LOWER BOUNDS ON THE COMPLEXITY OF POLYTOPE RANGE SEARCHING

Bernard Chazelle

CS-TR-166-88

June 1988

Lower Bounds on the Complexity of Polytope Range Searching

Bernard Chazelle

Department of Computer Science
Princeton University
Princeton, NJ 08544

Abstract: Polytope range searching is a central problem in multimensional searching, with applications to computer graphics, robotics, and database design. In its most elementary form, the problem can be stated as follows: Given a collection P of n weighted points in Euclidean d-space and a simplex q, compute the cumulative weight of $P \cap q$. The points are given once and for all and can be preprocessed. The simplex q, however, forms a query which must be answered on-line. We assume that the weights are chosen in a commutative semigroup and that the time to answer a query includes only the number of arithmetic operations performed by the algorithm. We prove that if m units of storage are available then the worst-case query time is $\Omega(n/\sqrt{m})$ in 2-space, and more generally, $\Omega((n/\log n)/m^{1/d})$ in d-space, if $d \geq 3$. These bounds also hold with high probability for a random set of points (drawn uniformly in the d-cube) and remains true if the queries are restricted to congruent copies of a fixed simplex. In the course of our investigation we also establish results of independent interest regarding a generalization of Heilbronn's problem.

A preliminary version of this work has appeared in the proceedings of the 28th Annual IEEE Symposium on Foundations of Computer Science (1987), 1–10.

The author wishes to acknowledge the National Science Foundation for supporting this research in part under Grant CCR–8700917.

1. Introduction

Orthogonal range searching and simplex range searching have received a great amount of attention recently. But whereas the former problem is nearing a definitive solution, the complexity of simplex range searching has remained elusive. To state the problem simply, suppose that we are given n points in Euclidean d-space and m units of computer memory. How should we organize the memory to be in a position to answer the following type of queries efficiently: Given an arbitrary simplex, how many points lie inside? Natural variants of the problem call for reporting the points in question, or more generally, computing some useful function defined on their power-set. There is abundant practical application to motivate research on this problem [5,7,11,12,16,19,21,23]. Also of great interest is the central theoretical question lying underneath: What is the most efficient way of organizing information to support a given class of queries? What takes this question apart from the classical problem of searching a linear list is the power of redundancy. While oversupply of memory space is usually of marginal interest when searching a linear list, it is often the key to efficiency in multidimensional searching. For this reason, the main research activity in that area has been the investigation of space-time tradeoffs.

The main result of this paper is a family of lower bounds on the space-time complexity of simplex range searching. We prove that the worst-case query time is $\Omega(n/\sqrt{m})$, for 2-space, and more generally, $\Omega((n/\log n)/m^{1/d})$ in d-space, for $d \geq 3$. Recall that n is the number of points and m is the amount of storage available. These bounds hold with high probability for a random point-set (from a uniform distribution in the d-cube) and thus are valid in the worst case as well as on the average. Interestingly, they still hold if the queries are restricted to congruent copies of a fixed simplex or even a fixed slab.

What is the practical significance of these lower bounds? It appears that little gain in query time can be expected unless the storage is enormous. In practice, therefore, the naive algorithm—which involves checking each of the n points individually—is probably the method of choice. To make our point a little more evident, consider the complexity of simplex range searching in 11-space. With only linear storage the worst-case query time is in $\Omega(n^{0.9})$. For a query time in $O(\sqrt{n})$, one would need $\Omega(n^5)$ storage, and a whopping $\Omega(n^{10})$, if a polylogarithmic query time were desired. Furthermore, our average-case result shows that the lower bound is hardly determined by some pathological input configuration but rather by random point-sets.

Our complexity results are established in the *arithmetic model* for range searching (Fredman [10,11], Yao [20], Chazelle [4]). The assumptions of the model are very weak and any lower bound in it can be trusted to hold on any reasonable sequential machine (which, in particular, allows bucketing, hashing, etc.) How close do our lower bounds come to meeting known upper bounds? It has been shown in (Chazelle and Welzl [5]) that simplex range searching on n points in d-space can be performed in $O(n^{1-1/d}\alpha(n))$ query time and $O(n)$ storage, where α is a functional inverse of Ackermann's function. This upper bound matches our lower bound very closely. It must be mentioned, however, that the upper bound holds in the arithmetic model. Its main interest, therefore, is to tell us that to obtain significantly higher lower bounds one will have to change the model of computation. On a random access machine [2] supplied with linear storage, the best upper bound on the query time to date is $O(\sqrt{n}\log n)$ in 2-space (Chazelle and Welzl [5]) and $O(n^{d(d-1)/(d(d-1)+1)+\varepsilon})$ in d-space, for any $d \geq 3$ and any fixed $\varepsilon > 0$ (Haussler and Welzl [12]). Earlier results were obtained in (Willard [19], Edelsbrunner and Welzl [7]). See also (Cole and Yap [6]) for variants of the problem.

Our results constitute the first (nontrivial) family of lower bounds for simplex range searching in the static case. These complement an earlier lower bound for the dynamic version of the problem: Fredman [11] established that a sequence of n insertions, deletions, and halfplane range queries may require $\Omega(n^{4/3})$ time. His ingenious proof technique rests on the fact that a single deletion may invalidate a large segment of the data structure: indeed, any precomputed cumulative weight which involves a point to be deleted becomes useless after the deletion since a semigroup has no inverse and no quick update is therefore possible. Obviously, we must use a different line of reasoning.†

The main novelty of our approach is to reduce space-time tradeoffs for range searching to fundamental inequalities in integral geometry. To achieve this goal we need a fairly heavy machinery which we build in three principal stages: first, we define a model for static range searching (section 2) which places the problem within the scope of bipartite Ramsey theory (section 3). The complexity of a given problem is then fully described by certain properties of its so-called *characteristic graph.* This involves two distinct tasks: proving integral-geometric inequalities about the query space (section 4.2) and studying various uniformity criteria for random point-sets (section 4.3). Incidentally, these investigations lead to results of independent interest regarding an intriguing generalization of Heilbronn's problem (Moser [17]). The lower bounds for simplex range searching are established in section 4.4.

2. A Combinatorial Framework

We describe a graph-theoretic model for range searching. The emphasis of this model is the arithmetic complexity of a problem, that is, the maximum number of operations needed to answer any query. The model purposely ignores the cost of searching the memory for the information needed during the computation. In this way, lower bounds can be trusted to hold on any sequential computer. Of course, from a practical viewpoint, upper bounds set in that model may not necessarily have much meaning, except to indicate how good or how bad a certain lower bound might be. The *arithmetic model*—as it is customarily called—originates in (Fredman [10,11]) for the dynamic case and (Yao [20]) for the static case.

† Interestingly, we can use our lower bound to strengthen Fredman's result by removing the use of deletions. We can prove that a sequence of n insertions, followed by n queries may require on the order of $n^{4/3}$ time in the worst case. To see this, insert the n points used to prove our static lower bound and then ask the hardest query repeatedly, n times. If less than $n^{4/3}$ time is required for all these operations then the storage used (in the model chosen) is $O(n^{4/3})$, therefore from our lower bound each query takes $\Omega(n/\sqrt{n^{4/3}})$ time, hence a total of $\Omega(n^{4/3})$ time. ∎

The main purpose of this section is to introduce a general technique for proving lower bounds (the *Core Lemma*). The basic idea is to relate the static complexity of a range searching problem to the existence of large complete bipartite subgraphs in its ***characteristic graph***. This graph is similar (although not identical) to what Burkhard et al. [3] call the ***semantic graph***: it provides a combinatorial characterization of any range searching problem.

A. Some terminology. In the following, $\mathcal{N}$ will denote the set of natural numbers $\{0,1,2,\ldots\}$, and for any integer $n > 0$, $[1\ldots n]$ will be the set $\{1,2,\ldots,n\}$. We use $\mathcal{C}_d$ to denote the d-cube $[0,1]^d$. As a shorthand, we say that a finite set of points P in a compact set K is ***random in*** K if each of its points has been drawn randomly from a uniform distribution in K (we assume mutual independence). In general, K will be $\mathcal{C}_d$. Finally, we introduce the notion of a ***faithful semigroup*** (Yao [20]). Let $(S,+)$ be a commutative semigroup with an operation denoted $+$. We say that $(S,+)$ is ***faithful*** if for each $n > 0$, $\emptyset \subset T_1, T_2 \subseteq [1\ldots n]$, $T_1 \neq T_2$, and every sequence of integers $\alpha_i, \beta_j > 0$ $(i \in T_1,\ j \in T_2)$, the equation

$$\sum_{i\in T_1} \alpha_i s_i = \sum_{j \in T_2} \beta_j s_j$$

is not an identity (that is, cannot be satisfied for all assignments of the variables $s_1, \ldots, s_n$). Note that this definition does not prohibit idempotence or identities of the form $s_1 + 2s_2 = 3s_1 + 4s_2$. For example, $(\mathcal{N},+)$, $(\mathcal{N},\max)$, and $(\{0,1\}, or)$ are faithful, but $(\{0\}, or)$ and $(\{0,1\}, \textit{exclusive or})$ are not.

B. Range searching. Let $(S,+)$ be a faithful commutative semigroup. We define a ***query space*** $\mathcal{Q}$ to be any collection (finite or infinite) of subsets $q \subseteq \Re^d$, called ***queries***. For example, $\mathcal{Q}$ might be the set of all hyperrectangles, simplices, balls in Euclidean d-space, etc. Let $P = \{p_1, \ldots, p_n\}$ be a set of n points in $\mathcal{C}_d$, and let ζ be an assignment of each point p to a semigroup value in S. We define a function $answ\colon \mathcal{Q} \rightarrow S$ as follows:

$$answ(q) = \sum_{p \in P\cap q} \zeta(p).$$

If $P \cap q$ is empty then we write $answ(q) = null$, which is a special symbol *not* in S. In practice, the semigroup can be chosen as $(\mathcal{N},+)$ for counting the number of points in the desired query, $(2^P, \cup)$ for reporting the points in question, $(\{0,1\}, or)$ for testing if there are any points in the query, etc. To summarize, a range searching problem $\mathcal{P}$ is specified by a quadruple $(S, \mathcal{Q}, P, \zeta)$ consisting of a semigroup, a query space, a finite set of points, and a weight function. We say that $\mathcal{P}$ is of ***size*** (n,p), if $|P| = n$ and $|\{\, P \cap q \mid q \in \mathcal{Q}\,\}| = p$.

C. The model of computation. Let $s_1, \ldots, s_n$ be n variables with values in S. A *generator* $g(s_1, \ldots, s_n)$ is a linear form $\sum_{1 \le i \le n} \alpha_i s_i$, where the α_i's are nonnegative integers (not all 0). For example, $2s_1 + 0s_2 + s_3$ stands for $s_1 + s_1 + s_3$. A *storage scheme* Γ for $\mathcal{P}$ of *size* m is a collection of m generators $\{g_1, \ldots, g_m\}$ satisfying the following property: for any $q \in \mathcal{Q}$ such that $P \cap q \neq \emptyset$, there exist $K \subseteq [1 \ldots m]$ and a set of labeled integers $\{\, \beta_k > 0 \mid k \in K \,\}$ such that the relation

$$answ(q) = \sum_{k \in K} \beta_k g_k(\zeta(p_1), \ldots, \zeta(p_n)) \tag{2.1}$$

holds for *any* weight function ζ over P. This means that a storage scheme can be dependent on the particular semigroup under consideration and also take advantage of any property which P may enjoy: however, it must hold for any assignment of semigroup values to P.

