Projective Clustering in High Dimensions using Core-Sets

Sariel Har-Peled* Kasturi R. Varadarajan†

January 31, 2003

Abstract

In this paper, we show that there exists a small core-set for the problem of computing the "smallest" radius k-flat for a given point-set in $\mathbb{R}^d$. The size of the core-set is *dimension independent*. Such small core-sets yield immediate efficient algorithms for finding the $(1+\varepsilon)$-approximate smallest radius k-flat for the points in $dn^{O(k^6/\varepsilon^5 \log(1/\varepsilon))}$ time. Furthermore, we can use it to $(1+\varepsilon)$-approximate the the smallest radius k-flat for a prespecified fraction of the given points, in the same running time. Our algorithm can also be used for computing the min-max such coverage of the point-set by j flats, each one of them of dimension k.

No previous efficient approximation algorithms were known for those problems in high-dimensions, when $k > 1$ or $j > 1$.

1 Introduction

Clustering is one of the central problems in computer-science. It is related to unsupervised learning, classification, databases, spatial range-searching, data-mining, etc. As such, it received a lot of attention in computer-science in the last twenty years. There is a large literature on this topic with numerous variants, see [DHS01, Hoc96]. In the projective clustering problem, one wants to find a covering of points by j k-flats. Of course, such a covering might not exist, and we are satisfied with the best fit by j k-flats. For example, for $k = 1$, we are interested in covering a set of points in $\mathbb{R}^d$ by j cylinders of equal radius, such that the radius is as small as possible. This problem is NP-Complete [MT82] even for $k = 1$.

Projective clustering has numerous applications. See [AP00] for extensive bibliography on this problem. For example, such a projective clustering implies that the point-set can be indexed as j point-sets where each of them is only k-dimensional. This is a considerable saving when the dimension is very high, as most efficient indexing structures have exponential dependency on the dimension. Thus, finding such a cover might result in a substantial performance improvements for various database applications.

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

† Department of Computer Science, University of Iowa, kvaradar@cs.uiowa.edu

Another, related problem, is the following fundamental data-mining problem: Given a set P of points in $\mathbb{R}^d$, and parameters $r \geq 0, k > 0$ find a k flat that is within a distance r of as many points of P as possible.

Such clustering problems become harder as the dimension increases, and unfortunately, such applications of projective clustering are usually set in high dimensions. The intractability arises because most algorithms for such problems tend to have running time exponential in the dimension. It is thus interesting to come up with fast algorithms that have only polynomial dependency on the dimension. A popular approach for doing so is to use dimension reduction techniques (especially the Johnson-Lindenstrauss lemma [JL84] and its various extensions, see [Ind01]). However, it is not clear that the projective clustering problem is dimension reducible.

Given a set P of n points in d dimensions (d might be as large as n), we present efficient algorithms for the following problems:

- **Approximate k-flat**: We compute in $dn^{O(k^6/\varepsilon^5 \log(1/\varepsilon))}$ time a k-flat that is within a distance of $(1+\varepsilon)\mathcal{RD}_k(P)$ from each point in P, where $\mathcal{RD}_k(P)$ is the minimum over all k-flats $\mathcal{F}$ of $\max_{p \in P} \mathrm{dist}(p, \mathcal{F})$.

  The only previous result of this type is the result of Bădoiu *et al.* [BHI02], which solves the problem in $n^{O(1/\varepsilon^2 \log 1/\varepsilon)}$ time, for $k = 1$. Although their algorithm is faster than ours in this case, it is considerably more involved. For example, it makes use of convex programming. While our algorithms also use convex programming, this is not essential, and we can get algorithms without convex programming with a modest increase in running time so that they are strongly polynomial. Also, it is not clear if the algorithm of Bădoiu *et al.* can be extended to the case of $k > 1$ or the problem with outliers.

- **Approximate k-flat with outliers**: Given an integer $\mu \leq n$ and a parameter $\varepsilon > 0$, we compute in $dn^{O(k^6/\varepsilon^5 \log(1/\varepsilon))}$ time a k-flat that is within a distance of $\leq (1+\varepsilon)\mathcal{RD}_k(P,\mu)$ from at least μ points in P, where $\mathcal{RD}_k(P,\mu)$ is the minimum, over all k-flats $\mathcal{F}$, of the distance of the μ-th closest point in P from $\mathcal{F}$.

- **j approximate k-flat**: Given a parameter $\varepsilon > 0$, and j, we compute in $dn^{O(k^6 j/\varepsilon^5 \log(1/\varepsilon))}$ time j k-flats such that each point in P is within a distance $(1+\varepsilon)\mathcal{RD}_k^j(P)$ from at least one of the flats, where $\mathcal{RD}_k^j(P)$ is the minimum over all sets of j k-flats $\mathcal{F}_1, \ldots, \mathcal{F}_j$ of $\max_{p \in P} \min_{i \leq j} \mathrm{dist}(p, \mathcal{F}_i)$.

- **j approximate k-flat with outliers**: Same as above with outliers.

Note that these results are interesting only when the dimension $d = \Omega(k^6/\varepsilon^6)$, as all those problems can be solved directly using direct enumerations of all possible solutions. Nevertheless, the authors believe that the ability to solve those problems in time which depends only polynomially on the dimension is quite surprising. To our knowledge, this the first efficient $(1+\varepsilon)$-approximation algorithms in high-dimensions for those problems for $k > 1$ or $j > 1$.

Our results rely on a scheme, obtained by extending a new technique of Bădoiu *et al.* [BHI02], that extracts a small subset of points that "represents" this point-set ε-well as far as those projective clustering problems are concerned. The surprising property of those sets is that their size is *independent* of the dimension. The existence of such core-sets for various approximation problems was known before, but their size depended polynomially or exponentially on the dimension [MOP01, ADPR00, Har01, HV01, IT00]. (It must be pointed out, however, that as in [BHI02] the core sets in this paper satisfy much weaker requirements.)

The paper is organized as follows: In Section 2, we introduce our basic technique. In Section 3 we prove the existence of small core-set for the k-flat problem. In Section 4 we present our algorithms. Concluding remarks are presented at Section 5.

2 Preliminaries

Let $\mathrm{dist}(p, q)$ denote the Euclidean distance between two points $p, q \in \mathbb{R}^d$. We define the distance $\mathrm{dist}(P, Q)$ between two point sets $P, Q \subseteq \mathbb{R}^d$ to be $\min_{p \in P, q \in Q} \mathrm{dist}(p, q)$.

