Computing the Carathéodory Number of a Point

Sergey Bereg* Mohammadreza Haghpanah*

Abstract

Carathéodory's theorem says that any point in the convex hull of a set P in $\mathbb{R}^d$ is in the convex hull of a subset P' of P such that $|P'| \leq d+1$. For some sets P, the upper bound $d+1$ can be improved. The best upper bound for P is known as the *Carathéodory number* [2, 15, 17]. In this paper, we study a computational problem of finding the smallest set P' for a given set P and a point p. We call the size of this set P', the *Carathéodory number of a point p* or CNP. We show that the problem of deciding the Carathéodory number of a point is NP-hard. Furthermore, we show that the problem is k-LDT-hard. We present two algorithms for computing a smallest set P', if CNP$= 2, 3$.

Bárány [1] generalized Carathéodory's theorem by using $d+1$ sets (colored sets) such that their convex hulls intersect. We introduce a Colorful Carathéodory number of a point or CCNP which can be smaller than $d+1$. Then we extend our results for CNP to CCNP.

1 Introduction

The well-known Carathéodory's theorem deals with the convex hull of a set P, denoted by $\mathsf{conv}(P)$.

Theorem 1 (Carathéodory's theorem [8, 13]) *Let P be a set of points in $\mathbb{R}^d$ and p be a point in $\mathsf{conv}(P)$. Then there is a subset P' of P consisting of at most $d+1$ points such that $p \in \mathsf{conv}(P')$.*

Sometimes there is a set P' of smaller size such that $p \in \mathsf{conv}(P')$, see Figure 1 for an example. We define a *Carathéodory number of a point.*

Definition 2 *For a set of points $P \subset \mathbb{R}^d$ and a point $p \in \mathsf{conv}(P)$,* Carathéodory number of p with respect to P*, denoted by $C(P,p)$, is the smallest integer k such that $p \in \mathsf{conv}(P')$ for a subset $P' \subseteq P$ of size k.*

Carathéodory's theorem guarantees that for every set of points $P \subset \mathbb{R}^d$ and $p \in \mathsf{conv}(P)$, $C(P,p)$ is well-defined and $C(P,p) \leq d+1$. This is related to the well-known concept of the *Carathéodory number of a set* that is the smallest integer k such that, for any point $p \in \mathsf{conv}(P)$, there is a subset P' of P consisting of at most k points such that $p \in \mathsf{conv}(P')$. Equivalently, it can be defined using $C(P,p)$ as follows.

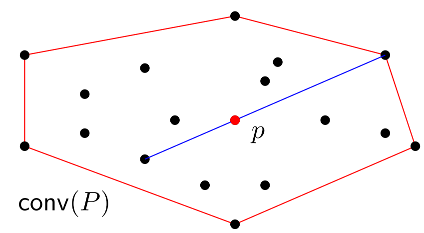

Figure 1: Point $p \in \mathsf{conv}(P)$ with $C(P,p) = 2$.

Definition 3 *For a set of points $P \subset \mathbb{R}^d$,* Carathéodory number of P*, denoted by $C(P)$, is the largest integer k where there exists a point $p \in \mathsf{conv}(P)$ such that $C(P,p) = k$.*

The Carathéodory number of a set is being studied for more than 90 years [15, 17], in a more general setting. The Carathéodory number of any set $P \subset \mathbb{R}^d$ is at most $d+1$ by Carathéodory's theorem. For a compactum $P \subset \mathbb{R}^d$, Bárány and Karasev [2] found sufficient conditions to have Carathéodory number less than $d+1$. Kay and Womble [22] showed a relation between the Carathéodory, Helly, and Radon numbers. Recently, Ito and Lourenço [19] showed an upper bound for the Carathéodory number of a set. Much research has been done on the Carathéodory number for some specific sets. Sierksma [27] studied the Carathéodory number for convex-product-structures, Naldi [25] for Hilbert cones of quadratic forms and binary forms. Bui and Karasev [7] showed the Carathéodory number for arbitrary gauge set K in $\mathbb{R}^d$ is greater than $d-1$. Also, the Carathéodory number for several graph convexities is studied in graph theory [4, 11, 12].

In this paper, we are interested in computing the Carathéodory number of a point. We found the following characterization of the Carathéodory number of a set in $\mathbb{R}^d$. This characterization of the Carathéodory number could be known but we were not able to find it[1]. Recall that the *affine hull* of a set S is the smallest affine

*University of Texas at Dallas, Richardson, TX 75080, USA. {besp,Mohammadreza.Haghpanah}@utdallas.edu

[1] We found that the upper bound of the Carathéodory number of a set follows from Proposition 1.15(ii) [28], see the proof in Section 2.

set containing S (a set A is *affine* if, for any $a, b \in A$, the line passing through a and b is also contained in A). We denote it by $\mathsf{aff}(S)$. The dimension of an affine set S, denoted by $\dim(S)$ is the dimension of its underlying linear subspace.

Proposition 4 *The Carathéodory number of any non-empty set $P \subseteq \mathbb{R}^d$ is equal to* $\dim(\mathsf{aff}(P)) + 1$.

The Carathéodory number of a finite set in $\mathbb{R}^d$ can be computed using Proposition 4. In this paper, we study the computational problem of finding the Carathéodory number of a point with respect to a finite set.

