Electronic Colloquium on Computational Complexity, Report No. 13 (2005)

Theory and Application of Width Bounded Geometric Separator*

Bin Fu

Department of Computer Science
University of New Orleans, New Orleans, LA 70148, and
Research Institute for Children
200 Henry Clay Avenue, New Orleans, LA 70118.
fu@cs.uno.edu

Abstract

We introduce the notion of width bounded geometric separator, develop the techniques for its existence as well as algorithm, and apply it to obtain a $2^{O(\sqrt{n})}$ time exact algorithm for the disk covering problem, which seeks to determine the minimal number of fixed size disks to cover n points on plane, and was proven to be NP-complete [14]. Applying our separator to a class of NP-hard problems on disk graphs, we also greatly improve the exact algorithm for maximum independent set problem on disk graph to $2^{O(\sqrt{n})}$ from $n^{O(\sqrt{n})}$ [4, 1]. For a constant $a > 0$ and a set of points Q on the plane, an a-wide separator is the region between two parallel lines of distance a that partitions Q into Q_1 (in the left side of the region), S (inside the region), and Q_2 (in the right side of the region). If the distance is at least one between every two points in the set Q with n points, called 1-separated set, there is an a-wide separator that partitions Q into Q_1, S and Q_2 such that $|Q_1|, |Q_2| \leq (2/3)n$, and $|S| \leq 1.2126a\sqrt{n}$. As the separator for grid points gives sub-exponential time algorithm for the protein folding problem in the HP-model [15], the separator for 1-separated set provides a tool for studying the protein folding problem in more realistic model.

1. Introduction

The geometric separator has applications in many problems (e.g. [28, 8, 7, 36]). It plays an important role when developing divide and conquer algorithm for geometric problems. Lipton and Tarjan [27] showed the well known geometric separator for planar graphs. They proved every n vertices planar graph has at most $\sqrt{8n}$ vertices whose removal separates the graph into two disconnected parts of size at most $\frac{2}{3}n$. Their $\frac{2}{3}$-separator was improved to $\sqrt{6n}$ by Djidjev [10], $\sqrt{5n}$ by Gazit [16], $\sqrt{4.5n}$ by Alon, Seymour and Thomas [5] and $1.97\sqrt{n}$ by Djidjev and Venkatesan [11]. Spielman and Teng [38] showed a $\frac{3}{4}$-separator with size $1.82\sqrt{n}$ for planar graphs. The separators for more general graphs were derived in [17, 6, 35]. Some other forms of the geometric separators were studied by Miller, Teng, Thurston, and Vavasis [32, 33, 31] and Smith and Wormald [37]. Assume each input point is covered by a regular geometric object such as circle, rectangle, etc. If every point on the plane is covered by at most k objects, it is called k-thick. Some $O(\sqrt{k \cdot n})$ size separators and their algorithms were derived in [32, 33, 31, 37].

The planar graph separators were applied in deriving some $2^{O(\sqrt{n})}$-time algorithms for certain NP-hard problems on planar graph by Lipton, Tarjan [28], Ravi and Hunt [36]. Those problems include computing the numbers of maximum independence set, minimum vertex covers and three-colorings of a planar graph, and the number of satisfying truth assignments to a planar 3CNF formula [26]. In [37], their separators were applied in deriving $n^{O(\sqrt{n})}$-time algorithms for some geometric problems such as the planar Traveling Salesman and Steiner Tree problems on the plane. The separators were applied to parameterized independent set problem for planar graph by Alber, Fernau and Niedermeier [2, 3] and disk graph by Alber and Fiala [4].

*This research is supported by Louisiana Board of Regents fund under contract number LEQSF(2004-07)-RD-A-35.

ISSN 1433-8092

We introduce the concept of width bounded separator. For a set of points Q on the plane, an a-wide separator is the region between two parallel lines of distance a, which partitions the set Q into two balanced subsets and measures its size with the number of points from Q in the strip region. Our width bounded separator concept is geometrically natural, and can achieve much smaller constant c for its size upper bound $c\sqrt{n}$ than the previous approaches that we just mentioned above. Fu and Wang [15] developed a method for deriving sharper upper bound separator for grid points via controlling the distance to the separator line. They proved that for a set of n grid points on the plane, there is a separator that has $\leq 1.129\sqrt{n}$ points and has $\leq \frac{2}{3}n$ points on each side. It was used to obtain the first sub-exponential time algorithm for the protein folding problem in the HP model. This paper not only generalizes the results of [15], but also substantially improves the techniques in [15].

We would like to mention our new technical developments in this paper. 1) In order to apply the separator to more general geometric problems with arbitrary input points other than grid points, we use weighted points in Euclid space and the sum of weights to measure the quality of separator instead of counting the number of points close to it. We introduce the local binding method to merge some nearby points into a grid point. This method is combined with our separator in deriving a $2^{O(\sqrt{n})}$ time algorithm for the well-known disk covering problem, which seeks to determine the minimal number of fixed size discs to cover n points on the plane. To our knowledge, this is the first algorithm for the disk covering problem with running time bounded by an exponential with a sublinear exponent. This method can also obtain $2^{O(\sqrt{n})}$ time algorithms for a class of NP-hard problems on disk graph. For example, we greatly improve the exact algorithm for maximum independent set problem on disk graph to $2^{O(\sqrt{n})}$ time from $n^{O(\sqrt{n})}$ [4, 1]. 2) We will handle the case of high dimension. In [15], it uses the angle ratio $\frac{\theta}{\pi}$ to characterize the probability for a point p to have distance $\leq a$ to a random line through a point o, where θ is the angle between the two lines through o that the point p has distance $= a$ to both of them. We develop a new area ratio method to replace the previous angle ratio method [15] when deriving higher dimensional separator. 3) In order to study protein folding problem in more general model than the grid model like HP model [25], we develop a similar separator theorem for a set of points with distance at least 1 between any two of them, called 1-separated set, we establish the connection between this problem and the famous fixed size discs packing problem. The discs packing problem on 2D was well solved in the combinatorial geometry (see [39]). The 3D case, which is the Kepler conjecture, has a very long proof (see [34, 21]). It is still a very elusive problem at higher dimensions. Our Theorem 15 shows how the separator sizes depends on the packing density. 4) We develop a simple polynomial time algorithm to find the width-bounded separator at fixed dimensional space. This is a starting point for the algorithms finding width bounded geometric separator, and is enough for the applications to the exact algorithms for some NP-complete geometric problems.

The paper spends section 3 to prove some existence theorems for width bounded separator. Section 4 gives polynomial time algorithm for finding width bounded separator. Section 5 describes the $2^{O(\sqrt{n})}$-time algorithm for disk covering problem and maximum independent set problem on disk graph by using the results from sections 3 and 4 .

2. Overview of our methods

We describe our techniques in the 2-dimensional case. For a set of arbitrary points Q on the plane, we also consider another set of grid points P on the plane. Each point $p \in P$ is assigned a positive weight which is used to measure the density for the points of set Q near p. A good separator line L partitions the set Q into two balanced parts. Furthermore, it is also expected to have small a number of points from Q close it. This will be measured by the sum of weights of the points of P close to L. This approach makes it more flexible than counting the number of points of Q close to L. The point set Q has a center point (see Lemma 1) such that every line through it has a balanced partition for Q. For a random line L through the center point, the expected sum of weights of points P close to it is maximal when all points of P stay at the least circle with center at o. Furthermore, the points of P with larger weights are closer to the center than the points with smaller weights. When the number of different weights is small, it gives us an easy way to compute such an expectation, which is used as the upper bound for the quality of the best separator.

Finding an a-wide separator is straightforward with $O(n^3)$-time. Let each point $p \in P$ be covered with a insulation circle with radius $a/2$ and center at p. A good separator line L can be moved until it touches points of Q or is tangent to some insulation circles with radius $a/2$ and center at points of P. The movement neither crosses any point of Q nor enters any new insulation circle of P. It does not change the balance result and size measure of the separator. Trying all lines that pass through points of Q or are tangent to the insulation circles with center at points of P, it can be done in $O(n^3)$ time.

For covering a set of points Q on the plane, the set P is a set of grid points that have points from Q close to each of them. A grid point p's is assigned weight i if there are 2^i to 2^{i+1} points of Q on the 1×1 grid square with p as the center. A balanced separator line for Q also has small sum of weights $(O(\sqrt{n}))$ for the points of P near the line. This gives at most $2^{O(\sqrt{n})}$ ways to cover all points of Q close to the separator line and decompose the problem into two problems Q_1 and Q_2 that can be covered independently. It makes the total time to be $2^{O(\sqrt{n})}$.

3. Theory of width-bounded separators on the d-dimension

Throughout the section 3 we assume the dimensional number d is fixed. We will use the following well known fact that can be easily derived from Helly theorem (see [18, 39]), and will be used to obtain our width bounded separator.

