Discrete Comput Geom 2:127-151 (1987)

Discrete & Computational Geometry

ε-Nets and Simplex Range Queries

David Haussler[1]* and Emo Welzl[2]

[1] Department of Computer and Information Sciences, University of California at Santa Cruz, Santa Cruz, CA 95064, USA

[2] Institutes for Information Processing, Technical University of Graz, A-8010 Graz, Austria

Communicated by David Dobkin

Abstract. We demonstrate the existence of data structures for half-space and simplex range queries on finite point sets in d-dimensional space, $d \geq 2$, with linear storage and $O(n^\alpha)$ query time,

$$\alpha = \frac{d(d-1)}{d(d-1)+1} + \gamma \qquad \text{for all} \quad \gamma > 0.$$

These bounds are better than those previously published for all $d \geq 2$. Based on ideas due to Vapnik and Chervonenkis, we introduce the concept of an ε-net of a set of points for an abstract set of ranges and give sufficient conditions that a random sample is an ε-net with any desired probability. Using these results, we demonstrate how random samples can be used to build a partition-tree structure that achieves the above query time.

1. Introduction

Rapid processing of geometric range queries has proven to be of fundamental importance in computational geometry, both as an end in itself and as a technique in the efficient solution of other geometric problems. Yao and Yao [23] have

* D. Haussler gratefully acknowledges the support of ONR grant N00014-86-K-0454.

recently demonstrated that a wide variety of range query problems can be reduced to half-space query problems, the basic form of which may be described as follows: given a set of n points in d-dimensional Euclidean space E^d, find a data structure that uses linear storage such that the *number* of points in any query half-space can be determined quickly (i.e., in sublinear time). This problem is called the (*linear storage*) *half-space counting problem.* A common variant of this problem is the *reporting problem*, in which the *set* of points in the half-space has to be determined.

The first sublinear time bounds for half-space counting queries with linear storage were given by Willard [20], who showed that queries are possible in $O(n^\alpha)$ time in E^2 for $\alpha \cong 0.774$. Subsequently, Edelsbrunner and Welzl [14] improved this to $\alpha \cong 0.695$. In E^3, the first nontrivial bound is Yao's [22] ($\alpha \cong 0.936$), which is followed by Dobkin and Edelsbrunner [7] ($\alpha \cong 0.916$), Edelsbrunner and Huber [12] ($\alpha \cong 0.909$), and Dobkin *et al.* [8] ($\alpha \cong 0.899$). Shortly after Cole [6] showed $\alpha \cong 0.977$ in E^4, Yao and Yao [23] gave a generalized version of this result, showing that $\alpha = [\log(2^d - 1)]/d$ can be achieved for all $d \geq 2$. This bound is the best published for $d \geq 4$. In this paper we exhibit a data structure that allows linear storage half-space counting queries in $O(n^\alpha)$ time in E^d for

$$\alpha = \frac{d(d-1)}{d(d-1)+1} + \gamma \qquad \text{for any} \quad \gamma > 0,$$

which improves on previous bounds for all $d \geq 2$; specific bounds are: $\alpha \cong 0.667$ in two dimensions, $\alpha \simeq 0.857$ in three dimensions, and $\alpha \simeq 0.923$ in four dimensions. The technique also works with the same asymptotic bounds for counting queries when ranges are simplices in d dimensions. Bounds for reporting are similar to those for counting, except that the number of points reported must be added to the time bound. It should be noted that better bounds are possible for reporting in two dimensions (specifically $O(\log n + t)$ time, where t is the number of points reported [3]), but these techniques only work for half-planes.

Our techniques are fundamentally similar to previous techniques employed for range queries. A partition tree is defined so that a recursive divide-and-conquer strategy can be efficiently applied to any query. The main difference is that our construction is probabilistic, using random sampling to build each level of the partition tree. It is related to the technique used by Clarkson to build efficient data structures for nearest neighbor queries [4].[1]

The construction of the partition tree in E^2 can be described as follows:

Given a set A of n points in the plane, create a root node and choose at random a subset N of A of size ν. Form the line arrangement consisting of

[1] Recently, other applications of random sampling in computational geometry have been found, including better partition trees for the half-space range query problem where query time must be $O(\log n)$ but nonlinear storage is allowed [5].

all lines defined by pairs of points in N. For each cell in this arrangement, create a child of the root that contains the number of points of A that lie in this cell. Now proceed recursively for each of these children, creating a subtree for the points in its cell until each cell contains less than ν points.

This tree is queried in the usual manner. Given a half-plane determined by a line, the point counts from all cells at the first level of the tree that are completely contained in the half-plane are summed, and recursive calls are made for any cells that are cut by the line. The trick in establishing the time bound is to choose ν such that with high probability:

(1) the total number of points in all cells that are intersected by any line is less than εl for some small positive ε, where l is total number of points at the current level of recursion, and
(2) the total number of cells intersected is reasonably small (this number is bounded by $O(\nu^2)$).

Thus, in contrast to previous techniques, we do not rely on subdivisions of the points into parts of certain sizes. We show that it is enough to choose ν to be $\lfloor c(1/\varepsilon)(\log(1/\varepsilon)+\log(1/\delta))\rfloor$, for some constant c, to get (1) with probability at least $1-\delta$ at any internal node of the tree. The key point is that for half-plane queries, the number ν is essentially a constant that is independent of l, depending only on the parameters ε and δ. As ε approaches zero, we obtain the asymptotic results stated above with arbitrarily small γ, and as δ approaches zero, this happens with arbitrarily high probability. It should be noted that the partition tree is guaranteed to give the correct answer to any query independently of the choice of δ. It is only the time bound for queries that is achieved with probability depending on δ. On the other hand, this shows that for every point set and every $\gamma > 0$ there *exists* a partition tree that gives the claimed time bound. Moreover, as in [4], we can modify the construction algorithm so that at each internal node we repeatedly draw a random sample of size ν until we get property (1). Since property (1) can be checked in $O(l)$ time (the constant depending on ν) and ν can be chosen such that property (1) holds with probability $\frac{1}{2}$ for each draw, this leads to an $O(n \log n)$ expected time algorithm for constructing a partition tree for n points that is guaranteed to realize the claimed time bound. We have not been able to show that such a tree can be efficiently constructed by a non-probabilistic algorithm.

In establishing the existence of the number ν described above, we build on concepts due to Vapnik and Chervonenkis on uniformly approximating classes of events by their empirical distributions [17]. In extending their results, we introduce a new geometrical concept that may be of independent interest. Let R be a class of ranges in E^d for some $d \geq 1$. Elements of R can be any subsets of E^d. Given a finite point set $A \subseteq E^d$ and $\varepsilon \geq 0$, an *ε-net* of A for R is a set of points $N \subseteq A$ such that N contains a point in r for every $r \in R$ with $|A \cap r|/|A| > \varepsilon$. For example, if $\varepsilon = 0$ and R is the set of half-spaces, then the smallest ε-net of A for R is the set of all extreme points of A. It follows that when all points of

A are extreme, the smallest 0-net of A for half-spaces is A itself. We show that this cannot occur for $\varepsilon > 0$ and R the set of d-dimensional half-spaces. In fact, for any $\varepsilon > 0$ and any finite point set $A \subseteq E^d$, there exists an ε-net of A for half-spaces with at most $\lceil (8(d+1)/\varepsilon) \log(8(d+1)/\varepsilon) \rceil$ points. (Here, and throughout the paper, logs are base 2.)

More generally, we characterize the classes of ranges for which there exists a function $f(\varepsilon)$ for $\varepsilon \neq 0$ such that any finite point set A has an ε-net of size $f(\varepsilon)$, independently of the size of A. These are precisely the classes of ranges with finite Vapnik–Chervonenkis dimension, known as Vapnik–Chervonenkis classes [17], [9], [19], [1]. From this characterization result, it follows that if there exists a function $f(\varepsilon)$ such that any finite point set has an ε-net for R of size $f(\varepsilon)$, then in fact any finite point set has an ε-net for R of size at most $\lceil (8d/\varepsilon) \log(8d/\varepsilon) \rceil$, where d is the Vapnik–Chervonenkis dimension of R. Since the Vapnik–Chervonenkis dimension of the class of half-spaces in E^d is $d+1$, the above-cited result is a special case of this theorem. Moreover, we show that if R has Vapnik–Chervonenkis dimension $d < \infty$ then for any $\varepsilon, \delta > 0$ and any finite point set A, if at least $\max((4/\varepsilon)\log(2/\delta), (8d/\varepsilon)\log(8d/\varepsilon))$ points are drawn independently at random from A then these points form an ε-net of A for R with probability at least $1-\delta$. This latter result is used to obtain the crucial number ν for our partition tree construction. Using the related notion of an *ε-approximation* (directly from [17]), we also point out trivial data structures of constant size that give approximate solutions to the counting problem for half-spaces in constant time (compare [13]).

Since the Vapnik–Chervonenkis dimension of half-spaces and balls in d-dimensions is $d+1$ [9], the probabilistic aspects of Clarkson's RPO construction for nearest-neighbor queries [4] can also be derived from the general properties of ε-nets and Vapnik–Chervonenkis classes. These concepts, in slightly more general form, have also proven useful in the investigation of learning algorithms for concepts defined by geometrical regions in feature space [2]. It is our expectation that these concepts, along with the ε-approximation concept, will find other applications in computational geometry as well.