Problem 5 (ComputingCNP)

> **Given** *a set of points P in $\mathbb{R}^d$ and a point $p \in$* $\mathsf{conv}(P)$.
>
> **Compute** *a subset P' of P such that (i) $p \in$* $\mathsf{conv}(P')$ *and (ii) the size of P' is minimized.*

We show that the decision version of ComputingCNP is NP-hard if the dimension d is part of the input. Furthermore, we show that the problem is k-LDT-hard if dimension d is fixed. We present two algorithms for ComputingCNP when $C(P, p) = 2, 3$.

Bárány [1] generalized Carathéodory's theorem by using $d+1$ sets (colored sets) such that their convex hulls intersect. As in [24], we call these sets *color classes* and we call a set of $d+1$ elements, one from each color class, a *colorful choice*.

Theorem 6 (Colorful Carathéodory theorem [1]) *Let $\mathcal{P} = \{P_1, P_2, \ldots, P_{d+1}\}$ be a collection of sets of points in $\mathbb{R}^d$ and p be a point such that $p \in \cap_{i=1}^{d+1} \mathsf{conv}(P_i)$. Then there is a colorful choice P' such that $p \in \mathsf{conv}(P')$.*

It is known that the number of color classes in Theorem 6 cannot be reduced. One may ask whether the number of colors in set P' can be reduced. Sometimes there is a set P' of size smaller than $d+1$ such that $p \in \mathsf{conv}(P')$, see Figure 2 for an example. In this paper, we define a *Colorful Carathéodory number of a point*. We call a set of at most $d+1$ elements, one from color class, a *rainbow*, i.e. a rainbow is a subset of a colorful choice for $\mathcal{P}$. We use notation $[k] = \{1, 2, \ldots, k\}$.

Definition 7 *Let $\mathcal{P} = \{P_1, P_2, \ldots, P_{d+1}\}$ be a collection of sets of points in $\mathbb{R}^d$ and p be a point such that $p \in \cap_{i=1}^{d+1} \mathsf{conv}(P_i)$. The* Colorful Carathéodory number *of p with respect to $\mathcal{P}$, denoted by $CC(\mathcal{P}, p)$, is the smallest size of a rainbow P' for $\mathcal{P}$ such that $p \in \mathsf{conv}(P')$.*

The colorful Carathéodory theorem guarantees that $CC(\mathcal{P}, p) \leq d+1$. In this paper, we also propose to study a new problem of computing $CC(\mathcal{P}, p)$.

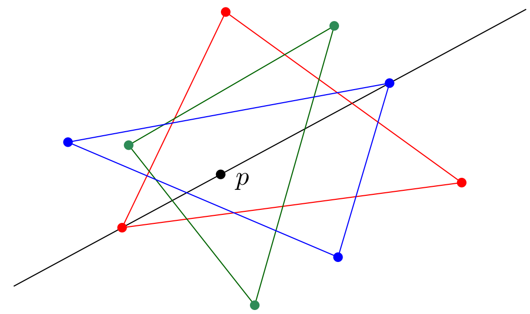

Figure 2: Three sets capturing p in the plane. There is a 2-colorful choice using one red point and one blue point.

Problem 8 (ComputingCCNP)

> **Given** *a collection $\mathcal{P} = \{P_1, P_2, \cdots, P_{d+1}\}$ of sets of points in $\mathbb{R}^d$ and a point $p \in \cap_{i=1}^{d+1} \mathsf{conv}(P_i)$.*
>
> **Compute** *a rainbow P' of the smallest size such that $p \in \mathsf{conv}(P')$.*

Related work. Bárány and Onn [3] describe an approximation algorithm to find a colorful set P' such that point p has distance at most ϵ from $\mathsf{conv}(P')$. Barman [6] showed that a rainbow P' of size $O(\gamma^2/\epsilon^2)$ for $\gamma = \max_{x \in P} \|x\|$ such that the distance between p and $\mathsf{conv}(P')$ is at most ϵ. Mulzer and Stein [24] studied a different approximation using m-colorful sets. A set P' is *m-colorful* if $P_i \cap P' \leq m$ for each color set P_i. Mulzer and Stein [24] give a polynomial algorithm to find a $\lceil \epsilon d \rceil$-colorful choice P' such that $p \in \mathsf{conv}(P')$ for some fixed $\epsilon > 0$. Meunier *et al.* [23] show that the problem of finding a colorful choice is $PPAD$ and PLS.

ComputingCNP is related to the sparse linear regression problem [18, 26] where a $d \times n$ matrix M and a vector $q \in \mathbb{R}^d$ are given and the task is to find a k-sparse vector τ minimizing $\|q - M\tau\|_2$. Natarajan [26] proved NP-hardness of this problem. Har-Peled, Indyk and Mahabadi [18] presented an algorithm with $n^{k-1}S(n, d, \varepsilon)$ space and $n^{k-1}T_Q(n, d, \varepsilon)$ query time where $S(n, d, \varepsilon)$ denotes the preprocessing time and space used by a $(1+\varepsilon)$-ANN (approximate nearest-neighbor) data-structure, and $T_Q(n, d, \varepsilon)$ denotes the query time. Recently, Cardinal and Ooms [9] studied the sparse regression problem and found a $O(n^{k-1} \log^{d-k+2} n)$-time randomized $(1+\varepsilon)$-approximation algorithm for this problem with d and ε constant.

Our results can be summarized as follows.

- We characterize the Carathéodory number of a finite set of distinct points in $\mathbb{R}^d$ (Proposition 4).
- We introduce new problems ComputingCNP and ComputingCCNP for computing $C(P, p)$ and $CC(\mathcal{P}, p)$. We show that DecidingCNP, the decision version of ComputingCNP, is

  - NP-hard (Theorem 10) if the dimension d is part of the input,
  - is k-LDT-hard if dimension d is fixed (Theorem 13).

