Halving Lines and Measure Concentration in the Plane

Rom Pinchasi*

November 29, 2008

Abstract

Given a set of n points in the plane and a collection of k halving lines of P $\ell_1, \ldots, \ell_k$ indexed according to the increasing order of their slopes, we denote by $d(\ell_j, \ell_{j+1})$ the number of points in P that lie above ℓ_{j+1} and below ℓ_j. We prove an upper bound of $3nk^{1/3}$ for the sum $\sum_{j=1}^{k-1} d(\ell_j, \ell_{j+1})$. We show how this problem is related to the halving lines problem and provide several consequences about measure concentration in $\mathbb{R}^2$.

1 Introduction

We will be motivated by the following theorem from [BPZ08] about measure concentration in the plane:

Theorem 1 ([BPZ08])**.** *For every $\epsilon > 0$ there exists $\alpha(\epsilon) > 0$ such that for every continuous probability measure μ in the plane one can find two lines ℓ_1 and ℓ_2 that meet at an angle of $\alpha(\epsilon)$ such that the measure of each of the two quadrants determined by ℓ_1 and ℓ_2 of angle $\pi - \alpha(\epsilon)$ is at least $\frac{1}{2} - \epsilon$.*

Let $f(\epsilon)$ denote the maximum possible value of $\alpha(\epsilon)$ in the statement of Theorem 1. The result in [BPZ08] implies $f(\epsilon) = \Omega(\epsilon^3)$. In the sequel we will improve on this bound and relate $f(\epsilon)$ to several problems, concerned with halving lines, where we obtain further results.

Problem 1. Find good lower bounds to $f(\epsilon)$ in terms of ϵ.

Problem 1 is closely related to the following problem:

Definition 1. Given two lines ℓ_1 and ℓ_2 in the plane, we denote by $Wedge(\ell_1, \ell_2)$ the region which consists of all points that lie above ℓ_2 and below ℓ_1 (see Figure 1).

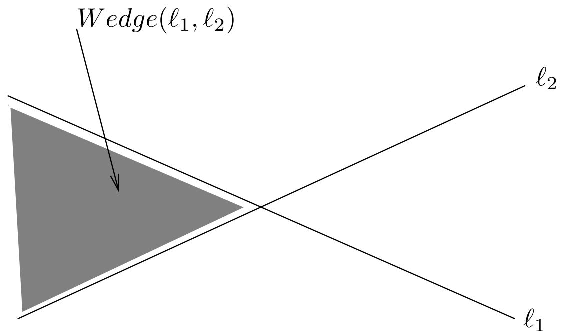

Figure 1: $Wedge(\ell_1, \ell_2)$.

Definition 2. We denote by $q(k)$ the minimum number with the following property. Let μ be any given continuous probability measure in the plane and let $\ell_1, \ldots, \ell_k$ be a collection of k halving lines for μ, indexed according to the increasing order of their slopes. Then there is $0 < j < k$ such that the measure of $Wedge(\ell_j, \ell_{j+1})$ is at most $q(k)$.

[1]This research was supported by a Grant from the Israel Science Foundation (grant No. 938/06)
*Mathematics Dept., Technion—Israel Institute of Technology, Haifa 32000, Israel. room@math.technion.ac.il.

Problem 2. Find good upper bounds for $q(k)$ in terms of k.

The simple relation between Problem 1 and Problem 2 is indicated in the the next proposition.

Proposition 1. $f(q(k)) \geq \frac{\pi}{k}$.

Proof. Let μ be any given continuous probability measure in the plane. For every $j = 1, \ldots, k-1$ let ℓ_j be the halving line for μ in the direction of the vector $(\cos(\frac{j\pi}{k} - \frac{\pi}{2}), \sin(\frac{j\pi}{k} - \frac{\pi}{2}))$. By the definition of $q(k)$ there is an index i between 1 and $k-2$ such that the measure of $Wedge(\ell_i, \ell_{i+1})$ is at most $q(k)$. Because both ℓ_i and ℓ_{i+1} are halving lines for μ, then ℓ_i and ℓ_{i+1} are two lines that meet at an angle of $\frac{\pi}{k}$ and the measure of each of the quadrants determined by ℓ_i and ℓ_{i+1} of angle $\pi - \frac{\pi}{k}$ is greater than or equal to $\frac{1}{2} - q(k)$. ∎

In order to investigate the functions $f(\epsilon)$ and $q(k)$ we will need to consider a discrete analogue of the function $q(k)$.

Throughout this paper P will denote a fixed set of n points in general position in the plane, where n is an even number. A line ℓ is called a halving line of P if ℓ does not pass through any point of P and it divides the set P into two parts each has cardinality $n/2$. We will assume that no two points of P have the same x-coordinate.

We denote by $K(P)$ the complete geometric graph whose set of vertices is P. That is, $K(P)$ consists of vertices that are the points in P and every two points in P are connected by a straight line segment. For an edge $e = (p, q)$ in $K(P)$, we call p the *left endpoint* of e if the x-coordinate of p is smaller than that of q. In this case q is called the *right endpoint* of the edge e. An edge $e = (p, q)$ of $K(P)$ is called a *halving edge* of P if the line through p and q divides the set $P \setminus \{p, q\}$ into two parts, each with cardinality $n/2 - 1$. We denote the geometric graph whose vertices are the points of P and whose edges are the halving edges of P by $G(P)$.

