The CSI Journal on
Computer Science and Engineering
Vol. 17, No. 1, 2019
Pages 49-54
Regular Paper

On The Spread of a Uniformly Distributed Point Set

Behnam Iranfar Mohammad Farshi Amir Mesrikhani

Combinatorial and Geometric Algorithms Lab, Department of Mathematical Sciences,
Yazd University, Yazd, Iran.

Abstract

Let P be a set of n points in $\mathbb{R}^d$, where d is a constant. The spread (or aspect ratio) of P is the ratio of the distance between the farthest pair of points in P to the distance between the closest pair of points in P. This ratio is an essential factor that appears in many geometric problems. In this paper, we show that the expected value of the spread of a set of n uniformly distributed point sets is $\Theta(n^{2/d})$
Keywords: spread, uniform distribution, expected value.

1. Introduction

The spread of the input point set P, which is the ratio of the diameter and the closest pair distance and denoted by $\varphi(P)$, has appeared in the size of many data structures and time and space complexity of some algorithms in computational geometry such as several t-spanner algorithms, Voronoi diagrams algorithms, the Well-Separated Pair Decomposition (WSPD), and the quadtree algorithm. In several applications, the input data has some kind of distribution in practice. So, it is crucial to analyze $\varphi(P)$ for a uniformly distributed point set P to find a reasonable bound for it.

In this paper, we study $ClosestPair(P_n)$, i.e. the distance between the closest pair in P_n, and $Diameter(P_n)$, i.e. the distance between the farthest pair in P_n, for any set of uniformly distribution points set P_n. We show that the expected value of $ClosestPair(P_n)$ is $\Theta(n^{-2/d})$, and the expected value of the spread is $\Theta(n^{2/d})$.

1.1. Related work

A t-spanner of a set P of points is a graph with vertex set P, such that for each pair $p, q \in P$, there is some path in the graph such that its length is at most t times the Euclidean distance between p and q. In the following, we review the time and space complexity of a well-known t-spanner algorithm that depends on the spread of the input point set. In 2006, Gao *et al.* [1] propose a $(1+\varepsilon)$-spanner with maximum degree bounded by $O(\log\varphi(P)/\varepsilon^d)$, and total weight bounded by $O(\log\varphi(P)/\varepsilon^{d+1} \times MST)$, where MST is the weight of the minimum spanning tree of P and $\varepsilon > 0$ is a real number. The size of the proposed algorithm was in $O(n)$ and insertion and deletion of each point in the spanner take $O(\log\varphi(P)/\varepsilon^d)$ time. The spanner can be constructed in $O(n\log\alpha/\varepsilon^d)$ time. They also proposed a kinetic spanner algorithm which is efficient, responsive, local, and compact, when the aspect ratio of the points has a reasonable bound. Specifically, the total number of events in maintaining G is $O(n^2\log\varphi(P))$ under pseudo-algebraic motion. Each event takes $O(\log\varphi(P)/\varepsilon^d)$ time, and each flight plan change uses $O(\log\varphi(P)/\varepsilon^d)$ time. Each point is involved in at most $O(\log\varphi(P)/\varepsilon^d)$ certificates.

In computational geometry, Voronoi diagrams, and their duals (Delaunay triangulation) have appeared in many applications [2]. There is an output-sensitive algorithm for computing the Voronoi diagram of a set of n points in $\mathbb{R}^d$, with $O(k\log n\log\varphi(P))$ time complexity, where k is the complexity of the output [3]. Erickson [4] proposed an algorithm for Delaunay triangulation of a set P of n points in $\mathbb{R}^3$ with $O((\varphi(P))^3)$ time complexity.

The doubling parameter $\lambda \in \mathbb{N}$ is the smallest integer such that for every $r \geq 0$ and every $p \in P$, the ball with center p and radius r, can be covered by at most λ balls of radius $r/2$. Talwar [5] gave an algorithm to compute a well-separated pair decomposition (WSPD) with separation parameter $s \geq 1$ of size $O(s^{\log\lambda} n \log\varphi(R))$, where $\varphi(P)$ is the spread of P.

Let M be a set of points. A set $A \subseteq M$ is a dense subset of M if for any point $v \in M$ there exists a point in A sufficiently close to v, that means, any open set that contains v, must contain a point from A too. Erickson [4] studied the spread of a dense subset of $\mathbb{R}^d$. He proved that the spread is $O(n^{1/d})$.