Lemma 1. *For an n-element set P in d-dimensional space, there is a point q with the property that any half-space that does not contain q, covers at most $\frac{d}{d+1}n$ elements of P. (Such a point q is called a centerpoint of P).*

Definition 2. For two points p_1, p_2 in the d-dimensional Euclid space R^d, $\text{dist}(p_1, p_2)$ is the Euclid distance between p_1 and p_2. For a set $A \subseteq R^d$, $\text{dist}(p_1, A) = \min_{q \in A} \text{dist}(p_1, q)$. For $a > 0$ and a set A of points on d-dimensional space, if the distance between every two points is at least a, the set A is called *a-separated*. For $\epsilon > 0$ and a points set $Q \subseteq R^d$, an *ϵ-net* of Q is another points set $P \subseteq R^d$ such that each point in Q has distance $\leq \epsilon$ to some point in P. We say P is a net of Q if P is an ϵ-net of Q for some constant $\epsilon > 0$ (that does not depend on the size of Q if it can be very large). A net set is usually an 1-separated set such as grid points set. A weight function $w : P \to [0, \infty)$ is often used to measure the points density of Q near each point of P. Let $f : R^d \to R$ be continuous and piece-wise smooth function. Its *surface* $L(f) = \{v \in R^d | f(v) = 0\}$, a $d-1$-dimensional manifold, characterizes a regular geometric shape such as ball, square, plane, etc. A *hyperplane* in R^d through a fixed point $p_0 \in R^d$ is defined by equation $(p - p_0) \cdot v = 0$, where v is normal vector of the plane and " ." is the regular inner product ($u \cdot v = \sum_{i=1}^{d} u_i v_i$ for $u = (u_1, \cdots, u_d)$ and $v = (v_1, \cdots, v_d)$). Let F be a class of functions $f : R^d \to R$, whose surfaces represent a class of simple geometric shapes. For $Q \subseteq R^d$ with net $P \subseteq R^d$, constant $a > 0$, and the weight function $w : P \to [0, \infty)$, an *a-wide-F-separator* is determined by the surface $L(f)$ for some $f \in F$, which has two measurements for its quality of separation: (1) $\text{balance}(L(f), Q) = \frac{\max(|Q_1|, |Q_2|)}{|Q|}$, where $Q_1 = \{q \in Q | f(q) < 0\}$ and $Q_2 = \{q \in Q | f(q) > 0\}$; and (2)$\text{measure}(L(f), P, \frac{a}{2}, w)$, where $\text{measure}(A, P, x, w) = \sum_{p \in P, \text{dist}(p, A) \leq x} w(p)$. In particular, an *$a$-$F$-wide separator* is simply called an *a-wide separator* when F is the set of all linear functions $f : R^d \to R$, whose surface is a hyper plane. Sometimes, if f is fixed, we use $\text{balance}(L, Q)$ and $\text{measure}(L, P, \frac{a}{2}, w)$ to represent $\text{balance}(L(f), Q)$ and $\text{measure}(L(f), P, \frac{a}{2}, w)$ respectively.

3.1. Volume, area, integrations and probability

We need some integrations for computing volume and surface area size at high dimensions. Some of the materials can be found in standard calculus books. We will treat the case of any fixed dimension. The reader is recommended to understand the cases $d = 2$ and 3 first. We use the standard polar transformation

$$\begin{aligned}
x_d &= r \cos \theta_{d-1}; \\
x_{d-1} &= r \sin \theta_{d-1} \cos \theta_{d-2}; \\
&\cdots\cdots \\
x_2 &= r \sin \theta_{d-1} \sin \theta_{d-2} \cdots \sin \theta_2 \cos \theta_1; \\
x_1 &= r \sin \theta_{d-1} \sin \theta_{d-2} \cdots \sin \theta_2 \sin \theta_1.
\end{aligned}$$

It is a smooth one-one and onto map from $[0, R] \times [0, \pi] \times \cdots \times [0, \pi] \times [0, 2\pi]$ to the d-dimensional ball of radius R with center at the origin. The Jacobian form is

$$J_d(r, \theta_{d-1}, \cdots, \theta_1) = \frac{\partial(x_d, x_{d-1}, \cdots, x_1)}{\partial(r, \theta_{d-1}, \cdots, \theta_1)} = \begin{vmatrix} \frac{\partial x_d}{\partial r} & \frac{\partial x_{d-1}}{\partial r} \cdots \frac{\partial x_1}{\partial r} \\ \frac{\partial x_d}{\partial \theta_{d-1}} & \frac{\partial x_{d-1}}{\partial \theta_{d-1}} \cdots \frac{\partial x_1}{\partial \theta_{d-1}} \\ \cdots & \\ \frac{\partial x_d}{\partial \theta_1} & \frac{\partial x_{d-1}}{\partial \theta_1} \cdots \frac{\partial x_1}{\partial \theta_1} \end{vmatrix}.$$

We can easily see the recursive equation that $J_d(r, \theta_{d-1}, \cdots, \theta_1) = r \cdot (\sin \theta_{d-1})^{d-2} \cdot J_{d-1}(r, \theta_{d-2}, \cdots, \theta_1)$ for $d > 2$. This gives the explicit expression: $J_d(r, \theta_{d-1}, \cdots, \theta_1) = r^{d-1} \cdot (\sin \theta_{d-1})^{d-2} \cdot (\sin \theta_{d-2})^{d-3} \cdots (\sin \theta_2)$. Let $B_d(R, o)$ be the d-dimensional ball of radius R and center o. The volume of d-dimensional ball of radius R is

$$\begin{aligned} V_d(R) &= \int_{B_d(R,o)} 1 d_z = \int_0^R \int_0^\pi \cdots \int_0^\pi \int_0^{2\pi} |J_d(r, \theta_{d-1}, \cdots, \theta_2, \theta_1)| d_r d_{\theta_{d-1}} \cdots d_{\theta_2} d_{\theta_1} && (1) \\ &= \begin{cases} \frac{2^{(d+1)/2} \pi^{(d-1)/2}}{1 \cdot 3 \cdots (d-2) \cdot d} R^d & \text{if } d \text{ is odd} \\ \\ \frac{2^{d/2} \pi^{d/2}}{2 \cdot 4 \cdots (d-2) \cdot d} R^d & \text{otherwise} \end{cases} && (2) \end{aligned}$$

Let the d-dimensional ball have the center at o. We also need the integration as follows:

$$\int_{B_d(R,o)} \frac{1}{\text{dist}(z, o)} d_z = \int_0^R \int_0^\pi \cdots \int_0^\pi \int_0^{2\pi} \frac{|J_d(r, \theta_{d-1}, \cdots, \theta_2, \theta_1)|}{r} d_r d_{\theta_{d-1}} \cdots d_{\theta_2} d_{\theta_1} = \frac{d}{(d-1)R} V_d(R) \quad (3)$$

Let $V_d(r) = v_d \cdot r^d$, where v_d is constant for fixed dimensional number d. In particular, $v_1 = 2, v_2 = \pi$ and $v_3 = \frac{4\pi}{3}$. Define $A_d(h, R) = \{(x_1, \cdots, x_d) | \sum_{i=1}^d x_i^2 \le R^2 \text{ and } 0 \le x_1 \le h\}$, which is a horizontal cross section of d-dimensional half ball. The volume of $A_d(h, R)$ at d-dimensional space is calculated by

$$U_d(h, R) = \int_0^h V_{d-1}(\sqrt{R^2 - x_1^2}) \, d_{x_1} = v_{d-1} \int_0^h \left(\sqrt{R^2 - x_1^2}\right)^{d-1} d_{x_1} \quad (4)$$

The surface area size of 3D ball ($4\pi R^2$) is the derivative of its volume ($\frac{4}{3}\pi R^3$). The boundary length of circle ($2\pi R$) is the derivative of its area size (πR^2). This fact can be extended to higher dimensional ball and the cross section of ball. The surface area size of $B_d(R, o)$ is $W_d(R) = \frac{\partial V_d(R)}{\partial R} = d \cdot v_d \cdot R^{d-1}$. The side surface of $A_d(h, R)$ is $\{(x_1, \cdots, x_d) | \sum_{i=1}^d x_i^2 = R^2 \text{ and } 0 \le x_1 \le h\}$. Its area size is

$$S_d(h, R) = \frac{\partial U_d(h, R)}{\partial R} = (d-1) v_{d-1} R \int_0^h \left(\sqrt{R^2 - x_1^2}\right)^{d-3} d_{x_1}$$

When R is fixed and h is small, we have $S_d(h, R) = v_{d-1} \cdot (d-1) \cdot R^{d-2} \cdot h + O(h^2)$. For a parameter $a > 0$, the probability that a d-dimensional point p to have a distance $\le a$ to a random plane through origin will be determined. This probability at dimension 3 was not well treated in [15].

