Discrete Comput Geom 35:177–191 (2006)
DOI: 10.1007/s00454-005-1200-4

Discrete & Computational
Geometry

k-Sets in Four Dimensions*

Jiří Matoušek,[1] Micha Sharir,[2,3] Shakhar Smorodinsky,[3] and Uli Wagner[1]

[1]Department of Applied Mathematics and
Institute for Theoretical Computer Science (ITI), Charles University,
Prague 1, Czech Republic
{matousek,uli}@kam.mff.cuni.cz

[2]School of Computer Science, Tel Aviv University,
Tel Aviv, Israel
michas@tau.ac.il

[3]Courant Institute of Mathematical Sciences, New York University,
New York, NY 10012, USA
shakhar@cims.nyu.edu

Abstract. We show, with an elementary proof, that the number of halving simplices in a set of n points in $\mathbb{R}^4$ in general position is $O(n^{4-2/45})$. This improves the previous bound of $O(n^{4-1/13^4})$. Our main new ingredient is a bound on the maximum number of halving simplices intersecting a fixed 2-plane.

1. Introduction

Let S be a finite set of $n \geq d+1$ points in $\mathbb{R}^d$ and let k be an integer parameter, $1 \leq k \leq n-1$. A *k-set* of S is a k-element subset of S that can be strictly separated from its complement by a hyperplane. The *k-set problem* asks for sharp bounds on the maximum number $F_k^{(d)}(n)$ of k-sets of any set of n points in $\mathbb{R}^d$. The dimension d is usually considered to be a constant, while $k, n \to \infty$. It is not hard to see that the number

* Work by Micha Sharir has been supported by NSF Grants CCR-97-32101 and CCR-00-98246, by a grant from the U.S.–Israeli Binational Science Foundation, by a grant from the Israel Science Fund, Israeli Academy of Sciences, for a Center of Excellence in Geometric Computing at Tel Aviv University, and by the Hermann Minkowski–MINERVA Center for Geometry at Tel Aviv University. Work by Uli Wagner has been supported by the COMBSTRU network of the European Union.

of k-sets is maximized for point sets *in general position*, i.e., such that no $d+1$ points lie in a common hyperplane. In this setting, the following variant of the problem turns out to be essentially equivalent and technically more convenient to study: An oriented $(d-1)$-dimensional simplex σ spanned by d points of S is called a *j-facet* of S, for $0 \le j \le n-d$, if there are exactly j points of S in the positive open halfspace determined by σ. We denote the number of j-facets of S by $G_j(S)$ and seek sharp bounds on the maximum $G_j^{(d)}(n)$ of the numbers $G_j(S)$ over all sets S of n points in general position in $\mathbb{R}^d$. In dimension 2, the number of k-sets of S is equal to $G_{k-1}(S)$, and in dimension 3, it is equal to $\frac{1}{2}(G_{k-2}(S) + G_{k-1}(S)) + 2$; see [3]. In higher dimensions, there are no longer any exact linear relations between these numbers, but for any fixed dimension d, the numbers $F_k^{(d)}(n)$ and $G_k^{(d)}(n)$ lie within constant multiplicative factors of one another (see, e.g., [12]).

A special case arises when $n-d$ is even and $j = (n-d)/2$. Then $G_{(n-d)/2}^{(d)}(n)$ counts the maximum possible number of so-called *halving facets* of S. If we reverse the orientation of a halving facet, we obtain again a halving facet. Thus, we can forget about the orientation and just speak of the underlying unoriented simplices, which are called *halving simplices*. Bounds on the number of halving simplices can be translated to bounds on the number of j-facets for any j, so it is sufficient to study the former quantity. More precisely, if the maximum number of halving facets of n points in dimension d can be bounded by $O(n^{d-c_d})$ for some constant $c_d > 0$, this implies a bound of $G_j^{(d)}(n) = O(n^{\lfloor d/2 \rfloor}(j+1)^{\lceil d/2 \rceil - c_d})$ for all j, see [1].

The study of k-sets and j-facets began more than 30 years ago [9], [11], and tight bounds on the above quantities are still elusive, even in the plane, where the maximum number of halving edges is known to be at most $O(n^{4/3})$ [6], and at least $\Omega(n \cdot 2^{c\sqrt{\log n}})$ for some constant c [16]. In three dimensions the upper bound is $O(n^{5/2})$ [14], and the lower bound is $\Omega(n^2 \cdot 2^{c\sqrt{\log n}})$. In fact, in any dimension d, the known lower bound is $\Omega(n^{d-1} \cdot 2^{c\sqrt{\log n}})$, which is obtained by "lifting" the two-dimensional construction of [16]. In $d \ge 4$ dimensions, the known upper bounds become considerably weaker, and are of the form $O(n^{d-\delta_d})$, where $\delta_d = 1/(4d-3)^d$, leaving a fairly big gap between the upper and lower bounds. Moreover, the proof of these bounds uses the so-called colored Tverberg theorem, for which there is no known elementary proof; the only known proof, given in [17], uses methods from algebraic topology. See [12] for a review of this approach.

In this paper we study the problem in four dimensions, and obtain the first elementary derivation of an upper bound on the number of halving simplices, which also considerably improves the previous upper bound mentioned above. Specifically, we show that the number of halving simplices in a set of n points in $\mathbb{R}^4$ is $O(n^{4-2/45})$. As mentioned above, this implies a bound of $O(n^2(k+1)^{2-2/45})$ for the number of k-facets (and k-sets) for all k.

As in essentially all known proofs of upper bounds for the number of halving simplices in any dimension, we only use a simple local property of halving simplices, the so-called *antipodality property*, first observed by Lovász [11], which we define in more detail below.

