

# Approximating Extent Measures of Points\*

Pankaj K. Agarwal<sup>†</sup>

Sariel Har-Peled<sup>‡</sup>

Kasturi R. Varadarajan<sup>§</sup>

November 5, 2003

## Abstract

We present a general technique for approximating various descriptors of the extent of a set  $P$  of  $n$  points in  $\mathbb{R}^d$  when the dimension  $d$  is an arbitrary fixed constant. For a given extent measure  $\mu$  and a parameter  $\varepsilon > 0$ , it computes in time  $O(n + 1/\varepsilon^{O(1)})$  a subset  $Q \subseteq P$  of size  $1/\varepsilon^{O(1)}$ , with the property that  $(1 - \varepsilon)\mu(P) \leq \mu(Q) \leq \mu(P)$ . The specific applications of our technique include  $\varepsilon$ -approximation algorithms for (i) computing diameter, width, and smallest bounding box, ball, and cylinder of  $P$ , (ii) maintaining all the previous measures for a set of moving points, and (iii) fitting spheres and cylinders through a point set  $P$ . Our algorithms are considerably simpler, and faster in many cases, than previously known algorithms.

## 1 Introduction

Motivated by a variety of applications, considerable work has been done on measuring various descriptors of the extent of a set  $P$  of  $n$  points in  $\mathbb{R}^d$ . We refer to such measures as *extent measures* of  $P$ . Roughly speaking, an extent measure of  $P$  either computes certain statistics of  $P$  itself or it computes certain statistics of a (possibly nonconvex) geometric shape (e.g. sphere, box, cylinder, etc.) enclosing  $P$ . Examples of the former include computing the  $k$ th largest distance between pairs of points in  $P$ , and the examples of the latter include computing the smallest radius of a sphere (or cylinder), the minimum volume (or surface area) of a box, and the smallest width of a slab (or a

---

\*Research by the first author is supported by NSF under grants CCR-00-86013, EIA-98-70724, EIA-01-31905, and CCR-02-04118, and by a grant from the U.S.–Israel Binational Science Foundation. Research by the second author is supported by NSF CAREER award CCR-0132901. Research by the third author is supported by NSF CAREER award CCR-0237431

A preliminary version of the paper appeared as: (i) P. K. Agarwal and S. Har-Peled, Maintaining the approximate extent measures of moving points, *Proc. 12th ACM-SIAM Sympos. Discrete Algorithms*, pages 148–157, 2001, and (ii) S. Har-Peled and K. R. Varadarajan, Approximate shape fitting via linearization, *Proc. 42nd Annu. IEEE Sympos. Found. Comput. Sci.*, pages 66–73, 2001.

<sup>†</sup>Department of Computer Science, Box 90129, Duke University, Durham NC 27708-0129; pankaj@cs.duke.edu; <http://www.cs.duke.edu/~pankaj/>

<sup>‡</sup>Department of Computer Science, DCL 2111; University of Illinois; 1304 West Springfield Ave., Urbana, IL 61801; sariel@uiuc.edu; <http://www.uiuc.edu/~sariel/>

<sup>§</sup>Department of Computer Science, The University of Iowa, Iowa City, IA 52242-1419; kvaradar@cs.uiowa.edu; <http://www.cs.uiowa.edu/~kvaradar/>

spherical or cylindrical shell) that contain  $P$ . Although  $P$  is assumed to be stationary in most of the work done so far, there has been some recent work on maintaining extent measures of a set of moving points [4].

Shape fitting, a fundamental problem in computational geometry, computer vision, machine learning, data mining, and many other areas, is closely related to computing extent measures. A widely used shape-fitting problem asks for finding a shape that best fits  $P$  under some “fitting” criterion. A typical criterion for measuring how well a shape  $\gamma$  fits  $P$ , denoted as  $\mu(P, \gamma)$ , is the maximum distance between a point of  $P$  and its nearest point on  $\gamma$ , i.e.,  $\mu(P, \gamma) = \max_{p \in P} \min_{q \in \gamma} d(p, q)$ . Then one can define the extent measure of  $P$  to be  $\mu(P) = \min_{\gamma} \mu(P, \gamma)$ , where the minimum is taken over a family of shapes (such as points, lines, hyperplanes, spheres, etc.). For example, the problem of finding the minimum radius sphere (resp. cylinder) enclosing  $P$  is the same as finding the point (resp. line) that fits  $P$  best, and the problem of finding the smallest width slab (resp. spherical shell, cylindrical shell)<sup>1</sup> is the same as finding the hyperplane (resp. sphere, cylinder) that fits  $P$  best.

The exact algorithm for computing extent measures are generally expensive, e.g., the best known algorithms for computing the smallest volume bounding box containing  $P$  in  $\mathbb{R}^3$  require  $O(n^3)$  time. Consequently, attention has shifted to developing approximation algorithms [10, 34]. Despite considerable work, no unified theory has evolved for computing extent measures approximately. Ideally, one would like to argue that for any extent measure  $\mu$  and for any given parameter  $\varepsilon$ , there exists a subset  $Q \subset P$  of size  $1/\varepsilon^{O(1)}$  so that  $\mu(Q) \geq (1 - \varepsilon)\mu(P)$ . No such result is known except in a few special cases. It is known that an arbitrary convex body  $C$  can be approximated by a convex polytope  $Q$  so that the Hausdorff distance between  $C$  and  $Q$  is at most  $\varepsilon \cdot \text{diam}(C)$  and so that  $Q$  is either defined as the convex hull of a set of  $1/\varepsilon^{O(1)}$  points or the intersection of a set of  $1/\varepsilon^{O(1)}$  halfspaces. If the given extent measure  $\mu$  of  $P$  is the same as that of  $\mathcal{CH}(P)$ , (e.g., diameter and width), then one can approximate  $\mathcal{CH}(P)$  by  $Q$ , compute  $\mu(Q)$ , and argue that  $\mu(Q)$  approximates  $\mu(P)$ . Although this approach has been used for computing a few extent measures of  $P$  [10, 16], it does not work if the extent measure  $\mu$  is defined in terms of a nonconvex shape (such as spherical shell) containing  $P$ .

This paper is a step toward the aforementioned goal of developing a unified theory for approximating extent measures. We introduce the notion of an  $\varepsilon$ -kernel of a point set  $P$ . Roughly speaking, a subset  $Q \subseteq P$  is called an  $\varepsilon$ -kernel of  $P$  if for every slab  $W$  containing  $Q$ , the expanded slab  $(1 + \varepsilon)W$  contains  $P$ . We present an  $O(n + 1/\varepsilon^{d-1})$ -time algorithm for computing an  $\varepsilon$ -kernel of  $P$  of size  $O(1/\varepsilon^{d-1})$  or an  $O(n + 1/\varepsilon^{3(d-1)/2})$ -time algorithm for computing an  $\varepsilon$ -kernel of size  $O(1/\varepsilon^{(d-1)/2})$ . These algorithms are improved variants of the algorithm described in [10] for a specific optimization problem. We call an extent measure  $\mu$  *faithful* if there exists a constant  $\alpha > 0$  such that for any  $\varepsilon$ -kernel  $Q$  of  $P$ ,  $\mu(Q) \geq (1 - \alpha\varepsilon)\mu(P)$ . The algorithm for computing an  $\varepsilon$ -kernel immediately gives an  $O(n + 1/\varepsilon^{O(1)})$  time algorithm for computing faithful measures approximately. This approach was used previously for some faithful measures [10, 16, 34] and we merely state it here in a general context. In order to handle unfaithful measures, we introduce the notion of an  $\varepsilon$ -kernel for a family of functions. Let  $\mathcal{F}$  be a family of  $(d - 1)$ -variate functions. We

---

<sup>1</sup>A *slab* is a region lying between two parallel hyperplanes; a *spherical shell* is the region lying between two concentric spheres; a *cylindrical shell* is the region lying between two coaxial cylinders.

define the *extent* of  $\mathcal{F}$  at a point  $x \in \mathbb{R}^{d-1}$  to be  $\mathfrak{E}_{\mathcal{F}}(x) = \max_{f \in \mathcal{F}} f(x) - \min_{f \in \mathcal{F}} f(x)$ . We call a subset  $\mathcal{G} \subseteq \mathcal{F}$  an  $\varepsilon$ -kernel of  $\mathcal{F}$  if  $\mathfrak{E}_{\mathcal{G}}(x) \geq (1 - \varepsilon)\mathfrak{E}_{\mathcal{F}}(x)$  for all  $x \in \mathbb{R}^{d-1}$ . Using our result on  $\varepsilon$ -kernel of points and the linearization technique, we show that we can compute in  $O(n + 1/\varepsilon^{O(1)})$  time an  $\varepsilon$ -kernel of  $\mathcal{F}$  of size  $O(1/\varepsilon^{r\sigma})$  if each  $f_i$  is of the form  $g_i^{1/r}$ , where  $g_i$  is a polynomial,  $r$  is a positive integer,  $\sigma = \min\{d - 1, k/2\}$ , and  $k$  is the dimension of linearization for  $g_i$ 's (see Section 4 for the definition of  $k$ ). Our algorithms for computing  $\varepsilon$ -kernels can be adapted to handle insertions and deletions of points (or functions) efficiently, see Section 5. If we only insert points, we can maintain an  $\varepsilon$ -kernel using only  $(\log(n)/\varepsilon)^{O(1)}$  space.

We show that many extent-measure problems can be formulated as computing  $\min_x \mathfrak{E}_{\mathcal{F}}(x)$ , where  $\mathcal{F}$  is obtained by transforming each input point to a function. Specific applications of our technique include the following:

**Spherical shell problem.** Given a point  $x$  in  $\mathbb{R}^d$  and two real numbers  $0 \leq r \leq R$ , the *spherical shell*  $\sigma(x, r, R)$  is the closed region lying between the two concentric spheres of radii  $r$  and  $R$  with  $x$  as their center, i.e.,

$$\sigma(x, r, R) = \left\{ p \in \mathbb{R}^d \mid r \leq d(x, p) \leq R \right\},$$

where  $d(x, p)$  is the Euclidean distance between the points  $p$  and line  $x$ . The *width* of  $\sigma(x, r, R)$  is  $R - r$ . In the *approximate spherical shell* problem, we are given a set  $P$  of  $n$  points and a parameter  $\varepsilon > 0$ , and we want to compute a spherical shell containing  $P$  whose width is at most  $(1 + \varepsilon)$  times the width of the minimum-width spherical shell containing  $P$ .

This problem, motivated by applications in computational metrology, has been widely studied; see [2, 8, 16] and the references therein. The best known exact algorithm runs in  $O(n^{3/2+\delta})$  time in  $\mathbb{R}^2$ , for any  $\delta > 0$ , and in  $O(n^2)$  time in  $\mathbb{R}^3$ . The best known  $\varepsilon$ -approximation algorithm, proposed by Chan [16], takes about  $O(n + 1/\varepsilon^{d^2/4})$  time. Our technique leads to an  $O(n + 1/\varepsilon^{3d})$ -time algorithm for the  $d$ -dimensional approximate spherical-shell problem, thereby improving Chan's algorithm.

**Cylindrical shell problem.** Given a line  $\ell$  in  $\mathbb{R}^d$  and two real numbers  $0 \leq r \leq R$ , the *cylindrical shell*  $\Sigma(\ell, r, R)$  is the closed region lying between two co-axial cylinders of radii  $r$  and  $R$  with  $\ell$  as their axis, i.e.,

$$\Sigma(\ell, r, R) = \left\{ p \in \mathbb{R}^d \mid r \leq d(\ell, p) \leq R \right\},$$

where  $d(\ell, p)$  is the Euclidean distance between the point  $p$  and line  $\ell$ . The *width* of  $\Sigma(\ell, r, R)$  is  $R - r$ .

In the *approximate cylindrical shell* problem, we are given a set  $P$  of  $n$  points and a parameter  $\varepsilon > 0$ , and we want to compute a cylindrical shell containing  $P$  whose width is at most  $(1 + \varepsilon)$  times the width of the minimum-width cylindrical shell containing  $P$ .

Agarwal *et al.* [3] present an algorithm that computes the exact minimum-width cylindrical shell for a set of  $n$  points in  $\mathbb{R}^3$  in  $O(n^5)$  time. They also present an algorithm that runs in roughly  $O(n^2)$  time and computes a shell whose width is at most 26 times the optimal. For this problem, our technique gives an  $\varepsilon$ -approximation algorithm that runs in  $O(n + 1/\varepsilon^{O(d^2)})$  time in  $\mathbb{R}^d$ , a significant improvement over their algorithm.

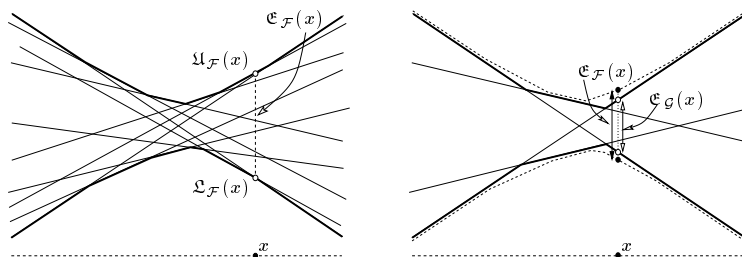
**Maintaining faithful measures of moving points.** Let  $P$  be a set of  $n$  points in  $\mathbb{R}^d$ , each point moving independently. Many applications call for maintaining extent measures of  $P$  as the points move with time. For example, various indexing structures, which answer range-searching queries or nearest-neighbor queries on  $P$ , need an algorithm for maintaining the smallest axes-parallel box containing  $P$  [1, 31, 32]. Agarwal *et al.* [4] have described kinetic data structures for maintaining a number of extent measures of points moving in the plane. They also show that most of these extent measures are expensive to maintain — the diametral pair of a set of points, each moving with a fixed velocity in the plane, can change  $\Omega(n^2)$  times, and no subcubic bound is known on the number of triples defining the smallest enclosing ball of a set of points moving in the plane. This has raised the question whether faster approximation algorithms exist for maintaining an extent-measure of a set of moving points.

For any  $\varepsilon > 0$ , we say that a subset  $Q \subseteq P$  is an  $\varepsilon$ -kernel of  $P$  if  $Q(t)$  is an  $\varepsilon$ -kernel of  $P(t)$  for every  $t$ . We show that our techniques can compute an  $\varepsilon$ -kernel of size  $1/\varepsilon^{O(1)}$ . For instance, given any set  $P$  of points in  $\mathbb{R}^d$  with linear motion, our technique can compute, in  $O(n + 1/\varepsilon^{2d})$  time, an  $\varepsilon$ -kernel  $Q \subseteq P$  of size  $O(1/\varepsilon^{2d})$ . It follows that for any faithful measure  $\mu$ , there exists a constant  $\alpha > 0$  such that  $(1 - \alpha\varepsilon)\mu(P(t)) \leq \mu(Q(t))$  for every  $t$ . We can simply maintain  $\mu$  for just the subset  $Q$ . We can thus efficiently maintain an  $\varepsilon$ -approximation to all of the following measures of  $P$ : diameter, minimum-radius enclosing ball, width, minimum-volume bounding box of arbitrary orientation, directional width. If we want to maintain an  $\varepsilon$ -approximation to the smallest axes-parallel box enclosing  $P$ , the size of  $Q$  can be reduced to  $O(1/\sqrt{\varepsilon})$ , for any fixed dimension. These results generalize to algebraic motion and to unfaithful measures such as minimum-width spherical/cylindrical shell. Our scheme can also allow efficient insertions into and deletions from the set  $P$ . Note that the  $\varepsilon$ -kernel does not change with time unless the trajectory of a point changes. These results must be contrasted with the schemes for maintaining the exact extent measures, which require at least quadratic updates.

**Maintaining faithful measures in a streaming model.** Motivated by various applications, the need for analyzing and processing massive data in real time has led to a flurry of activity related to performing computations on a *data stream*. The goal is to maintain a summary of the input data using little space and processing time, as the data objects arrives. The efficiency of an algorithm in this model is measured in terms of the size of the working space and the time spent on performing the computation on a new data object. See [18, 23, 24, 26, 29] and references therein for recent algorithms developed in the data-stream model. Our technique can be adapted to maintain various extent measures approximately in the streaming model. Specifically, an  $\varepsilon$ -kernel of size  $O(1/\varepsilon^{(d-1)/2})$  of a stream of points in  $\mathbb{R}^d$  can be maintained using a data structure of size  $O(\log^d(n)/\varepsilon^{(d-1)/2})$  that spends  $O(1/\varepsilon^{d-1})$  amortized time to process each new point. The same result holds for  $\varepsilon$ -kernels of linear functions. Consequently, we can maintain an  $\varepsilon$ -approximation of the width of a stream of points in  $\mathbb{R}^d$  in amortized  $O(1/\varepsilon^{d-1})$  time using  $O(\log^d(n)/\varepsilon^{(d-1)/2})$  space. Similar results can be obtained for various other problems such as maintaining the minimum-width spherical or cylindrical shell containing the point set.

