

# Tight Lower Bounds for Halfspace Range Searching

[Extended Abstract]

Sunil Arya<sup>\*</sup>

Department of Computer  
Science and Engineering  
The Hong Kong University of  
Science and Technology  
Clear Water Bay, Kowloon,  
Hong Kong  
arya@cse.ust.hk

David M. Mount<sup>†</sup>

Department of Computer  
Science and Institute for  
Advanced Computer Studies  
University of Maryland  
College Park, Maryland 20742  
mount@cs.umd.edu

Jian Xia

Department of Computer  
Science and Engineering  
The Hong Kong University of  
Science and Technology  
Clear Water Bay, Kowloon,  
Hong Kong  
piper@cse.ust.hk

## ABSTRACT

We establish two new lower bounds for the halfspace range searching problem: Given a set of  $n$  points in  $\mathbb{R}^d$ , where each point is associated with a weight from a commutative semigroup, compute the semigroup sum of the weights of the points lying within any query halfspace. Letting  $m$  denote the space requirements, we prove a lower bound for general semigroups of  $\tilde{\Omega}(n^{1-1/(d+1)}/m^{1/(d+1)})$  and for integral semigroups of  $\tilde{\Omega}(n/m^{1/d})$ .

Our lower bounds are proved in the semigroup arithmetic model. Neglecting logarithmic factors, our result for integral semigroups matches the best known upper bound due to Matoušek. Our result for general semigroups improves upon the best known lower bound due to Brönnimann, Chazelle, and Pach. Moreover, Fonseca and Mount have recently shown that, given uniformly distributed points, halfspace range queries over idempotent semigroups can be answered in  $O(n^{1-1/(d+1)}/m^{1/(d+1)})$  time in the semigroup arithmetic model. As our lower bounds are established for uniformly distributed point sets, it follows that they also resolve the computational complexity of halfspace range searching over idempotent semigroups in this important special case.

## Categories and Subject Descriptors

F.2.2 [Analysis of Algorithms and Problem Complexity]: Nonnumerical Algorithms and Problems—*Geometrical problems and computations*

<sup>\*</sup>The work of the first and third author was supported by the Research Grants Council, Hong Kong, China under project number 610106.

<sup>†</sup>This author's work was supported in part by the National Science Foundation under grant CCR-0635099 and the Office of Naval Research under grant N00014-08-1-1015.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

SCG'10, June 13–16, 2010, Snowbird, Utah, USA.

Copyright 2010 ACM 978-1-4503-0016-2/10/06 ...\$10.00.

## General Terms

Algorithms, Theory

## Keywords

Range searching, lower bounds, idempotence

## 1. INTRODUCTION

Given a set of points  $P$  in  $\mathbb{R}^d$ , preprocess these points so that given any halfspace  $\eta$ , the number of points of  $P$  lying within  $\eta$  can be computed efficiently. This problem, called *halfspace range counting*, which is so simple to state, has proven to be among the most challenging problems in the field of computational geometry. It is arguably the most fundamental search problem in affine geometry and is a building block in a number of other search problems, such as simplex range searching and spherical range searching.

This problem is typically posed in a weighted setting. Each point is associated with a weight from a commutative semigroup, and the output of the query is the semigroup sum of the points lying within the halfspace. Semigroup properties can be exploited by the data structure and may affect the computational complexity of the problem. A semigroup is *idempotent* if  $x + x = x$  for all semigroup elements  $x$ . It is *integral* if, for all nonzero semigroup elements  $x$  and all natural numbers  $k \geq 2$ , the  $k$ -fold sum  $x + \dots + x$  is not equal to  $x$  [11]. For example,  $(\mathbb{R}, \min)$  and  $(\{0, 1\}, \vee)$  are both idempotent, and  $(\mathbb{N}, +)$  is integral.

Idempotence is relevant because of the way that most range searching algorithms work. At preprocessing time the algorithm implicitly computes the semigroup sum of a number of suitably chosen subsets of  $P$ , which we call *generators* (also known as *canonical subsets*). To answer a query  $\eta$ , the algorithm determines an (ideally small) subset of generators whose union is equal to  $P \cap \eta$  and then returns their total sum. If the semigroup is idempotent these subsets may overlap, but for integral semigroups they must be disjoint. Because of the constraint of disjointness, one would expect that range searching over integral semigroups should be harder than for idempotent semigroups.

This problem has an extensive history within computational geometry. Early work focused on data structures of linear size. A number of methods based on partition-trees were proposed, including methods by Willard [19], Edels-

brunner and Welzl [10], and Yao [21]. A major breakthrough was achieved by Haussler and Welzl [14], who applied the concepts of VC-dimension and  $\varepsilon$ -nets to obtain a data structure with query time  $O(n^\alpha)$ , where  $\alpha \approx 1 - \frac{1}{d(d-1)+1}$ . In a subsequent sequence of papers, a number of new techniques were introduced, resulting in linear-space (or nearly linear space) algorithms by Welzl [18], Chazelle and Welzl [8], Matoušek and Welzl [17], and Chazelle *et al.* [7]. This culminated in a solution by Matoušek [15], which achieves linear space and the optimal query time of  $O(n^{1-\frac{1}{d}})$ . Very recently, Chan [5] has shown that this same query time can be achieved with a conceptually simpler and more flexible approach based on partition trees. For further information, see the excellent surveys by Agarwal and Erickson [1] and Matoušek [16]. All of these results apply in the most general case of integral semigroups, and so apply to idempotent semigroups as well.

Given the relatively high query time, it is natural to consider space-time tradeoffs. By precomputing more generators, it should be possible to cover an arbitrary range by a smaller number of generators, and so achieve lower query times. Letting  $m$  denote the space requirements, where  $n \leq m \leq n^d$ , the best upper bound for halfspace range searching is due to Matoušek [15] and Chan [5], who have shown that using  $m$  units of space, queries can be answered in  $O(n/m^{1/d})$  time. This bound holds in both the standard real RAM model as well as in the semigroup arithmetic model.

Lower bounds for this problem were established by Brönnimann, Chazelle, and Pach [4] in the semigroup arithmetic model. We refer to this paper throughout as BCP. They established a lower bound on the query time for general semigroups of  $\tilde{\Omega}(n^{1-\frac{d-1}{d(d+1)}}/m^{1/d})$  using  $m$  units of space. This bound holds for both idempotent and integral semigroups. The gap between the BCP lower bound and Matoušek's upper bound was investigated by Arya *et al.* [2], and the lower bound was strengthened to  $\tilde{\Omega}(n/m^{\frac{d+1}{d^2+1}})$  for the special case of integral semigroups.

In this paper, we take a substantial step towards achieving the goal of resolving the complexity of halfspace range searching by reducing the gap between the upper and lower bounds to logarithmic factors. Because of the weakness of the semigroup arithmetic model (which considers only the number of generators needed to answer a query, and not the time to determine these generators), it is natural to ignore logarithmic factors and focus on polynomial factors of  $n$  and  $m$ . In particular, we establish the following two new lower bounds for halfspace range searching in the semigroup arithmetic model:

- We establish an improved lower bound for general semigroups of  $\tilde{\Omega}(n^{1-1/(d+1)}/m^{1/(d+1)})$ .
- We establish an improved lower bound for integral semigroups of  $\tilde{\Omega}(n/m^{1/d})$ .

The latter result relies on the assumption that the convex hulls of the generator subsets used to answer a query are disjoint. Although this assumption does not hold for Matoušek's data structure [15], which is based on hierarchical cuttings, it is satisfied by Chan's partition-tree construction [5]. (We conjecture that our bound holds even without this assumption.)

To understand the significance of these results, observe that, neglecting logarithmic factors, our lower bound for integral semigroups matches the best known upper bounds to date.

For general semigroups, it is useful to observe that, as with BCP, our lower bounds are established for uniformly distributed point sets. In contrast to many geometric problems, uniformly distributed point sets appear to represent a fairly hard case for halfspace range searching. Fonseca and Mount [9] recently showed that, given uniformly distributed points, halfspace range queries over idempotent semigroups can be answered in  $O(n^{1-1/(d+1)}/m^{1/(d+1)})$  time in the semigroup arithmetic model. Thus, for the semigroup arithmetic model, our results resolve the computational complexity of halfspace range searching over idempotent semigroups in the case of uniformly distributed point sets.

