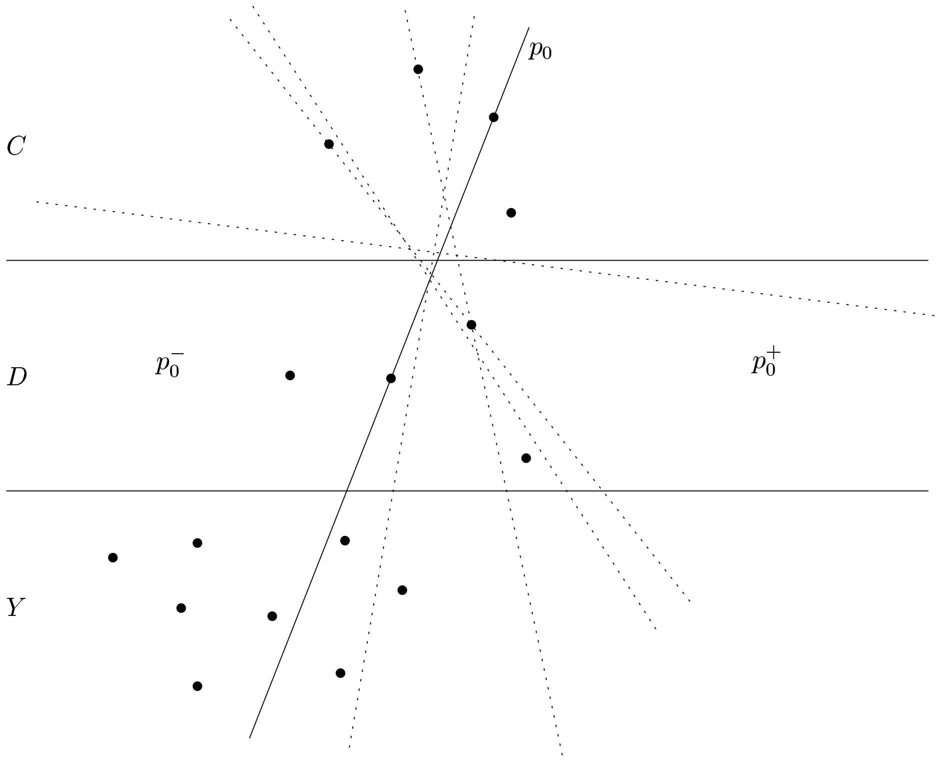

Figure 4: Case 1 in the proof of Claim 2(ii)

$C^{\text{rem}} = D^{\text{rem}} = Y^{\text{rem}} = \emptyset$.

Otherwise p contains a point $c \in C$ and a point $d \in D$ either in the position p_0 or immediately before it. In these cases we set

$C^{\text{rem}} = \{c\}$, $D^{\text{rem}} = \{d\}$,

$C^1 = C \cap p_0^+$, $C^2 = C \cap p_0^-$,

$D^1 = D \cap p_0^-$, $D^2 = D \cap p_0^+$,

and we appropriately choose a partition $Y = Y^1 \cup Y^2 \cup Y^{\text{rem}}$ such that $|Y^{\text{rem}}| = N, Y^1 \subset p_0^+, Y^2 \subset p_0^-$.

Case 2: the sets C, D, Y appear in the order C, Y, D from top to bottom.

We take a vertical line to the left of $C \cup D \cup Y$ and sweep it to the right so that it cuts C and D in the same proportions at any moment. We define p^+ and p^- as above and stop as soon as $|Y \cap (p \cup p^-)| \geq N \cdot |C \cap p^+|$. Then we continue similarly as in Case 1. Details are left to the reader. $\square$

A *weak* $(s+1)$*-configuration* is a $(2s+3)$-tuple $(\mathcal{C}_1, \ldots, \mathcal{C}_{s+1}; C_1, \ldots, C_{s+1}; Y)$ such that $(\mathcal{C}_1, \ldots, \mathcal{C}_s; C_1, \ldots, C_s; Y \cup \mathcal{C}_{s+1})$ is an s-configuration and C_{s+1} is one of the clusters of a convex $(k+1)$-clustering $\mathcal{C}_{s+1}$ of size $|\mathcal{C}_{s+1}| = |\mathcal{C}_1|$.

