On the Planar Spherical Depth and Lens Depth

David Bremner *   Rasoul Shahsavarifar *

Abstract

For a distribution function F on $\mathbb{R}^d$ and a point $q \in \mathbb{R}^d$, the *spherical depth* $SphD(q;F)$ is defined to be the probability that a point q is contained inside a random closed hyperball obtained from a pair of points from F. The *lens depth* $LD(q;F)$ is defined analogously using hyperlens instead of hyperball in the definition of spherical depth. The spherical depth $SphD(q;S)$ (*lens depth* $LD(q;S)$) is also defined for an arbitrary data set $S \subseteq \mathbb{R}^d$ and point $q \in \mathbb{R}^d$. This definition is based on counting all of the closed hyperballs (hyperlenses), obtained from pairs of points in S, that contain q. The straightforward algorithm for computing the spherical depth (lens depth) in dimension d takes $O(dn^2)$. The main result of this paper is an optimal algorithm for computing the planar (bivariate) spherical depth. The algorithm takes $O(n \log n)$ time. By reducing the problem of *Element Uniqueness*, we prove that computing the spherical depth (lens depth) requires $\Omega(n \log n)$ time. Some geometric properties of spherical depth (lens depth) are also investigated in this paper. These properties indicate that *simplicial depth* (SD) is linearly bounded by spherical depth and lens depth (in particular, $LD \geq SphD \geq \frac{2}{3} SD$). To illustrate these relationships, some experimental results are provided. In these experiments on random point sets, the bounds of $SphD \geq 2\,SD$ and $LD \geq 1.2\,SphD$ are achieved.

1 Introduction

The rank statistic tests play an important role in univariate non-parametric statistics. If one attempts to generalize the rank tests to the multivariate case, the problem of defining a multivariate order statistic will occur. It is not clear how to define a multivariate order or rank statistic in a meaningful way. One approach to overcome this problem is to use the notion of data depth. Data depth measures the centrality of a point in a given data set in non-parametric multivariate data analysis. In other words, it indicates how deep a point is located with respect to the data set.

Over the last decades, various notions of data depth such as *halfspace depth* (Hotelling, 1929, [9, 17]; Tukey, 1975, [19]), *simplicial depth* (Liu, 1990, [11]) *Oja depth* (Oja, 1983, [15]), *regression depth* (Rousseeuw and Hubert, 1999, [16]), and others have emerged as powerful tools for non-parametric multivariate data analysis. Most of them have been defined to solve specific problems in data analysis. They are different in application, definition, and geometry of their central regions (regions with the maximum depth). Some notable research on the algorithmic aspects of planar data depth can be found in [1, 2, 4, 5, 6, 7, 12, 14, 16].

In 2006, Elmore, Hettmansperger, and Xuan [8] defined another notion of data depth named *spherical depth*. It is defined as the probability that point q is contained in a closed random hyperball with the diameter $\overline{x_i x_j}$, where x_i and x_j are two random points from a common distribution function F. These closed hyperballs are known as influence regions of the spherical depth function. In 2011, Liu and Modarres [13], modified the definition of influence region, and defined *lens depth*. Each lens depth influence region is defined as the the intersection of two hyperballs $B(x_i, d(x_i,x_j))$ and $B(x_j, d(x_i,x_j))$. These influence regions of spherical depth (lens depth) are the multidimensional generalization of *Gabriel circles* (lunes) in the definition of the *Gabriel Graph* (*Relative Neighbourhood Graph*) [13, 18]. Spherical depth and lens depth have some nice properties including affine invariance, symmetry, maximality at the centre and monotonicity. All of these properties are explored in [8, 13, 20].

Although we focus on the planar case here, a notable characteristic of the spherical depth (lens depth) is that its time complexity grows linearly in the dimension d while for most other data depths the time complexity grows exponentially. To the best of our knowledge, the current best algorithm for computing the spherical depth (lens depth) is the straightforward algorithm which takes $O(dn^2)$.

In this paper, we present an $O(n \log n)$ algorithm for computing the spherical depth in $\mathbb{R}^2$. Furthermore, we reduce the problem of Element Uniqueness to prove that computing the spherical depth (lens depth) of a query point requires $\Omega(n \log n)$ time. We also investigate some geometric properties of spherical

*Faculty of Computer Science, University of New Brunswick, Fredericton, NB, Canada, {bremner,ra.shahsavari}@unb.ca

depth and lens depth. These properties lead us to bound the simplicial depth, spherical depth, and lens depth of a point in terms of one another. Finally, some experiments are provided to illustrate the relationship between spherical depth, lens depth, and simplicial depth.