- We present two algorithms in Section 4 for ComputingCNP when $C(P, p) = 2, 3$.
- Then we extend our results for ComputingCNP to ComputingCCNP in Section 5.

2 Proof of Proposition 4

Let P be a finite set of distinct points in $\mathbb{R}^d$. Theorem 4 states that

$$C(P) = \dim(\mathsf{aff}(P)) + 1.$$

Let $d' = \dim(\mathsf{aff}(P))$.

First, we will prove that $C(P) \geq d'+1$. There exists a set Q of $d'+1$ points of P which are affinely independent. Then $S = \mathsf{conv}(Q)$ is the $(d'+1)$-dimensional simplex. Consider the set A defined as

$$A = \bigcup_{P' \subset P, |P'| = d'} \mathsf{aff}(P').$$

Set A is the union of $\binom{n}{d'}$ sets each of dimension smaller than d'. Therefore $A \cap S \neq S$. For any point $p \in S \setminus A$, we have $C(P, p) \geq d'+1$. Therefore $C(P) \geq d'+1$.

Second, we show that $C(P) \leq d'+1$. This follows from Proposition 1.15(ii) [28] if we write n points of P as a $d \times n$ matrix X.

Proposition. Let $X \in \mathbb{R}^{d\times n}$ and $x \in \mathbb{R}^d$. If $x \in \mathsf{conv}(X)$, then $x \in \mathsf{conv}(X')$ holds for a subset $X' \subseteq X$ of at most $\mathsf{rank}\binom{\mathbb{1}}{X} = \dim(\mathsf{conv}(X)) + 1$ vectors in X.

3 Hardness of ComputingCNP

First, we state the decision problem corresponding to ComputingCNP.

Problem 9 (DecidingCNP)

Given *a set of points $P \subset \mathbb{R}^d$, a point $p \in \mathsf{conv}(P)$ and an integer $k \leq |P|$.*

Decide *whether $C(P, p) \leq k$.*

Observe that DecidingCNP can be solved in polynomial time if dimension d is a constant. We show that it is NP-hard if d is part of the input.

Theorem 10 DecidingCNP *is NP-hard.*

Proof. We reduce the following problem to DecidingCNP.

Problem 11 (ExactCoverBy3-Sets)

Given *a set $X = \{1, 2, 3, \ldots, m\}$ such that 3 is a divisor of m and a collection $\mathcal{S} = \{T_1, T_2, \ldots, T_n\}$ where $T_i \subset X$ and $|T_i| = 3$, for $1 \leq i \leq n$.*

Decide *whether there exists a subset $\mathcal{S}'$ of $\mathcal{S}$ such that $\mathcal{S}'$ is a partition of X, i.e. sets in $\mathcal{S}'$ are disjoint and their union is X.*

Problem ExactCoverBy3-Sets is a variant of ExactCover [21]. This problem is also known to be NP-complete [20].

For an instance of ExactCoverBy3-Sets, a set $X = \{1, 2, 3, \ldots, m\}$ such that 3 is a divisor of m and a collection $\mathcal{S} = \{T_1, T_2, \ldots, T_n\}$ where $T_i \subset X$ and $|T_i| = 3$ for $1 \leq i \leq n$, we construct an instance of DecidingCNP as follows. Set $k = m/3$, $p = (1, 1, \ldots, 1) \in \mathbb{R}^m$ and $P = \{p_1, p_2, \ldots, p_n\}$ where $p_i = (p_{i,1}, p_{i,2}, \ldots, p_{i,m}) \in \mathbb{R}^m$ and

$$p_{i,j} = \begin{cases} k & \text{if } j \in T_i, \\ 0 & \text{otherwise.} \end{cases}$$

We show that there exists an exact cover for set X if and only if there exists a subset $P' \subset P$ of size k where $p \in \mathsf{conv}(P')$.

$\Longrightarrow$) Suppose that $\mathcal{S}'$ is a partition for X. Then for every $j \in X$ there exists unique $T_i \in \mathcal{S}'$ with $j \in T_i$. Set P' as the set of all points p_i such that $T_i \in \mathcal{S}'$. For any $j \in [m]$, there is exactly one point $p_i \in P'$ with $p_{i,j} = k$ and $p_{i',j} = 0$ for all other points $p_{i'} \in P'$. Therefore the j-th coordinate of $\sum_{p_i \in P'} p_i$ is equal to k and $\sum_{p_i \in P'} p_i = kp$. Hence, $p \in \mathsf{conv}(P')$.

$\Longleftarrow$) Suppose that $p \in \mathsf{conv}(P')$, i.e. $\sum_{p_i \in P'} \lambda_i p_i = p$. Then each $\lambda_i \leq \frac{1}{k}$, otherwise some coordinate of $\sum_{p_i \in P'} \lambda_i p_i$ is greater than 1. We have $\sum_{p_i \in P'} \lambda_i = 1$ and each $\lambda_i \geq 0$. Since $|P'| = k$, each λ_i must be equal to $1/k$. Let $\mathcal{S}'$ be the set of all T_i such that $p_i \in P'$. Then, $\mathcal{S}'$ is a partition of X. □

Now, suppose that the dimension d is fixed. We show that DecidingCNP is k-LDT-hard.