Proof of Lemma 1. Let $X = (\mathcal{C}_1, \ldots, \mathcal{C}_s; C_1, \ldots, C_s; Y)$ be a non-empty s-configuration. By Claim 1, there is a closed strip S such that Y lies in S and $C_1 \cup \ldots \cup C_s$ lies outside of S. We may suppose that the lines bounding S are horizontal. Let S^+ (resp. S^-) be the open halfplane of points above (resp. below) S.

The size of Y is $|X| - s(k+1)\frac{|X|}{K} \geq \frac{|X|}{2}$. Therefore, by Theorem 4, Y contains a convex $(k+1)$-clustering $\mathcal{C}_{s+1}$ of size $(k+1)\frac{|X|/2}{K/2} = (k+1)\frac{|X|}{K}$. Fix a cluster C_{s+1} in $\mathcal{C}_{s+1}$.

In the sequel, the weak $(s+1)$-configuration $X = (\mathcal{C}_1, \ldots, \mathcal{C}_{s+1}; C_1, \ldots, C_{s+1}; Y \setminus \mathcal{C}_{s+1})$ will be partitioned into a bounded number of $(s+1)$-configurations and a small remaining set.

We apply Claim 2(i) on the sets $C_{s+1}, Y \setminus \mathcal{C}_{s+1}$. We obtain partitions $C_{s+1} = C^1_{s+1} \cup C^2_{s+1} \cup C^{\text{rem}}_{s+1}$ and $Y \setminus \mathcal{C}_{s+1} = Y^1 \cup Y^2 \cup Y^{\text{rem}}$ such that $C^1_{s+1} \cup Y^2$ is separated from $C^2_{s+1} \cup Y^1$ by a horizontal line, $|C^1_{s+1}| = \frac{|Y^1|}{K-(s+1)(k+1)}$, $|C^2_{s+1}| = \frac{|Y^2|}{K-(s+1)(k+1)}$, $|C^{\text{rem}}_{s+1}| \leq 1$, $|Y^{\text{rem}}| \leq K - (s+1)(k+1)$.

For each $i = 1, \ldots, s$, we take an arbitrary partition $C_i = C^1_i \cup C^2_i \cup C^{\text{rem}}_i$ with $|C^x_i| = |C^x_{s+1}|$ for each $x \in \{1, 2, \text{rem}\}$. Finally, for each $i = 1, \ldots, s+1$, we partition the convex $(k+1)$-clustering $\mathcal{C}_i$ into three convex $(k+1)$-clusterings $\mathcal{C}^1_i, \mathcal{C}^2_i, \mathcal{C}^{\text{rem}}_i$ arbitrarily so that, for each $x \in \{1, 2, \text{rem}\}$, C^x_i is one of the clusters of $\mathcal{C}^x_i$.

The two obtained $(2s+3)$-tuples $X^i = (\mathcal{C}^i_1, \ldots, \mathcal{C}^i_{s+1}; C^i_1, \ldots, C^i_{s+1}; Y^i)$, $i = 1, 2$, are weak $(s+1)$-configurations forming a partition of $X \setminus X^{\text{rem}}$, where $|X^{\text{rem}}| \leq K$.

Both X^1 and X^2 satisfy the assumptions in the following lemma:

Claim 3 *Let $X = (\mathcal{C}_1, \ldots, \mathcal{C}_{s+1}; C_1, \ldots, C_{s+1}; Y)$ be a weak $(s+1)$-configuration such that the sets C_{s+1}, Y are separated from each other by a horizontal line and also each $C_i, i = 1, \ldots, s$, is separated from $C_{s+1} \cup Y$ by a horizontal line. Further, let $w(X)$ denote the number of indices $i \in \{1, \ldots, s\}$ such* *that $C_i \subset S^+$ and the triple C_i, C_{s+1}, Y does not have the same-type property. Then X can be partitioned "componentwise" into two weak $(s+1)$-configurations X_j with $w(X_j) \le w(X)/2$ and a remaining set X_{rem} of size at most K. (A partition of $X = (\mathcal{C}_1, \ldots; Y)$ into $X_1, X_2, X_{\mathrm{rem}}$ is* componentwise, *if $X^j = (\mathcal{C}_1^j, \ldots; Y^j)$, $j = 1, 2, \mathrm{rem}$, and $\mathcal{C}_i = \mathcal{C}_i^1 \cup \mathcal{C}_i^2 \cup \mathcal{C}_i^{\mathrm{rem}}$, $C_i = C_i^1 \cup C_i^2 \cup C_i^{\mathrm{rem}}$, $Y = Y^1 \cup Y^2 \cup Y^{\mathrm{rem}}$, $i = 1, \ldots, s+1$.)*