The quadtree of $P_n \subset \mathbb{R}^d$ is a rooted tree such that every internal node has 2^d children. Each node in a quadtree corresponds to a d-dimensional hypercube. To construct a quadtree, it starts with the bounding box of the input points. For each node, the children's hypercube is produced by dividing the parent's hypercube into 2^d equal hypercubes, and partition point set in the node into these hypercubes. Recursively construct quadtrees for each hypercube with its corresponding points set. The recursion stops when all hypercubes contain at most one point. The time complexity of this algorithm depends on the height of the tree. Let $ClosestPair(P_n)$ be the distance between the closest pair in P_n and L be the length of the initial hypercube. When we descend from one node to one of its children, the length of the hypercubes halves. Therefore, in level i, the length is $\frac{L}{2^i}$, and the maximum distance between points of the corresponding points set is $\frac{\sqrt{d}L}{2^i}$. The dividing process continues until the closest pair of points is detected. So, the height of the tree is $O(\log\frac{\sqrt{d}L}{ClosestPair(P_n)})$. Obviously, $\sqrt{d}L$ is the diameter of the initial hypercube, so the height of the tree is $O(\log\varphi(P_n))$.

One way to measure the uniformity of a set of points in the bounding box is the gap ratio. The gap is the ratio between the radius of the largest empty circle (do not have any point in it) to the distance between the closest pair of points. Teramoto *et al*. [6] introduced the gap ratio and showed that by inserting k points, the upper bound of the gap ratio is $2^{\lceil k/2 \rceil / \lceil k/2 \rceil + 1}$ in one-dimensional space. Using a Voronoi diagram, they proposed the upper bound 2 for two-dimensional space. If the coordinates of the points are integer numbers, then inserting k points in $[0, n]$ yield two as the upper bound of the gap ratio[7,8].

Agarwal *et al*. [9] studied the diameter of a set of points in $\mathbb{R}^d$. They proposed a $(1-\varepsilon)$-approximation algorithm with $O((\log n + 1/\varepsilon^{d-1})/\varepsilon^{(d-1)/2})$ amortized time and $O(\log^d(n)/\varepsilon^{(d-1)/2})$ space complexity.
Also, Indyk *et al*. [10] studied geometric matching under noise. They proposed an algorithm with time complexity depends on n and the spread.

2. Problem definition and notations

Let $P_n = \{p_1, \ldots, p_n\}$ be a set of $n \geq 2$ points in $\mathbb{R}^d$, where d is a constant. The spread of P_n is $\varphi(P_n) = \frac{Diameter(P_n)}{ClosestPair(P_n)}$, where $Diameter(P_n)$ and $ClosestPair(P_n)$ are the farthest and closest distance of the pairs of points in P_n, respectively.

In the following, let $P_n = \{p_1, p_2, \ldots, p_n\}$ be a set of $n \geq 2$ points uniformly distributed in a d-dimension unit hypercube. The volume of a d-ball (V_d) with radius r is $V_d(R) = \gamma_d \times r^d$, where γ_d is a constant.

Uniform distribution: Let R be a set in $\mathbb{R}^d$ and $vol(R)$ denote the volume of R that is finite. The density function

$$\begin{cases} f(x_1, \ldots, x_d) = \frac{1}{vol(R)} & if\ (x_1, \ldots, x_d) \in R \\ 0 & otherwise \end{cases}$$

is the uniform distribution on R. For any set $A \subseteq R$, we have
$P(A) = \int \ldots \int_A f(x_1, \ldots, x_d) dx_1 \ldots dx_d =$
$\int \ldots \int_A \frac{1}{vol(R)} dx_1 \ldots dx_d = \frac{vol(A)}{vol(R)}$

3. The expected spread

In this section, we analyze the expected value of $\varphi(P_n)$, where P_n is a set of n uniformly distributed points in a unit d-dimensional hypercube. We use induction to calculate the lower and upper bound of the expected value of the closest pair. So, we have to know the expected value of the distance between two points that lies uniformly in the unit hypercube for the base case. The following lemma computes this value
.