Problem 12 (k-LDT)

Given *a set of $A \subset \mathbb{R}$ and a k-variate linear function $\phi(x_1, x_2, \ldots, x_k) = \alpha_0 + \sum_{i=1}^{k} \alpha_i x_i$ where $\alpha_0, \alpha_1, \ldots, \alpha_k \in \mathbb{R}$.*

Decide *whether there exists $x = (x_1, x_2, \ldots, x_k) \in A^k$ where x is a root of ϕ.*

k-LDT-hardness implies k-SUM-hardness and many problems are known to be k-SUM-hard, see for example [5, 16]. Erickson [14] proved any algorithm in r-linear decision tree model for k-LDT problem has $\Omega(n^{\lceil \frac{k}{2} \rceil})$ time complexity.

Theorem 13 DecidingCNP *for a fixed dimension d is k-LDT-hard.*

Proof. We show a linear-time reduction of k-LDT to DecidingCNP. Let $A = \{a_1, a_2, \ldots, a_n\}$ be a set of real numbers and linear function $\phi(x_1, x_2, \ldots, x_k) = \alpha_0 + \sum_{i=0}^{t} \alpha_i x_i$ be an instance of k-LDT problem. An instance of DecidingCNP must contain a set P, a point p, and an integer k' (it could be different from k in k-LDT). We construct an instance of DecidingCNP. We choose $k' = k+1$. We construct set P as follows.

Let $\{e_1, e_2, \ldots, e_{k+1}\}$ be the standard basis of $\mathbb{R}^{k+1}$, i.e. $e_i = (e_{i,1}, e_{i,2}, \ldots,\ e_{i,k+1})$ where $e_{i,j} = 1$ if $j = i$ and $e_{i,j} = 0$ otherwise. For every $x_i \in A$, $1 \le i \le n$, we construct k points in $\mathbb{R}^{k+1}$, $y_{i,j}$, for $1 \le j \le k$, as follows

$$y_{i,j} = e_j + \alpha_j x_i e_{k+1}.$$

We also define $p = (-\alpha_0 e_{k+1} + \sum_{i=1}^{k} e_i)/k$.

$\Longrightarrow$) Suppose there is k integer $i_1, i_2, \ldots, i_k$ where $1 \le i_j \le n$ for $1 \le j \le k$ such that $\phi(x_{i_1}, x_{i_2}, \ldots, x_{i_k}) = 0$. Consider set P' of k points $y_{i_1,1}, y_{i_2,2}, \ldots, y_{i_k,k}$. It implies that $\sum_{j=1}^{k} \lambda y_{i_j,j} = p$ where $\lambda = \frac{1}{k}$. Therefore, $p \in \mathsf{conv}(P')$ and $C(P,p) \le k$.

$\Longleftarrow$) Suppose there exist k pairs of integers $(i_1, j_1), (i_2, j_2), \ldots, (i_k, j_k)$ where $1 \le i_t \le n$, $1 \le j_t \le k$, for $1 \le t \le k$, such that $p \in \mathsf{conv}(p_1, p_2, \ldots, p_k)$ where $p_1 = y_{i_1,j_1}, p_2 = y_{i_2,j_2}, \ldots, p_k y_{i_k,j_k}$. Therefore, there exists λ_t, for $1 \le t \le k$, such that $0 \le \lambda_t \le 1$, for $0 \le t \le k$ and $\sum_{t=1}^{k} \lambda_t = 1$ and

$$\sum_{t=1}^{k} \lambda_t p_t = p. \tag{1}$$

We claim that for every pair of integers t_1 and t_2 where $1 \le t_1 < t_2 \le k$, $j_{t_1} \ne j_{t_2}$, otherwise there exists an integer m such that $1 \le m \le k$ and $m \notin \{j_1, j_2, \ldots, j_k\}$. Then the m-th coordinate of all points $p_1, p_2, \ldots, p_k$ is zero. Then $p_m = 0$ contradicting the choice of p. Therefore, $j_1, j_2, \ldots, j_k$ is a permutation of $1, 2, \ldots, k$. By reordering points $p_1, p_2, \ldots, p_k$ we assume that $j_t = t$ for $1 \le t \le k$.

By taking mth coordinate, $1 \le m \le k$, Equation (1) implies

$$\sum_{t=1}^{k} \lambda_t p_{t,m} = \lambda_m p_{m,m} = \lambda_m = 1/k.$$

Now take $(k+1)$th coordinate in Equation (1)

$$\sum_{t=1}^{k} \lambda_t p_{t,k+1} = \sum_{t=1}^{k} \frac{1}{k} p_{t,k+1} = \sum_{t=1}^{k} \frac{1}{k} \alpha_{j_t} x_{i_t} = \frac{-\alpha_0}{k}.$$

Hence, $\alpha_0 + \sum_{t=1}^{k} \alpha_t x_{i_t} = 0$ which is the solution for k-LDT. □

4 Algorithms for ComputingCNP

We present two algorithms for ComputingCNP when $C(P,p) = 2, 3$.

4.1 $C(P,p) = 2$ in $\mathbb{R}^d$

One can easily decide in $O(n)$ time whether $C(P,p) = 1$. In this section we discuss the problem of deciding whether $C(P,p) = 2$. We assume that dimension $d \ge 2$ is a constant.