For a (non-vertical) line ℓ not passing through any point of P we denote by $B(\ell)$ the set of all points of P that lie below ℓ. We say that two lines ℓ_1 and ℓ_2 that do not pass through any point of P are equivalent if $B(\ell_1) = B(\ell_2)$.

It is readily seen that if ℓ_1 and ℓ_2 are two halving lines for P that are equivalent, then every halving lines ℓ whose slope lies between the slopes of ℓ_1 and ℓ_2 is also equivalent to ℓ_1 and ℓ_2. This follows immediately if one considers the arrangement of lines dual to the set of points in P. Then the set of all halving lines for P which are equivalent forms a face in that arrangement. This allows us to define equivalence classes of halving lines for the set P and order them according to the slopes of representatives of these equivalence classes.

Let $h_1, \ldots, h_m$ be representatives of all equivalence classes of halving lines for P ordered according to the increasing order of their slopes. For $i = 1, \ldots, m$ we denote by $[h_i]$ the equivalence class of the line h_i.

Definition 3. For a halving line ℓ of P we denote by $s(\ell)$ the index of the equivalence class to which ℓ belongs. That is, $s(\ell) = i$ iff $\ell \in [h_i]$.

Definition 4. Let ℓ_1 and ℓ_2 be two halving lines of P such that the slope of ℓ_1 is smaller than the slope of ℓ_2. We denote by $d(\ell_1, \ell_2)$ the number of points in P that lie below ℓ_1 and above ℓ_2. (See Figure 2.)

Observe that according to this definition, if ℓ_1 and ℓ_2 are two halving lines such that $d(\ell_1, \ell_2) = 0$ then ℓ_1 and ℓ_2 are equivalent, that is, $s(\ell_1) = s(\ell_2)$.

Definition 5. We denote by $g(n, k)$ the minimum number such that for any set P of n points in the plane and any collection of k halving lines of P $\ell_1, \ldots, \ell_k$ indexed according to the increasing order of their slopes, we have $\sum_{j=1}^{k-1} d(\ell_j, \ell_{j+1}) \leq g(n, k)$. (See Figure 2 to have a picture of what is going on.)

Problem 3. Find good upper bound for $g(n, k)$ in terms of n and k.

In Section 3, we prove our main theorem which gives a bound for $g(n, k)$:

Theorem 2. *Let P be a set of n points in the plane and assume that n is even. Let $\ell_1, \ldots, \ell_k$ be k halving lines for the set P, indexed according to the increasing order of their slopes. Then $\sum_{j=1}^{k-1} d(\ell_j, \ell_{j+1}) \leq 3nk^{1/3}$.*

One can consider also the continuous analogue of the function $g(n, k)$ that we will denote by $g(k)$ (the parameter n will not play a role in the definition):

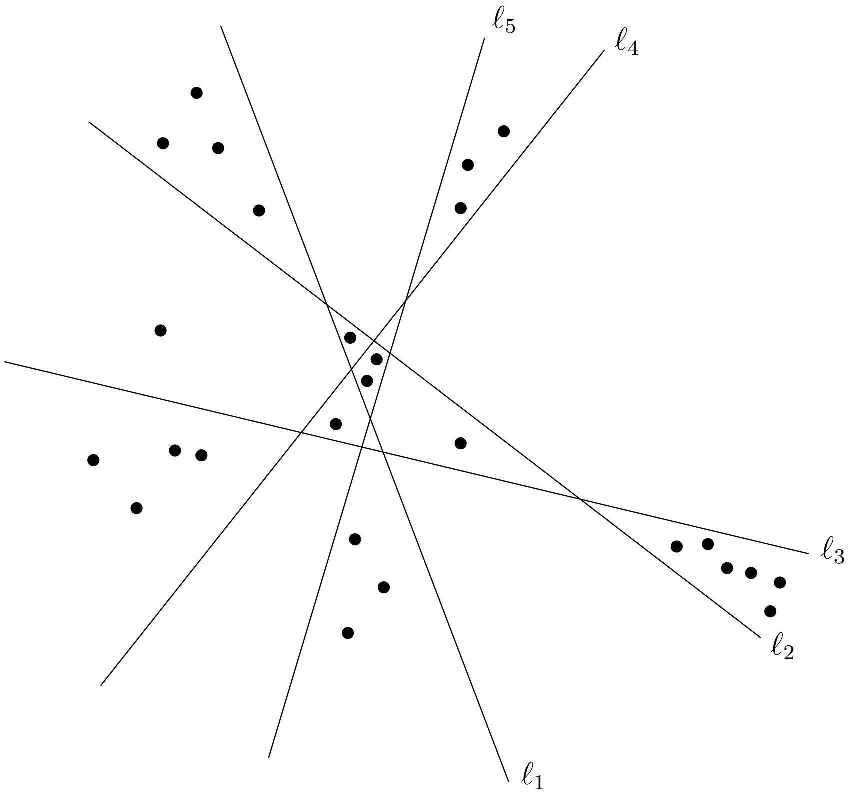

Figure 2: $d(\ell_1, \ell_2) = 4$, $d(\ell_2, \ell_3) = 6$, $d(\ell_3, \ell_4) = 4$, $d(\ell_4, \ell_5) = 3$.

Definition 6. We denote by $g(k)$ the minimum number such that for every continuous probability measure μ in the plane and every k halving lines for μ, $\ell_1, \ldots, \ell_k$, indexed according to the increasing order of their slopes, we have $\sum_{j=1}^{k-1} \mu(Wedge(\ell_j, \ell_{j+1})) \leq g(k)$.