Lemma 3. *Let $a > 0$ be a constant. Let p and o be the two points on d-dimensional space, the probability that p has distance $\le a$ to a random plane through o is in $[\frac{h_d \cdot a}{\text{dist}(p,o)} - \frac{c_0}{\text{dist}^2(p,o)}, \frac{h_d \cdot a}{\text{dist}(p,o)} + \frac{c_0}{\text{dist}^2(p,o)}]$, where $h_d = \frac{2(d-1)v_{d-1}}{d \cdot v_d}$ and c_o are constants for fixed d. In particular, $h_2 = \frac{2}{\pi}$ and $h_3 = 1$.*

Proof: Let o be the origin $(0, \cdots, 0)$. The point p can be moved to an axis via rotation that does not change the probability. Let's assume the point $p = (x_1, 0, \cdots, 0)$, where $x_1 = \text{dist}(p, o)$. For an unit vector $v = (v_1, \cdots, v_d)$ with $v_1 \ge 0$ in d-dimensional space, the plane through the origin with normal vector v is defined $u \cdot v = 0$, where $\cdot$ represents the regular inner product between two vectors. The distance between p to the plane is $|p \cdot v| = x_1 v_1$. If $x_1 v_1 \le a$, it implies $v_1 \le \frac{a}{x_1}$. The area size of $\{(v_1, \cdots, v_d) | \sum_{i=1}^d v_i^2 = 1 \text{ and } 0 \le v_1 \le \frac{a}{x_1}\}$ is $S_d(\frac{a}{x_1}, 1)$. The probability that p has distance $\le a$ to a random plane through the origin is $\frac{S_d(\frac{a}{x_1}, 1)}{\frac{1}{2} \cdot W_d(1)} = h_d \cdot \frac{a}{\text{dist}(p,o)} + O(\frac{1}{\text{dist}^2(p,o)})$. ∎

3.2. Width Bounded Separator

Definition 4. The *diameter* of a region R is $\sup_{p_1,p_2\in R} \text{dist}(p_1,p_2)$. A (b,c)*-partition* of d-dimensional space makes the space as the disjoint unions of regions $P_1, P_2, \cdots$ such that each P_i, called a *regular region*, has volume equal to b and the diameter of each P_i is $\le c$. A (b,c)*-regular point set* A is a set of points on a d-dimensional space with (b,c)-partition $P_1, P_2, \cdots$ such that each P_i contains at most one point from A. For two regions A and B, if $A \subseteq B$ ($A \cap B \neq \emptyset$), we say B *contains* (*intersects* resp.) A.

Lemma 5. *Assume $P_1, P_2, \cdots$ form a (b,c)-partition on d-dimensional space. We have (i) every d-dimensional ball of radius r intersects at most $\frac{v_d\cdot(r+c)^d}{b}$ regular regions; (ii) every d-dimensional ball of radius r contains at least $\frac{v_d\cdot(r-c)^d}{b}$ regular regions; (iii) every d-dimensional ball of radius $\left(\frac{nb}{v_d}\right)^{\frac{1}{d}}+c$ contains at least n (b,c)-regular regions in it; and (vi) every d-dimensional ball of radius $\left(\frac{nb}{v_d}\right)^{\frac{1}{d}}-c$ intersects at most n (b,c)-regular regions.*

Proof: (i) If a (b,c)-regular region P_i intersects a ball C of radius r at center o, the regular region P_i is contained by the ball C' of radius $r+c$ at the same center o. The number of regular regions contained by C' is no more than the volume size of the ball C' divided by b. (ii) If a regular region P_i intersects a ball C' of radius $r-c$ at center o, P_i is contained in the ball C of radius r at the same center o. The number of those regular regions intersecting C' is at least the volume size of the ball C' divided by b. (iii)Apply $r = (\frac{bn}{v_d})^{\frac{1}{d}} + c$ to (ii). (vi)Apply $r = (\frac{bn}{v_d})^{\frac{1}{d}} - c$ to (i). ∎

Definition 6. Let $a > 0$, p and o be two points in d-dimensional space. Define $Pr_d(a, p_0, p)$ to be the probability that the point p has $\le a$ perpendicular distance to a random hyper plane L through the point p_0. Define function $f_{a,p,o}(L) = \begin{cases} 1 & \text{if } p \text{ has distance} \le a \text{ to the hyper plane } L \text{ through } o; \\ 0 & \text{otherwise.} \end{cases}$

The expectation of function $f_{a,p,o}$ is $E(f_{a,p,o}) = Pr_d(a, o, p)$. Assume $P = \{p_1, p_2, \cdots, p_n\}$ is a set of n points in R^d and each p_i has weight $w(p_i) \ge 0$. Define function $F_{a,P,o}(L) = \sum_{p\in P} w(p) f_{a,p,o}(L)$. We give an upper bound for the expectation $E(F_{a,P,o})$ for $F_{a,P,o}$ in the lemma below.

Lemma 7. *Let $a, b, c > 0$ be constants and $\delta > 0$ be a small constant. Assume that $P_1, P_2, \cdots$ form a (b,c) partition in R^d. Let $w_1 > w_2 > \cdots > w_k > 0$ be positive weights, and $P = \{p_1, \cdots, p_n\}$ be the d-dimensional (b,c)-regular points set in R^d. Let w be a mapping from P to $\{w_1, \cdots, w_k\}$ and n_j be the number of points $p \in P$ with $w(p) = w_j$. Let o be a fixed point in R^d (a center point). For a random hyper plane passing through o, we have $E(F_{a,P,o}) \le (\frac{d\cdot h_d\cdot v_d}{(d-1)\cdot b}+\delta)\cdot a \cdot \sum_{j=1}^k w_j (r_j^{d-1} - r_{j-1}^{d-1}) + c_2 \sum_{j=1}^{k-1} w_{j+1} \cdot r_j^{d-2} + c_1 \cdot w_1$, where (1) $r_0 = 0$ and $r_i (i > 0)$ is the least radius such that $B_d(r_i, o)$ intersects at least $\sum_{j=1}^i n_j$ regular regions, (2) c_1 and c_2 are constants for fixed d, and (3) h_d and v_d are constants defined in section 3.1.*

Proof: Assume $p = (x, y)$ is a point of P and L is a random plane passing through the center $o = (x_0, y_0)$. Let C be the ball of radius r and center o such that C covers all points in P. Let C' be the ball of radius $r' = r + c$ and the same center o. It is easy to see every regular region with a point in P is inside C'. The probability that the point p has distance $\le a$ to L is $\le h_d \cdot \frac{a}{\text{dist}(o,p)} + \frac{c_0}{\text{dist}(o,p)^2}$ (by Lemma 3).

Let $\epsilon > 0$ be a small constant which will be determined later. Select constant R_0 to be large enough such that for every point p with $dist(o,p) \ge R_0$, $\frac{1}{\text{dist}(o,p')} + \frac{c_0}{h_d\cdot a\cdot \text{dist}(o,p')^2} < \frac{1+\epsilon}{\text{dist}(o,p)}$ for every point p' with $\text{dist}(p', p) \le c$. Let P_1 be the set of all points p in P such that $\text{dist}(o,p) < R_0$. For each point $p \in P_1$, $Pr_d(a, o, p) \le 1$. For every point $p \in P - P_1$, $Pr_d(a, o, p) \le h_d \cdot \frac{a}{\text{dist}(o,p)} + \frac{c_0}{\text{dist}(o,p)^2} < \frac{h_d\cdot a(1+\epsilon)}{\text{dist}(o,p)}$.

$$E(F_{a,P,o}) = E(\sum_{i=1}^n w(p_i)\cdot f_{a,p_i,o}) = \sum_{i=1}^n w(p_i)\cdot E(f_{a,p_i,o}) = \sum_{j=1}^k w_j \sum_{w(p_i)=w_j} E(f_{a,p_i,o}) \quad (5)$$

$$= \sum_{j=1}^k w_j \sum_{w(p_i)=w_j} Pr_d(a, o, p_i) < \sum_{j=1}^k w_j \sum_{w(p_i)=w_j} \frac{h_d\cdot a\cdot(1+\epsilon)}{b} \cdot \frac{1}{\text{dist}(o,p_i)} \cdot b \quad (6)$$

