Linear-Size Approximate Voronoi Diagrams*

Sunil Arya† Theocharis Malamatos†

Abstract

Given a set S of n points in $\mathbb{R}^d$, a (t, ϵ)-*approximate Voronoi diagram (AVD)* is a partition of space into constant complexity cells, where each cell c is associated with t representative points of S, such that for any point in c, one of the associated representatives approximates the nearest neighbor to within a factor of $(1+\epsilon)$. The goal is to minimize the number and complexity of the cells in the AVD. We show that it is possible to construct an AVD consisting of $O(n/\epsilon^d)$ cells for $t = 1$, and $O(n)$ cells for $t = O(1/\epsilon^{(d-1)/2})$. In general, for a real parameter $2 \le \gamma \le 1/\epsilon$, we show that it is possible to construct a (t, ϵ)-AVD consisting of $O(n\gamma^d)$ cells for $t = O(1/(\epsilon\gamma)^{(d-1)/2})$. The cells in these AVDs are cubes or differences of two cubes. All these structures can be used to efficiently answer approximate nearest neighbor queries. Our algorithms are based on the well-separated pair decomposition and are very simple.

1 Introduction

Given a set S of n points in $\mathbb{R}^d$, the Voronoi diagram is a partition of space into cells, such that each cell consists of all points closer to a particular point of S than to any other. Voronoi diagrams are fundamental geometric objects and have a rich literature. They have numerous applications in areas such as pattern recognition and classification, machine learning, robotics, and graphics. Many of these applications are in high dimensions but, unfortunately, the complexity of Voronoi diagrams can be as high as $n^{\lceil d/2 \rceil}$ in d dimensions. This has led researchers to investigate the problem of constructing subdivisions that approximate the Voronoi diagram.

Vleugels and Overmars [12] presented an algorithm for approximating the Voronoi diagram of a disjoint set of convex sites in $\mathbb{R}^d$, and applied it to retraction motion planning. Their focus is on practical applicability, rather than on obtaining good asymptotic bounds. Har-Peled [7] considered this problem from the perspective of worst-case size, when the input is a set of points. Before stating his result, we present some definitions.

For a real parameter $\epsilon > 0$, we say that a point $p \in S$ is an ϵ-*nearest neighbor (ϵ-NN)* of a point $q \in \mathbb{R}^d$, if the distance between q and p is at most $(1 + \epsilon)$ times the distance between q and its nearest neighbor in S. We assume that distances are measured in the Euclidean metric. An *approximate Voronoi diagram (AVD)* of S is defined to be a partition of space into cells, where each cell c is associated with a *representative* $r_c \in S$, such that r_c is an ϵ-NN for all the points in c [7]. We generalize this idea in a natural way to allow for $t \ge 1$ representatives to be stored with each cell, and require that for any point in the cell, one of these t representatives is an ϵ-NN. We refer to such a decomposition as a (t, ϵ)-approximate Voronoi diagram. The goal is to minimize the size (i.e., the number of cells) of the AVD. Throughout, we will require that the cells in the AVD have constant combinatorial complexity.

Har-Peled [7] showed that it is possible to construct a $(1, \epsilon)$-AVD of $O(\frac{n}{\epsilon^d}(\log n)\log\frac{n}{\epsilon})$ size. A cell in this subdivision is the difference of two cubes (the inner cube is optional). Moreover, after preprocessing, the structure can be used to answer ϵ-NN queries in $O(\log(n/\epsilon))$ time, where the constant factor is only quadratic in dimension. This is a significantly better query time than that achieved by any previous algorithm. Recently, Sabharwal et al. [11] have given an alternative construction that reduces the size by a logarithmic factor.

In this paper we present the following results. First, we show that it is possible to construct a $(1, \epsilon)$-AVD of $O(n/\epsilon^d)$ size, which significantly improves upon the results mentioned above. Our construction is based on the well-separated pair decomposition [3] and is much simpler. As in Har-Peled's construction, a cell in this subdivision is the difference of two axis-aligned cubes, and ϵ-NN queries can be answered using this structure in $O(\log(n/\epsilon))$ time. We also present a lower bound of $\Omega(n/\epsilon^{d-1})$ on the size of a $(1, \epsilon)$-AVD, assuming that the cells are differences of two axis-aligned hyperrectangles. Thus, under this assumption, the size of our construction is nearly optimal.

Second, we generalize our construction to tackle the case when more than one representative is allowed. Given a real parameter $2 \le \gamma \le 1/\epsilon$, we show that it is possible to construct a (t, ϵ)-AVD of $O(n\gamma^d)$ size, where $t = O(1/(\epsilon\gamma)^{(d-1)/2})$. As a byproduct of our approach,

*This research was supported by the Research Grants Council of Hong Kong, China (Project no. HKUST 6158/98E).

†Department of Computer Science, The Hong Kong University of Science and Technology, Clear Water Bay, Kowloon, Hong Kong. Email: {arya,tmalamat}@cs.ust.hk.

we obtain a family of data structures that can answer ϵ-NN queries in $O(\log(n\gamma) + 1/(\epsilon\gamma)^{(d-1)/2})$ time using space $O(n(\gamma/\epsilon)^{(d-1)/2}\gamma)$. Chan [4] showed that ϵ-NN queries could be answered in $O((1/\epsilon)^{(d-1)/2}\log n)$ time using a data structure of space $O((1/\epsilon)^{(d-1)/2} n \log n)$. By setting γ to two, we obtain a data structure that answers queries in $O(\log n + 1/\epsilon^{(d-1)/2})$ time using space $O((1/\epsilon)^{(d-1)/2} n)$. Thereby, we improve upon Chan's result, both in terms of space and query time.

We mention some other tradeoffs between space and query time that are known for the approximate nearest neighbor problem. Arya et al. [2] and, later, Duncan et al. [6] provided data structures that achieve $O((1/\epsilon)^d \log n)$ query time and use $O(n)$ space (independent of ϵ). Recently, Kushilevitz et al. [9] and Indyk and Motwani [8] have obtained algorithms that eliminate exponential dependencies on dimension in both query time and space. The space required by their algorithms is polynomial in d and n, and the query time is polynomial in $\log n, d$, and $1/\epsilon$. However, the space grows exponentially with $1/\epsilon$.

2 Preliminaries

Throughout we assume that the dimension d is a fixed constant, and the constants hidden in the asymptotic bounds may depend on d (but not on ϵ or γ).

Let x and y denote any two points in $\mathbb{R}^d$. We use $|xy|$ to denote the Euclidean distance between x and y, $\overline{xy}$ to denote the segment joining x and y, and $\overrightarrow{xy}$ to denote the vector from x to y.

We denote by $b(x, r)$ a ball of radius r centered at x, i.e, $b(x, r) = \{y : |xy| \leq r\}$. For a ball b and any positive real γ, we use γb to denote the ball with the same center as b and whose radius is γ times the radius of b, and $\overline{b}$ to denote the set of points that are not in b.

Let X and X' be two sets of points. We say that X' is δ-dense for X if, for any point $x \in X$, there is a point $x' \in X'$ such that $|xx'| \leq \delta$.

We briefly review the notions of well-separated pair decomposition and balanced box-decomposition trees, as they play an important role in our constructions.

