No-Dimensional Tverberg Partitions Revisited

Sariel Har-Peled
Department of Computer Science, University of Illinois, Urbana, IL, USA

Eliot W. Robson
Department of Computer Science, University of Illinois, Urbana, IL, USA

Abstract

Given a set $P \subset \mathbb{R}^d$ of n points, with diameter Δ, and a parameter $\delta \in (0,1)$, it is known that there is a partition of P into sets $P_1, \ldots, P_t$, each of size $O(1/\delta^2)$, such that their convex hulls all intersect a common ball of radius $\delta\Delta$. We prove that a random partition, with a simple alteration step, yields the desired partition, resulting in a (randomized) linear time algorithm (i.e., $O(dn)$). We also provide a deterministic algorithm with running time $O(dn \log n)$. Previous proofs were either existential (i.e., at least exponential time), or required much bigger sets. In addition, the algorithm and its proof of correctness are significantly simpler than previous work, and the constants are slightly better.

We also include a number of applications and extensions using the same central ideas. For example, we provide a linear time algorithm for computing a "fuzzy" centerpoint, and prove a no-dimensional weak ε-net theorem with an improved constant.

2012 ACM Subject Classification Theory of computation → Computational geometry

Keywords and phrases Points, partitions, convex hull, high dimension

Digital Object Identifier 10.4230/LIPIcs.SWAT.2024.26

Related Version *Full Version*: `https://arxiv.org/abs/2306.01678` [12]

Funding Partially supported by NSF AF award CCF-2317241.

Acknowledgements The authors thank Ken Clarkson and Sandeep Sen for useful discussions.

1 Introduction

Centerpoints

A point c is an *α-centerpoint* of a set $P \subseteq \mathbb{R}^d$ of n points, if all closed halfspaces containing c also contain at least αn points of P. The parameter α is the *centrality* of c, while αn is its **Tukey depth**. The centerpoint theorem [17], which is a consequence of Helly's theorem [14], states that a $1/(d+1)$-centerpoint (denoted $\bar{\mathsf{c}}_P$) always exists.

In two dimensions, Jadhav and Mukhopadhyay [16] presented an $O(n)$ time algorithm for computing a 1/3-centerpoint (but not the point of maximum Tukey depth). Chan et al. [4] presented an $O(n \log n + n^{d-1})$ algorithm for computing the point of maximum Tukey depth (and thus also a $1/(d+1)$-centerpoint). It is believed that $\Omega(n^{d-1})$ is a lower bound on solving this problem exactly, see [4] for details and history.

This guarantee of $1/(d+1)$-centerpoint is tight, as demonstrated by placing the points of P in $d+1$ small, equal size clusters (mimicking weighted points) in the vicinity of the vertices of a simplex. Furthermore, the lower-bound of $\lceil n/(d+1) \rceil$ is all but meaningless if d is as large as $n-1$.

Approximating centrality

A randomized $\widetilde{O}(d^9)$ time algorithm for computing a (roughly) $1/(4d^2)$ centerpoint was presented by Clarkson et al. [9], and a later refinement by Har-Peled and Jones [11] improved this algorithm to compute a (roughly) $1/d^2$ centerpoint in $\widetilde{O}(d^7)$ time, where $\widetilde{O}$ hides polylog terms. Miller and Sheehy [19] derandomized the algorithm of Clarkson et al., computing a

$\Omega(1/d^2)$ centerpoint in time $n^{O(\log d)}$. Developing an algorithm that computes a $1/(d+1)$-centerpoint in polynomial time (in d) in still open, although the existence of such an algorithm with running time better than $\Omega(n^{d-1})$ seems unlikely, as mentioned above.

Tverberg partitions

Consider a set P of n points in $\mathbb{R}^d$. Tverberg's theorem states that such a set can be partitioned into $k = \lfloor n/(d+1) \rfloor$ subsets, such that all of their convex-hulls intersect. Specifically, a point in this common intersection is a $1/(d+1)$-centerpoint. Indeed, a point p contained in the convex-hulls of the k sets of the partition is a k/n-centerpoint, as any halfspace containing p must also contain at least one point from each of these k subsets. Refer to the surveys [10] and [3] for information on this and related theorems.

This theorem has an algorithmic proof [20], but its running time is $n^{O(d^2)}$. To understand the challenge in getting an efficient algorithm for this problem, observe that it is not known, in strongly polynomial time, to decide if a point is inside the convex-hull of a point set (i.e., is it $1/n$-centerpoint?). Similarly, for a given point p, it is not known how to compute, in weakly or strongly polynomial time, the centrality of p. Nevertheless, a Tverberg partition is quite attractive, as the partition itself (and its size) provides a compact proof (i.e., lower bound) of its centrality. If the convex-combination realization of p inside each of these sets is given, then its k/n-centrality can be verified in linear time.

There has been significant work trying to compute Tverberg partitions with as many sets as possible while keeping the running time polynomial. The best polynomial algorithms currently known (roughly) match the bounds for the approximate centerpoint mentioned above. Specifically, it is known how to compute a Tverberg partition of size $O\big(n/(d^2 \log d)\big)$ (along with a point in the common intersection) in weakly polynomial time. See [13] and references therein.