Definition 2.1 A *k-flat* is an affine subspace of dimension k. For a point-set $P \subseteq \mathbb{R}^d$, and a k-flat $\mathcal{F}$, the *radius* of $\mathcal{F}$ is

$$\mathcal{RD}(P, \mathcal{F}) = \max_{p \in P} \mathrm{dist}(p, \mathcal{F}).$$

The *k-th outer radius* of P is

$$\mathcal{RD}_k(P) = \min_{\mathcal{F} \text{is a } k\text{-flat}} \mathcal{RD}(P, \mathcal{F}).$$

The quantity $\mathcal{RD}_0(P)$ is the radius of the smallest enclosing ball of P and $2\mathcal{RD}_{d-1}(P)$ is the width of P. See [GK92] for related results. In particular, just computing the width of a simplex in d dimension is NP-Hard [GK93, GK94].

For a set $C \subseteq \mathbb{R}^d$, let $\Delta(C) = max_{p,q \in C} \mathrm{dist}(p, q)$ denote the diameter of C.

Definition 2.2 For an affine subspace $\mathcal{F}$ of $\mathbb{R}^d$, let $\mathcal{F}^\perp$ be the linear subspace orthogonal to $\mathcal{F}$, and let $\mathcal{TP}_\mathcal{F}$ be the linear mapping projecting $\mathbb{R}^d$ into $\mathcal{F}^\perp$. The mapping $\mathcal{TP}_\mathcal{F}$ maps $\mathcal{F}$ into a point in $\mathcal{F}^\perp$.

Definition 2.3 Let U be any set of points in $\mathbb{R}^d$, and $\varepsilon > 0$ be a parameter. We say that a subset V of points in $\mathcal{CH}(U)$ is an *ε-net* for U if for any flat $\mathcal{F}$ that intersects $\mathcal{CH}(U)$, there is a $v \in V$ such that $\mathrm{dist}(v, \mathcal{F}) \leq \frac{\varepsilon}{2}\mathcal{RD}(U, \mathcal{F})$.

Lemma 2.4 *If V is an ε-net for U and π is a projection from $\mathbb{R}^d$ to $\mathbb{R}^{d'}$, then $\pi(V)$ is an ε-net for $\pi(U)$.*

Proof: Indeed, let $\mathcal{F}_\pi$ be any k-flat that intersects $\mathcal{CH}(\pi(U))$. Let $\mathcal{F} = \pi^{-1}(\mathcal{F}_\pi)$ be the affine subspace of all points in the original set that are being mapped to $\mathcal{F}_\pi$ by π. Clearly, for any point $x \in \mathbb{R}^d$, we have $\mathrm{dist}(x, \mathcal{F}) = \mathrm{dist}(\pi(x), \mathcal{F}_\pi)$, and in particular, $\mathcal{RD}(U, \mathcal{F}) = \mathcal{RD}(\pi(U), \mathcal{F}_\pi)$. Thus, there exists a point $y \in V$, such that $\mathrm{dist}(y, \mathcal{F}) \leq (\varepsilon/2)\mathcal{RD}(U, \mathcal{F})$. In particular, $\pi(y) \in \pi(V)$, and $\mathrm{dist}(\pi(y), \mathcal{F}_\pi) \leq (\varepsilon/2)\mathcal{RD}(\pi(U), \mathcal{F}_\pi)$, namely $\pi(V)$ is an ε-net from $\pi(U)$. ■

We are given a small point-set U in high-dimensions, and we want to perform a search over $\mathcal{CH}(U)$. The problem is that $\mathcal{CH}(U)$ contains infinitely many points. In the following lemma, we generate a small grid $A(U)$ (i.e., ε-net) that "represents" $\mathcal{CH}(U)$ as far as our search for a good k-flat is involved.

Lemma 2.5 *Let U be a set of points in $\mathbb{R}^d$ and $\varepsilon > 0$ be a parameter. We can compute an ε-net $A(U)$ for U in $(|U|^{2.5}/\varepsilon)^{O(|U|)}$ time. The cardinality of $A(U)$ is $(|U|^{2.5}/\varepsilon)^{O(|U|)}$.*

Proof: Let H the $(M-1)$-dimensional affine subspace spanned by U (notice that $M \leq |U|$), and let $\mathcal{E} \subseteq H$ be an ellipsoid such that $\mathcal{E}/(M+1)^2 \subseteq \mathcal{CH}(U) \subseteq \mathcal{E}$, where $\mathcal{E}/(M+1)^2$ is the scaling down of $\mathcal{E}$ around its center by a factor of $1/(M+1)^2$. Such an ellipsoid exists (a stronger version of this statement is known as John theorem), and can be computed in polynomial time in $|U|$ [GLS88, Section 4.6]. Let $\mathcal{B}$ be the minimum bounding box of $\mathcal{E}$ which is parallel to the main axises of $\mathcal{E}$. We claim, that $\mathcal{B}/\sqrt{M}$ is contained inside $\mathcal{E}$. Indeed, there exists a linear transformation $\mathcal{T}$ that maps $\mathcal{E}$ to a unit ball $\mathcal{S}$. The point $\mathbf{q} = (1/\sqrt{M}, 1/\sqrt{M}, \ldots, 1/\sqrt{M})$ lies on the boundary of this sphere. Clear, $\mathcal{T}^{-1}(\mathbf{q})$ is a corner of $\mathcal{B}/\sqrt{M}$, and is on the boundary of $\mathcal{E}$. In particular,

$$\Delta(\mathcal{B}) = \sqrt{M}\Delta(\mathcal{B}/\sqrt{M}) \leq \sqrt{M}\Delta(\mathcal{E}) \leq \sqrt{M}(M+1)^2\Delta(\mathcal{E}/(M+1)^2) \leq \sqrt{M}(M+1)^2\Delta(U).$$

For any flat $\mathcal{F}$, the same arguments works for the projection of those entities by $\mathcal{TP}_{\mathcal{F}}$. In particular,

$$\Delta(\mathcal{TP}_{\mathcal{F}}(\mathcal{B})) \leq \sqrt{M}(M+1)^2\Delta(\mathcal{TP}_{\mathcal{F}}(U)) \leq 2\sqrt{M}(M+1)^2\mathcal{RD}(U,\mathcal{F}).$$

Next, we partition $\mathcal{B}$ into a grid, where each grid cell is a translated copy of $\mathcal{B}_\varepsilon = (\varepsilon/2)\mathcal{B}/(2\sqrt{M}(M+1)^2)$. This grid has $V = (M^{2.5}/\varepsilon)^{O(M)}$ vertices, and let $A(U)$ denote this set of vertices.