2 Spherical Depth and Lens Depth

Definition: The spherical (lens) influence region of x_i and x_j in $\mathbb{R}^d$ is a closed hyperball (hyperlens) defined as follows:

$$Sph(x_i, x_j) = \left\{ t \mid d(t, \frac{x_i + x_j}{2}) \leq \frac{d(x_i, x_j)}{2} \right\}$$

$$L(x_i, x_j) = \{ t \mid \max \{ d(t, x_i), d(t, x_j) \} \leq d(x_i, x_j) \},$$

where $d(.,.)$ is the Euclidean distance. Figures 1 and 2 show the $Sph(x_i, x_j)$ and $L(x_i, x_j)$ in $\mathbb{R}^2$, respectively.

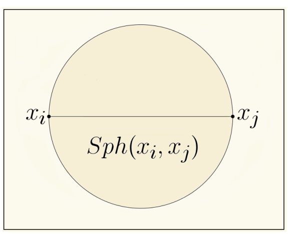

Figure 1: $Sph(x_i, x_j)$

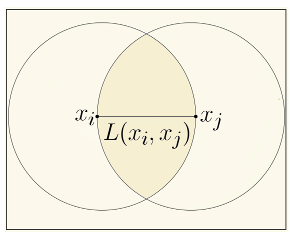

Figure 2: $L(x_i, x_j)$

Definition: For $S = \{x_1, ..., x_n\} \subset \mathbb{R}^d$ and $q \in \mathbb{R}^d$. The spherical (lens) depth of a q with respect to S, is defined as a proportion of $Sph(x_i, x_j)$ ($L(x_i, x_j)$), $1 \leq i < j \leq n$ that contain q. Using the indicator function $\boldsymbol{I}$, these definitions can be represented by (1) and (2).

$$SphD(q; S) = \frac{1}{\binom{n}{2}} \sum_{1 \leq i < j \leq n}^{n} I(q \in Sph(x_i, x_j)) \quad (1)$$

$$LD(q; S) = \frac{1}{\binom{n}{2}} \sum_{1 \leq i < j \leq n}^{n} I(q \in L(x_i, x_j)) \quad (2)$$

2.1 Algorithms for Computing the Spherical Depth of a Query Point

The current best algorithm for computing the spherical depth of a point $q \in \mathbb{R}^d$ with respect to a data set $S = \{x_1, x_2, ..., x_n\} \subseteq \mathbb{R}^d$ is the brute force algorithm. This naive algorithm needs to check all of the $\binom{n}{2}$ spherical influence regions obtained from the data points to figure out how many of them contain q. Checking all of the spherical influence regions causes the naive algorithm to take $\Theta(dn^2)$. Instead of counting, we focus on the geometric aspects of the spherical influence regions. These geometric properties lead us to develop an optimal $O(n \log n)$ algorithm for the computation of the spherical depth of q.
A proof of the following lemma which is a generalization of Thales' theorem[1] can be found in the Appendix.

Lemma 1 *For arbitrary points a, b, and t in $\mathbb{R}^2$, $t \in Sph(a, b)$ if and only if $\angle atb \geq \frac{\pi}{2}$.*

Algorithm: Using Lemma 1, we present an algorithm to compute the spherical depth of a query point $q \in \mathbb{R}^2$ with respect to $S = \{x_1, x_2, ..., x_n\} \subseteq \mathbb{R}^2$. This algorithm is summarized in the following steps.

- **Translating the points:** Suppose that T is a translation by $(-q)$. We apply T to translate q and all data points into their new coordinates. Obviously, $T(q) = O$.
- **Sorting the translated data points:** In this step we sort the translated data points based on their angles in their polar coordinates. After doing this step, we have S_T which is a sorted array of the translated data points.
- **Calculating the spherical depth:** Suppose that $x_i(r_i, \theta_i)$ is the i^{th} element in S_T. For x_i, we define the arrays O_i and N_i as follows:

$$O_i = \left\{ j \mid x_j \in S_T, \frac{\pi}{2} \leq |\theta_i - \theta_j| \leq \frac{3\pi}{2} \right\} \quad (3)$$

$$N_i = \{1, 2, ..., n\} \setminus O_i.$$

Thus the spherical depth of q with respect to S can be computed by:

$$SphD(q; S) = SphD(0; S_T) = \frac{1}{2} \sum_{1 \leq i \leq n} |O_i|. \quad (4)$$

To present a formula for computing $|O_i|$, we define f_i and l_i as follows:

$$f_i = \begin{cases} \min N_i - 1 & \text{if } \frac{\pi}{2} < \theta_i \leq \frac{3\pi}{2} \\ \min O_i & \text{otherwise} \end{cases}$$

$$l_i = \begin{cases} \max N_i + 1 & \text{if } \frac{\pi}{2} < \theta_i \leq \frac{3\pi}{2} \\ \max O_i & \text{otherwise.} \end{cases}$$

[1]Thales' theorem: If a, b, and c are points on a circle where $\overline{ac}$ is a diameter of the circle, then $\angle abc$ is a right angle

Figures 3 and 4 illustrate O_i, N_i, f_i, and l_i in two different cases. Considering the definitions of f_i and l_i,

$$|O_i| = \begin{cases} f_i + (n - l_i + 1) & \text{if } \frac{\pi}{2} < \theta_i \leq \frac{3\pi}{2} \\ l_i - f_i + 1 & \text{otherwise.} \end{cases}$$

This allows us to compute $|O_i|$ using a pair of binary searches.