No-dimensional Tverberg theorem

Adiprasito et al. [1] proved a no-dimensional variant of Tverberg's theorem. Specifically, for $\delta \in (0,1)$, they showed that one can partition a point set P into sets of size $O(1/\delta^2)$, such that the convex-hulls of the sets intersect a common ball of radius $\delta \operatorname{diam}(P)$. Their result is existential and does not yield an efficient algorithm. However, as the name suggests, it has the attractive feature that the sets in the partition have size that does not depend on the dimension.

Choudhary and Mulzer [7] gave a weaker version of this theorem, but with an efficient algorithm. Speculatively, given a set $P \subset \mathbb{R}^d$ of n points, and a parameter $\delta \in (0,1)$, P can be partitioned, in $O(nd \log k)$ time, into $k = O(\delta\sqrt{n})$ sets $P_1, \ldots, P_k$, each of size $\Theta(\sqrt{n}/\delta)$, such that there is a ball of radius $\delta \operatorname{diam}(P)$ that intersects the convex-hull of P_i for every i. Note that the later (algorithmic) result is significantly weaker than the previous (existential) result, as the subsets have to be substantially larger.

Thus, the question remains: Can one compute a no-dimensional Tverberg partition with the parameters of Adiprasito et al. [1] in linear time?

Centerball via Tverberg partition

As observed by Adiprasito et al. [1], a no-dimensional Tverberg partition readily implies a no-dimensional centerpoint result, where the central point is replaced by a ball. Specifically, they showed that one can compute a ball of radius $\delta \operatorname{diam}(P)$ such that any halfspace containing it contains $\Omega(\delta^2 n)$ points of P.

Centroid and sampling

The **centroid** of a point set P is the point $\overline{\mathsf{m}}_P = \sum_{p \in P} p / |P|$. The **1-mean** price of clustering P, using q, is the sum of squared distances of the points of P to q, that is $f(q) = \sum_{p \in P} \|p - q\|^2$. It is not hard to verify that f is minimized at the centroid $\overline{\mathsf{m}}_P$. A classical observation of Inaba et al. [15] is that a sample R of size $O(1/\delta^2)$ of points from P is δ-close to the global centroid of the point set. That is, $\|\overline{\mathsf{m}}_P - \overline{\mathsf{m}}_R\| \leq \delta \operatorname{diam}(P)$ with constant probability. Applications of this observation to k-means clustering and sparse coresets are well known, see Clarkson [8, Section 2.4] and references therein.

Our results

We show that the aforementioned observation of Inaba et al. implies the no-dimensional Tverberg partition. Informally, for a random partition of P into sets of size $O(1/\delta^2)$, most of the sets are in distance at most $\delta \operatorname{diam}(P)$ from the global centroid of P. By folding the far sets (i.e., "bad"), into the close sets (i.e., "good"), we obtain the desired partition. The resulting algorithm has (expected) linear running time $O(dn)$.

For the sake of completeness, we prove the specific form of the 1-mean sampling observation [15] we need in Lemma 3 – the proof requires slightly tedious but straightforward calculations. The linear time algorithm for computing the no-dimensional Tverberg partition is presented in Theorem 6, which is the main result of this paper.

In the other extreme, one wants to split the point set into two sets of equal size while minimizing their distance. We show that a set P with $2n$ points can be split (in linear time) into two sets of size n, such that (informally) the expected distance of their centroids is $\leq \operatorname{diam}(P)/\sqrt{n}$. The proof of this is even simpler (!), and the bound is tight; see Lemma 9. We present several applications:

(I) No-dimensional Centerball. In Section 3.1, we present a no-dimensional generalization of the centerpoint theorem. As mentioned above, this was already observed by Adiprasito et al. [1], but our version can be computed efficiently.

(II) Weak ε-net. A new proof of the no-dimensional version of the weak ε-net theorem with improved constants, see Section 3.2.

(III) Derandomization. The sampling mean lemma (i.e., Lemma 3) can be derandomized to yield a linear time algorithm, see Lemma 16. The somewhat slower version, Lemma 15, is a nice example of using conditional expectations for derandomization. Similarly, the halving scheme of Lemma 9 can be derandomized in a fashion similar to discrepancy algorithms [18, 5]. The derandomized algorithm, presented in Lemma 17, has linear running time $O(dn)$.

This leads to a deterministic $O(dn \log n)$ time algorithm for the no-dimensional Tverberg partition, see Lemma 18. The idea is to repeatedly apply the halving scheme, in a binary tree fashion, till the point set is partitioned into subsets of size $O(1/\delta^2)$. Both the running time and constants are somewhat worse than the randomized algorithm of Theorem 6, but it is conceptually even simpler, avoiding the need for an alteration step.

As an extra, another neat implication of the observation of Inaba et al. [15] is the dimension free version of Carathéodory's theorem [17], which we present in the full version.

Simplicity

While simplicity is in the eyes of the beholder, the authors find the brevity of the results here striking compared to previous work. In particular, our presentation here is longer than strictly necessary, as we reproduce proofs of previous known results, such as Lemma 3 and its variant Lemma 9, so our work is self contained.

