Comparison of Convex Hulls and Box Hulls*

Endre Boros[†] Vladimir Gurvich[‡] Ying Liu[§]

Abstract

A *convex hull* of a set of points X is the minimal convex set containing X. A *box* B is an interval $B = \{\mathbf{x}|\mathbf{x} \in [\mathbf{a}, \mathbf{b}], \mathbf{a}, \mathbf{b} \in \mathbb{R}^n\}$. A *box hull* of a set of points X is defined to be the minimal box containing X. Because both convex hulls and box hulls are closure operations of points, classical results for convex sets can naturally be extended for box hulls. We consider here the extensions of theorems by Carathéodory, Helly and Radon to box hulls and obtain exact results.

Let $\mathcal{A}_d$ be a family of closed sets in $\mathbb{R}^d$ which is closed under intersection. We can associate to such an intersection closed family a natural closure operation by defining the *$\mathcal{A}_d$-hull* of a subset $X \subseteq \mathbb{R}^d$ as the intersection of all subsets $A \in \mathcal{A}^d$ which contain X. In other words, the $\mathcal{A}_d$-hull of X is the unique minimal set $A \in \mathcal{A}_d$ containing X.

We define the *Helly number* of an intersection closed family $\mathcal{A}_d$ as the least positive integer $h = h(\mathcal{A}_d)$ such that all sets of a subfamily $\mathcal{F} \subseteq \mathcal{A}_d$ have a common point whenever every h sets of $\mathcal{F}$ have one. The *Radon number* of $\mathcal{A}_d$ is defined as the least positive integer $r = r(\mathcal{A}_d)$ such that for any set $X \subseteq \mathbb{R}^d$ with at least r points there exists a partition of X into two subsets, the $\mathcal{A}_d$-hulls of which intersect. The *Carathéodory number* of $\mathcal{A}_d$ is defined as the least positive integer $c = c(\mathcal{A}_d)$ such that for any set $X \subseteq \mathbb{R}^d$ and any point $\mathbf{p}$ from the $\mathcal{A}_d$-hull of X, there exists a subset $Y \subseteq X$ of at most c points, the $\mathcal{A}_d$-hull of which also contains $\mathbf{p}$.

A well-known example for an intersection closed family is the family $\mathcal{C}_d$ of convex sets in $\mathbb{R}^d$. Let us recall that a set of points $S \subseteq \mathbb{R}^d$ is called *convex* if for any two points $\mathbf{y}, \mathbf{z} \in S$, all the points on the line segment between

*This research was partially supported by the National Science Foundation (Grant IIS-0118635). The last two authors are also grateful for the partial support of DIMACS.

[†]RUTCOR, Rutgers University, 640 Bartholomew Road, Piscataway, NJ 08854 USA. E-mail:boros@rutcor.rutgers.edu

[‡]RUTCOR, Rutgers University, 640 Bartholomew Road, Piscataway, NJ 08854 USA. E-mail:gurvich@rutcor.rutgers.edu

[§]School of Science, University of Ontario Institute of Technology, 2000 Simcoe St. N., Oshawa, ON L1H 7K4 Canada. E-mail:yingliu@rutcor.rutgers.edu

y and z belong to S. Clearly, the intersection of two convex sets is convex. It is customary to denote the $\mathcal{C}_d$-hull of a set $X \subseteq \mathbb{R}^d$ by $conv(X)$ and call it the *convex hull* of X. The classical theorems of Helly[13], Radon[17], and Carathéodory[5, 6] on convex sets are fundamental theorems of combinatorial geometry and convex analysis, which were discovered in the beginning of twentieth century. Using the notations introduced above, these results can be stated as: $h(\mathcal{C}_d) = d+1$, $r(\mathcal{C}_d) = d+2$, and $c(\mathcal{C}_d) = d+1$.

In this paper we would like to determine the same parameters for another intersection closed family, the family of boxes. A *box* B is an interval $B = box(\mathbf{a}, \mathbf{b}) = \{\mathbf{x} | \mathbf{a} \leq \mathbf{x} \leq \mathbf{b}\} \subseteq \mathbb{R}^d$, where $\mathbf{a} \leq \mathbf{b}$ are vectors in $\mathbb{R}^d$, and the inequalities are componentwise. We denote by $\mathcal{B}_d$ the family of boxes in $\mathbb{R}^d$. It is easy to see that this family is also closed under intersection. Let us denote the $\mathcal{B}_d$-hull of a subset $X \subseteq \mathbb{R}^d$ by $box(X)$ and call it the *box hull* of X. It is easy to see that $box(X) = box(\min X, \max X)$ where $\min X$ and $\max X$ are the componentwise minimum and maximum of the points in X. We call a box B *spanned* by X if $B = box(X)$.