## 2. TECHNIQUES

Our proofs are structurally similar to those in BCP. Their proof is based on a number of sophisticated geometric tools, such as the isoperimetric inequality, the slicing lemma, and Macbeath regions. In the BCP proof the generators used to answer a query are allowed to overlap. Arya *et al.* [2] showed how to enhance their proof, so it yields a better lower bound in the integral case, by exploiting the fact that the generators used to answer a query must be disjoint.

Our paper makes two key technical contributions. First, by building on the ideas in BCP, we design new slicing lemmas specifically tailored to the requirements of the idempotent and integral cases. Second, we describe a much better way of taking advantage of disjointness than given in [2], leading to a nearly optimal bound in the integral case.

Before discussing our contributions in more detail, we need to first briefly recall the essential ingredients of the BCP proof. We let  $\mathbb{U} = [0, 1]^d$  denote the unit hypercube in  $\mathbb{R}^d$ , and for any body  $K$  in  $\mathbb{R}^d$ , we let  $\mu(K)$  denote its Lebesgue measure. We will assume some familiarity with the semigroup arithmetic model. (See Section 3 for more information about this model.) The data point instance used in the proof consists of a set of uniformly distributed points in  $\mathbb{U}$ . This makes it possible to relate the volume of a convex subset of  $\mathbb{U}$  to the number of data points it contains.

The complexity of halfspace range searching arises from the difficulty of covering points close to the boundary of the query halfspace. To formalize this idea, they define a suitable set of query halfspaces and a suitable *region of interest* for each query halfspace, which is defined to be a thin slab close to the boundary of the halfspace. Recall that the goal in the arithmetic model is to relate the space (number of generators) to the query time (maximum number of generators needed to answer any query). In order to establish this relationship, a probability distribution is defined on the set of query ranges, and a lower bound is obtained by showing that if a generator covers a large number of points in the region of interest for some range, then it cannot be useful in this manner for many query ranges.

In order to achieve query time  $t$ , the storage scheme must have at most  $t$  generators that cover all the points of  $P$  that lie in the region of interest  $R_H$  associated with the query halfspace  $H$ . It is possible to define a generator's *usefulness* to  $H$ , based on the number of such points it covers. The analysis proceeds by establishing an upper bound on the expected usefulness of any generator. In turn, this implies

a lower bound on the total number of generators needed to achieve the given query time.

Techniques are presented in the BCP paper for bounding the usefulness of a generator. They involve an *isoperimetric inequality* and a *slicing lemma*. Their isoperimetric inequality bounds the probability that a fixed convex body is contained within a random query slab, as a function of the volume of the convex body and the thickness of the slab. Since  $P$  is uniformly distributed, a generator  $G$  that is useful for some halfspace  $H$  enjoys the property that its convex hull  $K$  has a large volume of intersection with  $R_H$ . Their slicing lemma involves breaking the convex hull  $K$  of the generator into a small collection of convex bodies called *Macbeath regions*, all contained within  $K$ , which serve as witnesses to usefulness. In particular, if  $G$  is useful to  $H$ , then at least one of these bodies lies within the slab  $R_H$ , and it has volume at least  $\Omega(\mu(K \cap R_H))$ .

A critical quantity in the BCP analysis is the number of convex bodies into which we need to decompose the convex hull  $K$  of a generator in order to satisfy the aforementioned property. Define a *cap* to be the intersection of a convex body and a halfspace. We require that, for any cap  $C$  of  $K$  of sufficiently large volume and limited thickness, one of these bodies is contained within  $C$  and has volume  $\Omega(\mu(C))$ . BCP’s slicing lemma produces a decomposition of  $K$  into convex bodies by considering only the volume condition on the caps. Our key observation is that many of the bodies produced by their slicing lemma are actually too thick to lie inside the caps that arise in their proof, and as a result the size of their decomposition is much larger than necessary. We develop a new slicing lemma that is sensitive to *both* the volume and the thickness of the caps, and, as a consequence, we are able to show that a much smaller number of bodies suffice. This approach leads to an improved lower bound in the case of idempotent (and hence general) semigroups.

We now discuss our ideas for the integral case. As mentioned earlier, Arya *et al.* [2] showed how to exploit the disjointness of the generators to improve the BCP lower bound. Their proof relies on the following simple but critical observation: because of disjointness, the total volume of the convex hulls of the generators used to answer a query must be  $O(1)$ . In turn, they show that this implies that, for a typical generator  $G$  used to answer a query  $H$ , the volume of intersection of its convex hull  $K$  with the slab  $R_H$  is at least  $\Omega(\mu(K) \cdot \Delta_0)$ , where  $\Delta_0$  denotes the slab thickness. They use this fact together with the BCP slicing lemma to derive an improved lower bound in the integral case. While this approach is perhaps the simplest way of incorporating disjointness into the BCP framework, it leads to a lower bound that falls short of Matoušek’s upper bound [15].

In this paper, we close this gap by exploiting the disjointness of generators in a more sophisticated manner. More precisely, we show that, for a typical generator  $G$  used to answer a query  $H$ , the volume of intersection of its convex hull  $K$  with the slab  $R_H$  is at least  $\tilde{\Omega}(\text{breadth}(K, H) \cdot \Delta_0)$ , where  $\text{breadth}(K, H)$  denotes the maximum cross-sectional area of a surface formed by intersecting  $K$  with any hyperplane parallel to  $\partial H$ . In order to bound the usefulness of generators based on this new criterion, we need to develop a slicing lemma specifically tailored to it. In contrast with both the BCP framework and the aforementioned work by Arya *et al.* [2], in which each convex hull of a generator is decomposed into a polynomial number of bodies, our slic-

ing lemma decomposes each hull into only a polylogarithmic number of pieces. This improvement turns out to be critical and allows us to obtain a lower bound in the integral case that matches Matoušek’s upper bound, up to logarithmic factors.

### 3. PRELIMINARIES

Before presenting our results we begin with some general definitions and assumptions. Throughout we assume that the dimension  $d$  is a fixed constant greater than 1. Unless otherwise stated, we will use the term “constant” to refer to any fixed quantity, which may depend on  $d$  but not on  $n$ . To avoid specifying many real-valued constants that arise in our analyses, we will often hide them using asymptotic notation. For positive real  $x$ , we use the notation  $O(x)$  (resp.,  $\Omega(x)$ ) to mean a quantity whose value is at most (resp., at least)  $cx$  for an appropriately chosen constant  $c$ . We use the notation  $\tilde{\Omega}(x)$  in place of  $\Omega(x)$  to ignore factors depending logarithmically on  $n$ . It will also be convenient in our analysis to use the notations  $\preceq$  (resp.,  $\succeq$ ) to denote less than (resp., greater than) up to constant factors.

Let  $(S, +)$  be a commutative semigroup. Let  $P$  be a set of  $n$  points in  $\mathbb{R}^d$ , and let  $w : P \rightarrow S$  be a function that assigns a semigroup value in  $S$  to each point in  $P$ . For any subset  $G$  of  $P$ , we define its weight  $w(G) = \sum_{p \in G} w(p)$ , where the summation is taken over the semigroup. Let  $\mathcal{Q}$  denote the set of query ranges in the range space. In the range searching problem, we are required to preprocess  $P$  so that for any query range  $\eta \in \mathcal{Q}$ , we can efficiently compute  $w(P \cap \eta)$ .

