

# Approximate Data Depth Revisited

David Bremner \*

Rasoul Shahsavari \*<sup>\*</sup>

## Abstract

Halfspace depth and  $\beta$ -skeleton depth are two types of depth functions in nonparametric data analysis. The halfspace depth of a query point  $q \in \mathbb{R}^d$  with respect to  $S \subset \mathbb{R}^d$  is the minimum portion of the elements of  $S$  which are contained in a halfspace which passes through  $q$ . For  $\beta \geq 1$ , the  $\beta$ -skeleton depth of  $q$  with respect to  $S$  is defined to be the total number of  $\beta$ -skeleton influence regions that contain  $q$ , where each of these influence regions is the intersection of two hyperballs obtained from a pair of points in  $S$ . The  $\beta$ -skeleton depth introduces a family of depth functions that contain *spherical depth* and *lens depth* if  $\beta = 1$  and  $\beta = 2$ , respectively. The main results of this paper include approximating the planar halfspace depth and  $\beta$ -skeleton depth using two different approximation methods. First, the halfspace depth is approximated by the  $\beta$ -skeleton depth values. For this method, two dissimilarity measures based on the concepts of *fitting function* and *Hamming distance* are defined to train the halfspace depth function by the  $\beta$ -skeleton depth values obtained from a given data set. The goodness of this approximation is measured by a function of error values. Secondly, computing the planar  $\beta$ -skeleton depth is reduced to a combination of some range counting problems. Using existing results on range counting approximations, the planar  $\beta$ -skeleton depth of a query point is approximated in  $O(n \text{ poly}(1/\varepsilon, \log n))$ ,  $\beta \geq 1$ . Regarding the  $\beta$ -skeleton depth functions, it is also proved that this family of depth functions converge when  $\beta \rightarrow \infty$ . Finally, some experimental results are provided to support the proposed method of approximation and convergence of  $\beta$ -skeleton depth functions.

## 1 Introduction

Data depth is a method to generalize the concept of rank in the univariate data analysis to the multivariate case. Data depth measures the centrality of a data point with respect to a dataset, and it gives a center-outward ordering of data points. In other words, applying a data depth on a dataset generates a partial ordered set (poset) of the data points. A poset is a set together with a partial ordering relation which is

reflexive, antisymmetric and transitive. Over the last decades various notions of data depth such as *halfspace depth* (Hotelling [15, 28]; Tukey [30]), *simplicial depth* (Liu [19]) *Oja depth* (Oja [21]), *regression depth* (Rousseeuw and Hubert [22]), and others have been introduced in the area of non-parametric multivariate data analysis. These depth functions are different in application, definition, and complexity of computation. Among the different notions of data depth, we focus on halfspace depth and a recently defined data depth named  $\beta$ -skeleton depth (Yang and Modarres [32]).

In 1975, Tukey generalized the definition of univariate median and defined the halfspace median as a point in which the halfspace depth is maximized, where the halfspace depth is a multivariate measure of centrality of data points. Halfspace depth is also known as Tukey depth or location depth. In general, the halfspace depth of a query point  $q$  with respect to a given data set  $S$  is the smallest portion of data points that are contained in a closed halfspace through  $q$  [6, 30]. The halfspace depth function has various properties such as vanishing at infinity, affine invariance, and decreasing along rays. These properties are proved in [10]. Many different algorithms for the computation of halfspace depth in lower dimensions have been developed elsewhere [6, 7, 8, 24]. The bivariate and trivariate case of halfspace depth can be computed exactly in  $O(n \log n)$  and  $O(n^2 \log n)$  time [23, 29], respectively. However, computing the halfspace depth of a query point with respect to a data set of size  $n$  in dimension  $d$  is an NP-hard problem if both  $n$  and  $d$  are part of the input [16]. Due to the hardness of the problem, designing efficient algorithms to compute (or approximate) the halfspace depth of a point remains an interesting task in the research area of data depth [5, 9]. Some results on  $\varepsilon$ -approximation of halfspace depth can be found in [1, 3, 14].

In 2017, Yang and Modarres introduced a family of depth functions called  $\beta$ -skeleton depth, indexed by a single parameter  $\beta \geq 1$ . The  $\beta$ -skeleton depth of a query point  $q \in \mathbb{R}^d$  with respect to a given data set  $S$  is defined as the portion of  $\beta$ -skeleton influence regions that contain  $q$ . The influence regions of  $\beta$ -skeleton depth are the multidimensional generalization of lunes in the definition of the  $\beta$ -skeleton graph [17]. A notable characteristic of the  $\beta$ -skeleton depth is related to its

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

time complexity that grows linearly in the dimension  $d$  whereas no polynomial algorithms (in the dimension) in higher dimensions are known for most other data depths. To the best of our knowledge, the current best algorithm for computing the  $\beta$ -skeleton depth in higher dimension  $d$  is the straightforward algorithm which takes  $\Theta(dn^2)$ . The authors, in their previous work [26], improved this bound for the planar  $\beta$ -skeleton depth. They developed an  $O(n^{3/2+\epsilon})$  algorithm for all values of  $\beta \geq 1$ , and a  $\Theta(n \log n)$  algorithm for the special case of  $\beta = 1$ . Spherical depth (Elmore, Hettmansperger, and Xuan [11]) and lens depth (Liu and Modarres [20]) can be obtained from  $\beta$ -skeleton depth by considering  $\beta = 1$  and  $\beta = 2$ , respectively. It is known that the  $\beta$ -skeleton depth function is monotonic, maximized at the center, and vanishing at infinity. The  $\beta$ -skeleton depth function is also orthogonally (affinely) invariant if the Euclidean (Mahalanobis) distance is used to construct the  $\beta$ -skeleton influence regions [11, 20, 31, 32].

The concept of data depth is widely studied by statisticians and computational geometers. Some directions that have been considered by researchers include defining new depth functions, improving the complexity of computations, computing both exact and approximate depth values, and computing depth functions in higher dimensions. Two surveys by Aloupis [2] and Small [28] can be referred to as overviews of data depth from a computational geometer's and a statistician's point of view, respectively.

In this paper, different methods are presented to approximate the halfspace and  $\beta$ -skeleton depth functions. Computing the  $\beta$ -skeleton depth is reduced to a combination of range counting problems. Using different range counting approximations in [4, 13, 27], the planar  $\beta$ -skeleton depth ( $\beta \geq 1$ ) of a given point is approximated in  $O(n \text{ poly}(1/\epsilon, \log n))$  query time. Furthermore, we propose a technique to approximate the halfspace depth using the  $\beta$ -skeleton depth. In this method, two dissimilarity measures based on the concepts of *fitting function* and *Hamming distance* are defined to train the halfspace depth function by the  $\beta$ -skeleton depth values obtaining from a given data set. The goodness of approximation can be measured by the sum of square of error values. We also show that  $\beta$ -skeleton depth functions converge when  $\beta \rightarrow \infty$ . Finally, some experimental results are provided regarding our proposed method of approximation.