Ideally, we would like S to be rich enough to simulate (i.e., to map homomorphically onto) the semigroup $(P^*, \cup)$ of all nonempty subsets of P. But this would exclude too many important semigroups, so we move this requirement over to the storage scheme. By insisting that a scheme should work for all weight assignments, we are in effect no longer dealing with S itself but with the additive semigroup of n-variate linear forms over S. Faithfulness can then be called upon to ensure that the semigroup of linear forms is, indeed, rich enough. Given a linear form $\sum_{1 \le i \le n} \alpha_i s_i$, call the set of points $\{\, p_i \mid \alpha_i \neq 0 \,\}$ its *cluster*.† By means of this correspondence, the semigroup generated by the elementary forms $(s_1, \ldots, s_n) \mapsto s_i$ $(1 \le i \le n)$ maps homomorphically onto the semigroup $(P^*, \cup)$. Thus, the meaning of (2.1) is that any set of the form $P \cap q$ can be expressed as a union of clusters. Note that this union need not be disjoint. Of course, the irrelevance of the weight function allows us to say that a storage scheme is defined not only for $\mathcal{P}$, but more generally, for $(S, \mathcal{Q}, P)$.

Next, we define the complexity of a storage scheme. Given $q \in \mathcal{Q}$, let K be the smallest set such that (2.1) is true. We define $t(P, \Gamma, q) = |K|$, and we say that Γ is a (t, m)-scheme for $\mathcal{P}$, if $t \ge \max_{q \in Q} t(P, \Gamma, q)$. If $\mathcal{P}$ is now considered as one element in an infinite family (as P and n vary), we define the time complexity of this family as the function $t(n, m)$, where

$$t(n, m) = \max_{|P| = n} \min_{|\Gamma| = m} \max_{q \in \mathcal{Q}} t(P, \Gamma, q).$$

By abuse of notation, we will refer to $t(n, m)$ as the time complexity of $\mathcal{P}$ (when the notion of a family is understood). We also define the expected time complexity of $\mathcal{P}$ as

$$\bar{t}(n, m) = E_{|P| = n} \min_{|\Gamma| = m} \max_{q \in \mathcal{Q}} t(P, \Gamma, q),$$

where P is random in $\mathcal{C}_d$. We do not average over $\mathcal{Q}$ because the query space cannot always be assumed to admit a natural probability measure.

† To make this definition independent of the fact that S is faithful, we should regard a cluster as being associated with a formal linear form (no pun intended). Otherwise, problems arise if S is not faithful and a linear form can be expressed over two different sets of variables.

D. The Graph model. We begin with some terminology. Let $H \subseteq V \times W$ be a bipartite graph. We denote the number of edges of H by $|H|$. For any $w \in W$, let $N_H(w)$ denote the set $\{v \in V \mid (v,w) \in H\}$. By extension, if $U \subseteq W$ then $N_H(U) = \bigcup_{w \in U} N_H(w)$. A bipartite graph $C \subseteq V \times Z$ is called a *cover* of H if for every $w \in W$ there exists a subset $Z_w \subseteq Z$ such that $N_H(w) = N_C(Z_w)$. A subset Z_w of minimum size is called a *min-cover* of w. If the cardinality of no min-cover exceeds t, we say that C is a $(t, |Z|)$-cover of H. The graph C is called a *disjoint* $(t, |Z|)$-cover if

(i) for each $w \in W$, there exists $Z_w \subseteq Z$ such that all the sets in $\{ N_C(z) \mid z \in Z_w \}$ are pairwise disjoint, and

(ii) the maximum value, over all $w \in W$, of the size of the smallest Z_w satisfying (i) does not exceed t.

In light of our discussion of generators and clusters the meaning of all this should be obvious. We can use H to model a range searching problem, with the V-nodes acting as points and the W-nodes as queries. A cover C corresponds to a storage scheme, with the Z-nodes acting as generators and the $N_C(z)$'s ($z \in Z$) as clusters. To conclude this string of definitions, we refer to a *rectangle* of H as any complete bipartite subgraph $V' \times W' \subseteq H$: the *width* and *height* of the rectangle are respectively $|V'|$ and $|W'|$.

Given a range searching problem $\mathcal{P} = (S, \mathcal{Q}, P, \zeta)$ of size (n,p), the set $\{ P \cap q \mid q \in \mathcal{Q} \}$ partitions $\mathcal{Q}$ into p equivalence classes. Let $q_1, \ldots, q_p$ be representatives of each class. We define the *characteristic graph* of $\mathcal{P}$ as a bipartite graph $H \subseteq V \times W$, where $V = \{v_1, \ldots, v_n\}$ and $W = \{w_1, \ldots, w_p\}$, and an edge connects v_i and w_j if and only if $p_i \in q_j$. The graph fully describes the combinatorial nature of the range searching problem in question. Conversely, it is easily seen that any bipartite graph can be regarded as the characteristic graph of some range searching problem: for example, P may consist of n distinct points in the plane and $\mathcal{Q}$ is chosen as the set of polygonal curves. We are now in a position to formalize the relationship between schemes and covers, and present a methodology for proving space-time tradeoffs. Generators are to schemes what clusters are to covers. The following lemma uses faithfulness to establish that link.

Lemma 2.1. *Let $\mathcal{P}$ be a range searching problem and let H be its characteristic graph. If $\mathcal{P}$ admits a (t,m)-scheme then H admits a (t,m)-cover. Conversely, if H admits a disjoint (t,m)-cover then $\mathcal{P}$ admits a (t,m)-scheme.*

Proof: Suppose that $\mathcal{P}$ admits a (t,m)-scheme, and let $\Gamma = \{g_1, \ldots, g_m\}$ be the storage scheme in question. If $g_k(s_1, \ldots, s_n) = \sum_{1 \le i \le n} \alpha_{k,i} s_i$, we can rewrite this linear form as $\sum_{i \in N_k} \alpha_{k,i} s_i$, where $N_k = \{ i \mid \alpha_{k,i} > 0 \}$. We define a bipartite graph $C \subseteq V \times Z$, where $Z = \{z_1, \ldots, z_m\}$, by placing an edge between v_i and z_j if and only if $i \in N_j$. We now show that C is a (t,m)-cover of H. Let w_j be an arbitrary vertex of W and let $A_j = \{ i \mid p_i \in q_j \}$. It suffices to establish the existence of $Z_j \subseteq Z$, where $|Z_j| \le t$ and

$$N_H(w_j) = N_C(Z_j). \tag{2.2}$$

Since Γ is a (t,m)-scheme for P we have

$$answ(q_j) = \sum_{k \in B_j} \beta_k g_k(\zeta(p_1), \ldots, \zeta(p_n)),$$

where B_j is a subset of $[1 \ldots m]$ of size $\leq t$. Since the equality above holds for any weight function and S is a commutative semigroup, we can write

$$\sum_{i \in A_j} s_i = \sum_{k \in B_j} \beta_k g_k(s_1, \ldots, s_n) = \sum_{k \in C_j} \gamma_k s_k,$$

where $C_j = \bigcup_{k \in B_j} N_k$. Because of faithfulness we have $A_j = C_j$, therefore

$$N_H(w_j) = \{ v_i \mid i \in A_j \} = \bigcup_{k \in B_j} \{ v_i \mid i \in N_k \} = \bigcup_{k \in B_j} N_C(z_k),$$

which establishes (2.2), since $|B_j| \leq t$. The first part of the lemma is now proven. We omit the second part, which is straightforward. ∎

Now that range searching problems have been couched as combinatorial questions about bipartite graphs, we are ready to describe the lower bound proof technique which underlies much of what follows. Although the technique tends to weaken somewhat on problems of low complexity (e.g., orthogonal range queries), it is, we believe, a powerful tool for determining the complexity of "hard" problems, such as simplex range searching or problems defined by random characteristic graphs. The starting point is the observation that, informally, clusters are "good" if they are big and can be used by many representative queries. Translated in the language of covers, this means that the characteristic graph contains big rectangles: their widths tell us how big the clusters can be and their heights indicate how many representative queries they can help to answer.

The following result formalizes the relationship between the space complexity of a range searching problem $\mathcal{P}$ and the presence of large rectangles in its characteristic graph H. We define $\mathcal{A}(x)$ to be the largest "area" of a rectangle of H whose width is no less than $x > 0$:

$$\mathcal{A}(x) = \max \{ xh \mid H \text{ has a rectangle of width} \geq x \text{ and height } h \}.$$

Lemma 2.2. (The Core Lemma) – *Let H be the characteristic graph of a range searching problem of size (n,p). If H has a (t,m)-cover then $m \geq \frac{1}{2}|H|/\mathcal{A}\left(\frac{|H|}{2pt}\right)$.*

Proof: Using the previous notation, let $C \subseteq V \times Z$ be a (t,m)-cover of H and let $\{ C_w \subseteq Z \mid w \in W \}$ be a complete collection of min-covers. Form the graph G by removing from H each edge in the set

$$\bigcup \left\{ N_C(z) \times \{w\} \;\middle|\; w \in W,\, z \in C_w,\, |N_C(z)| \leq \frac{|H|}{2pt} \right\}.$$

Since C is a (t,m)-scheme and $|W| = p$, the resulting graph G contains at least half the edges of H. But to cover the sets $N_G(w)$ $(w \in W)$ only subsets $N_C(z)$ of size $> \frac{|H|}{2pt}$ are now used. Therefore Z must have at least $|G|/\mathcal{A}\left(\frac{|H|}{2pt}\right)$ vertices. ∎

3. How Hard Can Range Searching Be?

Any range searching problem of size (n,p) admits two trivial solutions: an (n,n)-scheme and a $(1,p)$-scheme. Two natural questions arise: (1) Is it always possible to improve over the two naive solutions? (2) What is the complexity of the *hardest* range searching problem? Answering these questions will help us assess the relative position of other range searching problems on the complexity ladder. Theorem 3.1 says that a small speed-up in query time can always be achieved with an amount of storage almost but not quite maximum: in other words, the worst of all possible worlds. Surprisingly, this result is in fact optimal, as a Ramsey-type argument (Theorem 3.2) can be used to show.

Theorem 3.1. *For any range searching problem of size (n,p), with $p > n$, there exists a (t,m)-scheme, where $t = O\left(n/\left\lceil\log\frac{p}{n}\right\rceil\right)$† and $m = O\left(p/\left\lceil\log\frac{p}{n}\right\rceil\right)$.*

Proof: We follow a strategy used in (Yao and Yao [22], Burkhard et al. [3]). Let $\alpha = \left\lceil\log\frac{p}{n}\right\rceil$ and, as usual, let $H \subseteq V \times W$ denote the characteristic graph of the range searching problem, with $V = \{v_1, \ldots, v_n\}$. For each i such that $0 \le i \le \lfloor(n-1)/\alpha\rfloor$, define V_i as the set $\{v_{i\alpha+1}, \ldots, v_{\min(n,(i+1)\alpha)}\}$. We construct a cover $C \subseteq V \times Z$ as follows. Originally, Z is empty: for each i between 0 and $\lfloor(n-1)/\alpha\rfloor$, consider each nonempty subset A of V_i in turn, and perform the following operations: add a new vertex z to Z and augment C with the edges of $A \times \{z\}$. It is easily verified that C is a disjoint (t,m)-cover of H, where $t \le 1 + \lfloor(n-1)/\alpha\rfloor$, and

$$m \le (2^\alpha - 1)\left(1 + \lfloor(n-1)/\alpha\rfloor\right).$$

Since $p \le 2^n$ we easily derive that $t = O\left(n/\left\lceil\log\frac{p}{n}\right\rceil\right)$ and $m = O\left(p/\left\lceil\log\frac{p}{n}\right\rceil\right)$. Lemma 2.1 completes the proof. ∎

Theorem 3.2. *There is a constant $c > 0$ such that the following is true. Given any integer function $p = p(n)$ $(n < p \le 2^n)$ there exists a class of range searching problems of size (n,p) for which any (t,m)-scheme with $t \le cn/\log\frac{p}{n}$ also satisfies $m = \Omega(p)$.*