Our lower bound proofs are in the *semigroup arithmetic model* [12, 20]. We briefly recall some of the basic elements of this model for online range searching problems. Lower bounds in this model assume that the semigroup is *faithful*, meaning that any two identically equal linear forms have the same set of variables [6]. For example,  $(\mathbb{N}, +)$ ,  $(\mathbb{R}, \min)$ , and  $(\{0, 1\}, \vee)$  are faithful, but  $(\{0, 1\}, + \text{ mod } 2)$  is not. Given a set  $\{x_1, \dots, x_n\}$  of  $n$  variables over  $S$ , a *generator*  $G(x_1, \dots, x_n)$  is a linear form  $\sum_{i=1}^n \alpha_i x_i$ , where the values  $\alpha_i$  are nonnegative integers, not all zero. A *storage scheme* is a set of generators  $\{G_1, \dots, G_m\}$  satisfying the following property. For any query range  $\eta \in \mathcal{Q}$ , there exists a set  $I_\eta \subseteq \{1, \dots, m\}$  and a set of labeled nonnegative integers  $\{\beta_i : i \in I_\eta\}$  such that

$$w(P \cap \eta) = \sum_{i \in I_\eta} \beta_i G_i(w(p_1), \dots, w(p_n)),$$

for any weight function  $w$ . The *query time* for  $\eta$  is defined to be the size of smallest such set  $I_\eta$ . In this model the *space* is the number of generators in the storage scheme. Intuitively, the generators correspond to partial sums that have been precomputed in the data structure. The query time in the semigroup arithmetic model counts the minimum number of semigroup operations on these generators needed to answer a query. The time for auxiliary operations, such as determining which generators to use, is ignored.

For our purposes, it will be convenient to identify a generator with the subset of  $P$  corresponding to the variables in the linear form with positive integer coefficients. Given an idempotent semigroup, the time to answer a query  $\eta$  is the size of the smallest set of generators whose union is  $P \cap \eta$ . More formally, let  $\mathcal{G}$  be any storage scheme consisting of  $m$  gener-

ators. For any range  $\eta \in \mathcal{Q}$ , define  $A_\eta \subseteq \mathcal{G}$  to be the smallest subset of generators of  $\mathcal{G}$  such that  $\bigcup_{G \in A_\eta} G = P \cap \eta$ . The worst-case query time for a storage scheme  $\mathcal{G}$  is the maximum of  $|A_\eta|$  over all ranges  $\eta \in \mathcal{Q}$ . Given a space bound  $m$ , the worst-case query time complexity in the semigroup arithmetic model is the minimum worst-case query time over all storage schemes  $\mathcal{G}$  of size  $m$  [6]. We can adapt this to integral semigroups by adding the restriction that, when taking unions in answering a query, the generators, viewed as subsets of  $P$ , are pairwise disjoint.

Let us define some other terms that will be used throughout the chapter. Consider a compact (closed and bounded) convex body  $K$ , a halfspace  $H$ , and a vector  $u$  orthogonal to  $\partial H$ . The *width* of  $K$  in the direction  $u$ , denoted  $\text{width}(K, u)$ , is defined to be the orthogonal distance between the two supporting hyperplanes for  $K$  that are orthogonal to  $u$ . The *breadth* of  $K$  orthogonal to  $u$ , denoted  $\text{breadth}(K, u)$ , is defined to be the maximum cross-sectional area of  $K$  orthogonal to  $u$ . Note that for our purposes, width is always a one-dimensional measure, while breadth is a  $(d-1)$ -dimensional measure. As a convenience, we also define  $\text{width}(K, H) = \text{width}(K, u)$  and  $\text{breadth}(K, H) = \text{breadth}(K, u)$ .

The closed region in  $\mathbb{R}^d$  bounded by two parallel  $(d-1)$ -dimensional hyperplanes is called a *slab*, and its *width* (or *thickness*) is the orthogonal distance between these hyperplanes. Let  $\Delta > 0$  be a real parameter. Given any hyperplane  $J$ , let  $S^\Delta(J)$  denote the slab consisting of points in  $\mathbb{R}^d$  whose distance from  $J$  is at most  $\Delta$ . This slab has width  $2\Delta$ . Given a halfspace  $H$ , let  $\partial H$  denote its bounding hyperplane. For any halfspace  $H$ , let  $S^\Delta(H) = S^\Delta(\partial H) \cap H$  denote the slab consisting of points in  $H$  that are within distance  $\Delta$  of the bounding hyperplane of  $H$ .

Consider a compact, convex body  $K$  and a halfspace  $H$  that has a nonempty intersection with  $K$ . Let  $u$  be a vector orthogonal to  $\partial H$ . The intersection  $C = K \cap H$  is called the *cap* of  $K$  generated by  $H$ . We refer to  $u$ , the vector orthogonal to  $\partial H$ , as the *defining direction* of cap  $C$ . The *width* (or *thickness*) of a cap, denoted  $\text{width}(C)$ , is defined to be the width of  $C$  along its defining direction  $u$ .

## 4. LOWER BOUND

The main results of this chapter are summarized in the following theorem. Our first result for general semigroups improves the lower bound of Brönnimann, Chazelle, and Pach [4]. Moreover, since our bounds are established for uniformly distributed points, Fonseca and Mount's result mentioned in the introduction implies that this result is almost optimal for this distribution and idempotent semigroups. Our second result is almost optimal in the worst-case for integral semigroups in light of Matoušek's upper bound [15].

**THEOREM 4.1.** *Let  $d \geq 2$  be a fixed dimension. Consider a range space consisting of all halfspaces and a weight function over any faithful semigroup. Then for all sufficiently large  $n$ , using  $m \geq n$  units of storage,*

- (i) *For general (and hence idempotent) semigroups, the worst-case query time in the semigroup arithmetic model is at least*

$$\Omega\left(\frac{n^{1-\frac{1}{d+1}}}{m^{\frac{1}{d+1}}} \cdot \frac{1}{\log^{1-\frac{1}{d+1}} n}\right) = \tilde{\Omega}\left(\frac{n^{1-\frac{1}{d+1}}}{m^{\frac{1}{d+1}}}\right).$$

- (ii) *If the semigroup is integral, then under the assumption that convex hulls of the generators used to answer a query are disjoint, the query time is at least*

$$\Omega\left(\frac{n}{m^{\frac{1}{d}}} \cdot \frac{1}{\log^{1+\frac{2}{d}} n}\right) = \tilde{\Omega}\left(\frac{n}{m^{\frac{1}{d}}}\right).$$

### 4.1 New Slicing Lemmas

In this section, we establish our new slicing lemmas tailored to the requirements of the idempotent and integral cases. We will need the following lemma, involving important convex bodies called *Macbeath regions*, to help us in that task. This lemma is an easy consequence of Lemma 2.7 in BCP. Recall that for any body  $K$  in  $\mathbb{R}^d$ , we let  $\mu(K)$  denote its Lebesgue measure.

**LEMMA 4.1.** *Consider a compact convex body  $K \subset \mathbb{R}^d$  and a parameter  $0 < v \leq \mu(K)$ . There exists a collection of pairwise disjoint convex bodies  $K_1, K_2, \dots \subseteq K$  satisfying the following properties:*

- (i)  $\mu(K_i) = \Omega(v)$  for all  $i$ .  
(ii) Let  $H$  be a halfspace, and let  $C$  be the cap  $K \cap H$ . If  $\mu(C) \geq v$ , then there exists  $j$  such that  $K_j \subseteq C$ .

Using the above lemma, we now prove our slicing lemma that will be useful in the idempotent case. It differs from the one given in BCP in that it incorporates a condition on the cap width and, thereby, produces a smaller number of convex bodies in cases of interest.

**LEMMA 4.2.** *Consider a compact convex body  $K \subset \mathbb{R}^d$  of surface area  $O(1)$ , and two parameters  $v > 0$  and  $\Delta > 0$ . There exists a collection of  $O(\Delta/v)$  convex bodies  $K_1, K_2, \dots \subseteq K$  satisfying the following property: Let  $H$  be a halfspace and let  $C$  be the cap  $K \cap H$ . If  $\text{width}(C) \leq \Delta$  and  $\mu(C) \geq v$ , then there exists  $j$  such that  $K_j \subseteq C$  and  $\mu(K_j) = \Omega(\mu(C))$ .*

**PROOF.** Let  $K_\Delta \subseteq K$  denote the region consisting of points in  $K$  that are within distance  $\Delta$  of  $K$ 's boundary. Observe that  $\mu(K_\Delta) = O(\Delta)$  since the surface area of  $K$  is  $O(1)$ .

