

# Oja Centers and Centers of Gravity<sup>☆</sup>

Dan Chen

*School of Computer Science, Carleton University*

Olivier Devillers

*INRIA Sophia Antipolis Méditerranée, France.*

John Iacono

*Department of Computer and Information Science, Polytechnic University*

Stefan Langerman

*Département d'Informatique, Université Libre de Bruxelles*

Pat Morin

*School of Computer Science, Carleton University*

---

## Abstract

*Oja depth* (Oja 1983) is a generalization of the median to multivariate data that measures the centrality of a point  $x$  with respect to a set  $S$  of points in such a way that points with smaller Oja depth are more central with respect to  $S$ . Two relationships involving Oja depth and centers of mass are presented. The first is a form of Centerpoint Theorem which shows that the center of mass of the convex hull of a point set has low Oja depth. The second is an approximation result which shows that the center of mass of a point set approximates a point of minimum Oja depth.

---

<sup>☆</sup>This work was initiated during the Workshop on Computational Geometry 2006 in Caldes de Malavella. The authors wish to thank Ferran Hurtado and the organizers for the opportunity of working on this topic.

*Email addresses:* `dchen4@connect.carleton.ca` (Dan Chen), `Olivier.Devillers@sophia.inria.fr` (Olivier Devillers), `jiacono@poly.edu` (John Iacono), `stefan.langerman@ulb.ac.be` (Stefan Langerman), `morin@scs.carleton.ca` (Pat Morin)

## 1. Introduction

Given a set  $S$  of  $n$  points in  $\mathbb{R}^d$ , the *Oja depth* [9] of a point  $x \in \mathbb{R}^d$  is

$$d(x, S) = \sum_{y_1, \dots, y_d \in \binom{S}{d}} v(x, y_1, \dots, y_d) ,$$

where  $v(p_1, \dots, p_{d+1})$  denotes the volume of the simplex whose vertices are  $p_1 \dots p_{d+1}$ .<sup>1</sup> A point in  $\mathbb{R}^d$  with the minimum Oja depth is called an *Oja center*.

### 1.1. New Results

In this paper we consider relationships between centers of mass of certain sets and Oja depth. The *center of mass* of a finite point set  $S \subset \mathbb{R}^d$  is the average of those points,

$$c(S) = |S|^{-1} \sum_{x \in S} x .$$

If  $P \subset \mathbb{R}^d$  is a bounded object of non-zero volume, the center of mass of  $P$  is

$$c(P) = \frac{\int_{x \in P} x \, dx}{v(P)} .$$

In this paper, we prove the following results about the Oja depth of an  $n$  point set  $S$ , whose convex hull  $A$  has unit volume and that has an Oja center  $x$ :

$$d(c(A), S) \leq \binom{n}{d} / (d + 1) , \tag{1}$$

$$d(c(S), S) \leq (d + 1) d(x, S) . \tag{2}$$

The bound in (1) is not known to be tight. The bound in (2) is tight, up to a lower-order term, for some point sets  $S$ .

---

<sup>1</sup>In Oja's original definition, the sum is normalized by dividing by  $\binom{|S|}{d}$ . We omit this here since it changes none of our results and clutters our formulas.

### 1.2. Related Results

Our first result, (1), is a form of *Centerpoint Theorem* that upper-bounds the Oja depth of  $c(A)$ , and hence also the Oja depth of  $x$ , in terms of the volume of the convex hull of  $S$ . Previously, centerpoint theorems were known for other depth functions such as Tukey depth [7, 10, 12] and simplicial depth [2, 3, 6]. To the best of our knowledge, this is the first such result for Oja depth.

Our next result, (2), can be viewed in two ways:

1. The first is a linear-time algorithm to find a point whose depth is a constant factor approximation of the depth of the Oja center. In 1-d, Oja depth is minimized by the median, which can be found in  $O(n)$  time. However, in 2-d, the best known algorithm for minimizing Oja depth exactly takes  $O(n \log^3 n)$  time [1]. Approximation algorithms for minimizing Oja depth, based on uniform grids and sampling from  $\binom{S}{d}$ , are given by Ronkainen, Oja, and Orponen [11]; this algorithm and several others are implemented in the R-package, OjaNP [5]. However, in pathological cases, their approximation algorithm is not guaranteed (or even likely) to find a point that closely approximates the Oja center, either in terms of distance or in terms of its Oja depth.<sup>2</sup>
2. Another view of (2) is that it gives insight into the Oja depth function and the Oja center. In some sense, it tells us that the Oja center is not terribly different from the center of mass of  $S$ , since the center of mass of  $S$  minimizes, to within a constant factor, the Oja depth function.