Let $\mathcal{F}$ be any flat intersecting $\mathcal{CH}(U)$. We claim that one of the points in $A(U)$ is in distance $\leq \frac{\varepsilon}{2}\mathcal{RD}(U,\mathcal{F})$ from $\mathcal{F}$. Indeed, let z be any point in $\mathcal{CH}(U) \cap F$. Let $\mathcal{B}''_\varepsilon$ be the grid cell containing z, and let v be one of its vertices. Clearly,

$$\begin{aligned}
\mathrm{dist}(v,\mathcal{F}) &\leq \|\mathcal{TP}_{\mathcal{F}}(v)\mathcal{TP}_{\mathcal{F}}(z)\| \leq \Delta(\mathcal{TP}_{\mathcal{F}}(\mathcal{B}''_\varepsilon)) \\
&= \frac{\varepsilon}{2} \cdot \frac{1}{2\sqrt{M}(M+1)^2}\Delta(\mathcal{TP}_{\mathcal{F}}(\mathcal{B})) \leq \frac{\varepsilon}{2}\mathcal{RD}(U,\mathcal{F}),
\end{aligned}$$

which establishes our claim. ■

Remark 2.6 We can guarantee that the set $A(U)$ lies completely inside $\mathcal{CH}(U)$ by modifying the above algorithm slightly. We compute the grid as before. For each grid cell C, we compute a point $x_C \in C \cap \mathcal{CH}(U)$ if such a point exists. Arguing as above, it is easy to verify that the resulting set of points is indeed an ε-net. The size of the ε-net remains the same as above. To compute x_C, we compute a triangulation of $\mathcal{CH}(U)$ in $O(|U|^{|M|})$ time, and check if each feature of the triangulation intersects C. It is easy to verify that doing this for each grid cell C does not increase the running time by too much. In fact the overall running time is the same as in Lemma 2.5. We also point out that the requirement that $A(U)$ lie in $\mathcal{CH}(U)$ is only needed to make our proofs simpler.

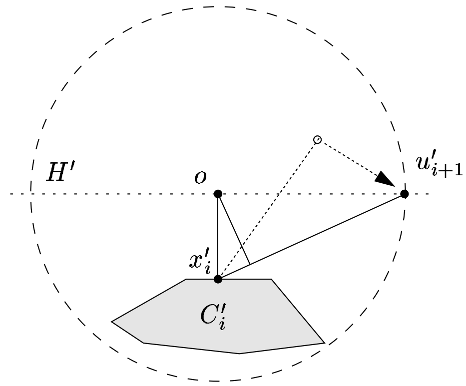

Figure 1: The distance between o and x'_i is maximized when u'_{i+1} lies on H', and ou'_{i+1} is as long as possible.

Remark 2.7 In Lemma 2.5 we had used convex-programming techniques to find the "good" ellipsoid $\mathcal{E}$. If we insist on simplicity and the constant induced by the dimension is not that crucial, we can just use the simple algorithm of Barequet and Har-Peled [BH01]. We will get an ε-net for U of size $|U|^{O(|U|^2)}/\varepsilon^{O(|U|)}$ in $|U|^{O(|U|^2)}/\varepsilon^{O(|U|)}$ time. This will make all the algorithms in this paper strongly polynomial.

The following lemma, shows that we can always find a small set of points U, such that a near-optimal flat must cross the convex hull of this point-set. In particular, together with the usage of ε-nets, this implies that we can find a polynomial number of candidate points one of which must be on a near-optimal k-flat. To do so, we will enumerate all possible such subsets U, and for each such subset we will generate a set of possible candidate points by computing an ε-net of U.

Lemma 2.8 *Let P be a point-set in $\mathbb{R}^d$, $\mathcal{F}$ be a k-flat that intersects $\mathcal{CH}(P)$, $\mathcal{R}_{\mathcal{F}} = \mathcal{RD}(P, \mathcal{F})$, and let $0 < \varepsilon < 1$ be a parameter. Then there exists a subset U of P of size $O(1/\varepsilon^2 \log(1/\varepsilon))$ such that* $\mathrm{dist}(\mathcal{CH}(U), \mathcal{F}) \leq (\varepsilon/2)\mathcal{R}_{\mathcal{F}}$.

Furthermore, if $A(U)$ is an ε-net for U, then $\mathrm{dist}(\mathcal{F}, A(U)) \leq \varepsilon\mathcal{R}_{\mathcal{F}}$.

Proof: We construct U iteratively. Let u_1 be any point of P. Let $C_{i-1} = \mathcal{CH}(U_{i-1})$, where $U_{i-1} = \{u_1, \ldots, u_{i-1}\}$. If $\mathrm{dist}(C_{i-1}, \mathcal{F}) \leq (\varepsilon/2)\mathcal{R}_{\mathcal{F}}$ we had found $U = U_{i-1}$. Otherwise, let x_{i-1} be the closest point on C_{i-1} to $\mathcal{F}$.

Let $\mathcal{F}^{\perp}$ be the linear subspace orthogonal to $\mathcal{F}$, and let $\mathcal{TP}_{\mathcal{F}}$ be the linear mapping projecting $\mathbb{R}^d$ into $\mathcal{F}^{\perp}$. The mapping $\mathcal{TP}_{\mathcal{F}}$ maps $\mathcal{F}$ into a point o in $\mathcal{F}^{\perp}$, and P is being mapped into a set P' which is contained inside a ball B' of radius $\mathcal{R}_{\mathcal{F}}$ centered at o.

Clearly, finding a point in C_i in distance $\leq (\varepsilon/2)\mathcal{R}_{\mathcal{F}}$ from $\mathcal{F}$, is equivalent to finding a point $x \in C'_i = \mathcal{TP}_{\mathcal{F}}(C_i)$ in distance $\leq (\varepsilon/2)\mathcal{R}_{\mathcal{F}}$ from o. Furthermore, o is contained inside the $\mathcal{CH}(P')$, the distance between C'_i and o is achieved by $x'_i = \mathcal{TP}_{\mathcal{F}}(x_i)$, and $\|x'_i o\| \geq (\varepsilon/2)\mathcal{R}_{\mathcal{F}}$. In particular, let H' be a hyperplane in $\mathcal{F}^{\perp}$ passing through o and perpendicular to $x'_i o$. There must be a point u'_{i+1} of P' that is on the other side of H' than x_i. Let u_{i+1} be the corresponding point in P.