## 2 Halfspace Depth

**Definition:** The halfspace depth of a query point  $q \in \mathbb{R}^d$  with respect to a given data set  $S = \{x_1, \dots, x_n\} \subseteq \mathbb{R}^d$  is defined as the minimum portion of points of  $S$  con-

tained in any closed halfspace that has  $q$  on its boundary. Using the notation of  $HD(q; S)$ , the above definition can be presented by (1).

$$HD(q; S) = \frac{2}{n} \min\{|S \cap H| : H \in \mathbb{H}\}, \quad (1)$$

where  $2/n$  is the normalization factor<sup>1</sup>,  $\mathbb{H}$  is the class of all closed halfspaces in  $\mathbb{R}^d$  that pass through  $q$ , and  $|S \cap H|$  denotes the number of points within the intersection of  $S$  and  $H$ . As illustrated in Figure 1,  $HD(q_1; S) = 6/13$  and  $HD(q_2; S) = 0$ , where  $S$  is a given set of points in the plane and  $q_1, q_2$  are two query points not in  $S$ .

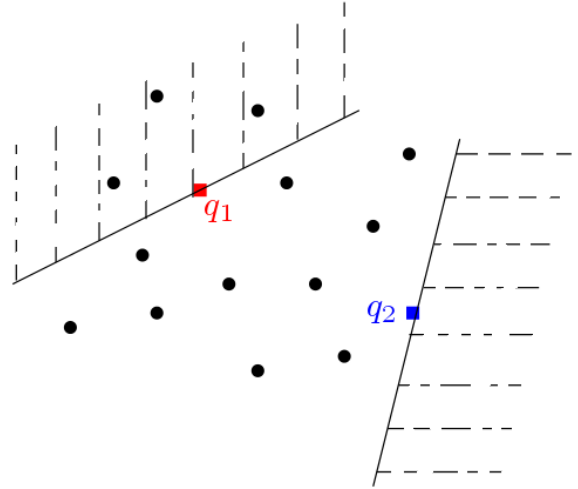


Figure 1: Two examples of halfspace depth in the plane

## 3 $\beta$ -skeleton Depth

**Definition:** For  $1 \leq \beta \leq \infty$ , the  $\beta$ -skeleton influence region of  $x_i$  and  $x_j$  ( $S_\beta(x_i, x_j)$ ) is defined as follows:

$$S_\beta(x_i, x_j) = B(c_i, r) \cap B(c_j, r), \quad (2)$$

where  $r = \frac{\beta}{2} \|x_i - x_j\|$ ,  $c_i = \frac{\beta}{2} x_i + (1 - \frac{\beta}{2}) x_j$ , and  $c_j = (1 - \frac{\beta}{2}) x_i + \frac{\beta}{2} x_j$ .

In the case of  $\beta = \infty$ , the  $\beta$ -skeleton influence region is well defined, and it is a slab defined by two halfspaces. Figure 2 shows the  $\beta$ -skeleton influence regions for different values of  $\beta$ .

**Definition:** Let  $S = \{x_1, \dots, x_n\}$  be a set of points in  $\mathbb{R}^d$ . For the parameter  $1 \leq \beta \leq \infty$ , the  $\beta$ -skeleton depth of a query point  $q \in \mathbb{R}^d$  with respect to  $S$ , is defined as the proportion of the  $\beta$ -skeleton influence regions  $S_\beta(x_i, x_j)$ ,  $1 \leq i < j \leq n$  that contain  $q$ . Using the

<sup>1</sup>Instead of the normalization factor  $1/n$  which is common in literature, we use the normalization factor  $2/n$  in order to let the depth of 1 to be achievable.

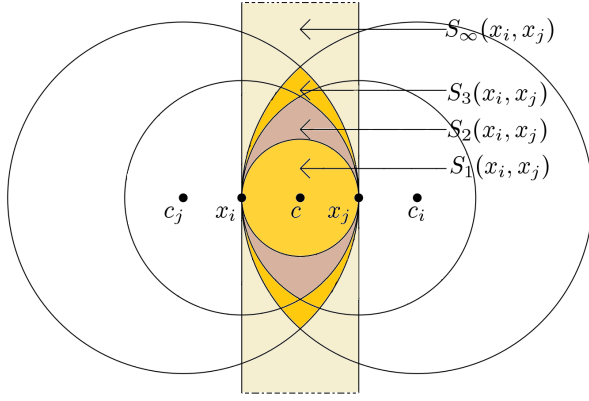


Figure 2: The  $\beta$ -skeleton influence regions defined by  $x_i$  and  $x_j$  for  $\beta=1, 2, 3$ , and  $\beta = \infty$ , where  $c = \frac{x_i+x_j}{2}$ ,  $c_i = \frac{3}{2}x_i + (1 - \frac{3}{2})x_j$ , and  $c_j = (1 - \frac{3}{2})x_i + \frac{3}{2}x_j$

indicator function  $I$ , this definition can be represented by Equation (3).

$$SkD_\beta(q; S) = \frac{1}{\binom{n}{2}} \sum_{1 \leq i < j \leq n} I(q \in S_\beta(x_i, x_j)) \quad (3)$$

It can be verified that  $q \in S_\beta(x_i, x_j)$  is equivalent to the inequality of  $\frac{\beta}{2} \|x_i - x_j\| \geq \max\{\|q - c_i\|, \|q - c_j\|\}$ . The straightforward algorithm for computing the  $\beta$ -skeleton depth of  $q \in \mathbb{R}^d$  takes  $\Theta(dn^2)$  time because the above inequality should be checked for all  $1 \leq i, j \leq n$ .

## 4 Dissimilarity Measures

In this section, two different types of dissimilarity measures for depth functions are introduced.

### 4.1 Fitting Functions and Dissimilarity Measures

To determine the dissimilarity between two vectors  $U = (u_1, \dots, u_n)$  and  $V = (v_1, \dots, v_n)$ , the idea of fitting functions can be applied. Considering the goodness measures of fitting functions, assume that  $f$  is the best function fitted to  $U$  and  $V$  which means that  $u_i = f(v_i) \pm \delta_i$ . Let  $\xi_i = u_i - \bar{U}$ , where  $\bar{U}$  is the average of  $u_i$  ( $1 \leq i \leq n$ ). We define the dissimilarity measure between  $U$  and  $V$  ( $d_E(U, V)$ ) to be a function of  $\delta_i$  and  $\xi_i$  as follows:

$$d_E(U, V) = 1 - r^2, \quad (4)$$

where  $r^2$  is the coefficient of determination [25] which is defined by:

$$r^2 = \frac{\sum_{i=1}^n (\xi_i^2 - \delta_i^2)}{\sum_{i=1}^n \xi_i^2}. \quad (5)$$

Since  $r^2 \in [0, 1]$ ,  $d_E(U, V) \in [0, 1]$ . A smaller value of  $d_E(U, V)$  represents more similarity between  $U$  and  $V$ .