## 2. Oja Center and Mass Center of A

In this section, we relate the Oja depth of the center of mass of the convex hull of  $S$  to the volume of the convex hull of  $S$ . Throughout this section,  $A$  denotes the convex hull of  $S$  and we assume, without loss of generality, that  $v(A) = 1$ .

Our upper-bound is based on the following *central identity*: For any disjoint sets  $X, Y \subseteq \mathbb{R}^d$  with  $v(X \cup Y) > 0$ ,

$$c(X \cup Y) = \frac{v(X) c(X) + v(Y) c(Y)}{v(X \cup Y)}.$$

---

<sup>2</sup>This follows from the fact that the value of the Oja depth function and the location of the Oja center can be arbitrarily different for two sets  $S_1$  and  $S_2$  that differ in only  $d$  points [8].

We first give an inductive proof of our result for point sets in  $\mathbb{R}^2$ , and then give a proof for point sets in  $\mathbb{R}^d$  that uses tools from convex geometry.

### 2.1. An Upper Bound in $\mathbb{R}^2$

The following result shows that, for a convex polygon  $E$  (e.g.,  $E = A$ ), it is not possible to form an overly-large triangle that has  $c(E)$  as one of its vertices:

**Lemma 1.** *Let  $A$  be a convex polygon and let  $p_1$  and  $p_2$  be any two points in  $A$ . Then  $v(p_1, p_2, c(A)) \leq v(A)/3$ .*

*Proof.* Assume, without loss of generality, that  $p_1p_2$  is horizontal and that  $c(A)$  is above  $p_1p_2$ . We may assume that  $p_1p_2$  is edge of  $A$  since, otherwise, we can remove the part of  $A$  that is below  $p_1p_2$ . This decreases  $v(A)$  and moves  $c(A)$  further away from the segment  $p_1p_2$ , which increases  $v(c(A), p_1, p_2)$ .

The proof is by induction on the number of vertices of  $A$ . If  $A$  is a triangle then one can easily verify the result. Therefore, assume  $A$  has  $n \geq 4$  vertices. Consider an edge  $ab$  of  $A$  where  $a \neq p_2$  is adjacent to  $p_1$  and let  $c \neq a$  be adjacent to  $b$  (see Figure 1). Since  $A$  has 4 or more vertices, we may assume that the  $y$ -coordinate of  $b$  is not smaller than the  $y$ -coordinate of  $a$ . Otherwise we can reverse the roles of  $p_1$  and  $p_2$  and redefine  $a$  and  $b$  with respect to the new  $p_1$ .

Draw a ray  $r$  whose origin is at  $p_1$  and such that the triangle  $t_1$  supported by  $p_1a$ ,  $ab$  and  $r$  and the triangle  $t_2$  supported by  $r$ ,  $ab$ , and the line through  $bc$  have the same area. Such a ray is guaranteed to exist by a standard continuity argument that starts with  $r$  containing  $a$  and rotates about  $p_1$  until  $r$  contains  $b$ .

Now, convert  $A$  into a polygon  $A'$  by removing  $t_1$  and adding  $t_2$ . This does not change the area of  $A$ . Furthermore, since  $t_1$  and  $t_2$  are separated by a horizontal line, with  $t_2$  above  $t_1$ , this implies that the  $c(A')$  has a larger  $y$ -coordinate than  $c(A)$ , so  $v(c(A), p_1, p_2) \leq v(c(A'), p_1, p_2)$ . Note, also, that  $A'$  has  $n - 1$  vertices so, by induction,

$$v(c(A), p_1, p_2) \leq v(c(A'), p_1, p_2) \leq v(A')/3 = v(A)/3 ,$$

completing the proof. □

**Theorem 1.** *Let  $S$  be a set of points in  $\mathbb{R}^2$  whose convex hull,  $A$ , has unit area. Then  $d(c(A), S) \leq n^2/6$ .*