2. Geometric Fundamentals

For $d \geq 1$, let E^d denote d-dimensional *Euclidean space.* We will use h^* to denote one of the two open half-spaces bounded by a hyperplane h in E^d. An open half-space h^* is called *positive* if either h is vertical or h^* is the open half-space above h, i.e., h^* intersects the positive vertical axis in a half-line. By $\bar{h}^*$ we denote the closure of h^*. H_d^* and H_d^+ denote the set of all open half-spaces in E^d and the set of all positive open half-spaces in E^d, respectively. If N is a set of at least d points in general position in E^d, then $H_d(N)$ denotes the set of all hyperplanes that contain d of the points in N.

For $d \geq 0$, and $n \geq 0$ integers, $\Phi_d(n)$ is defined as follows: $\Phi_d(0) = 1$ for all $d \geq 0$, $\Phi_0(n) = 1$ for all $n \geq 0$, and $\Phi_d(n) = \Phi_d(n-1) + \Phi_{d-1}(n-1)$ for $d, n \geq 1$.

Proposition 2.1. $\Phi_d(n) = \sum_{k=0}^{d} \binom{n}{k}$ *if* $d < n$, *otherwise* $\Phi_d(n) = 2^n$.

We assume familiarity with basic notions about arrangements of hyperplanes in E^d; in particular, we use the notion of a cell in an arrangement of a set H of hyperplanes, which can be defined as a maximal connected subset in the dissection of E^d induced by H. See, e.g., [11] for a general treatment of arrangements.

Proposition 2.2. *Let H be a set of n hyperplanes in E^d. Then the number of cells in the arrangement of H is at most $\Phi_d(n)$ and if the hyperplanes in H are in general position, i.e., no two hyperplanes are parallel and no $d+1$ have a common point, then the number of cells in the arrangement equals $\Phi_d(n)$. If h is any other hyperplane in E^d then the number of cells in the arrangement of H intersected by h is at most $\Phi_{d-1}(n)$.*

3. Finite Dimensional Range Spaces, ε-Nets and ε-Approximations

In this section we introduce abstract range spaces and give upper bounds on the number of points needed to form ε-nets and ε-approximations for sets in these spaces. The key concepts and proof techniques of this section are based on the pioneering work of Vapnik and Chervonenkis [17].

Definition. A *range space* S is a pair (X, R), where X is a set and R is a set of subsets of X. Members of X are called *elements* or *points* of S and members of R are called *ranges* of S. S is *finite* if X is finite.

The critical combinatorial parameter associated with a range space is its dimension, introduced in the following definition.

Definition. Let $S = (X, R)$ be a range space and let $A \subseteq X$ be a finite set of elements of S. Then $\Pi_R(A)$ denotes the set of all subsets of A that can be obtained by intersecting A with a range of S, i.e.,

$$\Pi_R(A) = \{A \cap r : r \in R\}.$$

If $\Pi_R(A) = 2^A$, then we say that A is *shattered* by R. The *Vapnik–Chervonenkis dimension of* S (or simply the *dimension* of S) is the largest integer d such that there exists a subset A of X of cardinality d that is shattered by R. If no such maximal d exists, we say the dimension of S is infinite.

We give now a number of examples that illustrate the notion of Vapnik–Chervonenkis dimension.

Example 1. Suppose that $S = (X, R)$ where R is finite. In particular, this case occurs whenever X is finite. Since it requires 2^d distinct ranges in R to shatter d points in X, it follows that the cardinality of the largest subset of X that is shattered by R is at most $d = \lfloor \log |R| \rfloor$. Hence the dimension of S is finite, and this is an upper bound on it.

Example 2. Consider the range space $S_1 = (E^1, H_1^+)$ consisting of the real line and all open half-lines that are unbounded on the right. For any two points a and b with $a < b$, only the three subsets $\emptyset$, $\{b\}$, and $\{a, b\}$ of $\{a, b\}$ can be formed by intersections with ranges in H_1^+, hence no two points of S_1 are shattered. Since it is obvious that a singleton set is shattered, this implies that the dimension of S_1 is 1. If we extend S_1 to (E^1, H_1^*) by adding half-lines unbounded on the left, it is readily verified that two points can be shattered, but three points cannot, thus the dimension is 2.

These results generalize to higher dimensions, so that for any $d \geq 1$, (E^d, H_d^+) is of dimension d and (E^d, H_d^*) is of dimension $d+1$. To see this, consider the dual image of a set A of n points in E^d (see, e.g., [11]). This yields a set of n hyperplanes that partition E^d into $\leq \Phi_d(n)$ cells, with equality whenever the points in A are in general position. Each of these cells corresponds to a unique intersection of a half-space in H_d^+ with A. Since $\Phi_d(n) = 2^n$ for $n = d$ and $\Phi_d(n) < 2^n$ for all $n > d$, this implies that the dimension of (E^d, H_d^+) is d. The other bound follows from an easy extension of this argument.

There are a great variety of natural "geometric" range spaces of finite dimension, e.g., disks or triangular regions in the plane or their higher-dimensional counterparts. The next example, however, shows that there are also nongeometric examples of finite-dimensional range spaces.

Example 3. Let G be an undirected (possibly infinite) graph with node set N. For each node x, let r_x be the set of neighbors of x in G plus x itself. By S_G we denote the range space (N, R) with $R = \{r_x : x \in N\}$. If G is a planar graph then S_G is of dimension at most 4. This can be shown by demonstrating that a shattered subset of five nodes in N forces the graph G to contain a subgraph that is a homeomorphic image of the complete graph on five nodes.

We give another example which shows that there are natural "geometric" range spaces of infinite dimension.

Example 4. Consider the range space (E^2, C), where C is the set of convex polygonal regions in the plane. It is clear that every finite subset A of E^2 on a circle is shattered by C: for each $A' \subseteq A$ the convex hull r of A' is a range in C with $A \cap r = A'$. This shows that the dimension of (E^2, C) is infinite.

When (X, R) is of finite dimension, Dudley calls R a *Vapnik-Chervonenkis Class* (VCC) [9], [19]. Dudley's notion of the *Vapnik–Chervonenkis number* of R corresponds to the dimension of (X, R) plus one. Translating into our terminology, Dudley shows that whenever (X, R) is of finite dimension, then (X, R_k) is also of finite dimension, where R_k is the set of all Boolean combinations formed from at most k ranges in R. Thus, for example, since the set C_k of convex k-gons in E^d for fixed $k > d$ is formed by k-fold intersections of half-spaces, (E^d, C_k) is of finite dimension for any finite k. We give bounds on the dimension of related spaces below. Dudley and Wenocur also prove more general results that imply that the range space formed by the set of all half-spaces bounded by polynomial curves of fixed degree also has finite dimension.

The function $\Phi_d(n)$ plays a fundamental role in all range spaces of finite dimension. The following lemma has been proven independently in [16] and [18].

Lemma 3.1. *Let (X, R) be a finite range space of dimension d with $|X| = n$. Then $|R| \leq \Phi_d(n)$.*

Proof. The assertion is trivially true for $d = 0$ and $n = 0$. Assume the assertion is true for any finite range space of dimension at most $d-1$, and for any range space of dimension d with at most $n-1$ elements, for some $d \geq 1$ and $n \geq 1$.

Let (X, R) be a range space of dimension d with $|X| = n$ and $x \in X$. Consider the range spaces $S - x = (X - \{x\}, R - x)$, where $R - x = \{r - \{x\}: r \in R\}$ and $S^{(x)} = (X - \{x\}, R^{(x)})$, where $R^{(x)} = \{r \in R: x \notin r, r \cup \{x\} \in R\}$. Obviously $S - x$ is of dimension at most d; hence, by assumption, $|R - x| \leq \Phi_d(n-1)$. We show that $S^{(x)}$ is of dimension at most $d-1$.

Let A be a subset of $X - \{x\}$ that can be shattered by $R^{(x)}$. Then it is easy to see that $A \cup \{x\}$ can be shattered by R. (For $A' \subseteq A$ there is an $r \in R^{(x)}$ with $A' = A \cap r$. Since $x \notin r$, $A' = (A \cup \{x\}) \cap r$ and $A' \cup \{x\} = (A \cup \{x\}) \cap (r \cup \{x\})$, where both r and $r \cup \{x\}$ are in R.) Since $A \cup \{x\}$ can be shattered by R, $|A \cup \{x\}| \leq d$, so $|A| \leq d-1$. Thus $S^{(x)}$ is of dimension at most $d-1$.

Since $S^{(x)}$ is of dimension at most $d-1$, by assumption, $|R^{(x)}| \leq \Phi_{d-1}(n-1)$. Observing that $|R| = |R - x| + |R^{(x)}|$, this yields $|R| \leq \Phi_d(n-1) + \Phi_{d-1}(n-1) = \Phi_d(n)$. □

As in Example 2, taking X to be a set of n points in general position in E^d and R the set of all intersections of X with positive half-spaces shows that the above lemma is the best possible. This lemma extends to arbitrary range spaces of finite dimension as follows.

Theorem 3.2. *Let (X, R) be a range space of dimension d. For every finite subset A of X, $|\Pi_R(A)| \leq \Phi_d(|A|)$.*

Proof. We need only observe that if $A \subseteq X$, where (X, R) is a range space of dimension d, then $(A, \Pi_R(A))$ is a finite range space of dimension at most d, and use the above lemma. □

The original motivation for the study of Vapnik-Chervonenkis classes was to determine the classes of sets whose probability measures could be uniformly approximated by random sampling. Since we are only concerned with finite point sets in this paper, we make the following:

Definition. Let (X, R) be a range space and A a finite subset of X. For any ε, $0 \leq \varepsilon \leq 1$, and $V \subseteq A$, V is an *ε-approximation of A* (for R) if for all $r \in R$, $\left| |A \cap r|/|A| - |V \cap r|/|V| \right| \leq \varepsilon$.

