SERIE B — INFORMATIK

Cutting Dense Point Sets in Half◇

Herbert Edelsbrunner*
Pavel Valtr**
Emo Welzl***

B 94–06
March 1994

Abstract

A *halving hyperplane* of a set S of n points in $\mathbb{R}^d$ contains d affinely independent points of S so that equally many of the points off the hyperplane lie in each of the two half-spaces. We prove bounds on the number of halving hyperplanes under the condition that the ratio of largest over smallest distance between any two points is at most $\delta n^{1/d}$, δ some constant. Such a set S is called *dense*.
In $d = 2$ dimensions, the number of halving lines for a dense set can be as much as $\Omega(n \log n)$, and it cannot exceed $O(n^{5/4}/\log^* n)$. The upper bound improves over the current best bound of $O(n^{3/2}/\log^* n)$ which holds more generally without any density assumption. In $d = 3$ dimensions we show that $O(n^{7/3})$ is an upper bound on the number of halving planes for a dense set. The proof is based on a metric argument that can be extended to $d \geq 4$ dimensions, where it leads to $O(n^{d-\frac{2}{d}})$ as an upper bound for the number of halving hyperplanes.

◇The research by H. Edelsbrunner is partially supported by the National Science Foundation, under grant ASC-9200301 and the Alan T. Waterman award, grant CCR-9118874. The research by P. Valtr is supported by the Deutsche Forschungsgemeinschaft, under grant We 1265/2-1. Part of the research by P. Valtr and E. Welzl has been carried out during a visit to the Tel Aviv University supported by the German-Israeli Foundation for Scientific Research and Development.
*Dept. Comput. Sci., Univ. Illinois at Urbana-Champaign, Urbana, Illinois 61801, USA
**Graduiertenkolleg "Algorithmische Diskrete Mathematik", Freie Univ. Berlin, D-14195 Berlin, Germany and Dept. Applied Math., Malostranské nám. 25, Charles Univ., 118 00 Praha 1, Czech Republic.
***Institut für Informatik, Fachbereich Mathematik und Informatik, Freie Universität Berlin, Takustr. 9, D-14195 Berlin, Germany

1 Introduction

Dense point sets are sets with bounded ratio of largest over smallest distance between points. We present extremal properties of dense point sets for the number of halving hyperplanes and for the stabbing number of geometric d-uniform hypergraphs. In this section, we introduce and discuss notions and results.

Dense Sets. Let $\delta > 0$. A set S of n points in $\mathbb{R}^d$ is *δ-dense* if the ratio of the maximum over the minimum distance between two points in S is at most $\delta n^{1/d}$. We use *dense* short for δ-dense if δ is a constant. Note that values of δ that admit arbitrarily large δ-dense sets cannot be arbitrarily small. For example in $\mathbb{R}^2$, $\delta \geq \delta_0 = \sqrt{2\sqrt{3}/\pi} \approx 1.05$ is necessary and sufficient. Arbitrarily large δ_0-dense sets can be constructed as the intersection of a triangle grid with a disk. For our purposes, it is possible to suppose the minimum distance is 1, so S is δ-dense if the diameter of S is at most $\delta n^{1/d}$.

Various combinatorial extremal problems previously considered for arbitrary point sets have recently been studied for dense sets. These include convex and empty convex subsets [2, 20, 21], separation discrepancy [1], mutual avoidance and crossing families [21]. The complexity of the convex hull problem for δ-dense sets was determined in [21].

The investigation of dense point sets is motivated by the common discrepancy between the complexity of algorithms in the worst case and in practical cases. The complexity of a geometric algorithm typically depends on certain combinatorial parameters associated with the geometric data, and it is often the case that these parameters obtain their extrema only for bizarre and rare data sets. In particular, extremal point sets constructed in the literature often exhibit large distance ratios, sometimes exponential in the number of points. This is related to the fact that there are combinatorial types of point configurations that require the distance ratio be at least double-exponential in the number of points, see [14, 15]. Such sets are unlikely to occur in practical applications. As a consequence, researchers spend a substantial amount of time and effort battling difficulties that are too rare to be relevant in practical applications. The notion of density can be seen as an attempt to make the theoretical analysis of algorithms more relevant to practice.

There are two related concepts. Random sets of points uniformly distributed in the unit square, and sets of points with small integer coordinates (grid points). Dense point sets behave differently from such sets, at least in some respect. For example, a random set of n points in the unit square is likely to have a pair at distance $O(\frac{1}{n})$, and is therefore not dense. Similarly, a dense set of n points in the plane can include the vertices of a convex m-gon, $m = \Omega(\sqrt{n})$, which is impossible for any set of points with integer coordinates between 0 and some constant times $\sqrt{n}$.

It is perhaps worthwhile to mention that dense point sets have been considered in computer graphics [6], and that they commonly appear in nature (eye-photoreceptor

distributions, minimum distance constraints for molecules, atoms of proteins, etc.).

Halving hyperplanes. A notion that is closely related to and somewhat more general than a halving hyperplane is that of a k-set. A subset $T \subseteq S$ of a finite set $S \subseteq \mathbb{R}^d$ is a *k-set* if $\operatorname{card} T = k$ and there is a half-space so that T is the intersection of S with this half-space. The number of k-sets, for some fixed k, arises in the analysis of geometric algorithms, see e.g. [7, 8, 12]. The number of k-sets obtains its maximum when the points are in general position, which we henceforth assume.