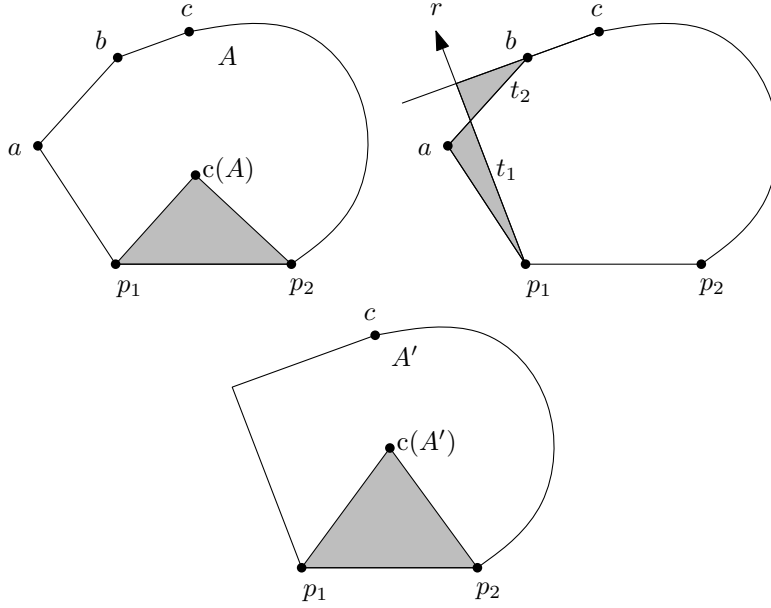


Figure 1: The proof of Lemma 1.

*Proof.* Applying Lemma 1, we immediately get

$$d(c(A), S) = \sum_{p_1, p_2 \in \binom{S}{2}} v(p_1, p_2, c(A)) \leq \binom{n}{2} / 3 < \frac{n^2}{6} ,$$

as required. □

## 2.2. An Upper Bound in $\mathbb{R}^d$

In this section, we extend our bound to  $\mathbb{R}^d$ . However, it seems difficult to apply induction to polytopes in  $\mathbb{R}^d$ , so we resort to some tools from convex geometry.

Let  $A$  be a convex body in  $\mathbb{R}^d$ , where  $d \geq 2$ . Suppose  $A$  lies between parallel hyperplanes  $x_1 = a$  and  $x_1 = b$ , where  $a < b$ . For each  $x$  with  $a \leq x \leq b$ , let  $A_x$  be the intersection of  $A$  with the hyperplane  $x_1 = x$ , and define  $r_x$  by the equation

$$\omega_{d-1} r_x^{d-1} = v_{d-1}(A_x) ,$$

where  $v_{d-1}(X)$  denotes the  $(d-1)$ -dimensional volume of  $X$  and  $\omega_{d-1}$  is the  $(d-1)$ -dimensional volume of the unit  $(d-1)$ -ball. In this way,  $r_x$  is the

radius of a  $(d-1)$ -ball whose  $(d-1)$ -dimensional volume is the same as that of  $A_x$ . For each  $a \leq x \leq b$ , let  $C_x$  be defined by the equation

$$C_x = \{(x, x_2, \dots, x_d) : x_2^2 + \dots + x_d^2 \leq r_x^2\}.$$

Then the set

$$C = \cup\{C_x : a \leq x \leq b\}$$

is called the *Schwarz rotation-symmetrized* of  $A$  in the  $x_1$ -axis. For example, in  $\mathbb{R}^3$ ,  $C$  is a stack of disks perpendicular to, and centered on, the  $x$ -axis. Each disk has the same area as the corresponding slice of  $A$ .