Theorem 3.3 [17]. *There is a positive real constant c with the following property: if (X, R) is a range space of dimension d, $A \subseteq X$ is a finite set and ε and δ are real numbers, $0 < \varepsilon, \delta \leq 1$, then a random sample V of A formed by at least*

$$\frac{c}{\varepsilon^2}\left(d \log \frac{d}{\varepsilon} + \log \frac{1}{\delta}\right)$$

independent draws from A is an ε-approximation of A for R with probability at least $1-\delta$.

Clearly Theorem 3.3 immediately implies that if (X, R) is of dimension d, $A \subseteq X$ and $\varepsilon > 0$ then there exists an ε-approximation of A for R of size at most

$$\frac{c}{\varepsilon^2}\left(d \log \frac{d}{\varepsilon}\right) + 1.$$

Example 5. Let A be a set of n points in E^2. Since the dimension of (E^2, H_2^+) is 2, the results in [17, Theorem 2] show that there exists a 0.01-approximation V of A for positive half-planes (and thus for all half-planes) with $|V| = 2{,}525{,}039$. That is, for every half-plane h^*, $|A \cap h^*|$ and $(|V \cap h^*|/|V|)n$ differ by at most 1% of n.

This leads to a simple data structure for point sets that answers half-plane counting queries to any desired accuracy εn, where n is the size of the point set, in constant time. We use an ε-approximation V, count the fraction of points of V in the half-plane and multiply by n. Moreover, the set V can be found with arbitrarily high probability by simply drawing a random sample of the point set of appropriate size. Since Theorem 3.3 holds for any range space of finite dimension, similar results hold when half-planes are replaced by disks, triangular regions, etc., as well as their higher-dimensional counterparts. The drawback is that the constants, if derived from the results in [17], can be quite large.

Our main goal, however, is to find algorithms for half-space queries that give exact answers using linear space and sublinear time. To this end, we introduce a new concept related to that of an ε-approximation.

Definition. Let (X, R) be a range space, A a finite subset of X and $0 \leq \varepsilon \leq 1$. Then $R_{A,\varepsilon}$ denotes the set of all $r \in R$ that contain a fraction of the points in A of size greater than ε, i.e., such that $|A \cap r|/|A| > \varepsilon$. A subset N of A is an *ε-net of A* (for R) if N contains a point in each $r \in R_{A,\varepsilon}$.

It is readily seen that every ε-approximation of a set A is an ε-net of A. The following example exhibits a case where it is easy to show that small ε-nets exist for $\varepsilon > 0$.

Example 6. Consider n points A on a circle in E^2. For any $\varepsilon > 0$, an ε-net of A for half-planes can be found by choosing a subset N of A such that among any $\lfloor \varepsilon n \rfloor + 1$ consecutive points on the circle at least one point is in N. This can clearly be done using at most $\lceil 1/\varepsilon \rceil$ points.

We now give better bounds on the sizes of ε-nets for arbitrary finite-dimensional range spaces than those given above for ε-approximations.

For the following two lemmas, let (X, R) be a fixed finite range space of dimension d. The following will make our arguments clearer.

Notation. For any $\varepsilon > 0$, $R_{X,\varepsilon}$ will be abbreviated by R_ε. For $m \geq 1$, X^m denotes the m-fold cross product of X. A vector $x_1 \cdots x_m \in X^m$ will be denoted $\bar{x}$ when m is clear from the context. Similarly, $\bar{y}$ denotes $y_1 \cdots y_m$. For any $Z \subseteq X^m$, $P^m(Z)$ denotes the probability that a vector in Z is obtained in m independent draws with replacement from X, i.e., $P^m(Z) = |Z|/|X|^m$.

Definition. For any $m \geq 1$ and $\varepsilon > 0$, let Q_ε^m be the set of all m-vectors whose elements do not form an ε-net of X, i.e.,

$$Q_\varepsilon^m = \{\bar{x} \in X^m: \text{there exists } r \in R_\varepsilon \text{ such that } x_i \notin r,\ 1 \leq i \leq m\}.$$

Let

$$J_\varepsilon^{2m} = \{\overline{xy} \in X^{2m} \text{ (where } \bar{x}, \bar{y} \in X^m\text{): there exists } r \in R_\varepsilon \text{ such that } x_i \notin r,\ 1 \leq i \leq m, \text{ but } y_i \in r \text{ for at least } \varepsilon m/2 \text{ indices } i,\ 1 \leq i \leq m\}.$$

Lemma 3.4. *$P^m(Q_\varepsilon^m) < 2P^{2m}(J_\varepsilon^{2m})$ for all $\varepsilon > 0$ and $m \geq 8/\varepsilon$.*

Proof. For $r \in R_\varepsilon$ let

$$Z_r = \{\bar{y} \in X^m: y_i \in r \text{ for at least } \varepsilon m/2 \text{ indices } i,\ 1 \leq i \leq m\}.$$

We first claim that $P^m(Z_r) > \frac{1}{2}$ for all $r \in R_\varepsilon$. To establish this, we show that $P^m(\bar{Z}_r) < \frac{1}{2}$, where $\bar{Z}_r = X^m - Z_r$. Since $P^1(r) \geq \varepsilon$ for each $r \in R_\varepsilon$ and $\bar{y}$ is in $\bar{Z}_r$ only if $y_i \in r$ for fewer than $\varepsilon m/2$ indices i, $P^m(\bar{Z}_r)$ is maximized as $P^1(r)$ approaches ε. In this case, for random $\bar{y} \in X^m$ the expected number of indices i such that $y_i \in r$ is εm and the variance is $\varepsilon(1-\varepsilon)m$. Thus for each $\bar{y} \in \bar{Z}_r$, the number of y_i's in r differs by at least $\varepsilon m/2$ from the expected value. Hence by Chebyshev's inequality

$$P^m(\bar{Z}_r) \leq \frac{\varepsilon(1-\varepsilon)m}{(\varepsilon m/2)^2} < \frac{4}{\varepsilon m} \leq \frac{1}{2},$$

since $m \geq 8/\varepsilon$, establishing the claim.

Now consider a fixed $\bar{x} \in Q_\varepsilon^m$. By definition, there exists $r_{\bar{x}} \in R_\varepsilon$ such that $x_i \notin r_{\bar{x}}$, $1 \leq i \leq m$. From the above, it follows that $\bar{y} \in Z_{r_{\bar{x}}}$ for more than half of the $\bar{y} \in X^m$; hence $\overline{xy} \in J_\varepsilon^{2m}$ for more than half of the $\bar{y} \in X^m$. Thus $P^{2m}(J_\varepsilon^{2m}) > \frac{1}{2}P^m(Q_\varepsilon^m)$. □

Lemma 3.5. *If (X, R) is of finite dimension d and $\varepsilon > 0$, $P^{2m}(J_\varepsilon^{2m}) \leq \Phi_d(2m)2^{-\varepsilon m/2}$.*

Proof. For each j, $1 \le j \le (2m)!$, let π_j be a distinct permutation of the indices $1, \ldots, 2m$. For each $\bar{x} \in X^{2m}$, let $\Theta(\bar{x}) = |\{j: \pi_j(\bar{x}) \in J_\varepsilon^{2m}\}|$. It is easily verified that $P^{2m}(J_\varepsilon^{2m}) \le \max_{\bar{x} \in X^{2m}}(\Theta(\bar{x})/(2m)!)$. Consider a fixed $\bar{x} \in X^{2m}$. Let E be the set of distinct elements of X that appear in $\bar{x}$. For each permutation $\pi_j(\bar{x})$ in J_ε^{2m} there is a subset T of E that is a witness to the fact that $\pi_j(\bar{x}) \in J_\varepsilon^{2m}$ in the sense that there exists $r \in R_\varepsilon$ such that $T = r \cap E$, all occurrences of members of T (and so of r) appear in the second half of $\pi_j(\bar{x})$, and there are at least $\varepsilon m/2$ such occurrences. However, a given T can be a witness for only a small fraction of all permutations of $\bar{x}$. In particular, if there are l occurrences of members of T in $\bar{x}$ and $l \ge \varepsilon m/2$, then T is a witness for at most

$$\frac{\binom{m}{l}}{\binom{2m}{l}} = \frac{m(m-1)\cdots(m-l+1)}{2m(2m-1)\cdots(2m-l+1)} \le 2^{-l} \le 2^{-\varepsilon m/2}$$

of all permutations of $\bar{x}$ (and if $l < \varepsilon m/2$, then T is a witness for no permutation of $\bar{x}$). Since $|E| \le 2m$ and (X, R) is of dimension d, by Theorem 3.2 there are at most $\Phi_d(2m)$ distinct subsets of E induced by intersections with $r \in R_\varepsilon$. Hence, there are at most $\Phi_d(2m)$ distinct witnesses. It follows that

$$\frac{\Theta(\bar{x})}{(2m)!} \le \Phi_d(2m) 2^{-\varepsilon m/2}. \qquad \square$$

Directly from the above two lemmas we get the following:

Theorem 3.6. *If (X, R) is a range space of finite dimension d, A is a finite subset of X, $\varepsilon > 0$, and $m \ge 8/\varepsilon$ then a set $N \subseteq A$ obtained by m random independent draws from A fails to be an ε-net of A for R with probability less than*

$$2\Phi_d(2m) 2^{-\varepsilon m/2}.$$

Proof. This follows from Lemma 3.4 and Lemma 3.5 by observing that $(A, \Pi_R(A))$ is a finite range space of dimension at most d. $\square$

Corollary 3.7. *For any (X, R) of finite dimension d, finite $A \subseteq X$ and $0 < \varepsilon, \delta < 1$, if N is a subset of A obtained by*

$$m \ge \max\left(\frac{4}{\varepsilon}\log\frac{2}{\delta}, \frac{8d}{\varepsilon}\log\frac{8d}{\varepsilon}\right)$$

random independent draws, then N is an ε-net of A for R with probability at least $1 - \delta$.