Proof: Let π be a real $(0 < \pi < 1)$, and let $H \subseteq V \times W$ be a random bipartite graph ($|V| = n$ and $|W| = p$), where each edge (v,w) is chosen independently with probability π. A rectangle of H is called *wide* if its width α is at least $\ln(p/n)$ and its height is equal to $\lceil n/\alpha\rceil$. To rid the graph of wide rectangles we use a standard technique for removing forbidden subsystems [8]. Let $\chi(H)$ be the number of wide rectangles in H. We modify H by taking each wide rectangle in turn, and removing exactly one edge from it (which one does not matter). After at most $\chi(H)$ such operations

† All logarithms in this paper are taken to the base 2, unless specified otherwise.

we obtain a new graph G free of wide rectangles, with $|G| \geq |H| - \chi(H)$. Taking expectations we derive

$$E(|G|) \geq np\pi - \sum_{\ln(p/n) \leq \alpha \leq n} \binom{n}{\alpha} \binom{p}{\lceil n/\alpha \rceil} \pi^{\alpha \lceil n/\alpha \rceil}.$$

Using the inequalities $\binom{b}{a} < (eb/a)^a$, for $0 < a \leq b$, and $(n/\alpha)^\alpha \leq e^{n/e}$, for $1 \leq \alpha \leq n$, where $e = 2.718\ldots$, we derive that for n large enough,

$$\binom{n}{\alpha} \binom{p}{\lceil n/\alpha \rceil} \pi^{\alpha \lceil n/\alpha \rceil} < \left(\frac{p\alpha}{n}\right)^{2n/\alpha} \pi^n e^{\alpha + \lceil n/\alpha \rceil + n/e} < e^{(4+(2/\alpha)\ln(p/n)+\ln \pi)n}.$$

If $\pi = e^{-6}$, it then follows that for n large enough,

$$E(|G|) > np\pi - ne^{(6+\ln \pi)n} > np/e^7,$$

so there exists a bipartite graph $G \subseteq V \times W$, with at least np/e^7 edges and no wide rectangle. With respect to this graph, we have $\mathcal{A}(x) < 2n$, for $x \geq \ln(p/n)$, so from the *Core Lemma,* any (t, m)-cover such that $t \leq c\left(n/\log \frac{p}{n}\right)$ will satisfy $m > p/e^9$, for c small enough. From Lemma 2.1 and our earlier observation that a range searching problem can always be defined to have a prespecified characteristic graph, the proof is now complete. ∎

The comparison between the last two theorems is a little startling. On the one hand, a time speed-up is always possible without using maximum storage. However, trying to improve this speed-up by even a constant factor will immediately force upon us the use of maximum storage (up to within a constant factor). The conclusion to draw is that, in practice, hard range searching problems do not offer any viable alternative to the two naive algorithms.

Remark: There is an intriguing parallel between this result and a general update/query time tradeoff given in (Burkhard et al. [3]), where a similarly pessimistic result is proven optimal. The two situations cannot really be compared, however, because of the difference in settings: storage vs. query time here, as opposed to update time vs. query time in [3]. Without pursuing this digression too far, let us just point out one major difference between the static and the dynamic models. In the former, a cluster is charged unit cost, regardless of its size. In the dynamic model, however, a large cluster, although still charged unit cost, is in effect more costly than a small one because it is more exposed to enemy fire: if any of its points is updated the information provided by the cluster must be thrown away.

4. The Complexity of Simplex Range Searching

We begin by stating the main result of this section: *Simplex range searching on n points requires $\Omega(n/\sqrt{m})$ query time in two dimensions and $\Omega((n/\log n)/m^{1/d})$ query time in any dimension $d \geq 3$. These bounds hold for a random point-set (uniform distribution in the d-cube) with high probability, and thus are valid in the worst case as well as on the average.*

For technical reasons, queries will be slabs of fixed width instead of simplices (since slabs can always be clipped and triangulated, this will actually strengthen our results). The heart of the argument comes from the *Core Lemma*: a generator can be very useful to a few queries or it can be moderately useful to lots of queries, but it *cannot* be very useful to lots of queries. We assess the "effectiveness" of a generator by the Lebesgue measure of the convex hull of its associated cluster. Why? Suppose that we set our sights on a very low query time. Then, presumably, answering a random query will require the use of big clusters. Since the points are uniformly distributed in $\mathcal{C}_d$, big clusters occupy a lot of space and therefore can be used by only few queries. This suggests a tradeoff between the effectiveness of a generator and its ability to be used by many queries. One will notice the similarity of this reasoning with the *Core Lemma.*

Our approach has two components. We begin with an integral-geometric analysis of the containment property between a convex body and a slab. The goal is to produce a continuous analog of the discrete complexity tradeoff sought. To carry out the analogy we must argue that the size of a set of points can be bounded below by the measure of its convex hull (up to within a constant factor). This entails a study of pseudo-uniform point-sets. The questions raised are akin to a classical problem of Heilbronn (Moser [17]) to which we provide new answers.

In section 4.1 we define a measure for slab systems, and we prove its invariance under the group of motions. This will give us a convenient probability measure for queries to work with. In section 4.2 we argue that a large convex set cannot be *moved* too much within a given slab (in other words, a big cluster cannot be used by too many queries). Two fundamental lemmas are derived to formalize this concept. In section 4.3 we turn to the problem of approximating uniform point distributions. Several criteria of uniformity are investigated, one of which leads to new results on a generalization of Heilbronn's problem. Finally, section 4.4 puts all the results above together and derives the desired lower bounds.

4.1. Preliminaries

We begin with some geometric terminology. Let d be a fixed positive integer and let E^d denote Euclidean d-space. Unless specified otherwise, we will always assume that $d > 1$. We endow E^d with a Cartesian system of reference $(O, \vec{e}_1, \dots, \vec{e}_d)$, where $(\vec{e}_1, \dots, \vec{e}_d)$ forms an orthonormal basis $(\vec{e}_i \cdot \vec{e}_j = \delta_{ij})$. We define $E^d_+ = \{(x_1, \dots, x_d) \neq O \mid x_i \geq 0; 1 \leq i \leq d\}$. If $p = (x_1, \dots, x_d)$ and $q = (y_1, \dots, y_d)$ are two points of E^d, then we let $\langle p, q \rangle$ denote the inner product $\sum_{1 \leq i \leq d} x_i y_i$. Similarly, we put $|p| = \sqrt{\langle p, p \rangle}$.

The *width* (resp. *diameter*) of a compact convex set K is the smallest (resp. largest) distance between two distinct hyperplanes of support parallel to each other. The diameter of K is denoted $D(K)$: it is also defined as the greatest distance between two points in K. Finally, if P is a finite set of points in E^d then $\kappa(P)$ denotes its convex hull.

Let α be a real value $(0 < \alpha < 1/12)$ to be considered a parameter in the following. We define a *slab* as the closed region of E^d between any pair of parallel hyperplanes distant of each other by 2α. For any $q \in E^d \setminus \{O\}$, let S_q denote the slab

$$S_q = \{ p \in E^d \,:\, |\langle p, q\rangle - |q|^2| \leq \alpha |q| \}.$$

Using the notation of the exterior calculus, it is well-known (Santaló [18]) that the point-set density $dx_1 \wedge dx_2 \wedge \cdots \wedge dx_d$ is invariant under the group of motions (i.e., isometries). Given $X \subset E^d$, the integral

$$\lambda_d(X) = \int_X dx_1 \wedge \cdots \wedge dx_d$$

is the *measure* of X in E^d (provided that the integral exists in the Lebesgue sense). Next we define the measure μ of a set of slabs X:

$$\mu(X) = \int_X dS_q = \int_{S_q \in X} \frac{dy_1 \wedge \cdots \wedge dy_d}{|q|^{d-1}}$$

(again, provided that the integral exists). Since S_q is not defined for $q = O$, we assume that X does not contain slabs whose bisecting hyperplanes pass through the origin. However, the integral

$$\int_{0 \leq |q| < 1} \frac{dy_1 \wedge \cdots \wedge dy_d}{|q|^{d-1}}$$

converges, so we can extend the measure X over the set of all slabs. (The set of slabs whose bisecting planes pass through the origin has measure zero.)

Lemma 4.1. *The measure μ is invariant under the group of isometries in E^d.*

Proof: Note that the lemma can be easily checked directly for small values of d (obviously, a rather tedious task). Instead of checking the validity of the lemma, we will derive it by using Cartan's method of moving frames (Santaló [18]). Let $P_q = \{ p \in E^d \,|\, \langle p, q\rangle = |q|^2 \}$ be the bisecting hyperplane of S_q, and let $\vec{u}_1, \ldots, \vec{u}_{d-1}$ be an orthonormal basis for P_q. We define $\vec{u}_d$ as a unit vector normal to P_q such that $\det(\vec{u}_1, \ldots, \vec{u}_d) = 1$. Let $\mathcal{M}$ be the group of motions in E^d and let $\Im$ be the subgroup of motions that leave invariant the hyperplane P_q. We have a one-to-one correspondence between the hyperplanes of E^d and the elements of the homogeneous space $\mathcal{M}/\Im = \{ g\Im \,|\, g \in \mathcal{M} \}$: to each coset of the form $g\Im$ $(g \in \mathcal{M})$ corresponds the hyperplane gP_q, and conversely, to each hyperplane P_r corresponds the coset $g\Im$, where g is a motion that carries P_q to P_r. Following [18], finding an invariant density of hyperplanes is then reduced to finding an invariant density for $\mathcal{M}/\Im$. Because $\mathcal{M}/\Im$ is unimodular, it is guaranteed to admit an invariant density. This will be the desired density dL_{d-1} for the hyperplanes P_q, and hence for the slabs S_q. The theory of moving frames gives us the Pfaffian system $dq \cdot \vec{u}_d = 0$ and $d\vec{u}_d \cdot \vec{u}_i = 0$ $(1 \leq i \leq d-1)$. Therefore dL_{d-1} can be chosen as the differential exterior d-form

$$\bigwedge_{i<d} d\vec{u}_d \cdot \vec{u}_i \wedge dq \cdot \vec{u}_d.$$

Let $q = (y_1, \ldots, y_d)$; since $\vec{u}_d = q/|q|$, we have $d\vec{u}_d = dq/|q| + d(1/|q|)q$. From $\vec{u}_i \cdot \vec{u}_j = \delta_{ij}$ we derive

$$dL_{d-1} = \bigwedge_{i<d} (\frac{1}{|q|} dq \cdot \vec{u}_i) \wedge dq \cdot \vec{u}_d = \frac{1}{|q|^{d-1}} \bigwedge_{i \le d} \sum_{1 \le j \le d} dy_j (\vec{u}_i \cdot \vec{e}_j) = \frac{\Delta}{|q|^{d-1}} dy_1 \wedge \cdots \wedge dy_d,$$

where

$$\Delta = \begin{vmatrix} (\vec{u}_1 \cdot \vec{e}_1) & (\vec{u}_2 \cdot \vec{e}_1) & \ldots & (\vec{u}_d \cdot \vec{e}_1) \\ (\vec{u}_1 \cdot \vec{e}_2) & (\vec{u}_2 \cdot \vec{e}_2) & \ldots & (\vec{u}_d \cdot \vec{e}_2) \\ \vdots & \vdots & \ddots & \vdots \\ (\vec{u}_1 \cdot \vec{e}_d) & (\vec{u}_2 \cdot \vec{e}_d) & \ldots & (\vec{u}_d \cdot \vec{e}_d) \end{vmatrix} = 1.$$

This establishes that $dL_{d-1} = \frac{1}{|q|^{d-1}} dy_1 \wedge \cdots \wedge dy_d$, which is precisely the differential form claimed for μ. ∎

Now that we have the appropriate tools to handle random queries, we need some machinery to study the convex hulls of clusters. Arbitrary convex subsets of $\mathcal{C}_d$ are a little unwieldy. Fortunately, they can be approximated quite well by rectangular objects. We need some additional terminology. A *hyperrectangle* is the Cartesian product of d closed intervals in $\Re$. If each interval is of the same length then we have a *hypercube*. A *parallelotope* is the image of a hyperrectangle under an isometric mapping. Two parallelotopes are said to be parallel to each other if they are congruent modulo a homothetic transformation. It seems that the following equivalence result has been rediscovered many times over the years. (I thank J. Pach for pointing out this 1951 reference (Macbeath [15]) to me.)