Since two boxes intersect each other if and only if their projections on each axis intersect, and since the set of intervals on the real line has Helly number 2, it is easy to see that the Helly number of boxes in $\mathbb{R}^d$ is $h(\mathcal{B}_d) = 2$. This fact is used for the recognition of intersection graphs of boxes (see [16]).

Determining the Radon and Carathéodory numbers for the family $\mathcal{B}_d$ is not that immediate, and is the subject of this short paper. We show in Section 2.1 that $r(\mathcal{B}_d) = \Theta(\log d)$, and in Section 2.2 that $c(\mathcal{B}_d) = d$. Finally in Section 3 we consider further analogies between convex sets and box families.

1 Radon and Carathéodory numbers for the families of boxes

1.1 Radon number of the family of boxes

Given a finite set X of n points, let us define the ith *level set* of X (for $1 \leq i \leq n$) as the family $\mathcal{X}^i = \{S \mid |S| = i, S \subset X\}$ consisting of all subsets of size i. A set family $\mathcal{F}$ is called *Sperner* if no two sets of it contain one another. It is known that such a system on n points can contain at most $m = \binom{n}{\lfloor \frac{n}{2} \rfloor}$ sets, and that $\mathcal{X}^{\lfloor \frac{n}{2} \rfloor}$ is such a largest Sperner family (see [18]).

Given a box B spanned by the set of points $X = \{\mathbf{p}_1, \mathbf{p}_2, \ldots, \mathbf{p}_l\}$, let us denote by $P_j = (\mathbf{p}_{i_1}, \mathbf{p}_{i_2}, \ldots, \mathbf{p}_{i_l})$ the permutation of these points arranged in the increasing order of their jth coordinates. We shall call P_j the *j-order* of X. For a permutation P of X we write $P = (X_1, X_2)$ if X_1 and X_2 form a partition of X and all elements in X_1 precede all elements of X_2 in P.

Let us finally call a partition $X_1 \cup X_2$ of a subset $X \subseteq \mathbb{R}^d$ a *Radon partition* if $box(X_1) \cap box(X_2) \neq \emptyset$.

Lemma 1. *A partition $X_1 \cup X_2$ of $X \subseteq \mathbb{R}^d$ is not a Radon partition if and only if there exists an axis i such that for the i-order of X we have either $P_i = (X_1, X_2)$ or $P_i = (X_2, X_1)$.*

Proof. Let us observe that two boxes intersect if and only if their projections on each axis intersect. Then $box(X_1) \cap box(X_2) = \emptyset$ implies the existence of an axis i such that the projections of $box(X_1)$ and $box(X_2)$ on axis i do not intersect. This readily implies the claim. □

Now we are ready to prove the following:

Lemma 2. *For positive integers n and $d < \frac{1}{2}\binom{n}{\lfloor \frac{n}{2} \rfloor}$ we have $r(\mathcal{B}_d) \leq n$.*

Proof. To prove the statement we show a somewhat stronger claim, namely that if $d < \frac{1}{2}\binom{n}{\lfloor \frac{n}{2} \rfloor}$ then any set $X \subseteq \mathbb{R}^d$ of n points has a Radon partition into roughly equal cardinality parts.

For this let us consider a subset $X \subseteq \mathbb{R}^d$ and the Sperner family $\mathcal{X}^{\lfloor \frac{n}{2} \rfloor}$ consisting of all subsets of X of cardinality $\lfloor \frac{n}{2} \rfloor$.

Let us first observe that if the box hulls of Q and $X \setminus Q$ are disjoint for a subset $Q \in \mathcal{X}^{\lfloor \frac{n}{2} \rfloor}$, then there exists an axis i, by Lemma 1, for which either $P_i = (Q, X \setminus Q)$ or $P_i = (X \setminus Q, Q)$. Let us call in such a case the set Q a *half segment* of the i-order P_i.

Since an i-order can have at most 2 different half segments, whenever $\binom{n}{\lfloor \frac{n}{2} \rfloor} > 2d$ there must exist a subset $Z \in \mathcal{X}^{\lfloor \frac{n}{2} \rfloor}$ which is not a half segment of any of the i-orders of X, $i = 1, ..., d$. Thus, Z and $X \setminus Z$ provide a Radon partition for X, from which the lemma follows. □

To be able to show a lower bound on the Radon number of box families, we need a technical lemma first. Given a positive integer n, let $k = \lfloor \frac{n}{2} \rfloor$ and let us consider the k-th level family $\mathcal{X}^k$ of an n-element base set X. Let us further order the elements in each of the sets $K \in \mathcal{X}^k$, independently of each other. For such an ordered set $K = \{i_1, i_2, ..., i_k\}$ and a subset $S \subseteq X$, $|S| < k$ we say that S is an *initial segment* in K if $S = \{i_1, ..., i_{|S|}\}$.