Consider the set  $\mathcal{C}$  consisting of caps  $C$  formed by intersecting  $K$  with any halfspace, such that  $C$  satisfies the properties  $\text{width}(C) \leq \Delta$  and  $\mu(C) \geq v$ . We partition the set  $\mathcal{C}$  into groups based on the volume of the caps as follows. For  $i > 0$ , let  $\mathcal{C}_i$  be the subset of  $\mathcal{C}$  consisting of caps whose volume lies between  $2^{i-1}v$  and  $2^i v$ . We claim that there exists a collection of  $O(\Delta/(2^{i-1}v))$  convex bodies satisfying the properties mentioned in the lemma with respect to the caps in  $\mathcal{C}_i$ . The lemma then follows immediately from this claim by summing over all  $i$ .

To prove the claim, we apply Lemma 4.1, setting the parameter  $v$  in that lemma to  $2^{i-1}v$ . Clearly, this yields a collection  $\mathcal{K}_i$  of pairwise disjoint convex bodies, each contained within  $K$  and having volume  $\Omega(2^{i-1}v)$ . Further, for any cap  $C \in \mathcal{C}_i$ , one of these bodies is contained within  $C$  and has volume  $\Omega(\mu(C))$  (because the volume of each body is  $\Omega(2^{i-1}v)$  and  $\mu(C) \leq 2^i v$ ). Thus, the collection  $\mathcal{K}_i$  satisfies all the required properties with respect to the caps in  $\mathcal{C}_i$ . However, the number of bodies in  $\mathcal{K}_i$  may be too large. We fix this by eliminating any convex body in  $\mathcal{K}_i$  that is not contained in  $K_\Delta$ . Since the caps in  $\mathcal{C}_i$  have width at most  $\Delta$ , such a body is in any case of no use in satisfying a

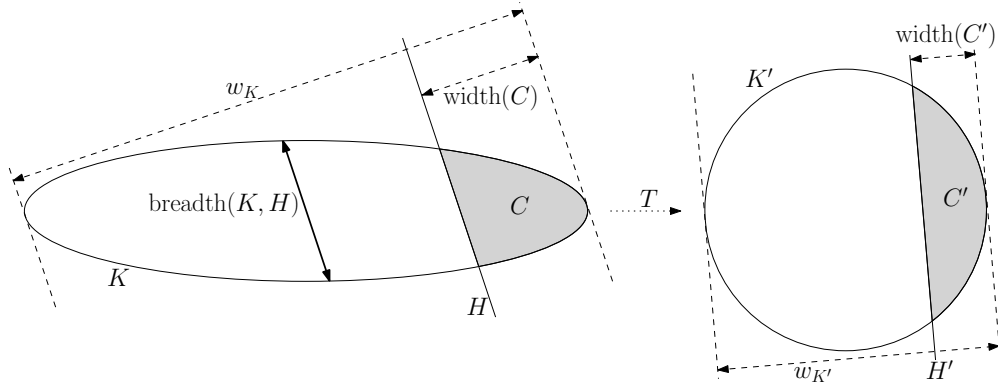


Figure 1: Proof of Lemma 4.4.

cap of  $C_i$ . Thus, the resulting collection continues to satisfy all the desired properties. Finally, since the bodies in  $\mathcal{K}_i$  are pairwise disjoint and are all contained in  $K_\Delta$ , we have  $|\mathcal{K}_i| = O(\mu(K_\Delta)/(2^{i-1}v)) = O(\Delta/(2^{i-1}v))$ . This completes the proof of the claim and the lemma.  $\square$

Before presenting our slicing lemma that will be useful for integral semigroups, we need to introduce a standard notion of fatness. A convex body  $K \subset \mathbb{R}^d$  is said to be  $\alpha$ -fat, for some  $\alpha \geq 1$ , if there exist two concentric balls  $B^+$  and  $B^-$  such that  $B^- \subseteq K \subseteq B^+$ , and the ratio between the radius of  $B^+$  and  $B^-$  is at most  $\alpha$ . Further, we say that  $K$  is fat if  $\alpha$  is a constant depending only on  $d$ . The following well-known result states that any convex body can be transformed by a volume-preserving affine transformation into a fat convex body (see e.g. [13]).

LEMMA 4.3. *Given any compact convex body  $K \subset \mathbb{R}^d$ , there exists a volume-preserving affine transform  $T$  such that  $T(K)$  is fat.*

The next lemma considers the effect of the above fattening transformation on the width of caps. We will find it useful in proving the slicing lemma.

LEMMA 4.4. *Consider any compact convex body  $K \subset \mathbb{R}^d$  of unit volume. Let  $H$  be a halfspace and let  $C$  denote the cap  $K \cap H$ . Let  $K' = T(K)$  and  $H' = T(H)$ , where  $T$  is the transformation of Lemma 4.3. Let  $C' = K' \cap H'$ . Then  $\text{width}(C') = \Theta(\text{breadth}(K, H) \cdot \text{width}(C))$ .*

PROOF. Let  $w_K$  denote the width of  $K$  along the defining direction of  $C$  and  $w_{K'}$  denote the width of  $K'$  along the defining direction of  $C'$ . (See Figure 1.) A straightforward geometric argument implies that  $\mu(K) = \Theta(\text{breadth}(K, H) \cdot w_K)$ . Since  $\mu(K) = 1$ , we have  $w_K = \Theta(1/\text{breadth}(K, H))$ . By Lemma 4.3,  $K'$  is fat and  $\mu(K') = 1$ . It follows that the width of  $K'$  in any direction is  $\Theta(1)$ . In particular, we have  $w_{K'} = \Theta(1)$ . Thus  $w_{K'}/w_K = \Theta(\text{breadth}(K, H))$ . By standard properties of affine transformations, we obtain  $\text{width}(C')/\text{width}(C) = w_{K'}/w_K$ . Thus  $\text{width}(C')/\text{width}(C) = \Theta(\text{breadth}(K, H))$ , as desired.  $\square$

We are now ready to prove the slicing lemma that will be applied in the integral case.

LEMMA 4.5. *Consider an  $\alpha$ -fat compact convex body  $K \subset \mathbb{R}^d$  and two parameters  $\beta \geq 1$  and  $\Delta > 0$ . There exists a collection of  $O(\beta \log \alpha)$  convex bodies  $K_1, K_2, \dots \subseteq K$  satisfying the following property: Let  $H$  be a halfspace and let  $C$  denote the cap  $K \cap H$ . If  $\text{width}(C) \leq \Delta$  and  $\mu(C) \geq \text{breadth}(K, H) \cdot \Delta/\beta$ , then there exists  $j$  such that  $K_j \subseteq C$  and  $\mu(K_j) = \Omega(\mu(C))$ .*

PROOF. Without loss of generality, we may assume that  $K$  has unit volume. (It is easy to check that it suffices to prove the lemma for  $K$  under this condition.) Consider the set  $\mathcal{C}$  consisting of caps  $C$  formed by intersecting  $K$  with some halfspace  $H$ , such that  $C$  satisfies the properties  $\text{width}(C) \leq \Delta$  and  $\mu(C) \geq \text{breadth}(K, H) \cdot \Delta/\beta$ . We partition the set  $\mathcal{C}$  into groups such that for all caps in a group, the breadth of  $K$  orthogonal to the defining direction of the cap differs by a factor of at most 2. Since  $K$  is  $\alpha$ -fat, it is clear that the ratio between the largest and smallest value of the breadth of  $K$  over all possible directions cannot exceed  $\alpha^{d-1}$ , and thus we can obtain a partitioning with at most  $O(\log \alpha)$  groups.

Let us focus on one such group of caps, for which the breadth of  $K$  orthogonal to the defining direction of any cap in it is between  $b$  and  $2b$ . We will show that there exists a collection of  $O(\beta)$  convex bodies satisfying the properties mentioned in the lemma with respect to all the caps in this group. Since the total number of groups is  $O(\log \alpha)$ , the lemma will then directly follow from this.