**Definition:** Let  $S = \{x_1, \dots, x_n\}$  be a finite set. It is said that  $P = (S, \preceq)$  is a partially ordered set (poset) if  $\preceq$  is a partial order relation on  $S$ , that is, for all  $x_i, x_j, x_t \in S$ : (a)  $x_i \preceq x_i$ ; (b)  $x_i \preceq x_j$  and  $x_j \preceq x_t$  implies that  $x_i \preceq x_t$ ; (c)  $x_i \preceq x_j$  and  $x_j \preceq x_i$  implies that  $x_i \equiv_p x_j$ , where  $\equiv_p$  is the corresponding equivalence relation.

Poset  $P = (S, \preceq)$  is called a chain if any two elements of  $S$  are comparable, i.e., given  $x_i, x_j \in S$ , either  $x_i \preceq x_j$  or  $x_j \preceq x_i$ . If there is no comparable pair among the elements of  $S$ , the corresponding poset is an anti chain. Figure 3 illustrates different posets with the same elements.

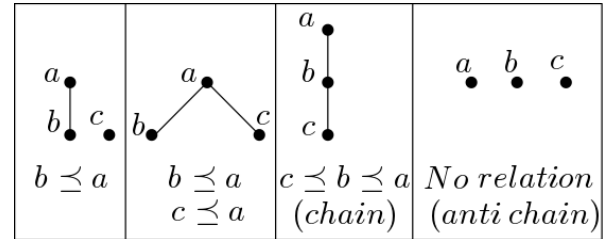


Figure 3: Different posets and relations among their elements

### 4.2 Dissimilarity Measures Between two Posets

The idea of defining the following distance comes from the proposed structural dissimilarity measure between posets in [12]. Let  $\mathbb{P} = \{P_t = (S, \preceq_t) | t \in \mathbb{N}\}$  be a finite set of posets, where  $S = \{x_1, \dots, x_n\}$ . For  $P_k \in \mathbb{P}$  we define a matrix  $M_{n \times n}^k$  by:

$$M_{ij}^k = \begin{cases} 1 & x_i \preceq_k x_j \\ 0 & \text{otherwise.} \end{cases}$$

We use the notation of  $d_c(P_f, P_g)$  to define a dissimilarity between two posets  $P_f, P_g \in \mathbb{P}$  as follows:

$$d_c(P_f, P_g) = \frac{\sum_{i=1}^n \sum_{j=1}^n |M_{ij}^f - M_{ij}^g|}{n^2 - n} \quad (6)$$

It can be verified that  $d_c(P_f, P_g) \in [0, 1]$ , where the closer value to 1 means the less similarity between  $P_f$  and  $P_g$ . This measure of similarity is a metric on  $\mathbb{P}$  because for all  $P_f, P_g, P_h \in \mathbb{P}$ ,

- $d_c(P_f, P_g) \geq 0$
- $d_c(P_f, P_g) = 0 \Leftrightarrow P_f = P_g$

- $d_c(P_f, P_g) = d_c(P_g, P_f)$
- $d_c(P_f, P_h) \leq d_c(P_f, P_g) + d_c(P_g, P_h)$ .

Proving these properties is straightforward. The proof of last property which is less trivial can be found in Appendix (see Lemma 3).

## 5 Approximation of Halfspace Depth

We propose a method, which is different from other approximation methods in the literatures, to approximate the halfspace depth. Motivated by statistical applications and machine learning techniques, we train the halfspace depth function using the values of another depth function. Among all depth functions, the  $\beta$ -skeleton depth is chosen to approximate the halfspace depth because it is easy to compute and its time complexity, i.e.  $O(dn^2)$ , grows linearly in higher dimension  $d$ .

### 5.1 Approximation of Halfspace Depth and Fitting Function

Suppose that  $S = \{x_1, \dots, x_n\}$  is a set of data points. By choosing some subsets of  $S$  as training samples, we consider the problem of learning the halfspace depth function using the  $\beta$ -skeleton depth values. Finally, by applying the cross validation techniques in machine learning, the best function  $f$  can be obtained such that  $HD(x_i; S) = f(SkD_\beta(x_i; S)) \pm \delta_i$ . The function  $f$  can be considered as an approximation function for halfspace depth, where the value of  $d_E(HD, SkD_\beta)$  that can be computed using Equation (4) is the error of approximation.

### 5.2 Approximation of Halfspace Depth and Poset Dissimilarity

In some applications, the structural ranking among the elements of  $S$  is more important than the depth value of single points. Let  $S = \{x_1, \dots, x_n\}$  be a set of points and  $D$  be a depth function. Applying  $D$  on  $x_i$  with respect to  $S$  generates a poset. In fact,  $P_D = (D(x_i; S), \leq)$  is a chain because for every  $x_i, x_j \in S$ , the values of  $D(x_i; S)$  and  $D(x_j; S)$  are comparable. For halfspace depth and  $\beta$ -skeleton depth, their dissimilarity measure of rankings can be obtained by Equation (6) as follows:

$$d_c(HD, SkD_\beta) = \frac{\sum_{i=1}^n \sum_{j=1}^n |M_{ij}^{HD} - M_{ij}^{SkD_\beta}|}{n^2 - n}.$$

The smaller value of  $d_c(HD, SkD_\beta)$ , the more similarity between  $HD$  and  $SkD_\beta$  in ordering the elements of  $S$ .

In sections 5.1 and 5.2, instead of  $\beta$ -skeleton, any other depth function can be considered to approximate

halfspace depth. Considering any other depth function, we can compute the goodness of approximation using dissimilarity measures  $d_E$  and  $d_c$ .

**Conjecture 1** *For two depth functions  $D_1$  and  $D_2$ , the small value of  $d_c(D_1, D_2)$  implies the small value of  $d_E(D_1, D_2)$  and vice versa.*

## 6 Approximation of $\beta$ -skeleton Depth

The convergence of  $\beta$ -skeleton depth functions when  $\beta \rightarrow \infty$  is investigated in this section. Furthermore, the problem of planar  $\beta$ -skeleton depth is reduced to a combination of disk and halfspace range counting problems. This reduction is applied to approximate the planar  $\beta$ -skeleton depth using range counting approximations.

### 6.1 Convergence of $\beta$ -skeleton Depth Functions

The following theorem helps understand the definition of  $SkD_\infty$  in Equation (3).