Lemma 3.1 *If x_1 and x_2 are two uniformly distributed points in the d-dimensional unite hypercube, then the expected value of the distance between x_1 and x_2 is $\sqrt{d}/3$.*

Proof. Since each coordinate of points is a random number between 0 and 1, we first consider the problem in the one-dimensional case. Assume Y is the distance between two uniformly distributed points between 0 and 1. The joint probability distribution of x_1 and x_2 is
$f_{x_1,x_2}(x_1, x_2) = 1 \quad where\ 0 \leq x_1 \leq 1, 0 \leq x_2 \leq 1.$
So, the density function of Y is

$$F_Y(y) = p(Y \leq y) = p(x_1 - x_2 \leq y)$$
$$= \begin{cases} \int_0^{1+y} \int_{x_1-y}^{1} dx_2 dx_1 & if\ -1 \leq y < 0 \\ 1 - \int_y^1 \int_0^{x_1-y} dx_2 dx_1 & if\ 0 \leq y \leq 1 \end{cases}$$
$$= \begin{cases} y^2/2 + y + 1/2 & if\ -1 \leq y < 0 \\ -y^2/2 + y + 1/2 & if\ 0 \leq y \leq 1 \end{cases}.$$

Therefore, the distribution function is $f_Y(y) = \begin{cases} y + 1 & if\ -1 \leq y < 0 \\ 1 - y & if\ 0 \leq y \leq 1 \end{cases}.$

If $Z = |Y|$, then by the cumulative distribution function [11], we have :

$$F_Z(z) = p(Z \leq z)$$
$$= p(|Y| \leq z)$$
$$= \int_0^z (1-y)dy + \int_{-z}^0 (y+1)dy$$
$$= 2z - z^2.$$

So, the distribution function of Z is

$$f_Z(z) = 2 - 2z (0 \le z \le 1).$$

The expected value of the distance between two points in $d = 1$ is

$$Exp(z) = \int_0^1 z(2-2z)dz = 1/3,$$

and thus ,

$$Exp(\sqrt{Z_1^2 + Z_2^2 + \cdots + Z_d^2}) = Exp(\sqrt{dZ^2}) = \sqrt{d}Exp(Z) = \sqrt{d}/3.$$

We compute the expected value of $Diameter(P_n)$ in the following lemma .

Lemma 3.2 *The expected value of the diameter of* P_n *belongs to* $\Theta(\sqrt{d})$ *.*

Proof. The diameter of a unit d-hypercube is $\sqrt{d}$. So, $Exp(Diameter(P_n)) = O(\sqrt{d})$. If we have only two points, then by Lemma 3.1 the expected value of the distance between two points is $\sqrt{d}/3$. Therefore, for $n = 2$ we have $Exp(Diameter(P_n)) = \Omega(\sqrt{d}/3)$, so, we have $Exp(Diameter(P_n)) = \Omega(\sqrt{d})$. The diameter of the unit d-dimensional hypercube is $\sqrt{d}$, therefore, $Exp(Diameter(P_n)) = \Theta(\sqrt{d})$.

Let x be a random point in $[-\alpha, \alpha]$ ($\alpha \in \mathbb{R}^+$). In the following lemma, we compute the probability density function of the distance between x and the center of the d-ball with constant radius α .

Lemma 3.3 *Let* x *be a random point in a* d*-ball with radius* α*. Then the distribution function of the distance between* x *and the center of the* d*-ball is*

$$f_R(r) = \frac{dr^{d-1}}{\alpha^d}.$$

Proof. Let $x_1, \ldots, x_d$ be a set of d independent random variables with uniform distribution in $[-\alpha, \alpha]$ ($\alpha \in \mathbb{R}, \alpha > 0$). Suppose $S = (x_1^2 + \cdots + x_d^2)^{d/2}$. By ignoring $x_i (1 \le i \le d)$ where $S = 0$ or $S \ge \alpha^d$, then S has a uniform distribution on $(0, \alpha^d)$. Since $x_1, \ldots, x_d$ have a uniform distribution, and valid points are placed in a d-ball with the radius α, so the value of S is in $(0, \alpha^d)$. By the cumulative distribution and for $s \in (0, \alpha^d)$, we have

$$F_S(s) = p(S \le s) = \frac{\gamma_d s}{\gamma_d \alpha^d} = \frac{s}{\alpha^d}.$$

So, the probability distribution function is $f_S(s) = \frac{1}{\alpha^d}$. Therefore ,

$$F_R(r) = p(R \le r) = p(\sqrt[d]{S} \le r) = p(S \le r^d) = \int_0^{r^d} \frac{1}{\alpha^d} ds = \frac{r^d}{\alpha^d}$$

So, we can conclude that $f_R(r) = \frac{dr^{d-1}}{\alpha^d}$.

Let P_k be a set of k points. The *critical* region of P_k is the union of k d-balls, each d-ball centered at a point in P_k and with radius $ClosestPair(P_k)$ (see Figure 1). If we add a point at the critical region of P_k, then the closest pair distance in the new set decreases. The following lemma presents a lower bound and an upper bound for the volume of the critical region.