Proof. We claim that for m of this size, the bound given in Theorem 3.6 is less than δ. Hence the sample N will be an ε-net of A with probability at least $1 - \delta$. This claim follows from Lemma 7 of [2], but, for completeness sake, we sketch the proof.

We establish this claim for the case $d > 1$. The proof for $d = 1$ is similar. It follows from Proposition 2.1 that for $d \geq 2$, $\Phi_d(2m) \leq (2m)^d$. Thus it suffices to show that $2(2m)^d \leq \delta 2^{\varepsilon m/2}$, which is equivalent to $\varepsilon m/2 \geq d \log(2m) + \log(2/\delta)$. The first bound of the two bounds on m implies $\varepsilon m/4 \geq \log(2/\delta)$. Thus it suffices to show that $\varepsilon m/4 \geq d \log(2m)$. If this inequality holds for some value of m, it will also hold for larger values, so suppose m is equal to the second bound in the statement of the theorem. We need only show that

$$2d \log \frac{8d}{\varepsilon} \geq d \log\left(\frac{16d}{\varepsilon} \log \frac{8d}{\varepsilon}\right),$$

which is equivalent to $4d/\varepsilon \geq \log(8d/\varepsilon)$ and this certainly holds. □

Although only the upper bound is used in what follows, for any finite-dimensional range space (X, R) we can, using the above corollary, give upper and lower bounds on the size of the smallest ε-net of any finite $A \subseteq X$ for R in terms of the dimension of (X, R), independent of the size of A. To this end, we make the following:

Definition. For any range space (X, R), $\varepsilon > 0$ and finite $A_0 \subseteq X$,

$$F((X, R), \varepsilon, A_0) = \min\{|N|: N \text{ is an } \varepsilon\text{-net of } A_0 \text{ for } R\},$$

$$f((X, R), \varepsilon) = \max\{F((X, R), \varepsilon, A): A \text{ is a finite subset of } X\}.$$

For any $\varepsilon > 0$ and $d \geq 1$,

$$f(d, \varepsilon) = \max\{f((X, R), \varepsilon): (X, R) \text{ is of dimension } d\}.$$

Theorem 3.8. *For any finite $d \geq 1$ and $0 < \varepsilon < 1$, $f(d, \varepsilon)$ exists and*

$$\left\lfloor \frac{d}{2} \right\rfloor \frac{1}{\varepsilon} - 1 \leq f(d, \varepsilon) \leq \left\lceil \frac{8d}{\varepsilon} \log \frac{8d}{\varepsilon} \right\rceil.$$

Proof. The above corollary shows that if $m = \lceil (8d/\varepsilon) \log(8d/\varepsilon) \rceil$, for any (X, R) of dimension d, a sample N obtained by m independent random draws of any finite $A \subseteq X$ will be an ε-net of A with probability greater than 0. Hence there always exists an ε-net of this size. For the lower bound, let $X = \{1, 2, \ldots, s\}$ and R_k be the set of all k-fold unions of intervals on X, where $k \geq 1$, i.e.,

$$R_1 = \{\{i, i+1, \ldots, j\}: 1 \leq i \leq j \leq s\}$$

and

$$R_k = \{r_1 \cup r_2: r_1 \in R_{k-1} \text{ and } r_2 \in R_1\} \qquad (k > 1).$$

Assume that $s \geq 2k+1$. We first claim that the dimension of (X, R_k) is $2k$. It is easily verified that the set $\{1, 2, \ldots, 2k\} \subseteq X$ is shattered by R_k, hence the dimension is at least $2k$. On the other hand, if $A = \{x_1, x_2, \ldots, x_{2k+1}\}$ is any subset of X of cardinality $2k+1$, where without loss of generality we assume that $x_i < x_{i+1}$, $1 \leq i \leq 2k$, then there is no range r in R_k such that $r \cap A = \{x_1, x_3, \ldots, x_{2k-1}, x_{2k+1}\}$. Hence A is not shattered by R_k and thus the dimension of (X, R_k) is $2k$.

Now let $N=\{x_1, x_2, \ldots, x_c\}$ be an ε-net of X for R_k for some $0<\varepsilon<1$ and $c \geq 1$, where, as above, we assume that $x_i < x_{i+1}$, $1 < i < c$. If $G \subseteq X$ is such that

$$G=\{1, 2, \ldots, x_1-1\},$$
$$G=\{x_i+1, x_i+2, \ldots, x_{i+1}-1\} \qquad (1 \leq i < c),$$

or

$$G=\{x_c+1, x_c+2, \ldots, s\},$$

then G will be called a *gap* (*generated by* N). (Some gaps may be empty). Since N is an ε-net of X for R_k,

$$\sum_{i=1}^{k} |G_i| \leq \varepsilon s,$$

where $G_1, \ldots, G_k$ are the k largest gaps generated by N (breaking ties arbitrarily). Since N generates $c+1$ gaps and the total size of all gaps is $s-c$,

$$\sum_{i=1}^{k} |G_i| \geq k\left(\frac{s-c}{c+1}\right).$$

Hence

$$k\left(\frac{s-c}{c+1}\right) \leq \varepsilon s,$$

which implies that

$$c \geq \left(\frac{k}{\varepsilon}-1\right)\left(\frac{s}{s+k/\varepsilon}\right).$$

Thus if s is large enough, we must have $c \geq k/\varepsilon - 1$. Since the dimension of (X, R_k) is $d=2k$, this shows that there exist finite range spaces of dimension d that have no ε-nets of size less than $d/(2\varepsilon)-1$, for all even $d \geq 2$. The lower bound given above follows easily. □

It is easily verified that $f((X, R), \varepsilon)$ does not exist for any ε, $0 \leq \varepsilon < 1$, if the dimension of (X, R) is infinite: whenever (X, R) has infinite dimension then for each $n \geq 1$ we can find a subset A_n of X of size n that is shattered by R. To obtain an ε-net of A_n it is clear that at least $(1-\varepsilon)n$ points must be used, otherwise there is a subset of A_n in $R_{A,\varepsilon}$ with no points in it. Clearly, this function grows arbitrarily large as n grows. Thus the existence of ε-nets of fixed size for arbitrarily large subsets is a characteristic property of finite-dimensional range spaces. While the sizes of ε-nets are considerably smaller than the sizes of ε-approximations obtained using Theorem 3.3, they are still quite large, as is evidenced in the following:

Example 7. For any set A of points in E^2 there is a 0.01-net N of A of size at most 5,355, i.e., a subset N such that every positive half-plane that contains at

least 1% of the points of A contains at least one point in N. This estimate can be obtained by using Theorem 3.6 directly, setting the bound given there to be less than 1. Since the Vapnik–Chervonenkis dimension of the set of all triangular regions in E^2 is 7, if the size of N is increased to 19,045, the same result holds for triangular regions. It is likely that these results actually hold for considerably smaller numbers.

We conclude this section by examining the relationship between the notion of an ε-net and the established notion of a centerpoint [21], [11] in combinatorial geometry. Let A be a set of n points in E^d. A point x (not necessarily in A) is called a *centerpoint* if every half-space that contains more than $[d/(d+1)]n$ points of A contains x as well. It is known that every point set in E^d has a centerpoint.

Let a set N be a *weak ε-net* of a point set A for a set of ranges R if it satisfies the conditions of an ε-net except that N is not necessarily a subset of A. Clearly, a point x is a centerpoint of a point set A in E^d if and only if $\{x\}$ is a weak $d/(d+1)$-net of A for half-spaces. Hence the existence of centerpoints implies the existence of weak $d/(d+1)$-nets for half-spaces of cardinality one in E^d. Moreover, there are always at most $d+1$ points N in A such that their convex hull contains a centerpoint. This set N then forms a (strong) $d/(d+1)$-net of A for half-spaces.

In general, it makes a significant difference whether or not we choose the points in the ε-net from the point set under consideration. For example, every finite point set A in E^d is contained in the convex hull of $d+1$ points. These points form a weak 0-net of A for half-spaces, while all of the extreme points of A are required for a (strong) 0-net.

4. Range Search Data Structures

We now describe linear storage data structures that support half-space and simplex counting and reporting queries for point sets in E^d. In our discussion of these data structures we treat only half-space counting queries; the extension to simplex counting is considered briefly at the end of this section and the extension to reporting queries will turn out to be obvious. In order to simplify the exposition, we restrict ourselves to point sets in general position in E^d.

Definition. Let $d \geq 2$ and $\nu > d$ be integers and $0 \leq \varepsilon \leq 1$ be a real number. A rooted tree T is called an (ε, ν)-*partition tree* for a finite point set A in general position in E^d, if the following hold:

(1) Every node p of T corresponds to some open region $reg(p)$ of E^d; $set(p)$ denotes $reg(p) \cap A$.
(2) The root r of T corresponds to $reg(r) = E^d$; hence $set(r) = A$.
(3) Let p be a node of T.

If $|set(p)| > \nu$, then p is an *internal node of* T and contains:

(i) The number $size(p) = |set(p)|$.
(ii) A set $points(p) \subseteq set(p)$ of cardinality ν.
(iii) The arrangement $arr(p)$ of all hyperplanes that contain d of the points in $points(p)$.
(iv) For each cell f in $arr(p)$ with $f \cap set(p) \neq \emptyset$, a pointer to a child p_f of p with $reg(p_f) = f \cap reg(p)$.

If $|set(p)| \leq \nu$, then p is a *leaf of* T and contains:

(i) The number $size(p) = |set(p)|$.
(ii) The set $points(p) = set(p)$.