**Theorem 2** (Webster [13]). *Let  $A$  be a convex body in  $\mathbb{R}^d$  ( $d \geq 2$ ) whose Schwarz rotation-symmetrized in the  $x_1$ -axis is  $C$ . Then  $C$  is a convex body having the same volume as  $A$ .*

**Lemma 2.** *Let  $P$  be a convex polytope in  $\mathbb{R}^d$  and let  $p_1, \dots, p_d$  be any  $d$  points in  $P$ . Then  $v(p_1, \dots, p_d, c(P)) \leq v(P)/(d+1)$ .*

*Proof.* Our proof of this lemma is based on the fact that the center of mass of a cone or pyramid with height  $\ell$  in  $\mathbb{R}^d$  has distance  $\frac{\ell}{d+1}$  from the base. To see this, we assume that base of the cone or pyramid contains the origin and is perpendicular to the  $x_1$ -axis. According to central identity, the  $x_1$ -value of the center of mass is

$$\frac{\int_0^\ell x_1 a(\ell - x_1)^{d-1} dx_1}{\frac{a\ell^{d-1}\ell}{d}},$$

where  $a$  is the constant involved in the computing the  $(d-1)$ -dimensional volume of the base. Solving this integration we get  $\frac{\ell}{d+1}$  [4].

Let  $h$  be a hyperplane that contains  $p_1, \dots, p_d$ . If  $c(P)$  is in  $h$ , or  $p_1, \dots, p_d$  are not in general position,  $v(p_1, \dots, p_d, c(P)) = 0$  and there is nothing to prove. Otherwise, rotate  $P$  to make  $h$  perpendicular to the  $x_1$ -axis with  $c(P)$  above  $h$ . If there is any volume of  $P$  below  $h$ , we can cut that part off from  $P$  to obtain a new polytope  $P'$ . The volume of  $P'$  will be less than that of  $P$ , and its center of mass  $c(P')$  will be above  $c(P)$ . In this way, if  $(P, p_1, \dots, p_d)$  is a counterexample to the lemma, then so is  $(P', p_1, \dots, p_d)$ .

The face,  $B$ , of  $P'$  in  $h$  contains  $p_1, \dots, p_d$ . If  $P'$  is a pyramid, we have

$$v(p_1, \dots, p_d, c(P')) \leq v(B, c(P')) = v(P')/(d+1) \leq v(P)/(d+1) .$$

If  $P'$  is not a pyramid, let  $q$  be a point on the  $x_1$  axis, above  $h$ , and such that the pyramid  $D$  with  $B$  as base and  $q$  as apex has the same volume as

$P'$ . Let  $C$  be the Schwarz rotation-symmetral of  $P'$  in the  $x_1$ -axis, and  $R$  be that of  $D$  (see Figure 2). Note that  $R$  is a conic pyramid. By Theorem 2,  $C$  is convex and  $v(C) = v(P') = v(R)$ . Let  $c$  be the intersection of the surfaces of  $C$  and  $R$  above  $B$ . Note that the surface of  $R$  is bounded by a collection of lines that pass through  $q$ . By convexity, each of these lines intersects  $C$  in at most 2 points. One of these points has the same  $x_1$ -coordinate as  $B$  and the other points lie on the boundary of a  $(d-1)$ -ball  $c$ .

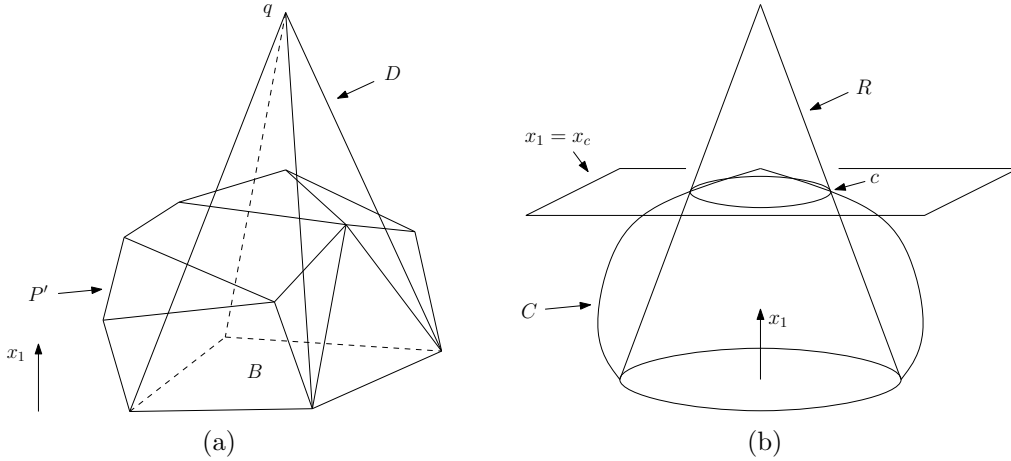


Figure 2: The Schwarz rotation-symmetral of  $P'$  and  $D$

Let  $x_c$  be the  $x_1$ -coordinate of  $c$ . Since  $C$  is convex and a rotation symmetral,  $C \setminus R$  is below the hyperplane  $x_1 = x_c$ . Furthermore,  $R \setminus C$  is above the hyperplane  $x_1 = x_c$ . By the central identity, we have