The problem of bounding the maximum number of k-sets possible for a set of n points turns out to be notoriously difficult, even in $d = 2$ dimensions. The largest k for which matching upper and lower bounds are known is 3, see [18]. The most difficult case seems to be when k is roughly $\frac{n}{2}$, which is the reason why this case receives special attention as the problem of bounding the number of *halving lines*. The best upper bounds for general point sets in $\mathbb{R}^2$ is $\mathrm{O}(n^{3/2}/\log^* n)$, see [17], and there are constructions of sets with $\Omega(n \log n)$ halving lines, [11, 13]. Our results for dense sets in $\mathbb{R}^2$ consist of an adaptation of the lower bound in [11] and of the upper bound $\mathrm{O}(n^{5/4}/\log^* n)$.

In three and higher dimensions, the problem seems even harder. The first non-trivial upper bound for $d = 3$ dimensions has been given in [5]. This bound has been improved in [4, 10], and the current best bound is $\mathrm{O}(n^{8/3})$, as derived in [9]. In $d \geq 4$ dimensions there are minute improvements over the trivial $\mathrm{O}(n^d)$ bound based on algebraic topology arguments about the non-embeddability of certain complexes, see [22]. We have $\mathrm{O}(n^{d-\frac{2}{d}})$ as an upper bound for dense sets in all $d \geq 3$ dimensions.

Outline. Section 2 shows the $\Omega(n \log n)$ lower bound on the number of halving lines is asymptotically unaffected by the density assumption. Section 3 gives a proof of the $\mathrm{O}(n^{\frac{5}{4}}/\log^* n)$ upper bound on the number of halving lines for a dense set of n points in $\mathbb{R}^2$. The crux of the argument is a reduction to k-sets for about $\sqrt{n}$ subsets of size at most about $\sqrt{n}$ each. Section 4 considers point sets in $\mathbb{R}^3$. The approach to proving an upper bound rests on an area argument combined with an observation about centroids of subsets of the set. This is extended to $d \geq 4$ dimensions in section 5. Alternatively, the metric argument can be combined with an upper bound on the stabbing number of geometric d-uniform hypergraphs known as Lovász' lemma [16]. This is discussed in section 6.

2 Lower bound in the plane

The constructions in [11, 13] establishing the $\Omega(n \log n)$ lower bound on the number of halving lines are not dense: with minimum distance 1 they require the maximum distance be at least some constant times n. This section shows that asymptotically the same lower bound can also be obtained for dense sets. We begin by modifying the construction in [11] so that the x_1-coordinates of the points are contiguous integers.

Lemma 1 *For any positive even integer m there are points $p_1, p_2, \ldots, p_m$ in $\mathbb{R}^2$ so that*

(i) for all i the x_1-coordinate of p_i is i, and

(ii) there are at least $\frac{m}{6}\log_3 \frac{m}{2}$ halving lines.

Proof: Assume first that $m = 2 \cdot 3^k$, $k \geq 0$, and denote the number of halving lines of a set S by $h(S)$. We construct sets S_k, $\mathrm{card}\,(S_k) = 2 \cdot 3^k$, with slightly higher number of halving lines than claimed, namely $h(S_k) = (k+1) \cdot 3^k = \frac{m}{2}\log_3 \frac{3m}{2}$.

For $k = 0$ we have $m = 2$, $S_0 = \{(1,0),(2,0)\}$, and $h(S_0) = 1$. We continue by induction. S_{k+1} is the union of three sets, T_0, T_1, T_2, each an affinely transformed copy of S_k as described below. By horizontal translation, along the x_1-axis, make sure the points in T_i have x_1-coordinates from $im+1$ through $(i+1)m$. Next, shrink the x_2-coordinates of all points in T_i by a sufficiently large factor. Then, move the points vertically, along the x_2-axis, so that the shrunk x_2-coordinates become vertical displacements with respect to a line, one for each T_i. The three lines meet in point $(\frac{2m+1}{2}, 0)$ and have slopes 0, 1, and -1. The line for T_0 is $x_2 = 0$, the one for T_1 is $x_2 = x_1 - \frac{2m+1}{2}$, and the one for T_2 is $x_2 = -x_1 + \frac{2m+1}{2}$. So in fact, the only difference between S_k and T_0 is in the x_2-coordinates which are scaled down for T_0; the points in T_1 and T_2 also experience horizontal translation and vertical displacement other than scaling. The shrinking factor for each set T_i is chosen sufficiently large so that any line through two points of T_i separates T_{i+1} and T_{i+2}, where indices are taken modulo 3.

Any halving line of T_i is also a halving line of S_{k+1}. In addition, S_{k+1} has $\frac{m}{2}$ halving lines connecting a point of T_i with a point of T_{i+1} each, for $i = 0, 1, 2$. It follows that

$$\begin{aligned} h(S_{k+1}) &= h(T_0) + h(T_1) + h(T_2) + 3m/2 \\ &= 3h(S_k) + 3^{k+1} \\ &= (k+2)3^{k+1}. \end{aligned}$$

To extend this construction to arbitrary positive even m observe that all halving lines of S_{k+1} have bounded slope. We can thus add equally many points above and below all these lines and retain them as halving lines. To satisfy the integer coordinate requirement, we add the points with contiguous integer x_1-coordinates, starting at $2 \cdot 3^k + 1$. After adding j points above and j points below all halving lines of S_k, we get a set of $m = 2 \cdot 3^k + 2j$ points, which we denote by P_m for later reference. We may assume $j < 2 \cdot 3^k$, so P_m has fewer than three times as many points than S_k, which implies the claimed lower bound for $h(P_m)$. □

The integer x_1-coordinates can be used to control the density of transformed copies of P_m. This is done in the lower bound construction below.

