HAL
open science

A Proof of the Oja Depth Conjecture in the Plane

Nabil Mustafa, Hans Raj Tiwary, Daniel Werner

To cite this version:

Nabil Mustafa, Hans Raj Tiwary, Daniel Werner. A Proof of the Oja Depth Conjecture in the Plane. Computational Geometry: Theory and Applications Computational Geometry @ ScienceDirect, 2014, 47 (6), pp.668-674. hal-01026341

HAL Id: hal-01026341

https://hal.science/hal-01026341

Submitted on 21 Jul 2014

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

A Proof of the Oja Depth Conjecture in the Plane

Nabil H. Mustafa
Dept. of Computer Science,
ESIEE Paris, Université Paris-Est, France.
mustafan@esiee.fr

Hans Raj Tiwary
Department of Mathematics,
Université Libre de Bruxelles, Belgium.
htiwary@ulb.ac.be

Daniel Werner *
Institut für Informatik,
FU Berlin, Germany.
daniel.werner@fu-berlin.de

Abstract

Given a set P of n points in the plane, the *Oja depth* of a point $x \in \mathbb{R}^2$ is defined to be the sum of the areas of all triangles defined by x and two points from P, normalized with respect to the area of the convex hull of P. The Oja depth of P is the minimum Oja depth of any point in $\mathbb{R}^2$. The *Oja depth conjecture* states that any set P of n points in the plane has Oja depth at most $n^2/9$. This bound would be tight as there are examples where it is not possible to do better. We present a proof of this conjecture. We also improve the previously best bounds for all $\mathbb{R}^d$, $d \geq 3$, via a different, more combinatorial technique.

1 Introduction

The general area of statistical data analysis includes designing measures (called data-depth measures) to succinctly capture the location, spread and variance of multivariate data. For example, for a set of points in $\mathbb{R}$, the notion of *mean* and *median* are two natural measures of its location. In particular, when the data consists of a finite set of points in Euclidean space $\mathbb{R}^d$, several notions for data depth have been proposed over the years. With each such measure, there come two questions: i) proving the existence of a point which suitably captures, with some guaranteed bounds, that measure and ii) devising efficient algorithms to compute this point.

Given a set P of n points in $\mathbb{R}^d$, some examples of data-depth measures are the following. *Location depth* (also called Tukey depth) of a point x is the minimum number of points of P lying in any halfspace containing x [Hod55, Tuk75, RRT99]. The Centerpoint Theorem [Rad46, Eck93] asserts that there always exists a point of location depth at least $n/(d+1)$, and that this value is tight. The point with the highest location depth w.r.t. to a point-set P is called the Tukey-median of P. The computational question of finding the Tukey-median of a point set has been studied extensively, and an optimal randomized

*This research was funded by Deutsche Forschungsgemeinschaft within the Research Training Group (Graduiertenkolleg) "Methods for Discrete Structures".

algorithm with expected running time $O(n \log n)$ is known in $\mathbb{R}^2$ [Cha04]. The best known deterministic algorithm for computing the Tukey median takes time $O(n \log^3 n)$ in $\mathbb{R}^2$ [LS03].

Another example of a statistical depth measure is *Simplicial depth* [Liu90], which for a point x is the number of simplices spanned by P that contain x. The First Selection Lemma [Bar82, Mat02] asserts that there always exists a point with Simplicial depth at least $c_d \cdot n^{d+1}$, where $c > 0$ is a constant depending only on d. The optimal value of c_d is known only for $d = 2$, where $c_2 = 1/27$ [BF84]. Determining the exact value of c_3 is still open, though it has been the subject of a flurry of work recently [BMRR10, BMN10, Gro10, MW11]. As for the computational question, the current-best algorithm computes a point with maximum simplicial depth in time $O(n^4)$ in $\mathbb{R}^2$ [ALST03].

Another well-studied measure, first proposed by Weber [Web29] in 1909, is the so-called L_1 median, where the depth of a point $q \in \mathbb{R}^2$ is defined to be the sum of the distances of q to the n input points. It is known that the point with the lowest such depth is unique in $\mathbb{R}^2$ and higher dimensions [Gal33]. Furthermore, it is also known that for $n > 3$ points the L_1 median cannot be computed exactly [Sca33], so the available algorithms only compute approximate solutions using gradients or iterations (See, for example, [Gow74] and [GS98]).

In this paper, we study another well-known measure called the *Oja depth* of a point-set.