The function $g(k)$ is directly related to $q(k)$ by observing that $q(k) \leq g(k)/(k-1)$ as follows directly from the definition of these two functions. $g(k)$ is also directly related to the function $g(n, k)$ as follows:

Claim 1. *1. For every even number n, $g(k) \geq g(n, k)/n$.*

2. $g(k) \leq \sup_n \{g(n, k)/n\}$.

Proof. Both parts follow quite easily. The first part is immediate. Assume that there is a construction of a set P of n points in the plane and a set of k halving lines of P $\ell_1, \ldots, \ell_k$, indexed according to the increasing order of their slopes, such that $\sum_{j=1}^{k-1} d(\ell_j, \ell_{j+1}) = g(n, k)$. Construct a probability measure μ by taking a small ball around each point of P with the uniform measure of $\frac{1}{n}$. If the balls are small enough so that no line among $\ell_1, \ldots, \ell_k$ crosses any of the balls, then

$$\sum_{j=1}^{k-1} \mu(Wedge(\ell_j, \ell_{j+1})) = \frac{1}{n} \sum_{j=1}^{k-1} d(\ell_j, \ell_{j+1}) = g(n, k)/n.$$

This proves that $g(k) \geq g(n, k)/n$.

To see the second part of the claim, define $z = \sup_n \{g(n, k)/n\}$. Let μ be a given continuous probability measure in the plane with halving lines $\ell_1, \ldots, \ell_k$, indexed according to the increasing order of their slopes. Consider the line arrangement determined by $\ell_1, \ldots, \ell_k$ and assume first that the measure μ of each face in that arrangement is a rational number.

Multiply the measure μ by a large enough integer B so that the measure of each face multiplied by B is an integer. Now position in each face F of the arrangement $\mu(F)B$ points. The resulting set of points consists of B points and $\ell_1, \ldots, \ell_k$ are halving lines for these set of points. It follows now that $g(k) \leq g(B, k)/B \leq z$.

If the measure μ of some of the faces is irrational, then one can approximate them by close enough rational numbers and use continuity arguments to conclude the theorem also in this case. ∎

In Section 5, we show some direct relations between Problems 1,2, and 3 and the 'Halving Lines' problem. As a consequence we deduce some nontrivial lower bounds for $g(n,k)$ and $g(k)$.

2 A first improvement for the upper bound on $f(\epsilon)$

In order to provide a first improvement for the lower bound of $f(\epsilon)$ we will need the following lemma.

Lemma 1. *Let P be a set of n points in the plane and let L be a set of weighted lines with a weight function $w : L \rightarrow \mathbb{R}^+$. Assume that no line in L passes through a point of P. For every two points $a, b \in P$ let $L_{a,b}$ denote the set of lines in L that separate a from b. If $\sum_{\ell \in L_{a,b}} w(\ell) \geq 1$ for every two distinct points $a, b \in P$, then $\sum_{\ell \in L} w(\ell) = \Omega(\sqrt{n})$.*

Proof. We use the following result from [C88, W92, M91] about spanning trees with low stabbing number: Given n points in the plane one on always construct a geometric spanning tree on the set of n points with the property that every line crosses $O(\sqrt{n})$ edges of the tree.

Construct such a tree T on the set of points of P and let $\ell_1, \ldots, \ell_m$ be all the lines in L that cross an edge of the tree. On one hand each of the lines $\ell_1, \ldots, \ell_m$ crosses only $O(\sqrt{n})$ edges of the tree while on the other hand each edge, and there are $n-1$ such edges, is crossed by lines whose total weight is at least 1. It follows that the total sum of the weights of $\ell_1, \ldots, \ell_m$ is at least $\Omega(\sqrt{n})$. ∎

The bound of $\Omega(\sqrt{n})$ cannot be improved in Lemma 1. To see this consider an arrangement L of $\sqrt{n}$ lines in general position in the plane. They determine roughly $n/2$ cells. Position a point in each of these cells and observe that every two such points are separated by a line in L. Now give every line in L a weight of 1.

Theorem 3. $q(k) = O(\frac{1}{\sqrt{k}})$.

Proof. Let μ be a continuous probability measure in the plane. Let $\ell_1, \ldots, \ell_k$ be k halving lines with respect to μ arranged according to the increasing order of their slopes. Observe that if ℓ and ℓ' are two halving lines with respect to the measure μ, then the measure of every two opposite wedges determined by ℓ and ℓ' is equal. Let $\epsilon = min_{1 \leq i < k} \mu(Wedge(\ell_i, \ell_{i+1}))$. Consider the faces in the arrangement determined by the lines $\ell_1, \ldots, \ell_k$ and position a point in each face with a weight that equals the measure μ of the region bounded by that face. Let P denote the set of all points thus defined and consider the dual plane. We get a set P of weighted lines with a total weight of 1 and a set L of k points. Observe that every two points in L are separated by lines with a total weight of at least ϵ. Hence, by Lemma 1 the total weight of all lines is $\Omega(\epsilon\sqrt{k})$. This implies that $\epsilon = O(\frac{1}{\sqrt{k}})$. ∎

Observe that Theorem 3 together with Proposition 1 immediately give an improved bound for $f(\epsilon)$. From Theorem 3 we know that there exists an absolute constant $c > 0$ such that $q(k) \leq \frac{c}{\sqrt{k}}$. Using Proposition 1 and the observation that $f(\epsilon)$ is monotone increasing in ϵ, we get:

$$f(\frac{c}{\sqrt{k}}) \geq f(q(k)) \geq \frac{\pi}{k}.$$

It follows now that for every $\epsilon > 0$ we have $f(\epsilon) = \Omega(\epsilon^2)$.