It is easy to see that the contribution to $E(F_{a,P,o})$ from the points in P_1 is $\leq w_1|P_1| \leq w_1 \cdot \frac{v_d(R_0+c)^d}{b} = w_1 c_1$ (by Lemma 5), where $c_1 = \frac{v_d(R_0+c)^d}{b}$. Next we only consider those points from $P - P_1$. The sum (6) is maximal when $dist(p,o) \leq dist(p',o)$ implies $w(p) \geq w(p')$. The ball C' is partitioned into k ring regions such that the j-th area is between $B_d(r_j, o)$ and $B_d(r_{j-1}, o)$ and is mainly used to hold those points with weight w_j. Notice that each regular region has diameter $\leq c$ and holds at most one point in P. It is easy to see that all points of $\{p_i | w(p_i) = w_j\}$ are located between $B_d(r_j, o)$ and $B_d(r_{j-1} - c, o)$ when (6) is maximal.

$$\sum_{w(p_i)=w_j} \frac{h_d \cdot a \cdot (1+\epsilon)}{b} \cdot \frac{1}{\mathrm{dist}(o,p_i)} \cdot b \leq \frac{h_d \cdot a \cdot (1+\epsilon)^2}{b} \int_{B_d(r_j,o)-B_d(r_{j-1}-c,o)} \frac{1}{\mathrm{dist}(o,z)} d_z \tag{7}$$

$$= \frac{h_d \cdot a \cdot (1+\epsilon)^2}{b} \int_{r_{j-1}-c}^{r_j} \int_0^{\pi} \cdots \int_0^{\pi} \int_0^{2\pi} \frac{J_d(r, \theta_{n-1}, \cdots, \theta_2, \theta_1)}{r} d_r d_{\theta_{n-1}} \cdots d_{\theta_2} d_{\theta_1} \tag{8}$$

$$= \frac{h_d \cdot a \cdot (1+\epsilon)^2}{b} \cdot \frac{d}{(d-1)} \cdot \left(\frac{V_d(r_j)}{r_j} - \frac{V_d(r_{j-1}-c)}{r_{j-1}-c} \right) \tag{9}$$

$$< \left(\frac{d \cdot h_d \cdot v_d}{(d-1) \cdot b} + \delta \right) \cdot a \cdot (r_j^{d-1} - r_{j-1}^{d-1}) + O(r_{j-1}^{d-2}). \tag{10}$$

Note: (8) → (9) → (10) follows from (3), and selecting ϵ small enough. ∎

Lemma 8. *Let o be a point on the plane, $a, b, c > 0$ be constants and $\epsilon, \delta > 0$ be small constants. Assume that $P_1, P_2, \cdots$ form a (b,c) partition in R^d. The weights $w_1 > \cdots > w_k > 0$ satisfy $k \cdot \max_{i=1}^k \{w_i\} = O(n^\epsilon)$. Let P be a set of n weighted (b,c)-regular points in a d-dimensional plane with $w(p) \in \{w_1, \cdots, w_k\}$ for each $p \in P$. Let n_j be the number of points $p \in P$ with $w(p) = w_j$ for $j = 1, \cdots, k$. We have $E(F_{o,P,a}) \leq (k_d \cdot (\frac{1}{b})^{\frac{1}{d}} + \delta) \cdot a \cdot \sum_{j=1}^k w_j \cdot n_j^{\frac{d-1}{d}} + O(n^{\frac{d-2}{d}+\epsilon})$, where $k_d = \frac{d \cdot h_d}{d-1} \cdot v_d^{\frac{1}{d}}$. In particular, $k_2 = \frac{4}{\sqrt{\pi}}$ and $k_3 = \frac{3}{2} \left(\frac{4\pi}{3}\right)^{\frac{1}{3}}$.*

Proof: Let r_j be the least radius such that the ball of radius r_j intersects at least $\sum_{i=1}^j n_i$ regular regions ($j = 1, \cdots, k$). By Lemma 5, $\left(\frac{(\sum_{i=1}^j n_i) b}{v_d}\right)^{\frac{1}{d}} - c \leq r_j \leq \left(\frac{(\sum_{i=1}^j n_i) b}{v_d}\right)^{\frac{1}{d}} + c$ for $j = 1, \cdots, k$.

$$r_j^{d-1} - r_{j-1}^{d-1} \leq \left(\left(\frac{(\sum_{i=1}^j n_i) b}{v_d} \right)^{\frac{1}{d}} + c \right)^{d-1} - \left(\left(\frac{(\sum_{i=1}^{j-1} n_i) b}{v_d} \right)^{\frac{1}{d}} - c \right)^{d-1} \tag{11}$$

$$= \left(\frac{b}{v_d} \right)^{\frac{d-1}{d}} \left((\sum_{i=1}^j n_i)^{\frac{d-1}{d}} - (\sum_{i=1}^{j-1} n_i)^{\frac{d-1}{d}} \right) + O((\sum_{i=1}^j n_i)^{\frac{d-2}{d}}) \tag{12}$$

$$= \left(\frac{b}{v_d} \right)^{\frac{d-1}{d}} n_j^{\frac{d-1}{d}} + O((\sum_{i=1}^j n_i)^{\frac{d-2}{d}}) \tag{13}$$

By Lemma 7, the lemma is proved. ∎

Definition 9. Let $a_1, \cdots, a_d > 0$ be positive constants. A $(a_1, \cdots, a_d)$-*grid regular partition* divides the d-dimensional space into disjoint union of $a_1 \times \cdots \times a_d$ rectangular regions. A $(a_1, \cdots, a_d)$-*grid regular point* is a corner point of a rectangular region. Under certain translation and rotation, each $(a_1, \cdots, a_d)$-grid regular point has coordinates $(a_1 t_1, \cdots, a_d t_d)$ for some integers $t_1, \cdots, t_d$.

Theorem 10. *Let $a, a_1, \cdots, a_d > 0$ be constants and $\epsilon, \delta > 0$ be small constants. Let P be a set of n $(a_1, \cdots, a_d)$-grid points in R^d, and Q be another set of m points in R^d with net P. Let $w_1 > w_2 \cdots > w_k > 0$ be positive weights with $k \cdot \max_{i=1}^k \{w_i\} = O(n^\epsilon)$, and w be a mapping from P to $\{w_1, \cdots, w_k\}$. There is a hyper plane L such that (1) each half space has $\leq \frac{d-1}{d} m$ points from Q, and (2) for the subset $A \subseteq P$*

containing all points in P with $\le a$ distance to L has the property $\sum_{p\in A} w(p) \le \left(k_d \cdot \left(\prod_{i=1}^d a_i\right)^{\frac{-1}{d}} + \delta\right) \cdot a \cdot \sum_{j=1}^k w_j \cdot n_j^{\frac{d-1}{d}} + O(n^{\frac{d-2}{d}+\epsilon})$ *for all large n.*

Proof: Let $b = \prod_{i=1}^d a_i$, $c = \sqrt{\sum_{i=1}^d a_i^2}$, and the point o be the center point of Q via Lemma 1. Apply Lemma 8. ∎

Corollary 11. *[15] Let Q be a set of n $(1,1)$-grid points on the plane. There is a line L such that each half plane has $\le \frac{2n}{3}$ points in Q and the number of points in Q with $\le \frac{1}{2}$ vertical distance to L is $\le 1.129\sqrt{n}$.*

Proof: Let all points of Q have weight 1, $k = 1, a = \frac{1}{2}$ and $P = Q$. Apply Theorem 10. ∎

Corollary 12. *Let Q be a set of n $(1,1,1)$-grid points on the 3D Euclid space. There is a plane L such that each half space has $\le \frac{3n}{4}$ points in Q and the number of points in Q with $\le \frac{1}{2}$ vertical distance to L is $\le 1.209n^{\frac{2}{3}}$.*

Corollaries 11 and 12 are the separators for the 2D and 3D grid graphes respectively. An edge connecting two neighbor grid points has distance 1. If two neighbor grid points are at different sides of the separator, one of them has distance $\le \frac{1}{2}$ to the separator. The work [15] for studying the protein folding in the HP-model lets us introduce the notion of width bounded separator. The atoms positions of real protein molecular do not follow the grid model like HP model [25]. In order to study the protein folding in more realistic model, we would like to obtain the separator for 1-separated set. The distance 1 represents the minimal distance between atoms in the protein. Such a separator may be useful for studying other geometric problems. The results in the next several sections do not depend on the results in the rest of this section.

Definition 13. Let $C = \{C_1, C_2, \cdots\}$ be a collection of balls in the d-dimensional Euclid space and let D be a region, C is called *packing* in D if $\cup_i C_i \subseteq D$ and no two of them have an interior point in common. The *density* of C with respect to D is defined as $d(C, D) = \frac{\sum_i A(C_i)}{A(D)}$, where the sum is over all i for which $C_i \cap D \ne \emptyset$ and $A(B)$ is the volume size of region B. Let $\overline{d}(C, D) = \lim_{r\to\infty} \sup d(C, D(r))$, where $D(r)$ is the ball of radius r centered at a fixed point o in D.