2 Approximate Tverberg partition via mean sampling

In the following, for two points $p, q \in \mathbb{R}^d$, let $pq = \langle p, q \rangle = \sum_{i=1}^{d} p[i]q[i]$ denote their dot-product. Thus, $p^2 = \langle p, p \rangle = \|p\|^2$. Let P be a finite set of points in $\mathbb{R}^d$ (but any metric space equipped with a dot-product suffices), and let $\overline{\mathsf{m}}_P = \sum_{p \in P} p / |P|$ denote the **centroid** of P. The **average price** of the 1-mean clustering of P is

$$\nabla(P) = \sqrt{\sum\nolimits_{p \in P} \|p - \overline{\mathsf{m}}_P\|^2 \,/\, |P|} \leq \operatorname{diam}(P). \tag{2.1}$$

The last inequality follows as $\overline{\mathsf{m}}_P \in \mathsf{CH}(P)$, and for any $p \in P$, we have $\|p - \overline{\mathsf{m}}_P\| \leq \operatorname{diam}(P)$. This inequality can be tightened.

▶ **Lemma 1.** *We have $\nabla(P) \leq \operatorname{diam}(P)/\sqrt{2}$, and there is a point set Q in $\mathbb{R}^d$, such that*

$$\nabla(Q) \geq \left(1 - \tfrac{1}{d}\right) \tfrac{1}{\sqrt{2}} \operatorname{diam}(Q)$$

(i.e., the inequality is essentially tight).

Proof. This claim only improves the constant in our main result, and the reader can safely skip reading the proof. Let P be a set of n points in $\mathbb{R}^d$, with $\Delta = \operatorname{diam}(P)$ and $\nabla = \nabla(P)$. Assume that $\overline{\mathsf{m}}_P = 0$, as the claim is translation invariant. That is $\sum_{q \in P} q = 0$, and

$$\beta = \sum_{p,q \in P} \langle p, q \rangle = \sum_{p \in P} \left\langle p, \sum\nolimits_{q \in P} q \right\rangle = \sum_{p \in P} \langle p, 0 \rangle = 0.$$

We have

$$n\nabla^2 = \sum_{p \in P} \|p\|^2 = \sum_{p,q \in P} \frac{\|p\|^2 + \|q\|^2}{2n} = \sum_{p,q \in P} \frac{\|p\|^2 - 2\langle p, q \rangle + \|q\|^2}{2n} + \frac{2\beta}{2n} = \sum_{p,q \in P} \frac{\|p - q\|^2}{2n}$$
$$\leq \sum_{p \in P, q \in P} \frac{\Delta^2}{2n} = \frac{n^2 \Delta^2}{2n}.$$

Implying that $\nabla^2 \leq \Delta^2/2$.

As for the lower bound, let e_i be the ith standard unit vector[1] in $\mathbb{R}^d$, and consider the point set $Q = \{e_1, \ldots, e_d\}$. We have that $\operatorname{diam}(Q) = \sqrt{2}$ and $\overline{\mathsf{m}}_Q = (1/d, \ldots, 1/d)$. Consequently,

$$\nabla(Q) = \sqrt{\tfrac{1}{|Q|} \sum\nolimits_{q \in Q} \|q - \overline{\mathsf{m}}_Q\|^2} = \sqrt{\tfrac{|Q|}{|Q|}\left((1 - 1/d)^2 + (d-1)/d^2\right)} = \sqrt{\frac{(d-1)^2 + d - 1}{d^2}}$$
$$= \sqrt{\frac{d-1}{d}} = \frac{\operatorname{diam}(Q)}{\sqrt{2}} \sqrt{1 - \frac{1}{d}} \geq \left(1 - \frac{1}{d}\right) \frac{1}{\sqrt{2}} \operatorname{diam}(Q). \qquad ◀$$

▶ **Definition 2.** *A subset $X \subseteq P$ is δ-**close** if the centroid of X is in distance at most $\delta \operatorname{diam}(P)$ from the centroid of P – that is, $\|\overline{\mathsf{m}}_X - \overline{\mathsf{m}}_P\| \leq \delta \operatorname{diam}(P)$.*

---

[1] That is, e_i is 0 in all coordinates except the ith coordinate where it is 1.

2.1 Proximity of centroid of a sample

The following is by now standard – a random sample of $O(1/\delta^2)$ points from P is δ-close with good probability, see Inaba et al. [15, Lemma 1]. We include the proof for the sake of completeness, as we require this somewhat specific form.