2. The Structure of the Proof

We begin by reviewing the notion of antipodal geometric (hyper)graphs. We formulate the definitions for general dimension. Later, we restrict our attention to dimensions 2 and 4.

Definition 2.1. A *geometric hypergraph* in $\mathbb{R}^d$ is a pair (S, T), where S is a finite set of points in general position, and T is a collection of simplices spanned by points from S. The elements of T are also called hyperedges. A geometric hypergraph is called k-uniform if all hyperedges have k vertices, i.e., if all hyperedges are simplices of dimension $k-1$. For a 2-uniform geometric hypergraph in $\mathbb{R}^2$ we drop the prefix "hyper" and just speak of a geometric graph and its edges.

We will often denote a simplex by an (unordered) list of its vertices. Thus, $p_1 \cdots p_k = \operatorname{conv}\{p_1, \ldots, p_k\}$, with the understanding that the points are affinely independent.

Definition 2.2. A d-uniform geometric hypergraph (S, T) in dimension d is called *antipodal* if the following holds for any $d-1$ points $p_1, \ldots, p_{d-1} \in S$: Whenever $a, b \in S$ are two distinct points such that both $ap_1 \cdots p_{d-1}, bp_1 \cdots p_{d-1} \in T$, then there is a third point $c \in S$ such that $cp_1 \cdots p_{d-1} \in T$ and such that the triangle abc intersects the affine hull of $p_1 \cdots p_{d-1}$ (see Fig. 1).

The crucial property is that the family of halving simplices of a point set has this property (see, e.g., [12]):

Lemma 2.3. *Let S be a finite set of points in $\mathbb{R}^d$ in general position, and let T be the set of halving simplices of S. Then (S, T) is an antipodal d-uniform geometric hypergraph.*

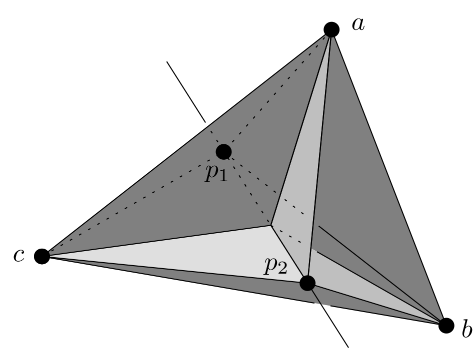

Fig. 1. The antipodality property in dimension 3.

Thus, our goal is to prove the following:

Theorem 2.4. *Let* (S, T) *be an antipodal* 4*-uniform geometric hypergraph in* $\mathbb{R}^4$, *with* $t := |T|$ *hyperedges and* $n := |S|$ *points. Then*

$$t = O(n^{4-2/45}).$$

The earlier proofs of upper bounds for the number of halving simplices consist of two steps. First, one shows that, for an arbitrary d-uniform geometric hypergraph in $\mathbb{R}^d$ with t hyperedges and n points, there is a line that intersects "many" of the hyperedges, namely, at least $\Omega(t^{s_d} n^{d(1-s_d)})$ many, where $s_2 = 2$, $s_3 = 3$, and $s_d = (4d-3)^d$ for $d \geq 4$. Then one shows that if the hypergraph is antipodal, then every line intersects only "few" hyperedges, namely, at most $O(n^{d-1})$ many (this latter observation is often referred to as *Lovász' lemma*; see [5] and [11]). We note, though, that the currently best bounds for $d = 2$ [6] and $d = 3$ [14] are derived using different techniques; see below.

The main new idea in our proof is to use 2-planes instead of lines. The proof then derives the following two lemmas, which stand in analogy to similar lemmas established in the preceding proofs, and which, together, immediately imply Theorem 2.4.

In the statements of the lemmas, a *generic* 2-plane is a plane π that lies in general position with respect to S. In particular, no point of S lies on π, no edge connecting two points of S meets π, and a triangle Δ spanned by S can meet π only at a single point that lies in the relative interior of Δ.

Lemma 2.5. *Let* (S, T) *be a* 4*-uniform geometric hypergraph in* $\mathbb{R}^4$, *with* $n = |S|$ *points and* $t = |T|$ *simplices. If* $t > Cn^{11/3}$, *for some absolute constant* $C > 0$, *then there is a generic two-dimensional plane* π *that intersects* $\Omega(t^3/n^8)$ *simplices of* T.

Lemma 2.6. *Let* (S, T) *be an antipodal* 4*-uniform geometric hypergraph on* n *points in* $\mathbb{R}^4$ *in general position. Then no generic* 2*-dimensional plane intersects more than* $O(n^{4-2/15})$ *simplices of* T.

Sections 3 and 4 are devoted to the respective proofs of these lemmas.

3. Selecting a Stabbing Plane

We need a result of Dey and Pach [8] on crossing simplices. Two simplices σ and τ in $\mathbb{R}^d$, of arbitrary dimensions $0 \leq \dim \sigma, \dim \tau \leq d$, are said to have a *nontrivial intersection* if their relative interiors intersect. They *cross* each other if they have a nontrivial intersection and their vertex sets are disjoint.