Lemma 4.2. (Macbeath, 1951) – *Given a compact convex set K in E^d, there exist two parallelotopes Π_1 and Π_2, such that $\Pi_1 \subseteq K \subseteq \Pi_2$ and $\lambda_d(\Pi_2)/d! \le \lambda_d(K) \le d^d \lambda_d(\Pi_1)$.*

4.2. Two Fundamental Lemmas on the Measure of Slabs

Let K be an arbitrary compact convex subset of E^d. Our main concern in this section is to show that the set of slabs

$$H(K) = \{ S_q \mid q \in E^d \setminus \{O\} \text{ and } K \subseteq S_q \}$$

shrinks fast enough as K grows. We will distinguish between the general case (Lemma 4.5) and a rather special case (Lemma 4.6) to be used later for simplex range searching in 2-space. The reason for this distinction is that we can obtain sharper lower bounds in the two-dimensional case by using more refined tools.

We start off our investigation by assuming that K is a hyperrectangle of the form $\prod_{1 \le i \le d} [1, \gamma_i]$, where $\gamma_1, \ldots, \gamma_d$ are d reals ≥ 1. We will actually find it more convenient to trade $H(K)$ for the smaller set

$$H^+(K) = H(K) \bigcap \{ S_q \mid q \in E^d_+ \}.$$

This substitution is fairly innocuous, as the following result shows.

Lemma 4.3. *For any hyperrectangle* $K = \prod_{1\le i\le d}[1,\gamma_i]$, *where* $\gamma_i \ge 1$, *we have the inequality* $\mu(H(K)) \le 2^d \mu(H^+(K))$.

Proof: The idea is to consider the symmetry group of the polytope K, and identify a subset of 2^d automorphisms which allow any point of E^d to be mapped to a point with nonnegative coordinates. The proof will immediately follow from the fact that these automorphisms carry K into itself and the measure μ is invariant under the group of isometries. Let $j = (j_1, \dots, j_d) \in \{-1,1\}^d$ and let g_j be the isometry mapping $p = (x_1, \dots, x_d) \in E^d$ to $g_j(p) = (z_1, \dots, z_d)$, where

$$\begin{pmatrix} z_1 \\ \vdots \\ z_d \end{pmatrix} = \begin{pmatrix} j_1 & 0 & \dots & 0 \\ 0 & j_1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & j_d \end{pmatrix} \times \begin{pmatrix} x_1 - \frac{\gamma_1+1}{2} \\ \vdots \\ x_d - \frac{\gamma_d+1}{2} \end{pmatrix} + \begin{pmatrix} \frac{\gamma_1+1}{2} \\ \vdots \\ \frac{\gamma_d+1}{2} \end{pmatrix}.$$

The transformations g_j are each the composition of a particular subset of plane symmetries. Since K is centrally symmetric about $(\frac{\gamma_1+1}{2}, \dots, \frac{\gamma_d+1}{2})$, each transformation g_j carries K into itself. Given a point $q = (y_1, \dots, y_d) \in E^d \setminus \{O\}$, let $\sigma(q)$ be the sign vector $(j_1, \dots, j_d)$, where $j_i = 1$ (resp. $j_i = -1$) if $y_i \ge 0$ (resp. $y_i < 0$). To complete the proof, we will show that $g_{\sigma(q)}$ maps S_q into $S_{\widehat{q}}$, where $\widehat{q} \in E^d_+$. The inverse transformation $g^{-1}_{\sigma(q)}$ maps a point $(x_1, \dots, x_d)$ into $(z_1, \dots, z_d)$, where

$$z_i = j_i x_i + (1 - j_i)(1 + \gamma_i)/2.$$

Consequently, the slab $S_{\widehat{q}}$ is the set of points $(x_1, \dots, x_d) \in E^d$ such that

$$\Big| \sum_{1\le i\le d} j_i x_i y_i + \frac{1}{2} \sum_{1\le i\le d} (1 - j_i)(1 + \gamma_i) y_i - \sum_{1\le i\le d} y_i^2 \Big| \le \alpha \sqrt{\sum_{1\le i\le d} y_i^2}.$$

We obtain an equivalent expression by noticing that

$$\beta = \Big(\sum_{1\le i\le d} (1 - j_i)(1 + \gamma_i) y_i \Big) \Big/ \sum_{1\le i\le d} y_i^2 \le 0,$$

and multiplying the previous equation by $1 - \beta/2$. This shows that $\widehat{q} = (\widehat{y}_1, \dots, \widehat{y}_d)$, where

$$\widehat{y}_i = (1 - \beta/2) j_i y_i \ge 0.$$

Noticing that $|\widehat{q}| = (1 - \beta/2)|q| > 0$, we can conclude that $\widehat{q} \in E^d_+$. ∎

Throughout this section the term "constant" refers to a quantity which may depend only on d, and not on α or any other parameter later defined. We will use c as a generic symbol to denote a constant, avoiding subscripts whenever we can. Sometimes, however, we will have to resort to subscripts to be able to distinguish between different constants.

Lemma 4.4. *For any dimension* $d > 1$, *there exists a constant* $c > 0$ *such that any hyperrectangle* $K = \prod_{1\le i\le d}[1,\gamma_i]$, *where* $\gamma_i > 1$, *satisfies the inequality* $\lambda_d(K) \cdot \mu(H^+(K)) < c\alpha^{d+1}$.

Proof: We will assume throughout this proof that $q = (y_1, \ldots, y_d) \in E_+^d$ and $K \subseteq S_q$. To begin with, observe that $|q|$ can be neither too large nor too small. Indeed, we have

$$\Big| \sum_{1 \le i \le d} y_i - |q|^2 \Big| \le \alpha |q|. \tag{4.1}$$

Since $x < 1 + x^2/2$ for all x, this implies that

$$|q|^2 - \alpha |q| \le \sum_{1 \le i \le d} y_i < d + |q|^2/2,$$

from which we derive

$$|q| < \alpha + \sqrt{\alpha^2 + 2d}.$$

Since $\alpha < 1/12$ and $d > 1$, it follows (conservatively) that $|q| < 3d$. Similarly, from (4.1) and the fact that $q \in E_+^d$ it follows that

$$|q|^2 + \alpha |q| \ge \sum_{1 \le i \le d} y_i \ge |q|,$$

hence $|q| \ge 1 - \alpha > 11/12 > 1/2$. To summarize, we have shown that

$$1/2 < |q| < 3d. \tag{4.2}$$

Let

$$\Delta_i(K) = \left\{ q = (y_1, \ldots, y_d) \in E_+^d \;\middle|\; K \subseteq S_q \text{ and } y_i > \frac{1}{2\sqrt{d}} \right\}.$$

From (4.2) it follows that q belongs to at least one of the sets $\Delta_i(K)$ $(1 \le i \le d)$, therefore

$$\mu(H^+(K)) \le \sum_{1 \le i \le d} M_i(K), \tag{4.3}$$

where

$$M_i(K) = \int_{\Delta_i(K)} \frac{dy_1 \wedge \cdots \wedge dy_d}{|q|^{d-1}}.$$

To estimate the value of $M_i(K)$, we set $i = 1$ without loss of generality, in order to avoid overburdening the notation. Let us consider the following change of variables: given $q = (y_1, \ldots, y_d)$, let $u_1 = \delta = |q|$, and for $i > 1$, let $u_i = y_i/\delta$. Note that the transformation acts bijectively between

$$\{ (y_1, \ldots, y_d) \in E^d \setminus \{O\} \mid y_1 \ge 0 \}$$

and

$$\{ (u_1, \ldots, u_d) \mid u_1 > 0 \text{ and } \sum_{1 < i \le d} u_i^2 \le 1 \}.$$

To compute its Jacobian, J_u, we notice that

$$\frac{\partial u_i}{\partial y_j} = \begin{cases} y_j/\delta & \text{if } i = 1; \\ 1/\delta - y_i^2/\delta^3 & \text{if } i > 1 \text{ and } i = j; \\ -y_i y_j/\delta^3 & \text{if } i > 1 \text{ and } i \neq j. \end{cases}$$

It follows that

$$J_u = \begin{vmatrix} y_1/\delta & y_2/\delta & \dots & y_d/\delta \\ -y_2y_1/\delta^3 & 1/\delta - y_2^2/\delta^3 & \dots & -y_2y_d/\delta^3 \\ \vdots & \vdots & \ddots & \vdots \\ -y_dy_1/\delta^3 & -y_dy_2/\delta^3 & \dots & 1/\delta - y_d^2/\delta^3 \end{vmatrix}.$$

We derive

$$J_u = \frac{\left(\prod_{1\le i\le d} y_i\right)^2}{y_1\delta^{3d-2}} \times \begin{vmatrix} +1 & +1 & \dots & +1 \\ -1 & (\delta/y_2)^2 - 1 & \dots & -1 \\ \vdots & \vdots & \ddots & \vdots \\ -1 & -1 & \dots & (\delta/y_d)^2 - 1 \end{vmatrix}.$$

The determinant above is made triangular by subtracting the first column from the others, which gives $J_u = y_1/\delta^d$. If $\Upsilon = \int_{q\in\Delta_1(K)} du_1 \wedge \cdots \wedge du_d$, we immediately derive

$$\Upsilon = \int_{(y_1,\dots,y_d)\in\Delta_1(K)} \frac{y_1}{(y_1^2 + \cdots + y_d^2)^{d/2}} dy_1 \wedge \cdots \wedge dy_d.$$

From (4.2) and the definition of $\Delta_1(K)$, we have

$$\Upsilon \geq \frac{1}{6d\sqrt{d}} M_1(K). \tag{4.4}$$

If $q \in \Delta_1(K)$ then we have $\left|\langle p, q\rangle - |q|^2\right| \leq \alpha|q|$, for both $p = (1,\dots,1)$ and $p = (\gamma_1,\dots,\gamma_d)$. Since $q \in E_+^d$ this implies that

$$-\alpha|q| \leq \sum_{1\le i\le d} y_i - |q|^2 \leq \sum_{1\le i\le d} \gamma_i y_i - |q|^2 \leq \alpha|q|,$$

and from (4.2)

$$\sum_{1\le i\le d} (\gamma_i - 1)y_i \leq 2\alpha|q| < 6d\alpha.$$

Because $\gamma_i > 1$, for each i $(1 \leq i \leq d)$, we have

$$0 \leq y_i < \frac{6d\alpha}{\gamma_i - 1}, \tag{4.5}$$

and hence from (4.2),

$$0 \leq u_i < \frac{12d\alpha}{\gamma_i - 1},$$

for $i > 1$. When $u_2, \ldots, u_d$ are fixed, u_1 always varies in an interval of length no greater than 2α. It easily follows that

$$\Upsilon < \frac{(12d\alpha)^d}{\prod_{1<i\leq d}(\gamma_i - 1)}. \tag{4.6}$$

Note that the integration domain of Υ assumes that $y_1 > \frac{1}{2\sqrt{d}}$. From (4.5) we also have $y_1 < 6d\alpha/(\gamma_1 - 1)$, therefore

$$\gamma_1 - 1 < 12d\sqrt{d}\alpha.$$

From (4.4) and (4.6) it then follows that

$$M_1(K) < 6d^2(12d)^{d+1}\alpha^{d+1}/\lambda_d(K).$$

Of course, the same inequality holds for any $M_i(K)$, so in view of (4.3) the proof is now complete. ∎

Lemmas 4.3 and 4.4 give us a tradeoff between the volume of a hyperrectangle and the measure of the slabs that contain it. We can call upon Lemma 4.2 to generalize this tradeoff to any compact convex set.