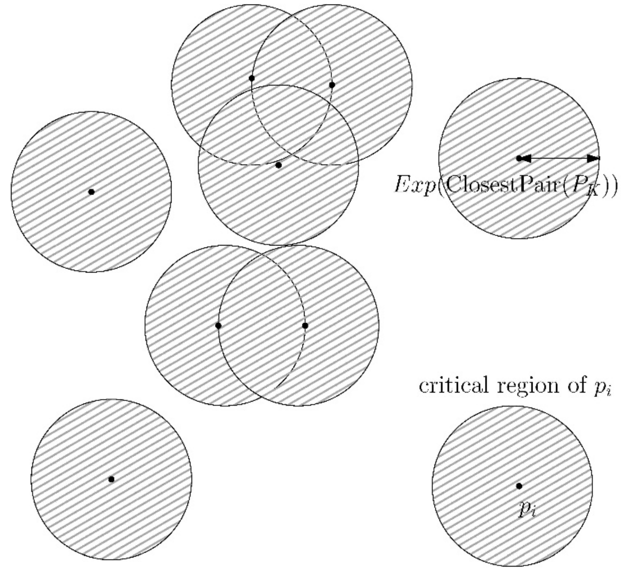

Figure 1: Critical regions of a set of points.

Lemma 3.4 *The expected volume of the critical region of a set of* k *uniformly distributed points in* $[0,1]^d$ *is between* $k\gamma_d(Exp(ClosestPair(P_k)))^d/2^d$ *and* $k\gamma_d(Exp(ClosestPair(P_k)))^d$ *.*

Proof. By the definition of the critical region (see Figure 1), we know that the maximum volume of the critical region has occurred when d-balls do not have any overlap. So, the maximum expected volume is

$$k \times \gamma_d(Exp(ClosestPair(P_k)))^d = k\gamma_d(Exp(ClosestPair(P_k)))^d.$$

The expected value of the distance between any two points is $Exp(ClosestPair(P_k))$. If we consider the radius of d-balls to be $\frac{Exp(ClosestPair(P_k))}{2}$, then none of them overlap. So, the minimum of the expected volume of the critical regions is $k \times \gamma_d(\frac{Exp(ClosestPair(P_k))}{2})^d = k\gamma_d 2^{-d}(Exp(ClosestPair(P_k)))^d$.

The following lemma gives an upper bound for $ClosestPair(P_n)$.

Lemma 3.5 *The expected value of* $ClosestPair(P_n)$ *is* $O(n^{-2/d})$ *.*

Proof. For proving the theorem, we show that there exists a constant c in $\mathbb{R}$ such that, for each n, we have

$$Exp(ClosestPair(P_n)) \le cn^{-2/d}.$$

We will prove it by induction on n.

Base case: By Lemma 3.1 for $n = 2$, the expected value of the distance between the two points is $\sqrt{d}/3$. If $c \ge \frac{2^{2/d}d}{3}$, then $\sqrt{d}/3 \le c \times 2^{-2/d}$.

Induction step: Suppose that for $n = k$ we have $Exp(ClosestPair(P_k)) \le ck^{-2/d}$.

For $n = k + 1$, we must show

that $Exp(ClosestPair(P_{k+1})) \le c(k+1)^{-2/d}$. Suppose we want to add $(k+1)$th point to P_k. Assume $(k+1)$th point is placed in the critical region. Therefore, the maximum value of $Exp(ClosestPair(P_{k+1}))$ is occurring when $(k+1)$th point is placed in a d-ball of the critical region that does not overlap with other d-balls in the critical region. So, by Lemma 3.3, the $Exp(ClosestPair(P_{k+1}))$ in this case is at most

$$\int_0^{ck^{-2/d}} \frac{dr^{d-1}}{(ck^{-2/d})^d} \times rdr = (c^{-1}k^{2/d})^d \int_0^{ck^{-2/d}} dr^d dr = \frac{d}{d+1} ck^{-2/d}.$$

If the volume of the critical region is minimized, then $ClosestPair(P_k)$ is maximized. By Lemma 3.4, the minimum volume of the critical region of P_k is

$$k \times \gamma_d (\frac{Exp(ClosestPair(P_k))}{2})^d = k \times \gamma_d (\frac{ck^{-2/d}}{2})^d = \gamma_d c^d 2^{-d} k^{-1}.$$