Theorem 3.1 (Dey and Pach). *There exists a real constant* $c(d) > 0$ *that depends only on* d, *so that the following holds. Let* (S, T) *be a* $(d+1)$*-uniform geometric hypergraph in* $\mathbb{R}^d$ *on* $n = |S|$ *points in general position. If* $t = |T| \geq c(d)\binom{n}{d}$, *then there are two simplices* $\tau_1, \tau_2 \in T$ *that cross each other.*

We also recall the following simple geometric fact:

Lemma 3.2. *If σ, τ are two simplices in $\mathbb{R}^d$ such that $\sigma \cap \tau \neq \emptyset$, then there are two faces σ' and τ' of σ and τ, respectively, such that $\dim \sigma' + \dim \tau' \leq d$ and σ' and τ' intersect nontrivially.*

Proof. Let σ' and τ' be faces of σ and τ, respectively, such that $\sigma' \cap \tau' \neq \emptyset$ and $\dim \sigma' + \dim \tau'$ is minimal among all such intersecting pairs of faces. Let $x \in \sigma' \cap \tau'$. If x were contained in the relative boundary of σ', say, then there would be a proper face σ'' of σ' still intersecting τ', contradicting the minimality of the dimensions. Thus, x lies in the relative interior of both σ' and τ'. Moreover, if $\dim \sigma' + \dim \tau' > d$, then the intersection of the affine hulls of σ' and τ' would contain a whole line through x. Moving along this line away from x until we first reach the relative boundary of σ' or of τ', we would again find a point x' contained in the intersection of the two simplices and in the relative boundary of at least one of them, thus reaching the same contradiction as before. □

We are now ready to prove Lemma 2.5. We project S and T orthogonally onto a generic hyperplane in $\mathbb{R}^4$, apply the following lemma to the resulting configuration, and then lift the resulting line ℓ orthogonally back to $\mathbb{R}^4$, to obtain the desired 2-plane π. (Actually, additional arguments, based on a slight perturbation of π, are needed to ensure that π is generic.)

Lemma 3.3. *There exists an absolute constant $C > 0$, such that if (S, T) is a 4-uniform geometric hypergraph in $\mathbb{R}^3$, with $n \geq 6$ points in general position, and $t \geq C\binom{n}{3}$ hyperedges, then there is a line ℓ that intersects at least $\Omega(t^3/n^8)$ simplices of T.*

Proof. We take $C := \max(2c(3), 1)$, where $c(3)$ is the constant in Theorem 3.1. First assume that $t > (C/2)\binom{n}{3}$. By Theorem 3.1, there exist two simplices $\tau_1, \tau_2 \in T$ that cross each other. By Lemma 3.2, there exist two faces σ_1 and σ_2 of τ_1 and τ_2, respectively, that cross as well, and such that $\dim \sigma_1 + \dim \sigma_2 \leq 3$. Moreover, since the points of S are in general position, we cannot have $\dim \sigma_1 + \dim \sigma_2 < 3$. Thus, up to symmetry, either σ_1 is an edge ab and σ_2 is a triangle xyz, or σ_1 is a point a contained in the interior of the 3-simplex $\sigma_2 = \tau_2 = xyzw$. In both cases there exists a *crossing edge–simplex pair* (ab, τ) consisting of an edge ab spanned by two points of S and a simplex $\tau \in T$, such that ab crosses τ. Indeed, in both cases, take τ to be τ_2. In the former case, ab is the edge provided by the preceding analysis, and in the latter case, ab is any edge that has a as one endpoint, and a point $b \in S$, which is different from a and from the four vertices of τ_2, as the other endpoint (such a b exists since we assume that $n \geq 6$).

Let x denote the number of such crossing edge–simplex pairs induced by (S, T). It follows that we have

$$x \geq t - \frac{C}{2}\binom{n}{3} \tag{1}$$

for any t and $n \geq 6$. This is trivially true if $t \leq c(3)\binom{n}{3}$. Otherwise, for each crossing edge–simplex pair (ab, τ), delete τ from T. Then the remaining set T' of simplices does not generate any crossing edge–simplex pair, so $t - x = |T'| \le c(3)\binom{n}{3}$.

It will be convenient to use (1) for all integers $n \ge 0$. For $n \le 3$ there are no hyperedges, so $t = 0$, $x = 0$, and (1) holds. For $n = 4, 5$ direct calculation shows that

$$t \le \binom{n}{4} \le \tfrac{1}{2}\binom{n}{3} \le \frac{C}{2}\binom{n}{3}$$

and (1) holds as well.

We now apply a standard random sampling argument to derive a stronger bound. For a parameter $\alpha \in (0, 1)$, to be specified later, let S_α be a random sample obtained by picking each point of S independently with probability α. Let T_α be the set consisting of those simplices $\tau \in T$ all of whose vertices are present in S_α. By the previous argument, the random variables $n_\alpha = |S_\alpha|$, $t_\alpha = |T_\alpha|$, and the random variable x_α counting the number of crossing edge–simplex pairs of (S_α, T_α) satisfy

$$x_\alpha \ge t_\alpha - \frac{C}{2}\binom{n_\alpha}{3}.$$

In particular, this inequality also holds for the expected values of these random variables. We have $\mathbf{E}[x_\alpha] = \alpha^6 x$, $\mathbf{E}[t_\alpha] = \alpha^4 t$, and $\mathbf{E}[\binom{n_\alpha}{3}] = \alpha^3\binom{n}{3}$ (in the last equality, both sides express the expected number of unordered triples in S_α). Hence

$$\alpha^6 x \ge \alpha^4 t - \frac{\alpha^3 C}{2}\binom{n}{3}.$$

Set $\alpha := C\binom{n}{3}/t$, which, by assumption, lies in $(0, 1)$. We conclude that there are

$$x \ge \frac{1}{\alpha^2} t - \frac{C}{2\alpha^3}\binom{n}{3} = \frac{t^3}{2C^2\binom{n}{3}^2} = \Omega\left(\frac{t^3}{n^6}\right)$$

crossing edge–simplex pairs for (S, T). Since there are $\binom{n}{2}$ edges ab, one of them participates in $\Omega(t^3/n^8)$ crossing edge–simplex pairs, so the line spanned by that edge intersects the asserted number of simplices of T. □

To complete the proof of Lemma 2.5, we note that the line ℓ produced by Lemma 3.3 is not generic, since it passes through two points $a, b \in S$. If we slightly perturb ℓ into generic position, we may lose intersections only with those simplices that are incident to either a or b (or both). The number of such simplices is at most $2\binom{n}{3} < n^3$. Hence, as long as $t \gg n^{11/3}$, the new generic line ℓ' will still intersect $\Omega(t^3/n^8)$ simplices of T. Lifting ℓ' back to 4-space, we obtain the desired generic 2-plane π.