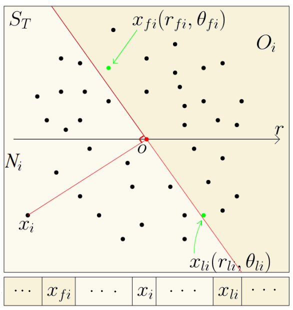

Figure 3: $\theta \in (\frac{\pi}{2}, \frac{3\pi}{2}]$.

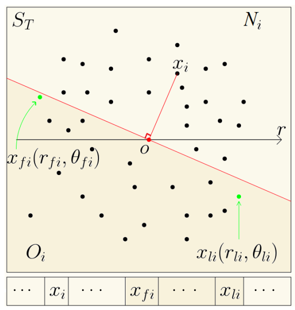

Figure 4: $\theta \notin (\frac{\pi}{2}, \frac{3\pi}{2}]$.

Time complexity: The first procedure in the algorithm takes $O(n)$ to translate q and all data points into the new coordinate system. The second procedure takes $O(n \log n)$ time. In this procedure, the loop iterates n times, and the sorting algorithm takes $O(n \log n)$. Due to using binary search for every O_i, the running time of the last procedure is also $O(n \log n)$. The rest of the algorithm contributes some constant time. In total, the running time of the algorithm is $O(n \log n)$.

Coordinate system: In practice it may be preferable to work in the Cartesian coordinate system. Sorting by angle can be done using some appropriate right-angle tests (determinants). Regarding the other angle comparisons, they can be done by checking the sign of dot products.

2.2 Lower Bound for Computing the the planar Spherical Depth and Lens Depth

We reduce the problem of Element Uniqueness[2] to the problem of computing the spherical depth and lens depth. It is known that the question of Element Uniqueness has a lower bound of $\Omega(n \log n)$ in the algebraic decision tree model of computation [3].

Theorem 2 *Computing the spherical depth of a query point in the plane takes $\Omega(n \log n)$ time.*

[2]Element Uniqueness problem: Given a set $A = \{a_1, a_2, ..., a_n\}$, is there a pair of indices i, j with $i \neq j$ such that $a_i = a_j$?

Proof. We show that finding the spherical depth allows us to answer the question of Element Uniqueness. Suppose that $A = \{a_1, a_2, ..., a_n\}$, for $n \geq 2$ is a given set of real numbers. We suppose all of the numbers to be positive (negative), otherwise we shift the points onto the positive X-axis. For every $a_i \in A$ we construct four points x_i, x_{n+i}, x_{2n+i}, and x_{3n+i} in the polar coordinate system as follows:

$$x_{(kn+i)} = \left(r_i, \theta_i + \frac{k\pi}{2}\right);\ 0 \leq k \leq 3,$$

where $r_i = \sqrt{1 + a_i^2}$ and $\theta_i = \tan^{-1}(1/a_i)$. Thus we have a set S of $4n$ points x_{kn+i}, for $1 \leq i \leq n$. The Cartesian coordinates of the points can be computed by:

$$x_{(kn+i)} = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}^k \begin{pmatrix} a_i \\ 1 \end{pmatrix};\ k = 0, 1, 2, 3.$$

See Figure 5.

We select the query point $q = (0, 0)$, and present an equivalent form of Equation (3) for O_j as follows:

$$O_j = \left\{x_k \in S \mid \angle x_j q x_k \geq \frac{\pi}{2}\right\},\ 1 \leq j \leq 4n, \quad (5)$$

We compute $SphD(q; S)$ in order to answer the Element Uniqueness problem. Suppose that every $x_j \in S$ is a unique element. In this case, $|O_j| = 2n + 1$ because, from (5), it can be figured out that the expanded O_j is as follows:

$$\begin{cases} \{x_{n+1}, .., x_{n+j}, x_{2n+1}, .., x_{3n}, x_{3n+j}, .., x_{4n}\}; & 1 \leq j \leq n \\ \{x_{2n+1}, .., x_{n+j}, x_{3n+1}, .., x_{4n}, x_{j-n}, .., x_n\}; & n < j \leq 2n \\ \{x_{3n+1}, .., x_{n+j}, x_1, .., x_n, x_{j-n}, .., x_{2n}\}; & 2n < j \leq 3n \\ \{x_1, .., x_{j-3n}, x_{n+1}, .., x_{2n}, x_{j-n}, .., x_{3n}\}; & \text{otherwise.} \end{cases}$$