$$c(R) = \frac{v(R \cap C) c(R \cap C) + v(R \setminus C) c(R \setminus C)}{v(R)},$$

and

$$c(C) = \frac{v(R \cap C) c(R \cap C) + v(C \setminus R) c(C \setminus R)}{v(C)}.$$

Since  $c(R \setminus C)$  is above  $c(C \setminus R)$ ,  $c(R)$  is above  $c(C)$ .

The centers of mass of  $P'$  and  $C$  have the same  $x_1$  value because, in the Schwarz rotation-symmetral,  $C_x$  has the same  $x_1$  value as  $A_x$ . So do the centers of mass of  $D$  and  $R$ . Since  $D$  is a pyramid, the convex hull of  $p_1, \dots, p_d$  is contained in  $B$ ,  $c(P)$  is below  $c(P')$ , and  $c(P')$  is below  $c(D)$ , we have

$$v(p_1, \dots, p_d, c(P')) \leq v(P')/(d+1) \leq v(P)/(d+1).$$

To see why the first inequality is true, observe that

$$\begin{aligned} v(p_1, \dots, p_d, c(P')) &\leq v(B, c(P')) = v(B, c(C)) \\ &\leq v(B, c(D)) = v(R)/(d+1) = v(P')/(d+1) . \end{aligned}$$

This completes the proof.  $\square$

The bound in Lemma 2 is tight, for example, when  $S$  consists of the  $d+1$  vertices of a simplex. Next we show how this relates to Oja depth:

**Theorem 3.** *Let  $S$  be a set of points in  $\mathbb{R}^d$  whose convex hull,  $A$ , has unit volume. Then  $d(c(A), S) \leq \binom{n}{d}/(d+1)$ .*

*Proof.*

$$\begin{aligned} d(c(A), S) &= \sum_{y_1, \dots, y_d \in \binom{S}{d}} v(c(A), y_1, \dots, y_d) \\ &\leq \binom{n}{d}/(d+1) , \end{aligned}$$

where the inequality is an application of Lemma 2.  $\square$

### 3. Oja Center and Mass Center of $S$

In this section, we show that the center of mass of  $S$  provides a constant-factor approximation to the point of minimum Oja depth. We begin by proving the result for point sets in 1 dimension:

**Lemma 3.** *For any finite set  $S \subset \mathbb{R}$ , and any  $x \in \mathbb{R}$ ,  $d(c(S), S) \leq 2 d(x, S)$ .*

*Proof.* Denote the elements of  $S$  by  $p_1, \dots, p_n$  in any order. Let the multiset  $S_i$  contain  $p_1, \dots, p_i$  as well as  $n-i$  copies of  $x$ . Let  $c_i = c(S_i)$ . We will show, by induction on  $i$ , that  $d(c_i, S_i) \leq 2 d(x, S_i)$  for all  $i \in \{0, \dots, n\}$ . This is sufficient, since  $S_n = S$ .

For the base case  $S_0$  consists of  $n$  copies of  $x$ , so  $c_0 = x$  and  $d(c_0, S_0) = 0 = 2 d(x, S_0)$ . Next, we assume that  $d(c_i, S_i) \leq 2 d(x, S_i)$  and prove that  $d(c_{i+1}, S_{i+1}) \leq 2 d(x, S_{i+1})$ . Note that