4. Bounding the Number of Simplices Stabbed by a 2-Plane

Let (S, T) be a 4-uniform geometric hypergraph in $\mathbb{R}^4$ and let π be a generic two-dimensional plane. Our goal is to bound the number of hyperedges $\tau \in T$ intersected by π, under the assumption that (S, T) satisfies the antipodality property. If a 3-simplex $\tau \in T$ and π intersect (necessarily generically), then the intersection is a line-segment of positive length whose endpoints lie in the relative interior of two triangles bounding τ.

Let E be the set of line segments $\{\tau \cap \pi \mid \tau \in T\}$, and let V be the set of endpoints of these edges. Then $G = (V, E)$ is a geometric graph in the plane π, but with a particular kind of vertex and edge labeling: Each point $q \in V$ is the intersection of π with some triangle spanned by a triple of points $a, b, c \in S$, and we label q by the triple abc. Similarly, each edge $e \in E$ is the intersection of π with some simplex spanned by four points $a, b, c, d \in S$, and we label e by the quadruple $abcd$. The (labels of the) endpoints of e are two distinct subtriples of $abcd$. The order of the indices (points of S) in the label is immaterial, but all indices are distinct. No two objects receive the same label. In particular, $m := |V| \leq \binom{n}{3}$. Moreover, rephrasing what has just been noted, if two points abc and xyz of V are connected by an edge, then the triples abc and xyz must share a common pair of indices, say, $a = x$ and $b = y$, and the edge is labeled by the quadruple $abcz$. We say that the geometric graph G has a *special n-labeling* if its vertices and edges are labeled in this manner (that is, vertices are labeled by distinct triples and each edge is labeled by the quadruple that is the union of the labels of its end vertices).

So far, everything holds for general 4-uniform geometric hypergraphs. It is time to exploit antipodality.

Lemma 4.1. *If (S, T) satisfies the antipodality property, then so does the graph $G = (V, E)$.*

Proof. Let us fix a triangle pqr of S that intersects the plane π at a point o, also labeled by pqr. We view o as the origin of $\mathbb{R}^4$ as well as of the plane π. Since the two-dimensional planes $\operatorname{aff}(pqr)$ and π intersect at the single point o, every point $x \in \mathbb{R}^4$ can be written uniquely as $x = x_\pi + x_{pqr}$ with $x_\pi \in \pi$ and $x_{pqr} \in \operatorname{aff}(pqr)$. Then the projection $x \mapsto x_\pi$ is a linear map $\mathbb{R}^4 \to \pi$, and a point x projects onto $x_\pi = o$ iff $x \in \operatorname{aff}(pqr)$.

Now, suppose that E contains two edges e, f incident to the point o. These edges correspond to simplices $apqr, bpqr \in T$ that are incident to the triangle pqr. By the antipodality of (S, T), there is a third simplex $cpqr \in T$ incident to pqr, such that $\operatorname{aff}(pqr)$ intersects the triangle abc. Since the simplex $cpqr$ intersects the plane π (the point o is contained in both), our genericity assumption implies that the intersection is a line segment. One endpoint of this line segment is the point o, and the other endpoint w arises as the intersection of some other facet of $cpqr$ with π, say $w = cpq \cap \pi$. Our genericity assumption now implies that w lies in the relative interior of cpq, and thus it can be written as $w = \nu c + \rho p + \eta q$, with $\nu + \rho + \eta = 1$ and $\nu, \rho, \eta > 0$. Applying our projection map, it follows that $w = \nu c_\pi$. Similarly, each of the edges e (labeled by $apqr$) and f (labeled by $bpqr$) have o as one of their endpoints, and the other endpoint is of the form $u = \lambda a_\pi$ and $v = \mu b_\pi$, respectively. It remains to observe that $abc \cap \operatorname{aff}(pqr) \neq \emptyset$ implies that $o \in a_\pi b_\pi c_\pi$, and since $\lambda, \mu, \nu > 0$ and we chose o as the origin of our coordinate system, this implies that $o \in uvw$, as asserted. □

In the remainder of this section, we derive an upper bound for the number of edges of an antipodal geometric graph $G = (V, E)$ with a special n-labeling as defined above.

The crossings between edges of G are of central importance in our analysis. We recall the following fundamental result, first proved by Ajtai et al. [2] and independently by Leighton [10] (see also [12]):

Theorem 4.2 (Crossing Lemma). *If $G = (V, E)$ is a simple graph, then in any drawing of G in the plane, there are at least $\Omega(|E|^3/|V|^2)$ crossings between the (not necessarily straight) arcs representing the edges of G, provided that $|E| \geq 4|V|$. Consequently, we always have $|E| = O(|V| + |V|^{2/3}X^{1/3})$, where X is the minimum number of crossings in any drawing of G.*

The proof of this theorem exploits the kind of probabilistic argument that we used in the proof of Lemma 3.3. For antipodal geometric graphs we also have the following result by Dey [6].

Lemma 4.3 (Dey). *The number of crossings between the edges of an antipodal geometric graph $G = (V, E)$ is at most $|V|^2$.*

Proof. We present the proof of the lemma, because later we need the notions of *convex* and *concave chains* that the proof exploits. We remark that the proof presented here is simpler than the original proof; as far as we know, the only previous mention of it is on p. 288 in [12].