We next bound the distance between o and $x'_i u'_{i+1}$. Clearly, this distance is maximized when when $x'_i o \perp ou'_{i+1}$, and $x'_i o$ is as long as possible, see Figure 2. Thus, we can assume that $x'_i o \perp ou'_{i+1}$ and $\|x'_i o\| = \mathcal{R}_{\mathcal{F}}$. It follows,

$$\left\|x'_i u'_{i+1}\right\| \geq \sqrt{\|x'_i o\|^2 + \left\|ou'_{i+1}\right\|^2} \geq \sqrt{((\varepsilon/2)^2+1)\mathcal{R}_{\mathcal{F}}^2} \geq (1+\varepsilon^2/8)\mathcal{R}_{\mathcal{F}},$$

as $\|x'_i o\| = \mathrm{dist}(C_i, o) \geq (\varepsilon/2)\mathcal{R}_{\mathcal{F}}$. By similarity of triangles, we have:

$$\begin{aligned}\mathrm{dist}(\mathcal{F}, C_{i+1}) &\leq \mathrm{dist}(o, x'_i u'_{i+1}) = \|ox'_i\| \cdot \frac{\left\|ou'_{i+1}\right\|}{\left\|x'_i u'_{i+1}\right\|} \leq \|ox'_i\| \cdot \frac{\mathcal{R}_{\mathcal{F}}}{(1+\varepsilon^2/8)\mathcal{R}_{\mathcal{F}}} \\ &\leq \left(1 - \frac{\varepsilon^2}{16}\right)\|ox'_i\| \leq \left(1 - \frac{\varepsilon^2}{16}\right)\mathrm{dist}(\mathcal{F}, x_i) \leq \left(1 - \frac{\varepsilon^2}{16}\right)\mathrm{dist}(\mathcal{F}, C_i)\end{aligned}$$

since $\varepsilon < 1$.

Overall, as $\mathrm{dist}(\mathcal{F}, x_1) \leq \mathcal{R}_{\mathcal{F}}$, it follows that $\mathrm{dist}(\mathcal{F}, C_M) \leq (\varepsilon/2)\mathcal{R}_{\mathcal{F}}$, for $M = O(1/\varepsilon^2 \log(1/\varepsilon))$. Let $U = U_M$ be the resulting point-set.

Let $\mathcal{F}'$ be a translation of $\mathcal{F}$ to the point of $\mathcal{CH}(U)$ closest to $\mathcal{F}$. It is easy to verify that $\mathrm{dist}(\mathcal{F}, \mathcal{F}') \leq \frac{\varepsilon}{2}\mathcal{R}_{\mathcal{F}}$. If $A(U)$ is an ε-net for U, $\mathrm{dist}(\mathcal{F}', A(U)) \leq \frac{\varepsilon}{2}\mathcal{R}_{\mathcal{F}}$. This implies that $\mathrm{dist}(\mathcal{F}, A(U)) \leq \mathrm{dist}(\mathcal{F}, \mathcal{F}') + \mathrm{dist}(\mathcal{F}', A(U)) \leq \varepsilon\mathcal{R}_{\mathcal{F}}$. ■

3 Good Core Sets Exist

Lemma 2.8 implies that we can essentially assume that we have at our hand a "small" list (i.e., $O(n^{1/\varepsilon^2 \log 1/\varepsilon})$) of points, such that one of those points lies on the optimal k-flat $\mathcal{F}_{opt}$. If we could just find additional k points that lie on $\mathcal{F}_{opt}$, then we would be done, as we could just inspect all possible k-flats induced by such a point-set, and take the best one. What we actually show is that we can find k such points on some near-optimal k-flat. To do so, we first solve the problem for $k = 1$, namely the case of a cylinder.

Lemma 3.1 *Let P be a set of points in $\mathbb{R}^d$, $\mathcal{F}$ be the k-flat that minimizes the maximum distance to the set P, namely $\mathcal{R}_{\mathcal{F}} = \mathcal{RD}(P, \mathcal{F}) = \mathcal{RD}_k(P)$, and let $\varepsilon > 0$ be a parameter. There is a subset $Q \subseteq P$ of $O(1/\varepsilon^2 \log(1/\varepsilon))$ points and a k-flat $\mathcal{G}$ that intersects $\mathcal{CH}(Q)$ such that $\mathcal{RD}(P, \mathcal{G}) \leq (1+\varepsilon)\mathcal{R}_{\mathcal{F}}$. Furthermore, we can assume that $\mathcal{G}$ passes through a point in an $\varepsilon/$-net of Q*

Proof: From the optimality of $\mathcal{F}$, it follows that $\mathcal{F}$ intersects $\mathcal{CH}(P)$. From Lemma 2.8, we get a set $Q \subseteq P$ of $O(1/\varepsilon^2 \log 1/\varepsilon)$ points such that $\mathcal{CH}(Q)$ contains a point s with $\mathrm{dist}(s, \mathcal{F}) \leq \varepsilon\mathcal{R}_{\mathcal{F}}$. We can also assume that s is in an ε-net of Q. Let $\mathcal{G}$ be the translate of $\mathcal{F}$ that passes through s. It is easy to see that $\mathcal{RD}(P, \mathcal{G}) \leq (1+\varepsilon)\mathcal{R}_{\mathcal{F}}$. ■

Proposition 3.2 *Let p be a point in $\mathbb{R}^d$ at distance x from the origin, and let ℓ_1 and ℓ_2 be two lines through the origin making an angle θ with each other. Then* $\mathrm{dist}(p, \ell_2) \leq \mathrm{dist}(p, \ell_1) + x \sin\theta$.

The following lemma, shows that we can find a core-set for a cylinder. Namely, we can find a small subset of points, so that it defines the axis of the approximate minimum radius cylinder. This is the critical step in our construction, as once we have such a core-set, we can

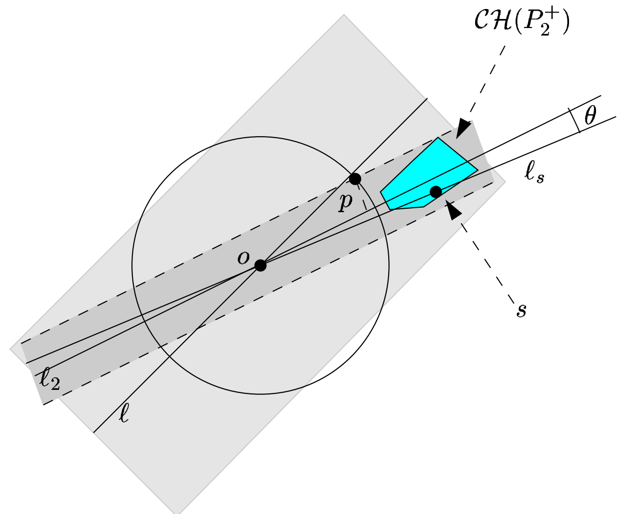

Figure 2: Proof of Lemma 3.3 for the short diameter case.

collapse the space along this line, and find the best $(k-1)$-dimensional flat in the collapsed space. Namely, applying this lemma repeatedly k times, will results in an ε-approximation to the optimal k-flat.