3 Proof of Theorem 2 and improved bounds for $f(\epsilon)$ and $q(k)$

We start with several propositions about halving lines and halving edges. Some of these propositions are of independent interest.

Proposition 2. *Let ℓ_1 and ℓ_2 be two halving lines for P such that $s(\ell_1) < s(\ell_2)$. Let $x \in P$ be any point that lies below ℓ_1 and above ℓ_2, then there exists a halving edge e of $G(P)$ such that the slope of e lies between the slopes of ℓ_1 and ℓ_2 and x is the left endpoint of e.*

Proof. Take a line ℓ' through x that is parallel to ℓ_1. Because ℓ_1 is a halving line for P and p lies below ℓ_1 it must be that there are less than $n/2$ points of p below ℓ'. We rotate ℓ' in the counterclockwise direction about x and keep track of the number of points of P that lie below ℓ'. By the time where ℓ' is parallel to ℓ_2 there are at least $n/2$ points of P below ℓ'. Therefore, there must be a time where there are $n/2-1$ points below ℓ' and immediately afterwards there are $n/2$ points below ℓ'. This implies that there must be a time where ℓ' has $n/2-1$ points of P below it and it passes through another point $y \in P$ that lies to the right of p. $e=(x,y)$ is the desired halving edge. (See Figure 3.) ∎

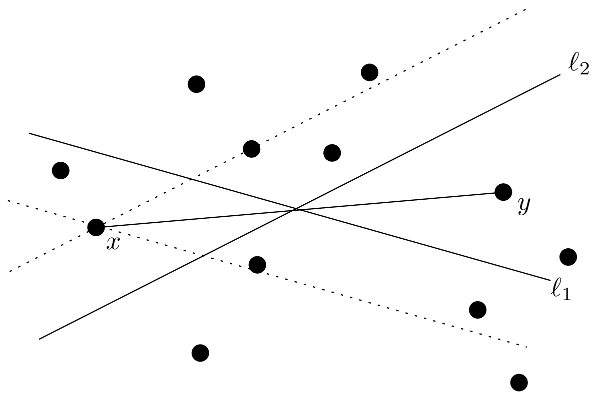

Figure 3: Proposition 2

Analogously to Proposition 2 we also have:

Proposition 3. *Let ℓ_1 and ℓ_2 be two halving lines for P such that $s(\ell_1) < s(\ell_2)$. Let $y \in P$ be any point that lies above ℓ_1 and below ℓ_2, then there exists a halving edge e of $G(P)$ such that the slope of e lies between the slopes of ℓ_1 and ℓ_2 and y is the right endpoint of e.*

Corollary 1. *Let ℓ_1 and ℓ_2 be two halving lines of P. If $s(\ell_2) = s(\ell_1)+1$, then there exists a unique point x that lies below ℓ_1 and above ℓ_2, and there exists a unique point y that lies above ℓ_1 and below ℓ_2. Moreover, $e=(x,y)$ is a halving edge of P.*

Proof. Because $s(\ell_2) = s(\ell_1)+1$, there can be at most one halving edge in $G(P)$ whose slope lies between the slopes of ℓ_1 and ℓ_2. The corollary now follows immediately from Proposition 2 and Proposition 3 ∎

Another immediate corollary from Proposition 2 is that for every two halving lines ℓ_1 and ℓ_2 of P $|s(\ell_2)-s(\ell_1)|$ is always greater than or equal to $d(\ell_1,\ell_2)$. This is because by Proposition 2, there are at least $d(\ell_1,\ell_2)$ halving edges whose slopes lie between the slopes of ℓ_1 and ℓ_2.

The next lemma will be very important for us and is one of our main tools in this paper.

Lemma 2. *Let ℓ_1 and ℓ_2 be two halving lines of P such that $d(\ell_1,\ell_2)=d>0$. Let $a_1,\dots,a_d$ be the points in P that lie below ℓ_1 and above ℓ_2. Let $b_1,\dots,b_d$ be the points in P that lie above ℓ_1 and below ℓ_2. Then there exists a permutation π on $\{1,\dots,d\}$ and d pairwise edge-disjoint x-monotone paths in $G(P)$ connecting a_i to $b_{\pi(i)}$ for $i=1,\dots,d$. Moreover, all these paths are composed only from edges whose slopes lie between the slopes of ℓ_1 and ℓ_2.*

Proof. Without loss of generality, assume that the slope of ℓ_1 is smaller than the slope of ℓ_2. We prove the proposition by induction on $s(\ell_2)-s(\ell_1)$. If $s(\ell_2)=s(\ell_1)+1$ the conclusion follows from Corollary 1.

If $s(\ell_2) > s(\ell_1)+1$, let ℓ be a halving line of P with $s(\ell_1) < s(\ell) < s(\ell_2)$.

Case 1. ℓ passes above the intersection point of ℓ_1 and ℓ_2. Observe that there must be points among $\{a_1,\dots,a_d\}$ that lie above ℓ, for otherwise $d(\ell_1,\ell)=0$ implying that $s(\ell_1)=s(\ell)$. Without loss of generality assume that $a_1,\dots,a_r$ lie above ℓ while $a_{r+1},\dots,a_d$ lie below ℓ, for some fixed integer r such that $0<r\le d$.