By choosing an appropriate coordinate system for the plane, we may assume that no edge in E is vertical. We consider an edge $uv \in E$ with left endpoint u and right endpoint v. If there exists an edge with left endpoint v and with slope larger than the slope of uv, then let vw be the edge that has the smallest slope among all such edges and we call vw the *convex successor* of uv; otherwise, the convex successor is not defined. The antipodality property guarantees that no two edges can have the same convex successor. Thus, if we define a convex chain as a maximal sequence $e_1, \ldots, e_k \in E$ such that each e_{i+1} is the convex successor of e_i, then these chains form a partition of the edge set E, and, clearly, each chain is an x-monotone convex polygonal curve. Note that if uv is the rightmost edge of a convex chain, uv must have the largest slope among all the edges with right endpoint v, for otherwise the antipodality property would imply that uv has a convex successor. Thus, every vertex v is the right endpoint of at most one convex chain, so there are at most $|V|$ such chains. Similarly, there are at most $|V|$ *concave chains*, which are defined analogously (by reversing the direction of the y-axis). If two edges in E cross, then we can extend one of them to a convex chain and the other one to a concave chain, and charge the crossing to the pair of chains. Since a convex curve and a concave curve can cross at most twice, and since a crossing in G can be charged to two different pairs of chains, the total number of crossings is at most $|V|^2$. □

By applying Lemma 4.3 to the graph of halving edges of n points in the plane, we see that there are only $O(n^2)$ crossings between the halving edges. Together with the Crossing Lemma, this yields a simplified proof of Dey's bound of $O(n^{4/3})$ for the number of halving edges.

In our setting, however, direct application of Lemma 4.3 does not yield a sharp bound: The number of vertices is $|V| = O(n^3)$, so Lemma 4.3 only implies that the number of crossings in G is $O(n^6)$. Combining this bound with the bound of the Crossing Lemma, we only obtain the trivial bound $|E| = O(n^4)$. We circumvent this difficulty as follows.

As noted above, each edge of G connects two points abc, abd whose labels share a common pair ab. For each pair of distinct points $a, b \in S$, we define G_{ab} to be the subgraph of all edges whose endpoints share the pair ab of point labels (recall that all labels are unordered). In addition, for each $a \in S$, we define G_a to be $\bigcup_{b \neq a} G_{ab}$. Thus every edge of G belongs to two distinct subgraphs of the form G_a, G_b, and to one subgraph of the form G_{ab}. Note that each graph G_{ab} has at most $n - 2$ vertices and each graph G_a has at most $\binom{n-1}{2}$ vertices. Note also that the degree in G (in G_a, G_{ab}, respectively) of any vertex $\xi = abc$ is at most $3(n - 3)$ ($2(n - 3)$, $n - 3$, respectively) since any neighbor of ξ in G must share two symbols out of $\{a, b, c\}$ and there are only $n - 3$ choices for the third symbol. (In G_a, the two shared symbols are a and one of b, c, and in G_{ab}, the shared symbols are a and b.)

We fix $a \in S$, and consider the graph G_a. Let $\xi = abc$ be a vertex of G_a, and let $d_a(\xi)$ denote its degree in G_a. Each edge of G_a incident to ξ can be classified as either a *b-edge*, if it is of the form (abc, abx), or as a *c-edge*, if it is of the form (abc, ayc). We call ξ *bichromatic* (with respect to a) if the number $d_{ab}(\xi)$ of its incident b-edges and the number $d_{ac}(\xi)$ of its incident c-edges are both at least $\lambda d_a(\xi)$. Here $\lambda \in (0, 1)$ is a parameter whose value will be determined later; we emphasize that it will depend on n. Otherwise, we call ξ *monochromatic* (again, with respect to a).

Let M_a (respectively, B_a) denote the number of edges of G_a that are incident to a monochromatic (respectively, bichromatic) vertex (an edge can contribute to both counts). We put $B := \sum_{a \in S} B_a$ and $M := \sum_{a \in S} M_a$. We first obtain an upper bound for B in terms of n and λ, and then we derive a relation among M, n, λ, and B. Finally, we optimize λ so that the resulting upper bound for $B + M$ is minimized.

A vertex $\xi = abc$ belongs to the three graphs G_a, G_b, G_c, and may be bichromatic in any subset of them. We put

$$B^* = \sum_{a \in S} \sum_{\xi \text{ bichromatic in } \mathrm{G_a}} d_a(\xi).$$

Note that an edge counted in B is counted at least once and at most four times in this sum, so $B^* = \Theta(B)$.

For each integer $k \geq 0$, let Π_k be the set of pairs (a, ξ) such that ξ is bichromatic in G_a and such that $2^{k-1} < d_a(\xi) \leq 2^k$. Since the maximum value of $d_a(\xi)$ is at most $2(n - 3)$, the maximum value of k is at most $\log(2n)$.

For each $a \in S$ and $0 \leq k \leq \log(2n)$, let $V_{a,k}$ denote the set of the vertices ξ of G_a with $(a, \xi) \in \Pi_k$, and put $D_{a,k} = \sum_{\xi \in V_{a,k}} d_a(\xi)$. Note that $D_{a,k} = \Theta(2^k |V_{a,k}|)$, and that $\sum_{k \geq 0} \sum_{a \in S} D_{a,k} = B^*$.

We construct a new graph $G^*_{a,k}$ from G_a as follows. For each bichromatic vertex $\xi = abc$ in $V_{a,k}$, for each b-edge (abc, abx) incident to ξ, and for each c-edge (abc, ayc) incident to ξ, with $x \neq y$, we form the 2-path (abx, abc, ayc), and regard it as a (drawing in π of an) edge (abx, ayc) in $G^*_{a,k}$. Note that an edge (abx, ayc) can be generated at most four times, since the middle vertex must be labeled by a, by one of b, x, and by one of c, y. We turn $G^*_{a,k}$ into a simple graph by retaining only one copy of each multiple edge, which we choose arbitrarily, and draw it as a 2-path that passes through the respective middle vertex, as prescribed above. The number of vertices of $G^*_{a,k}$ is $O(n^2)$.