Lemma 3.3 *Let P be a set of points in $\mathbb{R}^d$, ℓ be a line through the origin o, $\mathcal{RD}_\ell = \mathcal{RD}(P, \ell)$, and $\varepsilon > 0$ be a parameter. There is a subset $Q \subseteq P$ of $O(1/\varepsilon^5 \log 1/\varepsilon)$ points such that the linear subspace spanned by Q contains a line ℓ' through o such that $\mathcal{RD}(P, \ell') \leq (1+\varepsilon)\mathcal{RD}_\ell$.*

Furthermore, the line ℓ' can be assumed to pass through o and a point s different from o that lies in an $(\varepsilon^{5/2}/16)$-net of U, where U is obtained from Q by inverting some subset of Q.

Proof: Let $\mu \in P$ be the point furthest from the origin o and $\Delta = \|o\mu\|$.

- **Short diameter - $\mathcal{RD}_\ell \geq \varepsilon\Delta/8$:** Let $P_1 \subseteq P$ be the set of points at distance at most $(1+\varepsilon)\mathcal{RD}_\ell$ from the origin, and $P_2 = P \setminus P_1$ be the complement set. Let ℓ_2 be the line through the origin that minimizes $\mathcal{RD}(P_2, \ell_2)$. (If P_2 is empty, we simply let ℓ' be the line through o and any point in P_1.) Clearly, $\mathcal{RD}(P_2, \ell_2) \leq \mathcal{RD}_\ell$. See Figure 2. Let h be the hyperplane orthogonal to ℓ_2. Notice that the projection of any point $p \in P_2$ onto h is at distance at most $\mathcal{RD}(P_2, \ell_2) \leq \mathcal{RD}_\ell$ from o. Since $\|op\| \geq (1+\varepsilon)\mathcal{RD}_\ell$, the projection of p onto ℓ_2 is at distance at least $\sqrt{(1+\varepsilon)^2\mathcal{RD}_\ell^2 - \mathcal{RD}_\ell^2} \geq \sqrt{\varepsilon}\mathcal{RD}_\ell$ from o. Let us call one of the half-spaces bounded by h the positive half-space.

  Let P_2^+ be the point set obtained from P_2 by replacing any point $p \in P_2$ in the negative half-space by its inversion $-p$ in the positive half-space. It is easy to verify that the optimality of ℓ_2 for P_2 implies that ℓ_2 intersects the convex hull of P_2^+ (if not, we can "tilt" ℓ_2 so that $\mathcal{RD}(P_2, \ell_2)$ decreases). Applying Lemma 2.8 to P_2^+ and ℓ_2, we obtain a subset $Q^+ \subseteq P_2^+$ of $O(1/\varepsilon^5 \log 1/\varepsilon)$ points such that $\mathcal{CH}(Q^+)$ contains a point s at distance at most $\frac{\varepsilon^{5/2}}{8}\mathcal{RD}_\ell$ from ℓ_2. Furthermore, s can be assumed to be one of the points in an $(\varepsilon^{5/2}/8)$-net of Q^+. Since s is in $\mathcal{CH}(P_2^+)$, its projection t onto ℓ_2 is at distance at least $\sqrt{\varepsilon}\mathcal{RD}_\ell$ from o. Let ℓ' be the line that passes through o and s. The line ℓ' lies in the linear subspace spanned by the points $Q \subseteq P_2$ corresponding to Q^+ together with the origin. The angle θ between ℓ' and ℓ_2 satisfies

$$\tan\theta = \frac{\|st\|}{\|ot\|} \leq \frac{\varepsilon^{5/2}\mathcal{RD}_\ell/8}{\sqrt{\varepsilon}\mathcal{RD}_\ell} \leq \frac{\varepsilon^2}{8}.$$

For any $p \in P_2$, we have

$$\begin{aligned}\mathrm{dist}(p,\ell') &\leq \mathrm{dist}(p,\ell_2) + \Delta\sin\theta \leq \mathrm{dist}(p,\ell_2) + \Delta\tan\theta \leq \mathrm{dist}(p,\ell_2) + \frac{8\mathcal{RD}_\ell}{\varepsilon}\tan\theta \\ &\leq \mathrm{dist}(p,\ell_2) + \frac{8\mathcal{RD}_\ell}{\varepsilon}\cdot\frac{\varepsilon^2}{8} \leq \mathrm{dist}(p,\ell_2) + \varepsilon\mathcal{RD}_\ell \leq (1+\varepsilon)\mathcal{RD}_\ell.\end{aligned}$$

For any $p \in P_1$, $\mathrm{dist}(p,\ell') \leq \mathrm{dist}(p,o) \leq (1+\varepsilon)\mathcal{RD}_\ell$.

- **Long diameter -** $\mathcal{RD}_\ell \leq \varepsilon\Delta/8$: Let $P_1 \subseteq P$ be the set of points at distance at most $\varepsilon\Delta/4$ from the origin, and $P_2 = P \setminus P_1$ be the complement set. Let ℓ_2 be the line through the origin that minimizes the distance to P_2. Clearly $\mathcal{RD}(P_2,\ell_2) \leq \mathcal{RD}_\ell$. We now argue that the angle α between ℓ_2 and ℓ satisfies $\sin\alpha \leq 2r/\Delta$.

  Let ℓ_μ be the line through o and μ, where μ, we remind the reader, is the point of P furthest away from o. Let α_ℓ be the angle between ℓ and ℓ_μ, α_2 be the angle between ℓ_2 and ℓ_μ. Clearly, $\alpha \leq \alpha_\ell + \alpha_2$. Since $\mathrm{dist}(\mu,\ell), \mathrm{dist}(\mu,\ell_2) \leq \mathcal{RD}_\ell$, we know that $\sin\alpha_\ell, \sin\alpha_2 \leq \mathcal{RD}_\ell/\Delta$. Thus $\sin\alpha \leq \sin(\alpha_\ell+\alpha_2) \leq \sin\alpha_\ell + \sin\alpha_2 \leq 2\mathcal{RD}_\ell/\Delta$. This means that for any point $p \in P_1$,

  $$\mathrm{dist}(p,\ell_2) \leq \mathrm{dist}(p,\ell) + \varepsilon\frac{\Delta}{4}\sin\alpha \leq (1+\varepsilon/2)\mathcal{RD}_\ell.$$

  We conclude that $\mathrm{dist}(p,\ell_2) \leq (1+\varepsilon/2)\mathcal{RD}_\ell$ for any $p \in P$.