▶ **Lemma 3.** *Let P be a set of n points in $\mathbb{R}^d$, and $\delta \in (0,1)$ be a parameter. Let $R \subseteq P$ be a random sample of size r picked uniformly without replacement from P, where $r \geq \zeta/\delta^2$ and $\zeta > 1$ is a parameter. Then, we have $\mathbb{P}\big[\|\overline{\mathsf{m}}_P - \overline{\mathsf{m}}_R\| > \delta\nabla(P)\big] < 1/\zeta$.*

Proof. Let $P = \{p_1, \dots, p_n\}$. For simplicity of exposition, assume that $\overline{\mathsf{m}}_P = \sum_{i=1}^n \frac{1}{n} p_i = 0$, as the claim is translation invariant. For $\nabla = \nabla(P)$, and we have $\nabla^2 = \sum_{i=1}^n \frac{1}{n} p_i^2$. Let $Y = \sum_{p\in R} p = \sum_{i=1}^n I_i p_i$, where I_i is an indicator variable for p_i being in R. By linearity of expectations, we have

$$\mathbb{E}[Y] = \sum_{i=1}^n \mathbb{E}[I_i]\, p_i = \sum_{i=1}^n \frac{r}{n} p_i = r \sum_{i=1}^n \frac{1}{n} p_i = r\, \overline{\mathsf{m}}_P = 0.$$

Observe that, for $i \neq j$, we have

$$\mathbb{E}\big[I_i I_j\big] = \mathbb{P}\big[I_i = 1 \text{ and } I_j = 1\big] = \binom{n-2}{r-2} \Big/ \binom{n}{r} = \frac{(n-2)!}{(r-2)!(n-r)!} \cdot \frac{r!(n-r)!}{n!} = \frac{r(r-1)}{n(n-1)}. \tag{2.2}$$

By the above, and since $\mathbb{E}\big[I_i^2\big] = \mathbb{E}[I_i]$, we have

$$\begin{aligned}\mathbb{E}\big[\|Y\|^2\big] &= \mathbb{E}\big[\langle Y, Y\rangle\big] = \mathbb{E}\Big[\big(\sum_{i=1}^n I_i p_i\big)^2\Big] = \sum_{i=1}^n \mathbb{E}[I_i]\, p_i^2 + 2\sum_{i<j} \mathbb{E}\big[I_i I_j\big]\, p_i p_j \\ &= \sum_{i=1}^n \frac{r}{n} p_i^2 + 2 \sum_{i<j} \frac{r(r-1)}{n(n-1)} p_i p_j \leq r\nabla^2 + \frac{r(r-1)n}{n-1} \Big(\sum_{i=1}^n \frac{1}{n} p_i\Big)^2 = r\nabla^2, \end{aligned} \tag{2.3}$$

using the shorthand $p_i p_j = \langle p_i, p_j\rangle$ and $p_i^2 = \langle p_i, p_i\rangle$. As (i) $r = |R|$, (ii) $\overline{\mathsf{m}}_R = Y/|R| = Y/r$, (iii) $r \geq \zeta/\delta^2$, and (iv) by Markov's inequality, we have

$$\mathbb{P}\big[\|\overline{\mathsf{m}}_R\| > \delta\nabla\big] = \mathbb{P}\Big[\frac{\|Y\|}{r} > \delta\nabla\Big] = \mathbb{P}\Big[\|Y\|^2 > (r\delta\nabla)^2\Big] \leq \frac{\mathbb{E}[\|Y\|^2]}{(r\delta\nabla)^2} \leq \frac{r\nabla^2}{(r\delta\nabla)^2} = \frac{1}{r\delta^2} \leq \frac{1}{\zeta}.$$

◀

Lemma 3 readily implies the no-dimensional Carathéodory theorem, see the full version for details.

2.2 Approximate Tverberg theorem

We now present the key technical lemma that will allow us to prove an approximate Tverberg theorem.

▶ **Lemma 4.** *Let P be a set of n points in $\mathbb{R}^d$, and $\delta \in (0,1)$ be a parameter, and assume that $n \gg 1/\delta^4$. Let $\nabla = \nabla(P)$. Then, one can compute, in $O(nd/\delta^2)$ expected time, a partition of P into k sets $P_1, \dots, P_k$, and a ball $\mathcal{b}$, such that*

(i) $\forall i\ \ |P_i| \leq 4/\delta^2 + 6$,
(ii) $\forall i\ \ \mathsf{CH}(P_i) \cap \mathcal{b} \neq \emptyset$,
(iii) $\mathrm{radius}(\mathcal{b}) \leq \delta\nabla$, *and*
(iv) $k \geq n/(4/\delta^2 + 6)$.

Proof. Let $\mathcal{B} = \mathcal{B}(\overline{\mathsf{m}}_P, \delta\nabla)$. Let $\zeta = 2(1+\delta^2/8)$, and $M = \lceil \zeta/\delta^2 \rceil$. We randomly partition the points of P into $t = \lfloor n/M \rfloor > M$ sets $Q_1, \ldots, Q_t$, all of size either M or $M+1$ (this can be done by randomly permuting the points of P, and allocating each set a range of elements in this permutation). Thus, each Q_i, for $i \in [\![t]\!] = \{1, \ldots, t\}$, is a random sample according to Lemma 3 with parameter $\geq \zeta$. Thus, with probability $\geq 1 - 1/\zeta$, the set Q_i, for $i \in [\![t]\!]$, is δ-close – that is, $\left\| \overline{\mathsf{m}}_{Q_i} - \overline{\mathsf{m}}_P \right\| \leq \delta\nabla$, and Q_i is then considered to be *good*.