$$d(x, S_{i+1}) = d(x, S_i) + |p_{i+1} - x| .$$

Furthermore,

$$c_{i+1} = c_i + (p_{i+1} - x)/n ,$$



so

$$\begin{aligned}
& d(c_{i+1}, S_{i+1}) \\
&= d(c_i, S_i) + \sum_{q \in S_i} (|c_{i+1} - q| - |c_i - q|) + (|c_{i+1} - p_{i+1}| - |c_{i+1} - x|) \\
&\leq d(c_i, S_i) + n|p_{i+1} - x|/n + (|c_{i+1} - p_{i+1}| - |c_{i+1} - x|) \\
&\leq d(c_i, S_i) + 2|p_{i+1} - x| \\
&\leq 2d(x, S_i) + 2|p_{i+1} - x| \\
&= 2d(x, S_{i+1}) ,
\end{aligned}$$

as required.  $\square$

We remark that the above proof uses little more than triangle inequality. In particular, the same proof shows that the center of mass gives a 2-approximation for the Fermat-Weber center in any dimension.<sup>3</sup> Unfortunately, in higher dimensions, Oja depth does not enjoy this nice property.

**Theorem 4.** *For any finite set  $S \subseteq \mathbb{R}^d$ ,  $d(c(S), S) \leq (d+1)d(x, S)$  for any  $x \in \mathbb{R}^d$ .*

*Proof.* The case  $d = 1$  is covered by Lemma 3, so we assume  $d \geq 2$ . In this proof, we will make use of the fact that, for any  $d$ -simplex  $T$  with vertex set  $V_T$  and a point  $q \in \mathbb{R}^d$ ,

$$v(T) \leq \sum_{p_1, \dots, p_d \in \binom{V_T}{d}} v(p_1, \dots, p_d, q) , \quad (3)$$

since  $T$  is contained in the union of the simplices on the right hand side. Equality occurs if  $q$  is inside  $T$ .

Define  $S_i$  as in the proof of Lemma 3. Let  $S'$  be  $S_{i+1}$  with one occurrence of  $p_{i+1}$  removed. The induction and base case are the same as in Lemma 3. First, we have

$$\begin{aligned}
d(x, S_{i+1}) &= d(x, S_i) \\
&\quad + \sum_{Q \in \binom{S_i}{d-1}} v(x, p_{i+1}, Q).
\end{aligned} \quad (4)$$

---

<sup>3</sup>The Fermat-Weber center of a point set  $S$  in  $\mathbb{R}^d$  is the point  $x$  that minimizes  $\sum_{y \in S} \|x - y\|$ .

and

$$d(c_{i+1}, S_{i+1}) = d(c_i, S_i) + \sum_{P \in \binom{S_i}{d}} (v(c_{i+1}, P) - v(c_i, P)) \quad (5)$$

$$+ \sum_{Q \in \binom{S'_i}{d-1}} (v(c_{i+1}, p_{i+1}, Q) - v(c_{i+1}, x, Q)). \quad (6)$$

We denote by  $y^\perp$  the projection of a point  $y$  on a line perpendicular to

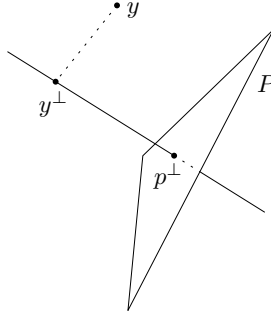


Figure 3: The projection of  $y$  and  $P$

the  $d - 1$  dimensional simplex  $P$  (see Figure 3), and by  $\|pq\|$  the Euclidean distance between points  $p$  and  $q$ . Let  $p$  be any point on  $P$ , then

$$\begin{aligned} |v(c_{i+1}, P) - v(c_i, P)| &= (1/d) v_{d-1}(P) \left| \|c_i^\perp p^\perp\| - \|c_{i+1}^\perp p^\perp\| \right| \\ &\leq (1/d) v_{d-1}(P) \|c_i^\perp c_{i+1}^\perp\| \\ &\leq (1/d) v_{d-1}(P) \|(1/n)x^\perp p_{i+1}^\perp\|. \end{aligned}$$

Then if  $x^\perp$  and  $p_{i+1}^\perp$  are on the same side of the hyperplane supporting  $P$ , we have

$$\begin{aligned} (1/d) v_{d-1}(P) \|(1/n)x^\perp p_{i+1}^\perp\| &\leq (1/nd) v_{d-1}(P) \left| \|x^\perp p^\perp\| - \|p_{i+1}^\perp p^\perp\| \right| \\ &\leq (1/n) |v(p_{i+1}, P) - v(x, P)| \\ &\leq (1/n) \sum_{Q \in \binom{P}{d-1}} v(x, p_{i+1}, Q) \end{aligned}$$