  As in the first case, let h be the hyperplane orthogonal to ℓ_2. Notice that the projection of any point $p \in P_2$ onto h is at distance at most $\mathcal{RD}_\ell$ from o. Since $\|op\| \geq \varepsilon\Delta/4$, the projection of p onto ℓ_2 is at distance at least $\varepsilon\Delta/4 - \mathcal{RD}_\ell \geq \varepsilon\Delta/8$ from o. As before, we argue that ℓ_2 intersects $\mathcal{CH}(P_2^+)$. Applying Lemma 2.8 to P_2^+ and ℓ_2, we obtain a subset $Q^+ \subseteq P_2^+$ of $O(1/\varepsilon^4\log(1/\varepsilon))$ points such that $\mathcal{CH}(Q^+)$ contains a point s at distance at most $\varepsilon^2\mathcal{RD}_\ell/16$ from ℓ_2. Furthermore, we may assume that s is in an $(\varepsilon^2/16)$-net of Q^+. Since s is in $\mathcal{CH}(P_2^+)$, its projection t onto ℓ_2 is at distance at least $\varepsilon\Delta/8$ from o. Let ℓ' be the line that passes through o and s. The line ℓ' lies in the linear subspace spanned by the points $Q \in P_2$ corresponding to Q^+ and the origin. The angle θ between ℓ' and ℓ_2 satisfies

  $$\tan\theta = \frac{\mathrm{dist}(s,t)}{\mathrm{dist}(o,t)} \leq \frac{\varepsilon\mathcal{RD}_\ell}{2\Delta}.$$

  For any $p \in P$, we have

  $$\mathrm{dist}(p,\ell') \leq \mathrm{dist}(p,\ell_2) + \Delta\sin\theta \leq (1+\varepsilon)\mathcal{RD}_\ell.$$

■

We next extend the cylinder argument of Lemma 3.3 to k-flats. Note, that we can assume that we a point o on the k-flat, by algorithmically using the proof of Lemma 2.8 to generate a list of candidate points to lie on the k-flat, and exhaustively check all of them.

Lemma 3.4 *Let P be a set of points in $\mathbb{R}^d$, $\mathcal{F}$ be a k-flat through the origin o, $\mathcal{R}_\mathcal{F} = \mathcal{RD}(P, \mathcal{F})$, and $\varepsilon > 0$ be a parameter. There is a subset $Q \subseteq P$ of $O(k/\varepsilon^5 \log 1/\varepsilon)$ points such that the linear subspace spanned by Q contains a k-flat $\mathcal{G}$ through o such that $\mathcal{RD}(P, \mathcal{G}) \leq (1+\varepsilon)^k \mathcal{R}_\mathcal{F}$.*

We can also compute a set of $n^{O(\nu)}$ k-flats so that for some k-flat $\mathcal{G}$ in this set, $\mathcal{RD}(P, \mathcal{G}) \leq (1+\varepsilon)^k \mathcal{R}_\mathcal{F}$, and the running time of this algorithm is $dn^{O(\nu)}$, where $\nu = O\big(\frac{k}{\varepsilon^5} \log(1/\varepsilon)\big)$. The only information this algorithm needs from P is the set of all subsets of P with $O(k/\varepsilon^5 \log 1/\varepsilon)$ points.

Proof: The proof is by induction on k. The base case $k = 1$ follows from Lemma 3.3. Let $k > 1$. Our first step is to show that there is a subset $Q_1 \subseteq P$ of $O(1/\varepsilon^5 \log 1/\varepsilon)$ points such that there is a k-flat $\mathcal{H}$ through o that satisfies the following two properties: (1) $\mathcal{RD}(P, \mathcal{H}) \leq (1+\varepsilon)\mathcal{R}_\mathcal{F}$, and (2) $\mathcal{H}$ intersects the linear subspace spanned by Q_1 in a line. Let us consider an orthonormal co-ordinate system in which the first k unit vectors $e_1, \ldots, e_k$ span the k-flat $\mathcal{F}$. Let $\pi : \mathbb{R}^d \to \mathbb{R}^{d-k+1}$ be the projection which takes a point $(x_1, \ldots, x_{k-1}, x_k, \ldots, x_d)$ to the point $(x_k, \ldots, x_d)$. The transformation π maps each k-flat $\mathcal{F}'$ in $\mathbb{R}^d$ that contains the vectors $e_1, \ldots, e_{k-1}$ bijectively to a line $\pi(\mathcal{F}')$ through the origin in $\mathbb{R}^{d-k+1}$ such that for any $p \in P$, $\mathrm{dist}(p, \mathcal{F}') = \mathrm{dist}(\pi(p), \pi(\mathcal{F}'))$. In particular, $\mathcal{F}$ is mapped to a line $\pi(\mathcal{F})$ through the origin such that $\mathrm{dist}(\pi(p), \pi(\mathcal{F})) \leq \mathcal{R}_\mathcal{F}$ for any $p \in P$. Applying Lemma 3.3 to $\pi(\mathcal{F})$, we see that there is a subset $Q_1 \subseteq P$ of $O(1/\varepsilon^5 \log 1/\varepsilon)$ points such that the linear subspace in $\mathbb{R}^{d-k+1}$ spanned by $\pi(Q_1)$ contains a line ℓ' through the origin such that $\mathrm{dist}(\pi(p), \ell') \leq (1+\varepsilon)\mathcal{R}_\mathcal{F}$ for any $p \in P$. Furthermore, we may assume that ℓ' passes through the origin and a point s' in an $(\varepsilon^{5/2}/16)$-net of U', where U' is obtained by inverting some of the points in $\pi(Q_1)$. But by Lemma 2.4, s' can be assumed to be $\pi(s)$ for some s in an $(\varepsilon^{5/2}/16)$-net of U, where U is obtained by inverting the corresponding set of points in Q_1. Let $\mathcal{H}$ be the k-flat in $\mathbb{R}^d$ containing $e_1, \ldots, e_{k-1}$ such that $\pi(\mathcal{H}) = \ell'$. We see that $\mathcal{H}$ intersects the linear subspace spanned by Q_1 in a line ℓ through o and s that $\mathrm{dist}(p, \mathcal{H}) \leq (1+\varepsilon)\mathcal{R}_\mathcal{F}$ for any $p \in P$.