Let t denote the number of points of P that lie above both ℓ_1 and ℓ_2, but below ℓ. We denote these points by $z_1,\dots,z_t$. There are $r-t$ points among $b_1,\dots,b_d$ that lie above ℓ_1 and below ℓ while the other $d-r+t$ points lie below ℓ_2 and above ℓ. Without loss of generality assume that $b_1,\dots,b_{r-t}$ are those points that lie above ℓ_1 and below ℓ. (See Figure 4.)

By the induction hypothesis, applied for the lines ℓ_1 and ℓ, there are r edge-disjoint x-monotone paths in $G(P)$ connecting each of $a_1, \ldots, a_r$ to a unique element among $z_1, \ldots, z_t, b_1, \ldots, b_{r-t}$. Moreover, the slope of every edge involved in these paths lies between the slopes of ℓ_1 and ℓ.

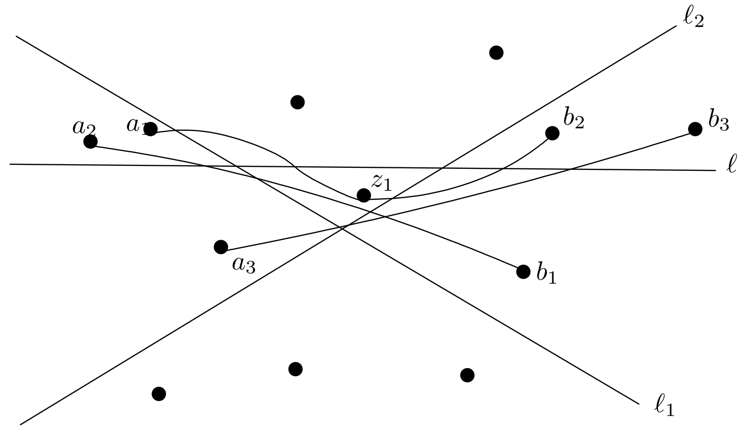

Figure 4: Lemma 2: $d = 3$, $t = 1$, $r = 2$.

Again by the induction hypothesis, applied this time to the lines ℓ and ℓ_2, there are $d - r$ edge-disjoint x-monotone paths in $G(P)$ connecting each of $a_{r+1}, \ldots, a_d, z_1, \ldots z_t$ to a unique element among $b_{r-t+1}, \ldots, b_d$. Moreover, the slope of every edge involved in these paths lies between the slopes of ℓ and ℓ_2.

Using these paths and compositions of these paths along the points $z_1, \ldots, z_t$, we obtain the desired result (see Figure 4).

Case 2. ℓ passes below the intersection point of ℓ_1 and ℓ_2. This case is very similar to Case 1. In fact it can be concluded from Case 1 by reflecting the plane about the x-axis. ▮

For the proof of Theorem 2 we will need the following definition:

Recall that $p_1, \ldots, p_n$ denote the points of P indexed according to the increasing order of their x-coordinates.

Definition 7. Let $e = (p_i p_j)$ be an edge in $K(P)$. We define the *x-length* of e as $|j - i|$.

Observe that the sum of the x-lengths of all edges in $G(P)$ equals to $n^2/4$. Indeed, this is a simple consequence to the fact that if ℓ is a line which divides the points of P into a set of k points and a set of $n - k$ points, then ℓ intersects exactly $\min(k, n - k)$ edges from $G(P)$ (see for example [L71]). Now, for $i = 1, \ldots, n - 1$ let W_i be a vertical line that is at equal distance from p_i and p_{i+1}. The observation follows now by noticing that the x-length of an edge (in $G(P)$) is equal to the number of lines from $W_1, \ldots, W_{n-1}$ that it crosses.

Proof of Theorem 2. Fix an index $1 \leq j < k$ and let $d_j = d(\ell_j, \ell_{j+1})$. Let $p_{i_{j,1}}, \ldots, p_{i_{j,d_j}}$ be the points of P that lie above ℓ_j and below ℓ_{j+1}. Because both ℓ_j and ℓ_{j+1} are halving lines for the set P, there are also d_j points that lie above ℓ_j and below ℓ_{j+1}. Let $p_{i'_{j,1}}, \ldots, p_{i'_{j,d_j}}$ denote those points. In our notation we assume that $i_{j,1} < \ldots < i_{j,d_j}$ and $i'_{j,1} < \ldots < i'_{j,d_j}$. Observe that we must have $i_{j,d_j} < i'_{j,1}$ because the slope of ℓ_j is smaller than the slope of ℓ_{j+1}. Define $z_j = i'_{j,\lceil d_j/2 \rceil} - i_{j,\lceil d_j/2 \rceil}$.

By Lemma 2, in $G(P)$ one can find edge-disjoint paths connecting each of $p_{i_{j,1}}, \ldots, p_{i_{j,d_j}}$ to a unique point among $p_{i'_{j,1}}, \ldots, p_{i'_{j,d_j}}$. Moreover, the collection of all edges in $G(P)$ that comprise these paths have their slopes laying between the slope of ℓ_j and the slope of ℓ_{j+1}. Observe that the sum of all x-lengths of these edges is equal to $\sum_{s=1}^{d_j} i'_{j,s} - \sum_{s=1}^{d_j} i_{j,s}$, and this difference, in turn, is greater than or equal to $z_j \lceil d_j/2 \rceil$.