Lemma 3. *There exists an ordering of the elements in each of the k-sets $K \in \mathcal{X}^k$ such that every subset $S \subseteq X$, $|S| < k$ is an initial segment in some of these k-sets.*

Proof. Let us consider the subsets of X of size at most k as vertices of a graph G, in which two sets S and S' are connected by an edge if $S \subseteq S'$ and $|S'| = |S| + 1$. In this graph the sets $\mathcal{X}^i$ and $\mathcal{X}^{i+1}$ induce a regular bipartite subgraph in which the vertices in $\mathcal{X}^{i+1}$ have smaller degrees than those in $\mathcal{X}^i$. Then a result of König [14] implies the existence of a matching in this bipartite graph, such that all vertices of $\mathcal{X}^i$ are matched to a vertex in $\mathcal{X}^{i+1}$. Let M_i denote the set of these matching edges, for $i = 1, ..., k-1$.

Then the union of these edges form a set of disjoint paths in G, such that one endpoint of each of these paths belong to $\mathcal{X}^k$, and all sets of size less than k are covered by them.

Let us consider such a path $S_1 \subseteq S_2 \subseteq \cdots \subseteq S_l$ with $|S_l| = k$. Let us then order the elements of S_l by listing first the elements of S_1 (in an arbitrary order), then putting the unique element of $S_2 \setminus S_1$ next, followed by the unique element of $S_3 \setminus S_2$, etc. Let us order the elements of the k-sets which do not belong to such a path arbitrarily.

We claim that the constructed ordering of the elements of k-sets have the stated property: Clearly, any subset S of size smaller than k belong to such a path, the endpoint K of which is a k-set, and by the above construction S appears as an initial segment in K. $\square$

Let us illustrate the above lemma on a small example. Let $n = 6$, $k = 3$, and $X = \{1, 2, 3, 4, 5, 6\}$. By choosing the matchings as in the proof above, we can arrive to the following system of paths, where we write each set in the order induced by these paths, as in the above proof.

$$
\begin{array}{c}
1 \to 12 \to 123 \\
2 \to 23 \to 236 \\
3 \to 35 \to 354 \\
4 \to 41 \to 416 \\
5 \to 52 \to 523 \\
6 \to 65 \to 654 \\
13 \to 134 \\
15 \to 154 \\
16 \to 165 \\
24 \to 241 \\
26 \to 264 \\
34 \to 342 \\
36 \to 365 \\
45 \to 452 \\
46 \to 463
\end{array}
$$

This way we obtained an ordering for 15 of the 20 3-sets of X, such that all subsets of size 2 or 1 appears as an initial segment in one of these. The ordering of the elements of the remaining 5 3-sets can be arbitrary.

Now, we are ready to state a lower bound for the Radon number of box families.

Lemma 4. *For positive integers n and* $d \geq \frac{1}{2}\binom{n}{\lfloor \frac{n}{2} \rfloor}$ *we have* $r(\mathcal{B}_d) \geq n+1$.

Proof. To prove this statement, we need to show that whenever $d \geq \frac{1}{2}\binom{n}{\lfloor \frac{n}{2} \rfloor}$ there exists an n element point set $X \subseteq \mathbb{R}^d$ for which a Radon partition does not exist. In other words, for which for every partition $X_1 \cup X_2 = X$ we have $box(X_1) \cap box(X_2) = \emptyset$. By Lemma 1 this latter is equivalent with the fact that the i-order of X for some axis i must have the form $P_i = (X_1, X_2)$ or $P_i = (X_2, X_1)$.

Thus, to prove the statement it is enough to construct an n-set X of points in $\mathbb{R}^d$ for which any subset $S \subseteq X$ of size at most $k = \lfloor \frac{n}{2} \rfloor$ appears as an initial segment or as a tail segment in one of the i-orders of X, $i = 1, ..., d$.

Let us also observe that given d permutations P_i, $i = 1, ..., d$ it is easy to construct a set of points $X \subseteq \mathbb{R}^d$ for which P_i is the i-th order of X, for $i = 1, ..., d$.

Therefore, to prove the lemma it is enough to construct d permutations of an n-element abstract set X for which every subset $S \subseteq X$ of size at most k appears as an initial or tail segment in one of these permutations.