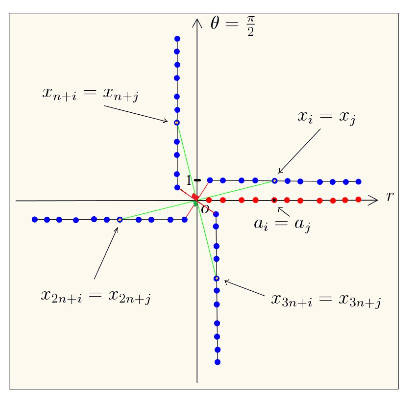

Figure 5: A representation of A, S, and duplications in these sets

Referring to Lemma 1 and Equation (4),

$$SphD(q;S) = \frac{1}{2}\sum_{1\le j\le 4n}(2n+1) = 4n^2+2n.$$

Now suppose that there exist some $i \neq j$ with $x_i = x_j$ in S. In this case, from (5), it can be seen that:

$$|O_{(kn+i) \bmod 4n}| = |O_{(kn+j) \bmod 4n}| = 2n+2,$$

where $k = 0, 1, 2, 3$ (see Figure 5). As an example, for $k = 0$, $|O_j| = |O_i| = 2n+2$ because the expanded form of these two sets is as follows: (without loss of generality, assume $i < j < n$)

$$O_i = O_j = \{x_{n+1}, .., x_{n+j}, x_{2n+1}, .., x_{3n}, \\ x_{3n+i}, x_{3n+j}, x_{3n+j+1}, .., x_{4n}\}.$$

Lemma 1 and Equation (4) imply that:

$$SphD(q;S) \ge \frac{1}{2}(8 + \sum_{1\le j\le 4n}(2n+1)) = 4n^2+2n+4.$$

Therefore the elements of A are unique if and only if the spherical depth of $(0,0)$ with respect to S is $4n^2+2n$. This implies that the computation of spherical depth requires $\Omega(n\log n)$ time. It is necessary to mention that the only computations in the reduction are the construction of S which takes $O(n)$ time. $\square$

Note: Instead of four copies of the elements of A, we could consider two copies of such elements to construct S. However, the depth calculation becomes more complicated in this case.

Theorem 3 *Computing the lens depth of a query point in the plane takes $\Omega(n\log n)$ time.*

Proof. Suppose that $B = \{b_1, b_2, ..., b_n\}$, for $n \ge 2$ is a given set of real numbers. Without loss of generality, let these numbers to be positive (see the proof of Theorem 2). We construct set $S = \{x_i, x_{n+i}\}$ of $2n$ points in the polar coordinate system such that $x_i = (b_i, 0)$ and $x_{n+i} = (b_i, \pi/3)$. See Figure 6. We select the query point $q = (0,0)$, and define L_j as follows:

$$L_j = \{x_k \in S \mid q \in L(x_j, x_k)\}, \; 1 \le j \le 2n. \quad (6)$$

Using Equation (6), the unnormalized form of Equation (2) can be presented by:

$$LD_S(q) = \frac{1}{2}\sum_{1\le j\le 2n}|L_j|. \quad (7)$$

We solve the problem of Element Uniqueness by computing $LD_S(q)$. Suppose that every $x_j \in S$ is a unique element. In this case, it can be verified that $L_j = \{x_{(n+j) \bmod 2n}\}$ (see Lemma 9 in the Appendix). Equation (7) implies that:

$$LD_S(q) = \frac{1}{2}\sum_{1\le j\le 2n}1 = n.$$

Now assume that there exists some $i \neq j$ with $x_i = x_j$ in S. In this case, $L_j = L_i = \{x_{(n+i) \bmod 2n}, x_{(n+j) \bmod 2n}\}$ and $L_{n+i} = L_{n+j} = \{x_{i \bmod 2n}, x_{j \bmod 2n}\}$ which means that

$$LD_S(q) = \frac{1}{2}\sum_{1\le j\le 2n}|L_j| = n+2.$$

In fact,

$$LD_S(q) = \frac{1}{2}\sum_{1\le j\le 2n}|L_j| = n+2c, \quad (8)$$

where c is the number of duplications in the elements of S. Therefore the elements of S are unique if and only if $c = 0$ in Equation 8. This implies that the computation of lens depth requires $\Omega(n\log n)$. Note that all of the other computations in this reduction take $O(n)$. $\square$

Note: This technique of reduction can be generalized to prove that computing a generalization of spherical and lens depth called $\beta-$skeleton depth ($\beta > 1$) [20] also requires $\Omega(n\log n)$ time.