(4) For every internal node p in T and every hyperplane h in E^d

$$\sum_{f \in F} |f \cap set(p)| \leq \varepsilon |set(p)|,$$

where F is the set of all cells f in $arr(p)$ with $f \cap h \neq \emptyset$.

Condition (4) in the definition of an (ε, ν)-partition tree is trivially satisfied for $\varepsilon = 1$. For smaller ε, we will see below that this condition leads to sublinear counting query times for half-spaces.

The objects contained in the nodes of the tree are used to support half-space counting queries, which will be performed by the procedure $ENUM(p, h^*)$, p a node in the tree and h^* a half-space. We assume the global variable $ANSW$ is set to zero prior to the query. Upon termination, $ANSW$ will hold the integer $|A \cap h^*|$. The query is accomplished by the call $ENUM(r, h^*)$, where r is the root of the (ε, ν)-partition tree storing the point set A. The procedure $ENUM(p, h^*)$ can be described as follows:

```
ENUM(p, h*):
  ANSW := ANSW + |points(p) ∩ h*|;
    if p is an internal node do:
      for each cell f in arr(p) for which child p_f exists do:
        if f ⊆ h*, then ANSW := ANSW + size(p_f);
        else if f is intersected by h, then ENUM(p_f, h*).
```

Note that we make the call $ENUM(p_f, h^*)$ under condition "f is intersected by h" and not under the condition "$reg(p_f)$ is intersected by h." Since, in general, $reg(p_f)$ is a proper subset of f, the latter rule would avoid some unnecessary recursive calls, i.e., calls on p_f when f is intersected by h^* but $reg(p_f)$ is not. The drawback is that this requires storing $reg(p)$ for each node, and these regions can become quite complex. Using the former rule, we simply store $arr(p)$, for example as a linked list of records, each of which corresponds to a cell f with $f \cap set(p) \neq \emptyset$. The record for f contains a pointer to the child p_f and a representation of f that allows one to test if f is intersected by h or contained in h^* for any hyperplane h. An appropriate choice for this representation uses $O(\nu^{d^2})$ storage for an internal node, and supports all the necessary computations for the node (excluding recursive calls) in $O(\nu^{d^2})$ time.

This leads to the following observations on the query time and space requirements for this structure.

Observation 4.1. Let $Q_T(p, h^*)$ be the query time required by the call $ENUM(p, h^*)$ in an (ε, ν)-partition tree T storing a point set in E^d for a half-space h^*.

(1) If p is an internal node then

$$Q_T(p, h^*) \leq O(\nu^{d^2}) + \sum_{f \in F} Q_T(p_f, h^*),$$

where F is the set of all cells f in $arr(p)$ with $f \cap h \neq \emptyset$.

(2) If p is a leaf then $Q_T(p, h^*) \leq O(\nu)$.

Observation 4.2. An (ε, ν)-partition tree uses $O(\nu^{d^2} n)$ storage to represent a set of n points in E^d in general position.

Lemma 4.3. *An (ε, ν)-partition tree T, $0 < \varepsilon < 1$, storing a set A of n points in E^d in general position, supports the computation of $|A \cap h^*|$ for every half-space h^* in $O(\nu^{d^2} n^\alpha)$ time with*

$$\alpha = \frac{\log_{1/\varepsilon} \Phi_{d-1}\left(\binom{\nu}{d}\right)}{\log_{1/\varepsilon} \Phi_{d-1}\left(\binom{\nu}{d}\right) + 1}.$$

Proof. By Observation 4.1, for a node p in T and a half-space h^*, the time required by the call $ENUM(p, h^*)$ is

$$Q_T(p, h^*) \leq O(\nu^{d^2}) + \sum_{f \in F} Q_T(p_f, h^*)$$

(where F is the set of all cells f in $arr(p)$ with $f \cap h \neq \emptyset$ if p is an internal node and F is empty otherwise). Moreover, we know that by Proposition 2.2

$$|F| \leq \Phi_{d-1}\left(\binom{\nu}{d}\right)$$

and by condition (4) of the definition of an (ε, ν)-partition tree

$$\sum_{f \in F} |set(p_f)| \leq \varepsilon |set(p)|.$$

Let $Q(0) = 0$ and for any integer $n \geq 1$, let $Q(n)$ be the maximal time required by a call $ENUM(p, h^*)$ for any node p in any (ε, ν)-partition tree storing a set A in E^d and any half-space h^*, where $|set(p)| \leq n$. Then, for all $n \geq 1$,

$$Q(n) \leq c_0 \nu^{d^2} + \max \sum_{i=1}^{m} Q(n_i),$$

where $c_0 > 0$ is some constant, $m = \Phi_{d-1}\left(\binom{\nu}{d}\right)$ and the maximum is taken over all nonnegative integers $(n_1, n_2, \ldots, n_m)$ with $\sum_{i=1}^{m} n_i \leq \varepsilon n$. It can be shown that this implies that $Q(n)$ is $O(\nu^{d^2} n^\alpha)$ (see the appendix), which gives the claimed time bound. □

It is clear that (ε, ν)-partition trees do not exist for every point set A in E^d and every ε and ν. For example, a $(0, \nu)$-partition tree exists for a set A in E^d (in general position) if and only if $\nu \geq |A|$. In the next few lemmas we show that for $\varepsilon > 0$, (ε, ν)-partition trees always exist with ν "reasonably small" compared with $1/\varepsilon$.

Lemma 4.4. *Let N be a set of $\nu > d$ points in E^d in general position and let h^* be a half-space in E^d.*

(1) *If $N \subseteq \bar{h}^*$ then there exist k half-spaces $h_1^*, h_2^*, \ldots, h_k^*$, where $k \leq d$, such that $h_i \in H_d(N)$ and $N \subseteq \bar{h}_i^*$ for each i, $1 \leq i \leq k$, and*

$$h \subseteq E^d - \bigcap_{i=1}^{k} h_i^*.$$

(2) *If both $N_1 = h^* \cap N$ and $N_0 = N - \bar{h}^*$ are nonempty, then there exist k half-spaces $h_1^*, h_2^*, \ldots, h_k^*$, where $k \leq d+1$, such that $h_i \in H_d(N)$, $N_1 \subseteq \bar{h}_i^*$ and $N_0 \cap h_i^* = \emptyset$ for each i, $1 \leq i \leq k$, and*

$$h \subseteq \bigcup_{i=1}^{k} \bar{h}_i^* - \bigcap_{i=1}^{k} h_i^*.$$

Proof. We outline here only the idea for (2). An illustration of this case for $d = 2$ is given in Fig. 1.

Let H^* be the set of all half-spaces g^* with $g \in H_d(N)$ such that $N_1 \subseteq \bar{g}^*$ and $N_0 \cap g^* = \emptyset$. Then it is easy to see that

$$h \subseteq \bigcup_{g^* \in H^*} \bar{g}^* - \bigcap_{g^* \in H^*} g^*.$$

Now the assertion can be seen as the dual formulation of Caratheodry's theorem (see [15], Theorem 2.3.5), which states that if a point x is in the convex hull of a set A in E^d, then there exists a subset A' of A such that $|A'| \leq d+1$ and x is in the convex hull of A'. □

Our existence proof for (ε, ν)-partition trees will use the interiors of regions such as those given in part (2) of the above lemma.

Definition. Let H^* be a set of open half-spaces in E^d. Then the *corridor* defined by H^*, denoted $corr(H^*)$, is the open region of the form

$$corr(H^*) = \bigcup_{h^* \in H^*} h^* - \bigcap_{h^* \in H^*} \bar{h}^*.$$

If $|H^*| \leq k$ then we call $corr(H^*)$ a *k-corridor.*

We obtain upper bounds on the Vapnik-Chervonenkis dimension of k-corridors in E^d from the following general result.

Lemma 4.5. *Assume $k \geq 1$ and (X, R) is a range space of dimension $d \geq 2$. Let R' be the set of all sets of the form $\bigcup_{i=1}^{k} r_i - \bigcap_{i=1}^{k} r_i$, where r_i is a range in R, $1 \leq i \leq k$. Then (X, R') has dimension less than $2dk \log(dk)$.*

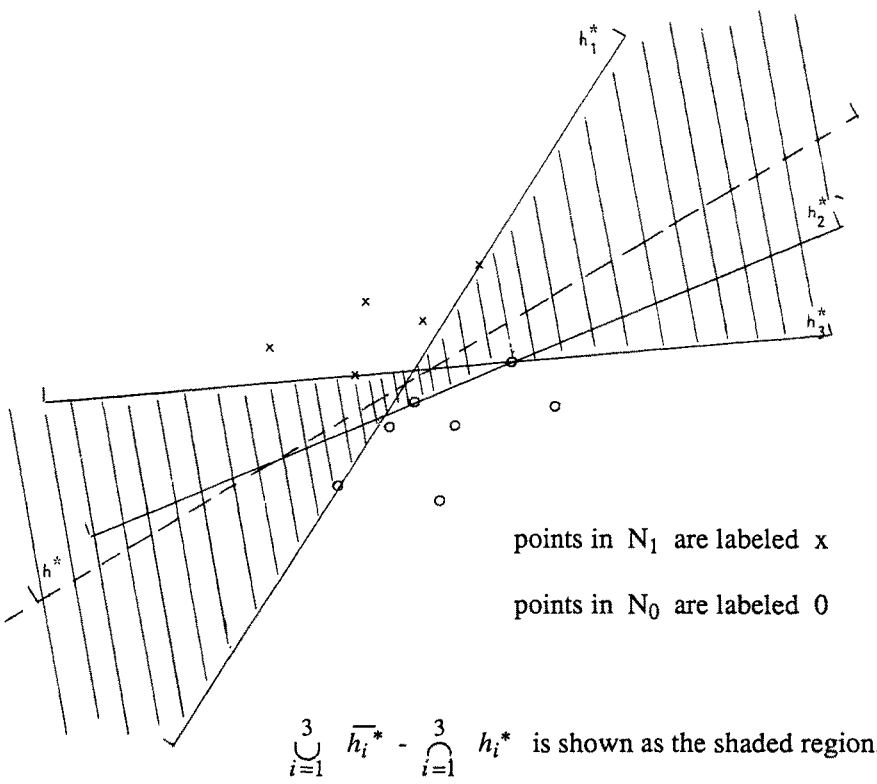

Fig. 1

Proof. Clearly we may assume that $k \geq 2$. Consider a finite set $A \subseteq X$ with $|A| = m \geq 3$. By Theorem 3.2, $|\Pi_R(A)| \leq \Phi_d(m)$. Every set in $\Pi_{R'}(A)$ is of the form $\bigcup_{i=1}^{k} r_i - \bigcap_{i=1}^{k} r_i$, with $r_i \in \Pi_R(A)$, $1 \leq i \leq k$. This shows that