Oja depth. Given a full-dimensional set P of n points in $\mathbb{R}^d$, the *Oja* depth [Oja83] of a point $x \in \mathbb{R}^d$ w.r.t. P, denoted Oja-depth(x,P), is defined to be the sum of the volumes of all $(d+1)$-simplices spanned by x and d points of P, normalized with respect to the volume of the convex hull of P. Formally, given a set $Q \subset \mathbb{R}^d$, let conv(Q) denote the convex hull of Q. Given a polyhedron $C \subset \mathbb{R}^d$, let vol(C) denote the d-dimensional volume of C. Then,

$$\text{Oja-depth}(x, P) = \sum_{\{y_1,\ldots,y_d\} \in \binom{P}{d}} \frac{\text{vol}(\text{conv}(x, y_1, \ldots, y_d))}{\text{vol}(\text{conv}(P))} \tag{1}$$

The Oja depth of P, denoted Oja-depth(P), is the minimum Oja depth over all $x \in \mathbb{R}^d$. A point that achieves this depth is called *Oja median* of P. Oja [Oja83] showed that such a point exists but it need not be unique and in general the set of all the points attaining the minimum define a convex region. Since Oja median is not necessarily unique, in this paper we will talk about the Oja depth of a point set instead of the depth of the Oja median. Also, from now onwards we assume w.l.o.g. that vol(conv$(P)) = 1$. See Figure 1 for an example of Oja depth of a random point-set in the plane.

Known bounds. It is known that for every point set P, we have

$$\text{Oja-depth}(P) \;\leq\; \binom{n}{d}$$

To see this, observe that any $(d+1)$-simplex spanned by points inside the convex hull of P can have volume at most 1, and so a trivial upper-bound for Oja depth of any $P \subset \mathbb{R}^d$ is $\binom{n}{d}$, achieved by picking any $x \in$ conv(P). Futhermore, one can construst a point set P such that

$$\text{Oja-depth}(P) \geq \left(\frac{n}{d+1}\right)^d.$$

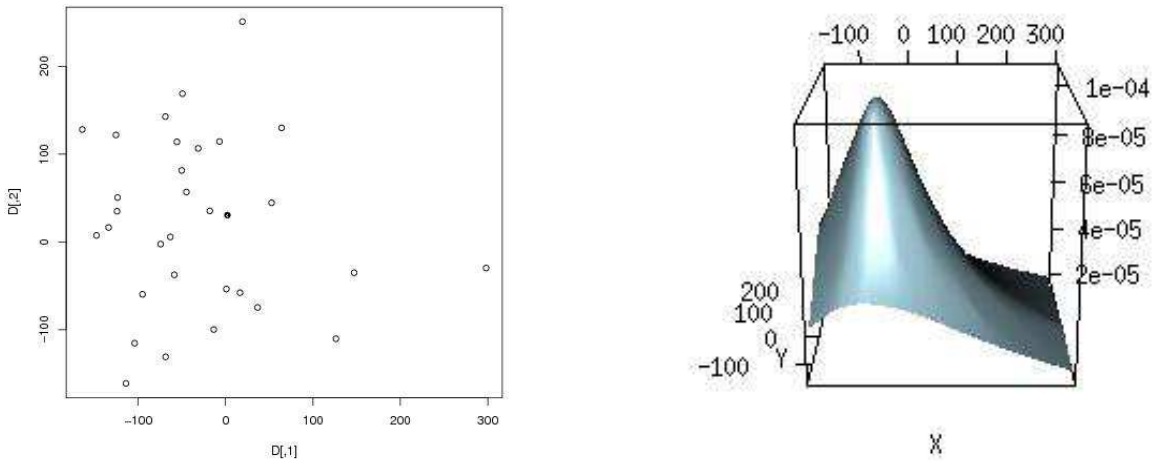


Figure 1: (a) Set of 30 points, together with a point (shaded black) with minimum Oja depth, (b) Contours for a variant (`http://cran.open-source-solution.org/web/packages/depth/depth.pdf`) of Oja depth for the point set on the left implemented in the statistical package **R**.

For this lower bound we construct P by placing $n/(d+1)$ points at each of the $(d+1)$ vertices of a unit-volume simplex in $\mathbb{R}^d$. It is easy to see that any point will have Oja depth at least $(n/(d+1))^d$.

The conjecture in [CDI$^+$10] is that the lower bound given above is tight.

Oja Depth Conjecture [CDI$^+$10]: For all sets $P \subset \mathbb{R}^d$ of n points, Oja-depth$(P) \leq (\frac{n}{d+1})^d$.