Otherwise if  $x^\perp$  and  $p_{i+1}^\perp$  are on different sides of the hyperplane supporting  $P$ , we have  $\|x^\perp p_{i+1}^\perp\| = \|x^\perp p^\perp\| + \|p_{i+1}^\perp p^\perp\|$ . In this case the two simplices

$Px$  and  $Pp_{i+1}$  are disjoint and the convex hull of  $Pxp_{i+1}$  is covered by the union of the simplices  $Qxp_{i+1}$  for  $Q \in \binom{P}{d-1}$  (see Figure 4), thus

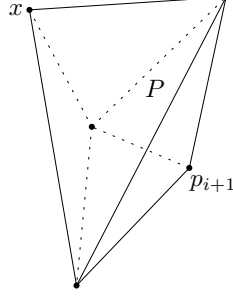


Figure 4: Point  $x$  and  $p_{i+1}$  are different sides of  $P$

$$\begin{aligned}
(1/d) v_{d-1}(P) \|(1/n)x^\perp p_{i+1}^\perp\| &\leq (1/nd) v_{d-1}(P) (\|x^\perp p^\perp\| + \|p_{i+1}^\perp p^\perp\|) \\
&\leq (1/n) (v(p_{i+1}, P) + v(x, P)) \\
&\leq (1/n) \sum_{Q \in \binom{P}{d-1}} v(x, p_{i+1}, Q)
\end{aligned}$$

We can now prove that (5)  $\leq$  (4) as follows:

$$\begin{aligned}
\sum_{P \in \binom{S_i}{d}} (v(c_{i+1}, P) - v(c_i, P)) &\leq \sum_{P \in \binom{S_i}{d}} |v(c_{i+1}, P) - v(c_i, P)| \\
&\leq (1/n) \sum_{P \in \binom{S_i}{d}} \sum_{K \in \binom{P}{d-1}} v(x, p_{i+1}, K) \\
&\leq ((n - (d - 1))/n) \sum_{K \in \binom{S_i}{d-1}} v(x, p_{i+1}, K) .
\end{aligned}$$

Next, we show that (6)  $\leq d \times$  (4). Applying (3),

$$\begin{aligned}
& \sum_{Q \in \binom{S'}{d-1}} (\mathbf{v}(c_{i+1}, p_{i+1}, Q) - \mathbf{v}(c_{i+1}, x, Q)) \\
& \leq \sum_{Q \in \binom{S'}{d-1}} \left( \mathbf{v}(x, p_{i+1}, Q) + \sum_{R \in \binom{Q}{d-2}} \mathbf{v}(x, p_{i+1}, c_{i+1}, R) \right) \\
& \leq \sum_{Q \in \binom{S_i}{d-1}} \mathbf{v}(x, p_{i+1}, Q) + (n-1-(d-2)) \sum_{R \in \binom{S_i}{d-2}} \mathbf{v}(x, p_{i+1}, c_{i+1}, R) .
\end{aligned}$$

Let  $\bar{\mathbf{v}}(p_1, \dots, p_{d+1})$  denote the signed volume of the simplex with vertices  $p_1, \dots, p_{d+1}$ . By linearity of the determinant we have

$$\begin{aligned}
\bar{\mathbf{v}}(x, p_{i+1}, c_{i+1}, R) &= (1/n) \sum_{y \in S_{i+1}} \bar{\mathbf{v}}(x, p_{i+1}, y, R) \\
&= (1/n) \sum_{y \in S_i} \bar{\mathbf{v}}(x, p_{i+1}, y, R)
\end{aligned}$$

Since the absolute value of the sum can be bounded by the sum of the absolute values, we get

$$\mathbf{v}(x, p_{i+1}, c_{i+1}, R) \leq (1/n) \sum_{y \in S_i} \mathbf{v}(x, p_{i+1}, y, R),$$

and thus

$$\sum_{R \in \binom{S_i}{d-2}} \mathbf{v}(x, p_{i+1}, c_{i+1}, R) \leq ((d-1)/n) \sum_{Q \in \binom{S_i}{d-1}} \mathbf{v}(x, p_{i+1}, Q).$$

Thus we can get (6)  $\leq d \times$  (4).