Lemma 4.5. *For any dimension $d > 1$, there exists a constant $c > 0$ such that, given any compact convex set K, we have $\lambda_d(K) \cdot \mu(H(K)) < c\alpha^{d+1}$.*

Proof: We can assume that $\lambda_d(K) > 0$, otherwise the lemma is obvious. Lemma 4.2 shows the existence of a parallelotope $\Pi \subseteq K$ such that

$$\lambda_d(K) \leq d^d \lambda_d(\Pi). \tag{4.7}$$

Since Π is not of measure zero, it is congruent to a hyperrectangle of the form $K' = \prod_{1\leq i\leq d}[1, \gamma_i]$, where $\gamma_i > 1$. Obviously, the two sets $H(\Pi)$ and $H(K')$ have the same measure. Note that the two sets might not be congruent, because S_q is defined for $q \neq O$, and $E^d \setminus \{O\}$ is not closed under the group of isometries. The difference, however, is simply a set of slabs whose bisecting hyperplanes pass through the origin, and thus has measure zero. From Lemmas 4.3 and 4.4, it follows that for some constant $c > 0$,

$$\lambda_d(\Pi) \cdot \mu(H(\Pi)) < c2^d\alpha^{d+1}.$$

From (4.7) we derive

$$\lambda_d(K) \cdot \mu(H(\Pi)) < c(2d)^d\alpha^{d+1}.$$

Since $\Pi \subseteq K$ the proof is now complete. ∎

We now turn to the special case where K is of measure 0. While Lemma 4.3 is still meaningful, Lemma 4.5 becomes trivial and must be modified a little. Recall that $D(K)$ denotes the diameter of the point-set K.

Lemma 4.6. *For any dimension $d > 1$, there exists a constant $c > 0$ such that given any compact convex set K we have $D(K) \cdot \mu(H(K)) < c\alpha^2$.*

Proof: We follow the proof of Lemma 4.4, assuming that K is a hyperrectangle of the form $\prod_{1 \le i \le d} [1, \gamma_i]$, where $\gamma_1 > 1$ and $\gamma_i = 1$, for $i > 1$. It suffices to show that $(\gamma_1 - 1)\mu(H(K)) < c\alpha^2$. From (4.5) we have $0 \le y_1 < 6d\alpha/(\gamma_1 - 1)$, and from (4.2) we derive that $0 \le y_j < 3d$, for $1 < j \le d$. Consider the case of $M_i(K)$, for $i > 1$. We have $0 \le u_1 < 12d\alpha/(\gamma_1 - 1)$, $u_i = |q|$, and for $j > 1$ and $j \ne i$, we have $0 \le u_j < 6d$. This shows that

$$\Upsilon < \frac{(2\alpha)(12d\alpha)(6d)^{d-2}}{\gamma_1 - 1},$$

and from (4.4)

$$M_i(K) < \frac{4\sqrt{d}(6d)^d\alpha^2}{\gamma_1 - 1}. \tag{4.8}$$

If we assume that $6d\alpha/(\gamma_1 - 1) < 1/(2\sqrt{d})$ then $\Delta_1(K)$ is empty, and hence $M_1(K) = 0$. From (4.3) and (4.8) we find

$$\mu(H^+(K)) < (6d)^{d+2}\alpha^2/(\gamma_1 - 1). \tag{4.9}$$

Is this inequality still true if we relax the assumption on γ_1? If $6d\alpha/(\gamma_1 - 1) \ge 1/(2\sqrt{d})$ then using the previous reasoning, we find $u_1 = |q|$, and $0 \le u_j < 6d$ $(1 < j \le d)$, from which it follows that

$$M_1(K) \le 6d\sqrt{d}(2\alpha)(6d)^{d-1} < (6d)^{d+2}\alpha^2/(\gamma_1 - 1).$$

From (4.9) we derive that

$$\mu(H^+(K)) < 2(6d)^{d+2}\alpha^2/(\gamma_1 - 1),$$

for all values of $\gamma_1 > 1$. Lemma 4.3 completes the proof. ∎

4.3. Approximating Uniform Point-Set Distributions

4.3.1. Introduction

We study the following discrepancy problem: Can we place n points in $\mathcal{C}_d = [0,1]^d$ so that every subset of $k > d$ points has a convex hull of measure at least proportional to k/n? This is the kind of result we need in order to argue that big clusters occupy a lot of space. Let us consider the case $d = 2$ and $k = 3$ for a moment. This is known as *Heilbronn's problem*: *what is the largest area, over all point-sets P of size n, of the smallest triangle formed by points in P?*

This problem has a rich history. If we look at the one-dimensional case for inspiration, we might expect the max-min area to be proportional to $1/n$. However, it has been shown by Komlós, Pintz, and Szemerédi [13] that any set of n points in $\mathcal{C}_2$ always contains a triangle of area less than $1/n^{8/7-\varepsilon}$, for any $\varepsilon > 0$. On the other hand, Komlós, Pintz, and Szemerédi [14] have shown the existence of point-sets with all $\binom{n}{3}$ triangles of area $\Omega\big((\log n)/n^2\big)$. See (Moser [17]) for a chronology of results on Heilbronn's problem.

At the other extreme—the case $k = \Omega(n)$—we certainly have what we expect: the max-min area of the convex hull of any subset of k distinct points is $\Theta(k/n)$: take the vertices of the largest regular n-gon inscribed in $\mathcal{C}_2$. A natural question is thus to determine a small function $k(n)$ for which the max-min area is $\Theta\big(k(n)/n\big)$. More generally, let

$$\Delta_d(n,k) = \max_{\substack{P \subseteq \mathcal{C}_d \\ |P|=n}} \min_{\substack{S \subseteq P \\ |S| \geq k}} \lambda_d\big(\kappa(S)\big),$$

where $d < k \leq n$; recall that $\kappa(S)$ denotes the convex hull of S. We know that $\Delta_d(n,n) = \Theta(n/n)$, but also that $\Delta_2(n,3) = o(3/n)$. We will use a probabilistic argument to prove that $\Delta_d(n,k) = \Theta(k/n)$, for any k such that $\log n \leq k \leq n$.

4.3.2. On a generalization of Heilbronn's problem

We begin with some terminology. Let ν be a positive integer. We say that a set P of n points in E^d is *ν-scattered* if, for every subset $S \subseteq P$ of size $k \geq \nu$, we have $\lambda_d\big(\kappa(S)\big) > \big(1/7^{7^d}\big)k/n$. As it turns out, a weaker version of this definition will allow us to sharpen our lower bounds for range searching in the case $d = 2$. Given any positive real ε, the set P is *weakly ε-scattered* if there exists a subset Q of P such that

(i) $|Q| \geq \varepsilon|P|$;

(ii) for every subset $S \subseteq Q$ of size $k > d$, we have $\lambda_d\big(\kappa(S)\big) > \varepsilon k/(10n)$.

Let R be a parallelotope in E^d of nonzero measure; the k-faces of R are called *vertices* if $k = 0$, and *edges* if $k = 1$. By *edge-length,* we refer to the Euclidean distance between the two vertices at the endpoints of an edge. There are at most d distinct edge-lengths, $a_1, \ldots, a_d$: the minimum value of a_i is the width of R, while $\sqrt{a_1^2 + \cdots + a_d^2}$ is its diameter. Let θ be a positive real: we define the *θ-pads* of R as a collection of 2^d parallelotopes parallel to R, of edge-length θ, attached to each vertex of R. More precisely, let g be an isometry (not necessarily unique) carrying R into some hyperrectangle $\prod_{1 \leq i \leq d} [0, a_i]$ $(a_i > 0)$. Each vertex v of the hyperrectangle is of the form $v = (j_1 a_1, \ldots, j_d a_d)$, where $(j_1, \ldots, j_d)$ is a *bit-vector* in $\{0,1\}^d$. We define the θ-pad of the vertex $g^{-1}(v)$ of R as the image under g^{-1} of the hypercube of edge-length θ centered at the point

$$\big(j_1 a_1 - (-1)^{j_1}\theta/2, \ldots, j_d a_d - (-1)^{j_d}\theta/2\big).$$

The notion of θ-pads is useful for approximating the set of all convex subsets of $\mathcal{C}_d$ by a finite number of canonical polytopes. We begin by listing a few interesting properties of θ-pads.

Lemma 4.7. *Let R be a parallelotope in E^d and let θ be a positive real. If S is a point-set which has at least one point in each θ-pad of R, then the convex hull of S contains R.*

Proof: Without loss of generality, we may assume that R is a hyperrectangle of the form $\prod_{1\le i\le d}[0,a_i]$. Let $s_1,\ldots,s_{2^d}$ be representative points of S in each θ-pad of R, with s_1 the point in the θ-pad of O. Any (closed) halfspace that contains O also contains at least one s_i. To see this, consider the hyperplanes

$$P_q = \{\, p\in E^d \mid \langle p,q\rangle = |q|^2 \,\},$$

and notice that for each sign assignment of the coordinates of q there is at least one desirable s_i. This shows that $O\in\kappa(\{s_1,\ldots,s_d\})$. By symmetry, the same is true of all the other vertices, therefore their convex hull, R, lies inside $\kappa(\{s_1,\ldots,s_d\})$. ∎

Lemma 4.8. *For any $d>1$ and any real ρ $(0<\rho\le 1)$, there exists a collection $\Im$ of convex sets such that* (i) $|\Im| < 5^{5^d}/\rho^{d2^d}$, (ii) *for each $C\in\Im$, we have $C\subseteq \mathcal{C}_d$ and $\lambda_d(C) < 5^{5^d}\rho$,* (iii) *given any convex set K in $\mathcal{C}_d$ of measure ρ, there exists some $C\in\Im$ which contains K.*

Proof: Let $\beta=\rho/d^{\frac{3d+1}{2}}$, and let $\mathcal{G}$ be the grid of points

$$\mathcal{G} = \{\,\pm i\beta \mid i\ge 0\,\}^d \cap [-2d,2d]^d.$$

We define $\Im$ as follows:

$$\Im = \{\, \kappa(S)\cap\mathcal{C}_d \mid S\subseteq\mathcal{G} \text{ and } |S|=2^d \text{ and } \lambda_d(\kappa(S)) < 3^d d!\rho \,\}.$$

We have $|\mathcal{G}| = (2\lfloor 2d/\beta\rfloor + 1)^d$, which implies that $|\mathcal{G}| < 5^d d^{3d(d+1)/2}/\rho^d$, since $\rho\le 1 < d$, and hence (i). Note that (ii) follows directly from the definition of $\Im$, so let us turn our attention to (iii).

Let K be a convex set in $\mathcal{C}_d$ of measure ρ, and let R (resp. r) be the circumscribed (resp. inscribed) parallelotope of least (resp. greatest) measure. From Lemma 4.2, it follows that

$$\lambda_d(R)/d! \le \rho \le d^d\lambda_d(r). \tag{4.10}$$

Let Ψ be the set of $(\beta\sqrt{d})$-pads of R. It is not difficult to show that each pad ψ of Ψ contains a grid point. Let ξ be the center of ψ and let $\mathcal{G}^*$ be the infinite grid $\{\,\pm i\beta \mid i\ge 0\,\}^d$. There is a point $\gamma\in\mathcal{G}^*$ within a distance $\sqrt{d}\beta/2$ of ξ, therefore γ lies within the pad ψ. We must now show that γ is actually a point of $\mathcal{G}$. By construction, no edge-length of R can exceed the diameter of K. Since K lies in $\mathcal{C}_d$ its diameter is at most $\sqrt{d}$. This implies that the diameter of R is at most d. Because the diameter of ψ is equal to $d\beta$, it follows that the distance from γ to O is at most $d+\sqrt{d}+d\beta$, which is less than $2d$, since $\rho\le 1$ and $d\ge 2$. This proves that γ belongs to $[-2d,2d]^d$, and hence to $\mathcal{G}$.