The well-separated pair decomposition. Let S be a set of n points in $\mathbb{R}^d$. We say that two sets of points X and Y are *well-separated* if they can be enclosed within two disjoint d-dimensional balls of radius r, such that the distance between the centers of these balls is at least αr, where $\alpha \geq 2$ is a real parameter called the *separation factor*. A *well-separated pair decomposition (WSPD)* of S is a set $\mathcal{P}_{S,\alpha} = \{(X_1, Y_1), \cdots, (X_m, Y_m)\}$ of pairs of subsets of S such that (i) for $1 \leq i \leq m$, X_i and Y_i are well-separated and (ii) for any distinct points $x, y \in S$, there exists a unique pair (X_i, Y_i) such that either $x \in X_i$ and $y \in Y_i$ or $x \in Y_i$ and $y \in X_i$. (We say that the pair (X_i, Y_i) *separates* x and y.) Callahan and Kosaraju [3] have shown that we can construct a WSPD containing $O(\alpha^d n)$ pairs in $O(n \log n + \alpha^d n)$ time. For each pair, their construction also provides the d-balls enclosing X_i and Y_i satisfying the separation criteria mentioned above.

The BBD tree. Let $U = [0, 1]^d$ denote a unit hypercube in $\mathbb{R}^d$. We define a *quadtree box* recursively as follows: (i) U is a quadtree box, and (ii) any hypercube obtained by splitting a quadtree box into 2^d equal parts is a quadtree box. The *size* of a quadtree box is its side length. A nice property of quadtree boxes is that any two quadtree boxes are either disjoint or one is contained inside the other.

The *balanced box-decomposition (BBD) tree* is a balanced 2^d-ary tree that compactly represents a hierarchical decomposition of space [2]. Each node of the tree is associated with a region of space called a *cell*, which is the difference of two quadtree boxes, an *outer box* and an (optional) *inner box*. The root of the tree is associated with U. The cell associated with any node is partitioned into disjoint cells, which are associated with the children of the node. (For details see [2].) We define the *size* of a cell to be same as the size of its outer box.

We will use the following facts in this paper. Given any collection $\mathcal{C}$ of quadtree boxes, we can store them in a BBD tree having $O(|\mathcal{C}|)$ nodes and $O(\log |\mathcal{C}|)$ depth. The time to construct this tree is $O(|\mathcal{C}| \log |\mathcal{C}|)$. The subdivision induced by its leaves is a refinement of the subdivision induced by the quadtree boxes in $\mathcal{C}$ and, for any point q, we can determine the leaf containing q in time proportional to the depth of the tree.

3 Approximate Voronoi Diagrams: Single Representative

Let S be a set of n points, and let $0 < \epsilon \leq 1/2$ be a real parameter. In this section, we show how to construct a $(1, \epsilon)$-approximate Voronoi diagram for S.

Let U' be a hypercube that encloses S, such that the distance of any point in S from the boundary of U' is at least D/ϵ, where D is the diameter of S. Note that for a query point outside U', any point in S works as an ϵ-NN, so it suffices to show how to construct the approximate Voronoi diagram inside U'. It will be convenient to assume that $U' = U = [0, 1]^d$, which can be easily ensured by scaling and translating the coordinate system.

We construct a WSPD $\mathcal{P}_{S,\alpha}$ for S, using separation factor $\alpha = 8$. Note that the number of pairs in $\mathcal{P}_{S,8}$ is $O(n)$. For each pair $P \in \mathcal{P}_{S,8}$, we compute a set of quadtree boxes $\mathcal{C}_P$ as follows. Let $P = (X, Y)$ and let x and y denote the centers of the balls enclosing X

and Y, respectively, that satisfy the separation criteria. Let $\ell = |xy|$, and let $\mathcal{B}_P$ denote the set of balls of radius $2^i\ell$ for $-2 \le i \le \lceil\log(1/\epsilon)+1\rceil$, centered at x and y. For a ball $b \in \mathcal{B}_P$, let $\mathcal{C}_b$ be the set of quadtree boxes overlapping b that have the largest size not exceeding $r_b\epsilon/(16d)$, where r_b denotes the radius of b. Note that $|\mathcal{C}_b| = O(1/\epsilon^d)$. Let $\mathcal{C}_P = \cup_{b\in\mathcal{B}_P}\mathcal{C}_b$ and $\mathcal{C} = \cup_{P\in\mathcal{P}_{S,8}}\mathcal{C}_P$. Clearly $|\mathcal{C}_P| = O(\frac{1}{\epsilon^d}\log\frac{1}{\epsilon})$ and $|\mathcal{C}| = O(\frac{n}{\epsilon^d}\log\frac{1}{\epsilon})$. Finally we store all the quadtree boxes in $\mathcal{C}$ in a BBD tree T. Note that the depth of T is $O(\log(n/\epsilon))$. With each leaf cell of T, we store a *representative* which is an $(\epsilon/4)$-NN of any point inside it.

We claim that the subdivision associated with the leaves of T along with the stored representatives is a $(1,\epsilon)$-approximate Voronoi diagram for S. The proof of this claim employs the following lemma.

LEMMA 3.1. *Let S be a set of n points in $\mathbb{R}^d$ and let $0 < \epsilon \le 1/2$ be a real parameter. Let x_1 be a point inside a d-cube c of size $(\epsilon/(4d))\cdot|x_1y_1|$, where y_1 denotes the nearest neighbor of x_1. If y_2 is an $(\epsilon/4)$-NN of some point x_2 inside c, then y_2 is an ϵ-NN of x_1.*

Proof Note that the diameter δ of cube c is no more that $(\epsilon/4)\cdot|x_1y_1|$. By the triangle inequality, we get

$$|x_1y_2| \le |x_2y_2| + \delta. \tag{3.1}$$

Since y_2 is an $(\epsilon/4)$-NN of x_2, it follows that

$$|x_2y_2| \le (1+\epsilon/4)\cdot|x_2y_1|. \tag{3.2}$$

Again, by the triangle inequality, we get

$$|x_2y_1| \le |x_1y_1| + \delta. \tag{3.3}$$

Using Eqs. (3.1), (3.2), and (3.3), and the bound on δ, we get $|x_1y_2| \le (1+3\epsilon/4+\epsilon^2/16)\cdot|x_1y_1|$. Since $\epsilon \le 1/2$, it follows that $|x_1y_2| \le (1+\epsilon)\cdot|x_1y_1|$. □

LEMMA 3.2. *The subdivision formed by the leaves of T, along with the representatives stored with the leaves, is a $(1,\epsilon)$-approximate Voronoi diagram for S.*

Proof In view of the transformation of the coordinate system mentioned above, all the points of S can be assumed to lie inside $U = [0,1]^d$. Let $q \in U$ be an arbitrary point. Let c denote the leaf cell that contains q and let x denote the representative stored with c. Let $y \in S$ denote the nearest neighbor of q. We will show that if $x \ne y$, then x is an ϵ-NN of q.

Let $P = (X,Y)$ be the pair in the WSPD $\mathcal{P}_{S,8}$ that separates x and y. Without loss of generality, assume that $x \in X$ and $y \in Y$. Let x' and y' denote the centers of the balls enclosing X and Y, respectively, that satisfy the separation criteria. Recall that the radius of these balls is at most $\ell/8$, where $\ell = |x'y'|$. We distinguish three cases: (1) $|qy'| \ge 2\ell/\epsilon$, (2) $\ell/4 \le |qy'| < 2\ell/\epsilon$, and (3) $|qy'| < \ell/4$.