Proof. Suppose that $C_{i_1}, \ldots, C_{i_r}$ are the clusters C_i contributing to $w(X)$, i.e. they are the clusters C_i in S^+ such that the triple C_i, C_{s+1}, Y does not have the same-type property. Moreover, suppose that they are ordered in the order in which they are seen (within an angle smaller than π) from any point in $C_{s+1} \cup Y$, see Fig. 5. The linear order $C_{i_1}, \ldots, C_{i_r}$ is well-defined because the family $\{C_1, \ldots, C_s, C_{s+1} \cup Y\}$ has the same-type property.

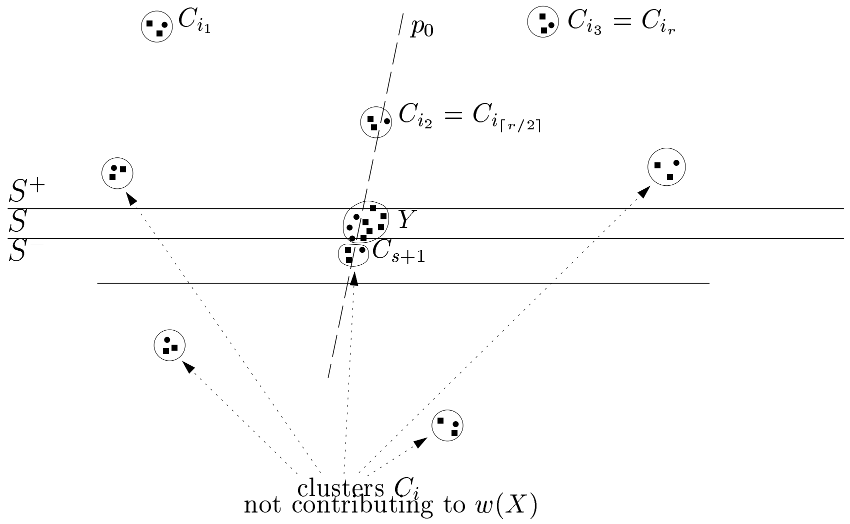

Figure 5: The partition $X = X_1 \cup X_2 \cup X_{\mathrm{rem}}$ in Claim 3 (the points of X_1 are denoted by squares, the points of X_2 by discs, X_{rem} is empty). The line p_0 is the line obtained in the application of Claim 2(ii). For technical reasons, we made Y small; in fact Y should contain majority of points of X.

We apply Claim 2(ii) on the triple of sets $C_{i_{\lceil r/2 \rceil}}, C_{s+1}, Y$. It gives us partitions $C_{i_{\lceil r/2\rceil}} = C^1_{i_{\lceil r/2\rceil}} \cup C^2_{i_{\lceil r/2\rceil}} \cup C^{\mathrm{rem}}_{i_{\lceil r/2\rceil}}$, $C_{s+1} = C^1_{s+1} \cup C^2_{s+1} \cup C^{\mathrm{rem}}_{s+1}$, $Y = Y^1 \cup Y^2 \cup Y^{\mathrm{rem}}$.

We further partition each set $C_i, i \neq i_{\lceil r/2\rceil}, s+1$, into three parts $C^1_i, C^2_i, C^{\mathrm{rem}}_i$ arbitrarily so that $|C^x_i| = |C^x_{s+1}|$ for each $x \in \{1, 2, \mathrm{rem}\}$. Finally, any partition of each $(k+1)$-clustering $\mathcal{C}_i$ into three $(k+1)$-clusterings $\mathcal{C}^1_i, \mathcal{C}^2_i, \mathcal{C}^{\mathrm{rem}}_i$ such that C^x_i is a cluster of $\mathcal{C}^x_i$ $(i = 1, \ldots, s+1, x = 1, 2, \mathrm{rem})$ finishes the required partition of X into two weak $(s+1)$-configurations $X_j = (\mathcal{C}^j_1, \ldots, Y^j)$, $j = 1, 2$, and a remaining set $\mathcal{C}^{\mathrm{rem}}_1 \cup \cdots \cup \mathcal{C}^{\mathrm{rem}}_{s+1} \cup Y^{\mathrm{rem}}$ of size at most K. (For each $j = 1, 2$ we have $w(X_j) \leq r/2 = w(X)/2$ because either only the clusters $C^j_{i_t}, 1 \leq t < \lceil r/2\rceil$, or only the clusters $C^j_{i_t}, \lceil r/2\rceil < t \leq r$, contribute to $w(X_j)$ – see Fig. 5.) $\square$