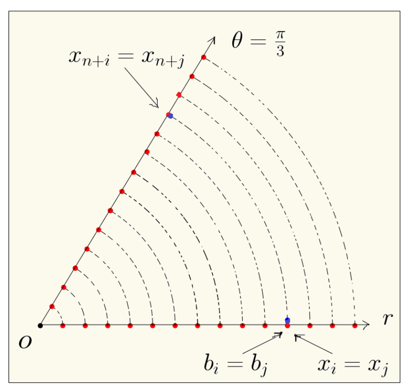

Figure 6: A representation of B, S, and duplications in these sets

3 Relationships Among Spherical Depth, Lens Depth, and Simplicial Depth

Theorem 4 *For $S \subset \mathbb{R}^d$ and $q \in \mathbb{R}^d$, $LD(q;S) \ge SphD(q;S)$.*

Proof. From the definition of the spherical (lens) influence regions of any arbitrary pair of points x_i and x_j

in S, it can be seen that $Sph(x_i, x_j) \subset L(x_i, x_j)$. Hence Equation (9) is sufficient to complete the proof.

$$\begin{aligned} SphD(q;S) &= \sum_{x_i,x_j \in S} I(q \in Sph(x_i,x_j)) \\ &\leq \sum_{x_i,x_j \in S} I(q \in L(x_i,x_j)) = LD(q;S) \end{aligned} \tag{9}$$

□

Definition: The simplicial depth of $q \in \mathbb{R}^d$ with respect to the data set $S = \{x_1, ..., x_n\} \subset \mathbb{R}^d$ is defined by:

$$SD(q;S) = \frac{1}{\binom{n}{d+1}} \sum_{(x_1,...,x_{d+1}) \in S} I(q \in Conv[x_1,...,x_{d+1}]), \tag{10}$$

where $Conv[x_1, ..., x_{d+1}]$ is a closed simplex formed by $d+1$ points of S [11].

Definition: For a point $q \in \mathbb{R}^2$ and a data set S consisting of n points in $\mathbb{R}^2$, we define $B_{in}(q;S)$ to be the set of all closed sphere areas, out of $\binom{n}{2}$ possible sphere areas, that contain q. We also define $S_{in}(q;S)$ to be the set of all closed simplices, out of $\binom{n}{3}$ possible closed simplices defined by S, that contain q.

Lemma 5 *Suppose that q is a point in a given convex hull H obtained from a data set S in $\mathbb{R}^2$. q is covered by the union of sphere areas defined by S.*

See the Appendix for a proof of this Lemma.

Lemma 6 *Suppose that $S = \{a, b, c\}$ is a set of points in $\mathbb{R}^2$. For every $q \in \mathbb{R}^2$, if $|S_{in}(q;S)| = 1$, then $|B_{in}(q;S)| \geq 2$.*

A proof of this Lemma can also be found in the Appendix. Another form of Lemma 6 is that if $q \in \triangle abc$, then q falls inside at least two sphere areas out of three sphere areas $Sph(a, b)$, $Sph(c, b)$, and $Sph(a, c)$.

Lemma 7 *For $S = \{x_1, ..., x_n\} \subset \mathbb{R}^2$,*

$$\frac{|B_{in}(q;S)|}{|S_{in}(q;S)|} \geq \frac{2}{n-2}.$$

Proof. We suppose that $Sph(x_i, x_j) \in B_{in}(q;S)$ (see Figure 7). There exist at most $(n-2)$ triangles in $S_{in}(q;S)$ such that x_ix_j is an edge of them. Let us consider $\triangle x_ix_jx_k$ from these triangles. Referring to Lemma 6, we know that q falls inside at least one of $Sph(x_i, x_k)$ and $Sph(x_j, x_k)$. It means that there exist at most $(n-2)$ triangles in $S_{in}(q;S)$ such that x_ix_k (respectively x_jx_k) is an edge of them. As can be seen, the triangle $\triangle x_ix_jx_k$ is counted at least two times, one time for $Sph(x_i, x_j)$ and one time for $Sph(x_i, x_k)$ (or $Sph(x_j, x_k)$). So, we can say that for every sphere area from $B_{in}(q;S)$, such as $Sph(x_i, x_j)$ there exist at most $\frac{(n-2)}{2}$ distinct triangles, triangles with only one common side, in $S_{in}(q;S)$. Consequently, (11) can be obtained.

$$\frac{|B_{in}(q;S)|}{|S_{in}(q;S)|} \geq \frac{2}{(n-2)} \tag{11}$$

□

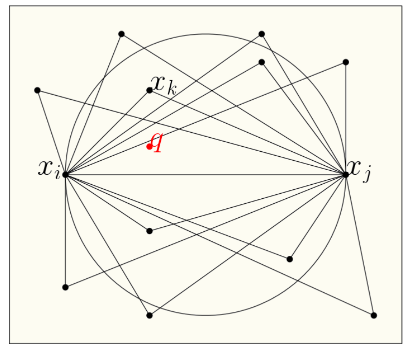

Figure 7: Sphere area $Sph(x_i, x_j)$ contains point q

Theorem 8 *For $q \in \mathbb{R}^2$ and a given data set S which consists of n points in $\mathbb{R}^2$, $SphD(q;S) \geq \frac{2}{3}\, SD(q;S)$.*