$$|\Pi_{R'}(A)| \leq |\Pi_R(A)|^k \leq (\Phi_d(m))^k < m^{dk}.$$

Hence, if $m^{dk} \leq 2^m$, then A cannot be shattered by R' and the dimension of (X, R') is at most $m-1$. It is easy to show that $m^{dk} < 2^m$ for $dk \geq 4$ and $m = 2dk \log(dk)$. □

Corollary 4.6. *The Vapnik-Chervonenkis dimension of the set of all k-corridors in E^d is at most $2(d+1)k \log((d+1)k)$.*

Lemma 4.7. *Let A be a set of n points in general position in E^d, let N be an ε-net of A for $(d+1)$-corridors and let h^* be a half-space. If F is the set of all cells in the arrangement formed by $H_d(N)$ that are intersected by h, then*

$$\sum_{f \in F} |f \cap A| \leq \varepsilon n.$$

Proof. Assume first that there are points in N on both sides of h and let $h_1^*, h_2^*, \ldots, h_k^*$, where $k \leq d+1$, be half-spaces with $h_i \in H_d(N)$ as they are described in Lemma 4.4(2). Then h is contained in the closure of $C = \bigcup_{i=1}^{k} h_i^* - \bigcap_{i=1}^{k} \bar{h}_i^*$, and C contains all cells in the arrangement of $H_d(N)$ that are intersected by h. Since $C \cap N = \emptyset$ and N is an ε-net of A for $(d+1)$-corridors, $|C \cap A| \leq \varepsilon n$. If one side of h contains no points in N, the result is established by a similar reasoning using Lemma 4.4(1). □

Lemma 4.8. *For each $d \geq 2$ there is a constant $c_d > 0$ such that an (ε, ν)-partition tree exists for every ε and ν, where $0 < \varepsilon < 1$ and $\nu = \lfloor c_d(1/\varepsilon)\log(1/\varepsilon) \rfloor$, for every finite point set A in E^d in general position.*

Proof. Let T be a $(1, \nu)$-partition tree for a finite point set A in E^d and let $0 < \varepsilon < 1$. From Lemma 4.7 we know that if for all internal nodes p in T, *points*(p) is an ε-net of *set*(p) for $(d+1)$-corridors, then T is an (ε, ν)-partition tree for A. By Theorem 3.8 and Corollary 4.6, there exists a constant c_d for each $d \geq 0$, such that for all ε, $0 < \varepsilon < 1$, there exists an ε-net for $(d+1)$-corridors of size $\nu = \lfloor c_d(1/\varepsilon)\log(1/\varepsilon) \rfloor$. □

Theorem 4.9. *Let $\gamma > 0$ and $d \geq 2$ be fixed. For every set A of n points in E^d there exists an $O(n)$ storage data structure that supports the computation of $|A \cap h^*|$ for every half-space h^* in $O(n^\alpha)$ time with*

$$\alpha = \frac{d(d-1)}{d(d-1)+1} + \gamma.$$

(The constants for storage and time depend on γ and d.)

Proof. Consider an (ε, ν)-partition tree T for a finite point set A in E^d with $\nu = \lfloor c_d(1/\varepsilon)\log(1/\varepsilon) \rfloor$ where c_d is the constant from the above lemma. Linear storage complexity follows from Observation 4.2. By Lemma 4.3, the time complexity for a query in T is in $O(\nu^{d^2} n^{\alpha_\varepsilon})$ where

$$\alpha_\varepsilon = \frac{\log_{1/\varepsilon} \Phi_{d-1}\left(\binom{\nu}{d}\right)}{\log_{1/\varepsilon} \Phi_{d-1}\left(\binom{\nu}{d}\right) + 1} \leq \frac{\log_{1/\varepsilon} \nu^{d(d-1)}}{\log_{1/\varepsilon} \nu^{d(d-1)} + 1}$$

$$\leq \frac{d(d-1)\log_{1/\varepsilon} c_d(1/\varepsilon)\log(1/\varepsilon)}{d(d-1)\log_{1/\varepsilon} c_d(1/\varepsilon)\log(1/\varepsilon) + 1}$$

$$= \frac{d(d-1)(1 + \log_{1/\varepsilon} c_d + \log_{1/\varepsilon}\log(1/\varepsilon))}{d(d-1)(1 + \log_{1/\varepsilon} c_d + \log_{1/\varepsilon}\log(1/\varepsilon)) + 1}.$$

Now we observe that $\log_{1/\varepsilon} c_d \to 0$ and $\log_{1/\varepsilon}\log(1/\varepsilon) \to 0$ as $\varepsilon \to 0$. Thus for any $\gamma > 0$,

$$\alpha_\varepsilon \leq \frac{d(d-1)}{d(d-1)+1} + \gamma$$

for small enough ε. □

As far as construction of (ε, ν)-partition trees is concerned, we briefly outline a probabilistic algorithm analogous to the algorithm in [4] that runs in $O(n \log n)$ expected time:

Let A be a set of n points in E^d.
If $|A| < \nu$ then create a leaf, else:
(1) Create an internal node p.
(2) Repeatedly select a random sample *points*(p) of A of size ν until you get one such that property (4) of the definition of an (ε, ν)-partition tree is satisfied.
(3) Construct the arrangement of $H_d(\textit{points}(p))$.
(4) Recursively build the subtrees for all nonempty cells of this arrangement using only the points of A that lie in the cell.

If ν is chosen such that the probability of getting an ε-net for $(d+1)$-corridors is at least $\frac{1}{2}$ then property (4) is satisfied with probability at least $\frac{1}{2}$ for any random draw (by Lemma 4.7) and thus the expected number of random draws for each internal node is at most 2. By Corollaries 3.7 and 4.6, this requires only that $\nu \geq c(1/\varepsilon)\log(1/\varepsilon)$ for some constant c depending on d. For each random draw, we can test if *points*(p) satisfies property (4) by constructing the arrangement of $H_d(\textit{points}(p))$ and checking, for each set of cells in this arrangement that can be intersected by a hyperplane, whether or not this set of cells contains at most $\varepsilon|\textit{set}(p)|$ points. Since ν is constant, the number of such sets of cells is constant, hence the checking procedure takes $O(|\textit{set}(p)|)$ time. To complete the timing analysis, we need only observe that the depth of an (ε, ν)-partition tree is $O(\log n)$ and for each level of the tree, the union of all *set*(p) over all nodes p at that level contains at most n points. It follows that the total expected construction time is $O(n \log n)$.

The extension of the result for half-space counting to simplex counting can be seen as follows: a simplex s^* in E^d is defined by $d+1$ hyperplanes. Thus the boundary s of s^* is a subset of the union of $d+1$ hyperplanes. If we use now an (ε, ν)-partition tree to answer a query for simplex s^* in E^d in the obvious way, then we recur in each cell intersected by s, and these cells contain at most a total of $(d+1)\varepsilon$ times the current number of points. Just as in the case of half-spaces, the constant $d+1$ is absorbed as ε approaches 0.

Finally, we mention the implication of the above result to a purely geometric problem as it was raised by Edelsbrunner [10]. The known bounds to this problem are the same as for the half-space range counting problem, as these problems are closely related. Thus we have also improved the exponents in this result for all dimensions d, $d \geq 2$.

Theorem 4.10. *For each $d \geq 2$ and $\gamma > 0$ there exist constants k_1 and k_2 such that for every finite point set A in E^d there is a cell complex $C(A)$ that partitions E^d and has the following properties*:

(i) *Each cell of $C(A)$ is a d-dimensional convex polytope.*
(ii) *No cell contains a point of A in its interior.*

(iii) *The number of cells in* $C(A)$ *is at most* $k_1|A|$.
(iv) *The maximum number of cells in* $C(A)$ *intersected by an arbitrary hyperplane is at most* $k_2|A|^{\alpha}$, *where*

$$\alpha = \frac{d(d-1)}{d(d-1)+1} + \gamma.$$

Proof. Consider an (ε, ν)-partition tree T for A in E^d that realizes $O(n^{\alpha})$ query time. Then the regions $reg(p)$, p a leaf in T, form a cell complex that obviously satisfies (i), (iii), and (iv) for appropriate constants. However, since every region considered contains at most ν points, it is also easy to ensure property (ii). □

5. Open Problems

One outstanding problem is to determine where the function $f(d, \varepsilon)$, which bounds the size of ε-nets in terms of the dimension d of the range space, actually lies between the $\Omega(d/\varepsilon)$ and $O((d/\varepsilon)\log(d/\varepsilon))$ bounds given in Theorem 3.8. Furthermore, while the existence of "small" ε-nets for spaces of finite dimension plays a crucial role in our results, we give only probabilistic algorithms for constructing them. Efficient deterministic algorithms for constructing such nets remain to be determined. Here they might be some tradeoffs in size versus time of computation. Perhaps more significantly, we have yet to determine the size of the smallest ε-nets possible for many natural range spaces of finite dimension.[2]

Acknowledgment

We would like to thank Herbert Edelsbrunner and Raimund Seidel for valuable discussions of these ideas.