Obviously, $Exp(ClosestPair(P_{k+1}))$ is equal to the multiplication of the probability that $(k+1)$th point lies in the critical region of P_k and the $Exp(ClosestPair(P_{k+1}))$ in the critical region plus multiplication of the probability that $(k+1)$th point lies outside the critical region of P_k and the $Exp(ClosestPair(P_k))$. So, the $Exp(ClosestPair(P_{k+1}))$ is at most

$$\frac{d}{d+1} ck^{-2/d} \times \gamma_d c^d 2^{-2d} k^{-1} + ck^{-2/d}(1 - \gamma_d c^d 2^{-d} k^{-1}).$$

Therefore, we must show that

$$\frac{d}{d+1} ck^{-2/d} . \gamma_d c^d 2^{-d} k^{-1} + ck^{-2/d}(1 - \gamma_d c^d 2^{-d} k^{-1}) \le c(k+1)^{-2/d}$$

By simplifying the inequality, we have

$$\frac{d}{d+1} \gamma_d c^d 2^{-d} k^{\frac{-2}{d}-1} + k^{-2/d} - \gamma_d c^d 2^{-d} k^{\frac{-2}{d}-1} \le (k+1)^{-2/d}$$

which is equivalent to

$$\frac{1}{k^{2/d}} - \frac{\gamma_d c^d 2^{-d}}{(d+1)k^{2/d+1}} \le \frac{1}{(k+1)^{2/d}}.$$

Since d, γ_d, and c are constant, for simplicity we define $c' = \frac{\gamma_d c^d 2^{-d}}{d+1}$. Therefore, we have

$$\frac{k-c'}{k^{2/d+1}} \le \frac{1}{(k+1)^{2/d}} \Rightarrow c' \ge k - k(\frac{k}{k+1})^{2/d}.$$

Since $d \ge 2$ and $k \ge 2$, we have $0 \le 2/d \le 1$ and $0 < \frac{k}{k+1} < 1$. So,

$$c' \ge k - k(\frac{k}{k+1}) \ge k - k(\frac{k}{k+1})^{2/d}$$

$$\Rightarrow c' \ge k - \frac{k^2}{k+1} \Rightarrow c' \ge \frac{k}{k+1} \Rightarrow c' \ge 1 \ge \frac{k}{k+1}$$

It is enough to set $c \ge \max\left\{\left(2\sqrt{\frac{d+1}{\gamma_d}}\right), \frac{2^{2/d}d}{3}\right\}$, and it completes the proof.

Similarly, we can obtain a lower bound for $Exp(ClosestPair(P_n))$ by the following lemma.

Lemma 3.6 *The expected value of* $ClosestPair(P_n)$ *is* $\Omega(n^{-2/d})$.

Proof. We show that there exists a constant c in $\mathbb{R}$, such that,

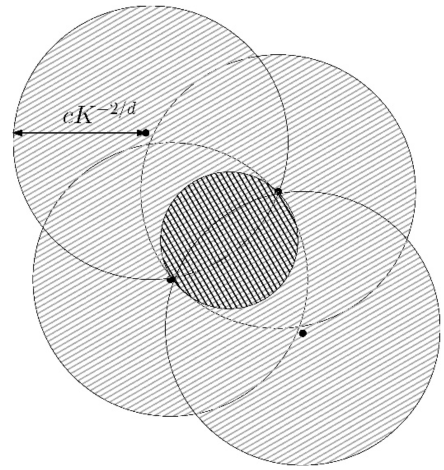

Figure 2: Illustration of the proof of Lemma 3.6.

for each n, $Exp(ClosestPair(P_n)) \ge cn^{-2/d}$.
We prove this by induction on n.

Base case: For $n = 2$, by Lemma 3.1, the expected value of the distance between the two points is $\sqrt{d}/3$. So, we have $\sqrt{d/6} \ge c.2^{-2/d}$. Therefore, by choosing $c \le 2^{2/d}\sqrt{d}/3$, the base case is correct.

Induction step: Suppose lemma is true for $n = k$, i.e. $Exp(closesetpair(P_k)) \ge ck^{-2/d}$. We must show that for $n = k+1$, we have $Exp(ClosestPair(P_{k+1})) \ge c(k+1)^{-2/d}$.