Theorem 14 *Let P be a set of n points in $\mathbb{R}^d$ and p be a point in $\mathbb{R}^d$, where $d \ge 2$ is fixed. One can decide whether $C(P,p) = 2$ and find the corresponding set P' in $O(n \log n)$ time which is optimal in the algebraic decision tree model.*

Proof. The task is to compute a subset $P' \subset P$ such that $|P'| \le 2$ and $p \in \mathsf{conv}(P')$ if it exists. We will describe an algorithm and prove the lower bound.

Algorithm. First, we decide whether $C(P,p) = 1$ in $O(n)$ time by searching p in P. Assume that $C(P,p) \ge 2$, i.e. $p_i \ne p$ for all $p_i \in P$. Compute normalized vectors $p'_i = \frac{p_i - p}{\|p_i - p\|}$. We sort points $p'_1, p'_2, \ldots, p'_n$ lexicographically and assume that they are in the lexicographic order, i.e. $p'_1 \preceq p'_2 \preceq \cdots \preceq p'_n$.

Since $C(P,p) \ne 1$, $C(P,p) = 2$ if and only if there exist two points $p_i, p_j \in P$ such that $p = ap_i + (1-a)p_j$ for some $1 \le i < j \le n$ and $0 < a < 1$ (Clearly, if $a = 0$ or $a = 1$ then $C(P,p) = 1$). This equation can be written as

$$0 = a(p_i - p) + (1-a)(p_j - p) \tag{2}$$

$$a(p_i - p) = (a-1)(p_j - p) \tag{3}$$

$$al_i p'_i = (a-1) l_j p'_j, \tag{4}$$

where $l_i = \|p_i - p\|$ and $l_j = \|p_j - p\|$. Since p'_i and p'_j are unit vectors, we have $|al_i| = |(a-1)l_j|$. Note that $al_i > 0$ and $(a-1)l_j < 0$. Equation (4) implies that $p'_i = -p'_j$. Conversely, if $p'_i = -p'_j$ then $p = ap_i + (1-a)p_j$ for $a = l_j/(l_i + l_j)$.

The algorithm performs binary search of $-p'_i$ in the sorted sequence $p'_1, p'_2, \ldots, p'_n$, for each $i \in [n]$. If $-p'_i$ is found then $-p'_i = p'_j$ for some p'_j. Note that j must be not equal to i since $-p'_i = p'_i$ would mean that $p'_i = 0$ but $\|p'_i\| = 1$. The time complexity of the above algorithm is $O(n \log n)$ where $n = |P|$.

Proof of the lower bound. We now prove that the lower bound on the time complexity for the problem of deciding $C(P,p) = 2$ is $\Omega(n \log n)$. We use the 2-SUM problem: Given n numbers, do any two of them sum to zero? Chan, Gasarch and Kruskal [10] proved that solving 2-SUM in algebraic decision tree model takes $\Omega(n \log n)$ time. Let $X = \{x_1, x_2, \ldots, x_n\}$ be set of integers in an instance of 2-SUM.

We construct a set P of n points in $\mathbb{R}^2$. If $d \geq 3$, we can use a 2-dimensional plane in $\mathbb{R}^d$ for the points in P. Assign point $p = (0, 0)$. Consider the map $\mu : \mathbb{R} \to \mathbb{R}^2$ defined as $\mu(x) := (\mathsf{sgn}(x), x)$ where

$$\mathsf{sgn}(x) = \begin{cases} x/|x| & \text{if } x \neq 0, \\ 0 & \text{otherwise.} \end{cases}$$

Find set $P = \mu(X)$ in $O(n)$ time, i.e. $P = \{p_i \mid p_i = \mu(x_i), i \in [n]\}$. To show that the reduction is correct, we prove the following claim. There are distinct integers i, j such that $x_i + x_j = 0$ if and only if there are two points $p_i, p_j \in P$ for which $p \in \mathsf{conv}(\{p_i, p_j\})$.

Suppose that $x_i + x_j = 0$ for some integers i, j. Then either $x_i = x_j = 0$ or $x_i x_j < 0$. If $x_i = x_j = 0$ then $p_i = (0, 0) \in P$, so $p \in \mathsf{conv}(\{p_i\})$. If $x_i x_j < 0$, then we can assume $x_i < 0 < x_j$. Then $0 \in \mathsf{conv}(\{p_i, p_j\})$ since $p_i = -p_j$.

Now suppose that there exists a subset $P' \subset P$ such that $|P'| = 2$ and $p \in \mathsf{conv}(P')$. If $p \in P'$, then $0 \in X$. If $p \notin P'$ and $P' = \{p_i, p_j\}$, then x_i and x_j have opposite signs (since $p_x = 0$). Then the convex combination $p = ap_i + (1-a)p_j$ must have $a = 1-a$ (using x-coordinates in $p = ap_i + (1-a)p_j$, $0 = a \cdot \mathsf{sgn}(x_i) + (1-a) \cdot \mathsf{sgn}(x_j)$). Then $a = 1/2$ and $x_i = -x_j$ (using y-coordinates in $p = ap_i + (1-a)p_j$). □

4.2 $C(P, p) = 3$ in $\mathbb{R}^3$

Theorem 15 *Let P be a set of n points in $\mathbb{R}^3$ and p be a point in $\mathbb{R}^3$. One can decide whether $C(P, p) = 3$ and find the corresponding set P' in $O(n^2 \log n)$ time.*

Let $P = \{p_1, \ldots, p_n\} \subset \mathbb{R}^3$. We denote by $\alpha_{i,j}$ the angle between vectors $\overrightarrow{pp_i}$ and $\overrightarrow{pp_j}$, i.e. $\cos \alpha_{i,j} = \frac{\overrightarrow{pp_i} \cdot \overrightarrow{pp_j}}{\|\overrightarrow{pp_i}\| \cdot \|\overrightarrow{pp_j}\|}$ and $0 \leq \alpha_{i,j} \leq \pi$. Let k be an integer in $\{1, 2, \ldots, n\}$. Consider the plane π_k passing through point p with $\overrightarrow{pp_k}$ is its normal vector. Let q_i be the projection of p_i on π_k, see Fig. 3. We apply the following algorithm.