Let Z be the number of bad sets in $Q_1, \ldots, Q_t$. The probability of a set to be bad is at most $1/\zeta$, and by linearity of expectations, $\mathbb{E}[Z] \leq t/\zeta$. Let $\beta = t(1+\delta^2/8)/\zeta = t/2$. By Markov's inequality, we have

$$\mathbb{P}\big[Z \geq t/2\big] = \mathbb{P}[Z \geq \beta] \leq \frac{\mathbb{E}[Z]}{\beta} \leq \frac{t/\zeta}{(1+\delta^2/8)t/\zeta} = \frac{1}{1+\delta^2/8} \leq 1 - \frac{\delta^2}{16}. \tag{2.4}$$

We consider a round of sampling *successful* if $Z < \beta = t/2$. The algorithm can perform the random partition and compute the centroid for all P_i in $O(nd)$ time overall. Since a round is successful with probability $\geq \delta^2/16$, after $\lceil 16/\delta^2 \rceil$ rounds, the algorithm succeeds with constant probability. This implies that the algorithm performs, in expectation, $O(1/\delta^2)$ rounds till being successful, and the overall running time is $O(nd/\delta^2)$ time in expectation.

In the (first and final) successful round, the number of bad sets is $< t/2$ – namely, it is strictly smaller than the number of good sets. Therefore, we can match each bad set B in the partition to a unique good set G, and replace both of them by a new set $X = G \cup B$. That is, every good set absorbs at most one bad set, forming a new partition with roughly half the sets. For such a newly formed set X, we have that

$$|X| = |B| + |G| \leq 2(M+1) \leq 2\lceil \zeta/\delta^2 \rceil + 2 = 2\left\lceil \frac{2(1+\delta^2/8)}{\delta^2} \right\rceil + 2 \leq 2\left(\frac{2}{\delta^2} + 2\right) + 2 \leq \frac{4}{\delta^2} + 6.$$

The point $\overline{\mathsf{m}}_G$ is in $\mathsf{CH}(G) \subset \mathsf{CH}(X)$, and $\overline{\mathsf{m}}_G$ is in distance at most $\delta\nabla$ from the centroid of P. Thus, all the newly formed sets in the partition are in distance $\leq \delta\nabla$ from $\overline{\mathsf{m}}_P$, and $\mathsf{CH}(X) \cap \mathcal{B} \neq \varnothing$.

Finally, we have that the number of sets in the merged partition is at least $k \geq \frac{n}{4/\delta^2+6}$. ◀

▶ Remark 5. The mysterious requirement that $n \gg 1/\delta^4$, in Lemma 4, is used in the partition implicitly – the number of sets in the partition needs to be even. Thus, one set might need to be absorbed in the other sets, or more precisely two sets, because of the rounding issues. Namely, we first partition the set into groups of size M, and we need at least $2M+2$ sets in the partition to have size M (one additional last set can have size smaller than M). Thus, the proof requires that $n \geq (2M+2)M + M = (2M+3)M$. This is satisfied, for example, if $n \geq 27/\delta^4$.

▶ **Theorem 6.** *Let P be a set of n points in $\mathbb{R}^d$, and $\delta \in (0, 1/\sqrt{2})$ be a parameter, and assume that $n \gg 1/\delta^4$. Then, one can compute, in $O(nd/\delta^2)$ expected time, a partition of P into sets $P_1, \ldots, P_k$, and a ball $\mathcal{B}$, such that*

- **(i)** $\forall i \;\; |P_i| \leq 2/\delta^2 + 6$,
- **(ii)** $\forall i \;\; \mathsf{CH}(P_i) \cap \mathcal{B} \neq \emptyset$,
- **(iii)** $\mathrm{radius}(\mathcal{B}) \leq \delta \mathrm{diam}(P)$, *and*
- **(iv)** $k \geq n/(2/\delta^2 + 6)$.

Proof. Let $\Delta = \mathrm{diam}(P)$. Use Lemma 4 with parameter $\sqrt{2}\delta$. Observe that

$$\mathrm{radius}(\mathcal{B}) \leq \sqrt{2}\delta\nabla \leq \sqrt{2}\delta(\Delta/\sqrt{2}) = \delta\Delta,$$

by Lemma 1, where $\nabla = \nabla(P)$.

Observe that the algorithm does not require the value of $\mathrm{diam}(P)$, but rather the value of $\nabla(P)$, which can be computed in $O(nd)$ time, see Eq. (2.1). ◀

▶ **Corollary 7.** *The expected running time of Theorem 6 can be improved to $O(nd)$, with two of the guarantees being weaker:*

(I) *The sets are bigger: $\forall i \ |P_i| \leq 3/\delta^2 + 9$.*
(II) *And there are fewer sets: $k \geq n/(3/\delta^2 + 9)$.*

Proof. We use Lemma 4 as before, but now requiring only third of the sets to be good, and merging triples of sets to get one final good set. The probability of success is now constant, as Eq. (2.4) becomes

$$\mathbb{P}\Big[Z \geq \frac{2}{3}t\Big] = \mathbb{P}\Big[Z \geq \frac{4}{3}\cdot\frac{t}{2}\Big] = \mathbb{P}\Big[Z \geq \frac{4}{3}\beta\Big] \leq \frac{\mathbb{E}[Z]}{(4/3)\beta} \leq \frac{3}{4}.$$

Namely, the partition succeeds with probability at least $1/4$, which implies that the algorithm is done in expectation after $O(1)$ partition rounds. ◀