To construct such permutations, let us consider first a graph G, the vertex set of which is $\mathcal{X}^k$, and in which two k-sets are connected by an edge iff they are disjoint. It is immediate to see that relabelling the elements of X induces an automorphism of this graph, moreover any vertex of G can be transformed to any other vertex by such an automorphism. This transitivity implies that there is a matching in G which is perfect if $|\mathcal{X}^k|$ is even, or missing only one vertex, if $|\mathcal{X}^k|$ is odd. This follows easily by the Edmonds-Gallai structure theorem (see [9, 10, 11]), as stated for instance in Exercise 3.2.5 in [15]. Let M denote the set of edges in such a maximum matching.

Let us also consider an ordering of the elements of the sets of $\mathcal{X}^k$ satisfying the claim of Lemma 3, and let us associate to every edge $e = (K, K')$ of M a permutation P_e of X constructed in the following way: Let us first list the elements of K, in the order obtained by Lemma 3. If n is odd, then let us place next the unique element of $X \setminus (K \cup K')$. Let us finally place the elements of K' in reverse order of the ordering obtained from Lemma 3. Furthermore, if $|\mathcal{X}^k|$ is odd, then to the unique set K^* not covered by the matching M we also associate a permutation P_{K^*} by listing first the elements of K^* in the order obtained by Lemma 3, followed by the rest of the elements in an arbitrary order.

We claim that the set of permutations we obtain in this way satisfy our claim: For an arbitrary subset $S \subseteq X$, $|S| < k$ we have by Lemma 3 a k-set K in which S is an initial segment. If $e = (K, K') \in M$ for some K', then S will be an initial segment in the permutation P_e. If $e' = (K', K) \in M$ for some K', then S will appear as a tail segment in the corresponding permutation $P_{e'}$. Finally, if $K = K^*$, then S is an initial segment in P_{K^*}.

Since the number of these permutations is $\lceil \frac{1}{2} \binom{n}{\lfloor \frac{n}{2} \rfloor} \rceil \le d$, the statement follows by simply adding $d - \lceil \frac{1}{2} \binom{n}{\lfloor \frac{n}{2} \rfloor} \rceil$ arbitrary permutations to this family. □

For example, by Lemma 4, in $\mathbb{R}^3$, there exist four points $\mathbf{p}_1, \mathbf{p}_2, \mathbf{p}_3, \mathbf{p}_4$ such that any partition of them is not a Radon partition. Let us construct these four points by the method of the proof in Lemma 4. Let $X = \{\mathbf{p}_1, \mathbf{p}_2, \mathbf{p}_3, \mathbf{p}_4\}$. Then

$$\begin{aligned} \mathcal{X}^1 &= \{\{\mathbf{p}_1\}, \{\mathbf{p}_2\}, \{\mathbf{p}_3\}, \{\mathbf{p}_4\}\}, \\ \mathcal{X}^2 &= \{\{\mathbf{p}_1\mathbf{p}_2\}, \{\mathbf{p}_1\mathbf{p}_3\}, \{\mathbf{p}_1\mathbf{p}_4\}, \{\mathbf{p}_2\mathbf{p}_3\}, \{\mathbf{p}_2\mathbf{p}_4\}, \{\mathbf{p}_3\mathbf{p}_4\}\}. \end{aligned}$$

We can use the following paths $\mathcal{X}^1 \to \mathcal{X}^2$ as in Lemma 3: $\mathbf{p}_1 \to \mathbf{p}_1\mathbf{p}_2$, $\mathbf{p}_2 \to \mathbf{p}_2\mathbf{p}_3$, $\mathbf{p}_3 \to \mathbf{p}_3\mathbf{p}_1$, and $\mathbf{p}_4 \to \mathbf{p}_4\mathbf{p}_1$. Since $|\mathcal{X}^2| = 6 = 2 \times 3$, we have a matching of size 3 in $\mathcal{X}^2$, and thus we obtain the following three permutations of X: $P_1 = (\mathbf{p}_1\mathbf{p}_2\mathbf{p}_3\mathbf{p}_4)$, $P_2 = (\mathbf{p}_3\mathbf{p}_1\mathbf{p}_4\mathbf{p}_2)$, $P_3 = (\mathbf{p}_4\mathbf{p}_1\mathbf{p}_3\mathbf{p}_2)$. These permutations have the property that for any subset $Y \subset X$, $(Y, X \setminus Y)$ is one of these permutations, after appropriately ordering the elements in Y and separately in $X \setminus Y$. Therefore, from these 3 permutations of X we can construct four points such that their projections on the three axes are in the order of these permutations, and consequently no partition of X is a Radon partition.