Since the partition in Claim 3 is componentwise, each of the subconfigurations X_1, X_2 obtained in Claim 3 satisfies the assumptions in Claim 3 and Claim 3 can be again applied on each of them. We iterate Claim 3 $\lfloor \log_2 s+1\rfloor$ times. We obtain a partition of a weak $(s+1)$-configuration X into (at most) $2^{\lfloor \log_2 s+1\rfloor} \leq 2s$ weak $(s+1)$-configurations X_j with $w(X_j) = 0$ (since after $\lfloor \log_2 s+1\rfloor$ iterations of Claim 3 we get $w(X_j) \leq w(X)/2^{\lfloor \log_2 s+1\rfloor} < s/s = 1$) and a remaining set of size at most $(2 \cdot 2^{\lfloor \log_2 s+1\rfloor} - 1)K < 4sK$. Analogously, we can consequently partition each X_j into at most $2s$ weak $(s+1)$-configurations X_{jl} and a remaining set of size at most $4sK$ such that $w(X_{jl}) = 0$ holds also if we replace S^+ by S^- in the definition of the parameter w. This means that the configurations X_{jl} are $(s+1)$-configurations.

Thus, each of the two weak $(s+1)$-configurations X^1, X^2 obtained above Claim 3 can be partitioned into at most $(2s)^2 = 4s^2$ $(s+1)$-configurations and a remaining set of size at most $(2s+1)4sK \leq 12s^2K$. Since X^1, X^2 form a partition of $X \setminus X^{\mathrm{rem}}$ (where $|X^{\mathrm{rem}}| \leq K$), we obtain a partition of X into $8s^2$ $(s+1)$-configurations and a remaining set of size at most $2 \cdot 12s^2K + K \leq 25s^2K$. This finishes the proof of Lemma 1. $\square$

5 Proof of Lemma 2

The proof of Lemma 2 relies on another famous result of Erdős and Szekeres [ES35]:

Theorem 5 (Erdős–Szekeres lemma) *Any sequences of $(k-1)^2+1$ distinct real numbers contains a monotone subsequence of length k.*

Proof of Lemma 2. Let $X = (\mathcal{C}_1, \ldots, \mathcal{C}_s, C_1, \ldots, C_s, Y)$ be an s-configuration, where $s = 2k^2$. By Claim 3 and by the pigeon-hole principle, there is a line l separating Y from at least k^2 of the clusters C_i. Let $C_{i_1}, \ldots, C_{i_{k^2}}$ be k^2 such clusters listed in the order in which they are seen (within an angle smaller than π) from any point of Y.

By the Erdős-Szekeres theorem on monotone subsequences (Theorem 5), the sequence $i_1, \ldots, i_{k^2}$ contains a monotone subsequence $j_1, \ldots, j_{k-1}$ of length $k-1$.

Claim 4 *The k sets $Y, C_{j_1}, \ldots, C_{j_{k-1}}$ are in convex position.*

Proof. Since Y is separated from $\bigcup_{i=1}^{k-1} C_{j_i}$ by a line, it suffices to show that for each $i = 1, \ldots, k-3$ the set $C_{j_{i+1}}$ cannot lie in the convex hull of $C_{j_i} \cup C_{j_{i+2}} \cup Y$. We use the sequential same-type property of the sequence $(C_1, \ldots, C_s, Y)$. If the sequence j_i, j_{i+1}, j_{i+2} is increasing, then $(C_{j_i}, C_{j_{i+1}}, C_{j_{i+2}}, Y)$ has the sequential same-type property. Otherwise the sequence j_i, j_{i+1}, j_{i+2} is decreasing and $(C_{j_{i+2}}, C_{j_{i+1}}, C_{j_i}, Y)$ has the sequential same-type property. In the first case the triple $\{C_{j_i}, C_{j_{i+1}}, C_{j_{i+2}} \cup Y\}$ has the same-type property, in the latter case the triple $\{C_{j_{i+2}}, C_{j_{i+1}}, C_{j_i} \cup Y\}$ has the same-type property. In either case, $C_{j_{i+1}}$ lies outside of the convex hull of $C_{j_i} \cup C_{j_{i+2}} \cup Y$. The lemma is proved. $\Box$

We now need the following corollary of Theorem 4 and a simple observation.