The current best upper bound [CDI$^+$10] is that the Oja depth of any set of n points is at most $\binom{n}{d}/(d+1)$. In particular, for $d = 2$, this gives $n^2/6$ whereas our result implies an upper bound of $n^2/9$. We would like to remark that both [CDI$^+$10] as well as we consider the center of mass of the given point set for proving our bounds, but our analysis offers better (tight) bound for the Oja-depth.

The Oja depth conjecture states the existence of a low-depth point, but given P, computing the *lowest*-depth point is also an interesting problem. In $\mathbb{R}^2$, Rousseeuw and Ruts [RR96] presented a straightforward $O(n^5 \log n)$ time algorithm for computing the lowest-depth point, which was then improved to the current-best algorithm with running time $O(n \log^3 n)$ [ALST03]. An approximate algorithm utilizing fast rendering systems on current graphics hardware was presented in [KMV06, Mus04]. For general d, various heuristics for computing points with low Oja depth were given by Ronkainen, Oja and Orponen [ROO03].

Our results. We present progress on the Oja depth conjecture. In Section 2, we present our main theorem (Theorem 2.6), which completely resolves the planar case. In particular we prove that for every set P of n points in $\mathbb{R}^2$ the center of mass of the convex hull of P has depth at most $\frac{n^2}{9}$.

In Section 3, using completely different (and more combinatorial) techniques for higher dimensions, we

also prove (Theorem 3.3) that every set P of n points in $\mathbb{R}^d$, $d \geq 3$, has Oja depth at most

$$\frac{2n^d}{2^d d!} - \frac{2d}{(d+1)^2(d+1)!}\binom{n}{d} + O(n^{d-1}).$$

This improves the previously best bounds by an order of magnitude.

2 A tight bound in $\mathbb{R}^2$

We now come to prove a tight bound for $\mathbb{R}^2$. First, let us give some basic definitions. The *center of mass* or *centroid* of a convex set X is defined as

$$c(X) = \frac{\int_{x \in X} x \, dx}{\text{area}\,(X)}.$$

For a discrete point set P, the center of mass of P is defined as the center of mass of the convex hull of P. When we talk about the *centroid of* P, we refer to the center of mass of the convex hull (not to be confused with the discrete centroid $\sum p/|P|$).

In this paper, we will bound the Oja depth of the centroid of a set. As we will see the Oja depth of the centroid is tight in the worst case. Our proof will rely on the following two known results.

Lemma 2.1. *[Winternitz [Bla23]] Every line through the centroid of a convex object has at most $\frac{5}{9}$ of the total area on either side.*

Lemma 2.2. *[CDI+10] Let P be a convex object with unit area and let c be its centroid. Then every simplex inside P which has c as a vertex has area at most $\frac{1}{3}$.*

To simplify matters, we will use the following proposition.

Proposition 1. *If we project an interior point $p \in P$ radially outwards from the centroid c to the boundary of the convex hull, the Oja depth of the point c does not decrease.*

Proof. First, observe that the center of mass does not change. It suffices to show that every triangle that has p as one of its vertices increases its area. Let $T := \Delta(c, p, q)$ be any triangle. The area of T is $\frac{1}{2}\|c - p\| \cdot h$, where h is the height of T with respect to $p - c$. If we move p radially outwards to a point p', h does not change, but $\|c - p'\| > \|c - p\|$. See Fig. 2. □

This implies that in order to prove an upper bound, we can assume that P is in convex position. Note that the aforementioned transformation brings the point only in weakly convex position, that is, some of the points lying on the boundary of the convex hull might not actually be vertices of the convex hull. This, however, is sufficient for our proof and for brevity we will use "convex" to mean "weakly convex".

From now on, let P be a set of points in convex postion, and let $c := c(\text{conv}(P))$ denote its center of mass as defined above. Further, let $p_1, \ldots, p_n$ denote the points sorted clockwise by angle from c. We define the *distance* of two points p_i, p_j as the difference of their position in this order (modulo n):

$$\text{dist}(p_i, p_j) := \min\{j - i \mod n, i - j \mod n\} \subseteq \{1, \ldots, \lfloor n/2 \rfloor\}.$$