Having defined $z_1, \ldots, z_{k-1}$, fix an integer threshold t, to be determined later, and let $J = \{1 \leq j \leq k - 1 \mid z_j \geq t\}$.

Observe that if $j \in J$, then there total sum of x-lengths of all edges in $G(P)$ whose slopes lie between the slope of ℓ_j and the slope of ℓ_{j+1}, is at least $t \lceil d_j/2 \rceil$. Because the sum of x-lengths of all edges in $G(P)$ is equal to $n^2/4$, we have that $\sum_{j \in J} t \lceil d_j/2 \rceil \leq n^2/4$. From here we deduce:

$$\sum_{j\in J} d_j \leq \frac{n^2}{2t}. \tag{1}$$

For every j and every $1 \leq s, s' \leq d_j$ the line through $p_{i_{j,s}}$ and $p_{i'_{j,s'}}$ has slope greater than that of ℓ_j and smaller than that of ℓ_{j+1}. Therefore, every $j \notin J$ gives rise to different $(\lceil d_j/2 \rceil)^2$ edges in $K(P)$ each has x-length at most t. However, the number of edges in $K(P)$ with x-length at most t is clearly smaller than tn. We thus have:

$$\sum_{j\notin J} (\lceil d_j/2 \rceil)^2 \leq tn. \tag{2}$$

From (2) it follows that

$$\sum_{j\notin J} d_j^2 \leq 4tn. \tag{3}$$

From (3) and the Cauchy-Schwartz inequality we obtain:

$$(\sum_{j\notin J} d_j)^2 \leq k \sum_{j\notin J} d_j^2 \leq 4tkn. \tag{4}$$

(4) now gives:

$$\sum_{j\notin J} d_j \leq 2\sqrt{tkn}. \tag{5}$$

Combining (1) and (5) we get:

$$\sum_{j=0}^{k-1} d_j = \sum_{j\in J} d_j + \sum_{j\notin J} d_j \leq \frac{n^2}{2t} + 2\sqrt{tkn}. \tag{6}$$

Finally, taking $t = \frac{n}{2^{2/3}k^{1/3}}$ in (6) we obtain the bound $\sum_{j=0}^{k-1} d_j \leq 3nk^{1/3}$. ∎

We can now obtain an improved bound for $f(\epsilon)$ and $q(k)$:

Theorem 4. *$f(\epsilon) = \Omega(\epsilon^{3/2})$ and $q(k) \leq \frac{6}{k^{2/3}}$.*

Proof. It follows from Claim 1 and Theorem 2 that $g(k) \leq 3k^{1/3}$. Let μ be a given continuous probability measure in the plane and let $\ell_1, \ldots, \ell_k$ be k halving lines for μ, indexed according to the increasing order of their slopes. It follows now from the definition of the number $g(k)$ and from the bound on $g(k)$ above that there exists an index $0 < i < k$ such that $\mu(wedge(\ell_i, \ell_{i+1})) \leq g(k)/(k-1) \leq \frac{6}{k^{2/3}}$. Therefore, by the definition of the function $q(k)$ we get, $q(k) \leq \frac{6}{k^{2/3}}$.

Using the monotonicity of $f(\epsilon)$ and Proposition 1, we have $f(\frac{6}{k^{2/3}}) \leq f(q(k)) \leq \frac{\pi}{k}$. It follows now from the monotonicity of f that $f(\epsilon) = \Omega(\epsilon^{3/2})$. ∎

4 A weaker (but tight) variant of Theorem 2

The following theorem can be considered as a variant of Theorem 2. It studies the case where $\ell_1, \ldots, \ell_k$ are not necessarily halving lines of P.

Theorem 5. *Let P be a set of weighted points in the plane with a total weight that is equal to n and let $\ell_1, \ldots, \ell_k$ be a collection of k lines not passing through any point of P, indexed according to the increasing order of their slopes. For every $1 \leq j < k$ let u_j^- denote the total weight of all points of P that lie above ℓ_{j+1} and below ℓ_j. Similarly, let u_j^+ denote the total weight of all points of P that lie below ℓ_{j+1} and above ℓ_j. Let $u_j = \min(u_j^-, u_j^+)$. Then $\sum_{i=j}^{k-1} u_j \leq n\sqrt{k/2}$.*

Proof. Let $p_1, \ldots, p_m$ denote the set of points in P and let w_i denote the weight of the point p_i for every $i = 1, \ldots, m$. We have

$$\sum_{i<j} w_i w_j = \frac{1}{2}(\sum_{i=1}^{m} w_i)^2 - \frac{1}{2}\sum_{i=1}^{m} w_i^2 = \frac{1}{2}(n^2 - \sum_{i=1}^{m} w_i^2).$$

For every $i < j$ there is at most one index t such that p_i and p_j are at two opposite wedges determined by ℓ_t and ℓ_{t+1}. On the other hand for every two consecutive lines ℓ_t and ℓ_{t+1} the sum of $w(p)w(q)$, taken over all pairs p and q in P such that p contributes to u_t^- and q contributes to u_t^+, is greater than or equal to u_t^2.