Let us consider another orthonormal co-ordinate system in which the first unit vector e_1 spans ℓ. Let $\sigma : \mathbb{R}^d \to \mathbb{R}^{d-1}$ be the projection that takes a point $(x_1, x_2, \ldots, x_d)$ to the point $(x_2, \ldots, x_d)$. The transformation σ maps each k-flat $\mathcal{H}'$ containing ℓ bijectively to a $(k-1)$-flat $\sigma(\mathcal{H}')$ through the origin such that $\mathrm{dist}(p, \mathcal{H}') = \mathrm{dist}(\sigma(p), \sigma(\mathcal{H}'))$. In particular, $\mathcal{H}$ is mapped to a $(k-1)$-flat $\sigma(\mathcal{H})$ through the origin such that $\mathrm{dist}(\sigma(p), \sigma(\mathcal{H})) \leq (1+\varepsilon)\mathcal{R}_\mathcal{F}$ for any $p \in P$. Applying the induction hypothesis, we see that there is a subset $Q_2 \subseteq P$ of $O((k-1)/\varepsilon^5 \log 1/\varepsilon)$ points such that the linear subspace in $\mathbb{R}^{d-1}$ spanned by $\sigma(Q_2)$ contains a $(k-1)$-flat $\mathcal{B}$ through the origin such that $\mathrm{dist}(\sigma(p), \mathcal{B}) \leq (1+\varepsilon)^k \mathcal{R}_\mathcal{F}$ for any $p \in P$. Let $\mathcal{G}$ be the k-flat in $\mathbb{R}^d$ containing ℓ such that $\pi(\mathcal{G}) = \mathcal{B}$. Clearly, $\mathrm{dist}(p, \mathcal{G}) \leq (1+\varepsilon)^k \mathcal{R}_\mathcal{F}$ for any $p \in P$. It is also easy to verify that $\mathcal{G}$ lies in the linear subspace spanned by Q_2 and ℓ, and therefore the linear subspace spanned by $Q = Q_1 \cup Q_2$. Q has $O(k/\varepsilon^5 \log 1/\varepsilon)$ points.

The algorithm to compute $\mathcal{G}$ is as follows. We first enumerate all possible candidates for the point s. There are $n^{O(1/\varepsilon^5 \log 1/\varepsilon)}$ possibilities for the set Q_1, $2^{|Q_1|}$ possibilities for the set U given Q_1, $\exp\big(O(1/\varepsilon^5 \log^2(1/\varepsilon))\big)$ possibilities for s in the $(\varepsilon^{5/2}/16)$-net $A(U)$ for U (Lemma 2.5). Once s is determined, the projection σ is also determined. For each possible s, we recursively apply the same algorithm to compute a good $(k-1)$-flat $\mathcal{B}$ through the origin for $\sigma(P)$. The flat $\mathcal{G}$ containing ℓ such that $\sigma(\mathcal{G}) = \mathcal{B}$ is a good k-flat for P. It is straightforward to verify that the running time of the algorithm is as stated. ■

4 Algorithms

The results in the previous section imply efficient approximation algorithms for the k-flat problem, the same problem with outliers, and the j k-flat problem. We now state these algorithms with their running times.

Theorem 4.1 *Let P be a set of n points in $\mathbb{R}^d$, $\mathcal{F}$ be the k-flat that realized $\mathcal{R}_\mathcal{F} = \mathcal{RD}_k(P)$ the minimum k-flat radius of P, and $\varepsilon > 0$ be a parameter. We can compute, in $dn^{O(\nu)}$ time, a set of $n^{O(\nu)}$ k-flats so that for some k-flat $\mathcal{G}$ in this set, $\mathcal{RD}(P, \mathcal{G}) \le (1+\varepsilon)^{k+1}\mathcal{R}_\mathcal{F}$ for every $p \in P$, where $\nu = O\big(\frac{k}{\varepsilon^5}\log(1/\varepsilon)\big)$. The only information this algorithm needs from P is all subsets of P with $O(k/\varepsilon^5 \log 1/\varepsilon)$ points. Furthermore, we can determine such a k-flat $\mathcal{G}$ using $O(nd)$ time for each candidate.*

Proof: The theorem follows from Lemma 3.1 and Lemma 3.4. In view of Lemma 3.1, we generate all subsets of P with $O(1/\varepsilon^2 \log 1/\varepsilon)$ points, and for each such subset Q, we compute an $(\varepsilon/2)$-net $A(Q)$ of Q. For each point $o \in A(Q)$, we perform a translation so that o becomes the origin, and then apply the algorithm of Lemma 3.4. ■

Theorem 4.2 *Let P be a set of n points in $\mathbb{R}^d$, $m \le n$ be given, $\mathcal{R}_\mathcal{F} \ge 0$ be the smallest value such that there is a k-flat $\mathcal{F}$ which is at distance at most $\mathcal{R}_\mathcal{F}$ from at least m of the points in P. We can compute, in $dn^{O(\nu)}$ time a k-flat $\mathcal{G}$ which is at distance at most $r(1+\varepsilon)^{k+1}$ from at least m of the points in P, where $\nu = O\big(\frac{k}{\varepsilon^5}\log(1/\varepsilon)\big)$.*

Proof: Let P' be the set of points at distance at most $\mathcal{R}_\mathcal{F}$ from $\mathcal{F}$. We can use the algorithm of Theorem 4.1 to compute a candidate set of $n^{O(\nu)}$ flats that contains the required k-flat g provided we can generate all possible subsets of P' with $O(k/\varepsilon^5 \log 1/\varepsilon)$ points. We simply try all possible subsets of P with $O(k/\varepsilon^5 \log 1/\varepsilon)$ points, and take the best flat among all candidates generated. ■

Theorem 4.3 *Let P be a set of n points in $\mathbb{R}^d$, and $\mathcal{RD} \ge 0$ be the smallest value such that there are j k-flats $\mathcal{F}_1, \ldots, \mathcal{F}_j$ for which $\min_{i \le j} \operatorname{dist}(p, \mathcal{F}_i) \le \mathcal{RD}$ for any $p \in P$. We can compute, in $dn^{O(\nu)}$ time a set of j k-flats $\mathcal{G}_1, \ldots, \mathcal{G}_j$ for which $\min_{i \le j} \operatorname{dist}(p, \mathcal{G}_i) \le (1+\varepsilon)^{k+1}\mathcal{RD}$ for any $p \in P$, where $\nu = O\big(\frac{j \cdot k}{\varepsilon^5}\log(1/\varepsilon)\big)$.*

Proof: Assign each point in P to the closest flat in $\{\mathcal{F}_1, \ldots, \mathcal{F}_j\}$. Let P_i denote the set of points assigned to $\mathcal{F}_i$. We can use the algorithm of Theorem 4.1 to generate a set of candidate k-flats C_i that is guaranteed to contain a k-flat $\mathcal{G}_i$ such that $\max_{p \in P_i} \operatorname{dist}(p, \mathcal{G}_i) \le (1+\varepsilon)^{k+1}\mathcal{R}_\mathcal{F}$. Since we do not know all possible subsets of P_i of size $O(k/\varepsilon^5 \log 1/\varepsilon)$, we simply try all possible subsets of P of size $O(k/\varepsilon^5 \log 1/\varepsilon)$. Once we have $C_1, \ldots, C_j$, we return the best set of j k-flats in $C_1 \times \cdots \times C_j$. ■

The arguments in the Section 3 immediately imply the following structural result.