Input: $p \in \mathbb{R}^3$ and $P = \{p_1, \ldots, p_n\} \subset \mathbb{R}^3$ such that $p \in \mathsf{conv}(P)$.

Output: Decide if $C(P, p) = 3$. If $C(P, p) = 3$, compute a subset $S \subset P$ such that $|S| = 3$ and $p \in \mathsf{conv}(S)$.

1. Check if $C(P, p) = 1$ or $C(P, p) = 2$ from Section 4.1. Stop if $C(P, p) \leq 2$.
2. For each point $p_k \in P$ do the following:
3. Compute plane π_k (it can be computed since $p \neq p_k$; otherwise $C(P, p) = 1$). Compute points q_i for all $i \in \{1, 2, \ldots, n\}, i \neq k$, see Fig. 3.
4. Compute set P_k as follows. Initialize $P_k = P \setminus \{p_k\}$, then prune P_k by repeating the following step.

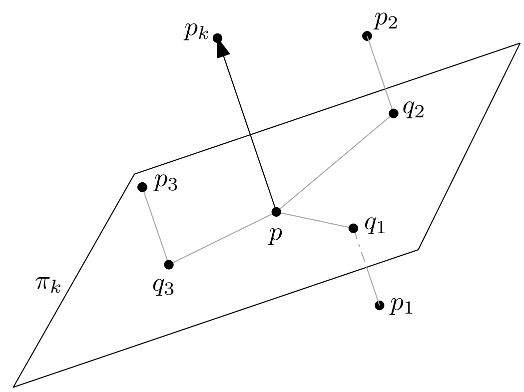

Figure 3: Plane π_k is orthogonal to vector $\overrightarrow{pp_k}$. Point $q_i, i = 1, 2, 3$ is the projection of point p_i onto the plane π_k.

5. **The pruning step**. Remove a point p_i from P_k, if there exists another point p_j in P_k such that
   (a) Vectors $\overrightarrow{pq_i}$ and $\overrightarrow{pq_j}$ have same direction and
   (b) $\alpha_{i,k} < \alpha_{j,k}$ (if $\alpha_{i,k} = \alpha_{j,k}$ remove either p_i or p_j from P_k).
6. Compute $Q_k = \{q_i \mid p_i \in P_k\}$. For each point $q_i \in Q_k$, use the binary search for $-q_i$ in Q_k as in the algorithm from Section 4.1. Suppose $C(Q_k, p) = 2$ and a set $Q' = \{q_i, q_j\}$ is found such that $p \in \mathsf{conv}(q_i, q_j)$. Check if $p \in \mathsf{conv}(\{p_i, p_j, p_k\})$ in $O(1)$ time. If $p \in \mathsf{conv}(\{p_i, p_j, p_k\})$ then output the solution $P' = \{p_i, p_j, p_k\}$. If $p \notin \mathsf{conv}(\{p_i, p_j, p_k\})$, check next point q_i in the loop.
7. If a solution is not found in Step 6, then there is no solution for $C(P, p) = 3$, so $C(P, p) = 4$.

First, we justify the pruning step.

Lemma 16 *Suppose $p \in \mathsf{conv}(\{p_i, p_j, p_k\})$ for some points $p_i, p_j, p_k \in \mathbb{R}^3$. If p_i or p_j is removed in the pruning step for p_k then there exist $p_{i'}, p_{j'} \in P_k$ such that $p \in \mathsf{conv}(\{p_{i'}, p_{j'}, p_k\})$.*

Proof. It suffices to prove the lemma if only one point from $\{p_i, p_j\}$ is removed in the pruning step. If both of them are removed, the argument can be used twice, see Fig. 4(b) for an example.

Suppose that point p_i is removed in the pruning step for p_k. Then there exists another point $p_{i'}$ in P such that

1. Vectors $\overrightarrow{pq_i}$ and $\overrightarrow{pq_{i'}}$ have same direction and
2. $\alpha_{i,k} \leq \alpha_{i',k}$

Then points p, p_i, p_j, p_k are coplanar. Without loss of generality we can assume $p_i, p_{i'}, p_j, p_k \in \mathbb{R}^2$, p_k is on