Let  $C = K \cap H$  denote any cap in this group. From the conditions in the lemma, we have  $\mu(C) \geq \text{breadth}(K, H) \cdot \Delta/\beta \geq b\Delta/\beta$  and  $\text{width}(C) \leq \Delta$ . Let  $K' = T(K)$  and  $H' = T(H)$ , where  $T$  is the transformation of Lemma 4.3. Let  $C' = K' \cap H'$ . Since  $T$  is volume-preserving, we have  $\mu(C') = \mu(C) \geq b\Delta/\beta$ . By Lemma 4.4, we have  $\text{width}(C') = \Theta(\text{breadth}(K, H) \cdot \text{width}(C))$ . Since  $\text{width}(C) \leq \Delta$  and  $\text{breadth}(K, H) \leq 2b$ , we obtain  $\text{width}(C') = O(b\Delta)$ .

Note that the surface area of  $K'$  is  $O(1)$ . We invoke Lemma 4.2 on  $K'$ , setting the parameter  $\Delta$  (upper bound on cap width) to a suitable constant times  $b\Delta$ , and parameter  $v$  (lower bound on cap volume) to  $b\Delta/\beta$ . We obtain a collection of  $O((b\Delta)/(b\Delta/\beta)) = O(\beta)$  convex bodies in  $K'$ , such that one of these bodies is inside any cap  $C'$  of  $K'$  corresponding to a cap  $C$  of  $K$  in the group under consideration, and its volume is  $\Omega(\mu(C'))$ . It follows that the set of bodies in the original space corresponding to these bodies

satisfy all the desired properties with respect to the caps in the group. This completes the proof of the lemma.  $\square$

## 4.2 Lower Bound: The Idempotent Case

We begin with the proof of Theorem 4.1(i), namely the lower bound in the idempotent case. As mentioned earlier, our proof in this case is a simple adaptation of the one in BCP, differing only in its use of a more suitable slicing lemma. We present details for the sake of completeness.

Let  $P$  be a set of  $n$  points in the unit hypercube  $\mathbb{U} = [0, 1]^d$ . Let  $K$  be any convex body contained within  $\mathbb{U}$  and let  $k = |P \cap K|$ . We say that  $P$  is *scattered* if the following two properties hold for some constant  $a > 1$ :

- (i) If  $\mu(K) \geq a \log n/n$ , then  $k \geq (n/a)\mu(K)$ .
- (ii) If  $k \geq \log n$ , then  $k \leq a n \mu(K)$ .

Intuitively, scattered point sets behave like uniformly distributed points in the sense that the volume of a convex body is related to the number of points it contains. The following lemma asserts the existence of such point sets. It is proved in [3] and is a simple variant of Lemma 2.1 in BCP.

LEMMA 4.6. *A random set of  $n$  points sampled uniformly and independently in  $\mathbb{U}$  is scattered with probability  $1 - o(1)$ .*

The data set  $P$  in the proof consists of a scattered set of  $n$  points in  $\mathbb{U}$ . We next describe the set of query ranges. Let  $O$  denote the origin of the coordinate system in  $\mathbb{R}^d$ . Let  $q$  be any point in  $\mathbb{R}^d \setminus \{O\}$ . Let  $H_q$  denote the halfspace containing the origin, whose bounding hyperplane passes through  $q$  and is orthogonal to segment  $Oq$ . Let  $B$  be the ball of radius  $1/4$  whose center coincides with the center of  $\mathbb{U}$ . We define the query set  $\mathcal{Q}$  to be the set of halfspaces  $\{H_q | q \in B\}$ .

Next, we define a probability measure over the set of queries. As in BCP, define the measure of any set  $\mathcal{H}$  of halfspaces as follows:

$$\int_{\mathcal{H}} dH_q = \int_{H_q \in \mathcal{H}} \frac{dx_1 \wedge \cdots \wedge dx_d}{\|q\|^{d-1}},$$

where  $q = (x_1, \dots, x_d)$ . The choice of this measure is based on the fact that it is invariant under rigid motions [6]. It is easy to see that the measure of the query set  $\mathcal{Q}$  is  $\Theta(1)$ . Therefore, restricted to the set  $\mathcal{Q}$ , we can interpret this measure as a probability density by dividing it by the total measure of  $\mathcal{Q}$ . In our analysis, we will assume that a query halfspace is selected from  $\mathcal{Q}$  according to this probability density (we will say that the halfspace is *random* over  $\mathcal{Q}$ ). Given any query halfspace  $H \in \mathcal{Q}$ , we have that for a suitable constant  $c_1$  and any sufficiently small  $\Delta$  (say  $\Delta \leq 1/4$ ),  $\mu(S^\Delta(H) \cap \mathbb{U}) \geq c_1 \Delta$ .

The following isoperimetric inequality, which was proved by Chazelle [6], bounds the probability that a random slab encloses a given convex body. We restate it slightly for convenience.

LEMMA 4.7. *Let  $K \subseteq \mathbb{U}$  be any compact convex body, let  $H$  be a random halfspace of  $\mathcal{Q}$ , and let  $\Delta > 0$  be a real parameter. Then*

$$\mu(K) \cdot \Pr[K \subseteq S^\Delta(H)] = O(\Delta^{d+1}).$$

Let  $\mathcal{G}$  be any set of  $m \geq n$  generators for  $P$ . Let  $t = t(n, m)$  denote the worst-case query time in the semigroup

arithmetic model over all the ranges of  $\mathcal{Q}$ . We will establish a lower bound on  $t$ , assuming that  $n$  is sufficiently large.

In BCP it is observed that the complexity of halfspace range searching stems from the difficulty of covering points that lie close to the boundary of the halfspace. To formalize the idea, for each halfspace  $H \in \mathcal{Q}$ , define a *region of interest*  $R_H \subseteq H$  to be  $S^{\Delta_0}(H)$ , where  $\Delta_0 = c_0 t \log n/n$  for a suitable constant  $c_0$ .

Let  $\Phi$  denote the expected number of points of  $P$  in the region of interest  $R_H$ . That is,

$$\Phi = \mathbb{E}[|P \cap R_H|].$$

We will compute lower and upper bounds on  $\Phi$ , which together will provide the desired lower bound on the query time  $t$ . We may assume that  $t = O(n/\log^2 n)$ , since otherwise Theorem 4.1 holds trivially.

We claim that, for a suitable choice of  $c_0$ ,  $|P \cap R_H| \geq 8t \log n$  for all  $H \in \mathcal{Q}$ . By the property of query halfspaces mentioned above, we have  $\mu(R_H \cap \mathbb{U}) \geq c_1 \Delta_0 = c_1 (c_0 t \log n/n)$ , if  $c_0 t \log n/n \leq 1/4$ . Since  $t = O(n/\log^2 n)$ , this condition holds for sufficiently large  $n$ . Setting  $c_0 = 8a/c_1$ , where  $a$  is the constant used in the scattered set properties, we have  $\mu(R_H \cap \mathbb{U}) \geq 8at \log n/n$ . Since  $t \geq 1$ , we can apply property (i) of scattered point sets to obtain  $|P \cap R_H| = |P \cap R_H \cap \mathbb{U}| \geq (n/a)\mu(R_H) \geq 8t \log n$ , as desired. The following lower bound on  $\Phi$  is now immediate.

LEMMA 4.8.  *$\Phi$  is at least  $\Omega(t \log n)$ .*

Next we compute an upper bound on  $\Phi$ . Towards this end, it is helpful to focus on those generators that are most efficient in covering the region of interest  $R_H$ . Intuitively, in order to achieve a query time of  $t$ , a typical generator should cover a fraction of roughly  $\Omega(1/t)$  of the points in  $R_H$ . This suggests the following definition. We say that a generator  $G \in \mathcal{G}$  is *absolutely fat* with respect to a query  $H \in \mathcal{Q}$  if  $|G \cap R_H| > \log n$  and  $G \subseteq H$ . The following lemma, implicit in the BCP proof, shows that a constant fraction of the points in  $R_H$  are covered by such generators.