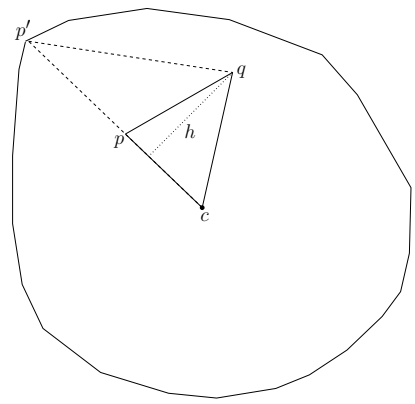
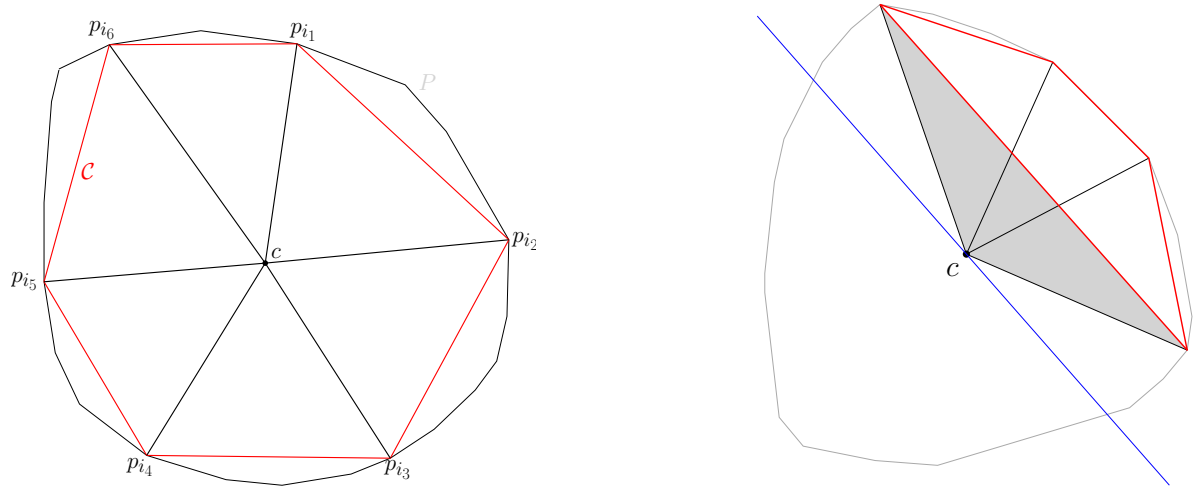


Figure 2: Moving points to the boundary increases the Oja depth

A triangle that is formed by c and two points at distance i is called an *i-triangle*, or *triangle of type i*. Observe that for each i, $1 \leq i < \lfloor n/2 \rfloor$, there are exactly n triangles of type i. Further, if n is even, then there are $n/2$ triangles of type $\lfloor n/2 \rfloor$, otherwise there are n. These constitute all possible triangles.

Let $C \subseteq P$, and let $\mathcal{C}$ be the boundary of the convex hull of C. This will be called a *cycle*. The length of a cycle is simply the number of elements in C. A cycle $\mathcal{C}$ of length i *induces i triangles* that arise by taking all triangles formed by an edge in $\mathcal{C}$ and the center of mass c (of $\text{conv}(P)$). The area induced by $\mathcal{C}$ is the sum of areas of these i triangles. See Fig. 3(a).



(a) A cycle and its induced triangles

(b) A case when c lies outside convex hull of $\mathcal{C}$

Figure 3: Cycles

The triangles induced by $C = P$ form a partition of $\text{conv}(P)$. Thus

Lemma 2.3. *The total area of all triangles of type* 1 *is exactly* 1.

The following shows that we can generalize this lemma to bound the total area induced by *any* cycle:

Lemma 2.4. *Let $\mathcal{C}$ be a cycle. Then $\mathcal{C}$ induces a total area of at most* 1.

Proof. We distinguish two cases.

Case 1: The centroid c lies in the convex hull of $\mathcal{C}$. In this case, all triangles are disjoint, so the area is at most one.

Case 2: c does not lie in the convex hull of $\mathcal{C}$ (see Fig. 3(b)). Then there is a line through c that has all the triangles induced by $\mathcal{C}$ on one side. Then we can remove one triangle – the one induced by the pair $\{p_{i_j}, p_{i_{j+1}}\}$ that has c on the left side (the gray triangle shown in the figure) – to get a set of disjoint triangles. By Lemma 2.1, the area of the remaining triangles can thus be at most $5/9$. By Lemma 2.2, the removed triangle has an area of at most $1/3$. Thus, the total area is at most $8/9$.

□

We now prove the key lemma, which is a general version of Lemma 2.3.