Suppose we have k points, and we want to add $(k+1)$th point. Let this point be in a critical region. We want to compute $Exp(ClosestPair(P_{k+1}))$. The minimum value of $Exp(ClosestPair(P_{k+1}))$ occurred if $(k+1)$th point is in a critical region that d-ball of the points overlap. Since the expected value of the distance between any two points is at least $ck^{-2/d}$, so if we consider a d-ball with radius $\frac{ck^{-2/d}}{2}$, then there are no points in it. So, in this case, the $Exp(ClosestPair(P_{k+1}))$ is smaller than the distance from $(k+1)$th point to the boundary of the d-ball (see Figure 2). So, if $(k+1)$th point is in a critical region, by Lemma 3.1 and Lemma 3.3,

$$Exp(ClosestPair(P_{k+1})) \ge \int_0^{\frac{ck^{-2/d}}{2}} \left(\frac{d \times r^{d-1}}{\left(\frac{ck^{-2/d}}{2}\right)^d}\right) \times \left(\frac{ck^{-2/d}}{2} - r\right) dr$$

$$= \left(\frac{2}{ck^{-2/d}}\right)^d \int_0^{\frac{ck^{-2/d}}{2}} \left(d \times r^{d-1}\left(\frac{ck^{-2/d}}{2} - r\right)\right) dr$$

$$= \left(\frac{2}{ck^{-2/d}}\right)^d \left(\frac{ck^{-2/d}}{2} \int_0^{\frac{ck^{-2/d}}{2}} d \times r^{d-1} dr - \int_0^{\frac{ck^{-2/d}}{2}} d \times r^d dr\right)$$

$$= \left(\frac{2}{ck^{-2/d}}\right)^d \frac{ck^{-2/d}}{2} \left(\frac{ck^{-2/d}}{2}\right)^d - \left(\frac{2}{ck^{-2/d}}\right)^d \times \frac{d}{d+1} \times \left(\frac{ck^{-2/d}}{2}\right)^{d+1}$$

$$= \frac{ck^{-2/d}}{2} - \frac{d}{d+1} \times \left(\frac{ck^{-2/d}}{2}\right)$$

$= \frac{ck^{-2/d}}{2(d+1)}.$

The minimum value of $Exp(ClosestPair(P_{k+1}))$ occurs when the volume of the critical region is maximized. The maximum volume of the critical region is
$k \times \gamma_d (Exp(ClosestPair(P_k)))^d = k \times \gamma_d (ck^{-2/d})^d = \gamma_d c^d k^{-1}.$

So, the $Exp(ClosestPair(P_{k+1}))$ is at least

$$\frac{ck^{-2/d}}{2(d+1)} \gamma_d c^d k^{-1} + ck^{-2/d}(1 - \gamma_d c^d k^{-1}).$$

Therefore, to prove the lemma, we must show that this value is greater or equal to $c(k+1)^{-2/d}$. We have
$$\frac{ck^{-2/d}}{2(d+1)} \gamma_d c^d k^{-1} + ck^{-2/d}(1 - \gamma_d c^d k^{-1}) \ge c(k+1)^{-2/d}.$$

By simplifying the inequality, we have
$$\frac{\gamma_d c^d k^{\frac{-2}{d}-1}}{2(d+1)} + k^{-2/d} - \gamma_d c^d k^{\frac{-2}{d}-1} \ge (k+1)^{-2/d},$$

which is equivalent to
$$\frac{1}{k^{2/d}} - \frac{(2d+1)\gamma_d c^d}{2(d+1)k^{\frac{2}{d}+1}} \ge \frac{1}{(k+1)^{2/d}}.$$
Since d, γ_d and c are constant, for simplicity we define $c'' = \frac{(2d+1)\gamma_d c^d}{2(d+1)}$, and we have
$$\frac{k - c''}{k^{\frac{2}{d}+1}} \ge \frac{1}{(k+1)^{2/d}} \Rightarrow c'' \le k(1 - (\frac{k}{k+1})^{2/d}).$$

We consider two cases, $d = 1$ and $d \ge 2$. For $d = 1$, we have
$$c'' \le k(1 - (\frac{k}{k+1})^2) \Rightarrow c'' \le \frac{2k^2 + k}{k^2 + 2k + 1}.$$

Since $k \ge 2$, it is enough to choose $c'' \le 1$. For $d \ge 2$, let $L = k + 1$, so, we have
$$c'' \le (L-1)(1 - (\tfrac{L-1}{L})^{2/d}) = (L-1)(1 - (1 - \tfrac{1}{L})^{2/d}).$$

Since $d \ge 2$ and $0 \le 2/d \le 1$, by Newton's generalized binomial theorem, we have $(1 - L)^{2/d} \ge 1 - \frac{2}{d} \times \frac{1}{L}$. So, we have $c'' \le (L-1)\left(1 - \left(1 - \frac{2}{d \times L}\right)\right) \Rightarrow c'' \le \frac{L-1}{d \times L}$ $L \ge 3$, since $k \ge 2$. It is clear that $c'' \le \frac{2}{3d}$. Since $1 \ge \frac{2}{3d}$, therefore for $d \ge 1$ it is enough to set $c'' \le \frac{2}{3d}$. So we have $\frac{(2d+1)\gamma_d c^d}{2(d+1)} \le \frac{2}{3d} \Rightarrow c \le \left(\frac{4(d+1)}{3d(2d+1)\gamma_d}\right)^{1/d}$.