As a consequence, by Lemma 2 and Lemma 4 we obtain the following Radon-type theorem for the family of boxes:

Corollary 1. *For any positive integer $n > 2$, and dimension d, if*

$$\frac{1}{2}\binom{n-1}{\lfloor \frac{n-1}{2} \rfloor} \le d < \frac{1}{2}\binom{n}{\lfloor \frac{n}{2} \rfloor}$$

then $r(\mathcal{B}_d) = n$.

Proof. By Lemma 2, if $d < \frac{1}{2}\binom{n}{\lfloor \frac{n}{2} \rfloor}$, then $r(\mathcal{B}_d) \le n$. On the other hand, by Lemma 4 we get that $r(\mathcal{B}_d) \ge n$ whenever $d \ge \frac{1}{2}\binom{n-1}{\lfloor \frac{n-1}{2} \rfloor}$. Therefore, we must have $r(\mathcal{B}_d) = n$. □

Let us remark that there is a one to one correspondence between the values of $r(\mathcal{B}_d)$ and the intervals of integers, as determined by the lower and upper bounds in Theorem 1. Clearly, these intervals form a partition of all positive integers. For example, $r(\mathcal{B}_d) = 3$ for $1 \le d < 2$; $r(\mathcal{B}_d) = 4$ for $2 \le d < 3$; $r(\mathcal{B}_d) = 5$ for $3 \le d < 5$; $r(\mathcal{B}_d) = 6$ for $5 \le d < 10$, etc.

Corollary 2. $r(\mathcal{B}_d) = \Theta(\log d)$.

Proof. This follows by elementary computations from the inequalities in Theorem 1 by using Sterling's formula $n! \approx \sqrt{2\pi n}(\frac{n}{e})^n$. □

1.2 Carathéodory number of the family of boxes

In this section we prove the following statement.

Theorem 1. *Let $X \subseteq \mathbb{R}^d$ be a finite set of cardinality $m \geq d > 1$. Then for any point $\mathbf{x} \in box(X)$, there exists a subset $Y \subset X$ of size at most d such that $\mathbf{x} \in box(Y)$.*

Proof. For any point $\mathbf{x} = (x_1, x_2, \ldots, x_d) \in box(X)$, we need to prove that there exists a box containing it, which is spanned by at most d points of X. We shall prove this statement by construction. A set $Y \subseteq X$ spanning such a box, containing $\mathbf{x}$ can be constructed as follows:

Preprocessing: Choose two points $\mathbf{a} = (a_1, \ldots, a_d)$ and $\mathbf{b} = (b_1, \ldots, b_d)$ from X, such that $a_1 \leq x_1 \leq b_1$. Such points must exist, since we have $\mathbf{x} \in box(X)$. If either $a_2 \leq x_2 \leq b_2$ or $b_2 \leq x_2 \leq a_2$ hold, then we set $Y \leftarrow \{\mathbf{a}, \mathbf{b}\}$. Otherwise, we have both a_2 and b_2 on one side of x_2. Assume, without any loss of generality that $a_2 \leq b_2 \leq x_2$. Since $\mathbf{x} \in box(X)$, there must exist a point $\mathbf{c} = (c_1, \ldots, c_d) \in X$ such that $x_2 \leq c_2$. If $x_1 \leq c_1$, then we set $Y \leftarrow \{\mathbf{a}, \mathbf{c}\}$, and otherwise we set $Y \leftarrow \{\mathbf{b}, \mathbf{c}\}$.

As a result, we have $|Y| = 2$ and $\min_{a \in Y} a_i \leq x_i \leq \max_{a \in Y} a_i$ for $i = 1, 2$.

Main Loop: For $i = 3, 4, ..., d$ we repeat the following:

> If there is no two points $\mathbf{p} = (p_1, \ldots, p_d)$, and $\mathbf{q} = (q_1, \ldots, q_d)$ in Y for which $p_i \leq x_i \leq q_i$, then the ith coordinates of all the points in Y are on the same side of x_i. Suppose for instance, without any loss of generality, that they are all smaller than x_i. Then, because $\mathbf{x} \in box(X)$, there must exist a point $\mathbf{u} = (u_1, \ldots, u_d) \in X$ such that $x_i \leq u_i$. Let us increment the set Y with $\mathbf{u}$ in this case, i.e. $Y \leftarrow Y \cup \{\mathbf{u}\}$.

After each step, for $i = 3, ..., d$, we have a set Y of size at most i (since we add at most one point to Y in every step) such that

$$\min_{\mathbf{a} \in Y} a_j \leq x_j \leq \max_{\mathbf{a} \in Y} a_j \quad \text{for} \quad j = 1, 2, ..., i.$$

Thus, at the end, the a set $Y \subseteq X$ is produced such that $\mathbf{x} \in box(Y)$ and $|Y| \leq d$. □

Let us remark that Theorem 1 was announced in [12] without proof.