The paper is organized as follows. In Section 2, we formally define  $\varepsilon$ -kernels for points and functions and make a few simple observations about them. In Section 3 we show that any set of

linear functions has an  $\varepsilon$ -kernel of small size. Section 4 shows that this property is also true for polynomials and related functions, using linearization. Section 5 shows that our technique can be dynamized. In Section 6, we apply these ideas to the problems mentioned above.



**Figure 1.** (i) Lower and upper envelopes and the extent of a family of linear functions; the extent at any point is the length of the vertical segment connecting lower and upper envelopes. (ii) An  $\varepsilon$ -kernel  $\mathcal{G}$  of  $\mathcal{F}$ ; dashed edges denote the envelopes of  $\mathcal{F}$ , and the thick lines denote the envelopes of  $\mathcal{G}$ .

## 2 Preliminaries

In this section we define the extent of functions, the directional-width of points, the  $\varepsilon$ -kernel of points and functions, and arrangements. We also establish some simple claims that will be useful later. Table 1 (cf. page 32) summarizes the notation used in this paper.

**Envelopes and extent.** Let  $\mathcal{F} = \{f_1, \dots, f_n\}$  be a set of  $n$   $(d-1)$ -variate functions defined over  $x = (x_1, \dots, x_{d-1}) \in \mathbb{R}^{d-1}$ . The *lower envelope* of  $\mathcal{F}$  is the graph of the function  $\mathfrak{L}_{\mathcal{F}} : \mathbb{R}^{d-1} \rightarrow \mathbb{R}$  defined as  $\mathfrak{L}_{\mathcal{F}}(x) = \min_{f \in \mathcal{F}} f(x)$ . Similarly, the *upper envelope* of  $\mathcal{F}$  is the graph of the function  $\mathfrak{U}_{\mathcal{F}} : \mathbb{R}^{d-1} \rightarrow \mathbb{R}$  defined as  $\mathfrak{U}_{\mathcal{F}}(x) = \max_{f \in \mathcal{F}} f(x)$ . The *extent*  $\mathfrak{E}_{\mathcal{F}} : \mathbb{R}^{d-1} \rightarrow \mathbb{R}$  of  $\mathcal{F}$  is defined as

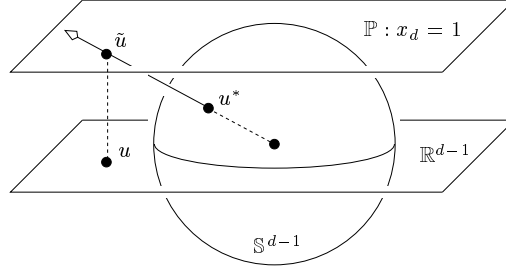
$$\mathfrak{E}_{\mathcal{F}}(x) = \mathfrak{U}_{\mathcal{F}}(x) - \mathfrak{L}_{\mathcal{F}}(x).$$

Let  $\varepsilon > 0$  be a parameter, and let  $\Delta$  be a subset of  $\mathbb{R}^{d-1}$ . We say that a subset  $\mathcal{G} \subseteq \mathcal{F}$  is an  $\varepsilon$ -kernel of  $\mathcal{F}$  within  $\Delta$  if

$$(1 - \varepsilon)\mathfrak{E}_{\mathcal{F}}(x) \leq \mathfrak{E}_{\mathcal{G}}(x)$$

for each  $x \in \Delta$ . Obviously,  $\mathfrak{E}_{\mathcal{G}}(x) \leq \mathfrak{E}_{\mathcal{F}}(x)$ , as  $\mathcal{G} \subseteq \mathcal{F}$ . If  $\Delta = \mathbb{R}^{d-1}$ , we say that  $\mathcal{G}$  is an  $\varepsilon$ -kernel of  $\mathcal{F}$ .

**Lemma 2.1** *Let  $\mathcal{F} = \{f_1, \dots, f_n\}$  be a family of  $(d-1)$ -variate functions,  $\varphi(x), \psi(x)$  two other  $(d-1)$ -variate functions, and  $\varepsilon > 0$  a parameter. Let  $\hat{f}_i(x) = \varphi(x) + \psi(x)f_i(x)$ , and set  $\hat{\mathcal{F}} = \{\hat{f}_i \mid 1 \leq i \leq n\}$ . If  $\mathcal{K}$  is an  $\varepsilon$ -kernel of  $\mathcal{F}$  within a region  $\Delta \subseteq \mathbb{R}^{d-1}$ , then  $\hat{\mathcal{K}} = \{\hat{f}_i \mid f_i \in \mathcal{K}\}$  is an  $\varepsilon$ -kernel of  $\hat{\mathcal{F}}$  within  $\Delta$ .*



**Figure 2.** Representing a direction in  $\mathbb{R}^d$ .

*Proof:* For any  $x \in \Delta$ ,

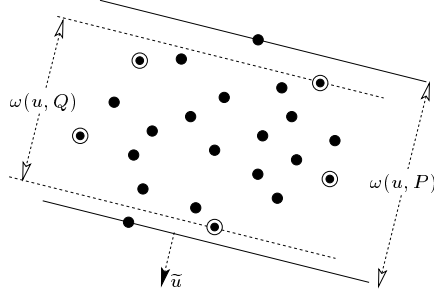
$$\begin{aligned}
(1 - \varepsilon)\mathfrak{E}_{\widehat{\mathcal{F}}}(x) &= (1 - \varepsilon) \left[ \max_{f_i \in \widehat{\mathcal{F}}} \widehat{f}_i(x) - \min_{f_i \in \widehat{\mathcal{F}}} \widehat{f}_i(x) \right] \\
&= (1 - \varepsilon) \left[ \max_{f_i \in \mathcal{F}} (\varphi(x) + \psi(x)f_i(x)) - \min_{f_i \in \mathcal{F}} (\varphi(x) + \psi(x)f_i(x)) \right] \\
&= (1 - \varepsilon)\psi(x) \left[ \max_{f_i \in \mathcal{F}} f_i(x) - \min_{f_i \in \mathcal{F}} f_i(x) \right] \\
&\leq \psi(x) \left[ \max_{f_i \in \mathcal{K}} f_i(x) - \min_{f_i \in \mathcal{K}} f_i(x) \right] \\
&= \max_{f_i \in \mathcal{K}} (\varphi(x) + \psi(x)f_i(x)) - \min_{f_i \in \mathcal{K}} (\varphi(x) + \psi(x)f_i(x)) \\
&= \mathfrak{E}_{\widehat{\mathcal{K}}}(x).
\end{aligned}$$

Hence  $\widehat{\mathcal{K}}$  is an  $\varepsilon$ -kernel of  $\widehat{\mathcal{F}}$ . ■

**Directions.** We will not distinguish between a vector in  $\mathbb{R}^d$  and the corresponding point in  $\mathbb{R}^d$ . Let  $\mathbb{P}$  denote the hyperplane  $x_d = 1$  in  $\mathbb{R}^d$ , and let  $\mathbb{S}^{d-1}$  represent the sphere of directions in  $\mathbb{R}^d$ . For an arbitrary vector  $v \in \mathbb{R}^d$ , we will use  $\phi(v) = v/\|v\| \in \mathbb{S}^{d-1}$  to denote the direction corresponding to  $v$ . Normally, a direction in  $\mathbb{R}^d$  is represented as a point in  $\mathbb{S}^{d-1}$ . Since we will not need to distinguish between directions  $x \in \mathbb{S}^{d-1}$  and  $-x \in \mathbb{S}^{d-1}$ , we can represent a direction  $u^* \in \mathbb{S}^{d-1}$  as a point  $u$  in  $\mathbb{R}^{d-1}$ , with the interpretation that  $\tilde{u} = (u, 1) \in \mathbb{P}$  is the central projection of the unit vector  $u^*$ ; see Figure 2. Namely,  $u^* = \phi(\tilde{u})$ . Though this representation has the drawback that the directions in  $\mathbb{S}^{d-1}$  lying in the hyperplane  $x_d = 0$  are not accounted for, we use this representation as it will be more convenient for our presentation.

**Directional width.** We can define the concept of extent for a set of points. For any non-zero vector  $x \in \mathbb{R}^d$  and a point set  $P \subseteq \mathbb{R}^d$ , we define

$$\bar{\omega}(x, P) = \max_{p \in P} \langle x, p \rangle - \min_{p \in P} \langle x, p \rangle,$$



**Figure 3.** A point set  $P$ , its  $\varepsilon$ -kernel  $Q$  (points with double circles), and their directional widths (scaled appropriately).

where  $\langle \cdot, \cdot \rangle$  is the inner product. For any set  $P$  of points in  $\mathbb{R}^d$  and any  $u \in \mathbb{R}^{d-1}$ , we define the *directional width* of  $P$  in direction  $u$ , denoted by  $\omega(u, P)$ , to be

$$\omega(u, P) = \bar{\omega}(\tilde{u}, P).$$

It is also called the *u-breadth* of  $P$ , see [22]. Let  $\varepsilon > 0$  be a parameter, and let  $\Delta \subseteq \mathbb{R}^{d-1}$ . A subset  $Q \subseteq P$  is called an  $\varepsilon$ -kernel of  $P$  within  $\Delta \subseteq \mathbb{R}^{d-1}$  if for each  $u \in \Delta$ ,

$$(1 - \varepsilon)\omega(u, P) \leq \omega(u, Q).$$

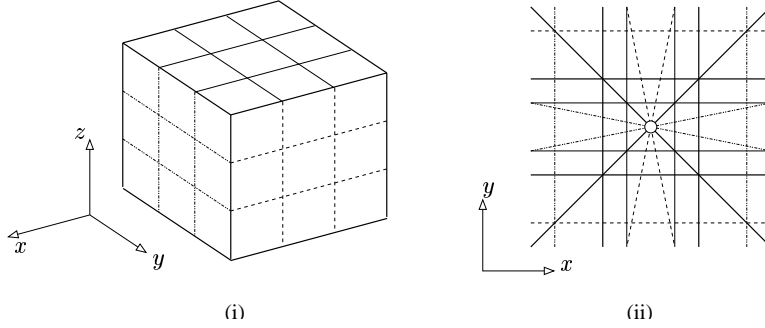
Clearly,  $\omega(u, Q) \leq \omega(u, P)$ . If  $\Delta = \mathbb{R}^{d-1}$ , we call  $Q$  an  $\varepsilon$ -kernel of  $P$ . Note that  $\omega(u, Q) \geq (1 - \varepsilon)\omega(u, P)$  if and only if for every  $0 \neq \lambda \in \mathbb{R}$ ,

$$\bar{\omega}(\lambda\tilde{u}, Q) \geq (1 - \varepsilon)\bar{\omega}(\lambda\tilde{u}, P). \quad (1)$$

**Arrangements.** The *arrangement* of a collection  $\mathcal{J}$  of  $m$  hyperplanes in  $\mathbb{R}^d$ , denoted as  $\mathcal{A}(\mathcal{J})$ , is the decomposition of the space into relatively open connected cells of dimensions  $0, \dots, d$  induced by  $\mathcal{J}$ , where each cell is a maximal connected set of points lying in the intersection of a fixed subset of  $\mathcal{J}$ . The complexity of  $\mathcal{A}(\mathcal{J})$  is defined to be the number of cells of all dimensions in the arrangement. It is well known that the complexity of  $\mathcal{A}(\mathcal{J})$  is  $O(m^d)$  [9]. A set  $\mathcal{J}$  of hyperplanes is *k-uniform* if  $\mathcal{J}$  consists of  $k$  families of either parallel hyperplanes or hyperplanes that share a  $(d - 2)$ -flat. In this case, each cell of  $\mathcal{A}(\mathcal{J})$  has at most  $2k$  facets. The notion of arrangements can be extended to a family of (hyper-)surfaces in  $\mathbb{R}^d$ . If  $\mathcal{G}$  is a family of  $m$  algebraic surfaces of bounded maximum degree, then the complexity of the arrangement is  $O(m^d)$ .

**Lemma 2.2** *For any  $\varepsilon > 0$ , there is a  $d(d - 1)$ -uniform set  $\mathcal{J}$  consisting of  $O(1/\varepsilon)$  hyperplanes in  $\mathbb{R}^{d-1}$  so that for any two points  $u, v$  lying in the (closure of the) same cell of  $\mathcal{A}(\mathcal{J})$ ,*

$$\|u^* - v^*\| \leq \varepsilon.$$



**Figure 4.** (i) Grid drawn on each facet of the unit cube  $\mathbb{C}$  ( $d = 3$ ). (ii) Lines in  $\mathcal{J}$ ; thick lines correspond to the grid lines of the edges of  $\mathbb{C}$ , and solid (resp. dashed, dashed-dotted) lines correspond to the grid on the face normal to the  $z$ -axis (resp.  $y$ -axis,  $x$ -axis). The grid lines parallel to the  $z$ -axis on  $\mathbb{C}$  map to the lines passing through the origin, and the grid lines parallel to the  $x$ -axis (resp.  $y$ -axis) map to lines parallel to the  $x$ -axis (resp.  $y$ -axis).

*Proof:* Partition the boundary of the hypercube  $\mathbb{C} = [-1, +1]^d$  in  $\mathbb{R}^d$  into small  $(d - 1)$ -dimensional “hypercubes” of diameter at most  $\varepsilon$ , by laying a uniform  $(d - 1)$ -dimensional axis-parallel grid on each facet of  $\mathbb{C}$ ; see Figure 4 (i). Each such grid is formed by  $(d - 1)$  families of parallel  $(d - 2)$ -flats. We extend each such  $(d - 2)$ -flat  $f$  into a  $(d - 1)$ -hyperplane  $\hat{f}$ , by considering the unique hyperplane that passes through it and the origin, and then intersect it with  $\mathbb{P}$  (namely, the resulting  $(d - 2)$ -flat lies on the hyperplane  $\mathbb{P} : x_d = 1$  in  $\mathbb{R}^d$  and as such can be regarded as a hyperplane in  $\mathbb{R}^{d-1}$ ). Because of symmetry, the  $(d - 2)$ -flats  $x_i = -1, x_j = \delta$  (i.e., the intersection of hyperplanes  $x_i = -1$  and  $x_j = \delta$ ) and  $x_i = 1, x_j = -\delta$  map to the same hyperplane, so it suffices to extend the  $(d - 2)$ -flats of the grid on the “front” facets of  $\mathbb{C}$ , i.e., the facets with  $x_i = 1$  for  $1 \leq i \leq d$ . We claim that the resulting set composed of  $d(d - 1)$  families of uniform hyperplanes is the desired set of hyperplanes.

Formally, let  $F(i, j, \beta)$  denote the  $(d - 2)$ -flat  $x_i = 1, x_j = \beta$  in  $\mathbb{R}^d$ . Set  $\gamma = \lceil \sqrt{d}/\varepsilon \rceil$ , and for integers  $i, j, l$ , let

$$\mathcal{F} = \{F(i, j, l/\gamma) \mid 1 \leq i \neq j \leq d, -\gamma \leq l \leq \gamma\}.$$

For a  $(d - 2)$ -flat  $F \in \mathcal{F}$  not passing through origin, let  $\eta(F)$  be the  $(d - 2)$ -hyperplane in  $\mathbb{R}^{d-1}$  defined as

$$\eta(F) = \{x \in \mathbb{R}^{d-1} \mid (x, 1) \in \text{aff}(F \cup \{0\}) \cap \mathbb{P}\}.$$

In other words,  $\eta(F)$  is the  $(d - 2)$ -hyperplane in  $\mathbb{R}^{d-1}$  corresponding to the intersection of  $\mathbb{P}$  with the  $(d - 1)$ -hyperplane  $\text{aff}(F \cup \{0\})$ . We set  $\mathcal{J} = \{\eta(F) \mid F \in \mathcal{F}\}$ . See Figure 4(ii). Clearly  $\mathcal{J}$  is a  $d(d - 1)$ -uniform family of hyperplanes because for any fixed pair  $i, j$ , either all hyperplanes  $\eta(F(i, j, l/\gamma))$  in  $\mathcal{F}$  are parallel or all of them pass through a  $(d - 3)$ -flat.

Let  $u, v \in \mathbb{R}^{d-1}$  be any two points on the same cell  $f$  of  $\mathcal{A}(\mathcal{J})$ . Let  $u'$  (resp.  $v'$ ) be the point where the line joining the origin and  $\tilde{u}$  (resp.  $\tilde{v}$ ) intersects  $\partial\mathbb{C}$ . Since  $u', v' \in \partial\mathbb{C}$ , we have  $\|u'\|, \|v'\| \geq 1$ . Our construction ensures that  $\|u' - v'\| \leq \varepsilon$ . These two facts easily imply that  $\|u^* - v^*\| \leq \varepsilon$ . Hence,  $\mathcal{A}(\mathcal{J}_{d-1})$  is the required partition.  $\blacksquare$

**Remark 2.3** An interesting open question is to obtain a tight bound on the minimum number of



uniform families of hyperplanes needed to achieve the partition of Lemma 2.2. Agarwal and Matoušek [6] have shown that the number of families is at least  $2d - 3$ , and they conjecture this bound to be tight.