So, it is enough to set $c \le \min\{(\frac{4(d+1)}{3d(2d+1)\gamma_d})^{1/d}, 2^{\frac{2}{d}}\sqrt{d}/3\}$ and the proof is complete .

Lemma 3.7 *Let P_n be a set of $n \ge 4$ uniformly distributed points in the unit d-dimensional hypercube. Then the expected value of $\varphi(P_n)$ is $\Theta(n^{2/d})$.*

Proof. By Lemma 3.2, we know that the expected value of the distance between any pair of points is $\Theta(\sqrt{d})$. Also, $Exp(Diameter(P_n)) = \Theta(1)$, since d is a constant. By Lemma 3.5 and Lemma 3.6, we have $Exp(ClosestPair(P_n)) = \Theta(n^{-2/d})$.

To decline the dependency between $Exp(Diameter(P_n))$ and $Exp(ClosestPair(P_n))$, we remove the closest pair points and compute $Exp(Diameter(P_n))$ again. The upper bound of $Exp(Diameter(P_n))$ is still $O(1)$, and the lower bound is $\Theta(1) - \Theta(n^{-2/d}) = \Theta(1)$. So, we have $Exp(\varphi(P_n)) = \frac{\Theta(1)}{\Theta(n^{-2/d})} = \Theta(n^{2/d})$.

Now, let P_n be a set of $n \ge 2$ uniformly distributed points in a d-dimensional hypercube with side length L. We have $Exp(ClosestPair(P_n)) = L\Theta(n^{-2/d})$ and $Exp(Diameter(P_n)) = L\Theta(\sqrt{d})$. So, we have the following lemma .

Lemma 3.8 *Let P_n be a set of $n \ge 4$ uniformly distributed points in a d-dimensional hypercube with side length L. Then, the expected value of $\varphi(P_n)$ is $\Theta(n^{2/d})$.*

In geometry, the inscribed hypercube of shape S is the largest hypercube that is inside S, and we call it *inhypercube*. The circumscribed hypercube or *circuhypercube* of shape S is the smallest hypercube that covers S entirely.

Theorem 3.9 *If M_n is a set of $n \ge 4$ uniformly distributed points in S and $S \in \mathbb{R}^d$ has a finite volume, then $Exp(\varphi(M_n))$ is $\Theta(n^{2/d})$.*

Proof. Let L_1 be the length of inhypercube of S and P_n be a set of $n \ge 4$ uniformly distributed points in a d-dimensional hypercube with side L_1. Then the $Exp(ClosestPair(P_n))$ is $L_1\Theta(2^{-2/d})$ and $Exp(Diameter(P_n))$ is $L_1\Theta(\sqrt{d})$ that are smaller than $Exp(ClosestPair(M_n))$ and $Exp(Diameter(M_n))$, respectively.

Let L_2 be the length of circuhypercube of S and G_n be a set of $n \ge 4$ uniformly distributed points in a d-dimensional hypercube with side L_2. Then the $Exp(ClosestPair(G_n))$ is $L_2\Theta(2^{-2/d})$ and $Exp(Diameter(G_n))$ is $L_2\Theta(\sqrt{d})$ that are greater than $Exp(ClosestPair(M_n))$ and $Exp(Diameter(M_n))$, respectively.

Therefore, the $Exp(\varphi(M_n))$ is between $\Omega(\frac{L_1}{L_2} n^{2/d})$ and $O(\frac{L_2}{L_1} n^{2/d})$. Since $\frac{L_1}{L_2}$ is a constant, the $Exp(\varphi(S_n))$ is $\Theta(n^{2/d})$.

4. Conclusion

In this paper, we studied the spread of a set of n uniformly distributed points, and we showed that the expected value of the spread is $\Theta(n^{2/d})$. For algorithms in which the spread appears in time or space complexity or another parameter, we can use the expected value of spread in the case that the input is uniformly distributed and get better results in this case. For example, Miller and Sheehy [12] propose an algorithm that computes the Voronoi diagram of a set of n points in $O(f \log n \log \Delta)$ time, where f is the output complexity and Δ

is the spread of the input point set. This gives a low time complexity when the input point set is uniformly distributed. Combining similar results with the results in this paper will result in very low time complexity results, see [2].