Theorem 2 *For any even $n \geq 2$ there is a 2-dense set of n points in $\mathbb{R}^2$ with at least $\frac{n}{12}\log_3 n - n$ halving lines.*

Proof: We may assume that $n \geq 3^{12}$, for otherwise the claim is void. Define $m = 2 \cdot \lfloor \frac{\sqrt{n}}{4} \rfloor$ and $M = 2 \cdot \lceil \sqrt{n} \rceil + 9$; so m is even and M is odd. Let Y be the annulus of points on and between the two circles with radii m and $2m$, both centered at the origin. For $1 \leq i \leq M$, let R_i be the half-line with angle $\frac{i-1}{M} \cdot 2\pi$ starting at the origin. The intersection $R_i \cap Y$ is a line segment of length m. For $\varepsilon > 0$ and $1 \leq i \leq M$ find an affine transformation Q_i of P_m so that

(i) $Q_i \subseteq Y$,

(ii) the distance of any $p \in Q_i$ from R_i is at most ε, and

(iii) the distance between any two points in Q_i is at least 1.

For example, shrink the x_2-coordinates of the points in P_m to within $[-\varepsilon, \varepsilon]$, rotate the set by $\frac{i-1}{M} \cdot 2\pi$ about the origin, and translate along R_i until all points lie in Y. Define $S = \bigcup_{i=1}^{M} Q_i$. For sufficiently small $\varepsilon > 0$, every halving line of Q_i is also a halving line of S.

Finally, we adjust the number of points from $\operatorname{card} S = M \cdot m = (2\lceil\sqrt{n}\rceil + 9)(2\lfloor\frac{\sqrt{n}}{4}\rfloor)$ to n. Notice that

$$n < n + \frac{\sqrt{n}}{2} - 18 \leq \operatorname{card} S \leq n + \frac{11\sqrt{n}}{2} < n + 12m.$$

We can delete points from 12 sets Q_i until a set $S' \subseteq S$ with $\operatorname{card} S' = n$ remains. By choosing half these points from Q_j, $1 \leq j \leq 6$, and the other half from Q_j, $\frac{M+1}{2} \leq j \leq \frac{M+11}{2}$, the halving lines of the remaining sets Q_i are still halving lines of S'. Therefore,

$$\begin{aligned} h(S') &\geq (M-12) \cdot h(P_m) \\ &\geq (2\sqrt{n} - 3) \cdot \frac{\sqrt{n}-4}{12} \log_3 \frac{\sqrt{n}-4}{4} \\ &\geq \frac{n}{12}\log_3 n - n. \end{aligned}$$

The diameter of S' is less than $4m$ and the number of points is $n \geq 4m^2$. The density of S' is therefore $\delta < \frac{4m}{\sqrt{4m^2}} = 2$. □

Remark. It is possible to improve the density factor by setting $m \approx \sqrt{\sqrt{3}/(4\pi)} \cdot \sqrt{n} \approx 0.37\sqrt{n}$ and $M \approx \frac{2}{\sqrt{3}} \cdot 2\pi m \approx \sqrt{4\pi/\sqrt{3}} \cdot \sqrt{n} \approx 2.69\sqrt{n}$ and by placing the m points of Q_i alternately at distances roughly $m + \frac{1}{4}$, $m + 1 + \frac{1}{4}$, $\cdots$, $2m - 1 + \frac{1}{4}$ and at distances roughly $m + \frac{3}{4}$, $m + 1 + \frac{3}{4}$, $\cdots$, $2m - 1 + \frac{3}{4}$ from the origin. The diameter is smaller than $4m$, so the density is at most $2\sqrt{\sqrt{3}/\pi} \approx 1.49$. The number of halving

lines is at least $\frac{n}{12}\log_3 n - n$. For any fixed $\delta > \sqrt{2\sqrt{3}/\pi} \approx 1.05$, we can construct a δ-dense set of even size n with at least $cn\log n - n$ halving lines, where $c = c(\delta) > 0$ is independent of n, as follows. Put $m_1 = \frac{1}{2}\delta\sqrt{n}$ and $m_2 = (\frac{1}{2}\delta - \varepsilon)\sqrt{n}$, where $\varepsilon = \varepsilon(\delta) > 0$ is sufficiently small. Let Y be the annulus of points on and between the two circles with radii m_1 and m_2, both centered at the origin. We find a set S of $4m_2(m_1 - m_2)$ points in the annulus Y similarly as in the proof of Theorem 2 so that S has at least $cn\log n - n$ halving lines. Let T be a set of points of a triangle grid with minimum distance 1 which is symmetric around the origin and does not contain the origin. Let T' be the intersection of T with the disk of radius $m_2 - 1$ centered at the origin. We may certainly assume that each of the $cn\log n - n$ halving lines of S cuts T' into two equal-sized parts. Then, $S \cup T'$ has at least $cn\log n - n$ halving lines. If $\varepsilon > 0$ was chosen sufficiently small, $S \cup T'$ has at least n points. Finally, we delete $(\mathrm{card}\,(S \cup T') - n)/2$ pairs of points of T', each pair symmetric around the origin, so that exactly n points remain in $S \cup T'$. This gives us a δ-dense set of size n with at least $cn\log n - n$ halving lines.