Finally, we resubstitute to obtain

$$\begin{aligned}
d(c_{i+1}, S_{i+1}) &\leq d(c_i, S_i) + (d+1) \sum_{Q \in \binom{S_i}{d-1}} \mathbf{v}(x, p_{i+1}, Q) \\
&\leq (d+1) d(x, S_i) + (d+1) \sum_{Q \in \binom{S_i}{d-1}} \mathbf{v}(x, p_{i+1}, Q) \\
&= (d+1) d(x, S_{i+1}) .
\end{aligned}$$

Thus we have  $d(c(S), S) \leq (d+1) d(x, S)$ , completing the proof.  $\square$

We remark that Lemma 3 and Theorem 4 are essentially the best possible. To see this, take the multiset  $S$  that contains  $n - d$  copies of the origin  $o$ , and each of the remaining  $d$  points has one different coordinate 1 and all other coordinates 0. In this case  $d(o, S) = 1/d!$  and  $d(c(S), S) = (d + 1 - O(d^2/n)) \times 1/d!$ .

#### 4. Conclusion

We have given several results relating Oja depth and centers of mass. There are several directions for future work.

We do not know of any point set that gives a lower-bound matching the upper-bound of Theorem 3. The best lower-bound we know is that placing  $n/(d + 1)$  points at each vertex of any  $d$ -simplex of unit volume yields to an Oja depth of  $n^d/(d + 1)^d$  for any point inside the simplex. For  $d = 2$ , for example, Theorem 3 implies  $d(x, S) \leq n^2/6 - O(n)$  where as the best lower bound (above) has  $d(x, S) \geq n^2/9$ . This construction leads us to the following conjecture:

**Conjecture 1.** *For any point set  $S \subset \mathbb{R}^d$  whose convex hull has unit volume, there exists  $x \in \mathbb{R}^d$ , such that  $d(x, S) \leq n^d/(d + 1)^d$*

- [1] G. Aloupis, S. Langerman, M. Soss, and G. Toussaint. Algorithms for bivariate medians and a Fermat-Torricelli problem for lines. *Computational Geometry: Theory and Applications*, 26(1):69–79, 2003.
- [2] I. Bárány. A generalization of Carathéodory’s Theorem. *Discrete Mathematics*, 40:141–150, 1982.
- [3] E. Boros and Z. Füredi. The maximal number of covers by the triangles of a given vertex set in the plane. *Geometrica Dedicata*, 17:69–77, 1984.
- [4] C. Edwards and D. Penney. *Calculus with analytic geometry*. Prentice Hall, Upper Saddle River, NJ USA, 5th edition, 1998.
- [5] D. Fischer, J. Mttinen, K. Nordhausen, and D. Vogel. *OjaNP: Multivariate Methods Based on the Oja Median and Related Concepts*, 2010. R package version 0.9-4 <http://cran.r-project.org/web/packages/OjaNP/index.html>.
- [6] R. Liu. On a notion of data depth based upon random simplices. *The Annals of Statistics*, 18:405–414, 1990.

- [7] J. Matoušek. *Lectures in Discrete Geometry*. Springer-Verlag, New York, NY, 2002.
- [8] A. Niinimaa, H. Oja, and M. Tableman. The finite-sample breakdown point of the Oja bivariate median and of the corresponding half-samples version. *Statistics & Probability Letters*, 10:325–328, 1990.
- [9] H. Oja. Descriptive statistics for multivariate distributions. *Statistics and Probability Letters*, 1(6):327–332, 1983.
- [10] J. Pach and P. K. Agarwal. *Combinatorial Geometry*. John Wiley & Sons, New York, NY, 1995.
- [11] T. Ronkainen, H. Oja, and P. Orponen. Computation of the multivariate Oja median. In R. Dutter and P. Filzmoser, editors, *International Conference on Robust Statistics (ICORS 2001)*, 2003.
- [12] J. Tukey. Mathematics and the picturing of data. In *International Conference of Mathematicians*, 1971.
- [13] R. Webster. *Convexity*. Oxford University Press, New York, USA, 1995.