Proof. From the definitions of spherical depth and simplicial depth, it is clear that:

$$\frac{SphD(q;S)}{SD(q;S)} = \frac{\frac{|B_{in}(q;S)|}{\binom{n}{2}}}{\frac{|S_{in}(q;S)|}{\binom{n}{3}}} = \frac{|B_{in}(q;S)|}{|S_{in}(q;S)|} \times \frac{(n-2)}{3}. \tag{12}$$

From (12) and Lemma 7, it can be seen that

$$\frac{SphD(q;S)}{SD(q;S)} \geq \frac{2}{3} \Rightarrow SphD(q;S) \geq \frac{2}{3}\, SD(q;S).$$

□

4 Experiments

To support Theorem 8 and Theorem 4, we compute the spherical depth, lens depth, and the simplicial depth of the points in three random sets Q_1, Q_2, and Q_3 with respect to data sets S_1, S_2, and S_3, respectively. The elements of Q_i and S_i are some randomly generated points (double precision floating point) within the square $A = \{(x, y)|x, y \in [-10, 10]\}$. The results of our experiments are summarized in Table 1. Every cell in the table represents the corresponding depth of q_i with respect to data set S_i, where $q_i \in Q_i$. The cardinalities of Q_is and S_is are as follows: $|Q_1| = 100$, $|S_1| = 750$, $|Q_2| = 750$, $|S_2| = 2500$, $|Q_3| = 2500$, $|S_3| = 10000$. As can be seen in Table 1, there are some gaps between experimental bounds for random points and the theoretical bounds. These gaps motivate us to do more research in this area.

	$(t_1; S_1)$		$(t_2; S_2)$		$(t_3; S_3)$	
	Min	Max	Min	Max	Min	Max
SD	0.00	0.25	0.00	0.25	0.00	0.24
$SphD$	0.01	0.50	0.00	0.50	0.00	0.50
LD	0.05	0.61	0.05	0.61	0.04	0.61
$\frac{SphD}{SD}$	2.00	∞	2.00	∞	2.03	∞
$\frac{LD}{SD}$	2.43	∞	2.44	∞	2.44	∞
$\frac{LD}{SphD}$	1.21	8.11	1.22	23.16	1.22	157.16

Table 1: Experimental results

5 Conclusion

In this paper, we developed an optimal $\Theta(n \log n)$ algorithm to compute the spherical depth of a bivariate query point with respect to a given data set in $\mathbb{R}^2$. In addition to the time complexity, the main advantage of this algorithm is it simplicity of implementation. To obtain a lower bound for computing the planar spherical (lens) depth, we reduced the Element Uniqueness problem to the computing of spherical (lens) depth. We also investigated some geometric properties which lead us to find some theoretical relationships (i.e. $SphD \geq \frac{2}{3}SD$ and $LD \geq SphD$) among spherical depth, lens depth, and simplicial depth. Finally, some experimental results (i.e. $SphD \geq 2SD$ and $LD \geq 1.2\, SphD$) are provided. More research on this topic is needed to figure out if the real bounds are closer to the experimental bounds or to the current theoretical bounds.

References

[1] Aloupis, Greg. On computing geometric estimators of location. *PhD thesis, McGill University*, Montreal, Canada, 2001.

[2] Aloupis, Greg and Langerman, Stefan and Soss, Michael and Toussaint, Godfried. Algorithms for bivariate medians and a Fermat–Torricelli problem for lines. *Journal of Computational Geometry*, Vol. 26, 69-79, 2003.

[3] Ben-Or, Michael. Lower bounds for algebraic computation trees. *Proceedings of the fifteenth annual ACM symposium on Theory of computing*, 80-86, 1983.

[4] Bremner, David and Chen, Dan and Iacono, John and Langerman, Stefan and Morin, Pat. Output-sensitive algorithms for Tukey depth and related problems. *Journal of Statistics and Computing*, Springer , Vol. 18, 259-266,2008.

[5] Chan, Timothy M. An optimal randomized algorithm for maximum Tukey depth. *Proceedings of the fifteenth annual ACM-SIAM symposium on Discrete algorithms*, 430-436, 2004.

[6] Chen, Dan. Algorithms for Data Depth. *PhD thesis, Carleton University*, Ottawa, Canada, 2013.

[7] Christmann, Andreas. Regression depth and support vector machine. *DIMACS series in discrete mathematics and theoretical computer science*, AMS, Vol. 72, 71, 2006.