Case 1: $|qy'| \ge 2\ell/\epsilon$.

By the triangle inequality, we get $|qx| \le |qy'| + |y'x'| + |x'x| \le |qy'| + 9\ell/8$, and $|qy| \ge |qy'| - |y'y| \ge |qy'| - \ell/8$. Using the facts that $0 < \epsilon \le 1/2$, and $|qy'| \ge 2\ell/\epsilon$, we can now easily show that $|qx| \le (1+\epsilon)\cdot|qy|$.

Case 2: $\ell/4 \le |qy'| < 2\ell/\epsilon$.

Note that there is a ball in $\mathcal{B}_P$ of radius at most $2|qy'|$ that overlaps q, which implies that there is a quadtree box $\hat{c} \in \mathcal{C}_P$ of size $s_{\hat{c}} \le (\epsilon/(8d))\cdot|qy'|$ that contains q. Since $|qy'| \ge \ell/4$ and $|qy'| \le |qy| + \ell/8$, it follows that $|qy'| \le 2|qy|$. Thus $s_{\hat{c}} \le (\epsilon/(4d))\cdot|qy|$. Further, by construction, x is an $(\epsilon/4)$-NN of some point $z \in \hat{c}$. Therefore, by Lemma 3.1, it follows that x is an ϵ-NN of q.

Case 3: $|qy'| < \ell/4$.

Recall that there is a ball in $\mathcal{B}_P$ centered at y' having radius $\ell/4$ that overlaps q. This implies that there is a quadtree box $\hat{c} \in \mathcal{C}_P$ of size $s_{\hat{c}} \le (\epsilon/(16d))\cdot(\ell/4)$ that contains q. Also, by construction, x is an $(\epsilon/4)$-NN of some point z in $\hat{c}$. Applying the triangle inequality, $|yz| \le |yy'| + |y'q| + |qz| \le \ell/8 + \ell/4 + \ell\epsilon/64$, and $|xz| \ge |y'x'| - |x'x| - |y'q| - |qz| \ge \ell - \ell/8 - \ell/4 - \ell\epsilon/64$. Since $\epsilon \le 1/2$, it follows that $|yz| \le 49\ell/128$ and $|xz| \ge 79\ell/128$. Thus $|xz|/|yz| \ge 79/49$. This contradicts the fact that x is an $(\epsilon/4)$-NN of z, given that ϵ is at most $1/2$. Hence, Case 3 cannot occur. This completes the proof. □

We summarize the main result of this section.

THEOREM 3.1. *Let S be a set of n points in $\mathbb{R}^d$, and let $0 < \epsilon \le 1/2$ be a real parameter. Then we can construct a $(1,\epsilon)$-approximate Voronoi diagram for S that consists of $O(\frac{n}{\epsilon^d}\log\frac{1}{\epsilon})$ regions, where each region is the difference of two cubes. Moreover, for any query point, we can return its ϵ-NN in $O(\log(n/\epsilon))$ time.*

In Section 4.1, we will give a method which can be used to reduce the size of the $(1,\epsilon)$-AVD in this theorem by a factor of $\log(1/\epsilon)$.

4 Approximate Voronoi Diagrams: Multiple Representatives

Let S be a set of n points in $\mathbb{R}^d$, and let $0 < \epsilon \le 1/2$ and $2 \le \gamma \le 1/\epsilon$ be two real parameters. In this section we show how to construct a (t,ϵ)-approximate Voronoi diagram for $t = O(1/(\epsilon\gamma)^{(d-1)/2})$. We will present two methods. The first method is easier to analyze and yields an AVD of size $O(n\gamma^d \log \gamma)$. The second method reduces the size to $O(n\gamma^d)$ by exploiting some additional observations. Before giving the constructions, we present four technical lemmas.

LEMMA 4.1. (Chan and Snoeyink [5]) *Let $\triangle xyz$ be a triangle with $\angle xzy = \theta, \angle yxz = \phi$, and $\angle xyz \geq \pi/2$. Then*

$$|xy| + |yz| \leq (1 + \sin\theta \sin\phi)|xz|.$$

LEMMA 4.2. *Let $\triangle xyz$ be a triangle with $\angle yxz = \theta$ and $\max(|xy|, |xz|) \geq |yz|$. Then $\theta \leq \pi/2$.*

Proof Follows from the fact that x must lie outside the circle with $|yz|$ as diameter. □

LEMMA 4.3. *Let $\triangle xyz$ be a triangle with $\angle yxz = \theta$. Then $sin\theta \leq |yz|/\max(|xy|, |xz|)$.*

Proof It follows from the law of sines that $\sin\theta/|yz| \leq 1/|xy|$ and $\sin\theta/|yz| \leq 1/|xz|$. Thus $\sin\theta \leq |yz|/\max(|xy|, |xz|)$. □

Given a set X of points and a point q, let $NN_q(X)$ be the distance to the nearest neighbor of q in X. (If there are no points in X, then $NN_q(X)$ is defined to be infinity.)

LEMMA 4.4. *Let S be a set of n points in $\mathbb{R}^d$. Let $0 < \epsilon \leq 1/2$ and $\gamma \geq 2$ be two real parameters. Let b_1 and b_2 denote two concentric balls of radius r and γr, respectively. There exists a set $R \subseteq S$ consisting of $\left(1 + O\left(\frac{1}{\sqrt{\epsilon\gamma}}\right)\right)^{d-1}$ points such that*

(i) for any point $q \in b_1$, $NN_q(R) \leq (1+\epsilon) \cdot NN_q(S \cap \overline{b_2})$, and

(ii) for any point $q \in \overline{b_2}$, $NN_q(R) \leq (1+\epsilon) \cdot NN_q(S \cap b_1)$.

Proof We only prove (i), since the proof of (ii) is similar. Let b_3 be the ball $(3/2)b_1$. First, we consider the case $\gamma > 16/\epsilon$. Let x be any point in b_1 and let n_x be any point of S that is its $(\epsilon/2)$-NN. Along the lines of Lemma 3.1, we can easily show that, for any point $q \in b_1$, $|qn_x| \leq (1 + \epsilon) \cdot NN_q(S \cap \overline{b_2})$. Thus (i) holds for $R = \{n_x\}$.

In the remainder we assume that $\gamma \leq 16/\epsilon$. Let R' be a set of points on the boundary of b_3 that is δ-dense for the boundary, where $\delta = r\sqrt{\epsilon\gamma}/16$. By standard results [13], we can find such a set R' of size $O(1/(\epsilon\gamma)^{(d-1)/2})$. For each point $x \in R'$, we let n_x denote any point of S that is its $(\epsilon/2)$-NN. We define $R = \{n_x : x \in R'\}$. We now show that R satisfies the property given in part (i) of the lemma.

Let q be a point in b_1. Let p denote the nearest neighbor of q among the points of $S \cap \overline{b_2}$. Let y denote the point of intersection of $\overline{qp}$ with the boundary of b_3. Let x be the point in R' that is closest to y. We will show that $|qn_x| \leq (1 + \epsilon)|qp|$, which will imply (i).