Lemma 2.5. *The total area of all triangles of type i is at most i.*

Proof. To prove this lemma for a fixed i, we will create n cycles. Each cycle will consist of one triangle of type i, and $n-i$ triangles of type 1 (counting multiplicities). We then determine the total area of these cycles and subtract the area of all 1-triangles. This will give the desired result.

Let $p_1, \ldots, p_n$ be the points ordered by angles from the centroid c. For $j = 1 \ldots n$, let $\mathcal{C}_j$ be the cycle consisting of the $n-i+1$ points $\{p_{j+i \mod n}, p_{j+i+1 \mod n}, \ldots, p_{j-1 \mod n}, p_{j \mod n}\}$. This is a cycle that consists of one triangle of type i (the one defined by the three points c, p_j, p_{j+i}), and $n-i$ triangles of type 1.

By Lemma 2.4, every cycle $\mathcal{C}_j$ induces an area of at most 1. If we sum up the areas of all cycles $\mathcal{C}_j$, $1 \leq j \leq n$, we thus get an area of at most n.

We now determine how often we have counted each triangle. Each i-triangle is counted exactly once. Further, for every cycle we count $n-i$ triangles of type 1. For reasons of symmetry, each 1-triangle is counted equally often. Indeed, each one is counted *exactly* $n-i$ times over all the cycles. By Lemma 2.3, their area is *exactly* $n-i$, which we can subtract from n to get the total area of the i-triangles:

$$\sum_{i\text{-triangle } T} \text{area}(T) =\leq n-(n-i) \cdot \left(\sum_{1\text{-triangle } T} \text{area}(T) \right) = n-(n-i) = i.$$

□

Now we prove the main result of this section:

Theorem 2.6. *Let P be any set of points in the plane with area(conv(P)) $= 1$, and let c be the centroid of conv(P). Then the Oja depth of c is at most $\frac{n^2}{9}$. Furthermore, the centroid can be computed in $O(n \log n)$ time.*

Proof. We will bound the area of the triangles depending on their type. For i-triangles with $1 \leq i \leq \lfloor n/3 \rfloor$, we will use Lemma 2.5. For i-triangles with $\lfloor n/3 \rfloor < i \leq \lfloor n/2 \rfloor$, this would give us a bound worse than $n/3$, so we will use Lemma 2.2 for each of these.

By Lemma 2.5, the sum of the areas of all triangles of type at most $\lfloor n/3 \rfloor$ is at most

$$\sum_{i=1}^{\lfloor n/3 \rfloor} i = \frac{\lfloor n/3 \rfloor (\lfloor n/3 \rfloor + 1)}{2} = \frac{1}{2} \left\lfloor \frac{n}{3} \right\rfloor^2 + \frac{1}{2} \left\lfloor \frac{n}{3} \right\rfloor .$$

If n is odd, there are $n (\lfloor n/2 \rfloor - \lfloor n/3 \rfloor)$ triangles remaining, n for each type j, $\lfloor n/3 \rfloor < j \leq \lfloor n/2 \rfloor$. If n is even, there are only $n/2$ triangles of type $n/2$ and so $n (\lfloor n/2 \rfloor - \lfloor n/3 \rfloor - 1/2)$ triangles remaining. In either case the number of remaining triangles is $n^2/2 - n \lfloor n/3 \rfloor - n/2$. For these we use Lemma 2.2 to bound the area of each by $1/3$. Thus, the area of these remaining triangles is at most $\frac{n^2}{6} - \frac{n}{3} \left\lfloor \frac{n}{3} \right\rfloor - \frac{n}{6}$.

So the Oja depth is at most $\frac{1}{2} \left\lfloor \frac{n}{3} \right\rfloor^2 + \frac{1}{2} \left\lfloor \frac{n}{3} \right\rfloor + \frac{n^2}{6} - \frac{n}{3} \left\lfloor \frac{n}{3} \right\rfloor - \frac{n}{6}$.

To complete the proof we distinguish the cases when n is of the form $3k, 3k+1$ or $3k+2$.