Corollary 1 *The set Y is a disjoint union of at most $K \log K$ convex k-clusterings and two remaining sets Y_0, Y_1 with $|Y_0| = \frac{|X|}{K}, |Y_1| \leq f(k)$.*

Proof. By Theorem 4, Y is a disjoint union of a convex k-clustering and a remaining set Y' of size at most $\max\{\left(1 - \frac{1}{K}\right)|Y|, f(k)\}$. Applying Theorem 4 on Y', we get a partition of the set Y into two convex k-clusterings and a remaining set of size at most $\max\{\left(1 - \frac{1}{K}\right)^2 |Y|, f(k)\}$.

After $K \log K$ applications of Theorem 4, we obtain a partition of Y into $K \log K$ convex k-clusterings and a remaining set Y_{rem} of size at most

$$\max\{\left(1 - \frac{1}{K}\right)^{K \log K} |Y|, f(k)\} \leq \max\{\frac{1}{K}|Y|, f(k)\} \leq \frac{|X|}{K} + f(k).$$

If $|Y_{\rm rem}| \geq \frac{\lfloor X \rfloor}{K}$ then we finish the whole partition by partitioning $Y_{\rm rem}$ into two subsets Y_0, Y_1, where $|Y_0| = \frac{\lfloor X \rfloor}{K}$.

Otherwise we repeatedly enlarge $Y_{\rm rem}$ by choosing an arbitrary of the $K \log K$ k-clusterings and moving one point from each of its clusters to Y_0. Each repetition of this step enlarges $|Y_{\rm rem}|$ by k, so we can achieve that $\frac{\lfloor X \rfloor}{K} \leq |Y_{\rm rem}| \leq \frac{\lfloor X \rfloor}{K} + (k-1)$, and then we make any partition $Y_{\rm rem} = Y_0 \cup Y_1$ with $|Y_0| = \frac{\lfloor X \rfloor}{K}$. $\square$

Observation 2 *Every convex $(k+1)$-clustering is a union of $k+1$ convex k-clusterings and a remaining set of size at most $k-1$.*

Proof. Let $\mathcal{C}$ be a convex $(k+1)$-clustering and let t be the size of a cluster in $\mathcal{C}$. Set $t_0 := t \bmod k$.

We can easily partition $\mathcal{C}$ into t_0 points and $k+1$ convex k-clusterings, each disjoint from one of the clusters of $\mathcal{C}$ and having size $k \cdot \lfloor t/k \rfloor$ or $k \cdot \lfloor t/k + 1 \rfloor$. $\square$

By Corollary 1, Y is a disjoint union of at most $K \log K$ convex k-clusterings and two remaining sets Y_0, Y_1 with $|Y_0| = \frac{\lfloor X \rfloor}{K}, |Y_1| \leq f(k)$. By Claim 4, the sets $Y_0, C_{j_1}, \ldots, C_{j_{k-1}}$ form a convex k-clustering. Further, each of the $k-1$ sets $\mathcal{C}_{j_1} \setminus C_{j_1}, \ldots, \mathcal{C}_{j_{k-1}} \setminus C_{j_{k-1}}$ is a convex k-clustering.

By Observation 2, each of the remaining $2k^2-(k-1)$ $(k+1)$-clusterings $\mathcal{C}_i$ can be partitioned into $k+1$ convex k-clusterings and a remaining set of at most $k-1$ points.