By the triangle inequality, we have $|qn_x| \leq |qx| + |xn_x|$. Since n_x is an $(\epsilon/2)$-NN of x, we have $|xn_x| \leq (1 + \epsilon/2)|px|$. Thus

$$|qn_x| \leq (1 + \epsilon/2)(|qx| + |px|). \quad (4.4)$$

In the triangle $\triangle qpx$, let θ denote $\angle pqx$ and ϕ denote $\angle qpx$. We claim that $\angle qxp \geq \pi/2, \sin\theta \leq \frac{\sqrt{\epsilon\gamma}}{8}$, and $\sin\phi \leq \frac{1}{4}\sqrt{\frac{\epsilon}{\gamma}}$. Assuming this claim for now and applying Lemma 4.1, we get $|qx| + |px| \leq (1 + \epsilon/32)|qp|$. Substituting this in Eq. (4.4), and noting that $\epsilon \leq 1/2$, we get

$$|qn_x| \leq (1 + \epsilon/2)(1 + \epsilon/32)|qp| \leq (1 + \epsilon)|qp|,$$

which is the desired result.

To prove the claim, consider $\triangle yqx$. Since R' is δ–dense for the boundary of b_3, we have $|xy| \leq \delta = r\sqrt{\epsilon\gamma}/16$. Also, $\max(|qx|, |qy|) \geq r/2$. Thus $|xy|/\max(|qx|, |qy|) \leq \sqrt{\epsilon\gamma}/8$. By Lemma 4.3, it follows that $\sin\theta \leq \sqrt{\epsilon\gamma}/8$. Using Lemma 4.2 and the fact that that $\gamma \leq 16/\epsilon$, it is easy to see that $\theta \leq \pi/6$. We next consider $\triangle ypx$. Note that $\max(|px|, |py|) \geq (\gamma - 3/2)r \geq \gamma r/4$, since $\gamma \geq 2$. Thus $|xy|/\max(|px|, |py|) \leq \frac{1}{4}\sqrt{\frac{\epsilon}{\gamma}}$. By Lemma 4.3, $\sin\phi \leq \frac{1}{4}\sqrt{\frac{\epsilon}{\gamma}}$. Applying Lemma 4.2 and noting that $\epsilon \leq 1/2$ and $\gamma \geq 2$, it follows that $\phi \leq \pi/6$. Since $\theta \leq \pi/6$ and $\phi \leq \pi/6$, we have $\angle qxp \geq \pi/2$, which completes the proof. □

We now describe the first method for constructing the approximate Voronoi diagram. This can be viewed as a generalization of the approach given in Section 3. Indeed, the construction of the subdivision itself is very similar, but the analysis of its properties, and the method used to assign representatives, are different.

Let $\gamma \geq 2$ and $\beta \geq 2$ be two real parameters. As in Section 3, we first transform the coordinate system to ensure that the points of S lie inside $U = [0, 1]^d$, and for any query point outside U, any point of S is an ϵ-NN. We then compute a WSPD $\mathcal{P}_{S,\alpha}$ for S using separation factor $\alpha = 4$. For each pair $P \in \mathcal{P}_{S,4}$, we compute a set of quadtree boxes as follows. Let $P = (X, Y)$, and let x and y denote the centers of the balls enclosing X and Y, respectively, that satisfy the separation criteria. Let $\ell = |xy|$, and let $\mathcal{B}_P$ denote the set of balls of radius $2^i\ell$ for $3 \leq i \leq \lceil\log\beta + 2\rceil$, centered at x and y. For a ball $b \in \mathcal{B}_P$, let $\mathcal{C}_b$ be the set of quadtree boxes overlapping b that have the largest size not exceeding $\Delta_b = r_b/(32\gamma d)$, where r_b denotes the radius of b. Note that $\mathcal{C}_b = O(\gamma^d)$. Let $\mathcal{C}_P = \cup_{b\in\mathcal{B}_P}\mathcal{C}_b$ and $\mathcal{C} = \cup_{P\in\mathcal{P}_{S,4}}\mathcal{C}_P$. Clearly $|\mathcal{C}_P| = O(\gamma^d \log\beta)$ and $|\mathcal{C}| = O(n\gamma^d \log\beta)$. Finally we store all the boxes in $\mathcal{C}$ in a BBD tree T. We will show that the subdivision induced by the leaves of T (for β to be specified later), along with suitably chosen representatives, is the desired approximate Voronoi diagram. Before describing how to choose the representatives, we need to identify a key property that is satisfied by this subdivision.

LEMMA 4.5. *Let S be a set of n points in $\mathbb{R}^d$, and let $\gamma \geq 2, \beta \geq 2$ be two real parameters. Then it is possible to construct a subdivision consisting of $O(n\gamma^d \log\beta)$ cells, where each cell c is the difference of two cubes and satisfies the following property. Let $c = c_O - c_I$, where c_O and c_I denote the outer and inner cube of c, respectively; let s denote the size of c; and let b_c be the ball of radius $sd/2$ whose center coincides with the center of c_O (note that $c \subseteq b_c$). Then either $|S\cap\gamma b_c| \leq 1$ or there exists a ball b'_c such that $S \cap \gamma b_c \subseteq b'_c$ and the ball $\beta b'_c$ does not overlap c.*

Proof We claim that the subdivision induced by the leaves of the BBD tree T satisfies the desired condition. The $O(n\gamma^d \log\beta)$ bound on the number of cells in the subdivision follows from the above discussion.

Let $c = c_O - c_I$ be any leaf cell. If $|S \cap \gamma b_c| \leq 1$, there is nothing to prove. So suppose that $|S\cap\gamma b_c| > 1$. Let x and y denote the farthest pair of points in $S\cap\gamma b_c$. Let b'_c be the ball of radius $|xy|$ centered at x. Clearly, all the points in $S \cap \gamma b_c$ are contained within b'_c.

It remains to show that the ball $\beta b'_c$ does not overlap c. To this end, consider the pair $P = (X, Y)$ in the WSPD $\mathcal{P}_{S,4}$ that separates x and y. Without loss of generality assume that $x \in X$ and $y \in Y$. Let x' and y' denote the centers of the balls enclosing X and Y, respectively, that satisfy the separation criteria. Let $\ell = |x'y'|$. Using the definition of well-separatedness and the triangle inequality, we have $|xx'| \leq \ell/4$, $|yy'| \leq \ell/4$, and $\ell/2 \leq |xy| \leq 3\ell/2$.

We distinguish three cases based on the closest distance L between cell c and x': (1) $L > 4\beta\ell$, (2) $L < 8\ell$, and (3) $8\ell \leq L \leq 4\beta\ell$.

Case 1: $L > 4\beta\ell$.

For any point $z \in c$, we have $|zx'| > 4\beta\ell$. By the triangle inequality, we get $|zx| \geq |zx'| - |xx'| > 4\beta\ell - \ell/4 > 31\beta\ell/8$, since $\beta \geq 2$. Since $|xy| \leq 3\ell/2$, we get $|zx| > 31\beta|xy|/12$. Thus, in this case, ball $\beta b'_c$ does not overlap c.

Case 2: $L < 8\ell$.