Hence, $\sum_{t=1}^{k-1} u_t^2 \leq \frac{1}{2}(n^2 - \sum_{i=1}^{m} w_i^2) \leq \frac{1}{2}n^2$. Therefore, using the Cauchy-Schwartz inequality, we conclude that $\sum_{i=1}^{k-1} u_i \leq n\sqrt{k/2}$. ∎

Observe that the result in Theorem 5 is best possible up to the constant multiplier. To see this, consider a set P of $m = \sqrt{k}$ points in general position in the plane and let $\ell_1, \ldots, \ell_{2\binom{m}{2}}$ be all the lines constructed by taking a line determined by two points in P and then slightly rotating the line both in the clockwise direction and in the counterclockwise direction about the midpoint of the segment determined by the two points of P on the original line. We thus get two lines very close to each line determined by P.

Now put n/m points very close to each point of P. Altogether we have n points and roughly k lines. It is easy to see that for every two lines that arise from one original line determined by P there are exactly n/m points that lie above one and below the other and vice-versa. Therefore, $\sum_{i=1}^{k-1} u_i$ in this case will be at least $\binom{m}{2}n/m \geq \frac{\sqrt{k}}{3}n$.

The result in Theorem 5 yields an alternative proof to Theorem 3. Indeed, assume to the contrary that $q(k) \geq \frac{1}{\sqrt{k}}$. Let μ be a continuous probability measure in the plane and let $\ell_1, \ldots, \ell_k$ be a collection of halving lines for μ, indexed according to their increasing slopes. Consider the arrangement determined by $\ell_1, \ldots, \ell_n$ and put a point at the middle of each face with an assigned weight of the μ measure of that face. It follows that the value of each u_i as defined in the statement of Theorem 5 is at least $\frac{1}{\sqrt{k}}$. Hence, $\sum_{i=1}^{k-1} u_i \geq (k-1)/\sqrt{k} > \sqrt{k/2}$, contradicting Theorem 5.

5 Direct connections with the 'Halving Lines' problem and lower bounds for $g(n, k)$ and $g(k)$

The function $g(n, k)$ is closely related to the problem of bounding the maximum number of halving lines of a point set. For an even number n let $h(n)$ denote the maximum possible number of distinct ways to halve a set of n points by a line. The problem of bounding from above and below the function $h(n)$ is one of the most celebrated open problems in combinatorial and computational geometry and was raised already in the early 70's (see [L71, ELSS73]). The best known upper bound for $h(n)$ currently known is due to Dey [D98]: $h(n) = O(n^{4/3})$. The best known lower bound for $h(n)$ was obtained by Tóth ([To01]): $h(n) = ne^{\Omega(\sqrt{\log n})}$.

Constructions of points sets with many halving lines gives rise to lower bounds for the functions $g(n, k)$ and $g(k)$ as we shall now see.

Claim 2. $g(n, k) \geq n\frac{k}{h^{-1}(k)}$.

Proof. Let $t = h^{-1}(k)$ and let $s = \frac{n}{t}$. Consider a configuration P' of t points in the plane with $h(t) = k$ pairwise non-equivalent halving lines. Let $\ell_1, \ldots, \ell_k$ be such k halving lines of P, indexed according to the increasing order of their slopes. We construct a set P of n points in the plane by taking s points very close to each point of P'. Then P consists of $s|P'| = st = n$ points. Observe that each of $\ell_1, \ldots, \ell_k$ is a halving line also of P. If we now denote by d_i the number of points of P that lie above ℓ_{i+1} and below ℓ_i, then it is easy to see that $d_i \geq s$. Indeed, this follows from the fact that ℓ_i and ℓ_{i+1} halve the set P' in two different ways and therefore the must be a point in P' (and in fact there is just one) that lies above ℓ_{i+1} and below ℓ_i.

It now follows that $\sum_{j=0}^{k-1} d_j \geq ks = n\frac{k}{h^{-1}(k)}$. ∎

Combining Claim 2 with the lower bound on the function $h(t)$ found by Tóth ([To01], namely $h(t) = te^{\Omega(\sqrt{\log t})}$, we obtain the following lower bound for the function $g(n, k)$:

Corollary 2.

$$g(n,k) \geq n e^{c\sqrt{\log k}},$$

for some absolute constant $c > 0$.

Consequently, by Claim 1, we know that $g(k) \geq g(n,k)/n$ for every even number n. We can now easily conclude:

Corollary 3.

$$g(k) \geq e^{c\sqrt{\log k}},$$

for some absolute constant $c > 0$. In particular $\lim_{k\to\infty} g(k) = \infty$.

As for bounding from above the function $g(n,k)$ in terms of the bound on the number of halving lines, we have only the following easy relation:

Claim 3. *For every k, $h(n) \geq g(n,k)$.*

Proof. This follows almost immediately from Proposition 2. Indeed, let P be a set of n points in the plane and let $\ell_1,\ldots,\ell_k$ be halving lines of P indexed according to the increasing order of their slopes such that $\sum_{j=1}^{k-1} d(\ell_j,\ell_{j+1}) = g(n,k)$. From Proposition 2 it follows that for every $1 \leq j \leq k-1$ there are at least $d(\ell_j,\ell_{j+1})$ edges in $G(P)$ with slopes that lie between the slopes of ℓ_j and ℓ_{j+1}. Therefore, $g(n,k) = \sum_{i=1}^{k-1} d(\ell_j,\ell_{j+1}) \geq h(n)$. ∎

Finding configuration of points with many halving lines can improve the lower bounds for $f(\epsilon)$, if the slope of the halving lines are 'well distributed'. This is illustrated in the next theorem.