3 Upper bound in the plane

Let $f_k(n)$ be the maximum number of k-sets in a set of n points in the plane, and let $f(n) = \max\{f_k(n) : 0 \le k \le n\}$. As mentioned in the introduction, $f(n) = \mathrm{O}(n^{\frac{3}{2}}/\log^* n)$ is the best known upper bound on $f(n)$. We show that δ-dense sets admit smaller bounds. In particular, the result implies that $\mathrm{O}(n^{\frac{5}{4}}/\log^* n)$ is an upper bound for dense sets.

Theorem 3 *The number of halving lines of a δ-dense set of n points in $\mathbb{R}^2$ is at most $\mathrm{O}(\delta\sqrt{n}\cdot f(3\delta\sqrt{n}))$.*

Proof: Let S be a δ-dense set of n points in $\mathbb{R}^2$. Consider a general direction $\gamma \in [0,\pi)$ and a line $\ell = \ell(\gamma)$ with direction γ so that equally many points of S lie on both side of ℓ. The length of $\ell \cap \mathrm{conv}\, S$ is at most the diameter of S, which is at most $\delta\sqrt{n}$. Let $R = R(\gamma)$ be the smallest rectangle that contains the $\frac{1}{2}$-neighborhood of $\ell \cap \mathrm{conv}\, S$; its length is at most $\delta\sqrt{n} + 1$, its width is 1, so its area is at most $\delta\sqrt{n} + 1$. For each point $p \in S \cap R$ consider the disk $D_p = \{x \mid |xp| < \frac{1}{2}\}$, where $|xp|$ is the Euclidean distance between x and p. The area of R covered by D_p is at least $\frac{\pi}{8}$. Since all such disks are disjoint, this implies that

$$\mathrm{card}\,(S \cap R) \le \frac{\delta\sqrt{n}+1}{\frac{\pi}{8}} < \frac{8}{\pi}\cdot\delta\sqrt{n} + 3.$$

We say a line h meets R *shortside* if h meets the two short sides, which have length 1 each. A halving line of S necessarily meets ℓ within $\mathrm{conv}\, S$ and thus within R. It follows that if h is a halving line then it meets R shortside if its direction differs from γ by at most x, where $(\delta\sqrt{n} + \frac{1}{2})\tan x = \frac{1}{2}$, or equivalently, $x = \arctan\frac{1}{2\cdot\delta\sqrt{n}+1}$.

Since $\delta\sqrt{n}$ certainly exceeds 1, we have

$$\arctan\frac{1}{2\delta\sqrt{n}+1} \geq \frac{1}{2\delta\sqrt{n}+1} \geq \frac{1}{3\cdot\delta\sqrt{n}} > \frac{\pi}{10\cdot\delta\sqrt{n}}.$$

Choose $m = \lceil 5\cdot\delta\sqrt{n}\rceil$ and assume that all directions $\gamma = 0, \frac{1}{m}\pi, \frac{2}{m}\pi, \ldots, \frac{m-1}{m}\pi$ are general in the sense that no two points of S lie on a line with such a direction. Two contiguous directions differ by less than $2x$, so each halving line meets at least one of the rectangles $R(\gamma)$ shortside.

To finish the argument observe that each rectangle, R, is met shortside by at most $f(\frac{8}{\pi}\delta\sqrt{n}+3)$ halving lines, because they are all split $S\cap R$ the same way. It follows that the total number of halving lines is at most

$$m\cdot f(\frac{8}{\pi}\delta\sqrt{n}+3) = \mathrm{O}(\delta\sqrt{n}\cdot f(3\delta\sqrt{n})),$$

as claimed. □

Remark. Let $f_k(n,\delta)$ be the maximum number of k-sets in a δ-dense set of n points in the plane, and let $f(n,\delta) = \max\{f_k(n,\delta) : 0\leq k\leq n\}$. Thus, Theorem 3 yields

$$f_k(n,\delta) = \mathrm{O}(\delta\sqrt{n}\cdot f(3\delta\sqrt{n})),$$

where $k = \frac{n}{2} - 1$. An analogous argument gives also

$$f(n,\delta) = \mathrm{O}(\delta\sqrt{n}\cdot f(3\delta\sqrt{n})).$$

Remark. Let us redefine $R = R(\gamma)$ in the proof of Theorem 3 to be the smallest rectangle that contains the ε-neighborhood of $\ell\cap\operatorname{conv} S$, where $\varepsilon > 0$ is a small positive constant. In this case, the points of $S\cap R$ are almost collinear. We call $S\subseteq\mathbb{R}^2$ of size n *δ-line-dense* if there is a line q such that the orthogonal projection of S to q is a δ-dense set in the 1-dimensional interpretation.