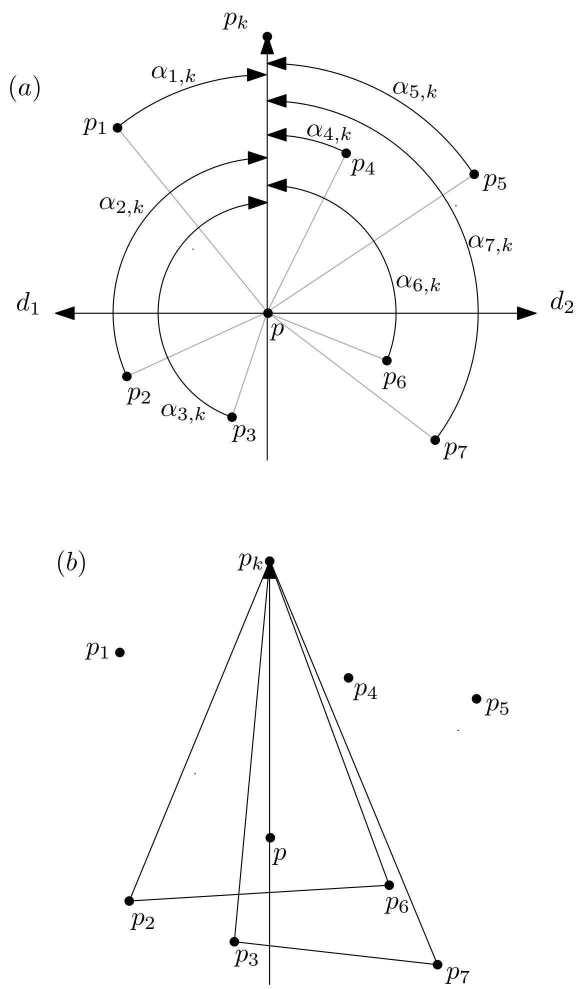

Figure 4: (a) The pruning step. Vectors $\overrightarrow{pq_i}, i = 1, 2, 3$ have the same direction d_1. Points p_1 and p_2 will be pruned based on α-angles. Vectors $\overrightarrow{pq_i}, i = 4, 5, 6, 7$ have the same direction $d_2 = -d_1$. Points $p_i, i = 4, 5, 6$ will be pruned based on α-angles. (b) Point $p \in \mathsf{conv}(\{p_k, p_2, p_6\})$ before the pruning for p_k and it is in $\mathsf{conv}(\{p_k, p_3, p_7\})$ after the pruning for p_k.

the y-axis and p is at the origin. We can assume that $x(p_i), x(p_{i'}) > 0$ and $x(p_j) < 0$. The necessary and sufficient condition for $p \in \mathsf{conv}(\{p_i, p_j, p_k\})$ is $\alpha_{i,k} + \alpha_{j,k} > \pi$. Since $\alpha_{i,k} \leq \alpha_{i',k}$, we have $\alpha_{i',k} + \alpha_{j,k} > \pi$. Therefore $p \in \mathsf{conv}(\{p_{i'}, p_j, p_k\})$ and the lemma follows. □

Time Complexity. Plane π_k is computed in $O(1)$ time. Projection of P onto π_k takes $O(n)$ time. The pruning step can be done in $O(n \log n)$ time by maintaining the sorted order of points q_i by the direction. Finally, Step 6 takes $O(n \log n)$ time since binary search takes $O(\log n)$ time and it is done for every point in Q_k. Therefore, the processing of p_k takes $O(n \log n)$ time and the total time is $O(n^2 \log n)$.

5 Hardness and Algorithms for ComputingCCNP

In this section we show that our results for ComputingCNP can be extended directly to ComputingCCNP.

5.1 Hardness

We show that DecidingCCNP (i.e. deciding if $CC(P, p) \leq k$), the decision version of ComputingCCNP, is NP-hard. There is a natural reduction from DecidingCNP problem to DecidingCCNP problem. Consider an instance of DecidingCNP, i.e. a set of points P in $\mathbb{R}^d$, a point $p \in \mathsf{conv}(P)$ and an integer k. We construct an instance of DecidingCCNP by taking $d+1$ copies of P, the color classes $\mathcal{P} = \{P_1, \ldots, P_{d+1}\}$ and by using the same point p and integer k. Clearly, this reduction can be computed in polynomial time. It remains to prove that $C(P, p) \leq k$ if and only if $CC(\mathcal{P}, p) \leq k$. If $C(P, p) \leq k$ then there exists a subset $P' = \{p_1, p_2, \ldots, p_k\}$ of P such that $p \in \mathsf{conv}(P')$. Then $CC(\mathcal{P}, p) \leq k$ by selecting p_i from color set P_i (i.e. P' is a rainbow for $\mathcal{P}$). Similarly, $CC(\mathcal{P}, p) \leq k$ implies $C(P, p) \leq k$. Therefore DecidingCCNP is NP-hard.

Similarly, one can prove that DecidingCCNP is k-LDT-hard if dimension d is fixed (we omit details due to lack of space).

5.2 $CC(\mathcal{P}, p) = 2$ in $\mathbb{R}^d$

We show that the algorithm from Section 4.1 can be modified for deciding if $CC(\mathcal{P}, p) = 2$ in $\mathbb{R}^d$ and computing the corresponding rainbow. In this problem, we have $d+1$ color classes, and they can be processed as follows. We normalize the vectors (of all colors) again, but this time there could be equal normal vectors of different colors. We store one vector for them and the list of their colors. Then the binary search is modified to select a vector of different color from the list.

The time complexity of this algorithm is $O(n \log n)$ where $n = \sum_{i=1}^{d+1} |P_i|$.

5.3 $CC(P, p) = 3$ in $\mathbb{R}^3$

We briefly (due to lack of space) show that the algorithm from Section 4.2 can be modified for deciding if $CC(\mathcal{P}, p) = 3$ in $\mathbb{R}^3$ and computing the corresponding rainbow. In step 2, we select p_k from $\cup P_i$. In step 3, we use the same colors for projected points. In the pruning step, if there are more than two points q_i and q_j with distinct colors with the same direction of $\overrightarrow{pq_i}$ and $\overrightarrow{pq_j}$, we store the two with the largest α-angles. In steps 6, we apply the algorithm for deciding $CC(\mathcal{P}, p) = 2$ instead of deciding $C(P, p) = 2$. The total time complexity of the algorithm is $O(n^2 \log n)$ where $n = \sum_{i=1}^{d+1} |P_i|$.