Let z be any point in $c \cap b(x', 8\ell)$. Recall that $b(x', 8\ell) \in \mathcal{B}_P$, and so there must be a quadtree box $\hat{c} \in \mathcal{C}_{b(x',8\ell)}$ of size $s_{\hat{c}} \leq 8\ell/(32\gamma d)$ that contains z. Since $\ell \leq 2|xy|$ and $|xy|$ is at most $sd\gamma$ (because both x and y are contained in the ball γb_c), we get $s_{\hat{c}} \leq s/2$. This is a contradiction since, by construction of the BBD tree T, $c \subseteq \hat{c}$ and c has size s. Thus this case cannot occur.

Case 3: $8\ell \leq L \leq 4\beta\ell$.

Clearly c must overlap a ball $b \in \mathcal{B}_P$ of radius r_b satisfying $r_b \leq 2L$. Let z be any point in $c \cap b$. By construction z must be contained in a quadtree box $\hat{c} \in \mathcal{C}_b$ of size $s_{\hat{c}} \leq r_b/(32\gamma d) \leq 2|zx'|/(32\gamma d)$. By the triangle inequality, we have $|zx'| \leq |zx| + |xx'|$. Further $|xx'| \leq \ell/4 \leq |zx'|/32$. Thus $|zx'| \leq 32|zx|/31$. Since z and x are both contained in γb_c, we have $|zx| \leq sd\gamma$. Thus $s_{\hat{c}} \leq 2s/31$, which is a contradiction for the same reason as in Case 2. Thus this case too cannot occur. This completes the proof. □

Given a real number $2 \leq \gamma \leq 1/\epsilon$, we construct the subdivision described in Lemma 4.5, for $\beta = \gamma$. Lemma 4.4 suggests the following approach for assigning representatives to the leaf cells. Let q be a point inside a leaf cell c. Let b_c and b'_c be the balls defined in Lemma 4.5. Since c is contained within the ball b_c, applying Lemma 4.4(i), it follows that we can find a set R'_c consisting of $O(1/(\epsilon\gamma)^{(d-1)/2})$ points such that $NN_q(R'_c) \leq (1+\epsilon)\cdot NN_q(S \cap \overline{\gamma b_c})$. For the points inside γb_c we proceed as follows. If $|S \cap \gamma b_c| \leq 1$, we define $R''_c = S \cap \gamma b_c$. Otherwise, by Lemma 4.5, $S \cap \gamma b_c \subseteq b'_c$ and $\gamma b'_c$ does not overlap c. Thus, applying Lemma 4.4(ii), it follows that we can find a set R''_c consisting of $O(1/(\epsilon\gamma)^{(d-1)/2})$ points such that $NN_q(R''_c) \leq (1+\epsilon)\cdot NN_q(S \cap b'_c) \leq (1+\epsilon)\cdot NN_q(S \cap \gamma b_c)$. Finally we assign $R_c = R'_c \cup R''_c$ to be set of representatives for c. Clearly, R_c has size $O(1/(\epsilon\gamma)^{(d-1)/2})$ and satisfies the desired property, namely, $NN_q(R_c) \leq (1+\epsilon)\cdot NN_q(S)$. In summary, we have shown that we can construct an $(O(1/(\epsilon\gamma)^{(d-1)/2}), \epsilon)$-approximate Voronoi diagram for S that consists of $O(n\gamma^d \log\gamma)$ cells. In the following subsection, we will show how to improve the bound on the size by a factor of $\log\gamma$.

4.1 Size Reduction

Our improved construction is exactly the same as the first one, except that the sizes of the quadtree boxes generated by each pair in the WSPD grow quadratically instead of linearly with distance from the pair. To be precise, we modify the value of the parameter Δ_b from $r_b/(32\gamma d)$ to $r_b^2/(256\ell\gamma d)$. With this change, for a ball $b \in \mathcal{B}_P$, we have $|\mathcal{C}_b| = O((\ell\gamma/r_b + 1))^d$. We set $\beta = \gamma$, as in the first construction. It is easy to see that $|\mathcal{C}_P| = \sum_{b\in\mathcal{B}_P}|\mathcal{C}_b| = O(\gamma^d)$. Thus $|\mathcal{C}| = O(n\gamma^d)$, and so the number of leaves in the resulting BBD tree

T is less than that obtained from the first method by a factor of $\log \gamma$. We will show that the leaves of T, along with suitable representatives, form a (t, ϵ)-AVD for $t = O(1/(\epsilon\gamma)^{(d-1)/2})$.

The next two lemmas use ideas similar to Lemmas 4.4 and 4.5, respectively.

LEMMA 4.6. *Let S be a set of n points in $\mathbb{R}^d$. Let $0 \le \epsilon \le 1/2$ be a real parameter. Let b_1 and b_2 be two disjoint balls of radius r_1 and r_2, respectively, whose minimum distance of separation is at least ℓ'. Further, suppose that $\ell' \ge \max(r_1, r_2)$. Then there exists a set $R \subseteq S$ consisting of $\left(1 + O\left(\frac{\sqrt{r_1 r_2}}{\ell' \sqrt{\epsilon}}\right)\right)^{d-1}$ points such that for any point $q \in b_1$, $NN_q(R) \le (1+\epsilon) \cdot NN_q(S \cap b_2)$.*

Proof We will assume that $r_1 \le r_2 \le \ell'$. The proof for the case $r_2 \le r_1 \le \ell'$ is similar and is omitted.

First, we consider the case $r_1 < r_2\epsilon/16$. Let x be any point in b_1 and let n_x be any point of S that is its $(\epsilon/2)$-NN. Along the lines of Lemma 3.1, we can easily show that, for any point $q \in b_1$, $|qn_x| \le (1+\epsilon) \cdot NN_q(S \cap b_2)$. Thus the lemma holds for $R = \{n_x\}$.

In the remainder we assume that $r_2\epsilon/16 \le r_1 \le r_2$. In order to describe the set R, we need some definitions. Let o_1 and o_2 denote the centers of the balls b_1 and b_2, respectively. The line $o_1 o_2$ intersects the boundary of b_1 at two points; let z_1 denote that point of intersection that is farther from o_2. Similarly line $o_1 o_2$ intersects the boundary of b_2 at two points; let z_2 denote that point of intersection that is closer to o_1. (See Fig. 1.) Let Λ_1 and Λ_2 denote the $(d-1)$-disks orthogonal to line $o_1 o_2$, centered at z_1 with radius r_1, and centered at z_2 with radius r_2, respectively. We define Π to be the truncated cone with bases Λ_1 and Λ_2 (that is, Π contains points that lie on any line segment joining a point in Λ_1 with a point in Λ_2). Let h be the hyperplane orthogonal to $\overline{o_1 o_2}$, and at distance $r_1 + \ell' r_1/(2r_2)$ from o_1 (towards o_2), and let Λ denote the $(d-1)$-disk $\Pi \cap h$. We choose R' to be any set of points on h that is δ-dense for Λ, where $\delta = \frac{\ell'}{16}\sqrt{\frac{\epsilon r_1}{r_2}}$. Note that we can easily compute a set R' satisfying this property and having size $(1+O(r/\delta))^{d-1}$, where r denotes the radius of disk Λ. Finally, we put $R = \{n_x : x \in R'\}$, where n_x denotes any $(\epsilon/2)$-NN of x.