▶ Remark 8. The (existential) result of Adiprasito et al. [1, Theorem 1.3] has slightly worse constants, but it requires some effort to see, as they "maximize" the number of sets k (instead of minimizing the size of each set). Specifically, they show that one can partition P into k sets, with the computed ball having radius $(2+\sqrt{2})\sqrt{k/n}\,\mathrm{diam}(P)$ (intuitively, one wants k to be as large as possible). Translating into our language, we require that

$$(2+\sqrt{2})\sqrt{\frac{k}{n}} \leq \delta \implies (2+\sqrt{2})^2\frac{k}{n} \leq \delta^2 \implies k \leq n\frac{\delta^2}{(2+\sqrt{2})^2}.$$

Our result, on the other hand, states that k is at least (over-simplifying for clarify) $n\frac{\delta^2}{2}$ (for δ sufficiently small). Adiprasito et al. mention, as a side note, that their constant improves to $1+\sqrt{2}$ under certain conditions. Even then, the constant in the above theorem is better.

This improvement in the constant is small (and thus, arguably minor), but nevertheless, satisfying.

2.3 Tverberg halving

An alternative approach is to randomly halve the point set and observe that the centroids of two halves are close together. In this section, we show this line of thinking leads to various algorithms that can be derandomized efficiently. Foundational to this approach is the following lemma (which is a variant of Lemma 3).

▶ **Lemma 9.** *Let $U = \{u_1, \ldots, u_{2n}\}$ be a set of $2n$ points in $\mathbb{R}^d$ with $\Delta = \mathrm{diam}(U)$. For $i = 1, \ldots, n$, with probability $1/2$, let $p_i = u_{2i-1}, q_i = u_{2i}$, or otherwise, let $p_i = u_{2i}, q_i = u_{2i-1}$. Let $P = \{p_1, \ldots, p_n\}$ and $Q = \{q_1, \ldots, q_n\}$. For any parameter $t \geq 1$, we have $\mathbb{P}\Big[\|\overline{\mathsf{m}}_P - \overline{\mathsf{m}}_Q\| \geq \frac{t}{\sqrt{n}}\Delta\Big] \leq \frac{1}{t^2}$.*

Proof. This follows by adapting the argument used in the proof of Lemma 3, and the details are included here for the sake of completeness.

Let $v_i = u_{2i-1} - u_{2i}$. Consider the random variable $Y = \overline{\mathsf{m}}_P - \overline{\mathsf{m}}_Q = \sum_{i=1}^{n} \frac{X_i v_i}{n}$, where $X_i \in \{-1, +1\}$ is picked independently with probability half. We first observe that

(i) $\mathbb{E}[Y] = \sum_{i=1}^{n} \mathbb{E}[X_i]\, v_i/n = 0$,
(ii) $\mathbb{E}\big[X_i^2\big] = 1$, and
(iii) for $i < j$, $\mathbb{E}\big[X_i X_j\big] = 0$.

Thus, we have

$$\begin{aligned}\mathbb{E}\big[\|Y\|^2\big] = \mathbb{E}\big[\langle Y, Y\rangle\big] &= \mathbb{E}\Big[\Big\langle \sum\nolimits_{i=1}^{n} \frac{X_i v_i}{n}, \sum\nolimits_{i=1}^{n} \frac{X_i v_i}{n} \Big\rangle\Big] \\ &= \frac{1}{n^2}\sum_{i=1}^{n} \mathbb{E}\Big[X_i^2\Big]\, v_i^2 + 2\frac{1}{n^2}\sum_{i<j} \mathbb{E}\big[X_i X_j v_i v_j\big] \\ &= \frac{1}{n^2}\sum_{i=1}^{n} v_i^2 \le \frac{n\Delta^2}{n^2} = \frac{\Delta^2}{n}, \end{aligned} \tag{2.5}$$

since $\|v_i\| = \|u_{2i-1} - u_{2i}\| \le \operatorname{diam}(U) = \Delta$. By Markov's inequality, we have

$$\mathbb{P}\Big[\|Y\| > t\frac{\Delta}{\sqrt{n}}\Big] = \mathbb{P}\Big[\|Y\|^2 > t^2\frac{\Delta^2}{n}\Big] \le \frac{\mathbb{E}[\|Y\|^2]}{t^2\Delta^2/n} \le \frac{1}{t^2}. \qquad \blacktriangleleft$$

Remarks.

(A) Lemma 9 can be turned into an efficient algorithm using the same Markov's inequality argument used in Theorem 6. Specifically, for any parameter $\xi \in (0,1)$, one can compute a partition into two sets P and Q with $\|\overline{\mathsf{m}}_P - \overline{\mathsf{m}}_Q\| \le (1+\xi)\Delta/\sqrt{n}$, in $O(nd/\xi)$ expected time.

(B) Lemma 9 implies that there exists a partition P and Q of U such that

$$\|\overline{\mathsf{m}}_P - \overline{\mathsf{m}}_Q\| \le \Delta/\sqrt{n}.$$

Note that this is tight. To see this, let U be the standard basis of $\mathbb{R}^{2n}$, with its diameter $\Delta = \sqrt{2}$. For any partition P and Q of U with $|P| = |Q| = n$, we have that $\|\overline{\mathsf{m}}_P - \overline{\mathsf{m}}_Q\| = \sqrt{2\sum_{i=1}^{n} 1/n^2} = \sqrt{2/n} = \Delta/\sqrt{n}$.