References

[1] I. Bárány. A generalization of Carathéodory's theorem. *Discrete Mathematics*, 40(2-3):141–152, 1982.

[2] I. Bárány and R. Karasev. Notes about the Carathéodory number. *Discrete & Computational Geometry*, 48(3):783–792, 2012.

[3] I. Bárány and S. Onn. Colourful linear programming and its relatives. *Mathematics of Operations Research*, 22(3):550–567, 1997.

[4] R. M. Barbosa, E. M. Coelho, M. C. Dourado, D. Rautenbach, and J. L. Szwarcfiter. On the Carathéodory number for the convexity of paths of order three. *SIAM Journal on Discrete Mathematics*, 26(3):929–939, 2012.

[5] G. Barequet and S. Har-Peled. Polygon containment and translational in-hausdorff-distance between segment sets are 3sum-hard. *International Journal of Computational Geometry & Applications*, 11(04):465–474, 2001.

[6] S. Barman. Approximating Nash equilibria and dense bipartite subgraphs via an approximate version of Caratheodory's theorem. In *Proc. 47th ACM symposium on Theory of computing*, pages 361–369. ACM, 2015.

[7] V. Bui and R. Karasev. On the Carathéodory number for strong convexity. *arXiv preprint arXiv:1806.10937*, 2018.

[8] C. Carathéodory. Über den variabilitätsbereich der koeffizienten von potenzreihen, die gegebene werte nicht annehmen. *Mathematische Annalen*, 64(1):95–115, 1907.

[9] J. Cardinal and A. Ooms. Sparse regression via range counting. In *17th Scandinavian Symposium and Workshops on Algorithm Theory, SWAT 2020, June 22-24, 2020, Tórshavn, Faroe Islands*, pages 20:1–20:17, 2020.

[10] A. C. Chan, W. I. Gasarch, and C. P. Kruskal. Refined upper and lower bounds for 2-SUM. `https://www.researchgate.net/publication/228515076_Refined_Upper_and_Lower_Bounds_for_2-SUM`.

[11] E. M. Coelho, M. C. Dourado, D. Rautenbach, and J. L. Szwarcfiter. The Carathéodory number of the P3 convexity of chordal graphs. *Discrete Applied Mathematics*, 172:104–108, 2014.

[12] M. C. Dourado, D. Rautenbach, V. F. Dos Santos, P. M. Schäfer, and J. L. Szwarcfiter. On the Carathéodory number of interval and graph convexities. *Theoretical Computer Science*, 510:127–135, 2013.

[13] J. Eckhoff. Helly, Radon, and Carathéodory type theorems. In *Handbook of convex geometry*, pages 389–448. Elsevier, 1993.

[14] J. Erickson. Lower bounds for linear satisfiability problems. In *SODA*, pages 388–395, 1995.

[15] W. Fenchel. Über krümmung und windung geschlossener raumkurven. *Mathematische Annalen*, 101(1):238–252, 1929.

[16] A. Gajentaan and M. H. Overmars. On a class of $o(n^2)$ problems in computational geometry. *Computational geometry*, 5(3):165–185, 1995.

[17] O. Hanner and H. Rådström. A generalization of a theorem of Fenchel. *Proceedings of the American Mathematical Society*, 2(4):589–593, 1951.

[18] S. Har-Peled, P. Indyk, and S. Mahabadi. Approximate sparse linear regression. In *45th International Colloquium on Automata, Languages, and Programming, ICALP 2018, July 9-13, 2018, Prague, Czech Republic*, pages 77:1–77:14, 2018.

[19] M. Ito and B. F. Lourenço. A bound on the Carathéodory number. *Linear Algebra and its Applications*, 532:347–363, 2017.

[20] D. S. Johnson and M. R. Garey. *Computers and intractability: A guide to the theory of NP-completeness*. WH Freeman, 1979.

[21] R. M. Karp. Reducibility among combinatorial problems. In *Proc. of a Symposium on the Complexity of Computer Computations*, pages 85–103. Springer, 1972.

[22] D. Kay and E. W. Womble. Axiomatic convexity theory and relationships between the Carathéodory, Helly, and Radon numbers. *Pacific Journal of Mathematics*, 38(2):471–485, 1971.

[23] F. Meunier, W. Mulzer, P. Sarrabezolles, and Y. Stein. The rainbow at the end of the line—a ppad formulation of the colorful carathéodory theorem with applications. In *Proceedings of the Twenty-Eighth Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 1342–1351. SIAM, 2017.

[24] W. Mulzer and Y. Stein. Computational aspects of the colorful Carathéodory theorem. *Discrete & Computational Geometry*, 60(3):720–755, 2018.

[25] S. Naldi. Nonnegative polynomials and their Carathéodory number. *Discrete & Computational Geometry*, 51(3):559–568, 2014.

[26] B. K. Natarajan. Sparse approximate solutions to linear systems. *SIAM J. Comput.*, 24(2):227–234, 1995.

[27] G. Sierksma. Carathéodory and Helly-numbers of convex-product-structures. *Pacific Journal of Mathematics*, 61(1):275–282, 1975.

[28] G. M. Ziegler. *Lectures on Polytopes*, volume 152 of *Graduate Texts in Mathematics*. Springer-Verlag, Heidelberg, 1994.