**Duality.** Let  $\mathcal{H} = \{h_1, \dots, h_n\}$  be a family of  $(d - 1)$ -variate linear functions and  $\varepsilon > 0$  a parameter. We define a *duality* transformation that maps the  $(d - 1)$ -variate function (or a hyperplane in  $\mathbb{R}^d$ )  $h : x_d = a_1x_1 + a_2x_2 + \dots + a_{d-1}x_{d-1} + a_d$  to the point  $h^* = (a_1, a_2, \dots, a_{d-1}, a_d)$  in  $\mathbb{R}^d$ . Let  $\mathcal{H}^* = \{h^* \mid h \in \mathcal{H}\}$ . The following lemma is immediate from the definition of duality.

**Lemma 2.4** *Let  $\mathcal{H} = \{h_1, \dots, h_n\}$  be a family of  $(d - 1)$ -variate linear functions and  $\varepsilon > 0$  a parameter. A subset  $\mathcal{K}^* \subseteq \mathcal{H}^*$  is an  $\varepsilon$ -kernel of  $\mathcal{H}^*$  within a  $(d - 1)$ -dimensional region  $\Delta \subseteq \mathbb{R}^{d-1}$  if and only if  $\mathcal{K}$  is an  $\varepsilon$ -kernel of  $\mathcal{H}$  within  $\Delta$ .*

### 3 Approximating the Extent of Linear Functions

In this section we describe algorithms for computing an  $\varepsilon$ -kernel of a set of linear functions whose size depends only on  $\varepsilon$  and  $d$ . We first show that if we can compute an  $\varepsilon$ -kernel of a “fat” point set contained in the unit hypercube  $\mathbb{C} = [-1, +1]^d$ , then we can also compute an  $\varepsilon$ -kernel of an arbitrary point set. We then describe fast algorithms for computing  $\varepsilon$ -kernels of fat point sets. Finally, we use Lemma 2.4 to construct  $\varepsilon$ -kernels of the extent of linear functions.

**Reduction to a fat point set.** We begin by proving a simple lemma, which will be crucial for reducing the problem of computing an  $\varepsilon$ -kernel for an arbitrary point set to the same problem for a fat point set.

**Lemma 3.1** *Let  $T(x) = Mx$  be an affine transformation from  $\mathbb{R}^d$  to  $\mathbb{R}^d$ , where  $M \in \mathbb{R}^{d \times d}$  is non-singular, let  $P$  be a set of points in  $\mathbb{R}^d$ , and let  $\Delta$  be a  $(d - 1)$ -dimensional convex region in  $\mathbb{R}^{d-1}$ . Define*

$$\widehat{M}(\Delta) = \{u \in \mathbb{R}^{d-1} \mid u^* = \pm\phi(M^T \tilde{z}) \text{ for some } z \in \Delta\};$$

where  $u^* = \phi(\tilde{u})$  as defined above (see Table 1 for notations).<sup>2</sup>  $T(Q) \subseteq T(P)$  is an  $\varepsilon$ -kernel of  $T(P)$  within  $\Delta$  if and only if  $Q$  is an  $\varepsilon$ -kernel of  $P$  within  $\widehat{M}(\Delta)$ .

*Proof:* For any vector  $x \in \mathbb{R}^d$ ,

$$\langle x, Mp \rangle = x^T Mp = \langle M^T x, p \rangle.$$

Therefore for any  $z \in \mathbb{R}^{d-1}$ ,

$$\begin{aligned} \overline{\omega}(\tilde{z}, T(Q)) &= \max_{q \in Q} \langle \tilde{z}, Mq \rangle - \min_{q \in Q} \langle \tilde{z}, Mq \rangle \\ &= \max_{q \in Q} \langle M^T \tilde{z}, q \rangle - \min_{q \in Q} \langle M^T \tilde{z}, q \rangle \\ &= \overline{\omega}(M^T \tilde{z}, Q). \end{aligned}$$

---

<sup>2</sup>let  $\Delta^* = \{u^* \mid u \in \Delta\}$ . If  $T(\Delta^*)$  intersects the equator of  $\mathbb{S}^{d-1}$ , then  $\widehat{M}(\Delta)$  can consist of two unbounded regions.

Similarly, we have  $\bar{\omega}(\tilde{z}, T(P)) = \bar{\omega}(M^T \tilde{z}, P)$ .

Suppose  $T(Q)$  is an  $\varepsilon$ -kernel of  $T(P)$  within  $\Delta$ . Consider any  $u \in \widehat{M}(\Delta)$  and let  $z \in \Delta$  be such that  $u^* = \pm\phi(M^T \tilde{z})$ . Since  $T(Q)$  is an  $\varepsilon$ -kernel of  $T(P)$  within  $\Delta$  we have

$$\bar{\omega}(\tilde{z}, T(Q)) \geq (1 - \varepsilon)\bar{\omega}(\tilde{z}, T(P)).$$

Hence,

$$\bar{\omega}(M^T \tilde{z}, Q) = \bar{\omega}(\tilde{z}, T(Q)) \geq (1 - \varepsilon)\bar{\omega}(\tilde{z}, T(P)) = (1 - \varepsilon)\bar{\omega}(M^T \tilde{z}, P).$$

Since  $u^* = \pm\phi(M^T \tilde{z})$ , we conclude using (1) that  $\omega(u, Q) \geq (1 - \varepsilon)\omega(u, P)$ . Thus  $Q$  is an  $\varepsilon$ -kernel of  $P$  within  $\widehat{M}(\Delta)$ .

Conversely, suppose  $Q$  is an  $\varepsilon$ -kernel of  $P$  within  $\widehat{M}(\Delta)$ . Let  $K \subseteq \Delta$  be the set of all points  $z$  such that  $M^T \tilde{z}$  lies on the hyperplane  $x_d = 0$ . Note that  $K$  is contained in a  $(d - 2)$ -dimensional hyperplane in  $\mathbb{R}^{d-1}$ . Consider any  $z \in \Delta \setminus K$ ; there is a  $u \in \widehat{M}(\Delta)$  such that  $u^* = \pm\phi(M^T \tilde{z})$ . Since  $Q$  is an  $\varepsilon$ -kernel of  $P$  within  $\widehat{M}(\Delta)$ , we have  $\omega(u, Q) \geq (1 - \varepsilon)\omega(u, P)$ . This implies (along with (1)) that

$$\bar{\omega}(M^T \tilde{z}, Q) \geq (1 - \varepsilon)\bar{\omega}(M^T \tilde{z}, P).$$

Hence

$$\bar{\omega}(\tilde{z}, T(Q)) = \bar{\omega}(M^T \tilde{z}, Q) \geq (1 - \varepsilon)\bar{\omega}(M^T \tilde{z}, P) = (1 - \varepsilon)\bar{\omega}(\tilde{z}, T(P)),$$

which means that  $T(Q)$  is an  $\varepsilon$ -kernel of  $T(P)$  within  $\Delta \setminus K$ . Since  $K$  is contained in a  $(d - 2)$ -dimensional ‘‘slice’’ of the  $(d - 1)$ -dimensional region  $\Delta$ , a standard limit argument implies that  $T(Q)$  is an  $\varepsilon$ -kernel of  $T(P)$  within  $K$  as well.  $\blacksquare$

We call  $P$   $\alpha$ -fat, for  $\alpha \leq 1$ , if there exists a point  $p \in \mathbb{R}^d$  and a hypercube  $\overline{\mathbb{C}}$  centered at origin so that  $p + \overline{\mathbb{C}} \supset \mathcal{CH}(P) \supset p + \alpha\overline{\mathbb{C}}$ .

**Lemma 3.2** *Let  $P$  be a set of  $n$  points in  $\mathbb{R}^d$ , and let  $\varepsilon$  be a parameter. We can find in  $O(n)$  time an affine transform  $T$  such that  $T(P)$  is  $\alpha_d$ -fat, where  $\alpha_d$  is a constant depending only on  $d$ .*

*Proof:* Using the algorithm of Barequet and Har-Peled [10], we compute in  $O(n)$  time two concentric, homothetic boxes  $B'$  and  $B$  such that

- (a)  $B$  is obtained from  $B'$  by scaling by a factor of at most  $a_d$ , a constant that depends only on  $d$ ,
- (b)  $B' \subseteq \mathcal{CH}(P) \subseteq B$ .

Let  $R \in \mathbb{R}^{d \times d}$  be a rotation transform such that  $R(B)$  is an axes-parallel box. Finally, let  $S$  be the scaling transform that maps  $R(B)$  to a translate of  $\mathbb{C}$ . Set  $T(x) = (S \cdot R)x$ . By construction, the point set  $P' = T(P)$  is  $a_d$ -fat. This completes the proof of the first part of the lemma. It is easy to verify that  $M = S \cdot R$  is non-singular.  $\blacksquare$

Lemmas 3.1 and 3.2 imply that it suffices to describe an algorithm for computing an  $\varepsilon$ -kernel of an  $\alpha$ -fat point set for some  $\alpha < 1$ . We assume that  $\mathbb{C} \supset P \supset [-\alpha, \alpha]^d$ ; this is no loss of generality because for any vector  $t \in \mathbb{R}^d$ , if  $Q$  is an  $\varepsilon$ -kernel of  $P$  within  $\{u\}$ , for some  $u \in \mathbb{R}^{d-1}$ , then  $Q + t$  is an  $\varepsilon$ -kernel of  $P + t$  within  $\{u\}$ . The following simple lemma, which follows immediately from the observation that for any  $u \in \mathbb{R}^d$  there is a point  $q \in \mathcal{CH}(P)$  such that  $\langle u, q \rangle \geq \alpha \|u\|$ , will be useful for our analysis.

**Lemma 3.3** Let  $P \subset \mathbb{C}$  be a set of  $n$  points in  $\mathbb{R}^d$ , which is  $\alpha$ -fat. For any  $x \in \mathbb{R}^d$ ,  $\bar{\omega}(x, P) \geq 2\alpha \|x\|$ .

**A weaker bound on  $\varepsilon$ -kernel.** Next, we prove a weaker bound on the size of an  $\varepsilon$ -kernel for a fat point set.

**Lemma 3.4** Let  $P$  be a  $\alpha$ -fat point set contained in  $\mathbb{C} = [-1, +1]^d$ , and let  $\varepsilon > 0$  be a parameter. Suppose  $P'$  is a point set with the following property: for any  $p \in P$ , there is a  $p' \in P'$  such that  $d(p, p') \leq \varepsilon\alpha$ . Then  $(1 - \varepsilon)\bar{\omega}(x, P) \leq \bar{\omega}(x, P')$  for any  $x \in \mathbb{R}^d$ .

*Proof:* By Lemma 3.3,  $\bar{\omega}(x, P) \geq 2\alpha \|x\|$ . Let  $p, q \in P$  be two points such that

$$\bar{\omega}(x, \{p, q\}) = \bar{\omega}(x, P) \geq 2\alpha \|x\|,$$

and let  $p', q' \in P'$  be two points such that  $d(p, p'), d(q, q') \leq \varepsilon\alpha$ .

Let  $w = p - q$  and  $w' = p' - q'$ . Then

$$\|w - w'\| \leq \|p - p'\| + \|q - q'\| \leq 2\varepsilon\alpha.$$

Moreover,

$$\begin{aligned} \bar{\omega}(x, \{p, q\}) &= \max\{\langle p, x \rangle, \langle q, x \rangle\} - \min\{\langle p, x \rangle, \langle q, x \rangle\} \\ &= |\langle p, x \rangle - \langle q, x \rangle| = |\langle w, x \rangle|. \end{aligned}$$

Similarly,  $\bar{\omega}(x, \{p', q'\}) = |\langle w', x \rangle|$ .

$$\begin{aligned} \bar{\omega}(x, P) - \bar{\omega}(x, P') &\leq \bar{\omega}(x, \{p, q\}) - \bar{\omega}(x, \{p', q'\}) \\ &= |\langle w, x \rangle| - |\langle w', x \rangle| \\ &\leq |\langle w - w', x \rangle| \leq \|w - w'\| \cdot \|x\| \\ &\leq 2\varepsilon\alpha \|x\| \\ &\leq \varepsilon\bar{\omega}(x, P). \end{aligned}$$

■

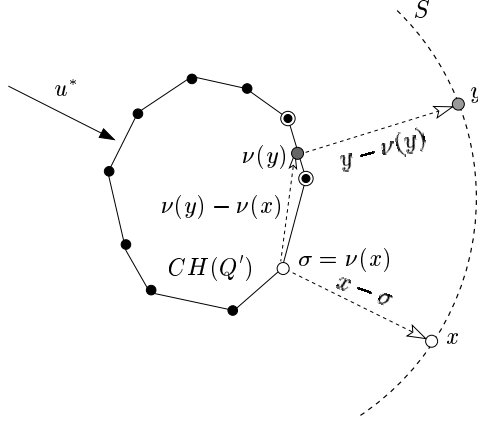
Using the above lemma, we can construct an  $\varepsilon$ -kernel of a fat point set as follows.

**Lemma 3.5** Let  $P$  be a  $\alpha$ -fat point set contained in  $\mathbb{C}$ . For any  $\varepsilon > 0$ , we can compute, in  $O(n + 1/(\alpha\varepsilon)^{d-1})$  time, a subset  $Q \subseteq P$  of  $O(1/(\alpha\varepsilon)^{d-1})$  points that constitutes an  $\varepsilon$ -kernel of  $P$ .

*Proof:* Let  $\delta$  be the largest value such that  $\delta \leq (\varepsilon/\sqrt{d})\alpha$  and  $1/\delta$  is an integer. Observe that  $\delta \geq (\varepsilon/(2\sqrt{d}))\alpha$ . We consider the  $d$ -dimensional grid  $\mathbb{Z}$  of size  $\delta$ . That is,

$$\mathbb{Z} = \{(\delta i_1, \dots, \delta i_d) \mid i_1, \dots, i_d \in \mathbb{Z}\}.$$

For each  $(d-1)$ -tuple  $I = (i_1, \dots, i_{d-1})$ , let  $C_I^+$  (resp.  $C_I^-$ ) be the highest (resp. lowest) cell (in  $x_d$ -direction) of  $\mathbb{Z}$  of the form  $[\delta i_1, \delta(i_1+1)] \times \dots \times [\delta i_{d-1}, \delta(i_{d-1}+1)] \times [\delta r, \delta(r+1)]$ ,  $r \in \mathbb{Z}$ , that contains a point of  $P$ ; if none of the cells in this column contains a point of  $P$ , we define  $C_I^-, C_I^+$  to



**Figure 5.** Illustration of the proof of Lemma 3.6;  $\sigma$  is the farthest vertex of  $\mathcal{CH}(Q')$  in direction  $u^*$ ; the two double-circles denote  $b(y)$  (for  $d = 2$ ).

be the empty set. Let  $B = \bigcup_I (C_I^- \cup C_I^+)$ . Since  $P \subseteq \mathcal{CH}(B)$ ,  $\omega(u, P) \leq \omega(u, \mathcal{CH}(B)) = \omega(u, B)$  for any  $u \in \mathbb{R}^{d-1}$ . Furthermore, we have  $B \subseteq \mathbb{C}$ .

For each  $(d-1)$ -tuple  $I$ , we choose one point from  $P \cap C_I^-$  and another point from  $C_I^+ \cap P$  (if  $C_I^+$  and  $C_I^-$  are not empty) and add both of them to  $Q$ . Since  $P \subseteq \mathbb{C} = [-1, +1]^d$ ,  $|Q| = O(1/(\alpha\varepsilon)^{d-1})$ ;  $Q$  can be constructed in  $O(n+1/(\alpha\varepsilon)^{d-1})$  time, assuming that the ceiling operation (i.e.,  $\lceil \cdot \rceil$ ) can be performed in constant time. For each grid cell  $C$  that contributes to  $B$ , we have chosen in  $Q$  one point from  $P \cap C$ . Therefore for every point  $p \in B$ , there is a point  $q \in Q$  with the property that  $d(p, q) \leq \varepsilon\alpha$ . Hence, by Lemma 3.4, for any  $u \in \mathbb{R}^{d-1}$ ,

$$(1 - \varepsilon)\omega(u, P) \leq (1 - \varepsilon)\omega(u, B) \leq \omega(u, Q),$$

thereby implying that  $Q$  is an  $\varepsilon$ -kernel of  $P$ . ■

**A stronger bound on  $\varepsilon$ -kernel.** Dudley [19] and Bronshteyn and Ivanov [13] have shown that given a convex body  $C$ , which is contained in a unit ball in  $\mathbb{R}^d$ , and a parameter  $\varepsilon > 0$ , one can compute a convex polytope  $C'$  so that the Hausdorff distance between  $C$  and  $C'$  is at most  $\varepsilon$ . Dudley represents  $C'$  as the intersection of  $O(1/\varepsilon^{(d-1)/2})$  halfspaces and Bronshteyn and Ivanov represent  $C'$  as the convex hull of a set of  $O(1/\varepsilon^{(d-1)/2})$  points. In the next lemma we use a variant of the construction in [13] to generate a set of  $O(1/\varepsilon^{(d-1)/2})$  points that forms an  $\varepsilon$ -kernel of  $P$ .