LEMMA 4.9. *For any query  $H \in \mathcal{Q}$ , a constant fraction of the points of  $P$  lying within  $R_H$  are covered by generators of  $\mathcal{G}$  that are absolutely fat with respect to  $H$ .*

PROOF. Given  $H \in \mathcal{Q}$ , consider the set  $A_H \subseteq \mathcal{G}$  of generators that allow us to answer query  $H$  within time  $t$ . Clearly,  $|A_H| \leq t$ . Let  $A_H^1$  denote the subset of  $A_H$  consisting of generators that are not absolutely fat with respect to  $H$ . Observe that the generators in the set  $A_H^1$  cover at most  $t \log n$  points of  $P \cap R_H$  (because there are at most  $t$  such generators and each covers at most  $\log n$  points of  $P \cap R_H$ ). Recalling that  $|P \cap R_H|$  is at least  $8t \log n$ , it follows that at least  $7/8$  of the points of  $P \cap R_H$  are covered by the generators in  $A_H \setminus A_H^1$ . As all these generators are absolutely fat with respect to  $H$ , the lemma follows.  $\square$

Consider any  $G \in \mathcal{G}$  and  $H \in \mathcal{Q}$ . We define the *usefulness* of  $G$  with respect to  $H$  to be  $|G \cap R_H|$  if  $G$  is absolutely fat with respect to  $H$ , and 0 otherwise. We denote this quantity by  $u(G, H)$ . Further, we define the *usefulness* of  $G$ , denoted  $u(G)$ , to be the expectation of  $u(G, H)$ , where the average is taken over the query distribution.

The above lemma implies that,

$$|P \cap R_H| \leq \sum_{G \in \mathcal{G}} u(G, H).$$

Therefore,

$$\begin{aligned}\Phi &= \mathbb{E}\left[|P \cap R_H|\right] \preceq \mathbb{E}\left[\sum_{G \in \mathcal{G}} u(G, H)\right] \\ &= \sum_{G \in \mathcal{G}} \mathbb{E}\left[u(G, H)\right] = \sum_{G \in \mathcal{G}} u(G).\end{aligned}$$

In the remainder, we compute an upper bound on  $u(G, H)$ , and hence on  $u(G)$ , with the help of our version of the slicing lemma and Chazelle's isoperimetric inequality. This will then allow us to upper bound  $\Phi$  by using the above equation.

**LEMMA 4.10.** *Consider any  $G \in \mathcal{G}$  and  $H \in \mathcal{Q}$  such that  $G$  is absolutely fat with respect to  $H$ . Let  $K = \text{conv}(G)$ . Then  $|G \cap R_H| \preceq n \mu(K \cap R_H)$  and  $\mu(K \cap R_H) = \Omega(\log n/n)$ .*

**PROOF.** By the definition of absolute fatness,  $|G \cap R_H| > \log n$ . Clearly,  $|P \cap K \cap R_H| \geq |G \cap R_H|$ . Therefore  $|P \cap K \cap R_H| > \log n$ . Applying property (ii) of scattered point sets, we have  $|P \cap K \cap R_H| \preceq n \mu(K \cap R_H)$ . Putting these inequalities together, the lemma follows.  $\square$

Let  $G, H, K$  be as in the above lemma. Since  $G$  is absolutely fat with respect to  $H$ , by definition,  $G \subseteq H$ , and so  $K \cap R_H$  is a cap of  $K$ . Clearly, this cap has width at most  $\Delta_0 = O(t \log n/n)$ , and its volume, by the above lemma, is  $\Omega(\log n/n)$ . Also, since  $K \subseteq \mathbb{U}$ , its surface area is  $O(1)$ . We apply our slicing lemma (Lemma 4.2) to  $K$ , setting  $v$  to a suitable constant times  $\log n/n$  and  $\Delta$  to  $\Delta_0$ . We obtain a collection of  $O(\Delta/v) = O(t)$  convex bodies  $K_1, K_2, \dots \subseteq K$  such that one of these bodies lies inside  $K \cap R_H$  and has volume  $\Omega(K \cap R_H)$ . Thus,

$$\mu(K \cap R_H) \preceq \sum_i \mu(K_i) \cdot \mathbb{I}[K_i \subseteq R_H],$$

where  $\mathbb{I}[K_i \subseteq R_H]$  is an indicator random variable that takes the value 1 if  $K_i \subseteq R_H$  and 0 otherwise. Applying Lemma 4.10, we obtain

$$u(G, H) = |G \cap R_H| \preceq n \mu(K \cap R_H) \preceq n \sum_i \mu(K_i) \cdot \mathbb{I}[K_i \subseteq R_H].$$

Since  $u(G, H) = 0$ , if  $G$  is not absolutely fat with respect to  $H$ , it follows that the inequality

$$u(G, H) \preceq n \sum_i \mu(K_i) \cdot \mathbb{I}[K_i \subseteq R_H]$$

holds for all  $G \in \mathcal{G}$  and  $H \in \mathcal{Q}$ . Taking expectation on both sides, we obtain

$$u(G) \preceq n \sum_i \mu(K_i) \cdot \Pr[K_i \subseteq R_H] \preceq n t \Delta_0^{d+1},$$

where we have used Lemma 4.7 (Chazelle's isoperimetric inequality) and the  $O(t)$  bound on the number of convex bodies derived from our slicing lemma.

Recalling that  $\Phi \preceq \sum_{G \in \mathcal{G}} u(G)$  and the total number of generators is  $|\mathcal{G}| = m$ , we obtain

$$\Phi = O(m n t \Delta_0^{d+1}).$$

Combining this with the lower bound of  $\Omega(t \log n)$  on  $\Phi$  derived earlier, substituting  $\Delta_0 = O(t \log n/n)$ , we obtain

$$t \log n \preceq m n t \left(\frac{t \log n}{n}\right)^{d+1},$$

which can be simplified to

$$t^{d+1} \succeq \frac{n^d}{m} \cdot \frac{1}{\log^d n}.$$

Therefore,

$$t = \Omega\left(\frac{n^{\frac{d}{d+1}}}{m^{\frac{1}{d+1}}} \cdot \frac{1}{\log^{\frac{d}{d+1}} n}\right) = \tilde{\Omega}\left(\frac{n^{1-\frac{1}{d+1}}}{m^{\frac{1}{d+1}}}\right).$$

This completes the proof of Theorem 4.1(i).

### 4.3 Lower Bound: The Integral Case

In this section we prove part (ii) of Theorem 4.1. As mentioned earlier, throughout this section we assume that the convex hulls of the generators used to answer a query are disjoint. This assumption does not seem to be too restrictive. As mentioned earlier, although it does not hold for Matoušek's data structure [15], it is satisfied by Chan's partition-tree structure [5].

Our proof is structurally similar to that of Arya *et al.* [2], who first showed how to enhance the BCP framework to take advantage of disjointness. The key difference is that we use a different formulation for capturing disjointness, which enables us to obtain better bounds. We now present the details.

We define the data set  $P$ , query set  $\mathcal{Q}$ , worst-case query time  $t$ , region of interest  $R_H$  for each halfspace  $H \in \mathcal{Q}$ , set of generators  $\mathcal{G}$ , and  $\Phi$ , exactly as in the idempotent case. To simplify the notation, for any generator  $G$ , we use  $\text{breadth}(G, H)$  to denote the quantity  $\text{breadth}(\text{conv}(G), H)$ . As before, we can establish a lower bound of  $\Omega(t \log n)$  on  $\Phi$ . A key aspect of our strategy is to use a different notion of generator *usefulness*. Towards this end, define a generator  $G \in \mathcal{G}$  to be *relatively fat* with respect to a query  $H \in \mathcal{Q}$  if  $|G \cap R_H| \geq (\text{breadth}(G, H) \cdot \Delta_0 \cdot n)/(c \log n)$ , for a suitable constant  $c$  to be specified later. Further, define generator  $G$  to be *fat* with respect to query  $H$  if  $G$  is absolutely fat with respect to  $H$  and either (i)  $G \subset R_H$  or (ii)  $G$  is relatively fat with respect to  $H$ . We can now strengthen Lemma 4.9 for the integral case as follows.

**LEMMA 4.11.** *Assume that the convex hulls of the generators used to answer a query are disjoint. For any query  $H \in \mathcal{Q}$ , a constant fraction of the points of  $P$  lying within  $R_H$  are covered by generators of  $\mathcal{G}$  that are fat with respect to  $H$ .*

The proof of Lemma 4.11 relies on the following important geometric observation.