Lemma 14. *Let $\epsilon > 0$ be a small constant. Let P' be a 1-separated set with n points, and P be another set in the R^d. The weight function $w(p) = 1$ for every $p \in P \cup P'$. Let ϵ be a small constant. Let o' and o be two points in R^d. The function $g : P \to P'$ be an one-one and onto mapping with* $\mathrm{dist}(o, p) \ge \frac{\mathrm{dist}(o', g(p))}{s}$ *for every $p \in P$. We have $E(F_{a,P,o}) \le (1+\epsilon)s \cdot E(F_{a,P',o'}) + c_4$ for all large n, where c_4 is a constant.*

Proof: Since $E(f_{a,p,o}) = Pr_d(a, o, p)$, by Lemma 3, we can select constant R_0 big enough such that $E(f_{a,p,o}) \le (1+\epsilon)s \cdot E(f_{a,g(p),o'})$ for all points $p \in P$ with $\mathrm{dist}(p, o) > R_0$. $E(f_{a,p,o}) \le 1$ for each point $p \in P$ with $\mathrm{dist}(p, o) \le R_0$. ∎

Theorem 15. *Assume the packing density for d-dimensional ball has $\overline{d}(C, R^d) \le D_d$. For every 1-separated set Q on the d-dimensional Euclid space, there is a hyper-plane L with* $\mathrm{balance}(L, Q) \le \frac{d-1}{d}$ *and the number of points with distance $\le a$ to L is $(2k_d \cdot \left(\frac{D_d}{v_d}\right)^{\frac{1}{d}} + o(1))a \cdot n^{\frac{d-1}{d}}$.*

Proof: Select $R_0 > 0$ large enough such that if $R > R_0$ and n balls of radius $r = \frac{1}{2}$ are packed into $B_d(R, o)$, then $\frac{n\cdot v_d r^d}{v_d R^d} < D_d + \epsilon$. Let $R_i = (1+\epsilon)^i R_0$ for $i = 1, 2, \cdots$. Let n_i be the number of balls packed in $B_d(R_i, o)$. Since $\frac{n_i \cdot v_d r^d}{v_d R_i^d} < D_d + \epsilon$, we have $n_i < \frac{R_i^d (D_d+\epsilon)}{r^d}$. Let $R_i' = \left(\frac{n_i b}{v_d}\right)^{\frac{1}{d}} + c < \left(\frac{\left(\frac{R_i^d(D_d+2\epsilon)}{r^d}\right)b}{v_d}\right)^{\frac{1}{d}} =$

$\frac{R_i}{r} \cdot \left(\frac{(D_d+2\epsilon)b}{v_d}\right)^{\frac{1}{d}}$. The ball $B_d(R'_i, o')$ contains $\geq n_i$ regular (b,c)-regions by Lemma 5. Let g be a function that maps every center point c_k of ball C_k of radius r in $B_d(R_i, o)$ to a point p'_k in one of the (b,c)-regular regions of $B_d(R'_i, o')$, and $g(c_k)$ and $g(c_j)$ belong to different (b,c) regular regions for $c_k \neq c_j$. For every ball center c_k between $B_d(R_i, o)$ and $B_d(R_{i+1}, o)$, it is mapped to a point p'_k in $B_d(R'_{i+1}, o')$. We have $\mathrm{dist}(o, c_k) \geq R_i = \frac{R_i}{R'_{i+1}} \cdot R'_{i+1} \geq \frac{\mathrm{dist}(o', p'_k)}{\frac{R'_{i+1}}{R_i}} \geq \frac{\mathrm{dist}(o', p'_k)}{s}$, where $s = \frac{1}{r} \cdot \left(\frac{(D_d+2\epsilon)b}{v_d}\right)^{\frac{1}{d}} \geq \frac{R'_{i+1}}{R_{i+1}}$. The theorem follows from Lemmas 1, 14 and 8. ∎

Corollary 16. *Let Q be a 1-separated set on the d-dimensional Euclid space. There is an a-wide separator L such that* $\mathrm{balance}(L, Q) \leq \frac{d-1}{d}$ *and the number of points with distance $\leq a/2$ to L is $(k_d \cdot (\frac{1}{v_d})^{\frac{1}{d}} + o(1)) \cdot a \cdot n^{\frac{d-1}{d}}$.*

Proof: The packing density is always ≤ 1. Apply Theorem 15. Notice that the distance to L should be $\leq a/2$ instead of a for the definition of a-wide separator. ∎

Theorem 17. *[39] Given a packing C with congruent copies of discs, $\overline{d}(C, R^2) \leq \frac{\pi}{\sqrt{12}}$.*

Corollary 18. *Let Q be a set of points on 2-dimensional plane. Every two points have distance ≥ 1. There is an a-wide separator L such that each half plane has at most $\frac{2}{3}n$ points in Q, and the number points in Q with distance $\leq a/2$ to L is $\leq 1.2126a\sqrt{n}$.*

Proof: Apply Theorem 15 with $D_2 = \frac{\pi}{\sqrt{12}}$. ∎

4. Algorithm for finding separator

In order to make sections 4 and 5 easy to follow, we focus on the 2-dimension case. We show there is an $O(n^3)$-time algorithm for finding separator in 2-dimensional plane. The technique for finding the separator can be easily extended to higher dimension after minor adjustments. The algorithm is essentially brute-force, but it is enough for its application in the next section. We have obtained almost linear time algorithm at 2-dimensional case, which is more involved and will be presented in the coming paper. For $a > 0$, $C_a(p)$ is the circle of radius a and center p.

Lemma 19. *Let constant $a > 0$, and small constant $\epsilon > 0$. Let P and Q be two sets of points on the plane. Let w be a mapping from P to $[0, +\infty)$. Let L' be a line on the plane. Then there is another line L such that* $\mathrm{measure}(L, P, a, w) \leq \mathrm{measure}(L', P, a+\epsilon)$ *and* $\mathrm{balance}(L, Q) = \mathrm{balance}(L', Q)$. *Furthermore L has one of the following properties: 1)L is through two points in Q or, 2)L is through one point from Q and is tangent to $C_{a+\epsilon}(p)$ for some $p \in P$ or, 3)L is tangent to both $C_{a+\epsilon}(p_1)$ and $C_{a+\epsilon}(p_2)$ for some $p_1, p_2 \in P$.*

Proof: We move the line L' along the direction vertical to itself until it touches a point q in Q or is tangent to a circle $C_{a+\epsilon}(p)$ for some $p \in P$. Rotate L' around q or $C_{a+\epsilon}(p)$ (keep the tangent relationship) until it touches another point q' in Q or tangent to $C_{a+\epsilon}(p')$ for some $p' \in P$. The movement of the line L' increases neither the number of points of Q in any of two sides of L', nor the number of points in P with $< a+\epsilon$ distance to L' on any side of L'. ∎

Theorem 20. *Let constant $a, a_1, a_2, \alpha > 0$ and small constants $\epsilon, \delta > 0$. Let P be a set of (a_1, a_2)-grid regular points and Q be another set of points of the plane. The weight function w is from P to $\{w_1, \cdots, w_k\}$, and the weights $w_1 > \cdots > w_k > 0$ have $k \cdot \max_{i=1}^{k}\{w_i\} = O(n^\alpha)$. There is an $O(n^3)$ time algorithm that finds a separator line L such that* $\mathrm{balance}(L, Q) \leq \frac{2}{3}$, *and* $\mathrm{measure}(L, P, a, w) \leq \left(\frac{k_2 \cdot a}{\sqrt{a_1 \cdot a_2}} + \delta\right) \sum_{i=1}^{k} w_i \sqrt{n_i} + O(n^\alpha)$ *for all large n.*

Proof: Let $\delta_1, \delta_2 > 0$ be small enough such that $k_2 \cdot \delta_1 + \delta_2\delta_1 + a\delta_1 < \delta$. By Theorem 10, there is a line L' with balance$(L', Q) \leq \frac{2}{3}$ and measure$(L', P, a + \delta_1, w) \leq (k_2 + \delta_2) \cdot (a + \delta_1) \cdot (\sum_{j=1}^{k} w_j \sqrt{n_j}) + O(n^\alpha) \leq (k_2 \cdot a + \delta) \cdot (\sum_{j=1}^{k} w_j \sqrt{n_j}) + O(n^\alpha)$. By Lemma 19 and the algorithm below (with $\epsilon = \delta_1$), it can be found in $O(n^3)$ time.