We thus have established the existence of a set S of 2^d points in $\mathcal{G}$, each of which lies in a distinct $(\beta\sqrt{d})$-pad of R. Since r lies inside $\mathcal{C}_d$, its maximum edge-length is at most $\sqrt{d}$, therefore the width of r, and hence the width of R, is at least $\lambda_d(r)/d^{\frac{d-1}{2}}$. From (4.10) it then follows that the width of R is at least $\rho\sqrt{d}/d^{3d/2}$. Let $l_1, \ldots, l_d$ be the edge-lengths of R; we have

$$l_i \geq \rho\sqrt{d}/d^{3d/2} > \beta\sqrt{d},$$

therefore

$$\lambda_d(\kappa(S)) \leq \prod_{1\leq i\leq d} (l_i + 2\beta\sqrt{d}) < 3^d\lambda_d(R).$$

From (4.10) we derive

$$\lambda_d(\kappa(S)) < 3^d d!\rho.$$

This proves that $\kappa(S) \cap \mathcal{C}_d \in \Im$. Lemma 4.7 shows that R lies within $\kappa(S)$, therefore $K \subseteq \kappa(S) \cap \mathcal{C}_d$. ∎

The previous lemma provides a polynomial-size approximation of the set of all convex subsets of $\mathcal{C}_d$ of measure ρ. This enables us to use discrete probabilistic techniques to study certain uniformity criteria for point-sets.

Lemma 4.9. *For any $d > 1$ and n sufficiently large, a random set of n points in $\mathcal{C}_d$ is $(\log n)$-scattered with probability greater than $1 - 1/n$.*

Proof: Let $c = 1/(2^{4^d}b)$, where $b = 5^{5^d}$ is the constant used in Lemma 4.8. Note that $c > 1/7^{7^d}$. Throughout the proof, we will use the notation of Lemma 4.8, with the value of ρ set to $3c(\log n)/n$. (Note that this assignment is valid, since for $n > 1$ we have $0 < \rho < 1$.) We shall also assume that n is larger than some appropriate constant. Let P be a random set of n points in $\mathcal{C}_d$, and let π be the probability that there exists a convex set $K \subseteq \mathcal{C}_d$ such that $k = |K \cap P| \geq \log n$ and $\lambda_d(K) \leq ck/n$. We can assume that the n points of P are distinct since this happens with probability 1. It is then possible to subdivide K into convex sets, each containing between $\log n$ and $2\log n + 1$ points. To do so, choose a line L which is not normal to any of the hyperplanes passing through a pair of points in $K \cap P$, and sort the projection of the points of $K \cap P$ onto L. Since there are no identical elements in the resulting list, we can partition it into sublists of size $\lceil \log n \rceil$ (except for the last one, which might be of lesser size). For each pair of adjacent sublists, find a point on L separating them and cut K by the hyperplane normal to L passing through the point. Of the pieces of K thus created, let K^* be the one of smallest measure. We have

$$\lambda_d(K^*) \leq 3c\frac{\log n}{n} = \rho,$$

therefore we can always enclose K^* inside a convex set $\subseteq \mathcal{C}_d$ of measure ρ. From Lemma 4.8, it follows that the collection $\Im$ contains at least one set C which encloses K^*, where

$$\lambda_d(C) < 3bc\frac{\log n}{n}. \tag{4.11}$$

Clearly, the set C contains at least $\log n$ points of P, therefore

$$\pi < \sum_{C \in \mathfrak{S}} \sum_{j \geq \log n} \binom{n}{j} \lambda_d^j(C)\big(1 - \lambda_d(C)\big)^{n-j}.$$

From (4.11) we have $n\lambda_d(C) < \log n$, therefore we can use the Chernoff bound [8] to approximate the tail of the binomial distribution. This yields

$$\pi < b\left(\frac{n}{3c\log n}\right)^{d2^d} \times \left(\frac{n\big(1-\lambda_d(C)\big)}{n - \log n}\right)^{n-\log n} \times \left(\frac{n\lambda_d(C)}{\log n}\right)^{\log n}. \tag{4.12}$$

Using Taylor's expansion, we have

$$\ln\big(1 - (\log n)/n\big) > -(\log n)/n - (\log n)^2/n^2,$$

for n large enough, therefore

$$n^2(n - \log n)^{n-\log n} > n^{n-\log n}. \tag{4.13}$$

On the other hand, it follows from (4.11) that

$$\big(n\lambda_d(C)/\log n\big)^{\log n} < 1/n^{4^d-2}. \tag{4.14}$$

Putting (4.12–4.14) together, we find the desired (conservative) upper bound $\pi < 1/n$. ∎

As an immediate corollary, we obtain this new result on the generalization of Heilbronn's problem. Whether $\log n$ can be replaced by anything smaller (asymptotically) is a very intriguing open problem.

Theorem 4.10. *The function $\Delta_d(n,k)$ is in $\Theta(k/n)$, for any k such that $\log n \leq k \leq n$.*

Proof: Because of Lemma 4.9 it suffices to show that $\Delta_d(n,k) = O(k/n)$. Given any set P of n points in $\mathcal{C}_d$, partition $\mathcal{C}_d$ into convex sets, each containing between k and $2k+1$ points (using, for example, the method given in Lemma 4.9). Now, consider the convex hull of the set of smallest measure. This set contains at least $\lceil k \rceil$ points and its measure is $O(k/n)$. ∎

Again, simplex range searching in 2-space requires a special treatment. What we need now is a result about uniform point distributions over the real line.

Lemma 4.11. *There exists a positive real $\varepsilon_0 < 1$ such that for any ε $(0 < \varepsilon < \varepsilon_0)$ and any $n > 2$, a random set of n points in $\mathcal{C}_1$ is weakly ε-scattered with probability greater than $1 - \varepsilon$.*

Proof: Consider the inequalities

$$1-\varepsilon/2<\left(1-\theta^2\right)e^{-\theta}<1-\theta/2<1-2\varepsilon/9. \tag{4.15}$$

We claim that there exists some real ε_0 $(0<\varepsilon_0<1/2)$ such that for any ε, where $0<\varepsilon<\varepsilon_0$, there exists θ which satisfies (4.15). To see this it suffices to notice that if $1-\varepsilon/2=\left(1-\theta^2\right)e^{-\theta}$ then $\theta=\varepsilon/2+O(\varepsilon^2)$, and

$$1-\theta/2-\left(1-\theta^2\right)e^{-\theta}=\theta/2+O(\theta^2)=\varepsilon/4+O(\varepsilon^2).$$

Let P be a random set of $n>2$ numbers in $[0,1]$. We say that $x\in P$ is *isolated* if there is no other number of P in $\left[x-\theta/(2n),x+\theta/(2n)\right]$. Let ν be the expected number of isolated points and let π be the probability that at least εn points of P are isolated. We have

$$\nu\leq(1-\pi)\varepsilon n+\pi n. \tag{4.16}$$

On the other hand, we have $\nu\geq n(1-\theta/n)^{n-1}$. Since $\theta<\varepsilon<1<n$, we have (Abramowitz and Stegun [1], pp.68)

$$\nu>n(1-\theta/n)^n=ne^{n\ln(1-\theta/n)}>ne^{-\theta/(1-\theta/n)}.$$

Using the inequalities $e^x\geq 1+x$ and $n>2\theta$, we derive

$$\nu>ne^{-\theta(1+2\theta/n)}\geq ne^{-\theta}\left(1-2\theta^2/n\right).$$

Using (4.15–4.16) and the inequalities $n>2$ and $\varepsilon<1/2$, we have

$$\pi\geq\frac{\nu/n-\varepsilon}{1-\varepsilon}>\frac{(1-\theta^2)e^{-\theta}-\varepsilon}{1-\varepsilon}>\frac{1-3\varepsilon/2}{1-\varepsilon}>1-\varepsilon.$$

On the other hand, the convex hull of any $k>1$ isolated points is an interval of length at least $\frac{1}{2}(k-1)\theta/n$, which from (4.15) is at least $\varepsilon k/(10n)$. ∎

4.4. The Lower Bounds on Simplex Range Searching

We are now in a position to attack our original problem. Let us recall our assumptions. The dimension d is at least 2, and the parameter α is a positive real less than $1/12$. Let m and n be two positive integers, and let Γ be a function mapping any set P of n points in $\mathcal{C}_d$ to a storage scheme for P of size $m>0$. When P and $\Gamma(P)$ are understood, we write t to denote the worst-case time complexity $\max_{q\in\mathcal{Q}}t(P,\Gamma(P),q)$. It will be important to keep in mind later on that t is actually a parameter depending on P and Γ. Ironically, the higher-dimensional case ($d\geq 3$) is easier to handle, so this is where we begin our investigation.

4.4.1. Range searching in 3-space and above

Assume that $d \geq 3$. Let $B_d(\rho)$ be the d-dimensional ball of radius ρ centered at $(1/2, \ldots, 1/2)$. We define the query space $\mathcal{Q}$ to be the set of slabs $\{ S_q \mid q \in B_d(1/4) \}$. We begin our investigation with a technical lemma saying that every query grabs a reasonable chunk of the d-cube: neither too big nor too small.

Lemma 4.12. *For any $d > 1$ there exists a constant $0 < c_1 < 1$ such that, for any $S_q \in \mathcal{Q}$, we have $c_1\alpha < \lambda_d(S_q \cap \mathcal{C}_d) < \alpha/c_1$.*

Proof: Because $S_q \cap B_d(1/2) \subseteq S_q \cap \mathcal{C}_d$ we have $\lambda_d(S_q \cap \mathcal{C}_d) > 2\alpha\lambda_{d-1}(B')$, where B' is the intersection of $B_d(1/2)$ with a hyperplane at distance $1/4 + \alpha$ from the center of $B_d(1/2)$. This implies that B' is a ball in E_{d-1} of radius $r = \sqrt{1/4 - (1/4 + \alpha)^2}$: its $(d-1)$-dimensional measure is therefore (Santaló [18])

$$\lambda_{d-1}(B') = \frac{2\pi^{(d-1)/2} r^{d-1}}{(d-1)\Gamma\big((d-1)/2\big)},$$

where Γ is the gamma function. Using simple approximations we easily verify that, since $\alpha < 1/12$, we have $\lambda_d(S_q \cap \mathcal{C}_d) > c\alpha$, where

$$c = \frac{\pi^{(d-1)/2}}{(d-1)3^d\Gamma\big((d-1)/2\big)}.$$

Conversely, the diameter of $\mathcal{C}_d$ is equal to $\sqrt{d}$, therefore $\lambda_d(S_q \cap \mathcal{C}_d) < 2\alpha\lambda_{d-1}(B^*)$, where B^* is a $(d-1)$-dimensional ball of radius $\sqrt{d}/2$. We derive

$$\lambda_d(S_q \cap \mathcal{C}_d) < \frac{\alpha\pi^{(d-1)/2} d^{(d-1)/2}}{(d-1)2^{d-3}\Gamma\big((d-1)/2\big)}.$$

∎

Given a set P of n points in $\mathcal{C}_d$, we say that a slab S_q is *heavy* if $S_q \in \mathcal{Q}$ and $|S_q \cap P| > c_1\alpha n/2$. We focus on heavy query slabs because they are both well positioned and reasonably filled with points of P. Our next result says that this focusing is legitimate when dealing with a random point-set P: a random query of $\mathcal{Q}$ is heavy with high probability.