To bound the size of R, observe that the intersection of a hyperplane orthogonal to line $o_1 o_2$ (that is, the axis of the truncated cone Π) with Π is a $(d-1)$-disk whose radius increases linearly as we traverse from z_1 to z_2. Thus

$$\begin{aligned} r &= r_1 + \frac{2r_1 + \ell' r_1/(2r_2)}{2r_1 + \ell'}(r_2 - r_1) \\ &\le r_1 + \frac{2r_1 + \ell' r_1/(2r_2)}{\ell'} r_2 \;\le\; \frac{7}{2} r_1, \end{aligned}$$

since $r_2 \le \ell'$. Recalling that $|R| \le |R'| = (1 + O(r/\delta))^{d-1}$, and using this bound on r, we easily obtain the bound on $|R|$ given in the statement of the lemma.

We next show that R satisfies the property given in the statement of the lemma. Let q be any point in b_1 and let p be its nearest neighbor among the points of $S \cap b_2$. It is easy to see that segment $\overline{qp}$ must intersect disk Λ. Let y denote this point of intersection. Since R' is δ-dense for Λ, there must be a point $x \in R'$ such that $|xy| \le \delta$. We will show that $|qn_x| \le (1+\epsilon)|qp|$.

By the triangle inequality, we have $|qn_x| \le |qx| + |xn_x|$. Since n_x is an $(\epsilon/2)$-NN of x, we have $|xn_x| \le (1+\epsilon/2)|px|$. Thus $|qn_x| \le (1+\epsilon/2)(|qx| + |px|)$. In $\triangle qpx$, let θ denote $\angle pqx$ and ϕ denote $\angle qpx$. We claim that $\angle qxp \ge \pi/2, \sin\theta \le \frac{1}{8}\sqrt{\frac{\epsilon r_2}{r_1}}$, and $\sin\phi \le \frac{1}{8}\sqrt{\frac{\epsilon r_1}{r_2}}$. Assuming this for now and applying Lemma 4.1, we get $|qx| + |px| \le (1+\epsilon/64)|qp|$. Thus $|qn_x| \le (1+\epsilon/2)(1+\epsilon/64)|qp| \le (1+\epsilon)|qp|$, as desired.

It remains only to prove the above claim. To this end, note that $|xy| \le \delta = \frac{\ell'}{16}\sqrt{\frac{\epsilon r_1}{r_2}}$. Also, $\max(|qx|, |qy|) \ge \ell' r_1/(2r_2)$ and $\max(|px|, |py|) \ge \ell' - \ell' r_1/(2r_2) \ge \ell'/2$. Applying Lemma 4.3 to $\triangle yqx$ and $\triangle ypx$, respectively, the desired bounds on $\sin\theta$ and $\sin\phi$ easily follow. Finally, noting that $r_2\epsilon/16 \le r_1 \le r_2$, $\epsilon \le 1/2$, and using Lemma 4.2 on $\triangle yqx$ and $\triangle ypx$, we can easily show that $\theta \le \pi/6$ and $\phi \le \pi/6$. Thus $\angle qxp \ge \pi/2$, which completes the proof. □

LEMMA 4.7. *Let S be a set of n points in $\mathbb{R}^d$, and let $\gamma \ge 2$ be a real parameter. Then it is possible to construct a subdivision consisting of $O(n\gamma^d)$ cells, where each cell c is the difference of two cubes and satisfies the following property. Let $c = c_O - c_I$, where c_O and c_I denote the outer and inner cube of c, respectively; let s denote the size of c; and let b_c be the ball of radius $r_1 = sd/2$ whose center coincides with the center of c_O (note that $c \subseteq b_c$) . Then one of the following three possibilities must hold:*

(i) $|S \cap \gamma b_c| \le 1$.

(ii) There exists a ball b'_c such that $S \cap \gamma b_c \subseteq b'_c$ and the ball $\gamma b'_c$ does not overlap c.

(iii) There exists a ball b'_c such that $S \cap \gamma b_c \subseteq b'_c$ and $\ell' \ge \max(r_1, r_2)$ and $\ell'/\sqrt{r_1 r_2} \ge 5\sqrt{\gamma}$. Here ℓ' denotes the minimum distance of separation between b_c and b'_c, and r_2 denotes the radius of ball b'_c.

Proof We claim that the subdivision induced by the leaves of the BBD tree T satisfies the desired condition.

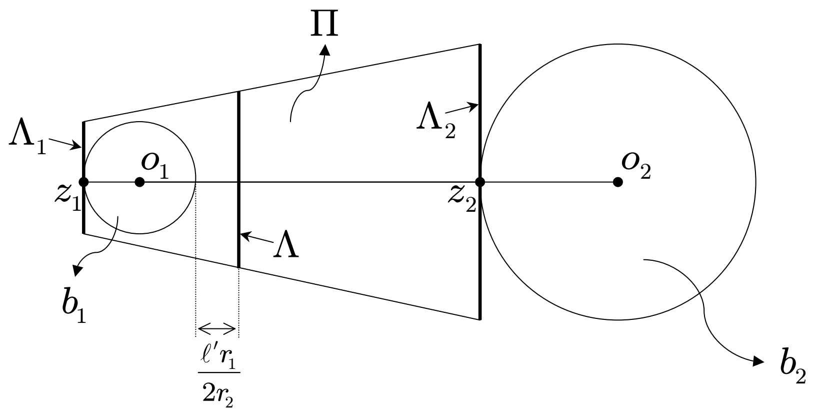

Fig. 1: Proof of Lemma 4.6.

The bound on the number of cells follows from the remarks given in the first paragraph of this subsection.

We borrow all the notation from the proof of Lemma 4.5, and set $\beta = \gamma$. The reader can easily check that the arguments given in the proof of Lemma 4.5 apply, except for a crucial difference in the argument for Case 3, which is the only case we will consider here.

Case 3: $8\ell \le L \le 4\gamma\ell$.

We will show that possibility (iii) holds in this case. To this end, we estimate r_1, r_2, and ℓ'.

Clearly c must overlap a ball $b \in \mathcal{B}_P$ of radius r_b satisfying $r_b \le 2L$. Let z be any point in $c \cap b$. By construction z must be contained in a quadtree box $\hat{c} \in \mathcal{C}_b$ of size $s_{\hat{c}} \le r_b^2/(k\ell\gamma d) \le 4L^2/(k\ell\gamma d)$, where $k = 256$. Since $c \subseteq \hat{c}$, this bound on $s_{\hat{c}}$ also applies to s. Thus $r_1 = sd/2 \le 2L^2/(k\ell\gamma)$. Since $L \le 4\gamma\ell$, we obtain $r_1 \le 8L/k$.

Further, $r_2 = |xy| \le 3\ell/2$. Since $L \ge 8\ell$, we obtain $r_2 \le 3L/16$.