Let us also note that this theorem implies that for every finite set $X \subseteq \mathbb{R}^d$ $(d > 1)$ of cardinality $m \geq d$ the box $box(\min X, \max X)$ is the union of a set of boxes, each of which is spanned by at most d points of X, i.e.,

$$box(\min X, \max X) = \bigcup_{\substack{X' \subseteq X \\ |X'| \leq d}} box(\min X', \max X').$$

Even though this statement is quite simple, it maybe useful in data analysis [4].

Theorem 1 above shows that $c(\mathcal{B}_d) \leq d$. The following easy example demonstrates that $c(\mathcal{B}_d) = d$.

Example: Let us consider the set X of d unit vectors in $\mathbb{R}^d$. Then the box hull of X contains the point $\mathbf{e} = (1, 1, \ldots, 1)$, however $\mathbf{e}$ is not contained in the box hull of any proper subset of X.

2 Dual boundedness of box families and convex sets

In this section we consider a further analogy between convex sets and boxes.

Let us consider two sets R and B of points of $\mathbb{R}^d$, called the sets of *red* and *blue* points. A subset $S \subseteq B$ is called *homogeneous* if $box(S) \cap R = \emptyset$. Let us denote by $\mathcal{A}_b = \mathcal{A}_b(B, R)$ the family of maximal homogeneous subsets of B, and let $\alpha_b = |\mathcal{A}_b|$. Similarly, if $box(S) \cap R \neq \emptyset$, then S is called *non-homogeneous*, and we denote by $\mathcal{B}_b = \mathcal{B}_b(B, R)$ the family of minimal non-homogeneous subsets of B, and set $\beta_b = |\mathcal{B}_b|$.

The problem of generating the family $\mathcal{A}_b$ is arising in data analysis (see e.g., [3, 8, 12]). Following the approach of [2], an incrementally efficient generation algorithm can be constructed to produce the sets in $\mathcal{A}_b \cup \mathcal{B}_b$. This procedure is an efficient way to generate $\mathcal{A}_b$ alone if

$$\beta_b \leq poly(|R \cup B|, d, \alpha_b)$$

holds for some polynomial $poly(\cdot)$. If the above inequality holds, the family $\mathcal{A}_b$ is called *dual bounded* (see [2]). Analogously, the family $\mathcal{B}_b$ is called *dual bounded* if $\alpha_b \leq poly(|R \cup B|, d, \beta_b)$ holds for some polynomial $poly(\cdot)$.

Replacing $box(S)$ by $conv(S)$ in the above definitions, we can introduce $\mathcal{A}_c$ and $\mathcal{B}_c$ as the maximal (resp. minimal) subsets of B the convex hull of which is homogeneous (resp. non-homogeneous). It was shown recently in [7] that if $|R| = 1$ and the points in $R \cup B \subseteq \mathbb{R}^d$ are in general position, then $\beta_c \leq (d+1)\alpha_c$. On the other hand, it was shown in [1] that neither $\mathcal{A}_c$ nor $\mathcal{B}_c$ are dual bounded, in general (i.e. if the points are not in general position).

In this paper we show that $\mathcal{A}_b$ and $\mathcal{B}_b$ are not dual bounded, either.

2.1 Dual Boundedness of Box Families

We shall illustrate by an example that β_b cannot be bounded polynomially by α_b.

Let us denote by $\mathbf{e}_j \in \mathbb{R}^d$ the jth unit vector, for $j = 1, ..., d$, and let $\mathbf{e} = \sum_{j=1}^{d} \mathbf{e}_j$.

Example 1: Let us define $R = \{\mathbf{e}\}$ as a one element set, and let $B = \{2\mathbf{e}_j, 3\mathbf{e}_j \mid j = 1, ..., d\}$ consisting of two blue points at value 2 and 3 along each axes, i.e., $|R| = 1$ and $|B| = 2d$.

Then any minimal set of blue points whose box hull contains $\mathbf{e}$ is an arbitrary set of d points of B from d different axes, implying $\beta_b = 2^d$. On the other hand $\alpha_b = d$, since all the blue points on any $d-1$ axes span a homogeneous box.

Thus, $\beta_b = 2^d$ cannot be bound by a polynomial of $\alpha_b = d$ and $|R \cup B| = 2d+1$.

This example is shown for $d = 2$ in Figure 1: The only red point is represented by an empty circle and the four blue points are represented by dark circles. More precisely, the only red point is at $(1,1)$ and the four blue points are $(2,0)$, $(3,0)$, $(0,2)$, and $(0,3)$.