**LEMMA 4.12.** *Consider a halfspace  $H \in \mathcal{Q}$  and any subset  $\mathcal{G}_H \subseteq \mathcal{G}$  consisting of generators whose convex hulls are disjoint. Further, suppose that each generator  $G \in \mathcal{G}_H$  satisfies the following properties: (i)  $G \subset H$ , (ii)  $G \cap S^{\Delta_0/2}(H) \neq \emptyset$ , and (iii)  $G \not\subset S^{\Delta_0}(H)$ . Then  $\sum_{G \in \mathcal{G}_H} \text{breadth}(G, H) = O(\log n)$ .*

**PROOF.** For the sake of illustration, assume that  $\partial H$  is "horizontal" and  $H$  is the halfspace consisting of points "below"  $\partial H$ . Let  $J_0$  denote the hyperplane parallel to  $\partial H$  and at distance  $\Delta_0/2$  below it. Further, for  $i \geq 1$ , let  $J_i$  denote the hyperplane parallel to  $J_0$  and at distance  $2^{i-3} \Delta_0$  below it. (See Figure 2.) Note that the number of hyperplanes in the set  $\{J_i : i \geq 0\}$  that intersect  $\mathbb{U} = [0, 1]^d$  is  $O(\log(1/\Delta_0)) = O(\log n)$ .

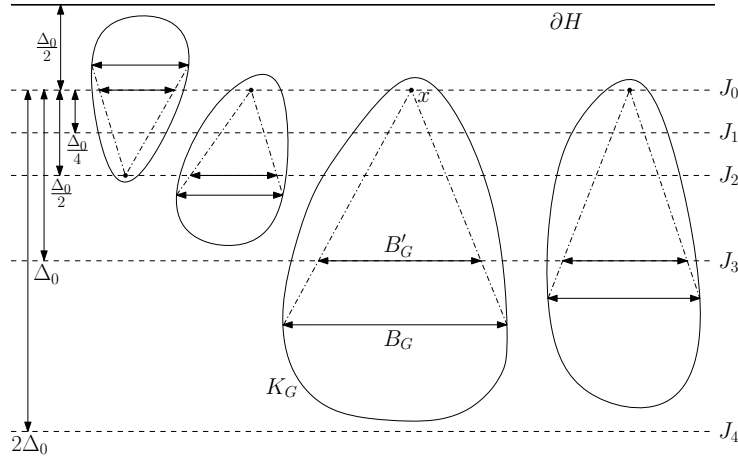


Figure 2: Proof of Lemma 4.12.

Let  $G$  be any generator of  $\mathcal{G}_H$ . Let  $K_G = \text{conv}(G)$ ,  $b_G = \text{breadth}(G, H)$ , and let  $B_G$  be the cross-section that defines  $b_G$  (note area of  $B_G$  is  $b_G$ ). Let  $\mathcal{G}'_H$  be the subset of  $\mathcal{G}_H$  consisting of generators  $G$  such that  $B_G$  lies below  $J_1$ . It will suffice to show that  $\sum_{G \in \mathcal{G}'_H} b_G = O(\log n)$ , because a similar reasoning will show that  $\sum b_G = O(1)$ , where the sum is over generators  $G \in \mathcal{G}_H$  such that  $B_G$  lies above  $J_1$ .

Consider any generator  $G \in \mathcal{G}'_H$ . Observe that  $K_G$  intersects both  $J_0$  and  $J_2$ , since  $G \cap S^{\Delta_0/2}(H) \neq \emptyset$  (Property (ii)) and  $G \not\subset S^{\Delta_0}(H)$  (Property (iii)). Suppose that  $B_G$  lies between the hyperplanes  $J_i$  and  $J_{i+1}$ , for some  $i \geq 1$ . Consider the pyramid formed by joining any fixed point  $x$  in  $K_G \cap J_0$  to all the points in the base  $B_G$ . Let  $B'_G$  be the surface formed by intersecting this pyramid with the plane  $J_i$ . It follows from convexity that  $B'_G \subseteq K_G$ . Also, since the distance of  $x$  from  $J_i$  is at least half the distance of  $x$  from  $J_{i+1}$ , it follows that the area of  $B'_G$  is  $\Omega(b_G)$ .

Finally, observe that since the convex hulls of the generators of  $\mathcal{G}_H$ , and hence  $\mathcal{G}'_H$ , are disjoint, it follows that the associated set of surface patches  $\{B'_G : G \in \mathcal{G}'_H\}$  are also disjoint. Since these patches lie on the portions of the hyperplanes  $J_i, i \geq 1$ , that lie inside  $\mathbb{U}$ , and the number of these hyperplanes that intersect  $\mathbb{U}$  is  $O(\log n)$ , it follows that  $\sum_{G \in \mathcal{G}'_H} b_G = O(\log n)$ . This completes the proof.  $\square$

We can now prove Lemma 4.11 by using the above observation. The proof relies crucially on showing that, among the set of generators used to answer query  $H$ , the ones that are neither contained in  $R_H$  nor relatively fat do not cover more than a constant fraction of the points of  $P \cap S^{\Delta_0/2}(H)$ . It can be shown that the convex hulls of such generators intersect both bounding hyperplanes of the slab  $S^\Delta(H) \setminus S^{\Delta_0/2}(H)$ . We show that the convex hull disjointness assumption implies that the sum of the breadths associated with these generators is at most  $O(\log n)$ . Recall that generators that are not relatively fat cover “few” points in  $R_H$  relative to their associated breadths. Using these two facts then allows us to obtain the desired bound on the number of points covered by these generators.

PROOF. (of Lemma 4.11) Given  $H \in \mathcal{Q}$ , consider the set  $A_H \subseteq \mathcal{G}$  of generators that allow us to answer the query  $H$  within time  $t$ . Clearly,  $|A_H| \leq t$ . Consider the slab

$S^{\Delta_0/2}(H)$ . Note that  $S^{\Delta_0/2}(H) \subset S^{\Delta_0}(H) = R_H$ . Just as we showed  $|P \cap R_H| \geq 8t \log n$  (after Lemma 4.7), we can show that  $|P \cap S^{\Delta_0/2}(H)| \geq 4t \log n$ . Applying property (ii) of scattered point sets, we can easily show that  $|P \cap R_H| = O(t \log n)$ . We claim that the generators of  $A_H$  that are not fat cover at most  $2t \log n$  points of  $P \cap S^{\Delta_0/2}(H)$ . (Here and throughout this proof, fatness, absolute fatness, and relative fatness of generators are understood to be with respect to  $H$ .) Clearly this would prove the lemma as it would imply that the fat generators of  $A_H$  cover at least  $2t \log n$  points of  $P \cap S^{\Delta_0/2}(H)$ , and hence a constant fraction of the points of  $P \cap R_H$ .

To see the claim, consider the subset  $A'_H \subseteq A_H$  consisting of generators that are not fat and cover some point of  $P \cap S^{\Delta_0/2}(H)$ . Let  $A^1_H$  denote the subset of  $A'_H$  consisting of generators that are not absolutely fat. By definition, each such generator covers at most  $\log n$  points of  $P \cap R_H$ . Therefore, the generators of  $A^1_H$  cover at most  $t \log n$  points of  $P \cap R_H$ , and hence of  $P \cap S^{\Delta_0/2}(H)$ .

Let  $A^2_H = A'_H \setminus A^1_H$ . Since any generator  $G \in A^2_H$  is absolutely fat but not fat, and covers some point of  $P \cap S^{\Delta_0/2}(H)$ , it is easy to see that it satisfies the following:

- (i)  $G \subset H$ ,
- (ii)  $G \cap S^{\Delta_0/2}(H) \neq \emptyset$ ,
- (iii)  $G \not\subset R_H$ , and
- (iv)  $G$  is not relatively fat (i.e.,  $|G \cap R_H| < (\text{breadth}(G, H) \cdot \Delta_0 \cdot n) / (c \log n)$ ).