Let $l_k(n,\delta)$ be the maximum number of k-sets in a δ-line-dense set of n points in the plane, and let $l(n,\delta) = \max\{l_k(n,\delta) : 0\leq k\leq n\}$. Analogous arguments as in sections 2 and 3 yield that for any fixed $\delta > \delta_0 = \sqrt{2\sqrt{3}/\pi}$ and $\delta' > 1$,

$$f(n,\delta) = \mathrm{O}(\sqrt{n}\cdot l(c_1\sqrt{n},\delta')) = \mathrm{O}(f(c_2 n,\delta)),$$

where c_1 and c_2 are two constants depending only on δ and δ'.

In particular, if the function $f(n,\delta)$ is smooth as a function of n (i.e., $f(cn,\delta) = \mathrm{O}(f(n,\delta))$ for any constant c) then

$$f(n,\delta) = \Theta(\sqrt{n}\cdot l(\sqrt{n},\delta')),$$

and determining the asymptotic behavior of the maximum number of halving lines is equally difficult for dense sets as for line-dense sets. We believe that the function $f(n,\delta)$ (and also the function $f(n)$) is smooth but have no proof at this moment.
Remark. Note that Theorem 3 can be bootstrapped. So if dense sets asymptotically maximize the number of halving lines, then this maximum is $\mathrm{O}(n\ \mathrm{polylog}\ n)$.

4 Upper bound in space

Let S be a dense set of n points in $\mathbb{R}^3$. As mentioned in the introduction, we derive an upper bound on the number of halving planes using an area argument for triangles spanned by points in S. In particular, we need a lower bound on the total area of a collection of m such triangles. We begin with two lemmas. The collection of subsets of S with size k is denoted by $\binom{S}{k}$.

Lemma 4 (Lemma A) *Let S be a set of n points with minimum distance 1 in $\mathbb{R}^3$. Then*

$$\sum_{\{p,q\}\in\binom{S}{2}} \frac{1}{|pq|} \leq 3 \cdot n^{\frac{5}{3}}.$$

Proof: Consider the ordered sequence of the $N = \binom{n}{2}$ distances: $x_1 \leq x_2 \leq \ldots \leq x_N$. The pigeon-hole principle implies that for each index i there is a point with at least $\frac{2i}{n}$ points at distance x_i or less. So we have at least $\frac{2i}{n} + 1$ points inside a ball of radius x_i. Because of the minimum distance assumptions we get at least $\frac{2i}{n} + 1$ pairwise disjoint open balls of radius $\frac{1}{2}$ each inside a ball of radius $x_i + \frac{1}{2}$. By dividing volumes we get

$$\frac{2i}{n} + 1 \leq \frac{(x_i + \frac{1}{2})^3}{(\frac{1}{2})^3} = 8(x_i + \frac{1}{2})^3 \leq 3^3 \cdot x_i^3.$$

It follows that $x_i > \frac{1}{3}(\frac{i}{n})^{\frac{1}{3}}$, and therefore

$$\sum_{\{p,q\}\in\binom{S}{2}} \frac{1}{|pq|} = \sum_{i=1}^{N} \frac{1}{x_i} < 3n^{\frac{1}{3}} \sum_{i=1}^{N} i^{-\frac{1}{3}} < 3 \cdot n^{2-\frac{1}{3}}.$$

□

Based on the bound for distances we can now prove a bound for areas of triangles.

Lemma 5 (Lemma B) *Let S be a set of n points with minimum distance 1 and maximum distance D in $\mathbb{R}^3$. Then the total area of $m > 12Dn^2$ different triangles is at least $\frac{1}{40}m^{\frac{3}{2}}/n^{\frac{5}{6}}$.*

Proof: For two points $p, q \in S$ consider all third points $r \in S$ so that the area of the triangle pqr is at most A and pq is its longest edge. The distance from r to the line through p and q is therefore at most $\frac{2A}{|pq|}$, and r lies between the two planes through p and q normal to pq. The number of points r that can possibly lie inside this cylinder can be bounded by a packing argument. After growing by $\frac{1}{2}$ in every direction, the volume of the cylinder is

$$V = (|pq| + 1)(\frac{2A}{|pq|} + \frac{1}{2})^2\pi \leq \frac{8\pi A^2}{|pq|} + 4\pi A + \frac{\pi|pq|}{2},$$

because $|pq| + 1 \leq 2|pq|$. It contains at most $\frac{6}{\pi}V$ pairwise disjoint open balls of radius $\frac{1}{2}$ each. Using Lemma A we get an upper bound on the number of triangles with area A or less:

$$\sum_{\{p,q\} \in \binom{S}{2}} \frac{6}{\pi} V < 144A^2 n^{\frac{5}{3}} + 12An^2 + \frac{3D}{2}n^2.$$

If $A_1 \leq A_2 \leq \ldots \leq A_M$ is the ordered sequence of the $M = \binom{n}{3}$ triangle areas, we can rewrite this as

$$i \leq 144A_i^2 n^{\frac{5}{3}} + 12A_i n^2 + \frac{3D}{2}n^2.$$

For $i > 6Dn^2$ this implies that at least one of $144A_i^2 n^{\frac{5}{3}}$ and $12A_i n^2$ exceeds $\frac{1}{2}(i - \frac{3D}{2}n^2) > \frac{3i}{4}$. The first inequality implies

$$A_i > \frac{i^{\frac{1}{2}}}{14n^{\frac{5}{6}}},$$

which is weaker than the second inequality provided $i > \frac{8}{7}n^{\frac{7}{3}}$. This is indeed the case if $i > 6Dn^2$, because $D > \frac{1}{2}n^{\frac{1}{3}}$, which can be proved by a straightforward packing argument. Finally, the total area of $m = 2i > 12Dn^2$ triangles is at least

$$i \cdot A_i > \frac{m}{2} \cdot \frac{(\frac{m}{2})^{\frac{1}{2}}}{14 \cdot n^{\frac{5}{6}}} > \frac{1}{40} \cdot \frac{m^{\frac{3}{2}}}{n^{\frac{5}{6}}}.$$