(C) As in the standard algorithm for computing a δ-net via discrepancy [5, 18], one can apply repeated halving to get the desired Tverberg partition until the sets are the desired size. This provides a method for a deterministic algorithm, which we present in Section 4.3.

3 Applications

3.1 No-dimensional centerball

We present an efficient no-dimensional centerpoint theorem; the previous version [1, Theorem 7.1] did not present an efficient algorithm.

▶ **Corollary 10** (No-dimensional centerpoint). *Let P be a set of n points in $\mathbb{R}^d$ and $\delta \in (0, 1/2)$ be a parameter, where n is sufficiently large (compared to δ). Then, one can compute, in $O(nd/\delta^2)$ expected time, a ball $\mathcal{b}$ of radius $\delta \operatorname{diam}(P)$, such that any halfspace containing $\mathcal{b}$ contains at least $\Omega(\delta^2 n)$ points of P.*

Proof. Follows by applying Theorem 6 and the observation that, for any halfspace containing the computed ball $\mathcal{b}$, it must also contain at least one point from each set of the partition $P_1, \ldots, P_k$, where $k = \Omega(\delta^2 n)$. Thus, the ball $\mathcal{b}$ is as desired. ◀

3.2 No-dimensional weak ε-net theorem

Originally given by Adiprasito et al. [1, Theorem 7.3], we prove a version of the no-dimensional weak ε-net theorem with an improved dependence on the parameters. For a sequence $Q = (q_1, \ldots, q_r) \in P^r$, let $\overline{\mathsf{m}}_Q = \sum_{i=1}^{r} q_i/r$. We reprove Lemma 3 under a slightly different sampling model.

▶ **Lemma 11.** *Let P be a set of n points in $\mathbb{R}^d$, and $\delta \in (0, 1/2)$ and $\zeta > 1$ be parameters. Let $r \geq \zeta/\delta^2$. For a random sequence $Q = (q_1, \ldots, q_r)$ picked uniformly at random from P^r, we have that $\mathbb{P}\Big[\big\|\overline{\mathsf{m}}_P - \overline{\mathsf{m}}_Q\big\| > \delta\Delta\Big] \leq 1/\zeta$, where $\Delta = \operatorname{diam}(P)$.*

Proof. The argument predictably follows the proof of Lemma 3, and the reader can safely skip reading it, as it adds little new. Assume that $\overline{\mathsf{m}}_P = \sum_{i=1}^n \frac{1}{n} p_i = 0$. Let $\nabla^2 = \sum_{i=1}^n \frac{1}{n} p_i^2$ and $Y = \sum_{i=1}^r q_i$. Then, $\mathbb{E}[Y] = \sum_{i=1}^r \mathbb{E}[q_i] = 0$. As $\|Y\|^2 = \langle Y, Y \rangle$, it follows that

$$\begin{aligned}\mathbb{E}\big[\|Y\|^2\big] &= \mathbb{E}\Big[\big(\sum_{i=1}^r q_i\big)^2\Big] = \sum_{k=1}^r \mathbb{E}\Big[q_k^2\Big] + 2\sum_{i<j} \mathbb{E}[q_i q_j] \\ &= \sum_{k=1}^r \sum_{i=1}^n \tfrac{1}{n} p_i^2 + 2\sum_{i<j} \mathbb{E}[q_i]\,\mathbb{E}[q_j] = r\nabla^2.\end{aligned}$$

Since $\overline{\mathsf{m}}_R = Y/r$, $r \geq \zeta/\delta^2$, and by Markov's inequality, we have

$$\mathbb{P}\big[\|\overline{\mathsf{m}}_R\| > \delta\nabla\big] = \mathbb{P}\bigg[\frac{\|Y\|}{r} > \delta\nabla\bigg] = \mathbb{P}\Big[\|Y\|^2 > (r\delta\nabla)^2\Big] \leq \frac{\mathbb{E}[\|Y\|^2]}{(r\delta\nabla)^2} \leq \frac{r\nabla^2}{(r\delta\nabla)^2} = \frac{1}{r\delta^2} \leq \frac{1}{\zeta}.$$

◀

A sequence $Q \in P^r$ **collides** with a ball $\mathscr{b}$ if $\mathscr{b}$ intersects $\mathsf{CH}(Q)$. In particular, if $\big\|\overline{\mathsf{m}}_P - \overline{\mathsf{m}}_Q\big\| \leq \delta\Delta$, then Q collides with the ball $\mathscr{b}(\overline{\mathsf{m}}_P, \delta\Delta)$, where $\Delta = \operatorname{diam}(P)$.