Case $n = 3k$:

$\frac{1}{2} \left\lfloor \frac{n}{3} \right\rfloor^2 + \frac{1}{2} \left\lfloor \frac{n}{3} \right\rfloor + \frac{n^2}{6} - \frac{n}{3} \left\lfloor \frac{n}{3} \right\rfloor - \frac{n}{6} = \frac{k^2}{2} + \frac{k}{2} + \frac{3k^2}{2} - k^2 - \frac{k}{2} = k^2 = \frac{n^2}{9}$

Case $n = 3k+1$:

$\frac{1}{2} \left\lfloor \frac{n}{3} \right\rfloor^2 + \frac{1}{2} \left\lfloor \frac{n}{3} \right\rfloor + \frac{n^2}{6} - \frac{n}{3} \left\lfloor \frac{n}{3} \right\rfloor - \frac{n}{6} = \frac{k^2}{2} + \frac{k}{2} + \frac{3k^2}{2} + k + \frac{1}{6} - k^2 - \frac{k}{3} - \frac{k}{2} - \frac{1}{6} = k^2 + \frac{2k}{3} \leq \frac{(3k+1)^2}{9} = \frac{n^2}{9}$

Case $n = 3k+2$:

$\frac{1}{2} \left\lfloor \frac{n}{3} \right\rfloor^2 + \frac{1}{2} \left\lfloor \frac{n}{3} \right\rfloor + \frac{n^2}{6} - \frac{n}{3} \left\lfloor \frac{n}{3} \right\rfloor - \frac{n}{6} = \frac{k^2}{2} + \frac{k}{2} + \frac{3k^2}{2} + 2k + \frac{2}{3} - k^2 - \frac{2k}{3} - \frac{k}{2} - \frac{1}{3} = k^2 + \frac{4k}{3} + \frac{1}{3} \leq \frac{(3k+2)^2}{9} = \frac{n^2}{9}$

Thus, the Oja depth of the centroid is at most $n^2/9$. Finally, note that $c = c(\text{conv}(P))$ can be computed in $O(n \log n)$ time by triangulating $\text{conv}(P)$.

□

3 Higher Dimensions

We now present improved bounds for the Oja depth problem in dimensions greater than two. Before the main theorem, we need the following two lemmas.

Lemma 3.1. *Given any set P of n points in $\mathbb{R}^d$ and any point $q \in \mathbb{R}^d$, any line l through q intersects at most $f(n,d)$ $(d-1)$-simplices spanned by P, where $f(n,d) = \frac{2n^d}{2^d d!} + O(n^{d-1})$.*

Proof. Project P onto the hyper plane H orthogonal to l to get the point set P' in $\mathbb{R}^{d-1}$. The line l becomes a point on H, say point l^*. Then l intersects the $(d-1)$-simplex spanned by $\{p_1, \ldots, p_d\}$ if and only if the convex hull of the corresponding points in P' contains the point l^*. By Barany [Bar82], given n points in $\mathbb{R}^d$, any point in $\mathbb{R}^d$ is contained in at most these many d-simplices:

$$\frac{2n}{n+d+1} \cdot \binom{(n+d+1)/2}{d+1} \text{ if } n-d \text{ is odd} \qquad \frac{2(n-d)}{n+d+2} \cdot \binom{(n+d+2)/2}{d+1} \text{ if } n-d \text{ is even}$$

Note that both the bounds above are equal within additive factor of $O(n^d)$, and simplifying the first, we get:

$$\frac{2n}{n+d+1} \cdot \binom{(n+d+1)/2}{d+1} \leq 2 \cdot \binom{(n+d+1)/2}{d+1} \leq \frac{2(n+d+1)^{d+1}}{2^{d+1}(d+1)!} \leq \frac{2n^{d+1}}{2^{d+1}(d+1)!} + O(n^d)$$

We apply this to P' in $\mathbb{R}^{d-1}$ to get the desired result. □

Lemma 3.2. *Given any set P of n points in $\mathbb{R}^d$, there exists a point q such that any half-infinite ray from q intersects at least $\frac{2d}{(d+1)^2(d+1)!}\binom{n}{d}$ $(d-1)$-simplices spanned by P.*

Proof. This follows directly from a recent result of Gromov [Gro10], who showed that given any set P, there exists a point q contained in at least $\frac{2d}{(d+1)(d+1)!}\binom{n}{d+1}$ simplices spanned by P. Now note that any half-infinite ray from q must intersect exactly one $(d-1)$-dimensional face of each simplex containing q and each such $(d-1)$-simplex can be counted at most $n-d$ times. Simplifying, we get the desired result. □

Remark: There have been several improvements [MW11, KMS12] (and still ongoing) after the initial paper of Gromov; however as these improvements are significant only for small constant dimensions, we prefer to give the above considerably simpler bound of Gromov. It is clear that any improvement in the above bound gives a corresponding improvement for our result.

Given a set P and a point q, call a simplex a q-simplex if it is spanned by q and d other points of P.