Since $z \in b_c$ and $x' \in b'_c$, by the triangle inequality, ℓ', the minimum distance of separation between b_c and b'_c is at least $L - 2r_1 - 2r_2 \ge L - 16L/k - 3L/8$. Substituting $k = 256$ in the above, we get $r_1 \le L/32, r_2 \le 3L/16$, and $\ell' \ge 9L/16$. Thus $\ell' \ge \max(r_1, r_2)$. Further,

$$\frac{\ell'}{\sqrt{r_1 r_2}} \ge \frac{9L/16}{\sqrt{\frac{2L^2}{256\ell\gamma} \cdot \frac{3\ell}{2}}} \ge 5\sqrt{\gamma},$$

as desired. □

We assign representatives to the leaves of T as follows. Let q be a point inside a leaf cell c. Let b_c and b'_c be the balls defined in Lemma 4.7. Since c is contained within the ball b_c, applying Lemma 4.4(i), it follows that we can find a set R'_c consisting of $O(1/(\epsilon\gamma)^{(d-1)/2})$ points such that $NN_q(R'_c) \le (1+\epsilon) \cdot NN_q(S \cap \overline{\gamma b_c})$. For the points inside γb_c we proceed as follows. Note that one of the three cases given in the statement of Lemma 4.7 must hold. If Case (i) holds (that is, $|S \cap \gamma b_c| \le 1$), then we define $R''_c = S \cap \gamma b_c$. If the Case (ii) (Case (iii)) holds then, by Lemma 4.4(ii) (Lemma 4.6), it follows that we can find a set R''_c consisting of $O(1/(\epsilon\gamma)^{(d-1)/2})$ points such that $NN_q(R''_c) \le (1+\epsilon) \cdot NN_q(S \cap b'_c) \le (1+\epsilon) \cdot NN_q(S \cap \gamma b_c)$. Finally we assign $R_c = R'_c \cup R''_c$ to be set of representatives for c. Clearly, R_c has size $O(1/(\epsilon\gamma)^{(d-1)/2})$ and satisfies the desired property, namely, $NN_q(R_c) \le (1+\epsilon) \cdot NN_q(S)$.

Given a query point q, we can determine the leaf containing q in $O(\log(n\gamma))$ time. By computing the distance from q for each of the stored representatives, we can answer queries in $O(\log(n\gamma) + 1/(\epsilon\gamma)^{(d-1)/2})$ time. We summarize the main result of this section.

THEOREM 4.1. *Let S be a set of n points in $\mathbb{R}^d$, and let $0 < \epsilon \le 1/2$ and $2 \le \gamma \le 1/\epsilon$ be two real parameters. Then we can construct an $(O(1/(\epsilon\gamma)^{(d-1)/2}), \epsilon)$-approximate Voronoi diagram for S that consists of $O(n\gamma^d)$ regions, where each region is the difference of two cubes. Moreover, for any query point, we can return its ϵ-NN in $O(\log(n\gamma) + 1/(\epsilon\gamma)^{(d-1)/2})$ time. Here the constants in the O-notation are independent of ϵ and γ.*

Based on this theorem we obtain a family of data structures that can answer ϵ-NN queries in $O(\log(n\gamma) + 1/(\epsilon\gamma)^{(d-1)/2})$ time using space $O(n(\gamma/\epsilon)^{(d-1)/2}\gamma)$. Setting γ to two we obtain the most space-efficient solution in this family, which we present in the following corollary.

COROLLARY 4.1. *Given a set S of n points in $\mathbb{R}^d$, we can answer ϵ-NN queries in $O(\log n + 1/\epsilon^{(d-1)/2})$ time* *using a data structure of size $O((1/\epsilon)^{(d-1)/2}n)$.*

Remark: Using ideas similar to those given in this subsection, we can reduce the size of the $(1,\epsilon)$-AVD in Theorem 3.1 to $O(n/\epsilon^d)$. We leave the details to the interested reader.

Remark: Instead of building the AVD by using the well-separated pair decomposition, an alternative method is to first build a *smoothed* box decomposition (BD) tree [10] and then suitably refine it. But this solution appears to be more complicated than the one we have presented here.

5 Lower Bound on Size of $(1,\epsilon)$-Approximate Voronoi Diagram

Throughout this section we will assume that ϵ is a sufficiently small constant (depending on dimension d). The following lemma is needed for the lower bound argument.

LEMMA 5.1. *Let $\epsilon \leq 1$ and let p,q be any two points in $\mathbb{R}^d$ such that $|pq| = 1$. Let h be the hyperplane orthogonal to $\overline{pq}$ and passing through its midpoint. Let x be a point at distance at most one from $\overline{pq}$, at distance 2ϵ from h, and on the same side of h as p. Then $|xq| > (1+\epsilon)|xp|$.*

Proof Let m be the midpoint of $\overline{pq}$ and x' be the orthogonal projection of x onto $\overline{pq}$. We have

$$\frac{|xq|^2 - |xp|^2}{|xp|^2} = \frac{(\overrightarrow{xq} + \overrightarrow{xp}) \cdot (\overrightarrow{xq} - \overrightarrow{xp})}{|xx'|^2 + |x'p|^2} \geq \frac{2\,\overrightarrow{xm} \cdot \overrightarrow{pq}}{|xx'|^2 + |mp|^2} \geq \frac{2\,|x'm|}{1+\frac{1}{4}} > 3\epsilon.$$

Adding one to both sides, we get $|xq|^2/|xp|^2 > 1 + 3\epsilon$. Since $\epsilon \leq 1$, it follows that $|xq| > (1+\epsilon)|xp|$. □

For the lower bound, let S consist of n points in the form of pairs $(p_i, q_i), 1 \leq i \leq n/2$, such that for each pair, the vector $\overrightarrow{p_i q_i}$ has all coordinates equal to $1/\sqrt{d}$. Further, for any two points in different pairs, the distance between them is at least 8. These conditions are obviously easy to ensure. Let $\mathcal{C}$ be the set of cells in a $(1,\epsilon)$-AVD for S. We assume that each cell of $\mathcal{C}$ is an axis-aligned hyperrectangle or the difference of two axis-aligned hyperrectangles. Our goal is to show that $|\mathcal{C}| = \Omega(n/\epsilon^{d-1})$. Observe that if a cell is the difference of two axis-aligned hyperrectangles, then we can partition it into a constant number of axis-aligned hyperrectangles. Thus, it suffices to prove that $|\mathcal{C}| = \Omega(n/\epsilon^{d-1})$, when the cells of $\mathcal{C}$ are axis-aligned hyperrectangles.

For a pair of points (p_i, q_i), let h_i be the hyperplane orthogonal to $\overline{p_i q_i}$ and passing through its midpoint, and let Γ_i be the cylinder consisting of points at distance at most one from $\overline{p_i q_i}$ and at distance at most 2ϵ from h_i. (See Fig. 2.) Since the distance between any two points belonging to different pairs is at least 8, it is easy to verify that only p_i or q_i can be the representative for a cell that intersects Γ_i (assuming $\epsilon < 1$). It follows that no cell can intersect both Γ_i and Γ_j, for $i \neq j$. Thus, it suffices to show that the number of cells of $\mathcal{C}$ that intersect Γ_i is $\Omega(1/\epsilon^{d-1})$.

To this end, define a cylinder Γ'_i to be the set of points at distance at most $(1 - 5\epsilon\sqrt{d})$ from $\overline{p_i q_i}$ and at distance at most 2ϵ from h_i. (Note that $\Gamma'_i \subseteq \Gamma_i$.) For sufficiently small ϵ, the volume of Γ'_i is clearly $\Omega(\epsilon)$. Let $c \in \mathcal{C}$ be any cell that intersects Γ'_i. We will show that the volume of the region $c \cap \Gamma'_i$ is at most $O(\epsilon^d)$, which will imply that the number of cells of $\mathcal{C}$ that intersect Γ'_i, and hence Γ_i, is $\Omega(1/\epsilon^{d-1})$.