The s-configuration X is partitioned into at most $K \log K + 1 + (k-1) + (2k^2-(k-1))(k+1) < 2K \log K$ convex k-clusterings and a remaining set of size at most $f(k) + (2k^2-(k-1))(k-1) < 4^k$. $\square$

6 Concluding remarks

1. Upper bounds on the constants $c(k), c'(k)$. Our proof gives Theorem 3 with $c(k) = k^{O(k^2)}$, $c'(k) = k^{O(k^2)}$. It is easy to lower $c'(k)$ to $f(k)-1$ (and keep $c(k) = k^{O(k^2)}$ at the same time). As long as the remaining set of points has size bigger than $f(k)-1$, it contains a convex k-clustering of size k, i.e., k points in convex position, and we may reduce the size of the remaining set by it. The number of convex k-clusterings will increase by at most $k^{O(k^2)} \cdot k^{-1} = k^{O(k^2)}$, thus staying at most $k^{O(k^2)}$. This bound on $c'(k)$ cannot be improved in general, since any set of $f(k)-1$ points with no k points in convex position contains no non-empty convex k-clustering.

If $|X|$ is sufficiently large, then $c'(k)$ can be lowered to $rr(k)$, the so-called *Ramsey-remainder for convex sets*, defined as the smallest number r such that any set of sufficiently many points in general position in the plane can be partitioned into k-tuples of points in convex position and a remaining set of size at most r (see [ETV96] for estimates on the numbers $rr(k)$).

2. An upper bound on ε_k and a lower bound on $c(k)$. Let $C_{k/2}$ be the set of $2^{k/2-2}$ points with no convex $(k/2)$-gon constructed by Erdős and Szekeres [ES60]. We may suppose that no pair of points in $C_{k/2}$ lies on a horizontal line. We replace each point of C_k by the same number L of points which are very close together, form a so-called L-cup (e.g., see [TV98] for the definition), and any line trough two of them is almost horizontal.

Then any convex k-clustering has all points of at least $\approx k/2$ clusters contained in the same L-cup, and therefore its size is at most $(\approx 2 \cdot 2^{2-k/2})$-fraction of our set. This gives an upper bound $\approx \left(\sqrt{1/2}\right)^k$ on the constant ε_k in Theorem 2 (it can be improved to $\approx (1/2+\varepsilon)^k$ by considering $C_{(1-\varepsilon)k}$ instead of $C_{k/2}$). Consequently, we get a lower bound $\approx \left(\sqrt{2}\right)^k$ on the constant $c(k)$ in Theorem 3 (it can be improved to $\approx (2-\varepsilon)^k$).

3. Higher dimensions. For any $d \geq 3$, an analogue of Theorem 3 in $\mathbb{R}^d$ holds. We simply project the set $X \subset \mathbb{R}^d$ orthogonally to a plane chosen so that the projection is a set in general position. Then we apply Theorem 3. The obtained partition gives us a partition of the original set X into at most $k^{O(k^2)}$ convex k-clusterings and a remaining set of size at most $k^{O(k^2)}$. In fact, if $|X|$ is divisible by k and sufficiently large, then we can partition the whole set X into convex k-clusterings because the Ramsey-remainder for convex sets is 0 in this case (see [K01]).

Acknowledgment. We would like to thank two anonymous referees for their detailed comments.

References

[BV98] I. Bárány and P. Valtr, A positive fraction Erdős–Szekeres theorem, *Discrete and Computational Geometry* 19 (1998), 335-342.

[ES35] P. Erdős and Gy. Szekeres, A combinatorial problem in geometry, *Compositio Math.* 2 (1935), 464-470.

[ES60] P. Erdős and Gy. Szekeres, On some extremum problems in elementary geometry, *Ann. Univ. Sci. Budapest* 3/4 (1960/61), 53-62.

[ETV96] P. Erdős, Zs. Tuza, and P. Valtr, Ramsey-remainder, *European Journal of Combinatorics* 17 (1996), 519-532.

[K01] Gy. Károlyi, Ramsey–remainder for convex sets and the Erdős–Szekeres theorem, *to appear in Discr. Appl. Math.*

[PS98] J. Pach and J. Solymosi, Canonical theorems for convex sets, *Discrete and Computational Geometry* 19 (1998), 427-435.

[P01] A. Pór, The partitioned version of the Erdős–Szekeres theorem, submitted.

[PV01] A. Pór and P. Valtr, manuscript in preparation.

[TV98] G. Tóth and P. Valtr, Note on the Erdős–Szekeres theorem, *Discrete and Computational Geometry* 19 (1998), 457-459.