Also, by the statement of the lemma, the convex hulls of the generators in  $A^2_H$  are disjoint. Clearly, the generators of  $A^2_H$  satisfy all the preconditions of Lemma 4.12, and therefore  $\sum_{G \in A^2_H} \text{breadth}(G, H) = O(\log n)$ . Since the generators of  $A^2_H$  are not relatively fat, it follows that the number of points of  $P \cap S^{\Delta_0/2}(H)$  covered by these generators is at most

$$\begin{aligned} \sum_{G \in A^2_H} |G \cap R_H| &\leq \sum_{G \in A^2_H} \frac{\text{breadth}(G, H) \cdot \Delta_0 \cdot n}{c \log n} \\ &= \frac{\Delta_0 n}{c \log n} \sum_{G \in A^2_H} \text{breadth}(G, H). \end{aligned}$$

We know that  $\sum_{G \in A_H^2} \text{breadth}(G, H) = O(\log n)$  and  $\Delta_0 = O(t \log n/n)$ , and so by choosing the constant  $c$  (in the definition of relatively fat generators) to be sufficiently large, we can ensure that the generators of  $A_H^2$  cover at most  $t \log n$  points of  $P \cap S^{\Delta_0/2}(H)$ . Combining the contributions from the generators of  $A_H^1$  and  $A_H^2$ , it follows that the generators of  $A_H$  together cover at most  $2t \log n$  points of  $P \cap S^{\Delta_0/2}(H)$ . This completes the proof.  $\square$

Next we define a notion of generator usefulness based on the above lemma, which will help us derive the lower bounds. Consider any  $G \in \mathcal{G}$  and  $H \in \mathcal{Q}$ . We define the *usefulness* of  $G$  with respect to  $H$ , denoted  $u(G, H)$ , as in the idempotent case, except that we use “fatness” in place of “absolute fatness”. As before, let  $u(G) = E[u(G, H)]$ . Arguing as before, but using Lemma 4.11 in place of Lemma 4.9, we obtain  $\Phi \leq \sum_{G \in \mathcal{G}} u(G)$ .

Consider any  $G \in \mathcal{G}$  and  $H \in \mathcal{Q}$  such that  $G$  is fat with respect to  $H$ . Let  $K = \text{conv}(G)$ . If  $G \subset R_H$  then by Lemma 4.10,

$$u(G, H) = |G \cap R_H| \leq n \mu(K \cap R_H) = n \mu(K).$$

Otherwise,  $G$  must be relatively fat with respect to  $H$ . In this case, applying Lemma 4.10, we have

$$\mu(K \cap R_H) \geq \frac{|G \cap R_H|}{n} \geq \frac{\text{breadth}(G, H) \cdot \Delta_0}{c \log n}.$$

Thus,  $K \cap R_H$  is a cap of  $K$  of width at most  $\Delta_0$  and volume  $\Omega((\text{breadth}(G, H) \cdot \Delta_0)/\log n)$ . Also, since  $K \subseteq \mathbb{U}$  and, by Lemma 4.10,  $\mu(K) \geq \mu(K \cap R_H) = \Omega(\log n/n)$ , it is easy to show that  $K$  is  $\alpha$ -fat for  $\alpha = O(n/\log n)$ . We apply our slicing lemma (Lemma 4.5) to  $K$ , setting  $\beta$  to a suitable constant times  $\log n$  and  $\Delta$  to  $\Delta_0$ . We obtain a collection of  $O(\beta \log \alpha) = O(\log^2 n)$  convex bodies  $K_1, K_2, \dots \subseteq K$  such that one of these bodies lies inside  $K \cap R_H$  and has volume  $\Omega(K \cap R_H)$ . Arguing as before, we obtain

$$u(G, H) \leq n \sum_i \mu(K_i) \cdot \mathbf{I}[K_i \subseteq R_H].$$

By definition,  $u(G, H) = 0$  if  $G$  is not fat with respect to  $H$ . Combining all the cases, we have

$$u(G, H) \leq n \left( \mu(K) \cdot \mathbf{I}[K \subseteq R_H] + \sum_i \mu(K_i) \cdot \mathbf{I}[K_i \subseteq R_H] \right).$$

Taking expectation on both sides and arguing as before we obtain  $u(G) \leq n \log^2(n) \Delta_0^{d+1}$ , which yields

$$\Phi = O\left(mn \log^2(n) \Delta_0^{d+1}\right).$$

Combining this with the  $\Omega(t \log n)$  lower bound on  $\Phi$ , we can now easily complete the proof of Theorem 4.1(ii).

## 5. REFERENCES

- [1] P. K. Agarwal and J. Erickson. Geometric range searching and its relatives. In B. Chazelle, J. E. Goodman, and R. Pollack, editors, *Advances in Discrete and Computational Geometry*, volume 223 of *Contemporary Mathematics*, pages 1–56. American Mathematical Society, Providence, RI, 1999.
- [2] S. Arya, T. Malamatos, and D. M. Mount. On the importance of idempotence. In *Proc. 38th Annu. ACM Sympos. Theory Comput.*, pages 564–573, 2006.
- [3] S. Arya, T. Malamatos, and D. M. Mount. The effect of corners on the complexity of approximate range searching. *Discrete Comput. Geom.*, 41(3):398–443, 2009.
- [4] H. Brönnimann, B. Chazelle, and J. Pach. How hard is halfspace range searching? *Discrete Comput. Geom.*, 10:143–155, 1993.
- [5] T. M. Chan. Optimal partition trees. In *Proc. 26th Annu. ACM Sympos. Comput. Geom.*, 2010. To appear.
- [6] B. Chazelle. Lower bounds on the complexity of polytope range searching. *J. Amer. Math. Soc.*, 2:637–666, 1989.
- [7] B. Chazelle, M. Sharir, and E. Welzl. Quasi-optimal upper bounds for simplex range searching and new zone theorems. *Algorithmica*, 8:407–429, 1992.
- [8] B. Chazelle and E. Welzl. Quasi-optimal range searching in spaces of finite VC-dimension. *Discrete Comput. Geom.*, 4(5):467–489, 1989.
- [9] G. D. da Fonseca and D. M. Mount. Approximate range searching: The absolute model. *Comput. Geom. Theory Appl.*, 43(4):434–444, 2009.
- [10] H. Edelsbrunner and E. Welzl. Halfplanar range search in linear space and  $O(n^{0.695})$  query time. *Inf. Process. Lett.*, 23(6):289–293, 1986.
- [11] J. Erickson. Space-time tradeoffs for emptiness queries. *SIAM J. Comput.*, 29:1968–1996, 2000.
- [12] M. L. Fredman. Lower bounds on the complexity of some optimal data structures. *SIAM J. Comput.*, 10:1–10, 1981.
- [13] M. Grötschel, L. Lovász, and A. Schrijver. *Geometric Algorithms and Combinatorial Optimization*, volume 2 of *Algorithms and Combinatorics*. Springer-Verlag, 2nd edition, 1993.
- [14] D. Haussler and E. Welzl. Epsilon-nets and simplex range queries. *Discrete Comput. Geom.*, 2:127–151, 1987.
- [15] J. Matoušek. Range searching with efficient hierarchical cuttings. *Discrete Comput. Geom.*, 10(2):157–182, 1993.
- [16] J. Matoušek. Geometric range searching. *ACM Comput. Surv.*, 26:421–461, 1994.
- [17] J. Matoušek and E. Welzl. Good splitters for counting points in triangles. In *Proc. 5th Annu. ACM Sympos. Comput. Geom.*, pages 124–130, 1989.
- [18] E. Welzl. Partition trees for triangle counting and other range searching problems. In *Proc. 4th Annu. ACM Sympos. Comput. Geom.*, pages 23–33, 1988.
- [19] D. E. Willard. Polygon retrieval. *SIAM J. Comput.*, 11:149–165, 1982.
- [20] A. C. Yao. On the complexity of maintaining partial sums. *SIAM J. Comput.*, 14:277–288, 1985.
- [21] F. F. Yao. A 3-space partition and its applications. In *Proc. 15th Annu. ACM Sympos. Theory Comput.*, pages 258–263, 1983.