Algorithm

Input $a, \epsilon > 0$, and points sets P, Q on the 2D plane;
$B = \infty$ and $L = \emptyset$;
for each $q_1 \in Q$ and $q_2 \in Q$
  let L_0 be the line through both q_1 and q_2
  if (balance$(L_0, Q) \leq \frac{2}{3}$ and (measure$(L_0, Q, a + \epsilon, w) < B$))
  then $B =$ measure$(L_0, Q, a + \epsilon, w)$ and $L = L_0$.
for each $p \in P$ and $q \in Q$
  let L_1, L_2 be the lines through q and tangent to $C_{a+\epsilon}(p)$
  for (i=1 to 2) if (balance$(L_i, Q) \leq \frac{2}{3}$ and (measure$(L_i, P, a + \epsilon, w) < B$))
    then $B =$ measure$(L_i, P, a + \epsilon, w)$ and $L = L_i$.
for each $p_1 \in P$ and $p_2 \in P$
  let L_1, L_2, L_3, L_4 be the lines tangent to both $C_{a+\epsilon}(p_1)$ and $C_{a+\epsilon}(p_2)$
  for (i=1 to 4) if (balance$(L_i, Q) \leq \frac{2}{3}$ and (measure$(L_i, P, a + \epsilon, w) < B$))
    then $B =$ measure$(L_i, P, a + \epsilon, w)$ and $L = L_i$.
Output L

End of the Algorithm

∎

5. Application of width bounded separator

In this section we apply our geometric separator to the well-known disk covering problem: Given a set of points on the plane, find the minimal number of discs with fixed radius to cover all of those points. Hochbaum and Maass [22] showed it has polynomial time approximation scheme, which was improved to $n^{O(l)}$-time with approximation ratio $(1 + \frac{1}{l})$ by Feder, Greene [13] and Gonzalez [19]. The d-dimensional ball covering problem is to cover n points on the d-dimensional Euclid space with minimal number of d-dimensional ball of fixed radius. This problem arises in the area of locating emergency facilities such that all customers are within a reasonable radius around the facility, as well as in the area of wireless computing, where a wireless network is modeled as a set of discs to cover a set of users.

Theorem 21. *There is a $2^{O(\sqrt{n})}$-time exact algorithm for the disk covering problem on the 2D plane.*

Proof: Assume Q is the set of n input points on the plane. Let's set up an $(1,1)$-grid regular partition. For an grid point $p = (i, j)$ (i and j are integers) on the plane, define grid$(p) = \{(x, y) | i - \frac{1}{2} \leq x < i + \frac{1}{2}, j - \frac{1}{2} < y \leq j + \frac{1}{2}\}$, which is a half close and half open 1×1 square. There is no intersection between grid(p) and grid(q) for two different grid points p and q. Assume the diameter of disk is 1. Our "local binding" method is to merge the points of $Q \cap$ grid(p) to the grid point p and assign certain weight to p to measure the Q points density in grid(p). Partition the point set Q into $Q(p_1), \cdots, Q(p_m)$ as follows

Partition(Q)

$m = 0$
Repeat
  select a point $q \in Q - \cup_{i=1}^{m} Q(p_i)$
  let p_{m+1} be the grid point with $q \in$ grid(p_{m+1})
  let $Q(p_{m+1})$ be the set $Q \cap ($grid$(p_{m+1}))$.
  $m = m + 1$
until $Q = \cup_{i=1}^{m} Q(p_i)$
Output $P = \{p_1, \cdots, p_m\}$ and $Q(p_1), \cdots, Q(p_m)$.

End of Partition

Let n_i be the grid points $p_j \in P$ with $g^{i-1} \le |Q(p_j)| < g^i$, where g is a constant $g > 1$. From this definition, we have

$$\sum_{i=1}^{\lceil \log_g n \rceil} g^i \cdot n_i \le g \cdot n, \tag{14}$$

where $\lceil x \rceil$ is the least integer $\ge x$. Let $P = \{p_1, \cdots, p_m\}$ be the set grid points derived from partitioning set Q in the algorithm above. Define function $w : P \to \{1, 2, \cdots, \lceil \log_g n \rceil\}$ such that $w(p) = i$ if $g^{i-1} \le |Q(p)| < g^i$.

Select small $\delta > 0$ and $a = \frac{3}{2} + \frac{\sqrt{2}}{2}$. By Theorem 20 we can get a line L on the plane such that $balance(L, Q) \le \frac{2}{3}$ and $\text{measure}(L, P, a, w) \le (k_2 \cdot a + \delta)(\sum_{i=1}^{\lceil \log_g n \rceil} i \cdot \sqrt{n_i})$. Let $J(L) = \{p | p \in P \text{ and } \text{dist}(q, L) \le \frac{1}{2} \text{ for some } q \in Q(p)\}$. After those points of Q with distance $\le \frac{1}{2}$ to the separator line L are covered, the rest of points of Q on the different sides of L can be covered independently. Therefore, the covering problem is solved by divide and conquer method as described by the algorithm below.

Algorithm

Input a set of points Q on the plane.
run Partition(Q) to get $P = \{p_1, \cdots, p_m\}$ and $Q(p_1), \cdots, Q(p_m)$
find a separator line L for P, Q with (by Theorem 20)
    balance$(L, Q) \le \frac{2}{3}$ and measure$(L, P, a, w) \le (k_2 \cdot a + \delta) \sum_{i=1}^{\lceil \log n \rceil} i\sqrt{n_i}$
for each covering to the points in Q with $\le 1/2$ distance to L
    let $Q_1 \subseteq Q$ be the those points on the left of L and not covered
    let $Q_2 \subseteq Q$ be the those points on the right of L and not covered
    recursively cover Q_1 and Q_2
    merge the solutions from Q_1 and Q_2
Output the optimal solution (with the minimal number of discs covering all points)

End of Algorithm

For each grid area grid(p_i), the number of discs containing the points in $Q(p_i)$ is no more than the number of discs covering the 3×3 area, which needs no more than $c_3 = (\left\lceil \frac{3}{\frac{\sqrt{2}}{2}} \right\rceil)^2 = 25$ discs. Two grid points $p = (i, j)$ and $p' = (i', j')$ are *neighbors* if $\max(|i - i'|, |j - j'|) \le 1$. For each grid point p, define $m(p)$ to be the neighbor grid point q of p (q may be equal to p) with largest weight $w(q)$. For a grid point $p = (i, j)$, the 3×3 region $\{(x, y) | i - \frac{3}{2} \le x < i + \frac{3}{2} \textit{ and } j - \frac{3}{2} < y \le j + \frac{3}{2}\}$ has $< 9 \times g^{w(m(p))}$ points in Q. The number of ways to put one disc covering at least one point in $Q(p)$ is $\le (9 \times g^{w(m(p))})^2$ (let each disc have two points from Q on its boundary whenever it covers at least two points). The number of ways to arrange $\le c_3$ discs to cover points in $Q(p)$ is $\le (9 \times g^{w(m(p))})^{2c_3}$. The total number of cases to cover all points with distance $\le \frac{1}{2}$ to L in $\cup_{p \in J(L)} Q(p)$ is

$$\le \prod_{p \in J(L)} (9 \cdot g^{w(m(p))})^{2c_3} = \prod_{p \in J(L)} 2^{(\log_2 9 + w(m(p)) \cdot \log_2 g) 2c_3} \le \prod_{p \in J(L)} 2^{2c_3(\log_2 9 + \log_2 g) w(m(p))} \tag{15}$$

$$= 2^{2c_3(\log_2 9 + \log_2 g) \sum_{p \in J(L)} w(m(p))} \le 2^{2c_3(\log_2 9 + \log_2 g) 9 \cdot \text{measure}(L, P, a, w)} \tag{16}$$

$$\le 2^{2c_3(\log_2 9 + \log_2 g) 9 (k_2 \cdot a + \delta)(\sum_{i=1}^{\lceil \log n \rceil} i \cdot \sqrt{n_i})} \tag{17}$$

This is because that for each grid point q, there are at most 9 grid points p with $m(p) = q$. Furthermore, for each $p \in J(L)$, p has distance $\le \frac{1}{2} + \frac{\sqrt{2}}{2} = a$ to L and $m(p)$ has distance $\le \frac{3}{2} + \frac{\sqrt{2}}{2}$ to L. Let the exponent of (17) be represented by $u = 2c_3(\log_2 9 + \log_2 g) 9 (k_2 \cdot a + \delta)(\sum_{i=1}^{\lceil \log n \rceil} i \cdot \sqrt{n_i})$. By the well known inequality $(\sum_{i=1}^{m} a_i \cdot b_i)^2 \le (\sum_{i=1}^{m} a_i^2) \cdot (\sum_{i=1}^{m} b_i^2)$,

$$(\sum_{i=1}^{\lceil \log_g n \rceil} i\sqrt{n_i})^2 \le (\sum_{i=1}^{\lceil \log_g n \rceil} \frac{i}{g^{i/2}} \cdot g^{i/2}\sqrt{n_i})^2 \le (\sum_{i=1}^{\lceil \log_g n \rceil} \frac{i^2}{g^i}) \cdot (\sum_{i=1}^{\lceil \log_g n \rceil} g^i n_i) \tag{18}$$