**Lemma 3.6** *Let  $P$  be a  $\alpha$ -fat point set in  $\mathbb{C}$ . For any  $\varepsilon > 0$ , we can compute, in  $O(n+1/(\alpha\varepsilon)^{3(d-1)/2})$  time, a subset  $Q \subseteq P$  of  $O(1/(\alpha\varepsilon)^{(d-1)/2})$  points that is an  $\varepsilon$ -kernel of  $P$ .*

*Proof:* Let  $S$  be the sphere of radius  $\sqrt{d} + 1$  centered at the center of the unit hypercube  $\mathbb{C}$  containing  $P$ . Notice that the distance between any point on the sphere and any point within the unit cube is at least 1. Using Lemma 3.5, we compute a set  $Q' \subseteq P$  of  $O(1/(\alpha\varepsilon)^{d-1})$  points that is an  $\frac{\varepsilon}{2}$ -kernel of  $P$ . Let  $\delta = \sqrt{\varepsilon\alpha/2}$ . We compute a set  $\mathcal{I}$  of  $O(1/\delta^{d-1}) = O(1/(\alpha\varepsilon)^{(d-1)/2})$  points on the sphere  $S$  such that for any point  $x$  on sphere  $S$  (e.g., using the construction in the

proof of Lemma 2.2), there is a point  $y \in \mathcal{I}$  such that  $\|x - y\| \leq \delta$ . For each point  $y \in \mathcal{I}$ , we then compute the point  $\nu(y)$  on  $\mathcal{CH}(Q')$  that is closest to  $y$ . Using the randomized algorithm of Gärtner [21], this can be done for each  $y$  in expected time  $O(|Q'|) = O(1/(\alpha\varepsilon)^{d-1})$ . Gärtner in fact shows that this is an LP-type problem, and therefore we can apply the technique of Chazelle and Matoušek [17] to obtain a deterministic algorithm with running time  $O(|Q'|)$ . These algorithms in fact returns a subset  $b(y) \subseteq Q'$  of at most  $d$  points such that  $\nu(y)$  is in the convex hull of  $b(y)$ . Set  $Q = \bigcup_{y \in \mathcal{I}} b(y)$ . It takes  $O(1/(\alpha\varepsilon)^{3(d-1)/2})$  time to compute  $Q$ , and  $|Q| = O(1/(\alpha\varepsilon)^{(d-1)/2})$ . We now argue that  $Q$  is an  $\varepsilon$ -kernel of  $P$ .

Fix a direction  $u \in \mathbb{R}^{d-1}$ , and let  $u^* \in \mathbb{S}^{d-1}$  be the unit vector  $\phi(\tilde{u})$ . Let  $\sigma \in Q'$  be the point that maximizes  $\langle u^*, q' \rangle$  over all  $q' \in Q'$ . Suppose the ray emanating from  $\sigma$  in direction  $u^*$  hits  $S$  at a point  $x$ . Then  $\sigma$  is the unique point on  $\mathcal{CH}(Q')$  nearest to  $x$ , i.e.,  $\sigma = \nu(x)$ , because the hyperplane normal to the vector  $\sigma - x$  supports  $\mathcal{CH}(Q')$  at  $\sigma$  and separates  $x$  from  $Q'$ . Moreover,

$$\frac{x - \nu(x)}{\|x - \nu(x)\|} = \phi(x - \nu(x)) = \phi(x - \sigma) = u^* \quad \text{and} \quad \|x - \nu(x)\| \geq 1 = \|u^*\|. \quad (2)$$

Let  $y \in \mathcal{I}$  be such that  $\|x - y\| \leq \delta$ . Since  $\nu(y)$  is the closest point to  $y$  in  $\mathcal{CH}(Q')$ , the hyperplane normal to  $y - \nu(y)$  and passing through  $\nu(y)$  separates  $y$  and  $\nu(x)$ , therefore

$$0 \leq \langle y - \nu(y), \nu(y) - \nu(x) \rangle. \quad (3)$$

See Figure 5. Note that for any  $a, b \in \mathbb{R}^d$ ,  $2\langle a, b \rangle \leq \|a\|^2 + \|b\|^2$ , therefore

$$\langle a, b \rangle - \|b\|^2 \leq \|a\|^2. \quad (4)$$

Now,

$$\begin{aligned} 0 \leq \max_{q' \in Q'} \langle u^*, q' \rangle - \max_{q \in Q} \langle u^*, q \rangle &\leq \langle u^*, \sigma \rangle - \langle u^*, \nu(y) \rangle = \langle u^*, \nu(x) - \nu(y) \rangle \\ &\leq \langle x - \nu(x), \nu(x) - \nu(y) \rangle \text{ (using (2))} \\ &\leq \langle x - \nu(x), \nu(x) - \nu(y) \rangle + \langle y - \nu(y), \nu(y) - \nu(x) \rangle \\ &\quad \text{(using (3))} \\ &\leq \langle x - \nu(x) - (y - \nu(y)), \nu(x) - \nu(y) \rangle \\ &= \langle x - y, \nu(x) - \nu(y) \rangle - \|\nu(x) - \nu(y)\|^2 \\ &\leq \|x - y\|^2 \quad \text{(using (4))} \\ &\leq \delta^2 = \alpha\varepsilon/2. \end{aligned}$$

Hence,

$$\max_{q \in Q} \langle \tilde{u}, q \rangle \geq \max_{q' \in Q'} \langle \tilde{u}, q' \rangle - \frac{\alpha\varepsilon}{2} \|\tilde{u}\|.$$

Similarly, we have

$$\min_{q \in Q} \langle \tilde{u}, q \rangle \leq \min_{q' \in Q'} \langle \tilde{u}, q' \rangle + \frac{\alpha\varepsilon}{2} \|\tilde{u}\|.$$

These two inequalities imply that  $\bar{\omega}(\tilde{u}, Q) \geq \bar{\omega}(\tilde{u}, Q') - \alpha\varepsilon \|\tilde{u}\|$ . Using Lemma 3.3, we obtain

$$\begin{aligned}\omega(u, Q) &= \bar{\omega}(\tilde{u}, Q) \\ &\geq \bar{\omega}(\tilde{u}, Q') - \alpha\varepsilon \|\tilde{u}\| \\ &\geq (1 - \varepsilon/2)\bar{\omega}(\tilde{u}, P) - (\varepsilon/2)\bar{\omega}(\tilde{u}, P) \\ &\geq (1 - \varepsilon)\bar{\omega}(\tilde{u}, P) = (1 - \varepsilon)\omega(u, P).\end{aligned}$$

■

Combining Lemmas 3.5 and 3.6 with Lemma 3.2, we obtain the following result.

**Theorem 3.7** *Let  $P$  be a point set in  $\mathbb{R}^d$ , and let  $\varepsilon > 0$  be a parameter. We can compute in  $O(n + 1/\varepsilon^{d-1})$  time an  $\varepsilon$ -kernel of  $P$  of size  $O(1/\varepsilon^{d-1})$ , or in  $O(n + 1/\varepsilon^{3(d-1)/2})$  time an  $\varepsilon$ -kernel of  $P$  of size  $O(1/\varepsilon^{(d-1)/2})$ .*

*Proof:* Using Lemma 3.2, we compute an affine transformation  $M$  such that  $M(P)$  is  $\alpha_d$ -fat. As mentioned above, we can assume that  $M(P) \subseteq \mathbb{C}$ . Using Lemma 3.5 or Lemma 3.6, we compute an  $\varepsilon$ -kernel  $M(Q)$  of  $M(P)$ . Lemma 3.1 (applied to  $M^{-1}$ ) immediately implies that  $P$  is an  $\varepsilon$ -kernel of  $Q$ . ■

Combining this theorem with Lemma 2.4, we obtain the following.

**Theorem 3.8** *Let  $\mathcal{H}$  be a set of  $n$   $(d-1)$ -variate linear functions, and let  $\varepsilon > 0$  be a parameter. We can compute in  $O(n + 1/\varepsilon^{d-1})$  time an  $\varepsilon$ -kernel of  $\mathcal{H}$  of size  $O(1/\varepsilon^{d-1})$ , or in  $O(n + 1/\varepsilon^{3(d-1)/2})$  time an  $\varepsilon$ -kernel of size  $O(1/\varepsilon^{(d-1)/2})$ .*

**A decomposition based bound.** Next, we show that we can decompose  $\mathbb{R}^{d-1}$  into cells so that a pair of points make up an  $\varepsilon$ -kernel for the entire point set within each cell of the decomposition.

**Lemma 3.9** *Let  $P$  be an  $\alpha$ -fat point set contained in  $\mathbb{C}$ , and let  $\varepsilon > 0$  be a parameter. We can compute, in  $O(n + 1/(\alpha\varepsilon)^{3(d-1)/2})$  time, a set  $\mathcal{J}$  of  $O(1/(\alpha\varepsilon))$   $d(d-1)$ -uniform hyperplanes in  $\mathbb{R}^{d-1}$  with the following property: for any cell  $\Delta \in \mathcal{A}(\mathcal{J})$ , there are two points  $p_\Delta, p'_\Delta$  such that  $\{p_\Delta, p'_\Delta\}$  is an  $\varepsilon$ -kernel of  $P$  inside  $\Delta$ .*

*Proof:* We first use Lemma 3.6 to compute a subset  $Q$  of  $O(1/(\alpha\varepsilon)^{(d-1)/2})$  points, which is an  $(\varepsilon/2)$ -kernel of  $P$ . We compute a set  $\mathcal{J}$  of  $O(1/(\alpha\varepsilon))$  hyperplanes, using Lemma 2.2, such that for any two points  $u, v$  in the same cell of  $\mathcal{A}(\mathcal{J})$ ,

$$\|u^* - v^*\| \leq \frac{\varepsilon}{4\sqrt{d}}\alpha.$$

We choose any point  $u_\Delta$  from each cell  $\Delta \in \mathcal{A}(\mathcal{J})$  and compute the points  $p_\Delta$  and  $p'_\Delta$ , by examining each point in  $Q$ , that achieve  $\max_{q \in Q} \langle u^*_\Delta, q \rangle$  and  $\min_{q \in Q} \langle u^*_\Delta, q \rangle$ , respectively. We associate the points  $p_\Delta$  and  $p'_\Delta$  with  $\Delta$ .

By Lemma 2.2,  $\mathcal{A}(\mathcal{J})$  can be computed in  $O(n + 1/(\alpha\varepsilon)^{d-1})$  time. We spend  $O(1/(\alpha\varepsilon)^{(d-1)/2})$  time at each cell  $\Delta \in \mathcal{A}(\mathcal{J})$  to compute  $p_\Delta, p'_\Delta$ . So the total running time of the algorithm is  $O(n + 1/(\alpha\varepsilon)^{3(d-1)/2})$ .

We now argue that  $\{p_\Delta, p'_\Delta\}$  is an  $\varepsilon$ -kernel of  $P$  within  $\Delta$ . Let  $u = u_\Delta$ ,  $p = p_\Delta$ ,  $p' = p'_\Delta$ . Let  $v$  be another point in  $\Delta$ , and let  $q$  and  $q'$  be the points in  $Q$  that achieve  $\max_{q \in Q} \langle v^*, q \rangle$  and  $\min_{q \in Q} \langle v^*, q \rangle$ , respectively. Since  $Q \subset \mathbb{C}$ ,  $\|p - q\| \leq 2\sqrt{d}$ .

$$\begin{aligned}
\langle v^*, p \rangle &= \langle u^*, p \rangle + \langle v^* - u^*, p \rangle \\
&\geq \langle u^*, q \rangle + \langle v^* - u^*, p \rangle \\
&= \langle v^*, q \rangle - \langle v^* - u^*, q \rangle + \langle v^* - u^*, p \rangle \\
&\geq \langle v^*, q \rangle - \|v^* - u^*\| \cdot \|p - q\| \\
&\geq \langle v^*, q \rangle - \frac{\varepsilon\alpha}{4\sqrt{d}} \cdot 2\sqrt{d} \\
&\geq \langle v^*, q \rangle - \frac{\varepsilon\alpha}{2} = \langle v^*, q \rangle - \frac{\varepsilon\alpha}{2} \|v^*\|.
\end{aligned}$$

Therefore  $\langle \tilde{v}, p \rangle \geq \langle \tilde{v}, q \rangle - (\varepsilon\alpha/2) \|\tilde{v}\|$ . By similar reasoning, we obtain that  $\langle \tilde{v}, p' \rangle \leq \langle \tilde{v}, q' \rangle + (\varepsilon\alpha/2) \|\tilde{v}\|$ . Subtracting this from the previous inequality, we get

$$\begin{aligned}
\omega(v, \{p_\Delta, p'_\Delta\}) &\geq \omega(v, Q) - \varepsilon\alpha \|\tilde{v}\| \geq (1 - \varepsilon/2)\omega(v, P) - (\varepsilon/2)\omega(v, P) \\
&\geq (1 - \varepsilon)\omega(v, P)
\end{aligned}$$

This completes the proof of the lemma. ■

**Theorem 3.10** *Let  $P$  be a set of  $n$  points in  $\mathbb{R}^d$ , and let  $\varepsilon > 0$  be a parameter. We can compute, in  $O(n + 1/\varepsilon^{3(d-1)/2})$  time, a set  $\mathcal{J}$  of  $O(1/\varepsilon)$   $d(d-1)$ -uniform hyperplanes in  $\mathbb{R}^{d-1}$  with the following property: for any cell  $\Delta \in \mathcal{A}(\mathcal{J})$ , there are two points  $p_\Delta, p'_\Delta$  such that  $\{p_\Delta, p'_\Delta\}$  is an  $\varepsilon$ -kernel of  $P$  inside  $\Delta$ .*

*Proof:* By Lemma 3.2, let  $T(x) = Mx$  be the affine transform so that  $T(P)$  is  $\alpha_d$ -fat. Using Lemma 3.9, we compute a set  $H$  of  $O(1/\varepsilon)$  hyperplanes in  $\mathbb{R}^{d-1}$  so that for any cell  $\Delta \in \mathcal{A}(H)$ , there are two points  $T(q_\Delta), T(q'_\Delta)$  such that  $\{T(q_\Delta), T(q'_\Delta)\}$  is an  $\varepsilon$ -kernel of  $T(P)$  within  $\Delta$ . For a hyperplane  $h \in H$ , let  $h'$  be the  $(d-1)$ -hyperplane containing  $\tilde{h} = \{\tilde{x} \mid x \in h\}$  and passing through origin, and let  $\hat{h} = M^T h' \cap \mathbb{P}$ . Let  $g'$  denote the  $(d-1)$ -hyperplane  $x_d = 0$  and let  $\hat{g} = M^T g' \cap \mathbb{P}$ . We set  $\mathcal{J} = \{\hat{h} \mid h \in H\} \cup \{\hat{g}\}$ . (We add  $\hat{g}$  because two “antipodal” unbounded cells in  $\mathcal{A}(H)$  may get merged into a single cell in  $\mathcal{A}(\mathcal{J} \setminus \{\hat{g}\})$ .) Clearly,  $\mathcal{J} \setminus \{\hat{g}\}$  is a  $d(d-1)$ -uniform family. It follows from the construction of  $H$  that  $\mathcal{J}$  is a  $d(d-1)$ -uniform family as well. It can be argued that any  $(d-1)$ -dimensional cell  $C$  in  $\mathcal{A}(\mathcal{J})$  is contained in  $\widehat{M}(\Delta)$  for some cell  $\Delta$  in  $\mathcal{A}(H)$ . If  $\{T(q_\Delta), T(q'_\Delta)\}$  is an  $\varepsilon$ -kernel of  $T(P)$  within  $\Delta$ , we associate  $q_\Delta$  and  $q'_\Delta$  with  $C$ . Lemma 3.1 implies that  $\{q_\Delta, q'_\Delta\}$  is an  $\varepsilon$ -kernel of  $P$  within  $C$ . For a lower-dimensional cell  $D$  in  $\mathcal{A}(\mathcal{J})$ , we choose a  $(d-1)$ -dimensional cell  $C$  in  $\mathcal{A}(\mathcal{J})$  such that  $D \subseteq \partial C$ . If  $q$  and  $q'$  are the points associated with  $C$ , we associate them with  $D$  as well. A standard limit argument shows that  $\{q, q'\}$  is an  $\varepsilon$ -kernel of  $P$  within  $D$ . ■

Finally, using Lemma 2.4 we conclude the following.

**Theorem 3.11** *Given a family  $\mathcal{H}$  of  $n$   $(d-1)$ -variate linear functions and a parameter  $\varepsilon > 0$ , we can compute in  $O(n + 1/\varepsilon^{3(d-1)/2})$  time a family  $\mathcal{J}$  of  $O(1/\varepsilon)$   $d(d-1)$ -uniform hyperplanes*

in  $\mathbb{R}^{d-1}$  with the following property: for each cell  $\Delta \in \mathcal{A}(\mathcal{J})$ , there are two associated linear functions,  $h'_\Delta, h''_\Delta \in \mathcal{H}$  such that  $\{h'_\Delta, h''_\Delta\}$  is an  $\varepsilon$ -kernel of  $\mathcal{H}$  inside  $\Delta$ .