□

The lower bound on the total area of m triangles is sufficient to prove an upper bound on the number of halving planes of a δ-dense set S. Assume n is odd and define $k = \frac{n-1}{2}$. For every subset $T \subseteq S$ of size k, let γ_T be the centroid, that is, $\gamma_T = \frac{1}{k}\sum_{p \in T} p$. Let S_k be the set of k-fold centroids, that is, $S_k = \{\gamma_T \mid T \subseteq S, \operatorname{card} T = k\}$. The convex hull of S_k is a three-dimensional convex polytope, P_k. Note that P_k is contained in the convex hull of S, and by density assumption, its projections to the three coordinate planes have area at most $\delta^2 n^{\frac{2}{3}}$ each. It follows that the surface area of P_k is less than $6\delta^2 n^{\frac{2}{3}}$.

We claim that if the plane, h, through points $p, q, r \in S$ is a halving plane then a homothetic copy of the triangle pqr is a face of P_k. Let $U \subseteq S$ be the subset of points on one side of h, so $\operatorname{card} U = k - 1$. For each $x \in \{p, q, r\}$, there is a plane that separates $U \cup \{x\}$ from $S - U - \{x\}$, and therefore $\gamma_{U \cup \{x\}}$ is a vertex of P_k. Furthermore, the centroids defined by $U \cup \{p\}$, $U \cup \{q\}$, and $U \cup \{r\}$ span a triangular face of P_k. This face is a homothetic copy of pqr, with scaling factor $\frac{1}{k}$. In summary, if the plane through points p, q, r is a halving plane then a homothetic copy of pqr is a face of P_k, and the area of pqr is k^2 times the area of the corresponding facet. Since the surface area of P_k is less than $6\delta^2 n^{2/3}$, the total area of all triangles corresponding to halving planes is less than $6k^2\delta^2 n^{2/3} < \frac{3}{2}\delta^2 n^{8/3}$. Lemma B thus implies the following upper bound on the number of halving planes.

Theorem 6 *The number of halving planes of an δ-dense set of n points in $\mathbb{R}^3$ is less than $20 \cdot \delta^{4/3} n^{7/3}$.*

5 Four and higher dimensions

The area argument of section 4 can be extended to $\mathbb{R}^d$, $d \geq 4$. To simplify the argument, only dense sets of n points in $\mathbb{R}^d$ are considered, so the diameter is $D \leq \delta n^{\frac{1}{d}}$ for some constant δ. Suitable positive constants are used throughout this section and denoted as c, with or without sub- and superscript. A set T of $\operatorname{card} T = k+1 \leq d+1$ points in $\mathbb{R}^d$ spans a k-simplex, $\sigma_T = \operatorname{conv} T$. Its k-dimensional Euclidean measure is denoted by $|\sigma_T|$. We let $s_k(A)$ denote the number of k-simplices spanned by the points whose measure is A or less. All logarithms are to the base 2.

Lemma 7 (Lemma C) *Let S be a dense set of n points in $\mathbb{R}^d$. Then $s_k(A)$ is bounded from above by*

$$\begin{cases} c \cdot nA^d & \text{for } k = 1, \\ c \cdot n^{1+\frac{\binom{k}{2}}{d}} A^d \log A + c' n^{k+\frac{k-1}{d}} & \text{for } 2 \leq k \leq d. \end{cases}$$

Proof: Let F be an $(i-1)$-flat, $1 \leq i \leq d$. The $(i-1)$-dimensional measure of its intersection with $\operatorname{conv} S$ is less than D^{i-1}. It follows that the d-dimensional measure of the cylinder of points at distance $z+\frac{1}{2}$ of less from F is less than $c_1(z+\frac{1}{2})^{d-i+1} D^{i-1}$. With a different constant, c_1', this is also an upper bound on the number of points in S at distance z or less from F.

Consider an arbitrary ordering, $p_0, p_1, \ldots, p_k$, of the vertices of a k-simplex σ_T, $T \subseteq S$. For $1 \leq i \leq k$, let x_i be the distance of p_i from the $(i-1)$-flat spanned by $p_0, \ldots, p_{i-1}$. If $z_1, z_2, \ldots, z_k$ is a sequence of distances so that $x_i \leq z_i$ for all i, then

$$|\sigma_T| = \frac{1}{k!} \prod_{i=1}^{k} x_i \leq \frac{1}{k!} \prod_{i=1}^{k} z_i. \tag{1}$$