Lemma 4.13. *There exists a constant $c > 0$, such that for any fixed real ε $(0 < \varepsilon < 1)$ and a random set of n points in $\mathcal{C}_d$ $(d > 1)$, the measure of the set of heavy slabs exceeds $\big(1 - c/(\alpha\varepsilon n)\big)\mu(\mathcal{Q})$ with probability greater than $1 - \varepsilon$.*

Proof: Let P be a random set of n points in $\mathcal{C}_d$ and let S_q be a slab of $\mathcal{Q}$. Put $\chi = |S_q \cap P|$ and $\sigma = S_q \cap \mathcal{C}_d$. The mean and variance of χ are respectively $n\lambda_d(\sigma)$ and $n\lambda_d(\sigma)\big(1 - \lambda_d(\sigma)\big)$. Let $\pi(q)$ be the probability that S_q is heavy with respect to a random P. Combining Lemma 4.12 and Chebyshev's inequality (Feller [9]), we find

$$1 - \pi(q) \leq \mathrm{Prob}\big(|\chi - n\lambda_d(\sigma)| \geq c_1\alpha n/2\big) \leq \frac{4n\lambda_d(\sigma)\big(1 - \lambda_d(\sigma)\big)}{c_1^2\alpha^2 n^2}.$$

Again from Lemma 4.12 it follows that

$$\pi(q) > 1 - \frac{4}{c_1^3 \alpha n}. \tag{4.17}$$

By Fubini's theorem, the expected value E of the measure of the set of heavy slabs is equal to $\int_{S_q \in \mathcal{Q}} \pi(q)\, dS_q$, which from (4.17) gives

$$E > \left(1 - \frac{4}{c_1^3 \alpha n}\right) \mu(Q). \tag{4.18}$$

On the other hand, we have

$$E \leq (1-p)\left(1 - \frac{5}{c_1^3 \alpha \varepsilon n}\right) \mu(Q) + p\mu(Q),$$

where p is the probability that the measure of the set of heavy slabs is at least $\big(1 - 5/(c_1^3 \alpha \varepsilon n)\big)\mu(\mathcal{Q})$. This inequality, combined with (4.18), shows that $p > 1 - \varepsilon$, which completes the proof. ∎

Let $p_1, \ldots, p_n$ be the points of P. Recall that each generator g of $\Gamma(P)$ is a linear form $\sum_{1 \leq i \leq n} \alpha_i s_i$ and that its cluster is the set $\{\, p_i \mid \alpha_i \neq 0 \,\}$. By abuse of notation we will refer to the clusters *of* $\Gamma(P)$. From the equivalence result of Lemma 2.1 we know that, for each $S_q \in \mathcal{Q}$, the set $S_q \cap P$ can be expressed as the union of at most t clusters. A heavy query contains $\Omega(\alpha n)$ points of P. To be answered in time t therefore requires the use of clusters of size $\Omega(\alpha n/t)$. Just as we chose to focus on heavy queries we will restrict our analysis to those "fat" clusters. Specifically, we say that a cluster is *fat* if it contains at least $\frac{1}{4} c_1 \alpha n/t$ points. For any $S_q \in \mathcal{Q}$, let $\nu(q)$ be equal to the number of points in $S_q \cap P$ which belong to at least one fat cluster lying entirely within S_q. (Note that these clusters may not necessarily be used in answering the query S_q.) Our next result says that with a random point-set P the average value of $\nu(q)$ (over all $S_q \in \mathcal{Q}$) is $\Omega(\alpha n)$.

Lemma 4.14. *There exists a constant $c > 0$ such that, for any fixed real ε $(0 < \varepsilon < 1)$ and a random set of n points in $\mathcal{C}_d$ $(d > 1)$, the inequality*

$$\int_{\mathcal{Q}} \nu(q)\, dS_q > \big(\alpha n/c - c/\varepsilon\big)\mu(\mathcal{Q})$$

holds true with probability greater than $1 - \varepsilon$.

Proof: Given a random set P and $S_q \in \mathcal{Q}$, let $C_1, \ldots, C_u$ be a set of clusters such that $u \leq t$ and $S_q \cap P = \bigcup_{1 \leq i \leq u} C_i$. By the pigeon-hole principle, the number of points of $S_q \cap P$ that belong to at least one cluster C_i such that $|C_i| \geq |S_q \cap P|/(2u)$ exceeds $|S_q \cap P|/2$. Suppose that S_q is heavy. Then because $u \leq t$, we have $\nu(q) > c_1 \alpha n/4$. From Lemma 4.13 it follows that with probability greater than $1 - \varepsilon$ we have

$$\int_{\mathcal{Q}} \nu(q)\, dS_q > \frac{1}{4} c_1 \alpha n \Big(1 - \frac{c}{\alpha \varepsilon n}\Big)\mu(\mathcal{Q}) > \Big(\alpha n/(4/c_1 + c_1 c/4) - (4/c_1 + c_1 c/4)/\varepsilon\Big)\mu(\mathcal{Q}),$$

where c is the constant of Lemma 4.13. ∎

Suppose now that P is $(\log n)$-scattered, with $n > 2^d$, and let S be a subset of P of size at least $\log n$. By definition, we have $\lambda_d(\kappa(S)) > (1/7^{7^d})|S|/n$. From Lemma 4.5 we derive that for some constant $a_1 > 0$,

$$|S| \cdot \mu(H(\kappa(S))) < a_1 \alpha^{d+1} n. \tag{4.19}$$

Put

$$\alpha = \left(\frac{\mu(\mathcal{Q})}{3a_1 cm}\right)^{1/d},$$

where c is the constant of Lemma 4.14, and assume that we have the following inequality:

$$mt^d < \frac{(c_1 n)^d \mu(\mathcal{Q})}{a_1 c (8 \log n)^d}. \tag{4.20}$$

Observe that since $m \geq n$ the condition $\alpha < 1/12$ is satisfied for any n large enough. Let $C_1, \ldots, C_\beta$ be the fat clusters of $\Gamma(P)$. From (4.20) we find that any fat cluster contains more than $\log n$ points. Using the idea behind the *Core Lemma*, we have the following key inequality

$$\int_{\mathcal{Q}} \nu(q)\, dS_q \leq \sum_{1 \leq i \leq \beta} |C_i| \mu(H(\kappa(C_i))).$$

From (4.19) it follows that

$$\int_{\mathcal{Q}} \nu(q)\, dS_q < a_1 \alpha^{d+1} \beta n.$$

Assume that

$$m < \frac{(\varepsilon n)^d \mu(\mathcal{Q})}{a_1 (4c)^{2d+1}}. \tag{4.21}$$

Then Lemmas 4.9 and 4.14 imply that for any n large enough and any ε $(0 < \varepsilon < 1)$ a random set P satisfies

$$\frac{\alpha n}{2c} \mu(\mathcal{Q}) < a_1 \alpha^{d+1} mn$$

with probability greater than $1 - \varepsilon - 1/n$. But this leads to a contradiction, so (4.20) or (4.21) must be false. Since $\mu(\mathcal{Q})$ is larger than some positive constant (independent of ε), we immediately derive the following result:

Lemma 4.15. *For any $d > 1$ and any ε $(0 < \varepsilon < 1)$ there exists a constant $c > 0$, such that for any $n > 1/c$, a random set of n points in $\mathcal{C}_d$ satisfies $mt^d > c(n/\log n)^d$ with probability greater than $1 - \varepsilon$.*

4.4.2. Range searching in 2-space

We use the notation of the previous section. A parallelotope is now simply called a rectangle. Given a slab S_q, we define R_q as the largest rectangle in $S_q \cap \mathcal{C}_2$ with two sides collinear with the bounding lines of S_q. Since $\alpha < 1/12$, R_q is well-defined and unique. As before the query space $\mathcal{Q}$ is the set of slabs $\{ S_q \mid q \in B_2(1/4) \}$. The proof of the following result is almost identical to that of Lemma 4.12, so we omit it.

Lemma 4.16. *There exists a constant* $0 < c_2 < 1$ *such that, for any* $S_q \in \mathcal{Q}$*, we have* $c_2\alpha < \lambda_2(R_q) < \alpha/c_2$.

We must strengthen the concept of heaviness by bringing into play the notion of weak scattering. Given a set P of n points in $\mathcal{C}_2$, we now say that a slab S_q is *ε-favorable* if

(i) $S_q \in \mathcal{Q}$,

(ii) $c_2\alpha n/2 < |R_q \cap P| < 2\alpha n/c_2$, and

(iii) the orthogonal projection of the points of $R_q \cap P$ on either bounding line of S_q is weakly $(c_2\varepsilon^2/2)$-scattered in E^1.

We now have the analog of Lemma 4.13, saying that if we have a random point-set P, then a random query is ε-favorable with high probability.

Lemma 4.17. *There exist two positive constants* c *and* ε_1*, such that for any real* ε $(0 < \varepsilon < \varepsilon_1)$ *and a random set of* n *points in* $\mathcal{C}_2$ $(\alpha n > c)$*, the measure of the set of* ε*-favorable slabs exceeds* $\big(1 - \varepsilon - c/(\alpha\varepsilon n)\big)\mu(\mathcal{Q})$ *with probability greater than* $1 - \varepsilon$.

Proof: Let $\pi(q)$ be the probability that S_q is ε-favorable, conditioned on $S_q \in \mathcal{Q}$. We have $\pi(q) = \pi_1(q)\pi_2(q)$, where $\pi_1(q)$ is the probability that $c_2\alpha n/2 < |R_q \cap P| < 2\alpha n/c_2$ and $\pi_2(q)$ is the conditional probability that the points of $R_q \cap P$ projected onto a bounding line of S_q are weakly $(c_2\varepsilon^2/2)$-scattered, given that $c_2\alpha n/2 < |R_q \cap P| < 2\alpha n/c_2$. Using Chebyshev's inequality we derive

$$\pi_1(q) \geq 1 - \frac{4}{c_2^3\alpha n}. \tag{4.22}$$

On the other hand, since the point distribution is uniform in $\mathcal{C}_2$, given a fixed subset S of P in R_q, the projection of S onto a bounding line of S_q is uniformly distributed along the corresponding side s of R_q. From Lemma 4.16, the length of s is at least $c_2/2 < 1$, so it follows from Lemma 4.11 that if $|R_q \cap P| > c_2\alpha n/2$, then for $\varepsilon < \varepsilon_1$ $(\varepsilon_1 > 0)$ and $\alpha n \geq 4/c_2$, the projection of $R_q \cap P$ is weakly $(c_2\varepsilon^2/2)$-scattered with probability greater than $1 - \varepsilon^2$. From (4.22) we then derive that the expected value Φ of the measure of the set of ε-favorable slabs satisfies

$$\Phi \geq (1 - \varepsilon^2)\left(1 - \frac{4}{c_2^3\alpha n}\right)\mu(Q). \tag{4.23}$$

But we also have

$$\Phi \leq (1 - p)\left(1 - \varepsilon - \frac{5}{c_2^3\alpha\varepsilon n}\right)\mu(Q) + p\mu(Q),$$

where p is the probability that the measure of the set of ε-favorable slabs is at least $\big(1 - \varepsilon - 5/(c_2^3\alpha\varepsilon n)\big)\mu(\mathcal{Q})$. Combining this inequality with (4.23), the lemma follows readily. ∎

Let P be a set of n points in $\mathcal{C}_2$, and let S_q be ε-favorable with respect to P. Then S_q contains a subset $Q \subseteq P$ of size $\geq c_2\varepsilon^2|R_q \cap P|/2$, with the following properties: let Q' be the orthogonal projection of Q onto a bounding line of S_q; then for every subset $S \subseteq Q'$ of at least two points, we have

$$\lambda_1(\kappa(S)) > \frac{c_2\varepsilon^2|S|}{20|R_q \cap P|}.$$

Since $|R_q \cap P| < 2\alpha n/c_2$, this implies that given any subset C of Q of at least two points, the diameter of C satisfies

$$D(C) > \frac{c_2^2\varepsilon^2|C|}{40\alpha n}. \tag{4.24}$$

The subset Q is called the *prime subset* of S_q. Since S_q is ε-favorable, we have

$$|Q| \geq c_2^2\varepsilon^2\alpha n/4. \tag{4.25}$$

Although Q is not necessarily unique we can always use a canonical ordering to make the prime subset unambiguously defined.