[8] Elmore, Ryan T and Hettmansperger, Thomas P and Xuan, Fengjuan. Spherical data depth and a multivariate median. *Journal of DIMACS Series in Discrete Mathematics and Theoretical Computer Science*, AMS, Vol. 72, 87-101,2006.

[9] Hotelling, Harold. Stability in competition. *The Collected Economics Articles of Harold Hotelling*, Springer, 50-63, 1990.

[10] Libeskind, Shlomo. Euclidean and transformational geometry: A deductive inquiry. *Jones & Bartlett Publishers*, 2008.

[11] Liu, Regina Y. On a notion of data depth based on random simplices. *Journal of The Annals of Statistics*, Vol. 18, 405-414, 1990.

[12] Liu, Regina Y and Serfling, Robert Joseph and Souvaine, Diane L. Data depth: robust multivariate analysis, computational geometry, and applications. *American Mathematical Soc.*, Vol. 72, 2006.

[13] Liu, Zhenyu and Modarres, Reza. *Lens data depth and median*. Journal of Nonparametric Statistics, Vol. 23, 1063-1974, 2011.

[14] Matousek, Jirí. Computing the center of planar point sets. *Journal of Computational Geometry: Papers from the DIMACS special year*, AMS, 221-230, 1991.

[15] Oja, Hannu. Descriptive statistics for multivariate distributions. *Journal of Statistics & Probability Letters*, Elsevier, Vol. 1, 327-332, 1983.

[16] Rousseeuw, Peter J and Hubert, Mia. Regression depth. *Journal of the American Statistical Association*, Taylor & Francis Group, Vol. 94, 388-402, 1999.

[17] Small, Christopher G. A survey of multidimensional medians. *Journal of International Statistical Review/Revue Internationale de Statistique*, JSTOR, 263-277,1990.

[18] Supowit, Kenneth J. The relative neighborhood graph, with an application to minimum spanning trees. *Journal of the ACM (JACM)*, Vol 30, N3, 428-448,1983.

[19] Tukey, John W. Mathematics and the picturing of data. *Proceedings of the international congress of mathematicians*. Vol. 2, 523-531,1975.

[20] Yang, Mengta. Depth Functions, Multidimensional Medians and Tests of Uniformity on Proximity Graphs. *PhD Thesis, The George Washington University, USA*, 2014.

Appendix

Lemma 1: For arbitrary points a, b, and t in $\mathbb{R}^2$, $t \in Sph(a,b)$ if and only if $\angle atb \geq \frac{\pi}{2}$.

Proof. If t is on the boundary of $Sph(a,b)$, the *Inscribed Angle Theorem* (Theorem 2.2 in [10]) suffices as the proof in both directions. For the rest of the proof, by $t \in Sph(a,b)$, we mean $t \in int\ Sph(a,b)$.

$\Rightarrow$) For $t \in Sph(a,b)$, suppose that $\angle atb < \pi/2$ (proof by contradiction). We continue the line segment $\overline{at}$ to cross the boundary of $Sph(a,b)$. Let t' be the crossing point (see Figure 8). Since $\angle atb < \frac{\pi}{2}$, then, $\angle btt'$ is greater than $\frac{\pi}{2}$. Let $\angle btt' = \frac{\pi}{2} + \epsilon_1; \epsilon_1 > 0$. From the Inscribed Angle Theorem, we know that $\angle at'b$ is a right angle. The angle $tbt' = \epsilon_2 > 0$ because $t \in Sph(a,b)$. Summing up the angles in $\triangle tt'b$, as computed in (13), leads to a contradiction. So, this direction of proof is complete.

$$\angle tt'b + \angle t'bt + \angle btt' \geq \frac{\pi}{2} + \epsilon_2 + (\frac{\pi}{2} + \epsilon_1) = \pi + \epsilon_1 + \epsilon_2 > \pi \quad (13)$$

$\Leftarrow$) If $\angle atb = \frac{\pi}{2} + \epsilon_1; \epsilon_1 > 0$, we prove that $t \in Sph(a,b)$. Suppose that $t \notin Sph(a,b)$ (proof by contradiction). Since $t \notin Sph(a,b)$, at least one of the line segments $\overline{at}$ and $\overline{bt}$ crosses the boundary of $Sph(a,b)$. Without loss of generality, assume that $\overline{at}$ is the one that crosses the boundary of $Sph(a,b)$ at the point t' (see Figure 9). Considering the Inscribed Angle Theorem, we know that $\angle at'b = \frac{\pi}{2}$ and consequently, $\angle bt't = \frac{\pi}{2}$. The angle $\angle t'bt = \epsilon_2 > 0$ because $t \notin Sph(a,b)$. If we sum up the angles in the triangle $\triangle tt'b$, the same contradiction as in (13) will be implied. □

Lemma 5: Suppose that q is a point in a given convex hull H obtained from a data set S in $\mathbb{R}^2$. q is covered by the union of sphere areas defined by S.