## 4 Approximating the Extent for Polynomials and Their Variants

**Extent of polynomials.** Let  $\mathcal{F} = \{f_1, \dots, f_n\}$  be a family of  $(d-1)$ -variate polynomials and  $\varepsilon > 0$  a parameter. We use the linearization technique [5, 33] to compute  $\varepsilon$ -kernels for  $\mathcal{F}$ .

Let  $f(x, a)$  be a  $(d+p-1)$ -variate polynomial,  $x \in \mathbb{R}^{d-1}$  and  $a \in \mathbb{R}^p$ , such that  $f_i(x) \equiv f(x, a^i)$  for some  $a^i \in \mathbb{R}^p$ . There always exists such a polynomial for  $\mathcal{F}$ . Suppose we can express  $f(x, a)$  in the form

$$f(x, a) = \psi_0(a) + \psi_1(a)\varphi_1(x) + \dots + \psi_k(a)\varphi_k(x), \quad (5)$$

where  $\psi_0, \dots, \psi_k$  are  $p$ -variate polynomials and  $\varphi_1, \dots, \varphi_k$  are  $(d-1)$ -variate polynomials. We define the map  $\varphi : \mathbb{R}^{d-1} \rightarrow \mathbb{R}^k$

$$\varphi(x) = (\varphi_1(x), \dots, \varphi_k(x)).$$

Then the image  $\Gamma = \{\varphi(x) \mid x \in \mathbb{R}^{d-1}\}$  of  $\mathbb{R}^{d-1}$  is a  $(d-1)$ -dimensional surface in  $\mathbb{R}^k$ , and for any  $a \in \mathbb{R}^p$ ,  $f(x, a)$  maps to a  $k$ -variate linear function

$$h_a(y_1, \dots, y_k) = \psi_0(a) + \psi_1(a)y_1 + \dots + \psi_k(a)y_k$$

in the sense that for any  $x \in \mathbb{R}^{d-1}$ ,  $f(x, a) = h_a(\varphi(x))$ . We refer to  $k$  as the *dimension* of linearization. The simplest way to express the polynomial  $f(x, a)$  in the form (5) is to write  $f$  as a sum of monomials in  $x_1, \dots, x_{d-1}$  with its coefficients being polynomials in  $a_1, \dots, a_p$ . Then each monomial in the  $x_1, \dots, x_{d-1}$  corresponds to one function  $\varphi_i$ , and its coefficient is the corresponding function  $\psi_i$ . However, this method does not necessarily give a linearization of the smallest dimension. For example, let  $f(x_1, x_2, a_1, a_2, a_3)$  be the square of the distance between a point  $(x_1, x_2) \in \mathbb{R}^2$  and a circle with center  $(a_1, a_2)$  and radius  $a_3$ , which is the 5-variate polynomial

$$f(x_1, x_2, a_1, a_2, a_3) = a_3^2 - (x_1 - a_1)^2 - (x_2 - a_2)^2.$$

A straightforward application of the above method yields a linearization of dimension 4. However,  $f$  can be written in the form

$$f(x_1, x_2, a_1, a_2, a_3) = [a_3^2 - a_1^2 - a_2^2] + [2a_1x_1] + [2a_2x_2] - [x_1^2 + x_2^2], \quad (6)$$

thus, setting

$$\begin{aligned} \psi_0(a) &= a_3^2 - a_1^2 - a_2^2, & \psi_1(a) &= 2a_1, & \psi_2(a) &= 2a_2, & \psi_3(a) &= -1, \\ \varphi_1(x) &= x_1, & \varphi_2(x) &= x_2, & \varphi_3(x) &= x_1^2 + x_2^2, \end{aligned}$$

we get a linearization of dimension 3. It corresponds to the well-known ‘‘lifting’’ transform to the unit paraboloid. Agarwal and Matoušek [5] describe an algorithm that computes a linearization of the smallest dimension.



Returning to the problem of computing an  $\varepsilon$ -kernel of  $\mathcal{F}$ , let  $\mathcal{H} = \{h_{a_i} \mid 1 \leq i \leq n\}$ . Let  $\mathcal{K}$  be an  $\varepsilon$ -kernel of  $\mathcal{H}$  within a region  $\Delta \in \mathbb{R}^k$ . Since  $f_i(x) = h_{a_i}(\varphi(x))$  for any  $x \in \mathbb{R}^{d-1}$ ,  $\mathcal{G} = \{f_i \mid h_{a_i} \in \mathcal{K}\}$  is an  $\varepsilon$ -kernel of  $\mathcal{F}$  within the region  $\varphi^{-1}(\Delta \cap \Gamma)$ , where  $\varphi^{-1}(\gamma) = \{x \in \mathbb{R}^{d-1} \mid \varphi(x) \in \gamma\}$ , for  $\gamma \in \mathbb{R}^k$ , is the pre-image of  $\gamma$  in  $\mathbb{R}^{d-1}$ . Hence, by Theorem 3.8, we obtain the following.

**Theorem 4.1** *Let  $\mathcal{F} = \{f_1, \dots, f_n\}$  be a family of  $(d-1)$ -variate polynomials that admits a linearization of dimension  $k$ , and let  $\varepsilon > 0$  be a parameter. We can compute an  $\varepsilon$ -kernel of  $\mathcal{F}$  of size  $O(1/\varepsilon^k)$  in time  $O(n + 1/\varepsilon^k)$ , or an  $\varepsilon$ -kernel of size  $O(1/\varepsilon^{k/2})$  in time  $O(n + 1/\varepsilon^{3k/2})$ .*

For a  $(k-1)$ -dimensional hyperplane  $h$  in  $\mathbb{R}^k$ , let  $h^{-1}$  denote the pre-image  $\varphi^{-1}(h \cap \Gamma)$  in  $\mathbb{R}^{d-1}$ ;  $h^{-1}$  is a  $(d-2)$ -dimensional algebraic variety, whose degree depends on the maximum degree of a polynomial in  $\mathcal{F}$  and on  $d$ . Using Theorem 3.11, we can prove the following.

**Theorem 4.2** *Let  $\mathcal{F} = \{f_1, \dots, f_n\}$  be a family of  $(d-1)$ -variate polynomials of bounded maximum degree that admits a linearization of dimension  $k$ , and let  $\varepsilon > 0$  be a parameter. We can compute in time  $O(n + 1/\varepsilon^{3k/2})$  a family  $\mathcal{G}$  of  $O(1/\varepsilon)$  algebraic varieties, whose degrees depend on  $d$  and the maximum degree of a polynomial in  $\mathcal{F}$ , so that for any cell  $\tau$  of  $\mathcal{A}(\mathcal{G})$ , there are two polynomials  $f_\tau, f'_\tau \in \mathcal{F}$  such that  $\{f_\tau, f'_\tau\}$  is an  $\varepsilon$ -kernel of  $\mathcal{F}$  within  $\tau$ .*

*Proof:* Let  $\mathcal{H}$  be the linearization of  $\mathcal{F}$  of dimension  $k$ . By Theorem 3.11, we can compute in  $O(n + 1/\varepsilon^{3k/2})$  time a set  $\mathcal{K}$  of  $O(1/\varepsilon)$   $k(k-1)$ -uniform hyperplanes in  $\mathbb{R}^k$  such that for any cell  $\Delta$  of  $\mathcal{A}(\mathcal{K})$ , there exist two hyperplanes  $h_\Delta, h'_\Delta$  such that  $\{h_\Delta, h'_\Delta\}$  is an  $\varepsilon$ -kernel of  $\mathcal{H}$  within  $\Delta$ . Set  $\mathcal{G} = \{h^{-1} \mid h \in \mathcal{K}\}$ . Each cell  $\tau$  in  $\mathcal{A}(\mathcal{G})$  is the pre-image  $\varphi^{-1}(\Delta \cap \Gamma)$  of some cell  $\Delta \in \mathcal{A}(\mathcal{H})$ . For each cell  $\tau \in \mathcal{A}(\mathcal{G})$ , which is the pre-image of  $\Delta \cap \Gamma$ , we set  $f_\tau = h_\Delta^{-1}$  and  $f'_\tau = h'^{-1}_\Delta$ . It is easily seen that  $\{f_\tau, f'_\tau\}$  is an  $\varepsilon$ -kernel of  $\mathcal{F}$  within  $\tau$ . ■

Since  $\bigcup_{\tau \in \mathcal{A}(\mathcal{G})} \{f_\tau, f'_\tau\}$  is an  $\varepsilon$ -kernel of  $\mathcal{F}$  and  $\mathcal{A}(\mathcal{G})$  has  $O(1/\varepsilon^{d-1})$  cells [9], combining this observation with Theorem 4.1 we can conclude the following.

**Theorem 4.3** *Let  $\mathcal{F} = \{f_1, \dots, f_n\}$  be a family of  $(d-1)$ -variate polynomials that admits a linearization of dimension  $k$ , and let  $\varepsilon > 0$  be a parameter. We can compute in time  $O(n + 1/\varepsilon^{3k/2})$  an  $\varepsilon$ -kernel of  $\mathcal{F}$  of size  $O(1/\varepsilon^\sigma)$ , where  $\sigma = \min\{d-1, k/2\}$ .*

Unlike an arrangement of hyperplanes, it is not known whether an arrangement of  $m$  algebraic surfaces in  $\mathbb{R}^d$ , each of constant degree, can be decomposed into  $O(m^d)$  Tarski cells.<sup>3</sup> However, such a decomposition is feasible for the surfaces in Theorem 4.1. Indeed, by construction in the proof of Lemma 2.2, each cell  $\Delta$  in  $\mathcal{A}(\mathcal{K})$  has  $O(1)$  faces, so its pre-image  $\varphi^{-1}(\Delta \cap \Gamma)$  also has  $O(1)$  complexity. We can further refine it into  $O(1)$  Tarski cells. Hence, we can decompose  $\mathcal{A}(\mathcal{G})$  into  $O(1/\varepsilon^{d-1})$  Tarski cells.

**Theorem 4.4** *Let  $\mathcal{F} = \{f_1, \dots, f_n\}$  be a family of  $(d-1)$ -variate polynomials that admits a linearization of dimension  $k$ , and let  $\varepsilon > 0$  be a parameter. We can compute in time  $O(n + 1/\varepsilon^{3k/2})$  a decomposition  $\Xi$  of  $\mathbb{R}^{d-1}$  into  $O(1/\varepsilon^{d-1})$  Tarski cells with the following property: for each cell  $\tau$  in  $\Xi$ , there are two polynomials  $f_\tau, f'_\tau \in \mathcal{F}$  such that  $\{f_\tau, f'_\tau\}$  is an  $\varepsilon$ -kernel of  $\mathcal{F}$  within  $\tau$ .*

<sup>3</sup>A  $k$ -dimensional semialgebraic set is called a *Tarski cell* if it is homeomorphic to a  $k$ -dimensional ball and it is defined by constant number of polynomial inequalities, each of which has bounded degree.

**Remark 4.5** Note that the results of Theorems 4.3 and 4.4 are somewhat surprising. In particular, it implies that if  $\mathcal{F}$  is a family of polynomials defined over a single variable (i.e.,  $d = 2$ ), then  $\mathcal{F}$  has an  $\varepsilon$ -kernel of size  $O(1/\varepsilon)$ . We use this observation in Theorem 6.7.

**Fractional powers of polynomials.** We now consider the problem of computing an  $\varepsilon$ -kernel of a family of functions  $\mathcal{F} = \{(f_1)^{1/r}, \dots, (f_n)^{1/r}\}$ , where  $r \geq 1$  is an integer and each  $f_i$  is a polynomial of some bounded degree. This case is considerably harder than handling polynomials because they can not be linearized directly. In certain special cases this can be overcome by special considerations of the functions at hand [2, 16]. We, however, prove here that it is enough to compute an  $O(\varepsilon^r)$ -kernel of the polynomials inside the roots. We need the following lemma.

**Lemma 4.6** *Let  $0 < \varepsilon < 1$  be a parameter,  $r \geq 2$  an integer, and let  $\delta = (\varepsilon/2(r-1))^r$ . If we have  $0 \leq a \leq A \leq B \leq b$  and  $B - A \geq (1 - \delta)(b - a)$ , then*

$$B^{1/r} - A^{1/r} \geq (1 - \varepsilon)(b^{1/r} - a^{1/r}).$$

*Proof:* First, observe that for any  $x, y$  and for any integer  $r \geq 0$ ,

$$x^r - y^r = (x - y)(x^{r-1} + x^{r-2}y + \dots + xy^{r-2} + y^{r-1}), \quad (7)$$

and for any  $0 \leq p \leq 1$ ,

$$x^p + y^p \geq (x + y)^p. \quad (8)$$

Using (7),

$$\begin{aligned} B^{1/r} - A^{1/r} &= (B - A) / \left( \sum_{i=0}^{r-1} A^{i/r} B^{1-(i+1)/r} \right) \\ &\geq (1 - \delta)(b - a) / \left( \sum_{i=0}^{r-1} A^{i/r} B^{1-(i+1)/r} \right) \\ &\geq (1 - \delta)(b^{1/r} - a^{1/r}) \left( \sum_{i=0}^{r-1} a^{i/r} b^{1-(i+1)/r} \right) / \left( \sum_{i=0}^{r-1} A^{i/r} B^{1-(i+1)/r} \right). \end{aligned}$$

Therefore, for  $0 \leq i < r$ ,

$$\begin{aligned} a^{i/r} b^{1-(i+1)/r} &\geq a^{i/r} B^{1-(i+1)/r} \\ &\geq a^{i/r} B^{1-(i+1)/r} + \delta^{i/r} B^{1-1/r} - \delta^{i/r} B^{1-1/r} \\ &\geq (a^{i/r} + (\delta B)^{i/r}) B^{1-(i+1)/r} - \delta^{i/r} B^{1-1/r} \\ &\geq (a + \delta B)^{i/r} B^{1-(i+1)/r} - \delta^{i/r} B^{1-1/r} \\ &\quad \text{(Using (8) since } i < r) \\ &\geq A^{i/r} B^{1-(i+1)/r} - \delta^{i/r} B^{1-1/r}. \end{aligned}$$

The last inequality holds because, by our assumption,

$$B - A \geq (1 - \delta)(b - a) \geq (1 - \delta)(B - a) \Rightarrow A - a \leq \delta(B - a) \leq \delta B.$$

Hence,

$$\begin{aligned}
\sum_{i=0}^{r-1} a^{i/r} b^{1-(i+1)/r} &\geq \sum_{i=1}^{r-1} (A^{i/r} B^{1-(i+1)/r} - \delta^{i/r} B^{1-1/r}) + B^{1-1/r} \\
&\geq \sum_{i=1}^{r-1} A^{i/r} B^{1-(i+1)/r} + (1 - (r-1)\delta^{1/r}) B^{1-1/r} \\
&\geq (1 - (r-1)\delta^{1/r}) \sum_{i=0}^{r-1} A^{i/r} B^{1-(i+1)/r}.
\end{aligned}$$

Putting everything together,

$$\begin{aligned}
B^{1/r} - A^{1/r} &\geq (1 - \delta)(b^{1/r} - a^{1/r})(1 - (r-1)\delta^{1/r}) \\
&\geq (1 - (\varepsilon/2(r-1))^r)(1 - \varepsilon/2)(b^{1/r} - a^{1/r}) \\
&\geq (1 - \varepsilon)(b^{1/r} - a^{1/r}).
\end{aligned}$$

■

Hence, by Lemma 4.6, we can obtain the following.

**Theorem 4.7** *Let  $\mathcal{F} = \{f_1, \dots, f_n\}$  be a family of  $(d-1)$ -variate polynomials that are non-negative for every  $x \in \mathbb{R}^{d-1}$ , let  $\varepsilon > 0$  be a parameter, and  $r \geq 1$  be an integer. For any  $\Delta \subseteq \mathbb{R}^{d-1}$ , if  $\mathcal{G}$  is an  $(\varepsilon/2(r-1))^r$ -kernel of  $\mathcal{F}$  within  $\Delta$ , then  $\{(f_i)^{1/r} \mid f_i \in \mathcal{G}\}$  is an  $\varepsilon$ -kernel of  $\{(f_i)^{1/r} \mid f_i \in \mathcal{F}\}$  within  $\Delta$ .*

Combining this with Theorem 4.1 and Theorem 4.3, we can prove the following:

**Corollary 4.8** *Let  $\mathcal{F} = \{(f_1)^{1/r}, \dots, (f_n)^{1/r}\}$  be a family of  $(d-1)$ -variate functions (over  $x = (x_1, \dots, x_{d-1}) \in \mathbb{R}^{d-1}$ ), where  $r \geq 2$  is an integer and each  $f_i$  is a polynomial that is non-negative for every  $x \in \mathbb{R}^{d-1}$ , and let  $\varepsilon > 0$  be a parameter. Suppose  $f_i$ 's admit a linearization of dimension  $k$ . We can compute an  $\varepsilon$ -kernel of  $\mathcal{F}$  of size  $O(1/\varepsilon^{rk})$  in time  $O(n + 1/\varepsilon^{rk})$ , or an  $\varepsilon$ -kernel of size  $O(1/\varepsilon^{r\sigma})$ , where  $\sigma = \min\{d-1, k/2\}$ , in  $O(n + 1/\varepsilon^{3rk/2})$  time.*

Similarly, by Theorem 4.4, we can prove the following.