Using the standard calculus (see appendix 8.1), $\sum_{i=1}^{\infty}\frac{i^2}{g^i}=\frac{g(g+1)}{(g-1)^3}$. By (18) and (14), $u\le e(g)\sqrt{n}$, where $e(g)=2c_3(\log_2 9+\log_2 g)(k_2\cdot a+\delta)\sqrt{\frac{g(g+1)}{(g-1)^3}}\cdot\sqrt{g}$. Let $T(n)$ be the maximal computational time of the algorithm for covering n points. The problem $T(n)$ is reduced to two problems $T(\frac{2}{3}n)$. We have $T(n)\le 2\cdot 2^{e(g)\sqrt{n}}T(\frac{2n}{3})\le 2^{\log_{3/2}n}2^{e(g)(1+\alpha+\alpha^2+\cdots)\sqrt{n}}=2^{e(g)(\frac{1}{1-\alpha})\sqrt{n}+\log_{3/2}n}=2^{O(\sqrt{n})}$, where $\alpha=\sqrt{\frac{2}{3}}$. ∎

Definition 22. We consider undirected graphs $G=(V,E)$, where V denotes the vertex set and E denotes the edge set. An *independent set* I of a graph $G=(V,E)$ is a set of pairwise nonadjacent vertices of a graph. An *vertex cover* C of a graph $G=(V,E)$ is a subset of vertices such that each edge in E has at least one end point in C. A *dominating set* D is a set of vertices such that the rest of the vertices in G has at least one neighbor in D. For a set of disks $D=\{C_{r_1}(p_1),C_{r_2}(p_2),\cdots,C_{r_n}(p_n)\}$, the *disk graph* is $G_D=(V_D,E_D)$, where vertices set $V_D=\{p_1,p_2,\cdots,p_n\}$ and $E_D=\{(p_i,p_j)|C_{r_i}(p_i)\cap C_{r_j}(p_j)\ne\emptyset\}$. DG is the class of all disk graphs. DG_σ is the class of all disk graphs G_D such that D is the set of disks $\{C_{r_1}(p_1),C_{r_2}(p_2),\cdots,C_{r_n}(p_n)\}$ with $\frac{\max_{i=1}^n r_i}{\min_{i=1}^n r_i}\le\sigma$.

Disk graphs have been used to model problems in several areas such as broadcast networks [20, 24], image processing [22] and VLSI design [29]. Several standard graph theoretic problems for GD_1 are NP-hard [9, 14, 30, 40]. The approximation schemes were developed for maximum independent set and and minimum vertex cover problems on GD_1 [23] and GD_σ [12]. The $n^{O(\sqrt{n})}$-time exact algorithm for the maximum independent set problem for DG_σ with constant σ was derived by Alber and Fiala [4] via parameterized approach, which was further simplified by Agarwal, Overmars and Sharir [1] for DG_1. We obtain $2^{O(\sqrt{n})}$-time algorithms for maximum independent set, minimum vertex cover, and minimum dominating set problems for DG_σ with constant σ. Their algorithms are similar each other. We only describe the algorithm for maximum independent set. We believe our method also works for many other problems on disk graph.

Theorem 23. *There is a $2^{O(\sigma\sqrt{n})}$ time algorithm for the maximum independent set problem for DG_σ.*

Proof: The algorithm is similar to that for the disk covering problem. We only describe the difference between them. Assume $G_D=(Q,E)$ is a disk graph such that D is a set of discs on the plane with diameters in the range $[1,\sigma]$. Choose $a=1.5\sigma+\frac{\sqrt{2}}{4}$. Let the plane form a $(\frac{\sqrt{2}}{2},\frac{\sqrt{2}}{2})$-grid regular partition. For each $(\frac{\sqrt{2}}{2},\frac{\sqrt{2}}{2})$-grid regular point $p=(u,v)$ ($u=\frac{i\sqrt{2}}{2}$ and $v=\frac{j\sqrt{2}}{2}$ for some integers i and j), define $\mathrm{grid}(p)=\{(x,y)|u-\frac{\sqrt{2}}{4}\le x<u+\frac{\sqrt{2}}{4},v-\frac{\sqrt{2}}{4}\le y<v+\frac{\sqrt{2}}{4}\}$. By Theorem 20, we can get a separator line L on the plane such that $balance(L,Q)\le\frac{2}{3}$ and $\mathrm{measure}(L,P,a,w)\le(\sqrt{2}\cdot k_2\cdot a+\delta)(\sum_{i=1}^{\lceil\log_g n\rceil}i\cdot\sqrt{n_i})$. Select all of the possible independent points in Q with distance $\le\sigma/2$ to the separator line L. Let $J(L)=\{p|p\in P$ and $\mathrm{dist}(q,L)\le\frac{\delta}{2}$ for some $q\in Q(p)\}$. For each $(\frac{\sqrt{2}}{2},\frac{\sqrt{2}}{2})$-grid regular point p, at most one point can be selected from $Q\cap\mathrm{grid}(p)$ to join the maximum independent set. The number of ways to select independent points with $\le\frac{\sigma}{2}$ distance to L is bounded by $\prod_{p\in J(L)}g^{w(p)}=g^{\sum_{p\in J(L)}w(p)}\le g^{(\sqrt{2}\cdot k_2\cdot a+\delta)(\sum_{i=1}^{\lceil\log_g n\rceil}i\cdot\sqrt{n_i})}=2^{O(\sqrt{n})}$. This gives that $T(n)=2^{O(\sigma\sqrt{n})}T(\frac{2n}{3})=2^{O(\sigma\sqrt{n})}$. ∎

6. Covering algorithm in higher dimension space

The methods above in sections 4 and 5 can be extended to high dimension with some slight adjustment.

Lemma 24. *Let constant $a,a_1,\cdots,a_d>0$ and small constant $\delta>0$. Let P be a set of $(a_1,\cdots,a_d)$-grid points and Q be another set of points on d-dimensional space. The weights $w_1>\cdots>w_k>0$ have $k\cdot\max_{i=1}^k\{w_i\}=o(n^\epsilon)$. There is an $O(n^{d+1})$ time algorithm that finds a separator such that* $\mathrm{balance}(L,Q)\le\frac{d-1}{d}$, *and* $\mathrm{measure}(L,P,a,w)\le\left(\frac{k_d}{(a_1\cdots a_d)^{1/d}}+\delta\right)a\sum_{i=1}^k w_i n_i^{\frac{d-1}{d}}+O(n^{\frac{d-2}{d}+\epsilon})$ *for all large n.*

Proof: (Sketch). For every integer pair $a, b \geq 0$ with $a + b = d$, select all possible a points $p_1, \cdots, p_a$ from P and all possible b points $q_1, \cdots, q_b$ from Q. Let the hyper-plane be through $q_1, \cdots, q_b$ and tangent to $B_d(a + \delta, p_i)$ $(i = 1, \cdots, a)$. ▮

Theorem 25. *There is a $2^{O(n^{1-1/d})}$-time algorithm for the ball covering problem in the d-dimensional space.*

Proof: (Sketch). By Theorem 10, the separator L has the properties: balance$(L, P) \leq \frac{d-1}{d}$, and measure$(L, Q, a, w) \leq (k_d \cdot a + \delta) \sum i \cdot n_i^{\frac{d-1}{d}}$. Use the well known Hölder inequality $(\sum_{i=1}^m a_i b_i) \leq (\sum_{i=1}^m a_i^k)^{\frac{1}{k}} (\sum_{i=1} b_i^{k'})^{\frac{1}{k'}}$, where $k, k' > 1$ and $\frac{1}{k} + \frac{1}{k'} = 1$. We have

$$\sum_{i=1}^{\lceil \log_g n \rceil} i \cdot n_i^{\frac{d-1}{d}} = \sum_{i=1}^{\lceil \log_g n \rceil} \left(\left(\frac{i}{g^{\frac{d-1}{d} i}} \right) \cdot (g^{\frac{d-1}{d} i} n^{\frac{d-1}{d}}) \right) \tag{19}$$

$$\leq \left(\sum_{i=1}^{\lceil \log_g n \rceil} \left(\frac{i}{g^{\frac{d-1}{d} i}} \right)^d \right)^{\frac{1}{d}} \cdot \left(\sum_{i=1}^{\lceil \log n \rceil} \left(g^{\frac{d-1}{d} i} \cdot n_i^{\frac{d-1}{d}} \right)^{\frac{d}{d-1}} \right)^{\frac{d-1}{d}} \tag{20}$$

$$= \left(\sum_{i=1}^{\lceil \log_g n \rceil} \left(\frac{i^d}{g^{(d-1)i}} \right) \right)^{\frac{1}{d}} \cdot \left(\sum_{i=1}^{\lceil \log n \rceil} (g^i n_i) \right)^{\frac{d-1}{d}} = O(n^{\frac{d-1}{d}}) \tag{21}$$

▮

7. Conclusions

We derive some width-bounded geometric separators that can be applied to some geometric problem with arbitrary input points like disk covering problem. We also proved the width-bounded separator for 1-separated points. It is believed that there will be more applications for this notion to other geometric problems. For example, the separator for 1-separated set may be useful tool for the protein folding problem in more general model other than the grid model [25], but we still do not have a suitable model for it.