Appendix

The function we consider is defined by

$$Q(0) = 0$$

and for $n \geq 1$

$$Q(n) = c + \max \sum_{i=1}^{m} Q(n_i),$$

where the maximum is over all nonnegative integers $(n_1, \ldots, n_m)$ with $\sum_{i=1}^{m} n_i \leq \varepsilon n$, $c > 0$, $m \geq 2$, and $0 < \varepsilon < 1$.

We first claim that

$$Q(n+1) \leq Q(n) + c \qquad \text{for all} \quad n \geq 0. \tag{1}$$

[2] For half-planes in E^2 we can efficiently construct ε-nets of size $O(1/\varepsilon)$ for any finite point set. However, our construction does not generalize to higher dimensions.

Since $Q(0)=0$ and $Q(1)=c$, this is true for $n=0$. Now assume that it is true for all $0\le k\le n$, where $n\ge 1$. By definition of Q, there exist nonnegative integers $\bar{n}_1,\ldots,\bar{n}_m$ such that $\sum_{i=1}^{m}\bar{n}_i\le\varepsilon(n+1)$ and

$$Q(n+1)=c+\sum_{i=1}^{m}Q(\bar{n}_i).$$

If $\bar{n}_i=0$ for all i, $1\le i\le m$, then $Q(n+1)=c$. Hence, since Q is nonnegative, $Q(n+1)\le Q(n)+c$. Otherwise, there exists i such that $\bar{n}_i\neq 0$. Assume without loss of generality that $i=1$. Then

$$\begin{aligned}
Q(n+1)&=c+\sum_{i=1}^{m}Q(\bar{n}_i)\\
&\le c+Q(\bar{n}_1-1)+c+\sum_{i=2}^{m}Q(\bar{n}_i)\qquad\text{(by i.h.)}\\
&\le c+\left(c+\max\sum_{i=1}^{m}Q(n_i)\right)\\
&\qquad\text{where the maximum is over all nonnegative}\\
&\qquad n_1,\ldots,n_m \text{ such that } \sum_{i=1}^{m}n_i\le\varepsilon n\\
&\qquad\left(\text{since } \bar{n}_1-1+\sum_{i=2}^{m}\bar{n}_i\le\varepsilon(n+1)-1<\varepsilon n\right)\\
&=c+Q(n).
\end{aligned}$$

This establishes the claim. It follows from (1) that

$$Q(n+1)+Q(0)\le Q(n)+Q(1).$$

Hence, whenever we have $n_1,\ldots,n_m\ge 0$, and there exist i and j such that $n_i=0$ and $n_j\ge 2$, we can replace n_i with 1 and n_j with n_j-1 without decreasing $\sum_{i=1}^{m}Q(n_i)$. Thus if $\varepsilon n<m+1$, $\max\sum_{i=1}^{m}(n_i)$ over $n_1,\ldots,n_m$ such that $\sum_{i=1}^{m}n_i\le\varepsilon n$ is achieved for $n_1=1,\ldots,n_{\lfloor\varepsilon n\rfloor}=1$, $n_{\lfloor\varepsilon n\rfloor+1}=0,\ldots,n_m=0$. Thus

$$Q(n)=c(\lfloor\varepsilon n\rfloor+1)\qquad\text{for all } n,\quad 1\le n<\frac{m+1}{\varepsilon}. \tag{2}$$

We want to consider a slightly different function to obtain our final result, namely

$$\gamma(0)=0,$$

$$\gamma(n)=\frac{m}{\lceil m/\varepsilon\rceil-1}(n-1)+1\qquad\text{for}\quad 1\le n\le\left\lceil\frac{m}{\varepsilon}\right\rceil$$

$$\left(m \text{ and } \varepsilon \text{ are such that } \frac{m}{\varepsilon}>1\right),$$

$$\gamma(n)=1+\max\sum_{i=1}^{m}\gamma(n_i)\qquad\text{for}\quad n>\left\lceil\frac{m}{\varepsilon}\right\rceil,$$

where, as above, the maximum is over all nonnegative integers $(n_1, \ldots, n_m)$ with $\sum_{i=1}^{m} n_i \leq \varepsilon n$. First we show that

$$Q(n) \leq c(1+\varepsilon)\gamma(n) \qquad \text{for} \quad n \geq 0 \tag{3}$$

then

$$\gamma(n) \leq \frac{1}{m-1}(mn^{\alpha}-1) \quad \text{for} \quad n \geq 1, \qquad \text{where} \quad \alpha = \frac{\log_{1/\varepsilon} m}{\log_{1/\varepsilon} m + 1}. \tag{4}$$

Clearly (3) is true for $n=0$. Assume $1 \leq n \leq \lceil m/\varepsilon \rceil$. Then $\varepsilon n \leq \varepsilon \lceil m/\varepsilon \rceil < \varepsilon(m/\varepsilon+1) = m+\varepsilon < m+1$ so by (2) it suffices to show

$$c(\lfloor \varepsilon n \rfloor + 1) \leq c(1+\varepsilon)\left(\frac{m}{\lceil m/\varepsilon \rceil - 1}(n-1)+1\right).$$

However,

$$\begin{aligned}(1+\varepsilon)\left(\frac{m}{\lceil m/\varepsilon \rceil - 1}(n-1)+1\right) &\geq (1+\varepsilon)\left(\frac{m}{m/\varepsilon}(n-1)+1\right) \\ &= (1+\varepsilon)(\varepsilon(n-1)+1) = (1+\varepsilon)(\varepsilon n - \varepsilon + 1) \\ &= (\varepsilon n - \varepsilon + 1 + \varepsilon^2 n - \varepsilon^2 + \varepsilon) = (\varepsilon n + 1 + \varepsilon^2 n - \varepsilon^2) \\ &\geq \varepsilon n + 1 \geq \lfloor \varepsilon n \rfloor + 1,\end{aligned}$$

and the result follows.

For $n > \lceil m/\varepsilon \rceil$ we do a simple proof by induction.

Let $\bar{n}_1, \ldots, \bar{n}_m$ be nonnegative integers such that $Q(n) = c + \sum_{i=1}^{m} Q(\bar{n}_i)$ and $\sum_{i=1}^{m} \bar{n}_i \leq \varepsilon n$. Then

$$\begin{aligned}Q(n) &= c + \sum_{i=1}^{m} Q(\bar{n}_i) \leq c + c(1+\varepsilon) \sum_{i=1}^{m} \gamma(\bar{n}_i) \\ &\leq c + c(1+\varepsilon) \cdot \max \sum_{i=1}^{m} \gamma(n_i) \leq c(1+\varepsilon)\left(1 + \max \sum_{i=1}^{m} \gamma(n_i)\right) \\ &= c(1+\varepsilon)\gamma(n),\end{aligned}$$

where, as above, the maximum is over all nonnegative integers $(n_1, \ldots, n_m)$ with $\sum_{i=1}^{m} n_i \leq \varepsilon n$. This concludes the proof of inequality (3).

In order to prove (4) we define the positive real function

$$f(x) = \frac{1}{m-1}(mx^{\alpha}-1),$$

where $x > 0$ and α is as above. Observe that the first derivative $f'(x)$ is positive for all $x > 0$ and that the second derivative $f''(x)$ is negative for all $x > 0$. Hence $f(x)$ is a monotone increasing convex function.

First, we show that

$$\gamma(n) \leq f(n) \qquad \text{for all } n, \quad 1 \leq n \leq \left\lceil \frac{m}{\varepsilon} \right\rceil. \tag{5}$$

We have $\gamma(1)=1, f(1)=1$,

$$\gamma\left(\left\lceil\frac{m}{\varepsilon}\right\rceil\right)=m+1 \quad \text{and} \quad f\left(\left\lceil\frac{m}{\varepsilon}\right\rceil\right)\geq f\left(\frac{m}{\varepsilon}\right)=m+1.$$

The last equality is verified by observing that

$$\alpha=\frac{\log_{1/\varepsilon} m}{\log_{1/\varepsilon} m+1}=\frac{\log_{1/\varepsilon} m}{\log_{1/\varepsilon}(m/\varepsilon)}=\log_{m/\varepsilon} m,$$

and thus

$$f\left(\frac{m}{\varepsilon}\right)=\frac{1}{m-1}\left(m\left(\frac{m}{\varepsilon}\right)^{\alpha}-1\right)=\frac{1}{m-1}(m^2-1)=m+1.$$

Thus $\gamma(1)\leq f(1)$ and $\gamma(\lceil m/\varepsilon\rceil)\leq f(\lceil m/\varepsilon\rceil)$. Moreover, γ grows linearly in n in the range $1\leq n\leq\lceil m/\varepsilon\rceil$. Hence (5) follows by the convexity of $f(x)$.