The number of k-simplices for which (1) holds is bounded from above by

$$c_2 n \prod_{i=1}^{k} [(z_i + \frac{1}{2})^{d-i+1} D^{i-1}].$$

The number of k-simplices with at least one x_i at most 1 is less than $c' \cdot n^k \cdot D^{k-1}$. For the others we can assume all z_i exceed 1, so the upper bound can be simplified to

$$c_2' n D^{\binom{k}{2}} \prod_{i=1}^{k} z_i^{d-i+1}. \tag{2}$$

For a given A, we can now derive an upper bound on $s_k(A)$, the number of k-simplices σ_T with $|\sigma_T| \leq A$. For $k = 1$ set $z_1 = A$, get $|\sigma_T| = x_1 \leq A$ from (1)

and $s_1(A) \leq c_2' n A^d$ from (2). This completes the proof of the first inequality in Lemma C. For $k \geq 2$ we consider all sequences z_1 though z_k so that $z_i > 1$, for all i, and $\prod_{i=1}^{k} = k!A$. Simplices are counted only if all their x_i's exceed 1; the others are covered by the second term on the right side of the second inequality in Lemma C. To consider all sequences we use induction and the fact that $\prod_{i=1}^{k-1} z_i = k!\frac{A}{z_k}$. So

$$s_k(A) < \int_{z=1}^{A} s_{k-1}(\frac{A}{z}) \cdot z^{d-k+1} \cdot D^{k-1} dz.$$

For $k = 2$ we get

$$\begin{aligned} s_2(A) &< c \cdot nA^d \cdot D \int_{z=1}^{A} \frac{1}{z} dz \\ &< c_2'' \cdot nA^d \log A \cdot D, \end{aligned}$$

and for $k \geq 3$ we get

$$\begin{aligned} s_k(A) &< c \cdot nA^d \log A \cdot D^{\binom{k}{2}} \int_{z=1}^{A} \frac{1}{z^{k-1} \log z} dz \\ &< c_2''' \cdot nA^d \log A \cdot D^{\binom{k}{2}}. \end{aligned}$$

□

The upper bound on $s_k(A)$ can be used to prove a lower bound on the total measure of a collection of k-simplices, all spanned by S.

Lemma 8 (Lemma D) *Let S be a dense set of n points in $\mathbb{R}^d$. The total k-dimensional measure of $m > c'' \cdot n^{k+\frac{k-1}{d}}$ k-simplices is at least*

$$\begin{cases} m^{1+\frac{1}{d}}/(c \cdot n^{\frac{1}{d}}) & \text{for } k = 1 \\ m^{1+\frac{1}{d}}/(c \cdot n^{\frac{2d+k(k-1)}{2d^2}} \log^{\frac{1}{d}} m) & \text{for } 2 \leq k \leq d. \end{cases}$$

Proof: Let $A_1 \leq A_2 \leq \ldots \leq A_M$ be the ordered sequence of the $M = \binom{n}{k+1}$ k-simplex measures. Assume $c'' > 4c'$ so the majority of any set of m k-simplices have all x_i exceeding 1, see the proof of lemma C. For $k = 1$, we can rewrite the result in lemma C as $i < c \cdot nA_i^d$, so

$$\sum_{i=1}^{m} A_i > \frac{1}{(cn)^{\frac{1}{d}}} \sum_{i=1}^{m} i^{\frac{1}{d}} > c_3 \cdot \frac{m^{1+\frac{1}{d}}}{n^{\frac{1}{d}}}.$$

For $k \geq 2$ and $i > \frac{m}{2} > 2c' n^{k+\frac{k-1}{d}}$, we get $i < 2c \cdot n^{\ell} A_i^d \log A_i$, where $\ell = 1 + \binom{k}{2}/d$. Thus

$$A_i > c_4 \cdot \frac{(i \cdot n^{-\ell})^{\frac{1}{d}}}{\log^{\frac{1}{d}}(i \cdot n^{-\ell})}.$$

For the total measure we get

$$\begin{aligned}\sum_{i=1}^{m} A_i &> c_4 \cdot \frac{1}{n^{\frac{\ell}{d}}} \sum_{i=2}^{m} \frac{i^{\frac{1}{d}}}{\log^{\frac{1}{d}} i} \\ &> c_4' \cdot \frac{m^{1+\frac{1}{d}}}{n^{\frac{\ell}{d}} \log^{\frac{1}{d}} m},\end{aligned}$$

because $\int_{x=2}^{m} \frac{x^a}{\log^a x} dx < c_4'' \cdot \frac{m^{1+a}}{\log^a m}$. □

Remark. For $d = 3$, $k = 2$, and D proportional to $n^{\frac{1}{3}}$, the bounds in Lemmas B and D agree up to the logarithmic factor. The larger D gets the further the bound in Lemma D lags behind the one in Lemma B.

For $k = d - 1$, Lemma D says that $m > c'' \cdot n^{d-\frac{2}{d}}$ $(d-1)$-simplices have total $(d-1)$-dimensional measure at least

$$\frac{m^{1+\frac{1}{d}}}{c \cdot n^{\frac{d^2-d+2}{2d^2}} \log^{\frac{1}{d}} m} > \frac{c_5}{\log^{\frac{1}{d}} n} \cdot n^{d+\frac{1}{2}-\frac{3}{2d}-\frac{3}{d^2}}.$$

The exponent of n exceeds $d - \frac{1}{d}$ for $d \geq 4$. This is worth noticing because the same argument on centroids which leads to Theorem 3 in section 4 implies that

$$c_6 D^{d-1} n^{d-1} < c_6' n^{d-\frac{1}{d}}$$

is an upper bound on the total measure of all $(d-1)$-simplices spanning halving hyperplanes. It follows that there cannot be $m > c'' \cdot n^{d-\frac{2}{d}}$ such $(d-1)$-simplices, because this would contradict lemma D.