References

[1] P. Agarwal, M. Overmars, and M. Sharir, Computing maximally separated sets in the plane and independent sets in the intersection graph of unit graphs, In proceedings of 15th ACM-SIAM symposium on discrete mathematics algorithms, ACM-SIAM 2004, pp. 509-518.

[2] J. Alber, H. Fernau, and R. Niedermeier, Graph separators: a parameterized view, In proceedings of 7th Internal computing and combinatorics conference 2001, pp. 318-327.

[3] J. Alber, H. Fernau, and R. Niedermeier, Parameterized complexity: exponential speed-up for planar graph problems, In Proceedings of 28st international colloquium on automata, languages and programming, 2001, pp. 261-272.

[4] J. Alber and J. Fiala, Geometric separation and exact solution for parameterized independent set problem on disk graphs, Journal of Algorithms, 52, 2, 2004, pp. 134-151.

[5] N. Alon, P. Seymour, and R. Thomas, Planar Separator, SIAM J. Discr. Math. 7,2(1990) 184-193.

[6] N. Alon, P. Seymour, and R. Thomas, A separator theorem for graphs with an excluded minor and its applications, In proceedings of the 22nd annual ACM symposium on theory of computing, ACM 1990, pp. 293-299.

[7] S. N. Bhatt and F. T. Leighton, A framework for solving VLSI graph layout problems, Journal of computer and systems science, 28,2(1982), pp. 300-343

[8] N. Chiba, T. Nishizeki and T. Saito, Applications of the Lipton and Tarjan planar separator theorem, Journal of information processing letters, 4,4(1981), 203-207.

[9] B. N. Clark, C. J. Colbourn, and D. S. Johnson, Unit disk graphs, Discrete mathematics, 86(1990), pp. 165-177.

[10] H. N. Djidjev, On the problem of partitioning planar graphs. SIAM journal on discrete mathematics, 3(2) June, 1982, pp. 229-240.

[11] H. N. Djidjev and S. M. Venkatesan, Reduced constants for simple cycle graph separation, Acta informatica, 34(1997), pp. 231-234.

[12] T. Erlebach, K. Jansen, and E. Seidel, Polynomial-time approximation schemes for geometric graphs, In proceedings of 12th ACM-SIAM symposium on discrete mathematics algorithms, ACM-SIAM 2001, pp.671-679.

[13] T. Feder and D. Greene, Optimal algorithms for approximate clustering, Proceedings of the 20th annual ACM symposium on the theory of computing, 1988, pp. 434-444.

[14] R. J. Fowler, M. S. Paterson, and S. L. Tanimoto, Optimal packing and covering in the plane are NP-complete, Information processing letters, 3(12), 1981, pp. 133-137.

[15] B. Fu and W. Wang, A $2^{O(n^{1-1/d}\log n)}$-time algorithm for d-dimensional protein folding in the HP-model, Proceedings of 31st international colloquium on automata, languages and programming, 2004, pp. 630-644.

[16] H. Gazit, An improved algorithm for separating a planar graph, manuscript, USC, 1986.

[17] J. R. Gilbert, J. P. Hutchinson, and R. E. Tarjan, A separation theorem for graphs of bounded genus, Journal of algorithm, (5)1984, pp. 391-407.

[18] R. Graham, M. Grötschel, and L. Lovász, Handbook of combinatorics (volume I), MIT Press, 1996

[19] T. Gonzalez, Covering a set of points in the multidimensional space, Information processing letters, 4(40) 1991, pp. 181-188.

[20] W. K. Hale, Frequency assignment: theory and applications, Proceedings of IEEE, vol. 68, 1980, pp. 1497-1514.

[21] T. C. Hales, A computer verification of the Kepler conjecture, Proceedings of the ICM, Beijing 2002, vol. 3, 795–804

[22] D. S. Hochbaum, and W. Maass, Approximation shcemes for covering and packing problems in image processing and VLSI, Journal of the ACM,1(32), 1985, pp. 130-136.

[23] H. B. Hunt, M. V. Marathe, V, Radhakrishnan, V. Radhakrishnan, S. S. Ravi, D. J. Rosenkrantz, and R. E. Stearns, NC-approximation schemes for NP-and PSPACE-hard problems for geometric graphs, Journal of algorithms, 26(1998), pp. 238-274.

[24] K. Kammerlander, C900- An advanced mobile radio telephone system with optimum frequency utilization, IEEE trans. selected areas in communication, v. 2, 1984, pp. 589-597.

[25] K. F. Lau and K. A. Dill, A lattice statistical mechanics model of the conformational and sequence spaces of proteins, Macromolecules, 22(1989), 3986-3997.

[26] D. Lichtenstein, Planar formula and their uses, SIAM journal on computing, 11,2(1982), pp. 329-343.

[27] R. J. Lipton and R. Tarjan, A separator theorem for planar graph, SIAM Journal on Applied Mathematics, 36(1979) 177-189.

[28] R. J. Lipton and R. Tarjan, Applications of a planar separator theorem, SIAM journal on computing, 9,3(1980), pp. 615-627.

[29] C. A. Mead and L. Conway, Introduction to VLSI systems, Addison-Wesley, Reading, MA, 1980.

[30] N. Meggido and K. Supowit, On the complexity of some common geometric location problems, SIAM journal on computing, v. 13, 1984, pp. 1-29.

[31] G. L. Miller, S.-H. Teng, and S. A. Vavasis, An unified geometric approach to graph separators. In 32nd annual symposium on foundation of computer science, IEEE 1991, pp. 538-547.

[32] G. L. Miller and W. Thurston, Separators in two and three dimensions, In 22nd Annual ACM symposium on theory of computing, ACM 1990, pp. 300-309.

[33] G. L. Miller and S. A. Vavasis, Density graphs and separators, In second annual ACM-SIAM symposium on discrete algorithms, ACM-SIAM 1991, pp. 331-336.

[34] D. J. Muder, A new bound on the local density of sphere packing, Discrete and comp. geom. 10(1993), 351-375.

[35] S. Plotkin, S. Rao, and W. D. Smith, Shallow excluded minors and improved graph decomposition, In 5th symp. on discrete algorithms, SIAM 1990, pp. 462-470.

[36] S. S. Ravi, H. B. Hunt III, Application of the planar separator theorem to computing problems, Information processing letter, 25,5(1987), pp. 317-322.

[37] W. D. Smith and N. C. Wormald, Application of geometric separator theorems, The 39th annual symposium on foundations of computer science,1998, 232-243.

[38] D. A. Spielman and S. H. Teng, Disk packings and planar separators, The 12th annual ACM symposium on computational geometry, 1996, pp.349-358.

[39] J. Pach and P. K. Agarwal, Combinatorial geometry, Wiley-Interscience Publication, 1995.

[40] D. W. Wong and Y. S. Kuo, A study of two geometric location problems, Information processing letters, v. 28, 1988, pp. 281-286.

8. Appendix

8.1. Sum of infinite sequences

Let x be a variable with $|x| < 1$. $f(x) = \sum_{i=1}^{\infty} x^i = \frac{x}{1-x}$. $f(x)' = \sum_{i=1}^{\infty} ix^{i-1} = \frac{1}{(1-x)^2}$. $x(f(x)') = \sum_{i=1}^{\infty} ix^i = \frac{x}{(1-x)^2}$. $(x(f(x)'))' = \sum_{i=1}^{\infty} i^2x^{i-1} = \frac{1+x}{(1-x)^3}$. $x(x(f(x)'))' = \sum_{i=1}^{\infty} i^2x^i = \frac{x(1+x)}{(1-x)^3}$. This shows that for $g > 1$, $\sum_{i=1}^{\infty} \frac{i^2}{g^i} = x(x(f(\frac{1}{g})'))' = \frac{g(g+1)}{(g-1)^3}$.

ECCC ISSN 1433-8092
http://www.eccc.uni-trier.de/eccc
ftp://ftp.eccc.uni-trier.de/pub/eccc
ftpmail@ftp.eccc.uni-trier.de, subject 'help eccc'