Theorem 4.4 *Let P be a set of n points in $\mathbb{R}^d$ and $\mathcal{R}_{\mathcal{F}}$ be the k-flat radius of P. For any $\varepsilon > 0$, there is a subset $Q \subseteq P$ of $b = O(k/\varepsilon^5 \log 1/\varepsilon)$ points such that for some k-flat $\mathcal{G}$ in the affine hull of Q, $\mathrm{dist}(p, \mathcal{G}) \leq (1+\varepsilon)^{k+1}\mathcal{R}_{\mathcal{F}}$ for any $p \in P$.*

It is possible to derive a reasonably efficient approximation algorithm for the k-flat center problem from this existential result alone. Although the resulting algorithm is not as efficient as the algorithms derived above, it is worth describing. Suppose we have guessed the right set Q (by trying all possible subsets of P with b points). We are then left with the problem of finding the k-flat $\mathcal{F}$ in the affine hull of Q that minimizes $\max_{p \in P} \mathrm{dist}(p, \mathcal{F})$. This problem can be written as an optimization problem in $b+1$ dimensions as follows. Let us construct an orthonormal co-ordinate system for $\mathbb{R}^d$ in which the first k unit vectors $e_1, \ldots, e_t$ span the affine hull of Q, where $t \leq b$. For any point $x \in \mathbb{R}^d$, let $\pi(x)$ denote its projection onto the affine hull of Q, and $w(x) = ||x - \pi(x)||$ denote the distance of p from the affine hull of Q. Consider the map σ that takes a point $x = (x_1, \ldots, x_t, x_{t+1}, \ldots, x_d)$ to the point $\sigma = (x_1, \ldots, x_t, w(x))$ in $\mathbb{R}^{d+1}$. The important property is that for any flat $\mathcal{F}$ in the affine hull of Q and any $x \in \mathbb{R}^d$, the distance of $\mathcal{F}$ from x is equal to the distance of the "image" of $\mathcal{F}$ from $\sigma(x)$. We are now left with the problem of finding the best k-flat $\mathcal{G}$ lying in a given hyperplane in $\mathbb{R}^{b+1}$ that minimizes $\max_{p \in \sigma(P)} \mathrm{dist}(p, \mathcal{G})$. This is an optimization problem in $b+1$ dimensions and can be solved exactly in $O(n^{bk})$ time by trying all possible extremal solutions.

This results in an $O(n^{bk})$ time algorithm for the original problem, which is worse than the algorithms described earlier. Indeed, the running time of the previous algorithm was (essentially) $O(n^b)$. This approach also extends to the problem with outliers and the projective clustering problem.

5 Conclusions

In this paper, we presented several fast algorithms for doing $(1+\varepsilon)$-approximate projective clustering. Our algorithm relied on the ability to compute small core-sets for those projective clustering problems. The most striking property of our algorithm is the fact that the running time depends only polynomially on the dimension.

There are several interesting questions for further research:

- Can one come up with an algorithm for approximate k-flat clustering in high-dimensions which would work well also in low-dimensions. Namely, it should be competitive with the currently fastest algorithms in low dimensions [HV01]?
- The existence of core-sets of size independent of the dimension and that depends only on the approximation quality is quite surprising. It would be interesting to extend the family of problems for which we know the existence of such core-sets.

- Can one use dimension reduction techniques [Mag01, JL84, Ach01] coupled with convex programming techniques [GLS88] to get faster algorithms for the problems discussed?

Acknowledgments

The authors thank Joachim Giesen for helpful comments on the manuscript.

References

[Ach01] D. Achlioptas. Database-friendly random projections. In *Proc. 20th ACM Sympos. Principles Database Syst.*, pages 274–281, 2001.

[ADPR00] N. Alon, S. Dar, M. Parnas, and D. Ron. Testing of clustering. In *Proc. 41th Annu. IEEE Sympos. Found. Comput. Sci.*, pages 240–250, 2000.

[AP00] P. K. Agarwal and C. M. Procopiuc. Approximation algorithms for projective clustering. In *Proc. 11th ACM-SIAM Sympos. Discrete Algorithms*, pages 538–547, 2000.

[BH01] G. Barequet and S. Har-Peled. Efficiently approximating the minimum-volume bounding box of a point set in three dimensions. *J. Algorithms*, 38:91–109, 2001.

[BHI02] M. Bădoiu, S. Har-Peled, and P. Indyk. Approximate clustering via core-sets. In *Proc. 34th Annu. ACM Sympos. Theory Comput.*, pages 250–257, 2002.

[DHS01] R. O. Duda, P. E. Hart, and D. G. Stork. *Pattern Classification.* Wiley-Interscience, New York, 2nd edition, 2001.

[GK92] P. Gritzmann and V. Klee. Inner and outer j-radii of convex bodies in finite-dimensional normed spaces. *Discrete Comput. Geom.*, 7:255–280, 1992.

[GK93] P. Gritzmann and V. Klee. Computational complexity of inner and outer j-radii of polytopes in finite-dimensional normed spaces. *Math. Program.*, 59:163–213, 1993.

[GK94] P. Gritzmann and V. Klee. On the complexity of some basic problems in computational convexity: I. containment problems. *Discrete Math.*, 136:129–174, 1994.

[GLS88] M. Grötschel, L. Lovász, and A. Schrijver. *Geometric Algorithms and Combinatorial Optimization*, volume 2 of *Algorithms and Combinatorics.* Springer-Verlag, Berlin Heidelberg, 2nd edition, 1988. 2nd edition 1994.

[Har01] S. Har-Peled. Clustering motion. In *Proc. 42nd Annu. IEEE Sympos. Found. Comput. Sci.*, pages 84–93, 2001.

[Hoc96] D. Hochbaum. Approximation algorithms for np-hard problems, 1996.

[HV01] S. Har-Peled and K. R. Varadarajan. Approximate shape fitting via linearization. In *Proc. 42nd Annu. IEEE Sympos. Found. Comput. Sci.*, pages 66–73, 2001.

[Ind01] P. Indyk. Algorithmic applications of low-distortion geometric embeddings. In *Proc. 42nd Annu. IEEE Sympos. Found. Comput. Sci.*, pages 10–31, 2001. Tutorial.

[IT00] P. Indyk and M. Thorup. Approximate 1-median. manuscript, 2000.

[JL84] W. B. Johnson and J. Lindenstrauss. Extensions of lipshitz mapping into hilbert space. *Contemporary Mathematics*, 26:189–206, 1984.

[Mag01] A. Magen. Dimensionality reductions that preserve volumes and distance to affine spaces, and its algorithmic applications. Submitted to STOC 2002, 2001.

[MOP01] N. Mishra, D. Oblinger, and L. Pitt. Sublinear time approximate clustering. In *Proc. 12th ACM-SIAM Sympos. Discrete Algorithms*, pages 439–447, 2001.

[MT82] N. Megiddo and A. Tamir. On the complexity of locating linear facilities in the plane. *Oper. Res. Lett.*, 1:194–197, 1982.