**Theorem 1** *If  $\beta \rightarrow \infty$ , all  $\beta$ -skeleton depth functions converge to  $SkD_\infty$ . In other words, for data set  $S = \{x_1, \dots, x_n\}$  and query point  $q$ ,*

$$\lim_{\beta \rightarrow \infty} \frac{SkD_\beta(q; S)}{SkD_{(\beta+1)}(q; S)} = 1. \quad (7)$$

**Proof.** Referring to (3), the definition of  $\beta$ -skeleton depth, it is enough to prove that if  $\beta \rightarrow \infty$ ,

$$\forall x_i, x_j \in S; S_\beta(x_i, x_j) = S_{(\beta+1)}(x_i, x_j). \quad (8)$$

It is proved that  $S_\beta(x_i, x_j) \subseteq S_{(\beta+1)}(x_i, x_j)$ , and  $u_\beta = u_{(\beta+1)}$  if  $\beta \rightarrow \infty$  (see Lemma 4 and 5 in Appendix), where  $u_\beta$  is a fixed intersection point of  $B(c_i, r)$  and  $B(c_j, r)$  (see figure 4). Since the influence regions of  $\beta$ -skeleton depth functions are closed and convex, the proof is complete if

$$\lim_{\beta \rightarrow \infty} \frac{A(S_{\beta ij})}{A(S_{(\beta+1)ij})} = 1, \quad (9)$$

where  $A(S_{\beta ij})$  is the area of  $S_\beta(x_i, x_j)$ . It can be verified that  $A(S_{\beta ij})$  is equal to  $\theta r^2 - da$ , where  $l = d(x_i, x_j)$ ,  $a = (l/2)\sqrt{(2\beta-1)}$ ,  $r = \beta l/2$ ,  $\theta = \cos^{-1}((\beta-1)/\beta)$ , and  $d = (\beta-1)l$ . See Figure 4.

$$\frac{A(S_{\beta ij})}{A(S_{(\beta+1)ij})} = \frac{\beta^2 \cos^{-1}[(\beta-1)/\beta] - (\beta-1)\sqrt{2\beta-1}}{(\beta+1)^2 \cos^{-1}[\beta/(\beta+1)] - \beta\sqrt{2\beta+1}}$$

Equation (9) is proved because

$$\lim_{\beta \rightarrow \infty} \cos^{-1}\left(\frac{\beta-1}{\beta}\right) = \lim_{\beta \rightarrow \infty} \cos^{-1}\left(\frac{\beta}{\beta+1}\right) = 0.$$

□

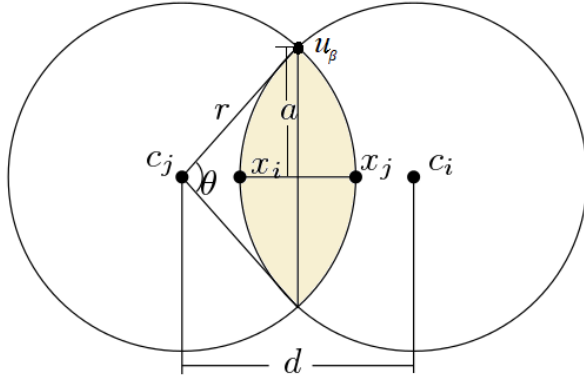


Figure 4: An illustration of  $d$ ,  $a$ ,  $r$ ,  $l$ , and  $\theta$ .

## 6.2 $\varepsilon$ -approximation of Planar $\beta$ -skeleton Depth

In this section, we prove that for  $S = \{x_1, \dots, x_n\} \subset \mathbb{R}^2$  and  $q \in \mathbb{R}^2$ , computing  $SkD_\beta(q; S)$  is equivalent to at most  $3n$  semialgebraic range counting problems. This result can be applied to approximate the planar  $\beta$ -skeleton depth in  $O(n \text{ poly}(1/\varepsilon, \log n))$  query time with  $O(n \text{ poly}(1/\varepsilon, \log n))$  storage and the same bound for preprocessing time.

Given a set  $S = \{x_1, \dots, x_n\} \subset \mathbb{R}^2$ , a semialgebraic range  $\tau$  with constant description complexity, and a parameter  $\varepsilon > 0$ ,  $S$  can be preprocessed into a data structure that helps to efficiently compute an approximated count  $\eta_\tau$  which satisfies the  $(\varepsilon, \tau)$ -approximation given by:

$$(1 - \varepsilon)|S \cap \tau| \leq \eta_\tau \leq (1 + \varepsilon)|S \cap \tau|.$$

Considering  $S$  and  $\varepsilon$  as introduced above, a short list of recent results for approximation of some semialgebraic range counting problems in  $\mathbb{R}^2$  is as follows. To obtain an  $\varepsilon$ -approximation of  $\beta$ -skeleton depth, we use these results as black boxes.

- **Halfspace:** With  $O(n \text{ poly}(1/\varepsilon, \log n))$  preprocessing time, one can construct a data structure of size  $O(n \text{ poly}(1/\varepsilon, \log n))$  such that for a given halfspace  $h$ , it outputs a number  $\eta_h$  that satisfies the  $(\varepsilon, h)$ -approximation with query time  $O(\text{poly}(1/\varepsilon, \log n))$ . For points in  $\mathbb{R}^3$ , the same results can be obtained [13].
- **Disk:** By standard lifting of planar points to the paraboloid in  $\mathbb{R}^3$ , one can reduce a disk range query in the plane to a halfspace range query in three dimensions [13].
- **Circular cap:** A circular cap is the larger part of a circle cut by a line. In near linear time,  $S$  can be preprocessed into a linear size data structure that returns the emptiness result of any circular cap

query in  $O(\text{poly} \log n)$ . With  $O(\varepsilon^{-2}T(n) \log n)$  preprocessing time and  $O(\varepsilon^{-2}S(n) \log n)$  storage, one can construct a data structure such that for a given circular cap  $\rho$ , it returns in  $O(\varepsilon^{-2}Q(n) \log n)$  time, a number  $\eta_\rho$  that satisfies the  $(\varepsilon, \rho)$ -approximation. Both of  $T(n)$  and  $S(n)$  are near linear and  $Q(n)$  is the time for an emptiness query of  $\rho$  [27].

**Definition:** For an arbitrary non-zero point  $a \in \mathbb{R}^2$  and parameter  $\beta \geq 1$ ,  $\ell(p)$  is a line that is perpendicular to  $\vec{a}$  at the point  $p = p(a, \beta) = (\beta - 1)a/\beta$ . This line forms two halfspaces  $H_o(p)$  and  $H_a(p)$ . The one that includes the origin is  $H_o(p)$  and the other one that includes  $a$  is  $H_a(p)$ .

**Definition:** For a disk  $B(c, r)$  with center  $c = c(a, \beta) = \beta a / (2(\beta - 1))$  and radius  $r = \|c\|$ ,  $B_o(c, r)$  is the intersection of  $H_o(p)$  and  $B(c, r)$ , and  $B_a(c, r)$  is the intersection of  $H_a(p)$  and  $B(c, r)$ , where  $\beta > 1$  and  $a$  is an arbitrary non-zero point in  $\mathbb{R}^2$ .

Figure 5 is an illustration of these definitions for different values of parameter  $\beta$ .

**Theorem 2** For arbitrary non-zero points  $a, b$  in  $\mathbb{R}^2$  and parameter  $\beta > 1$ ,  $b \in H_o(p) \setminus \{\text{int}B_o(c, r)\}$  if and only if the origin  $O = (0, 0)$  is contained in  $S_\beta(a, b)$ , where  $c = \beta a / (2(\beta - 1))$ ,  $r = \|c\|$ , and  $p = (\beta - 1)a/\beta$ .