Theorem 9 *The number of halving hyperplanes of a dense set of n points in $\mathbb{R}^d$, $d \geq 4$, is less than $c \cdot n^{d-\frac{2}{d}}$.*

6 Stabbing number

As described by Santaló [19], there is a motion-invariant measure on lines in $\mathbb{R}^3$ with the property that the measure of lines intersecting a triangle is twice its area. This measure can be used to extend Theorem 6 to the stabbing problem for triangles defined by a δ-dense set $S \subseteq \mathbb{R}^3$. The measure of lines intersecting the convex hull of S is bounded by $O(\delta^2 n^{2/3})$. If we take every line with the multiplicity of its stabbing number in a set of m triangles defined by S, then this measure is twice the total area of these triangles. Lemma B shows that this area is $\Omega(m^{3/2}/n^{5/6})$, provided $m > c\delta n^{7/3}$. So if the set of triangles has the property that every line stabs at most $O(n^2)$ of them, then $m = O(\delta^{4/3} n^{7/3})$, which is the same bound as in Theorem 6.

Lovász' lemma guarantees that this is indeed the case for triangles that span halving planes. Thus, the motion-invariant measure for lines together with Lovász' lemma gives another proof of Theorem 6 based on the area bound in Lemma B. The same can be said about $d \geq 4$ dimensions, Theorem 9, and Lemma D.

References

[1] R. Alexander. Geometric methods in the study of irregularities of distribution. *Combinatorica* **10** (1990), 115–136.

[2] N. Alon, M. Katchalski and W. R. Pulleyblank. The maximum size of a convex polygon in a restricted set of points in the plane. *Discrete Comput. Geom.* **4** (1989), 245–251.

[3] N. Alon and M. A. Perles. On the intersection of edges of a geometric graph by straight lines. *Discrete Math.* **60** (1986), 75–90.

[4] B. Aronov, B. Chazelle, H. Edelsbrunner, L. J. Guibas, M. Sharir and R. Wenger. Points and triangles in the plane and halving planes in space. *Discrete Comput. Geom.* **6** (1991), 435–442.

[5] I. Bárány, Z. Füredi and L. Lovász. On the number of halving planes. *In* "Proc. 5th Ann. Sympos. Comput. Geom., 1989", 140–144.

[6] R. L. Cook. Stochastic sampling in computer graphics. *ACM Trans. Graphics* **5** (1986), 51–72.

[7] B. Chazelle and F. P. Preparata. Halfspace range search: an algorithmic application of k-sets. *Discrete Comput. Geom.* **1** (1986), 83–93.

[8] R. Cole, M. Sharir and C. K. Yap. On k-hulls and related problems. *SIAM J. Comput.* **16** (1987), 61–77.

[9] T. K. Dey and H. Edelsbrunner. Counting triangle crossings and halving planes. *In* "Proc. 9th Ann. Sympos. Comput. Geom., 1993", 270–273.

[10] D. Eppstein. Improved bounds for intersecting triangles and halving planes. Rept. 91-60, Dept. Comput. Sci., Univ. California at Irvine, 1991.

[11] H. Edelsbrunner and E. Welzl. On the number of line separations of a finite set in the plane. *J. Combin. Theory, Ser. A* **38** (1985), 15–29.

[12] H. Edelsbrunner and E. Welzl. Constructing belts in two-dimensional arrangements with applications. *SIAM J. Comput.* **15** (1986), 271–284.

[13] P. Erdős, L. Lovász, A. Simmons and E. G. Straus. Dissection graphs of planar point sets. In *A Survey of Combinatorial Theory*, ed. J. N. Srivastava et al., 139–149, North-Holland, Amsterdam, 1973.

[14] J. E. Goodman, R. Pollack and B. Sturmfels. Coordinate representation of order types requires exponential storage. *In* "Proc. 21st Ann. ACM Sympos. Theory Comput., 1989, 405–410.

[15] J. E. Goodman, R. Pollack and B. Sturmfels. The intrinsic spread of a configuration in $\mathbb{R}^d$. *J. Amer. Math. Soc.* **3** (1990), 639–651.

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

[17] J. Pach, W. Steiger and E. Szemerédi. An upper bound on the number of planar k-sets. *Discrete Comput. Geom.* **7** (1992), 109–123.

[18] E. Ramos. The number of 3-sets of a finite point set in the plane. Rept. UIUCDCS-R-93-1843, Dept. Comput. Sci., Univ. Illinois at Urbana-Champaign, Illinois, 1993.

[19] L. A. Santaló. *Integral Geometry and Geometric Probability.* Addison-Wesley, Reading, Mass., 1976.

[20] P. Valtr. Convex independent sets and 7-holes in restricted planar point sets. *Discrete Comput. Geom.* **7** (1992), 135–152.

[21] P. Valtr. Planar point sets with bounded ratios of distances. Ph.D. thesis, Fachbereich Math., Freie Univ. Berlin, 1994.

[22] R. T. Živaljević and S. T. Vrećica. The colored Tverberg's problem and complexes of injective functions. *J. Combin. Theory, Ser. A* **61** (1992), 309–318.