**Theorem 4.9** *Let  $\mathcal{F} = \{(f_1)^{1/r}, \dots, (f_n)^{1/r}\}$  be a family of  $(d-1)$ -variate functions (over  $x = (x_1, \dots, x_{d-1}) \in \mathbb{R}^{d-1}$ ), where  $r \geq 2$  is an integer and each  $f_i$  is a polynomial that is non-negative for every  $x \in \mathbb{R}^{d-1}$ , and let  $\varepsilon > 0$  be a parameter. Suppose  $f_i$ 's admit a linearization of dimension  $k$ . We compute in  $O(n + 1/\varepsilon^{3rk/2})$  time, a decomposition  $\Xi$  of  $\mathbb{R}^{d-1}$  into  $O(1/\varepsilon^{r(d-1)})$  Tarski cells with the following property: for each cell  $\tau$  in  $\Xi$ , there are two associated functions  $f_\tau, f'_\tau \in \mathcal{F}$  such that  $\{f_\tau, f'_\tau\}$  is an  $\varepsilon$ -kernel of  $\mathcal{F}$  within  $\tau$ .*

## 5 Dynamization

In this section we show that we can adapt our algorithm for maintaining an  $\varepsilon$ -kernel of a set of points or a set of linear functions under insertions and deletions. We describe the algorithm for a set  $P$  of points in  $\mathbb{R}^d$ . We assume the existence of an algorithm  $\mathbb{A}$  that can compute a  $\delta$ -kernel of a subset  $S \subseteq P$  of size  $O(1/\delta^k)$  in time  $O(|S| + T_{\mathbb{A}}(\delta))$ ; obviously  $T_{\mathbb{A}}(\delta) \geq 1/\delta^k$ . We will use  $\mathbb{A}$  to maintain an  $\varepsilon$ -kernel dynamically. We first describe a dynamic data structure of (near) linear size that handles both insertions and deletions. Next, we describe another data structure that uses  $O((\log(n)/\varepsilon)^{O(1)})$  space and handles each insertion in  $O((1/\varepsilon)^{O(1)})$  amortized time.

**A fully dynamic data structure.** We assume that each point in  $P$  has a unique id. Using this id as the key, we store  $P$  in a 2-4-tree  $\mathbf{T}$  of height at most  $2 \log_2 n$ ; each point of  $P$  is stored at a leaf of  $\mathbf{T}$ . Some of the leaves of  $\mathbf{T}$  may be empty, i.e., they do not store any point of  $P$ .  $\mathbf{T}$  is periodically reconstructed, but otherwise the structure of  $\mathbf{T}$  is static—only the information stored at the nodes is updated as points are inserted and deleted.

For a node  $v \in \mathbf{T}$ , let  $P_v \subseteq P$  be the subset of points stored at the leaves in the subtree rooted at  $v$ . We also associate a subset  $Q_v \subseteq P_v$  with  $v$ , which is defined recursively, as follows. Set  $\delta = \varepsilon/3h$ , where  $h$  is the height of  $\mathbf{T}$ . If  $v$  is a leaf, then  $Q_v = P_v$ . For an internal node  $v$  with  $w$  and  $z$  as its children,  $Q_v$  is a  $\delta$ -kernel of  $Q_w \cup Q_z$  of size  $O(1/\delta^k)$ , computed using algorithm  $\mathbb{A}$ . Our construction ensures that for a node at height  $i$  (leaves have height 0),  $Q_v$  is an  $(\varepsilon i/(2h))$ -kernel of  $P_v$  since  $(1 + \varepsilon/3h)^i \leq (1 + \varepsilon i/(2h))$ . Therefore the subset  $Q_{\text{root}}$  associated with the root of  $\mathbf{T}$  is an  $(\varepsilon/2)$ -kernel of  $P$  of size  $O(1/\delta^k)$ . Finally, we maintain an  $(\varepsilon/3)$ -kernel  $Q$  of  $Q_{\text{root}}$  of size  $O(1/\varepsilon^k)$  using algorithm  $\mathbb{A}$ ;  $Q$  is an  $\varepsilon$ -kernel of  $P$ .

Suppose we want to delete a point  $p_i$  from  $P$ . We find the leaf  $z$  that stores  $p_i$ , delete  $p$  from that leaf and make that leaf empty. If the number of points in  $P$  becomes at most one-fourth the number of leaves, we reconstruct  $\mathbf{T}$  with half as many leaves as in the current tree, so that half of the leaves in the new tree are empty. Otherwise, we recompute  $Q_v$  at all ancestors  $v$  of  $z$  in a bottom-up manner. At each ancestor  $v$ , with  $x$  and  $w$  as its children, we compute, in  $O(T_{\mathbb{A}}(\delta))$  time, a  $\delta$ -kernel of  $Q_w \cup Q_x$  using algorithm  $\mathbb{A}$ . Finally, we recompute, in time  $O((1/\delta)^k + T_{\mathbb{A}}(\varepsilon))$ , an  $(\varepsilon/2)$ -kernel  $Q$  of  $Q_{\text{root}}$ . The total time spent is thus  $O(T_{\mathbb{A}}(\delta) \log n)$ . (In fact, one can stop this traversal up the tree when we encounter the first node  $z$ , such that  $Q_z$  does not contain  $p_i$ .) Next, suppose we want to insert a point  $p$ . If there is an empty leaf  $z$ , we insert  $p$  into  $z$  and update the information stored at the ancestors of  $z$ . If there is no empty leaf, we first reconstruct  $\mathbf{T}$  with twice as many leaves as the current tree, so that half of the leaves in the new tree are empty. We now insert  $p$  into some empty leaf and proceed as before.

Since  $\mathbf{T}$  is reconstructed after at least  $n_0/2$  updates, where  $n_0$  is the number of points in  $\mathbf{T}$  when it was last updated, the reconstruction costs  $O(T_{\mathbb{A}}(\varepsilon/6 \log n))$  amortized time per update operation. Hence, we obtain the following.

**Theorem 5.1** *Let  $P$  be a set of points in  $\mathbb{R}^d$ , and let  $\varepsilon > 0$  be a parameter. Suppose we can compute an  $\varepsilon$ -kernel of a subset  $S \subseteq P$  of size  $O(1/\varepsilon^k)$  in time  $O(|S| + T_{\mathbb{A}}(\varepsilon))$  time. Then we can maintain an  $\varepsilon$ -kernel of  $P$  of size  $O(1/\varepsilon^k)$  under insertion/deletion in amortized time  $O(T_{\mathbb{A}}(\varepsilon/6 \log n) \log n)$  per update operation.*

**Remark 5.2** A weakness of our approach is that insertion or deletion of a point can change the  $\varepsilon$ -kernel completely. It would be desirable to develop a dynamic data structure that causes  $O(1)$  change in the  $\varepsilon$ -kernel after insertion or deletion of a point.

**Corollary 5.3** *Let  $\mathcal{F}$  be a set of functions, and let  $\varepsilon > 0$  be a parameter. Suppose we can compute an  $\varepsilon$ -kernel of a subset  $\mathcal{G} \subseteq \mathcal{F}$  of size  $O(1/\varepsilon^k)$  in time  $O(|\mathcal{G}| + T_{\mathbb{A}}(\varepsilon))$  time, then we can maintain an  $\varepsilon$ -kernel of  $\mathcal{F}$  of size  $O(1/\varepsilon^k)$  under insertion/deletion in amortized time  $O(T_{\mathbb{A}}(\varepsilon/6 \log n) \log n)$  per update operation.*

**An insertion-only data structure.** Suppose we are receiving a stream of points  $p_1, p_2, \dots$  in  $\mathbb{R}^d$ . Given a parameter  $\varepsilon > 0$ , we wish to maintain an  $\varepsilon$ -kernel of the  $n$  points received so far. Note that our analysis is in term of  $n$ , the number of points inserted into the data structure. However,  $n$  does not need to be specified in advance. In particular, if  $n$  is specified in advance, a slightly simpler solution arises using the techniques described above. We assume without loss of generality that  $1/\varepsilon$  is an integer. We use the dynamization technique of Bentley and Saxe [12], as follows: Let  $P = \langle p_1, \dots, p_n \rangle$  be the sequence of points that we have received so far. For integers  $i \geq 1$ , let  $\rho_i = \varepsilon/ci^2$ , where  $c > 0$  is a constant, and set  $\delta_i = \prod_{l=1}^i (1 + \rho_l) - 1$ . We partition  $P$  into subsets  $P_0, P_1, \dots, P_u$ , where  $u = \lfloor \log_2 \varepsilon^k n \rfloor + 1$ , as follows.  $|P_0| = n \bmod 1/\varepsilon^k$ , and for  $1 \leq i \leq u$ , if the  $i$ th rightmost bit in the binary expansion of  $\lfloor \varepsilon^k n \rfloor$  is 1, then  $|P_i| = 2^{i-1}/\varepsilon^k$ , otherwise  $|P_i| = 0$ . Furthermore, if  $0 \leq i < j \leq u$ , the points in  $P_j$  arrived before any point in  $P_i$ . These conditions uniquely specify  $P_0, \dots, P_u$ . We refer to  $i$  as the *rank* of  $P_i$ . Note that for  $i \geq 1$ , there is at most one non-empty subset of rank  $i$ .

Unlike the standard Bentley-Saxe technique, we do not maintain each  $P_i$  explicitly. Instead, for each non-empty subset  $P_i$ , we maintain a  $\delta_i$ -kernel  $Q_i$  of  $P_i$ ; if  $P_i = \emptyset$ , we set  $Q_i = \emptyset$  as well. We also let  $Q_0 = P_0$ . Since

$$1 + \delta_i = \prod_{l=1}^i \left(1 + \frac{\varepsilon}{cl^2}\right) \leq \exp\left(\sum_{l=1}^i \frac{\varepsilon}{cl^2}\right) = \exp\left(\frac{\varepsilon}{c} \sum_{l=1}^i \frac{1}{l^2}\right) \leq \exp\left(\frac{\pi^2 \varepsilon}{6c}\right) \leq 1 + \frac{\varepsilon}{3}, \quad (9)$$

provided  $c$  is chosen sufficiently large,  $Q_i$  is an  $(\varepsilon/3)$ -kernel of  $P_i$ . Therefore,  $\bigcup_{i=0}^u Q_i$  is an  $(\varepsilon/3)$ -kernel of  $P$ . We define the *rank* of a set  $Q_i$  to be  $i$ . For  $i \geq 1$ , if  $P_i$  is non-empty,  $|Q_i|$  will be  $O(1/\rho_i^k)$ ; note that  $|Q_0| = |P_0| < 1/\varepsilon^k$ .

For each  $i \geq 0$ , we also maintain an  $\varepsilon/3$ -kernel  $K_i$  of  $\bigcup_{j \geq i} Q_j$ , as follows. Let  $u = \lfloor \log_2(\varepsilon^k n) \rfloor + 1$  be the largest value of  $i$  for which  $P_i$  is non-empty. We have  $K_u = Q_u$ , and for  $1 \leq i < u$ ,  $K_i$  is a  $\rho_i$ -kernel of  $K_{i+1} \cup Q_i$ . Finally,  $K_0 = Q_0 \cup K_1$ . The argument in (9) implies that  $K_i$  is an  $(\varepsilon/3)$ -kernel of  $\bigcup_{j \geq i} Q_j$ , and thus  $K_0$  is the required  $\varepsilon$ -kernel of  $P$ . The size of the entire data structure is

$$\begin{aligned} \sum_{i=0}^u (|Q_i| + |K_i|) &\leq |Q_0| + |K_0| + \sum_{i=1}^u O(1/\rho_i^k) \\ &= O(1/\varepsilon^k) + \sum_{i=1}^{\lfloor \log_2 \varepsilon^k n \rfloor + 1} O\left(\frac{i^{2k}}{\varepsilon^k}\right) = O\left(\frac{\log^{2k+1} n}{\varepsilon^k}\right). \end{aligned}$$

At the arrival of the next point  $p_{n+1}$ , the data structure is updated as follows. We add  $p_{n+1}$  to  $Q_0$  (and conceptually to  $P_0$ ). If  $|Q_0| < 1/\varepsilon^k$  then we are done. Otherwise, we promote  $Q_0$  to have rank 1. Next, if there are two  $\delta_j$ -kernels  $Q_x, Q_y$  of rank  $j$ , for some  $j \leq \lfloor \log_2 \varepsilon^k (n+1) \rfloor + 1$ , we compute a  $\rho_{j+1}$ -kernel  $Q_z$  of  $Q_x \cup Q_y$  using algorithm  $\mathbb{A}$ , set the rank of  $Q_z$  to  $j+1$ , and discard the sets  $Q_x$  and  $Q_y$ . By construction,  $Q_z$  is a  $\delta_{j+1}$ -kernel of  $P_z = P_x \cup P_y$  of size  $O(1/\rho_{j+1}^k)$  and  $|P_z| = 2^j/\varepsilon^k$ . We repeat this step until the ranks of all  $Q_i$ 's are distinct. Suppose  $\xi$  is the maximum rank of a  $Q_i$  that was reconstructed, then we recompute  $K_\xi, \dots, K_0$  in that order. That is, for  $\xi \geq i \geq 1$ , we compute a  $\rho_i$ -kernel of  $K_{i+1} \cup Q_i$  and set this to be  $K_i$ ; finally, we set  $K_0 = K_1 \cup Q_0$ .

For any fixed  $i \geq 1$ ,  $Q_i$  and  $K_i$  are constructed after every  $2^{i-1}/\varepsilon^k$  insertions, therefore the amortized time spent in updating  $Q$  after inserting a point is

$$\sum_{i=1}^{\lfloor \log_2 \varepsilon^k n \rfloor + 1} \frac{\varepsilon^k}{2^{i-1}} O\left(\frac{i^{2k}}{\varepsilon^k} + T_{\mathbb{A}}\left(\frac{\varepsilon}{ci^2}\right)\right) = O\left(\sum_{i=1}^{\lfloor \log_2 \varepsilon^k n \rfloor + 1} \frac{\varepsilon^k}{2^{i-1}} T_{\mathbb{A}}\left(\frac{\varepsilon}{ci^2}\right)\right).$$

If  $T_{\mathbb{A}}(x)$  is bounded by a polynomial in  $1/x$ , then the above expression is bounded by  $O(\varepsilon^k T_{\mathbb{A}}(\varepsilon))$ .

**Theorem 5.4** *Let  $P$  be a stream of points in  $\mathbb{R}^d$ , and let  $\varepsilon > 0$  be a parameter. Suppose that for any subset  $S \subseteq P$ , we can compute an  $\varepsilon$ -kernel of  $S$  of size  $O(1/\varepsilon^k)$  in  $O(|S| + T_{\mathbb{A}}(\varepsilon))$  time, where  $T_{\mathbb{A}}(\varepsilon) \geq 1/\varepsilon^k$  is bounded by a polynomial in  $1/\varepsilon$ . Then we can maintain an  $\varepsilon$ -kernel of  $P$  of size  $O(1/\varepsilon^k)$  using a data structure of size  $O(\log^{2k+1}(n)/\varepsilon^k)$ . The amortized time to insert a point is  $O(\varepsilon^k T_{\mathbb{A}}(\varepsilon))$ , and the worst case time is  $O((\log^{2k+1} n)/\varepsilon^k + T_{\mathbb{A}}(\varepsilon/\log^2 n) \log n)$ .*

**Remark 5.5** The exponent  $2k+1$  in the bounds of the above theorem can be improved to  $k+1+\delta$ , for any  $\delta > 0$ , by being more careful, but we feel this improvement is not worth the effort.

The following is an immediate corollary of Theorems 3.7 and 5.4.

**Corollary 5.6** *Let  $P$  be a stream of points in  $\mathbb{R}^d$ , and let  $\varepsilon > 0$  be a parameter. We can maintain an  $\varepsilon$ -kernel of  $P$  of size  $O(1/\varepsilon^{(d-1)/2})$  using a data structure of size  $O(\log^d(n)/\varepsilon^{(d-1)/2})$ . The amortized time spent at each point is  $O(1/\varepsilon^{d-1})$ .*

## 6 Applications

In this section we present a few specific applications of the results on  $\varepsilon$ -kernels obtained in Sections 3 and 4. We begin by describing approximation algorithms for computing faithful extent measures, and then showing that our technique can be extended to maintaining faithful measures of moving points. Next, we describe approximation algorithms for computing two nonfaithful measures, namely the minimum width of spherical and cylindrical shells that contain a set of points.