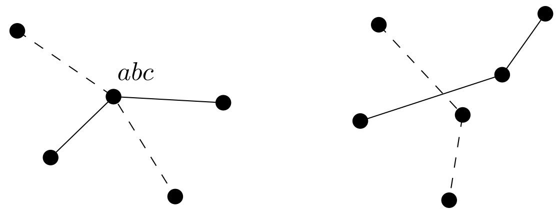

Fig. 2. On the left, the dashed path and the solid one cross improperly at the vertex abc. On the right, the two paths cross properly.

Consider a bichromatic vertex $\xi = abc \in V_{a,k}$ and choose the notation so that $d_{ab}(\xi) \le d_{ac}(\xi)$. Then we have $d_{ab}(\xi) \ge \lambda d_a(\xi)$ because ξ is bichromatic, and $d_{ac}(\xi) \ge \frac{1}{2} d_a(\xi)$ because $d_a(\xi) = d_{ab}(\xi) + d_{ac}(\xi)$. So the number of 2-paths of the form (abx, abc, ayc) is at least $d_{ab}(\xi)(d_{ac}(\xi) - 1)$, where we subtract 1 from $d_{ac}(\xi)$ to ensure that we count only 2-paths with $x \ne y$. Since all the degrees $d_a(\xi)$, for $\xi \in V_{a,k}$, lie in $[2^{k-1}+1, 2^k]$, we have $d_{ab}(\xi)(d_{ac}(\xi)-1) = \Omega(\lambda 2^k d_a(\xi))$. From this we get the following lower bound for the number of edges of $G^*_{a,k}$:

$$|E^*_{a,k}| \ge \Omega\left(\lambda 2^k \sum_{\xi \in V_{a,k}} d_a(\xi)\right) = \Omega(\lambda 2^k D_{a,k}). \tag{2}$$

Let X_a denote the number of crossings between the edges of G_a. We clearly have $\sum_a X_a \le 2X = O(n^6)$. Let $X^*_{a,k}$ denote the number of crossings between the edges of $G^*_{a,k}$, for the above plane embedding of that graph (after eliminating multiple edges). A pair of edges of $G^*_{a,k}$ can cross either *properly*, when they contain two respective crossing edges of G_a, or *improperly*, when they cross each other at a common middle vertex. See Fig. 2. Since the middle vertices of both edges in a crossing pair in $G^*_{a,k}$ are such that their d_a-degrees lie in $[2^{k-1}+1, 2^k]$, each crossing in G_a induces at most $O(2^{2k})$ proper crossings in $G^*_{a,k}$, for a total of $O(2^{2k} X_a)$ proper crossings. Similarly, the number of improper crossings is $O(|V_{a,k}| \cdot 2^{4k}) = O(2^{3k} D_{a,k})$. That is, we have

$$X^*_{a,k} = O(2^{3k} D_{a,k} + 2^{2k} X_a).$$

By the Crossing Lemma, the number of edges of $G^*_{a,k}$ is

$$|E^*_{a,k}| = O(n^2 + (n^2)^{2/3} (X^*_{a,k})^{1/3}),$$

which, combined with (2), yields

$$\lambda 2^k D_{a,k} = O(n^2 + n^{4/3} (2^{3k} D_{a,k} + 2^{2k} X_a)^{1/3}),$$

and this gives

$$D_{a,k} = O(\lambda^{-1} 2^{-k} n^2 + \lambda^{-1} n^{4/3} D_{a,k}^{1/3} + \lambda^{-1} 2^{-k/3} n^{4/3} X_a^{1/3}).$$

If the second term on the right dominates, then $D_{a,k} = O(\lambda^{-3/2} n^2)$. So we always have

$$\begin{aligned} D_{a,k} &= O(\lambda^{-1} 2^{-k} n^2 + \lambda^{-3/2} n^2 + \lambda^{-1} 2^{-k/3} n^{4/3} X_a^{1/3}) \\ &= O(\lambda^{-3/2} n^2 + \lambda^{-1} 2^{-k/3} n^{4/3} X_a^{1/3}) \end{aligned}$$

since $\lambda^{-3/2}n^2 > \lambda^{-1}2^{-k}n^2$. We fix a threshold integer parameter k_0, that we will shortly optimize, sum these bounds over all $k \geq 0$ and $a \in S$, and use Hölder's inequality, to obtain

$$\begin{aligned} B &= O\left(\sum_{k\geq 0}\sum_{a\in S} D_{a,k}\right) \\ &= O\left(\sum_{k\leq k_0}\sum_{a\in S} D_{a,k} + \sum_{k>k_0}\sum_{a\in S} D_{a,k}\right) \\ &= O\left(2^{k_0}\sum_{k\leq k_0}\sum_{a\in S}|V_{a,k}| + \lambda^{-3/2}n^3\log n + \lambda^{-1}2^{-k_0/3}n^{4/3}\left(\sum_{a\in S}X_a\right)^{1/3} n^{2/3}\right) \\ &= O(2^{k_0}n^3 + \lambda^{-3/2}n^3\log n + \lambda^{-1}2^{-k_0/3}n^4). \end{aligned}$$

We now fix the threshold parameter k_0 to minimize this bound. That is, we choose k_0 to be the integer that satisfies $2^{k_0} \leq \lambda^{-3/4}n^{3/4} < 2^{k_0+1}$, and obtain

$$B = O(\lambda^{-3/2}n^3\log n + \lambda^{-3/4}n^{15/4}).$$

The second term dominates the first term, provided that $\lambda > (\log^{4/3} n)/n$, which indeed will be the case for our choice of λ. Hence the total number of edges in all the graphs G_a that are incident to a bichromatic vertex is

$$B = O(\lambda^{-3/4}n^{15/4}). \tag{3}$$

We now turn to the analysis of the overall number M of edges incident to monochromatic vertices, by analyzing the number of edges in the individual refined subgraphs G_{ab}, which proceeds by exploiting the global convex/concave chain decomposition of the whole graph G.