**Proof.** First, we show that  $B_o(c, r)$  is a well-defined set meaning that  $\ell(p)$  intersects  $B(c, r)$ . We compute  $d(c, \ell(p))$ , the distance of  $c$  from  $\ell(p)$ , and prove that this value is not greater than  $r$ . It can be verified that  $d(c, \ell(p)) = d(c, p)$ . Let  $k = \beta / (2(\beta - 1))$ ; the following calculations complete this part of the proof.

$$\begin{aligned} d(c, p) &= d\left(\frac{\beta a}{2(\beta - 1)}, \frac{(\beta - 1)a}{\beta}\right) = d\left(ka, \frac{1}{2k}a\right) \\ &= \left(k - \frac{1}{2k}\right) \sqrt{(a_x^2 + a_y^2)} = \left(\frac{2k^2 - 1}{2k}\right) \|a\| \\ &\leq \frac{2k^2}{2k} \|a\| = k\|a\| = r \end{aligned}$$

We recall the definition of  $\beta$ -influence region given by  $S_\beta(a, b) = B(c_a, \frac{\beta}{2}\|a - b\|) \cap B(c_b, \frac{\beta}{2}\|a - b\|)$ , where  $c_a = \frac{\beta}{2}a + (1 - \frac{\beta}{2})b$  and  $c_b = \frac{\beta}{2}b + (1 - \frac{\beta}{2})a$ . Using this definition, the following equivalencies can be derived from  $O \in S_\beta(a, b)$ .

$$O \in S_\beta(a, b) \Leftrightarrow \frac{\beta\|a - b\|}{2} \geq \max\{\|c_a\|, \|c_b\|\} \Leftrightarrow$$

$$\begin{aligned} \beta\|a - b\| &\geq \max\{\|\beta(a - b) + 2b\|, \|\beta(b - a) + 2a\|\} \Leftrightarrow \\ \beta^2\|a - b\|^2 &\geq \max\{\|\beta(a - b) + 2b\|^2, \|\beta(b - a) + 2a\|^2\} \end{aligned}$$

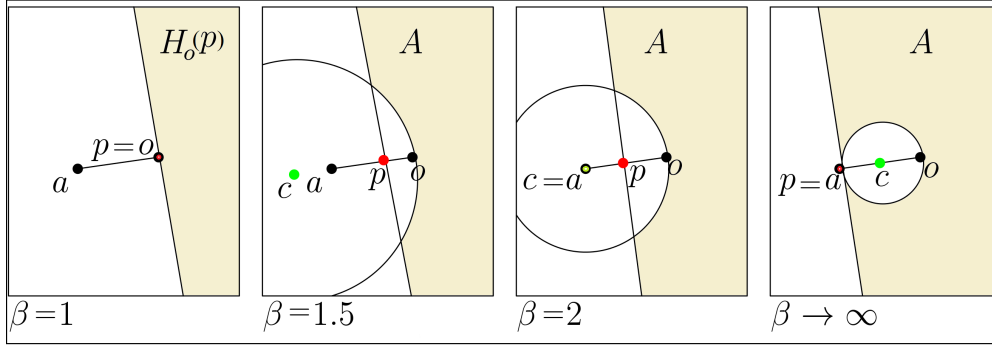


Figure 5: The  $H_o(p)$  and  $B(c, r)$  defined by  $a \in \mathbb{R}^2$  for  $\beta = 1, 1.5, 2$ , and  $\beta \rightarrow \infty$ , where  $A = H_o(p) \setminus \{intB_o(c, r)\}$

$$\Leftrightarrow 0 \geq \max\{b^2(1 - \beta) + \beta \vec{a} \cdot \vec{b}, a^2(1 - \beta) + \beta \vec{a} \cdot \vec{b}\}.$$

By solving these inequalities for  $(\beta - 1)/\beta$  which is equal to  $1/2k$ , we have:

$$\frac{1}{2k} \geq \max\left\{\frac{\vec{a} \cdot \vec{b}}{\|a\|^2}, \frac{\vec{a} \cdot \vec{b}}{\|b\|^2}\right\}. \quad (10)$$

For a fixed point  $a$ , the inequalities in Equation (10) determine one halfspace and one disk given by (11) and (12), respectively.

$$\frac{1}{2k} \geq \frac{\vec{a} \cdot \vec{b}}{\|a\|^2} \Leftrightarrow \vec{a} \cdot \vec{b} \leq \frac{1}{2k} \|a\|^2. \quad (11)$$

$$\begin{aligned} \frac{1}{2k} \geq \frac{\vec{a} \cdot \vec{b}}{\|b\|^2} &\Leftrightarrow b^2 - 2k \vec{a} \cdot \vec{b} \geq 0 \\ &\Leftrightarrow b^2 - 2k \vec{a} \cdot \vec{b} + k^2 a^2 \geq k^2 a^2 \\ &\Leftrightarrow (b - ka)^2 \geq (k\|a\|)^2. \end{aligned} \quad (12)$$

The proof is complete because for a point  $a$ , the set of all points  $b$  contained in the feasible region defined by Equations (11) and (12) is equal to  $H_o(p) \setminus \{intB_o(c, r)\}$ .  $\square$

**Note 1:** It can be verified that for the value of  $\beta = 2 + \sqrt{2}$ , the given halfspace in (11) passes through the center of the given disk in (12). See Lemma 6 in Appendix.

**Note 2:** The range counting approximations work after translating the points of  $S$  by  $-q$ .

For  $S = \{x_1, \dots, x_n\} \subset \mathbb{R}^2$  and  $x_i \in S$ , let  $h(x_i)$  be the given halfspace by (11), and  $B(x_i)$  be the given disk by (12). From Theorem 2, it can be deduced that we need to do

- $n$  halfspace range counting approximations to approximate  $SkD_1(q; S)$  because

$$SkD_1(q; S) = \frac{1}{2} \sum_{i=1}^n |h(x_i)|.$$

- $n$  halfspace and  $n$  disk range counting approximations to approximate  $SkD_\infty(q; S)$  because

$$SkD_\infty(q; S) = \frac{1}{2} \sum_{i=1}^n (|h(x_i)| - |B(x_i)|).$$

- $n$  halfspace,  $n$  disk, and  $n$  circular cap range counting approximations to approximate  $SkD_\beta(q; S)$ , for  $1 < \beta < 2 + \sqrt{2}$ , because

$$SkD_\beta(q; S) = \frac{1}{2} \sum_{i=1}^n (|h(x_i)| - |B(x_i)| + |\rho(x_i)|),$$

where  $\rho(x_i)$  is the circular cap obtained from  $b(x_i)$  cut by  $h(x_i)$ .

- $n$  halfspace and  $n$  circular cap range counting approximations to approximate  $SkD_\beta(q; S)$ , for  $2 + \sqrt{2} \leq \beta < \infty$ , because

$$SkD_\beta(q; S) = \frac{1}{2} \sum_{i=1}^n (|h(x_i)| - |\rho(x_i)|).$$

## 7 Experimental Results

In this section some experimental results are provided to support Section 5 and Theorem 1. We compute the planar halfspace depth and planar  $\beta$ -skeleton depth of  $q \in Q$  with respect to  $S$  for different values of  $\beta$ , where  $Q$  and  $S$  are sets of randomly generated points (with double precision floating point coordinates) of sizes 1000 and 2500, respectively. The results of our experiments are summarized in Table 1 and its corresponding figures provided in Appendix. The columns 2 – 4 of Table 1 show that the halfspace depth can be approximated by a quadratic function of the  $\beta$ -skeleton depth with a relatively small value of  $d_E(SkD, HD) \cong 0.003$ . In particular,  $HD \cong 4.3(SkD_\beta)^2 - 2.8(SkD_\beta) + 0.47$  if  $\beta \rightarrow \infty$ . Columns 5 – 7 include the experimental results to support Theorem 1.