Let h_i^1 and h_i^2 denote the hyperplanes obtained by translating h_i by distance 2ϵ towards p_i and q_i, respectively. Let f_i^1 and f_i^2 denote the faces of cylinder Γ_i contained in h_i^1 and h_i^2, respectively. By Lemma 5.1, it follows that for any point x on f_i^1, we have $|xq_i| > (1+\epsilon)|xp_i|$. By symmetry, any point y on f_i^2 satisfies $|yp_i| > (1+\epsilon)|yq_i|$. Thus, c cannot intersect both f_i^1 and f_i^2. We will need this fact later in the proof.

Let v^+ and v^- denote the vertices of c that have, among the vertices of c, the highest coordinate along all dimensions, and the lowest coordinate along all dimensions, respectively. Clearly v^+ lies on the same side of h_i^1 as q_i and v^- lies on the same side of h_i^2 as p_i. We claim that either v^+ or v^- must lie inside Γ_i. For the sake of contradiction, suppose that both v^+ and v^- lie outside Γ_i. Let y be any point in $c \cap \Gamma'_i$ and let z denote the point of intersection of $\overline{yv^+}$ with the boundary of Γ_i. Clearly z must lie either on f_i^2 or on the cylinderical face of Γ_i. Let θ be the angle that $\overrightarrow{yv^+}$ makes with the axis $\overrightarrow{p_i q_i}$ of cylinder Γ_i. Then it is easy to see that $|yz|$ is at most $4\epsilon/\cos\theta$ (recall that the height of cylinder Γ_i is 4ϵ). Further, by convexity of c, z lies inside c, and so $\cos\theta \geq 1/\sqrt{d}$. Thus $|yz| \leq 4\epsilon\sqrt{d}$. Since the minimum distance between the cylinderical faces of Γ'_i and Γ_i is $5\epsilon\sqrt{d}$, it follows that z must lie on f_i^2. Thus c intersects f_i^2. By a symmetric argument, we can show that c intersects f_i^1 as well. But this contradicts the fact shown earlier that c cannot intersect both f_i^1 and f_i^2.

We have thus proved that either $v^+ \in \Gamma_i$ or $v^- \in \Gamma_i$. Without loss of generality, suppose that $v^- \in \Gamma_i$. Clearly the volume of $c \cap \Gamma'_i$ is no more than the volume of $c \cap R$, where R denotes the infinite region contained between the hyperplanes h_i^1 and h_i^2. This volume is

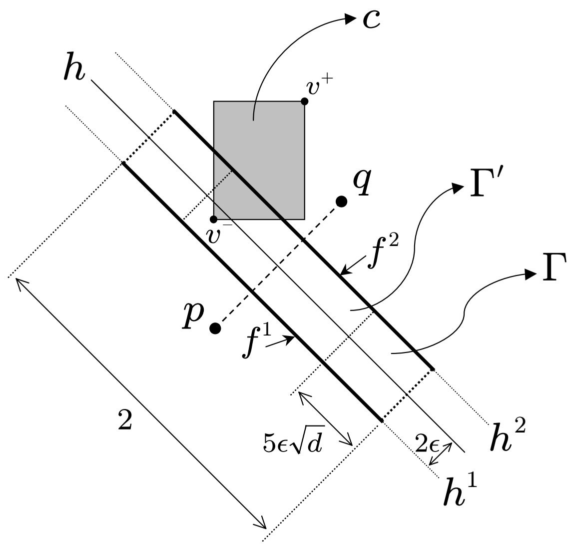

Fig. 2: Lower bound construction.

maximized in the limit as v^- touches h_i^1 and each edge of c incident to v^- extends upto h_i^2. This quantity is easily seen to be at most $(4\epsilon\sqrt{d})^d/2 = O(\epsilon^d)$, which completes the proof. We have established the following lower bound.

THEOREM 5.1. *Assuming that the cells of a $(1, \epsilon)$-approximate Voronoi diagram are axis-aligned hyperrectangles or differences of two axis-aligned hyperrectangles, its size in the worst case is $\Omega(n/\epsilon^{d-1})$.*

6 Conclusions

Along with David Mount, we have recently improved several of the results given in this paper [1]. In particular, we show that we can construct a (t, ϵ)-AVD consisting of $O(n\epsilon^{(d-1)/2}\gamma^{(3d-1)/2})$ cells for $t = O(1/(\epsilon\gamma)^{(d-1)/2})$. This yields a data structure of $O(n\gamma^d)$ space (including the space for representatives) that can answer ϵ-NN queries in time $O(\log(n\gamma) + 1/(\epsilon\gamma)^{(d-1)/2})$.

References

[1] S. Arya, T. Malamatos, and D. M. Mount. Manuscript, in preparation.

[2] S. Arya, D. M. Mount, N. Netanyahu, R. Silverman, and A. Y. Wu. An optimal algorithm for approximate nearest neighbor searching in fixed dimensions. *J. ACM*, 45:891–923, 1998.

[3] P. B. Callahan and S. R. Kosaraju. A decomposition of multidimensional point sets with applications to k-nearest-neighbors and n-body potential fields. *J. ACM*, 42:67–90, 1995.

[4] T. M. Chan. Approximate nearest neighbor queries revisited. *Discrete Comput. Geom.*, 20:359–373, 1998.

[5] T. M. Chan and J. Snoeyink. Algorithms for approximate nearest-neighbor queries. Manuscript, 1995.

[6] C. A. Duncan, M. T. Goodrich, and S. G. Kobourov. Balanced aspect ratio trees: Combining the advantages of k-d trees and octrees. In *Proc. 10th ACM-SIAM Sympos. Discrete Algorithms*, 1999.

[7] S. Har-Peled. A replacement for Voronoi diagrams of near linear size. In *Proc. 42 Annu. IEEE Sympos. Found. Comput. Sci.*, 2001.

[8] P. Indyk and R. Motwani. Approximate nearest neighbors: Towards removing the curse of dimensionality. In *Proc. 30th Annu. ACM Sympos. Theory Comput.*, pages 604–613, 1998.

[9] E. Kushilevitz, R. Ostrovsky, and Y. Rabani. Efficient search for approximate nearest neighbor in high dimemsional spaces. In *Proc. 30th Annu. ACM Sympos. Theory Comput.*, 1998.

[10] J. S. B. Mitchell, D. M. Mount, and S. Suri. Query-sensitive ray shooting. *Internat. J. Comput. Geom. Appl.*, 7(4):317–347, August 1997.

[11] Y. Sabharwal, S. Sen, and N. Sharma. Improved space bound for approximate Voronoi diagram. Manuscript, 2001.

[12] J. Vleugels and M. Overmars. Approximating Voronoi diagrams of convex sites in any dimension. *Internat. J. Comput. Geom. Appl.*, 8:201–222, 1998.

[13] A. C. Yao. On constructing minimum spanning trees in k-dimensional spaces and related problems. *SIAM J. Comput.*, 11(4):721–736, 1982.