The graph G is antipodal, and so it can be decomposed into $O(n^3)$ pairwise edge–disjoint x-monotone convex (or concave) chains, as in the proof of Lemma 4.3. Consider the graph G_{ab}, for a fixed pair of points $a, b \in S$. We delete from it all the edges that are adjacent to at least one bichromatic vertex either in G_a or in G_b, and we denote by G'_{ab} the resulting subgraph. We take each chain γ from the global collection of convex and concave G-chains, and extract from it all maximal contiguous subchains that consist exclusively of edges of G'_{ab}. Let C_{ab} denote the number of such ab-subchains. See Fig. 3.

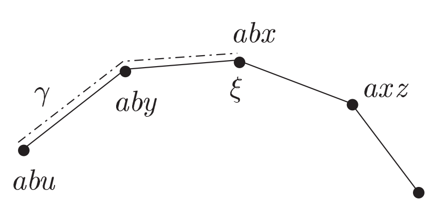

Fig. 3. The extraction of an ab-chain from a global chain.

Lemma 4.4. $\sum_{a,b} C_{ab} = O(\lambda M + \lambda^{-3/4} n^{15/4})$.

Proof. An ab-subchain γ may end, as we trace it from left to right, when the global chain c that contains γ ends. This happens a total of $O(n^3)$ times, over all subgraphs G_{ab}. Otherwise, we consider an "abrupt end" of such a chain γ at some node $\xi = abx$. If the next edge s along the global chain is also an ab-edge, it must be adjacent to a bichromatic vertex, and we charge the chain to s. Any such edge can be charged at most once by concave chains, and at most once by convex chains, because each charged edge belongs to a unique global convex chain c_1, and to a unique global concave chain c_2, and the charging subchain γ is contained in either c_1 or c_2. Hence, the total number of chains of this kind, over all subgraphs G_{ab}, is $O(B) = O(\lambda^{-3/4} n^{15/4})$.

We now suppose that the next edge s is not an ab-edge. The preceding node along the global chain c is of the form aby, and the succeeding node is of the form axz, say (the case where the succeeding node is of the form bxz is symmetric). See Fig. 3. The two edges (aby, abx), (abx, axz) belong to G_a, and we regard the terminal vertex $\xi = abx$ of γ as a vertex of that graph. By construction, ξ is monochromatic in G_a. Then either the number of b-edges incident to ξ, or the number of its incident x-edges, is at most $\lambda d_a(\xi)$. In the former case we charge γ to the edge (aby, abx), and in the latter case we charge it to the edge (abx, axz). As above, an edge can be charged at most once by convex chains, and at most once by concave chains, because it lies on the global chain that contains the charging subchain γ. The overall number of edges that are charged, over all subgraphs G'_{ab}, and thus the number of abrupt ends of the type under consideration, is at most λ times the number of edges incident to monochromatic vertices, that is, at most λM.

Since the first bound, $O(\lambda^{-3/4} n^{15/4})$, on the number of abrupt ends dominates the number $O(n^3)$ of global chain ends, the lemma follows. □

The graph G'_{ab} has $n-2$ vertices. Denote by e_{ab} the number of its edges, and by X_{ab} the number of crossings between its edges. As in the proof of Lemma 4.3, we have $X_{ab} = O(C_{ab}^2)$. By the Crossing Lemma, we have

$$e_{ab} = O(n + n^{2/3} X_{ab}^{1/3}) = O(n + n^{2/3} C_{ab}^{2/3}).$$

Summing these bounds over all graphs G'_{ab}, and using Hölder's inequality, we obtain

$$\begin{aligned} M &\le \sum_{a,b} e_{ab} \equiv e \\ &= O\left(n^3 + n^{2/3} \left(\sum_{a,b} C_{ab} \right)^{2/3} (n^2)^{1/3} \right) \\ &= O(n^3 + n^{4/3}(\lambda M + \lambda^{-3/4} n^{15/4})^{2/3}), \end{aligned}$$

and so

$$\begin{aligned} M &\le e = O(n^3 + n^{4/3}(\lambda^{2/3} M^{2/3} + \lambda^{-1/2} n^{5/2})) \\ &= O(n^{4/3}(\lambda^{2/3} M^{2/3} + \lambda^{-1/2} n^{5/2})). \end{aligned}$$

If the second term on the right dominates, then

$$e = O(\lambda^{-1/2} n^{23/6}). \tag{4}$$

If the first term dominates, then

$$M = O(n^{4/3} \lambda^{2/3} M^{2/3}),$$

or $M = O(\lambda^2 n^4)$. The bound (3) for B is dominated by the bound in (4), provided that $\lambda > 1/n^{1/3}$, which again will be the case for our choice of λ. We thus conclude that

$$e = O(\lambda^2 n^4 + \lambda^{-1/2} n^{23/6}).$$

Now we choose $\lambda = n^{-1/15}$, and note that it satisfies both constraints assumed along the way, namely, $\lambda > (\log^{4/3} n)/n$ and $\lambda > 1/n^{1/3}$. We thus obtain $e = O(n^{4-2/15})$. This complete the proof of Lemma 2.6, and thus also of our main Theorem 2.4. □

As mentioned in the Introduction, an upper bound on the number of halving sets translates into a bound on the number of k-sets (see [1]). Thus, we have obtained the following:

Corollary 4.5. *A set of n points in* $\mathbb{R}^4$ *has at most* $O(n^2(k+1)^{2-2/45})$ *many k-sets*, $0 \le k \le n$.