### 6.1 Approximating faithful extent measures

A function  $\mu(\cdot)$  defined over a finite set  $P$  of points is called a *faithful measure* if (i) for any  $P \subseteq \mathbb{R}^d$ ,  $\mu(P) \geq 0$ , and (ii) there exists a constant (depending on  $\mu$ )  $c \geq 0$ , so that for any  $\varepsilon$ -kernel  $Q$  of  $P$ ,

$(1 - c\varepsilon)\mu(P) \leq \mu(Q) \leq \mu(P)$ . Examples of faithful measures are common and include diameter, width, radius of the smallest enclosing ball, volume of the minimum bounding box, volume of  $\mathcal{CH}(P)$ , and surface area of  $\mathcal{CH}(P)$ . A common property of all these measures is that  $\mu(P) = \mu(\mathcal{CH}(P))$ . For a given point set  $P$ , a faithful measure  $\mu$ , and a parameter  $\varepsilon > 0$ , we can compute a value  $\bar{\mu}$ ,  $(1 - \varepsilon)\mu(P) \leq \bar{\mu} \leq \mu(P)$  by first computing an  $(\varepsilon/c)$ -kernel  $Q$  of  $P$  and then using an exact algorithm for computing  $\mu(Q)$ . Using Theorems 3.7 and 5.1 we obtain the following.

**Theorem 6.1** *Given a set  $P$  of  $n$  points in  $\mathbb{R}^d$ , a faithful measure  $\mu$  that can be computed in  $n^\alpha$  time, and a parameter  $\varepsilon > 0$ , we can compute, in time  $O(n + f(\varepsilon))$ , a value  $\bar{\mu}$  so that  $(1 - \varepsilon)\mu(P) \leq \bar{\mu} \leq \mu(P)$ , where  $f(\varepsilon) = \min \{1/\varepsilon^{\alpha(d-1)}, 1/\varepsilon^{3(d-1)/2} + 1/\varepsilon^{\alpha(d-1)/2}\}$ . Moreover,  $P$  can be stored in a dynamic data structure that can update  $\bar{\mu}$  in amortized time*

$$\min \left\{ \frac{\log^d n}{\varepsilon^{d-1}} + \frac{1}{\varepsilon^{\alpha(d-1)}}, \frac{\log^{3d/2-1/2} n}{\varepsilon^{3(d-1)/2}} + \frac{1}{\varepsilon^{\alpha(d-1)/2}} \right\}$$

if a point is inserted into or deleted from  $P$ .

For example, since the diameter of a set  $P$  of points in  $\mathbb{R}^d$  can be trivially computed in  $O(n^2)$  time, we can compute an  $\varepsilon$ -approximation of the diameter of  $P$  in  $O(n + 1/\varepsilon^{3(d-1)/2})$  time. Similarly, we can compute in  $O(n + 1/\varepsilon^3)$  time an  $\varepsilon$ -approximation of the volume of the smallest box enclosing a set of  $n$  points in  $\mathbb{R}^3$ , as the exact algorithms for these problems take  $O(n^3)$  time [10, 30]. For all of the measures mentioned in the beginning of this section, algorithms with similar running time (even slightly better in some cases) are already known [10, 16]. However, our technique is general and does not require us to carefully inspect the problem at hand to develop an approximation algorithm.

We can use Corollary 5.6 for maintaining faithful extent measures of a stream of points in  $\mathbb{R}^d$  using  $O(\log^d(n)/\varepsilon^{(d-1)/2})$  space. For instance, applying a result of Duncan et al. [20] that an  $\varepsilon$ -approximation to the width of a set of  $m$  points in  $\mathbb{R}^d$  can be computed in  $O(m/\varepsilon^{(d-1)/2})$  time, we can conclude the following.

**Theorem 6.2** *Given a parameter  $\varepsilon > 0$ , we can maintain an  $\varepsilon$ -approximation of the width of a stream of points in  $\mathbb{R}^d$  using  $O(\log^d(n)/\varepsilon^{(d-1)/2})$  space and spending  $O(1/\varepsilon^{d-1})$  amortized time at each incoming point.*

## 6.2 Maintaining faithful measures of moving points

Next we show that our technique can be extended to maintain various extent measures of a set of moving points. Let  $P = \{p_1, \dots, p_n\}$  be a set of  $n$  points in  $\mathbb{R}^d$ , each moving independently. Let  $p_i(t) = (p_{i1}(t), \dots, p_{id}(t))$  denote the position of point  $p_i$  at time  $t$ . Set  $P(t) = \{p_i(t) \mid 1 \leq i \leq n\}$ . If each  $p_{ij}$  is a polynomial of degree at most  $r$ , we say that the motion of  $P$  has *degree*  $r$ . We call the motion of  $P$  *linear* if  $r = 1$  and *algebraic* if  $r$  is bounded by a constant.

Given a parameter  $\varepsilon > 0$ , we call a subset  $Q \subseteq P$  an  $\varepsilon$ -kernel of  $P$  if for any direction  $u \in \mathbb{R}^{d-1}$ ,

$$(1 - \varepsilon)\omega(u, P(t)) \leq \omega(u, Q(t)) \quad \text{for all } t \in \mathbb{R}.$$

We first show that a small  $\varepsilon$ -kernel of  $P$  can be computed efficiently and then discuss how to use it to maintain a faithful measure of  $P$  approximately as the points move, assuming that the trajectories of points are algebraic and do not change over time. Finally, we show how to update the  $\varepsilon$ -kernel if we allow the trajectories of points to change or if we allow points to be inserted or deleted.

**Computing an  $\varepsilon$ -kernel.** First let us assume that the motion of  $P$  is linear, i.e.,  $p_i(t) = a_i + b_i t$ , for  $1 \leq i \leq n$ , where  $a_i, b_i \in \mathbb{R}^d$ . For a direction  $u = (u_1, \dots, u_{d-1}) \in \mathbb{R}^{d-1}$ , we define a  $d$ -variate polynomial

$$\begin{aligned} f_i(u, t) &= \langle p_i(t), \tilde{u} \rangle = \langle a_i + b_i t, \tilde{u} \rangle \\ &= \sum_{j=1}^{d-1} a_{ij} u_j + \sum_{j=1}^{d-1} b_{ij} \cdot (t u_j) + a_{id} + b_{id} t. \end{aligned}$$

Set  $\mathcal{F} = \{f_1, \dots, f_n\}$ . Then

$$\omega(u, P(t)) = \max_i \langle p_i(t), \tilde{u} \rangle - \min_i \langle p_i(t), \tilde{u} \rangle = \max_i f_i(u, t) - \min_i f_i(u, t) = \mathfrak{E}_{\mathcal{F}}(u, t).$$

Since  $\mathcal{F}$  is a family of  $d$ -variate polynomials, which admits a linearization of dimension  $2d - 1$  (there are  $2d - 1$  monomials), using Theorem 4.1, we conclude the following.

**Theorem 6.3** *Given a set  $P$  of  $n$  points in  $\mathbb{R}^d$ , each moving linearly, and a parameter  $\varepsilon > 0$ , we can compute an  $\varepsilon$ -kernel of  $P$  of size  $O(1/\varepsilon^{2d-1})$  in  $O(n + 1/\varepsilon^{2d-1})$  time, or an  $\varepsilon$ -kernel of size  $O(1/\varepsilon^{d-1/2})$  in  $O(n + 1/\varepsilon^{3(d-1/2)})$  time.*

If the degree of motion of  $P$  is  $r > 1$ , we can write the  $d$ -variate polynomial  $f_i(u, t)$  as:

$$f_i(u, t) = \langle p_i(t), \tilde{u} \rangle = \left\langle \sum_{j=0}^r a_{ij} t^j, \tilde{u} \right\rangle = \sum_{j=0}^r \langle a_{ij} t^j, \tilde{u} \rangle$$

where  $a_{ij} \in \mathbb{R}^d$ . A straightforward extension of the above argument shows that  $f_i$ 's admit a linearization of dimension  $(r + 1)d - 1$ . Using Theorems 4.1 and 4.3, we obtain the following.

**Theorem 6.4** *Given a set  $P$  of  $n$  moving points in  $\mathbb{R}^d$  whose motion has degree  $r > 1$  and a parameter  $\varepsilon > 0$ , we can compute an  $\varepsilon$ -kernel of  $P$  of size  $O(1/\varepsilon^{(r+1)d-1})$  in  $O(n + 1/\varepsilon^{(r+1)d-1})$  time, or of size  $O(1/\varepsilon^d)$  in  $O(n + 1/\varepsilon^{3((r+1)d-1)/2})$  time.*

**Remark 6.5** By Corollary 5.3, if we can compute in time  $O(n + T_{\mathbb{A}}(\varepsilon))$  an  $\varepsilon$ -kernel of size  $O(1/\varepsilon^k)$  of a set  $P$  of  $n$  moving points in  $\mathbb{R}^d$ , then we can update it in time  $O(((\log n)/\varepsilon)^k + T_{\mathbb{A}}((\log n)/\varepsilon) \log n)$  per insertion/deletion of a point.



**Kinetic data structures.** As in Section 6.1, we can use an  $\varepsilon$ -kernel of  $P$  to maintain various faithful extent measure of  $P$  approximately as the points in  $P$  move. Namely, we first compute an  $\varepsilon$ -kernel  $Q$  of  $P$  and then maintain the desired measure for  $Q$ . Note that  $Q$  does not depend on the underlying measure. Agarwal *et al.* [4] have described kinetic data structures for maintaining various extent measures, including diameter, width, area (or perimeter) of the smallest enclosing rectangle, of a set of points moving algebraically in the plane. Plugging their technique on  $Q$ , we can, for example, construct a kinetic data structure of size  $O(|Q|)$  that maintains a pair  $(q, q')$  with the property that

$$d(q(t), q'(t)) = \text{diam}(Q(t)) \geq (1 - \varepsilon) \text{diam}(P(t)).$$

The pair  $(q, q')$  is updated  $O(|Q|^{2+\delta})$  times, for any  $\delta > 0$ , and the data structure can be updated in  $O(\log |Q|)$  time at each such event. Similar bounds hold for width, area of the smallest enclosing rectangle, etc. Applying Theorem 6.3 for linear motion and Theorem 6.4 for higher-degree motion, we obtain the following:

**Theorem 6.6** *Let  $P$  be a set of  $n$  points moving in the plane, and let  $\varepsilon > 0$  be a parameter. If  $P$  is moving linearly, then after  $O(n + 1/\varepsilon^{9/2})$  preprocessing, we can construct a kinetic data structure of size  $O(1/\varepsilon^{3/2})$  so that an  $\varepsilon$ -approximation of diameter, width, or the area (or perimeter) of the smallest enclosing rectangle of  $P$  can be maintained. The data structure processes  $O(1/\varepsilon^{3+\delta})$  events, for an arbitrarily small constant  $\delta > 0$ , and each such event requires  $O(\log(1/\varepsilon))$  time. If the motion of  $P$  has degree  $r$ , then the preprocessing time is  $O(n + 1/\varepsilon^{3r+3/2})$ , the size of the data structure is  $O(1/\varepsilon^2)$ , and the number of events is  $O(1/\varepsilon^{4+\delta})$ .*

In some cases, the size of the  $\varepsilon$ -kernel that we use to maintain a faithful measure can be improved by reducing the problem to a lower dimensional problem. For example, let  $\mathcal{B}(t) = \mathcal{B}(P(t))$  denote the smallest orthogonal box containing  $P(t)$ , and let  $\mathcal{B}^\varepsilon(t) = (1 - 2\varepsilon)\mathcal{B}(t)$ , scaled with respect to the center of  $\mathcal{B}(t)$ . We call a box  $\hat{\mathcal{B}}(t)$  an  $\varepsilon$ -approximation of  $\mathcal{B}(t)$  if  $\mathcal{B}^\varepsilon(t) \subseteq \hat{\mathcal{B}}(t) \subseteq \mathcal{B}(t)$ . Let  $Q$  be an  $\varepsilon$ -kernel of  $P$ , then  $\mathcal{B}^\varepsilon(t) \subseteq \mathcal{B}(Q(t)) \subseteq \mathcal{B}(t)$ , therefore we can compute an  $\varepsilon$ -kernel of size  $O(1/\varepsilon^{d-1/2})$  (if points are moving linearly) and maintain its bounding box. However, one can do better using the following observation.

For  $1 \leq i \leq d$ , let  $P^j(t) = \{p_{ij}(t) \mid 1 \leq i \leq n\}$ . Then  $\mathcal{B}(t) = \beta_1(t) \times \cdots \times \beta_d(t)$ , where  $\beta_j(t)$  is the smallest interval containing  $P^j(t)$ . Hence, the problem of maintaining  $\mathcal{B}(t)$  reduces to maintaining the smallest interval containing  $P^j(t)$ , for each  $j \leq d$  (see also Remark 4.5). We thus compute an  $\varepsilon$ -kernel  $Q^j$  of each  $P^j$  and maintain the smallest interval containing  $Q^j$ ; the latter can be accomplished by maintaining the maximum and minimum of  $Q^j$ , using a kinetic tournament tree described in [11]. The data structure processes  $O(|Q^j| \log |Q^j|)$  events, and each event requires  $O(\log^2 |Q^j|)$  time. Since  $P^j(t)$  is a set of  $n$  points moving in  $\mathbb{R}$ , using Theorem 6.4 and putting everything together, we obtain the following.

**Theorem 6.7** *Let  $P$  be a set of  $n$  points moving in  $\mathbb{R}^d$ , and let  $\varepsilon > 0$  be a parameter. If  $P$  is moving linearly, then after  $O(n + 1/\varepsilon^{3/2})$  preprocessing, we can construct a kinetic data structure of size  $O(1/\sqrt{\varepsilon})$  that maintains an  $\varepsilon$ -approximation of the smallest orthogonal box containing  $P$ ; the data structure processes  $O((1/\sqrt{\varepsilon}) \log(1/\varepsilon))$  events, and each event takes  $O(\log^2(1/\varepsilon))$  time. If the motion of  $P$  has degree  $r > 1$ , then the preprocessing time is  $O(n + 1/\varepsilon^{3r/2})$ , the size of the data*

structure is  $O(1/\varepsilon)$ , the number of events is  $O((1/\varepsilon) \log(1/\varepsilon))$ , and each event takes  $O(\log^2(1/\varepsilon))$  time.

The data structures described above assume that the trajectories of each point is specified in the beginning and it remains fixed. However in most of the applications, we know only a part of the trajectory, and it changes with time. We can handle trajectory updates using the dynamization technique described in Section 5. Since the  $\varepsilon$ -kernel  $Q$  of  $P$  being maintained by our algorithm can change significantly after an update operation, we simply reconstruct the kinetic data structure on  $Q$ . If we can prove a bound on how much  $Q$  changes after an update operation, a kinetic data structure that supports efficient updates can improve the efficiency of our algorithm.

### 6.3 Minimum-width spherical shell

Let  $P = \{p_1, \dots, p_n\}$  be a set of  $n$  points in  $\mathbb{R}^d$ . As defined in Section 1, a spherical shell is (the closure of) the region bounded by two concentric spheres: the width of the shell is the difference of their radii. Let  $d(x, p)$  be the Euclidean distance between  $x$  and  $p$ , and let  $f_p(x) = d(x, p)$ . Set  $\mathcal{F} = \{f_{p_i} \mid p_i \in P\}$ . Let  $w(x, S)$  denote the width of the thinnest spherical shell centered at  $x$  that contains  $S$ , and let  $w^* = w^*(S) = \min_{x \in \mathbb{R}^d} w(x, S)$  be the width of the thinnest spherical shell containing  $S$ . Then

$$w(x, S) = \max_{p \in P} d(x, p) - \min_{p \in P} d(x, p) = \max_{f_p \in \mathcal{F}} f_p(x) - \min_{f_p \in \mathcal{F}} f_p(x) = \mathfrak{E}_{\mathcal{F}}(x).$$

Therefore,  $w^* = \min_{x \in \mathbb{R}^d} \mathfrak{E}_{\mathcal{F}}(x)$ . It thus suffices to compute an  $\varepsilon$ -kernel of  $\mathcal{F}$ . Set

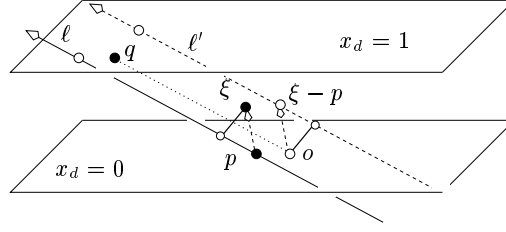
$$g_p(x) = f_p(x)^2 = \|x\|^2 - 2 \langle x, p \rangle + \|p\|^2.$$

As shown in Section 4 (for  $d = 2$ ),  $\mathcal{G} = \{g_{p_i} \mid p_i \in P\}$  admits a linearization of dimension  $d + 1$ . However, let  $g'_i(x) = g_{p_i}(x) - \|x\|^2$ . By Lemma 2.1, an  $\varepsilon$ -kernel of  $\mathcal{G}' = \{g'_1 \dots g'_n\}$  is also an  $\varepsilon$ -kernel of  $\mathcal{G}$ . Since  $\mathcal{G}'$  admits a linearization of dimension  $d$ , we can use Theorem 4.7 (with  $r = 2$ ) and Theorem 4.3 to compute an  $\varepsilon$ -kernel  $Q$  of  $\mathcal{F}$  of size  $O(1/\varepsilon^d)$  in  $O(n + 1/\varepsilon^{3d})$  time and then compute  $w^*(Q)$  in time  $1/\varepsilon^{O(d^2)}$  [2]. However, we can do better using Theorem 4.4. We construct in  $O(n + 1/\varepsilon^{3d})$  time a decomposition  $\Xi$  of  $\mathbb{R}^d$  into  $O(1/\varepsilon^{2d})$  Tarski cells along with two functions  $f_{\Delta}, f'_{\Delta}$  for each  $\Delta \in \Xi$  such that  $\{f_{\Delta}, f'_{\Delta}\}$  is an  $(\varepsilon/2)$ -kernel of  $\mathcal{F}$  within  $\Delta$ . For each cell  $\Delta \in \Xi$ , we compute  $w^*_{\Delta} = \min_{x \in \Delta} |f_{\Delta}(x) - f'_{\Delta}(x)|$ , and then compute  $\bar{w} = \min_{\Delta} w^*_{\Delta}$  as well as a point  $x^* \in \mathbb{R}^d$  that realizes  $\bar{w}$ . We return the smallest spherical shell centered at  $x^*$  that contains  $P$ . Note that  $w^* \geq \bar{w} \geq (1 - \varepsilon/2)\mathfrak{E}_{\mathcal{F}}(x^*)$ . Therefore

$$\mathfrak{E}_{\mathcal{F}}(x^*) \leq \frac{1}{1 - \varepsilon/2} \bar{w} \leq (1 + \varepsilon)w^*.$$

Hence, we obtain the following.