Conversely, we shall demonstrate by an example that $\mathcal{B}_b$ is not dual bounded either.

Example 2: Let us assume that $d = 2k$, consider the points $\mathbf{p}^{il} \in \mathbb{R}^d$, for $i = 1, \ldots, k$ and $l = 1, ..., 4$ defined by $\mathbf{p}^{i1} = \mathbf{e}_{2i-1} + \mathbf{e}$, $\mathbf{p}^{i2} = 3\mathbf{e}_{2i} - \mathbf{e}$, $\mathbf{p}^{i3} = -3\mathbf{e}_{2i-1} + \mathbf{e}$, and $\mathbf{p}^{i4} = -\mathbf{e}_{2i} - \mathbf{e}$, and define $B = \{\mathbf{p}^{il} \mid i = 1, ..., k,\ l = 1, ..., 4\}$.

Let us also consider the points $\mathbf{q}^{il} \in \mathbb{R}^d$, for $i = 1, \ldots, k$ and $l = 1, 2$ defined by $\mathbf{q}^{i1} = \frac{3}{2}(\mathbf{e}_{2i-1} + \mathbf{e}_{2i})$ and $\mathbf{q}^{i2} = -\frac{3}{2}(\mathbf{e}_{2i-1} + \mathbf{e}_{2i})$, and let $R = \{\mathbf{q}^{il} \mid i = 1, ..., k,\ l = 1, 2\}$.

We claim that in this case the minimal sets of non-homogeneous blue point sets are

$$\mathcal{B}_b = \left\{ \{\mathbf{p}^{i1}, \mathbf{p}^{i2}\}, \{\mathbf{p}^{i3}, \mathbf{p}^{i4}\} \mid i = 1, ..., k \right\}$$

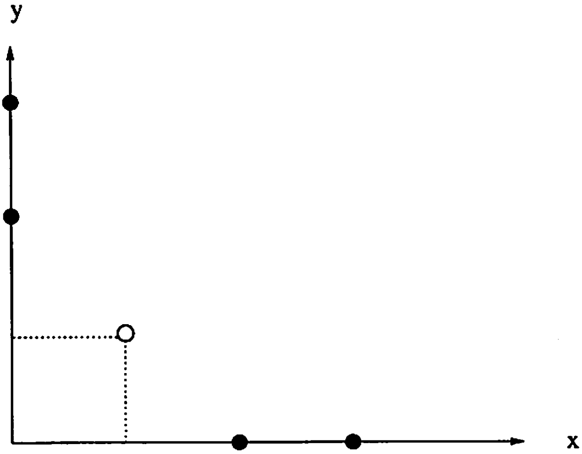

Figure 1: 2 dimensional case of Example 1

since any box hull of blue points containing the red point $\mathbf{q}^{i1}$ has to contain the both points $\mathbf{p}^{i1}$, and $\mathbf{p}^{i2}$, and any box hull of blue points containing the red point $\mathbf{q}^{i2}$ has to contain both blue points $\mathbf{p}^{i3}$, and $\mathbf{p}^{i4}$, for $i = 1, 2, \ldots, k$.

It is also easy to see that in this case the maximal sets of homogeneous blue point sets are

$$\mathcal{A}_b = \left\{ \bigcup_{i=1}^{k} \{\mathbf{p}^{ia_i}, \mathbf{p}^{ib_i}\} \mid (a_i, b_i) \in \{(1,4),(2,3),(1,3),(2,4)\},\ i = 1, ..., k \right\}$$

since, similarly to the above, it is easy to check that the box hull of each of these sets do not contain any red point and that these sets are maximal for this property.

Thus, we have for this example $\beta_b = |\mathcal{B}_b| = d$ and $\alpha_b = |\mathcal{A}_b| = 4^k = 2^d$.

This construction is shown in Figure 2 for $d = 2$.

For $d = 4$ the set of blue points are:

$$\begin{array}{llll} (\ 2,\ 1,\ 1,\ 1), & (\ -1,\ 2,\ -1,\ -1), \\ (\ -2,\ 1,\ 1,\ 1), & (\ -1,\ -2,\ -1,\ -1), \\ (\ 1,\ 1,\ 2,\ 1), & (\ -1,\ -1,\ -1,\ 2), \\ (\ 1,\ 1,\ -2,\ 1), & (\ -1,\ -1,\ -1,\ -2), \end{array}$$

and the set of red points are:

$$\begin{array}{ll} (\ 1.5,\ 1.5,\ 0,\ 0), & (\ -1.5,\ -1.5,\ 0,\ 0), \\ (\ 0,\ 0,\ 1.5,\ 1.5), & (\ 0,\ 0,\ -1.5,\ -1.5). \end{array}$$