**Acknowledgment:** The authors would like to thank Joseph O'Rourke and other attendees of CCCG2017 for inspiring discussions regarding the relationships between depth functions and posets.

## References

- [1] Peyman Afshani and Timothy M Chan. On approximate range counting and depth. *Discrete & Computational Geometry*, 42(1):3–21, 2009.
- [2] Greg Aloupis. Geometric measures of data depth. *DIMACS series in discrete mathematics and theoretical computer science*, 72:147, 2006.
- [3] Bagchi, Amitabha and Chaudhary, Amitabh and Epstein, David and Goodrich, Michael T. Deterministic sampling and range counting in geometric data streams. *ACM Transactions on Algorithms (TALG)*, 3(2), 2007.
- [4] Boris Aronov and Sariel Har-Peled. On approximating the depth and related problems. *SIAM Journal on Computing*, 38(3):899–921, 2008.
- [5] Boris Aronov and Micha Sharir. Approximate half-space range counting. *SIAM Journal on Computing*, 39(7):2704–2725, 2010.
- [6] David Bremner, Dan Chen, John Iacono, Stefan Langerman, and Pat Morin. Output-sensitive algorithms for tukey depth and related problems. *Statistics and Computing*, 18(3):259–266, 2008.
- [7] David Bremner, Komei Fukuda, and Vera Rosta. Primal-dual algorithms for data depth. *DIMACS series in discrete mathematics and theoretical computer science*, 72:147, 2006.
- [8] Timothy M Chan. An optimal randomized algorithm for maximum tukey depth. In *Proceedings of the fifteenth annual ACM-SIAM symposium on Discrete algorithms*, pages 430–436. Society for Industrial and Applied Mathematics, 2004.
- [9] Dan Chen, Pat Morin, and Uli Wagner. Absolute approximation of tukey depth: Theory and experiments. *Computational Geometry*, 46(5):566–573, 2013.
- [10] David L Donoho and Miriam Gasko. Breakdown properties of location estimates based on halfspace depth and projected outlyingness. *The Annals of Statistics*, pages 1803–1827, 1992.
- [11] Ryan T Elmore, Thomas P Hettmansperger, and Fengjuan Xuan. Spherical data depth and a multivariate median. *DIMACS Series in Discrete Mathematics and Theoretical Computer Science*, 72:87, 2006.
- [12] Marco Fattore, Rosanna Grassi, and Alberto Arcagni. Measuring structural dissimilarity between finite partial orders. In *Multi-indicator Systems and Modelling in Partial Order*, pages 69–84. Springer, 2014.
- [13] Sariel Har-Peled. *Geometric approximation algorithms*. Number 173. American Mathematical Soc., 2011.
- [14] Sariel Har-Peled and Micha Sharir. Relative  $(p, \varepsilon)$ -approximations in geometry. *Discrete & Computational Geometry*, 45(3):462–496, 2011.
- [15] Harold Hotelling. Stability in competition. In *The Collected Economics Articles of Harold Hotelling*, pages 50–63. Springer, 1990.
- [16] David S. Johnson and Franco P Preparata. The densest hemisphere problem. *Theoretical Computer Science*, 6(1):93–107, 1978.
- [17] David G Kirkpatrick and John D Radke. A framework for computational morphology. In *Machine Intelligence and Pattern Recognition*, volume 2, pages 217–248. Elsevier, 1985.
- [18] Regina Y Liu. *Data depth: robust multivariate analysis, computational geometry, and applications*, volume 72. American Mathematical Soc., 2006.
- [19] Regina Y Liu et al. On a notion of data depth based on random simplices. *The Annals of Statistics*, 18(1):405–414, 1990.
- [20] Zhenyu Liu and Reza Modarres. Lens data depth and median. *Journal of Nonparametric Statistics*, 23(4):1063–1074, 2011.
- [21] Hannu Oja. Descriptive statistics for multivariate distributions. *Statistics & Probability Letters*, 1(6):327–332, 1983.
- [22] Peter J Rousseeuw and Mia Hubert. Regression depth. *Journal of the American Statistical Association*, 94(446):388–402, 1999.
- [23] Peter J Rousseeuw and Ida Ruts. Algorithm as 307: Bivariate location depth. *Journal of the Royal Statistical Society. Series C (Applied Statistics)*, 45(4):516–526, 1996.
- [24] Peter J Rousseeuw and Anja Struyf. Computing location depth and regression depth in higher dimensions. *Statistics and Computing*, 8(3):193–203, 1998.
- [25] Douglas S Shafer and Z Zhang. *Beginning statistics*. Phylis-Barnidge publisher, 304, 2012.
- [26] Rasoul Shahsavari and David Bremner. Computing the planar  $\beta$ -skeleton depth. *arXiv preprint arXiv:1803.05970*, 2018.
- [27] Hayim Shaul. *Range Searching: Emptiness, Reporting, and Approximate Counting*. University of Tel-Aviv, 2011.
- [28] Christopher G Small. A survey of multidimensional medians. *International Statistical Review/Revue Internationale de Statistique*, pages 263–277, 1990.
- [29] Anja J Struyf and Peter J Rousseeuw. Halfspace depth and regression depth characterize the empirical distribution. *Journal of Multivariate Analysis*, 69(1):135–153, 1999.
- [30] John W Tukey. Mathematics and the picturing of data. In *Proceedings of the international congress of mathematicians*, volume 2, pages 523–531, 1975.
- [31] Mengta Yang. *Depth Functions, Multidimensional Medians and Tests of Uniformity on Proximity Graphs*. PhD thesis, The George Washington University, 2014.
- [32] Mengta Yang and Reza Modarres.  $\beta$ -skeleton depth functions and medians. *Communications in Statistics-Theory and Methods*, pages 1–17, 2017.