Theorem 3.3. *Given any set P of n points in $\mathbb{R}^d$, there exists a point q with Oja depth at most*

$$\frac{2n^d}{2^d d!} - \frac{2d}{(d+1)^2(d+1)!}\binom{n}{d} + O(n^{d-1}).$$

Proof. Let q be the point from Lemma 3.2. Assign a weight function, $w(r)$, to each point $r \in \text{conv}(P)$, where $w(r)$ is the number of q-simplices spanned by P and q that contain r. Then note that if r is contained in a q-simplex, spanned by, say, $\{q, p_{i_1}, \ldots, p_{i_d}\}$, then the half-infinite ray $\overrightarrow{qr}$ intersects the $(d-1)$-simplex spanned by $\{p_{i_1}, \ldots, p_{i_d}\}$. Therefore $w(r)$ is equal to the number of $(d-1)$-simplices intersected by the ray $\overrightarrow{qr}$. To upper-bound this, note that the ray starting from q but in the opposite direction to the ray $\overrightarrow{qr}$, intersects at least $\frac{2d}{(d+1)^2(d+1)!}\binom{n}{d}$ $(d-1)$-simplices (by Lemma 3.2). On the other hand, by Lemma 3.1, the entire line passing through q and r intersects at most $\frac{2n^d}{2^d d!} + O(n^{d-1})$ $(d-1)$-simplices spanned by P. These two together imply that the ray $\overrightarrow{qr}$ intersects at most $\frac{2n^d}{2^d d!} - \frac{2d}{(d+1)^2(d+1)!}\binom{n}{d} + O(n^{d-1})$ $(d-1)$-simplices spanned by P, and this is also an upper bound on $w(r)$. Finally, we have

$$\begin{aligned}
\sum_{|P'|=d} \text{vol}(\text{conv}(\{q\} \cup P')) &= \int_{x \in \text{conv}(P)} w(x)\, dx \\
&\leq \left(\frac{2n^d}{2^d d!} - \frac{2d}{(d+1)^2(d+1)!}\binom{n}{d} + O(n^{d-1})\right) \int_{x \in \text{conv}(P)} dx \\
&= \frac{2n^d}{2^d d!} - \frac{2d}{(d+1)^2(d+1)!}\binom{n}{d} + O(n^{d-1}),
\end{aligned}$$

finishing the proof. □

Remark: For approximate comparison, the previous-best result gives the upper-bound of $\binom{n}{d}/(d+1) \approx \frac{n^d}{(d+1)d!}$, while the above result, even after ignoring the much-smaller second term, gives the upper-bound of $\frac{2n^d}{2^d d!}$ – an exponential improvement.

4 Conclusion and further prospects

The technique we used to prove the two-dimensional case fails in higher dimensions due to our inability to characterize and count the number of sets of type i. As a triangle of type i in the planar case corresponds to a line which has exactly $i-1$ points on the outer side, looking at d-dimensional i-sets[1] seems to be a promising approach to adapt our two-dimensional technique to higher dimensions. Using this, one might be able to prove the d-dimensional analogue of Theorem 2.6:

Conjecture 4.1. *(Strong Oja Depth Conjecture) Let P be a set of points in $\mathbb{R}^d$. Then the center of mass of the convex hull of P has Oja depth at most*

$$\left(\frac{n}{d+1}\right)^d.$$

On the contrary, we do not expect to prove the exact bound using our combinatorial higher-dimensional technique.

Besides the (static) combinatorial questions, the main computational question is whether or not the Oja depth of a point set in $\mathbb{R}^d$ can be computed in polynomial time, or at least in time $O(f(d) \cdot n^c)$ for any computable function f. (Problems that admit algorithms with such a running time are called *fixed-parameter tractable*.) More precisely, we ask whether the following decision problem NP-hard: Given a set of points P in $\mathbb{R}^d$ a point x, and an integer N, is Oja-depth$(x, P) \leq N$?

Acknowledgments. This research was done during the DCG Special Semester in Lausanne. We thank the organizers Janos Pach and Emo Welzl, as well as the EPFL and Bernoulli Center staff. We also thank Günter Rote for pointing out to us a simplification in the proof of Lemma 2.5.

References

[ALST03] G. Aloupis, S. Langerman, M. A. Soss, and G. T. Toussaint. Algorithms for bivariate medians and a Fermat-Torricelli problem for lines. *Comput. Geom.*, 26(1):69–79, 2003.

[Bar82] I. Barany. A generalization of Caratheodory's theorem. *Discrete Mathematics*, 40:141–152, 1982.