**Theorem 6.8** *Given a set  $P$  of  $n$  points in  $\mathbb{R}^d$ , and a parameter  $\varepsilon > 0$ , we can find in  $O(n + 1/\varepsilon^{3d})$  time a spherical shell containing  $P$  whose width is at most  $(1 + \varepsilon)w^*(P)$ . We can also compute within the same time bound a subset  $Q \subseteq P$  of size  $O(1/\varepsilon^d)$  so that for any  $x \in \mathbb{R}^d$ ,  $w(x, Q) \geq (1 - \varepsilon)w(x, P)$ .*



**Figure 6.** Parametrization of a line  $\ell$  in  $\mathbb{R}^3$  and its distance from a point  $\xi$ ; the small hollow circle on  $\ell$  is the point closest to  $\xi$ .

#### 6.4 Minimum-width cylindrical shell

Let  $P = \{p_1, \dots, p_n\}$  be a set of  $n$  points in  $\mathbb{R}^d$ , and a parameter  $\varepsilon > 0$ . Let  $w^* = w^*(P)$  denote the width of the thinnest cylindrical shell, the region lying between two co-axial cylinders, containing  $P$ . Let  $d(\ell, p)$  denote the distance between a point  $p \in \mathbb{R}^d$  and a line  $\ell \subset \mathbb{R}^d$ . If we fix a line  $\ell$ , then the width of the thinnest cylindrical shell with axis  $\ell$  and containing  $P$  is  $w(\ell, P) = \max_{p \in P} d(\ell, p) - \min_{p \in P} d(\ell, p)$ . A line  $\ell \in \mathbb{R}^d$  not parallel to the hyperplane  $x_d = 0$  can be represented by a  $(2d - 2)$ -tuple  $(x_1, \dots, x_{2d-2}) \in \mathbb{R}^{2d-2}$ :

$$\ell = \{p + tq \mid t \in \mathbb{R}\},$$

where  $p = (x_1, \dots, x_{d-1}, 0)$  is the intersection point of  $\ell$  with the hyperplane  $x_d = 0$  and  $q = (x_d, \dots, x_{2d-2}, 1)$  is the orientation of  $\ell$  (i.e.,  $q$  is the intersection point of the hyperplane  $x_d = 1$  with the line parallel to  $\ell$  and passing through the origin). The lines parallel to the hyperplane  $x_d = 0$  can be handled separately by a simpler algorithm. The distance between  $\ell$  and a point  $\xi$  is the same as the distance of the line  $\ell' = \{(p - \xi) + tq \mid t \in \mathbb{R}\}$  from the origin; see Figure 6. The point  $y$  on  $\ell$  closest to the origin satisfies  $y = (p - \xi) + tq$  for some  $t$ , and at the same time  $\langle y, q \rangle = 0$ , which implies that

$$d(\ell, \xi) = \|y\| = \left\| (p - \xi) - \frac{\langle p - \xi, q \rangle q}{\|q\|^2} \right\|,$$

Define  $f_i(\ell) = d(\ell, p_i)$ , and set  $\mathcal{F} = \{f_i \mid p_i \in P\}$ . Then  $w^* = \min_{x \in \mathbb{R}^{2d-2}} \mathfrak{E}_{\mathcal{F}}(x)$ . (We assume for simplicity that the axis of the optimal shell is not parallel to the hyperplane  $x_d = 0$ .) Let  $f'_i(x) = \|q\|^2 \cdot f_i(x)$ , and set  $\mathcal{F}' = \{f'_1, \dots, f'_n\}$ . By Lemma 2.1, it suffices to compute an  $\varepsilon$ -kernel of  $\mathcal{F}'$ . Define  $g_i = f'_i(x)^2$ , and let  $\mathcal{G} = \{g_1, \dots, g_n\}$ . Then  $g_i$  is a  $(2d - 2)$ -variate polynomial and has  $O(d^2)$  monomials. Therefore  $\mathcal{G}$  admits a linearization of dimension  $O(d^2)$ . Now, proceeding as in the case of spherical shell but using Corollary 4.8 and Theorem 4.9, we can compute in  $O(n + 1/\varepsilon^{O(d^2)})$  time a set  $Q \subseteq P$  of  $1/\varepsilon^{O(d^2)}$  points so that for any line  $\ell$ ,  $w(\ell, P) \geq w(\ell, Q) \geq (1 - \varepsilon)w(\ell, P)$  as well as a cylindrical shell of width at most  $(1 + \varepsilon)w^*(P)$  that contains  $P$ . Hence, we conclude the following.

**Theorem 6.9** *Given a set  $P$  of  $n$  points in  $\mathbb{R}^d$  and a parameter  $\varepsilon > 0$ , we can compute in  $O(n + 1/\varepsilon^{O(d^2)})$  time a cylindrical shell containing  $P$  whose width is at most  $(1 + \varepsilon)w^*(P)$ . We can also*

compute within the same time bound a subset  $Q \subseteq P$  of size  $O(1/\varepsilon^{O(d^2)})$  so that for any line  $\ell$  in  $\mathbb{R}^d$ , we have  $w(\ell, Q) \geq (1 - \varepsilon)w(\ell, P)$ .

## 7 Conclusions

In this paper, we have presented a general technique for computing extent measures approximately. The new technique shows that for many extent measures  $\mu$ , one can compute in time  $O(n + 1/\varepsilon^{O(1)})$  a subset  $Q$  (called an  $\varepsilon$ -kernel) of size  $1/\varepsilon^{O(1)}$  and then simply compute  $\mu(Q)$ . Such a subset  $Q$  is computed by combining convex-approximation techniques with duality and linearization techniques. Specific applications of our technique include near-linear approximation algorithms for computing minimum-width spherical and cylindrical shells, a general technique for approximating faithful measures of stationary as well as moving points. Interestingly enough, the dynamization and streaming techniques presented in Section 5 are generic and seem to apply without too many additional assumptions whenever a small  $\varepsilon$ -kernel exists. We believe that there are numerous other applications of our technique.

To some extent, our algorithm is the ultimate approximation algorithm for such problems: It has linear dependency on  $n$ , and a polynomial dependency on  $1/\varepsilon$ . The existence of such a general (and fast) approximation algorithm is quite surprising. Subsequent to our work, several geometric approximation algorithms have been developed that compute a subset that shares some properties of an  $\varepsilon$ -kernel for specific problems. The term *core-set* is now commonly used to refer to such a subset [7, 14, 15, 25, 27, 28].

We conclude by mentioning a few open problems and recent developments in this area.

- (i) Our algorithms compute an  $\varepsilon$ -kernel whose size is exponential in  $d$ . Recently a few algorithms have been proposed that compute a core-set of size  $(d/\varepsilon)^{O(1)}$  for specific problems such as the smallest enclosing sphere or ellipsoid [14, 15, 27, 28]. However it is not clear whether these algorithms can be extended to a more general setting.
- (ii) A possible direction for future research is to investigate how practical is this technique, and to improve/simplify it further. In particular, it seems that faster algorithms should exist for the problems of approximating the diameter and width of a point set.
- (iii) Recently, Agarwal *et al.* [7] used the  $\varepsilon$ -kernel technique for computing  $k$  congruent cylinders of the minimum radius that contain a point set in  $\mathbb{R}^d$ . Whether similar techniques can be developed for other projective-clustering problems in high dimensions remains an open problem.
- (iv) Another interesting direction for further research is to extend this technique to handle outliers. Some progress in this direction is recently made in [25]. It would also be interesting to develop algorithms for shape-fitting when the quality of the fit is measured by the sum of squares of distances (instead of the maximum distance as in this paper). There are efficient algorithms if the shape is a linear subspace but little seems to be known for other shapes.

## Acknowledgments

The authors thank Imre Barany, Jeff Erickson, Cecilia Magdalena Procopiuc, and Micha Sharir for helpful discussions, and the two referees for many useful comments.

## References

- [1] P. K. Agarwal, L. Arge, and J. Erickson, Indexing moving points, *J. Computer and System Sciences*, 66 (2003), 207–243.
- [2] P. K. Agarwal, B. Aronov, S. Har-Peled, and M. Sharir, Approximation and exact algorithms for minimum-width annuli and shells, *Discrete Comput. Geom.*, 24 (2000), 687–705.
- [3] P. K. Agarwal, B. Aronov, and M. Sharir, Exact and approximation algorithms for minimum-width cylindrical shells, *Discrete Comput. Geom.*, 26 (2001), 307–320.
- [4] P. K. Agarwal, L. J. Guibas, J. Hershberger, and E. Veach, Maintaining the extent of a moving point set, *Discrete Comput. Geom.*, 26 (2001), 353–374.
- [5] P. K. Agarwal and J. Matoušek, On range searching with semialgebraic sets, *Discrete Comput. Geom.*, 11 (1994), 393–418.
- [6] P. K. Agarwal and J. Matoušek, Discretizing the space of directions, in preparation.
- [7] P. K. Agarwal, C. M. Procopiuc, and K. Varadarajan, Approximation algorithms for  $k$ -line center, *Proc. 10th Annu. Sympos. European Algorithms*, 2002, pp. 54–63.
- [8] P. K. Agarwal and M. Sharir, Efficient algorithms for geometric optimization, *ACM Comput. Surv.*, 30 (1998), 412–458.
- [9] P. K. Agarwal and M. Sharir, Arrangements and their applications, in: *Handbook of Computational Geometry* (J.-R. Sack and J. Urrutia, eds.), Elsevier Science Publishers B.V. North-Holland, Amsterdam, 2000, pp. 49–119.
- [10] G. Barequet and S. Har-Peled, Efficiently approximating the minimum-volume bounding box of a point set in three dimensions, *J. Algorithms*, 38 (2001), 91–109.
- [11] J. Basch, L. J. Guibas, and J. Hershberger, Data structures for mobile data, *J. Algorithms*, 31 (1999), 1–28.
- [12] J. L. Bentley and J. B. Saxe, Decomposable searching problems I: Static-to-dynamic transformation, *J. Algorithms*, 1 (1980), 301–358.
- [13] E. M. Bronshteyn and L. D. Ivanov, The approximation of convex sets by polyhedra, *Siberian Math. J.*, 16 (1976), 852–853.

- [14] M. Bădoiu and K. Clarkson, Smaller core-sets for balls, *Proc. 14th Annu. ACM-SIAM Sympos. Discrete Algorithms*, 2003, pp. 801–802.
- [15] M. Bădoiu, S. Har-Peled, and P. Indyk, Approximate clustering via core-sets, *Proc. 34th Annu. ACM Sympos. Theory Comput.*, 2002, pp. 250–257.
- [16] T. M. Chan, Approximating the diameter, width, smallest enclosing cylinder and minimum-width annulus, *Internat. J. Comput. Geom. Appl.*, 12 (2002), 67–85.
- [17] B. Chazelle and J. Matoušek, On linear-time deterministic algorithms for optimization problems in fixed dimension, *J. Algorithms*, 21 (1996), 579–597.
- [18] Y. Chen, G. Dong, J. Han, B. W. Wah, and J. Wang, Multi-dimensional regression analysis of time-series data streams, *Proc. 28th Intl. Conf. Very Large Data Bases*, 2002, pp. 323–334.
- [19] R. M. Dudley, Metric entropy of some classes of sets with differentiable boundaries, *J. Approx. Theory*, 10 (1974), 227–236.
- [20] C. A. Duncan, M. T. Goodrich, and E. A. Ramos, Efficient approximation and optimization algorithms for computational metrology, *Proc. 8th ACM-SIAM Sympos. Discrete Algorithms*, 1997, pp. 121–130.
- [21] B. Gärtner, A subexponential algorithm for abstract optimization problems, *SIAM J. Comput.*, 24 (1995), 1018–1035.
- [22] P. Gritzmann and V. Klee, Inner and outer  $j$ -radii of convex bodies in finite-dimensional normed spaces, *Discrete Comput. Geom.*, 7 (1992), 255–280.
- [23] S. Guha, N. Koudas, and K. Shim, Data-streams and histograms, *Proc. 33rd Annu. ACM Sympos. Theory Comput.*, 2001, pp. 471–475.
- [24] S. Guha, N. Mishra, R. Motwani, and L. O’Callaghan, Clustering data streams, *Proc. 41th Annu. IEEE Sympos. Found. Comput. Sci.*, 2000, pp. 359–366.
- [25] S. Har-Peled and Y. Wang, Shape fitting with outliers, *Proc. 19th Annu. ACM Sympos. Comput. Geom.*, 2003, pp. 29–38.
- [26] F. Korn, S. Muthukrishnan, and D. Srivastava, Reverse nearest neighbour aggregates over data streams, *Proc. 28th Intl. Conf. Very Large Data Bases*, 2002, pp. 814–825.
- [27] P. Kumar, J. Mitchell, and E. Yildirim, Computing core-sets and approximate smallest enclosing hyperspheres in high dimensions, *Proc. 5th Workshop Algorithm Eng. Exper.*, 2003, pp. 45–55.
- [28] P. Kumar and E. Yildirim, Approximating minimum volume enclosing ellipsoids using core sets, Unpublished manuscript, 2003.

- [29] J. I. Munro and M. S. Paterson, Selection and sorting with limited storage, *Theoretical Computer Science*, 12 (1980), 315–323.
- [30] J. O’Rourke, Finding minimal enclosing boxes, *Internat. J. Comput. Inform. Sci.*, 14 (1985), 183–199.
- [31] C. M. Procopiuc, P. K. Agarwal, and S. Har-Peled, STAR-Tree: An efficient self-adjusting index for moving objects, *Proc. 4th Workshop Algorithm Eng. Exper.* (D. M. Mount and C. Stein, eds.), *Lect. Notes in Comp. Sci.*, Vol. 2409, 2002, pp. 178–193.
- [32] S. Saltenis, C. S. Jensen, S. Leutenegger, and M. Lopez, Indexing the positions of continuously moving objects, *ACM SIGMOD Int. Conf. on Manag. of Data*, 2000, pp. 331–342.
- [33] A. C. Yao and F. F. Yao, A general approach to  $D$ -dimensional geometric queries, *Proc. 17th Annu. ACM Sympos. Theory Comput.*, 1985, pp. 163–168.
- [34] Y. Zhou and S. Suri, Algorithms for a minimum volume enclosing simplex in three dimensions, *SIAM J. Comput.*, 31 (2002), 1339–1357.

## Appendix: Summary of Notations

$\mathbb{R}^d$	$d$ -dimensional Euclidean space
$\mathbb{S}^{d-1}$	Unit sphere in $\mathbb{R}^d$
$\mathbb{P}$	$(d-1)$ -dimensional hyperplane $x_d = 1$
$\mathbb{C}$	$d$ -dimensional unit hypercube $[-1, +1]^d$
$\mathcal{U}_{\mathcal{F}}$	Upper envelope of $\mathcal{F}$
$\mathcal{L}_{\mathcal{F}}$	Lower envelope of $\mathcal{F}$
$\mathcal{E}_{\mathcal{F}}$	Extent of $\mathcal{F}$
$\mathcal{A}(\mathcal{J})$	Arrangement of $\mathcal{J}$
$\mathcal{CH}(S)$	Convex hull of $S$
$\phi(v)$	$v/\ v\ , v \in \mathbb{R}^d$
$\tilde{u}$	$(u, 1) \in \mathbb{P}, u \in \mathbb{R}^{d-1}$
$u^*$	$\phi(\tilde{u}) \in \mathbb{S}^{d-1}, u \in \mathbb{R}^{d-1}$ .
$\bar{\omega}(x, P)$	$\max_{p \in P} \langle x, p \rangle - \min_{p \in P} \langle x, p \rangle, x \in \mathbb{R}^d, P \subseteq \mathbb{R}^d$
$\omega(u, P)$	$\bar{\omega}(\tilde{u}, P), u \in \mathbb{R}^{d-1}, P \subseteq \mathbb{R}^d$

**Table 1.** Summary of notations used in the paper.