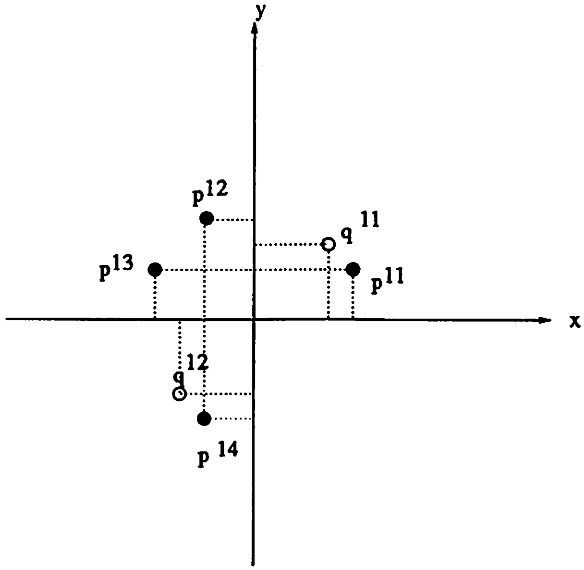

Figure 2: 2 dimensional case of Example 2

References

[1] E. Boros, K. Elbassion, V. Gurvich, and L. Khachiyan. On inequalities for geometric transversal hypergraphs related to their generation. manuscript.

[2] E. Boros, V. Gurvich, L. Khachiyan, and K. Makino. Dual-bounded generating problems: partial and multiple transversals of a hypergraph. *SIAM J. Comput.*, 30:2036–2050, 2001.

[3] E. Boros, P. Hammer, T. Ibaraki, and A. Kogan. Logical analysis of numerical data. *Mathematical Programming*, 79:163–190, 1997.

[4] E. Boros, P. L. Hammer, T. Ibaraki, A. Kogan, E. Mayoraz, and I. Muchnik. An implementation of logical analysis of data. *IEEE Transactions of Knowledge and Data Engineering*, 12(2):292–306, 2000.

[5] C. Carathéodory. Über den variabilitätsbereich der koeffizienten von potenzreihen , die gegebene wertenicht annehmen. *Math. Ann.*, 64:95–115, 1907.

[6] C. Carathéodory. Über den variabilitätsbereich der fourier'schen konstanten von positiven harmonischen funktionen. *Rend. Circ. Mat. Palermo*, 32:193–217, 1911.

[7] R. Collado, A. Kelmans, and D. Krasner. On convex polytops in the plane "containing" and "avoiding" zero. DIMACS Technical Report 2002-33, Rutgers University, 2002.

[8] J. Eckstein, P. Hammer, Y. Liu, M. Nediak, and B. Simeone. The maximum box problem and its application to data analysis. *Computational Optimization and Applications*, 23(3):285–298, 2002.

[9] J. Edmonds. Paths, trees and flowers. *Canad. J. Math.*, 17:449–467, 1965.

[10] T. Gallai. Kritische graphen ii. *Magyar Tud. Akad. Mat. Kutató. Int. Közl.*, 8:373–395, 1963.

[11] T. Gallai. Maximale systeme unabhängiger kanten. *Magyar Tud. Akad. Mat. Kutató. Int. Közl.*, 9:401–413, 1964.

[12] Peter L. Hammer, Y. Liu, B. Simeone, and S Szedmak. Saturated systems of homogeneous boxes and the logical analysis of numerical data. In *Studies in Classification, data analysis, and knowledge organization, Proceedings of GFKL2001 (accepted)*. Springer-Verlag, 2001.

[13] E. Helly. Über mengen konvexer körper m it gemeinschaftlichen punkten. *Jber. Deutsch. Math. Vereinig.*, 32:175–176, 1923.

[14] D. König. Über graphen and ihre anwendung auf determinantentheorie und mengenlehre. *Mathematische Annalen*, 77:453–465, 1916.

[15] L. Lovász and M.D. Plummer. In *Matching Theory*, volume 29. North-Holland, Amsterdam, New York, Oxford, Tokyo, 1986.

[16] T.A. McKee and F.R. McMorris. Topics in intersection graph theory. In *SIAM Monograph on Discrete Math. and Appl.* 1999.

[17] J. Radon. Mengen konvexer körper, die einen gemeinsamen punkt enthalten. *Math. Ann.*, 83:113–115, 1921.

[18] E. Sperner. Ein satz über untermengen einer endlichen menge. *Math. Z.*, 27:544–548, 1928.