There are some interesting problems to be pursued. One is analyzing the spread of a set of points with other distributions like Normal, Poisson, etc.

Acknowledgment

The authors would like to thank the reviewers for carefully reading the paper and suggesting several corrections that improved the paper.

References

[1] J. Gao, L. J. Guibas, A. Nguyen, Deformable spanners and applications, Computational Geometry 35 (1-2) (2006) 2–19.
[2] F. Aurenhammer, Voronoi diagrams-a survey of a fundamental geometric data structure, ACM Computing Surveys (CSUR) 23 (3) (1991) 345–405.
[3] G. L. Miller, D. R. Sheehy, A new approach to output-sensitive Voronoi diagrams and Delaunay triangulations, in: Proceedings of the twenty-ninth Annual Symposium on Computational Geometry, ACM, 2013, pp. 281–288.
[4] J. Erickson, Dense point sets have sparse Delaunay triangulations or "... but not too nasty'', Discrete & Computational Geometry 33 (1) (2005) 83–115.
[5] K. Talwar, Bypassing the embedding: algorithms for low dimensional metrics, in: Proceedings of the thirty-sixth annual ACM symposium on Theory of Computing, ACM, 2004, pp. 281–290.
[6] S. Teramoto, T. Asano, N. Katoh, B. Doerr, Inserting points uniformly at every instance, IEICE transactions on information and systems 89 (8) (2006) 2348–2356.
[7] T. Asano, Online uniformity of integer points on a line, Information Processing Letters 109 (1) (2008) 57–60.
[8] A. Bishnu, S. Desai, A. Ghosh, M. Goswami, S. Paul, Uniformity of point samples in metric spaces using gap ratio, SIAM Journal on Discrete Mathematics 31 (3) (2017) 2138–2171.
[9] P. K. Agarwal, S. Har-Peled, K. R. Varadarajan, Approximating extent measures of points, Journal of the ACM (JACM) 51 (4) (2004) 606–635.
[10] P. Indyk, R. Motwani, S. Venkatasubramanian, Geometric matching under noise: Combinatorial bounds and algorithms, in: ACM-SIAM Symposium on Discrete Algorithms (SODA), 1999, pp. 457–465.
[11] A. M. Mood, Introduction to the Theory of Statistics., McGraw-hill, 1950.
[12] G. L. Miller, D. R. Sheehy, A new approach to output-sensitive construction of Voronoi diagrams and Delaunay triangulations, Discrete Computational Geometry 52 (3) (2014) 476–491.

Behnam Iranfar received his BSc in Computer Science from Payam Noor University, Isfahan, Iran in 2007 and his MSc in Computer Science from Bahonar University, Kerman, Iran, in 2014. Now he is a Ph.D. student in the Department of Computer Science, Yazd University, Yazd, Iran. His research interests are computational geometry, massive data, and uncertain data and algorithms.
Email: b_iranfar@stu.yazd.ac.ir

Mohammad Farshi received his BSc in Computer Science from Yazd University, Yazd, Iran in 1996 and his MSc in Pure Mathematics from Shiraz University, Shiraz, Iran, in 1999, and his Ph.D. in Computer Science from Eindhoven University of Technology, Eindhoven, The Netherlands, in 2008. Between 1999 and 2004, he worked at Yazd University as an instructor and from 2009 he has been an Assistant Professor at the same university. His research interests are Computational Geometry, Geometric Spanners and Algorithms.
Email: mfarshi@yazd.ac.ir

Amir Mesrikhani received his B.S. degree in computer science from Kharazmi University, Tehran, Iran, in 2011, and his MSc degrees in Institute for Advanced Studies in Basic Sciences (IASBS), Zanjan, Iran in 2014, and his Ph.D. in computer science from Yazd University, Yazd, Iran in 2019. His research interests are computational geometry, massive data, and approximation algorithms.
Email: mesrikhani@stu.yazd.ac.ir

Paper Handling Data:

Submitted: 05.30.2019
Received in revised form: 06.29.2020
Accepted: 07.05.2020
Corresponding author: Mohammad Farshi
Combinatorial and Geometric Algorithms Lab,
Department of Mathematical Sciences, Yazd University,
Yazd, Iran.