Proof. It can be seen that there is at least one triangle, defined by the vertices of H, that contains q. We prove that the union of the sphere areas defined by such triangle contains q. See Figure 10. We prove this statement by contradiction. Suppose that q is covered by none of $Sph(a,b)$, $Sph(a,c)$, and $Sph(b,c)$. Therefore, Lemma 1 implies that none of the angles $\angle aqb$, $\angle aqc$, and $\angle bqc$ is greater than or equal to $\frac{\pi}{2}$ which is a contradiction because at least one of these angles should be at least $\frac{2\pi}{3}$ in order to get 2π as their sum. □

Lemma 6: Suppose that $S = \{a,b,c\}$ is a set of points in $\mathbb{R}^2$. For every $q \in \mathbb{R}^2$, if $|S_{in}(q;S)| = 1$, then $|B_{in}(q;S)| \geq 2$.

Proof. We prove the lemma by contradiction. By Lemma 5, $B_{in}(q;S) \geq 1$. Suppose that $|B_{in}(q;S)| = 1$. If q is located on the vertices of $\triangle abc$, it clear that $|B_{in}(q;S)| \geq 2$ thus, we suppose that q is not located on the vertices of $\triangle abc$. Without loss of generality, we suppose that q falls inside $Sph(a,b)$. For the rest of the proof, we focus on the relationships among the angles $\angle aqb$, $\angle cqa$, and $\angle cqb$ (see Figure 10). Since q is inside $\triangle abc$, $\angle aqb \leq \pi$. Consequently, at least one of $\angle cqa$ and $\angle cqb$ is greater than or equal to $\frac{\pi}{2}$.

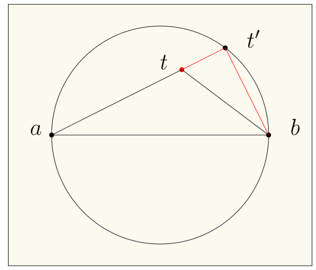

Figure 8: $t \in Sph(a,b)$

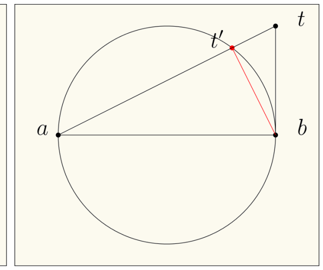

Figure 9: $t \notin Sph(a,b)$

So, Lemma 1 implies that q will fall inside at least one of $Sph(a,c)$ and $Sph(b,c)$. Hence, $|B_{in}(q;S)| = 1$ contradicts $|S_{in}(q;S)| = 1$. This means that the case $|B_{in}(q;S)| \geq 2$. As an illustration, in Figure 10, for the points inside the hatched area $|B_{in}(q;S)| = 3$.

□

Lemma 9 *$L_j = \{x_{(n+j) \bmod 2n}\}$ if every x_j $(1 \leq j \leq 2n)$ is a unique element in S , where $S = \{(b_i, 0), (b_i, \pi/3) \mid b_i > 0, 1 \leq i \leq n\}$, and $L_j = \{x_k \in S \mid q \in L(x_j, x_k)\}$.*

Proof. Suppose that $L_j = \{x_k, x_{(n+j) \bmod 2n}\}$ for some $x_k \in S$ $(k \neq j)$. We prove that such x_k does not exist. If $\angle x_j O x_k = 0$, it is obvious that $O \notin L(x_j, x_k)$ which means that x_k cannot be an element of L_j. For the case $\angle x_j O x_k = \pi/3$, let us assume that $O \in L(x_j, x_k)$ which is equivalent with $d(x_j, x_k) \geq d(O, x_k)$ and $d(x_j, x_k) \geq d(O, x_j)$. From the definitions $d(O, x_k) = b_k$, $d(O, x_j) = b_j$, and from the cosine formula, $d^2(x_j, x_k) = b_j^2 + b_k^2 - 2b_k b_j cos(\pi/3)$. Therefore,

$$d(x_j, x_k) \geq d(O, x_k) \Rightarrow b_j^2 - b_j b_k \geq 0 \Rightarrow b_j - b_k \geq 0 \Rightarrow b_j \geq b_k$$

and

$$d(x_j, x_k) \geq d(O, x_j) \Rightarrow b_k^2 - b_j b_k \geq 0 \Rightarrow b_k - b_j \geq 0 \Rightarrow b_k \geq b_j.$$

□

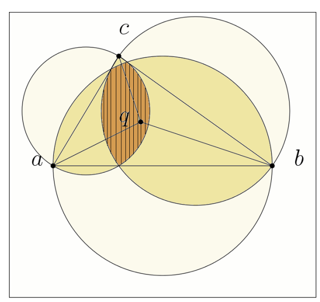

Figure 10: Triangle abc contains point q