Next, we replace $\nu(q)$ by the function $\xi(\varepsilon, q)$, defined as follows: if S_q is ε-favorable then $\xi(\varepsilon, q)$ is the number of points in the prime subset of S_q which share a cluster with at least another point. More precisely, let $\mathcal{F}$ be the set of clusters which lie entirely within S_q and contain at least two points in the prime subset of S_q; then

$$\xi(\varepsilon, q) = \Big|\bigcup_{C \in \mathcal{F}} (C \cap \text{prime subset of } S_q)\Big|.$$

If S_q is not ε-favorable then $\xi(\varepsilon, q) = 0$.

Lemma 4.18. *There exist two positive constants c and ε_1, such that for any real ε $(0 < \varepsilon < \varepsilon_1)$ and a random set of n points in $\mathcal{C}_2$, with $\alpha n/t > 1/(c\varepsilon^2)$, the inequality*

$$\int_{\mathcal{Q}} \xi(\varepsilon, q)\, dS_q > c\varepsilon^2\alpha n\Big(1 - \varepsilon - \frac{1}{c^2\alpha\varepsilon n}\Big)\mu(\mathcal{Q})$$

holds true with probability greater than $1 - \varepsilon$.

Proof: Let S_q be an ε-favorable slab, let Q be its prime subset, and let $C_1, \ldots, C_u$ be a set of clusters such that $u \leq t$ and $S_q \cap P = \bigcup_{1 \leq i \leq u} C_i$. All but at most u points of Q belong to clusters C_i each of which contains at least two points of Q. From (4.25) we derive that for $\alpha n/t > 8/(c_2^2\varepsilon^2)$,

$$\xi(\varepsilon, q) \geq c_2^2\varepsilon^2\alpha n/4 - t > c_2^2\varepsilon^2\alpha n/8.$$

Consequently, Lemma 4.17 shows that with probability greater than $1 - \varepsilon$ a random set P satisfies

$$\int_{\mathcal{Q}} \xi(\varepsilon, q)\, dS_q > \frac{1}{8}c_2^2\varepsilon^2\alpha n\left(1 - \varepsilon - \frac{c}{\alpha\varepsilon n}\right)\mu(\mathcal{Q}),$$

provided that $\alpha n > c$ and $\alpha n/t > 8/(c_2^2\varepsilon^2)$. Since $t \geq 1$ the lemma follows directly. ∎

Let $C_1, \ldots, C_m$ be the clusters of $\Gamma(P)$, and for each i ($1 \leq i \leq m$), let η_i be the maximum number of points in C_i to be in the prime subset of the same ε-favorable slab S_q. If this number is strictly less than 2, then we set $\eta_i = 0$. Clearly,

$$\int_{\mathcal{Q}} \xi(\varepsilon, q)\, dS_q \leq \sum_{1 \leq i \leq m} \eta_i \mu(H(\kappa(C_i))).$$

Using (4.24) for each i such that $\eta_i > 0$, as well as Lemma 4.6, we derive

$$\int_{\mathcal{Q}} \xi(\varepsilon, q)\, dS_q < \left(\frac{40\alpha n}{c_2^2 \varepsilon^2}\right) \sum_{1 \leq i \leq m} D(\kappa(C_i))\mu(H(\kappa(C_i))) < c_3 \alpha^3 mn/\varepsilon^2, \tag{4.26}$$

for some constant $c_3 > 0$.

Finally, put

$$\alpha = \frac{\varepsilon^2}{2}\sqrt{\frac{c\mu(\mathcal{Q})}{c_3 m}},$$

where c is the constant of Lemma 4.18. Note that $\alpha < 1/12$ for any n large enough (since $n \leq m$). Assume now that

$$mt^2 < \frac{c^3 \varepsilon^8 \mu(\mathcal{Q}) n^2}{4c_3} \tag{4.27}$$

and

$$m < \frac{c^5 \varepsilon^8 \mu(\mathcal{Q}) n^2}{4c_3}. \tag{4.28}$$

From (4.27) we have $\alpha n/t > 1/(c\varepsilon^2)$, therefore Lemma 4.18 and (4.26) show that with probability greater than $1 - \varepsilon$,

$$c\varepsilon^2 \alpha n\Big(1 - \varepsilon - \frac{1}{c^2 \alpha \varepsilon n}\Big)\mu(\mathcal{Q}) < c_3 \alpha^3 mn/\varepsilon^2,$$

for n large enough. From (4.28) we derive

$$c\varepsilon^2 \alpha n(1 - 2\varepsilon)\mu(\mathcal{Q}) < c_3 \alpha^3 mn/\varepsilon^2,$$

which gives a contradiction if, say, $\varepsilon < \min\{\varepsilon_1, 1/4\}$. This implies that (4.27) or (4.28) has to be false. Since $\mu(\mathcal{Q})$ is bounded below by a positive constant, with probability greater than $1 - \varepsilon$ we have $mt^2 > c_4 \varepsilon^8 n^2$, for $c_4 > 0$ and n large enough. Note that the condition $\varepsilon < \min\{\varepsilon_1, 1/4\}$ can be relaxed by choosing c_4 small enough.

Lemma 4.19. *There exists a constant $c > 0$, such that for any ε ($0 < \varepsilon < 1$) and $n > 1/c$, a random set of n points in $\mathcal{C}_2$ satisfies $mt^2 > c\varepsilon^8 n^2$ with probability greater than $1 - \varepsilon$.*

4.4.3. Summary of results and closing remarks

Let us recap the main results of this section (Lemmas 4.15 and 4.19) and state some immediate corollaries.

Lemma 4.20. *Let S be a faithful commutative semigroup; let $d > 1$ be a positive integer and ε any real $(0 < \varepsilon < 1)$. There exists a constant $c > 0$ such that the following is true. Let P be a random set of n points in $\mathcal{C}_d$ and let Γ be any storage scheme of size m for the range searching problem $(S, \mathcal{Q}, P)$, where $\mathcal{Q}$ is the set of all slabs of fixed (appropriately chosen) width in E^d. Then if n is large enough, with probability greater than $1 - \varepsilon$, the time complexity $t = \max_{q \in \mathcal{Q}} t(P, \Gamma, q)$ satisfies the inequality $mt^2 > cn^2$, for $d = 2$, and $mt^d > c(n/\log n)^d$, for $d > 2$. As a corollary, the worst-case and average-case time complexities satisfy*

$$t(n,m) \geq \bar{t}(n,m) = \Omega\big(n/\sqrt{m}\big)$$

for $d = 2$, and

$$t(n,m) \geq \bar{t}(n,m) = \Omega\Big(\frac{n}{\log n}/m^{1/d}\Big)$$

for $d \geq 3$.

Of course, these lower bounds also apply to simplex range searching, since a slab can always be clipped into a parallelotope without changing the nature of the problem, and a d-dimensional parallelotope can always be triangulated into at most $d!$ simplices. We can therefore state our result in a more illustrative manner.

Theorem 4.21. *Simplex range searching on n points requires $\Omega\big(n/\sqrt{m}\big)$ query time in two dimensions and $\Omega\big((n/\log n)/m^{1/d}\big)$ query time in any dimension $d \geq 3$, where m denotes the amount of storage available. These bounds hold for a random point-set, and therefore are valid in the worst case as well as on the average.*

As we mentioned in the introduction, simplex range searching on n points in d-space can be performed in $O\big(n^{1-1/d}\alpha(n)\big)$ query time and $O(n)$ storage, where α is a functional inverse of Ackermann's function (Chazelle [5]). This upper bound, which holds in the arithmetic model, matches our lower bound very closely. On a random access machine supplied with linear storage, the best upper bound on the query time to date is $O\big(\sqrt{n}\log n\big)$ in 2-space (Chazelle and Welzl [5]) and $O\big(n^{d(d-1)/(d(d-1)+1)+\varepsilon}\big)$ in d-space, for any $d \geq 3$ and any fixed $\varepsilon > 0$ (Haussler and Welzl [12]). An interesting open problem is to bridge the gap in higher dimensions and generalize the upper bounds to general space-time tradeoffs. Another intriguing question is to determine whether halfspace queries are as hard as simplex queries.

Acknowledgments: I would like to thank F.K. Chung, P. Erdös, R.L. Graham, A.M. Odlyzko, and E. Szemerédi for helpful discussions about Heilbronn's problem.

REFERENCES

1. Abramowitz, M., Stegun, I.A. *Handbook of mathematical functions*, Dover Publications, 1970.
2. Aho, A.V., Hopcroft, J.E., Ullman, J.D. *The design and analysis of computer algorithms*, Addison-Wesley, 1974.
3. Burkhard, W.A., Fredman, M.L., Kleitman, D.J. *Inherent complexity trade-offs for range query problems*, Theoretical Computer Science 16 (1981), 279–290.
4. Chazelle, B. *Lower bounds on the complexity of multidimensional searching*, Proc. 27th Annu. IEEE Symp. on Foundat. of Comput. Sci. (1986), 87–96.
5. Chazelle, B., Welzl, E. *Partition trees and stabbing numbers in Euclidean space*, Disc. Comp. Geom., to appear.
6. Cole, R., Yap , C.K. *Geometric retrieval problems*, Inform. and Control 63 (1984), 39–57.
7. Edelsbrunner, H., Welzl, E. *Halfplanar range search in linear space and* $O(n^{0.695})$ *query time*, Inform. Process. Lett. (1986).
8. Erdös, P., Spencer, J. *Probabilistic methods in combinatorics*, Academic Press, New York, 1974.
9. Feller, W. *An introduction to probability theory and its applications*, Vol.1, 3rd edition, John Wiley & Sons, New York, 1968.
10. Fredman, M.L. *A lower bound on the complexity of orthogonal range queries*, J. ACM, 28 (1981), 696–705.
11. Fredman, M.L. *Lower bounds on the complexity of some optimal data structures*, SIAM J. Comput. 10 (1981), 1–10.
12. Haussler, D., Welzl, E. *Epsilon-nets and simplex range queries*, Disc. Comp. Geom., 2, (1987), 127–151.
13. Komlós, J., Szemerédi, E., Pintz, J. *On Heilbronn's triangle problem*, J. London Math. Soc. 2, 24 (1981), 385–396.
14. Komlós, J., Szemerédi, E., Pintz, J. *A lower bound for Heilbronn's problem*, J. London Math. Soc. 2, 25 (1982), 13–24.
15. Macbeath, A.M. *A compactness theorem for affine equivalence classes of convex regions*, Canadian J. of Math. 3 (1951), 54–61.
16. Mehlhorn, K. *Data structures and algorithms 3: Multidimensional Searching and Computational Geometry*, Springer-Verlag (1984).
17. Moser, W.O.J. *Problems on extremal properties of a finite set of points*, Discrete Geometry and Convexity, Annals of the New York Academy of Sciences 440 (1985), 52–64.
18. Santaló, L.A. *Integral geometry and geometric probability*, Encyclopedia of Math. and its Applicat., Vol. 1, Ed. Gian-Carlo Rota, Addison-Wesley, Reading, Mass., 1976.
19. Willard, D.E. *Polygon retrieval*, SIAM J. Comput. 11 (1982), 149–165.
20. Yao, A.C. *On the complexity of maintaining partial sums*, SIAM J. Comput. 14 (1985), 277–288.
21. Yao, F.F. *A 3-space partition and its applications*, Proc. 15th Ann. ACM Sympos. Theory Comput. (1983), 258–263.

22. Yao, A.C., Yao, F.F. *On computing the rank function for a set of vectors*, Rep. No. UIUCDCS-R-75-699, Univ. of Illinois at Urbana-Champaign, 1975.
23. Yao, A.C., Yao, F.F. *A general approach to d-dimensional geometric queries*, Proc. 17th Annu. ACM Symp. on Theory of Comput. (1985), 163–168.