5. Discussion and Open Problems

Summarizing the technical ingredients of our proof, the first step is to find a 2-plane π that intersects (generically) many halving simplices, and the second step is to show that no 2-plane can intersect many halving simplices. The cross sections of these simplices within π form an antipodal geometric graph G. However, direct application of the Crossing Lemma to G fails to produce sharp bounds, because G has (potentially) too many crossings. However, G has a special labeling of its vertices and edges, and only nodes with "nearby" labels can be connected by an edge. We exploit this property by decomposing G into various subgraphs according to the labels of its edges and vertices, and apply the Crossing Lemma within each subgraph separately. The subgraphs, however, need no longer be antipodal, so we need an estimate on the number of convex and concave chains that cover their edges. This in turn is done by classifying vertices as being either bichromatic or monochromatic, and by carrying out a preliminary analysis that estimates the number of edges incident to bichromatic vertices. This bound is then used to estimate the number of chains in the decomposition of our subgraphs.

Clearly, the main open problem is to extend the new ideas to higher dimensions, where the first target is to do it in five dimensions. The difficulty in such an extension is twofold: Finding a 2-plane that crosses many halving simplices, and deriving an upper bound on the number of halving simplices that a 2-plane can cross. So far we have accomplished the first step, but not the second one. The difficulty is that the labels of the resulting geometric graph are more involved: vertices are labeled by quadruples and edges by quintuples, which tends to make the analysis that we have developed in this paper quite hard to extend.

Apparently, one needs some new ideas. Here are possible directions to look for them.

One possibility is to look for higher-dimensional flats that cross many halving simplices. Consider, for example, a three-dimensional cross section h. The resulting configuration in h is still antipodal, but it now consists of convex polygons with $O(1)$ sides, that meet at common edges, about which antipodality holds. The challenge is to develop new machinery that replaces/extends the convex chain decomposition and the Crossing Lemma.

We observe that the *only* property of halving simplices that we have used is the antipodality. In fact, almost all known proofs only use antipodality. Nevertheless, some progress has been made in the way antipodality is exploited. The earlier proofs use it only implicitly, via the Lovász Lemma. Starting with Dey's proof, antipodality is used more explicitly, via the convex and concave chain decomposition (see [6] and [14], and the present paper, for applications in two, three, and four dimensions, respectively). What we are seeking are new ways to exploit this property, say directly in three dimensions.

A more challenging direction is to find and exploit additional properties of the hypergraph of halving simplices beyond antipodality. In fact, Dey's bound is optimal for arbitrary antipodal geometric graphs. Strangely enough, the property that halving simplices are *halving* is not used at all (other than in proving that their hypergraph is antipodal).

References

1. P. Agarwal, B. Aronov, T. Chan, and M. Sharir, On levels in arrangements of lines, segments, planes, and triangles, *Discrete Comput. Geom.* **19** (1998), 315–331.
2. M. Ajtai, V. Chvátal, M.M. Newborn, and E. Szemerédi, Crossing-free subgraphs, *Ann. Discrete Math.* **12** (1982), 9–12.
3. A. Andrzejak and E. Welzl, In between k-sets, j-facets, and i-faces: (i, j)-partitions, *Discrete Comput. Geom.* **29** (2003), 105–131.
4. B. Aronov, B. Chazelle, H. Edelsbrunner, L. 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, *Combinatorica* **10** (1990), 175–183.
6. T. Dey, Improved bounds for planar k-sets and related problems, *Discrete Comput. Geom.* **19** (1998), 373–382.
7. T. Dey and H. Edelsbrunner, Counting triangle crossings and halving planes, *Discrete Comput. Geom.* **12** (1994), 281–289.
8. T. Dey and J. Pach, Extremal problems for geometric hypergraphs, *Discrete Comput. Geom.* **19** (1998), 473–484.
9. P. Erdős, L. Lovász, A. Simmons, and E.G. Strauss, Dissection graphs of planar point sets, in *Survey Combinat. Theory, Sympos., Colorado State Univ., Colorado* 1971, pp. 139–149, 1973.
10. F.T. Leighton, *Complexity Issues in VLSI*, Foundations of Computing Series, MIT Press, Cambridge, MA, 1983.
11. L. Lovász, On the number of halving lines, *Ann. Univ. Sci. Budapest Eötvös Sect. Math.* **14** (1971), 107–108.
12. J. Matoušek, *Lectures on Discrete Geometry*, Graduate Texts in Mathematics, Springer-Verlag, Berlin, 2002.

13. J. Pach and P.K. Agarwal, *Combinatorial Geometry*, Wiley-Interscience, New York, 1995.
14. M. Sharir, S. Smorodinsky, and G. Tardos, An improved bound for k-sets in three dimensions, *Discrete Comput. Geom.* **26** (2001), 195–204.
15. L. Székely, Crossing numbers and hard Erdős problems in discrete geometry, *Combin. Probab. Comput.* **6** (1997), 353–358.
16. G. Tóth, Point sets with many k-sets, *Discrete Comput. Geom.* **26** (2001), 187–194.
17. R.T. Živaljević and S.T. Vrećica, The colored Tverberg theorem and complexes of injective functions, *J. Combin. Theory, Ser. A* **61** (1992), 309–318.

Received September 7, 2004, *and in revised form June* 8, 2005. *Online publication October* 25, 2005.