Theorem 6. *If there exists a set P of n points and a collection of pairwise non-equivalent halving lines $\ell_1,\ldots,\ell_m$ for P such that the angle between any two halving lines with consecutive slopes is smaller than α, then $f(\frac{1}{3n}) \leq 2\alpha$.*

Proof. We construct a measure μ as follows. Take a very small ball around each point of P and define its uniform measure to be $1/n$. We take the balls so small that they do not intersect any of the lines $\ell_1,\ldots,\ell_m$.

Assume to the contrary that $f(\frac{1}{3n}) > 2\alpha$, then there are two lines ℓ and ℓ' that meet at an angle of 2α such that the measure of each of the quadrants of angle $\pi - 2\alpha$, determined by ℓ and ℓ' is bigger than $\frac{1}{2} - \frac{1}{3n}$.

Without loss of generality assume that both ℓ and ℓ' create an angle of α with the positive part of the x-axis such that the slope of ℓ is positive (see Figure 5).

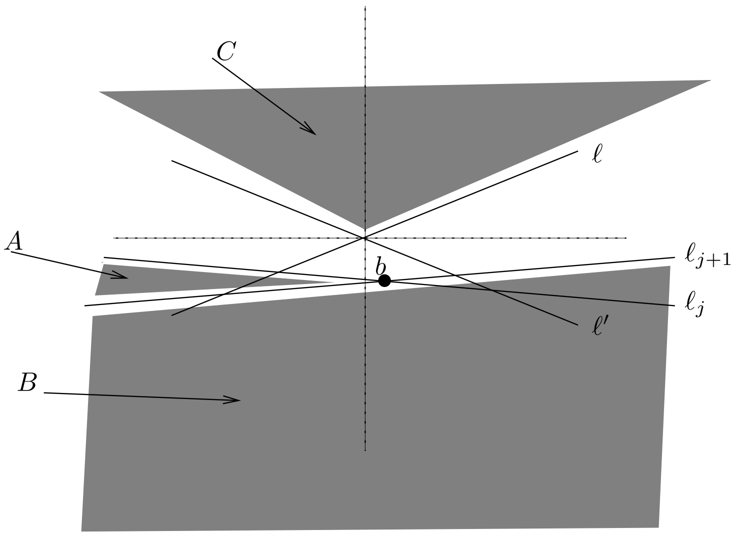

Figure 5: Theorem 6.

We assume that the lines $\ell_1, \ldots, \ell_m$ are indexed according to the increasing order of their slopes. Because ℓ and ℓ' meet at an angle of 2α and the angle between any two consecutive halving lines among $\ell_1, \ldots, \ell_m$ is at most α, there must exist two consecutive halving lines ℓ_j and ℓ_{j+1} in our collection such that both create an angle smaller than α with the positive part of the x-axis.

Let b denote the intersection point of ℓ_j and ℓ_{j+1}. If b lies above ℓ' and below ℓ, then the measure μ of $Wedge(\ell, \ell') \geq 1/n$ which is a contradiction, as μ is a probability measure. Similarly, if b lies above ℓ and below ℓ', the measure μ of $Wedge(\ell, \ell') \geq 1/n$ and again we get a contradiction. Assume therefore that b lies below both ℓ and ℓ' (the symmetric case where b lies above these two lines can be treated similarly).

Denote by A the region which consists of all points that lie below ℓ_j and above ℓ_{j+1}. Note that $\mu(A) \geq \frac{1}{n}$. Let B denote the region which consists of all points that lie below ℓ_{j+1}. As ℓ_{j+1} is a halving line for μ, we have $\mu(B) = \frac{1}{2}$. Let C denote the region which consists of all points that lie above both ℓ and ℓ'.

Notice that the regions A, B, and C are pairwise disjoint. Therefore, $\mu(C) \leq 1 - \mu(A) - \mu(B) = \frac{1}{2} - \frac{1}{n}$. This is a contradiction because we must have $\mu(C) \geq \frac{1}{2} - \frac{1}{3n}$. ∎

Acknowledgments. We wish to thank Eyal Ackerman for pointing out the current state of art regarding spanning trees with low stabbing numbers.

References

[BPZ08] I. Ben-Dan, R. Pinchasi, and R. Ziv, Points with large α-depth, *J. Combin. Theory, Ser. A.*, accepted.

[C88] B. Chazelle, Tight bounds on the stabbing number of spanning trees in Euclidean space. *Report CS-TR-155-88, Dept. Comput. Sci.*, Princeton Univ., Princeton, NJ, 1988.

[D98] T. K. Dey, Improved bounds for planar k-sets and related problems, *Discrete Comput. Geometry* **19** (1998), no. 3, 373–382.

[ELSS73] P. Erdős, L. Lovász, A. Simmons, and E. G. Straus, Dissection graphs of planar point sets, in *A survey of combinatorial theory* pp. 139–149. North-Holland, Amsterdam, 1973.

[L71] L. Lovász, On the number of halving lines, *Ann. Univ. Sci. Budapest. Eőtvős Sect. Math.* **14** (1971), 107–108.

[M91] J. Matousek, Spanning trees with low crossing number. *Inform. Theor. Appl.*, **25** (1991), 102–123.

[To01] G. Tóth, Point sets with many k-sets, *Discrete Comput. Geom.* **26** (2001) no. 2, 187–194.

[W92] E. Welzl, On spanning trees with low crossing numbers. In B. Monien and Th. Ottmann, editors, Data Structures and Efficient Algorithms, volume 594 of Springer LNCS, Berlin, 1992