## Appendix

**Lemma 3** For posets  $A, B, C$  in  $\mathbb{P} = \{P_t = (S, \preceq_t) | t \in \mathbb{N}\}$ ,  $d_c(A, B) \leq d_c(A, C) + d_c(C, B)$ , where

$$d_c(A, B) = \frac{\sum_{i=1}^n \sum_{j=1}^n |M_{ij}^A - M_{ij}^B|}{n^2 - n}$$

and

$$M_{ij}^k = \begin{cases} 1 & x_i \preceq_k x_j \\ 0 & \text{otherwise.} \end{cases}$$

**Proof.**

$$\begin{aligned} d_c(P_f, P_h) &= \frac{\sum_{i=1}^n \sum_{j=1}^n |M_{ij}^f - M_{ij}^h|}{n^2 - n} \\ &= \frac{\sum_{i=1}^n \sum_{j=1}^n |(M_{ij}^f - M_{ij}^g) + (M_{ij}^g - M_{ij}^h)|}{n^2 - n} \\ &\leq \frac{\sum_{i=1}^n \sum_{j=1}^n (|M_{ij}^f - M_{ij}^g| + |M_{ij}^g - M_{ij}^h|)}{n^2 - n} \\ &= \frac{\sum_{i=1}^n \sum_{j=1}^n |M_{ij}^f - M_{ij}^g|}{n^2 - n} + \frac{\sum_{i=1}^n \sum_{j=1}^n |M_{ij}^g - M_{ij}^h|}{n^2 - n} \\ &= d_c(P_f, P_g) + d_c(P_g, P_h). \end{aligned}$$

□

**Lemma 4** For  $\beta' > \beta \geq 1$  and  $a, b \in \mathbb{R}^2$ ,  $S_\beta(a, b) \subseteq S_{\beta'}(a, b)$ , where  $S_\beta(a, b)$  is the intersection of two disks  $B(C_{ab\beta}, R_{ab\beta})$  and  $B(C_{ba\beta}, R_{ba\beta})$ ,  $C_{ab\beta} = (\beta/2)(a - b) + b$ , and  $R_{ab\beta} = (\beta/2)d(a, b)$ .

**Proof.** To prove that  $B(C_{ab\beta}, R_{ab\beta}) \cap B(C_{ba\beta}, R_{ba\beta})$  is a subset of  $B(C_{ab\beta'}, R_{ab\beta'}) \cap B(C_{ba\beta'}, R_{ba\beta'})$ , it is enough to prove  $B(C_{ab\beta}, R_{ab\beta}) \subseteq B(C_{ab\beta'}, R_{ab\beta'})$  and  $B(C_{ba\beta}, R_{ba\beta}) \subseteq B(C_{ba\beta'}, R_{ba\beta'})$ . We only prove the first one, and the second one can be proved similarly. Suppose that  $\beta < \beta' = \beta + \varepsilon; \varepsilon > 0$ . It is trivial to check that two disks  $B(C_{ab\beta}, R_{ab\beta})$  and  $B(C_{ab\beta'}, R_{ab\beta'})$  meet at  $b$ . See Figure 6. Let  $t \neq b$  be an extreme point of  $B(C_{ab\beta}, R_{ab\beta})$ . This means that

$$\begin{aligned} d(t, C_{ab\beta}) &= R_{ab\beta} \Leftrightarrow d\left(t, \frac{\beta(a-b)}{2} + b\right) = \frac{\beta d(a, b)}{2} \\ &\Leftrightarrow \left| \frac{\beta(a-b)}{2} + (b-t) \right|^2 = \left( \frac{\beta(a-b)}{2} \right)^2 \\ &\Leftrightarrow |b-t|^2 - \beta(b-a) \cdot (b-t) = 0 \end{aligned}$$

This means  $(b-a) \cdot (b-t) \geq 0$ . Hence,

$$\begin{aligned} |b-t|^2 - \beta(b-a) \cdot (b-t) &= 0 \Leftrightarrow \\ |b-t|^2 - (\beta + \varepsilon)(b-a) \cdot (b-t) &< 0 \Leftrightarrow \\ |b-t|^2 - \beta'(b-a) \cdot (b-t) &< 0 \Leftrightarrow \\ d(t, C_{ab\beta'}) - R_{ab\beta'} &< 0 \end{aligned}$$

The last inequality means that  $t$  is an interior point of  $B(C_{ab\beta'}, R_{ab\beta'})$ . □

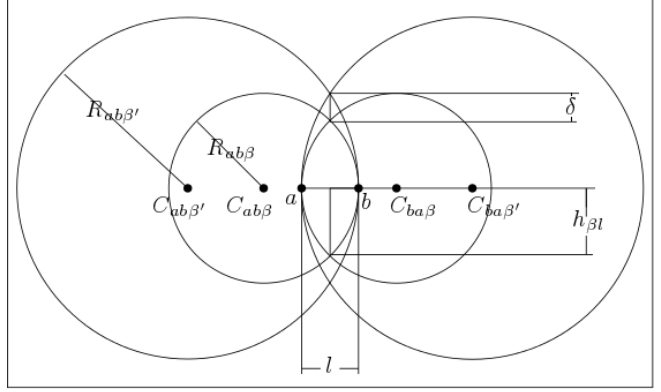


Figure 6:  $S_\beta(a, b)$  and  $S_{\beta'}(a, b)$ .

**Lemma 5** For  $a, b \in \mathbb{R}^2$ ,

$$\lim_{\beta \rightarrow \infty} \delta = \lim_{\beta \rightarrow \infty} (h_{(\beta+1)l} - h_{\beta l}) = 0, \quad (13)$$

where  $h_{\beta l} = (l/2)\sqrt{(2\beta-1)}$ . See Figure 6.

**Proof.** Instead of proving (13), we prove its equivalent form as follows:

$$\lim_{\beta \rightarrow \infty} \frac{h_{(\beta+1)l}}{h_{\beta l}} = \lim_{\beta \rightarrow \infty} \frac{(l/2)\sqrt{(2\beta+1)}}{(l/2)\sqrt{(2\beta-1)}} = \lim_{\beta \rightarrow \infty} \frac{\sqrt{(2\beta+1)}}{\sqrt{(2\beta-1)}} = 1.$$

□

**Lemma 6** For  $k = \beta/(2(\beta-1))$  and  $a, b \in \mathbb{R}^2$  ( $a$  is fixed and  $b$  is arbitrary), halfspace  $\vec{a} \cdot \vec{b} \leq (1/2k)\|a\|^2$  passes through the center of disk  $(b - ka)^2 \geq (k\|a\|)^2$  if  $\beta = 2 + \sqrt{2}$ .

**Proof.** It is enough to substitute  $b$  in the given halfspace with  $ka$  which is the center of the given disk.

$$\begin{aligned} \vec{a} \cdot (ka) &\leq \frac{1}{2k}\|a\|^2 \Rightarrow k\|a\|^2 \leq \frac{1}{2k}\|a\|^2 \Rightarrow 2k^2 \leq 1 \Rightarrow \\ 2\left(\frac{\beta}{2(\beta-1)}\right)^2 &\leq 1 \Rightarrow -\beta^2 + 4\beta - 2 \leq 0 \Rightarrow \beta = 2 \pm \sqrt{2} \end{aligned}$$

Since  $\beta \geq 1$ , the  $\beta = 2 + \sqrt{2}$  is valid. □



Table 1: Summary of experimental results

$x$	$HD = f(x)$	$d_E(x, HD)$	Figure	$f(x, SkD_\infty)$	$d_E(x, SkD_\infty)$	Figure
$SkD_1$	$3.103x^2 - 0.013x + 0.013$	0.0016	7	$0.942x + 0.288$	0.007	12
$SkD_2$	$2.71x^2 - 0.48x + 0.04$	0.0019	8	$0.858x + 0.23$	0.002	13
$SkD_3$	$2.92x^2 - 0.82x + 0.08$	0.0021	9	$0.86x + 0.192$	0.001	14
$SkD_{1000}$	$4.26x^2 - 2.76x + 0.47$	0.0031	10	$0.999x + 0.002$	0.0	15
$SkD_{10000}$	$4.27x^2 - 2.78x + 0.47$	0.0031	11	$1.0x + 0.0$	0.0	16