Second, we observe that

$$\gamma(n+1)\leq\gamma(n)+1 \qquad \text{for all } n,\quad 0\leq n<\left\lceil\frac{m}{\varepsilon}\right\rceil. \tag{6}$$

This is true for $n=0$. For $n\geq 1$, we have to show

$$\gamma(n+1)=\frac{m}{\lceil m/\varepsilon\rceil-1}\cdot n+1\leq\frac{m}{\lceil m/\varepsilon\rceil-1}\cdot(n-1)+1+1=\gamma(n)+1$$

$$\Leftrightarrow\ \frac{m}{\lceil m/\varepsilon\rceil-1}\leq 1$$

$$\Leftrightarrow\ m\leq\left\lceil\frac{m}{\varepsilon}\right\rceil-1 \qquad \left(\text{since } \frac{m}{\varepsilon}>1\right)$$

$$\Leftrightarrow\ m+1\leq\left\lceil\frac{m}{\varepsilon}\right\rceil \qquad \text{which holds, because } \varepsilon<1.$$

Third, we prove

$$\gamma(n+1)\leq\gamma(n)+1 \qquad \text{for} \quad n=\left\lceil\frac{m}{\varepsilon}\right\rceil. \tag{7}$$

First, we recall that $\gamma(\lceil m/\varepsilon\rceil)=m+1$. So we have to show that $\gamma(\lceil m/\varepsilon\rceil+1)\leq m+2$. To this end we first observe that

$$m\leq\left\lfloor\varepsilon\left(\left\lceil\frac{m}{\varepsilon}\right\rceil+1\right)\right\rfloor\leq m+1;$$

$$m=\lfloor m+\varepsilon\rfloor=\left\lfloor\varepsilon\left(\frac{m}{\varepsilon}+1\right)\right\rfloor\leq\left\lfloor\varepsilon\left(\left\lceil\frac{m}{\varepsilon}\right\rceil+1\right)\right\rfloor\leq\left\lfloor\varepsilon\left(\frac{m}{\varepsilon}+2\right)\right\rfloor$$

$$=\lfloor m+2\varepsilon\rfloor\leq m+1.$$

Hence, there exist $\bar{n}_1,\ldots,\bar{n}_m$ such that

$$\gamma\left(\left\lceil\frac{m}{\varepsilon}\right\rceil+1\right)=1+\sum_{i=1}^{m}\gamma(\bar{n}_i),$$

where $\sum_{i=1}^{m} \bar{n}_i = m$ or $\sum_{i=1}^{m} \bar{n}_i = m+1$ (because $\gamma(n) \leq \gamma(n+1)$) and $\bar{n}_i \neq 0$ for all i (because $\gamma(n+1)+\gamma(0) \leq \gamma(n)+\gamma(1)$ for $0 \leq n < \lceil m/\varepsilon \rceil$ by (6)). Thus either all $\bar{n}_i = 1$, whence $\gamma(\lceil m/\varepsilon \rceil + 1) = m+1$, or all $\bar{n}_i = 1$ but one, which is 2, whence

$$\gamma\left(\left\lceil \frac{m}{\varepsilon} \right\rceil + 1\right) = 1 + (m-1) + \left(\frac{m}{\lceil m/\varepsilon \rceil - 1} \cdot (1) + 1\right)$$

$$= m + 1 + \frac{m}{\lceil m/\varepsilon \rceil - 1} \leq m+2$$

as in the verification of equation (6). As a consequence, (7) holds.

We can now extend (6) and (7) by induction to

$$\gamma(n+1) \leq \gamma(n) + 1 \qquad \text{for all} \quad n \geq 0, \tag{8}$$

using (6) and (7) as a basis case and proceeding as in the verification of (1) above.

We conclude by showing

$$\gamma(n) \leq \frac{1}{m-1}(mn^{\alpha} - 1) \qquad \text{for} \quad n \geq 1.$$

By (5), we may assume that $n > \lceil m/\varepsilon \rceil$. Then by (8), for some $\bar{n}_1, \ldots, \bar{n}_m$, $\bar{n}_i \neq 0$, $1 \leq i \leq m$, where $\sum_{i=1}^{m} \bar{n}_i \leq \varepsilon n$ we have

$$\gamma(n) = 1 + \sum_{i=1}^{m} \gamma(\bar{n}_i) \leq 1 + \sum_{i=1}^{m} \frac{1}{m-1}(m\bar{n}_i^{\alpha} - 1) \qquad \text{(by i.h.)}$$

$$= 1 + \frac{m}{m-1} \cdot \sum_{i=1}^{m} \bar{n}_i^{\alpha} - \frac{m}{m-1} = 1 + \frac{m^2}{m-1} \cdot \frac{\sum_{i=1}^{m} \bar{n}_i^{\alpha}}{m} - \frac{m}{m-1}$$

$$\leq 1 + \frac{m^2}{m-1} \cdot \left(\frac{\sum_{i=1}^{m} \bar{n}_i}{m}\right)^{\alpha} - \frac{m}{m-1} \qquad \text{(since } \alpha \leq 1\text{)}$$

$$\leq 1 + \frac{m^2}{m-1} \cdot \left(\frac{\sum_{i=1}^{m} \bar{n}_i}{m}\right)^{\alpha} - \frac{m}{m-1} \qquad \text{(since } \alpha \leq 1\text{)}$$

$$\leq 1 + \frac{m^2}{m-1} \cdot \left(\frac{\varepsilon n}{m}\right)^{\alpha} - \frac{m}{m-1}$$

$$= 1 + \frac{m}{m-1} \cdot \left(m\left(\frac{\varepsilon n}{m}\right)^{\alpha}\right) - \frac{m}{m-1}$$

$$= \frac{m}{m-1} \cdot (n^{\alpha}) - \frac{m}{m-1} + 1 \qquad \text{(by def. of } \alpha\text{)}$$

$$= \frac{m}{m-1}(n^{\alpha}) - \frac{1}{m-1} = \frac{1}{m-1} \cdot (mn^{\alpha} - 1).$$

This proves (4) by induction.

We now summarize (3) and (4) as

$$Q(n) \leq \frac{c \cdot (1+\varepsilon)}{m-1} \cdot (mn^{\alpha} - 1)$$

$$\leq \frac{c(1+\varepsilon)}{m-1} \cdot mn^{\alpha} = c \cdot (1+\varepsilon) \cdot \left(\frac{m}{m-1}\right) \cdot n^{\alpha} \leq c \cdot 2 \cdot 2 \cdot n^{\alpha} = 4cn^{\alpha},$$

Hence $Q(n)$ is $O(cn^{\alpha})$.

References

1. P. Assouad, Densite et dimension, *Ann. Inst. Fourier (Grenoble)* **33** (1983), 233-282.
2. A. Blumer, A. Ehrenfeucht, D. Haussler, and M. Warmuth, Classifying learnable geometric concepts with the Vapnik-Chervonenkis dimension, *Proceedings of the 18th Symposium on Theory of Computation*, 273-282, 1986.
3. B. Chazelle, L. Guibas, and D. T. Lee, The power of geometric duality, *Proceedings of the 24th Symposium on Foundations of Computer Science*, 217-225, 1983.
4. K. Clarkson, "A probabilistic algorithm for the post office problem," *Proceedings of the 17th Symposium on Theory of Computation*, 175-185, 1985.
5. K. Clarkson, Further applications of random sampling to computational geometry, *Proceedings of the 18th Symposium on Theory of Computation*, 414-423, 1986.
6. R. Cole, Partitioning point sets in 4 dimensions, *Proceedings of the Colloquium on Automata, Language and Programming*, 111-119, Lecture Notes in Computer Science 194, Springer-Vérlag, Berlin, 1985.
7. D. Dobkin and H. Edelsbrunner, Organizing Points in Two and Three Dimensions, Technical Report F130, Technische Universitat Graz, 1984.
8. D. Dobkin, H. Edelsbrunner, and F. Yao, A 3-space partition and its applications, manuscript.
9. R. M. Dudley, Central limit theorems for empirical measures, *Ann. Probab.* **6** (1978), 899-929.
10. H. Edelsbrunner, Problem P110, *Bull. EATCS* **26** (1985), 239.
11. H. Edelsbrunner, *Algorithms in Combinatorial Geometry*, EATCS Monographs in Theoretical Computer Science, Springer-Verlag, Berlin, 1987.
12. H. Edelsbrunner and F. Huber, Dissecting Sets of Points in Two and Three-Dimensions, Technical Report F138, Technische Universitat Graz, 1984.
13. H. Edelsbrunner and E. Welzl, Constructing belts in two-dimensional arrangements with applications, *SIAM J. Comput.* **15** (1986), 271-284.
14. H. Edelsbrunner and E. Welzl, Halfplanar range search in linear space and $O(n^{0.695})$ query time, *Inform. Process. Lett.*, to appear.
15. Branko Gruenbaum, *Convex Polytopes*, Interscience, New York, 1967.
16. N. Sauer, On the density of families of sets, *J. Combin. Theory Ser. A* **13** (1972), 145-147.
17. V. N. Vapnik and A. Ya. Chervonenkis, On the uniform convergence of relative frequencies of events to their probabilities, *Theory Probab. Appl.* **16** (1971), 264-280.
18. V. N. Vapnik and A. Ya. Chervonenkis, *The Theory of Pattern Recognition*, Nauka, Moscow, 1974.
19. R. S. Wenocur and R. M. Dudley, Some special Vapnik-Chervonenkis classes, *Discrete Math.* **33** (1981), 313-318.
20. D. Willard, Polygon retrieval, *SIAM J. Comput.* **11** (1982), 149-165.
21. I. M. Yaglom and V. G. Bolyansky, *Convex Figures*, Holt, Rinehart and Winston, New York, 1961 (transl.).
22. F. Yao, A 3-space partition and it applications, *Proceedings of the 15th Symposium on Theory of Computation*, 258-263, 1983.
23. A. Yao and F. Yao, A general approach to d-dimensional geometric queries, *Proceedings of the 17th Symposium on Theory of Computation*, 163-169, 1985.

Received April, 1986, and in revised form November 21, 1986.