[BF84] E. Boros and Z. Furedi. The maximal number of covers by the triangles of a given vertex set on the plane. *Geom. Dedicata*, 17:69–77, 1984.

[1] Typically called k-sets.

[Bla23] W. Blaschke. *Vorlesungen uber Differentialgeometrie. II, Affine Differentialgeometrie.* Springer, 1923.

[BMN10] B. Bukh, J. Matousek, and G. Nivasch. Stabbing simplices by points and flats. *Discrete & Computational Geometry*, 43(2):321–338, 2010.

[BMRR10] A. Basit, N. H. Mustafa, S. Ray, and S. Raza. Hitting simplices with points in 3d. *Discrete & Computational Geometry*, 44(3):637–644, 2010.

[CDI+10] D. Chen, O. Devillers, J. Iacono, S. Langerman, and P. Morin. Oja medians and centers of mass. In *Proceedings of the 22nd Canadian Conference on Computational Geometry (CCCG2010)*, pages 147–150, 2010.

[Cha04] T. M. Chan. An optimal randomized algorithm for maximum Tukey depth. In *SODA*, pages 430–436, 2004.

[Eck93] J. Eckhoff. Helly, Radon and Carathéodory type theorems. In *Handbook of Convex Geometry*, pages 389–448. North-Holland, 1993.

[Gal33] L. Galvani. Sulla determinazione del centro di gravità e del centro mediano di una popolazione. *Metron*, 11(1):17–47, 1933.

[Gow74] J. C. Gower. The mediancentre. *Applied Statistics*, 23(3):466–470, 1974.

[Gro10] M. Gromov. Singularities, expanders and topology of maps. part 2: From combinatorics to topology via algebraic isoperimetry. *Geometric and Functional Analysis*, 20:416–526, 2010.

[GS98] C. Groß and T. Strempel. On generalizations of conics and on a generalization of the Fermat-Torricelli problem. *American Mathematical Monthly*, 105(8):732–743, 1998.

[Hod55] J. Hodges. A bivariate sign test. *Annals of Mathematical Statistics*, 26:523–527, 1955.

[KMS12] D. Kral, L. Mach, and J. S. Sereni. A new lower bound based on Gromov's method of selecting heavily covered points. *Discrete and Computational Geometry, to appear*, 2012.

[KMV06] S. Krishnan, N. H. Mustafa, and S. Venkatasubramanian. Statistical data depth and the graphics hardware. In *Data Depth: Robust Multivariate Analysis, Computational Geometry and Applications*, pages 223–250. DIMACS Series in Discrete Mathematics and Theoretical Computer Science, 2006.

[Liu90] R. Liu. A notion of data depth based upon random simplices. *The Annals of Statistics*, 18:405–414, 1990.

[LS03] S. Langerman and W. L. Steiger. Optimization in arrangements. In *STACS*, pages 50–61, 2003.

[Mat02] J. Matoušek. *Lectures in Discrete Geometry*. Springer-Verlag, New York, NY, 2002.

[Mus04] N. H. Mustafa. *Simplification, Estimation and Classification of Geometric Objects*. PhD thesis, Duke University, Durham, North Carolina, 2004.

[MW11] J. Matousek and U. Wagner. On Gromov's method of selecting heavily covered points. *Manuscript*, 2011.

[Oja83] H. Oja. Descriptive statistics for multivariate distributions. *Statistics and Probability Letters*, 1:327–332, 1983.

[Rad46] R. Rado. A theorem on general measure. *Journal London Mathematical Society*, s1-21:291–300, 1946.

[ROO03] T. Ronkainen, H. Oja, and P. Orponen. Computation of the multivariate oja median. In *R. Dutter and P. Filzmoser*. International Conference on Robust Statistics, 2003.

[RR96] P. Rousseeuw and I. Ruts. Bivariate location depth. *Applied Statistics*, 45:516–526, 1996.

[RRT99] P. J. Rousseeuw, I. Ruts, and J. W. Tukey. The bagplot: A bivariate boxplot. *The American Statistician*, 53(4):382–387, 1999.

[Sca33] D. Scates. Locating the median of the population in the united states. *Metron*, 11(1):49–65, 1933.

[Tuk75] J. Tukey. Mathematics and the picturing of data. In *Proc. of the international congress of mathematicians*, pages 523–531, 1975.

[Web29] A. Weber [translated by C. J. Friedrich from Weber's 1909 book]. *Theory of the Location of Industries*. The University of Chicago Press, 1929.