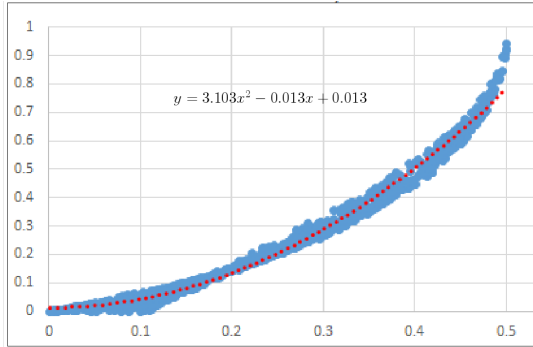


Figure 7:  $HD$  versus  $SkD_1$ .

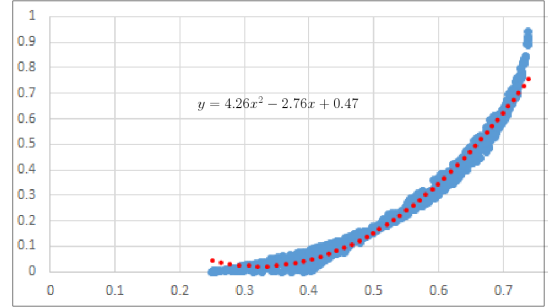


Figure 10:  $HD$  versus  $SkD_{1000}$ .

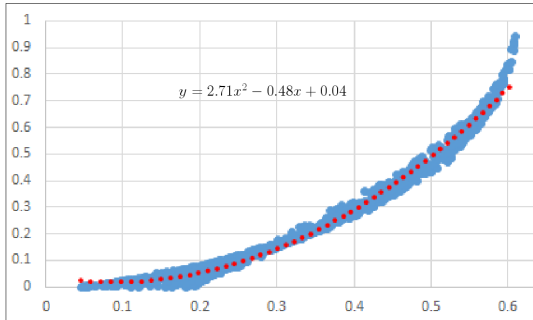


Figure 8:  $HD$  versus  $SkD_2$ .

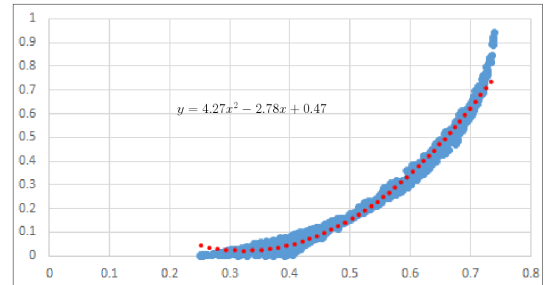


Figure 11:  $HD$  versus  $SkD_{10000}$ .

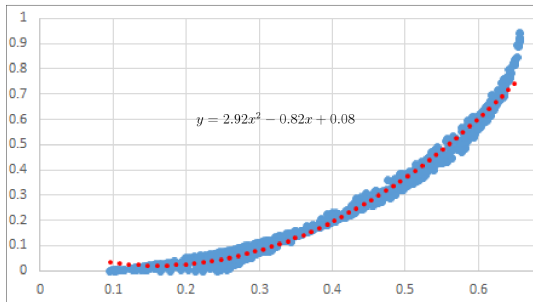


Figure 9:  $HD$  versus  $SkD_3$ .

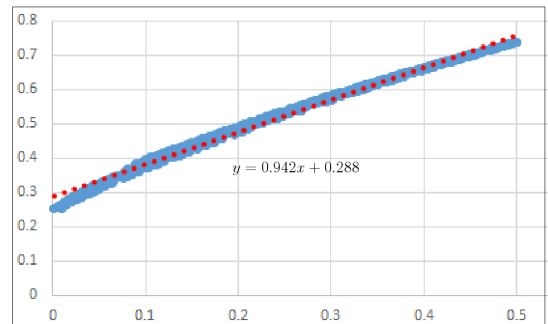


Figure 12:  $SkD_\infty$  versus  $SkD_1$ .

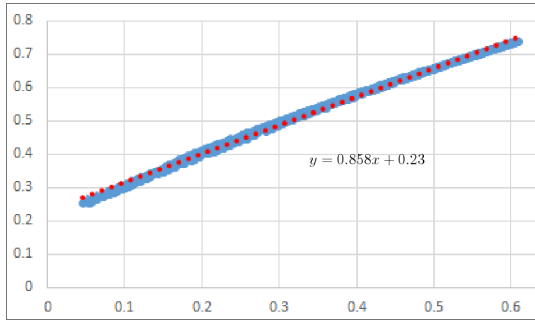


Figure 13:  $SkD_\infty$  versus  $SkD_2$ .

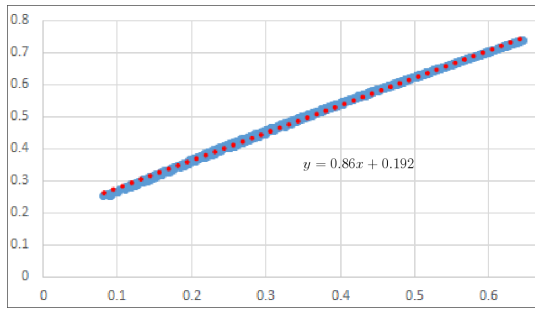


Figure 14:  $SkD_\infty$  versus  $SkD_3$ .

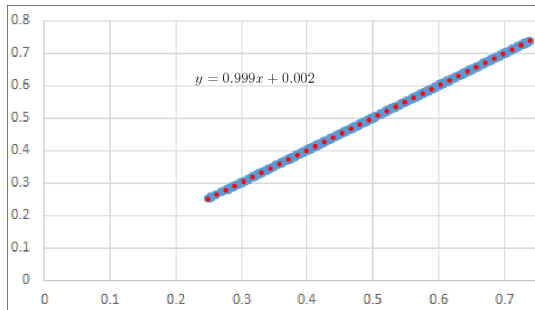


Figure 15:  $SkD_\infty$  versus  $SkD_{1000}$ .

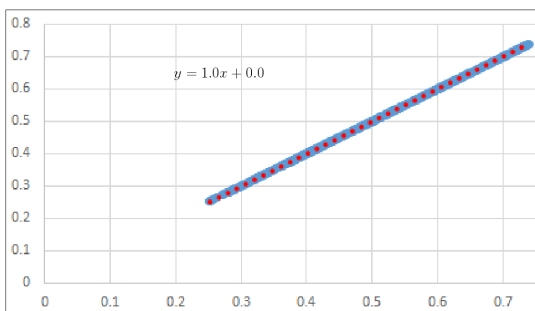


Figure 16:  $SkD_\infty$  versus  $SkD_{10000}$ .