▶ **Lemma 12** (Selection lemma)**.** *Let P be a set of n points in $\mathbb{R}^d$ and $\delta \in (0, 1)$ be a parameter. Let $r = \lceil 2/\delta^2 \rceil$. Then, the ball $\mathscr{b} = \mathscr{b}(\overline{\mathsf{m}}_P, \delta\Delta)$ collides with at least $n^r/2$ sequences of P^r.*

Proof. Taking $\zeta = 2$, by Lemma 11, a random r-sequence from P^r has probability at least half to collide with $\mathscr{b}$, which readily implies that this property holds for half the sequences in P^r. ◀

▶ **Theorem 13** (No-dimensional weak ε-net)**.** *Let P be a set of n points in $\mathbb{R}^d$, with diameter Δ, and $\delta, \varepsilon \in (0, 1)$ be parameters, where $2/\delta^2$ is an integer. Then, there exists a set $F \subset \mathbb{R}^d$ of size $\leq 2\varepsilon^{-2/\delta^2}$ balls, each of radius $\delta\Delta$, such that, for all $Y \subset P$, with $|Y| \geq \varepsilon n$, F contains a ball of radius $\delta\Delta$ that intersects $\mathsf{CH}(Y)$.*

Proof. Our argument follows Alon et al. [2]. Let $r = 2/\delta^2$. Initialize $F = \varnothing$, and let $\mathcal{H} = P^r$. If there is a set $Q \subset P$, with $|Q| \geq \varepsilon n$, where no ball of F intersects $\mathsf{CH}(Q)$, then applying Lemma 12 to Q, the algorithm computes a ball $\mathscr{b}$, of radius $\delta\Delta$, that collides with at least $(\varepsilon n)^r/2$ sequences of Q^r. The algorithm adds $\mathscr{b}$ to the set F, and removes from $\mathcal{H}$ all the sequences that collide with $\mathscr{b}$. The algorithm continues till no such set Q exists.

As initially $|\mathcal{H}| = n^r$, the number of iterations of the algorithm, and thus the size of F, is bounded by $\frac{n^r}{(\varepsilon n)^r/2} = 2/\varepsilon^r$. ◀

▶ Remark 14. In the version given by Adiprasito et al. [1, Theorem 7.3], the set F has size at most $(2/\delta^2)^{2/\delta^2}\varepsilon^{-2/\delta^2}$, while our bound is $2\varepsilon^{-2/\delta^2}$.

4 Derandomization

4.1 Derandomizing mean sampling

Lemma 3 can be derandomized directly using conditional expectations. We also present a more efficient derandomization scheme using halving in Section 4.2.

▶ **Lemma 15.** *Let P be a set of n points in $\mathbb{R}^d$. Then, for any integer $r \geq 1$, one can compute, in deterministic $O(dn^3)$ time, a subset $R \subset P$ of size r, such that $\|\overline{\mathsf{m}}_P - \overline{\mathsf{m}}_R\| \leq \nabla(P)/\sqrt{r} \leq \operatorname{diam}(P)/\sqrt{2r}$, where $\nabla = \nabla(P)$, see Eq. (2.1).*

Proof. We derandomize the algorithm of Lemma 3. We assume for simplicity of exposition that $\overline{\mathsf{m}}_P = 0$. Let R be a sample of size r without replacement from P, and let $I_i \in \{0, 1\}$ be the indicator for the event that $p_i \in R$.

Let $Y = \sum_{i=1}^{n} I_i p_i$. Then, $\overline{\mathsf{m}}_R = Y/r$, and thus $\|\overline{\mathsf{m}}_R - \overline{\mathsf{m}}_P\| = \|Y\|/r$. Consider the quantity

$$\beta = Z(x_1, \ldots, x_t) = \mathbb{E}\Big[\|Y\|^2 \;\Big|\; \mathcal{E}\Big], \qquad \mathcal{E} \equiv (I_1 = x_1, \ldots, I_t = x_t),$$

where the expectation is over the random choices of $I_{t+1}, \ldots, I_n$. At the beginning of the $(t+1)$th iteration, the values of $x_1, \ldots, x_t$ were determined in earlier iterations, and the task at hand is to decide what value to assign to x_{t+1} that minimizes $Z(x_1, \ldots, x_t, x_{t+1})$. Thus, the algorithm computes $\beta_0 = Z(x_1, \ldots, x_t, 0)$ and $\beta_1 = Z(x_1, \ldots, x_t, 1)$.

Using conditional expectations, Eq. (2.3) becomes

$$\beta = \mathbb{E}\Big[\|Y\|^2 \;\Big|\; \mathcal{E}\Big] = \sum_{i=1}^{n} \mathbb{E}[I_i \mid \mathcal{E}]\, p_i^2 + 2 \sum_{i<j} \mathbb{E}\big[I_i I_j \mid \mathcal{E}\big]\, p_i p_j. \tag{4.1}$$

Let $\alpha = \sum_{k=1}^{t} x_k$, and observe that $r - \alpha$ points are left to be chosen to be in R after $\mathcal{E}$. As such, arguing as in Eq. (2.2), for $i < j$, we have

$$\mathbb{E}[I_i \mid \mathcal{E}] = \begin{cases} x_i & i \leq t \\ \frac{r-\alpha}{n-t} & i > t, \end{cases} \qquad \text{and} \qquad \mathbb{E}\big[I_i I_j \mid \mathcal{E}\big] = \begin{cases} x_i x_j & i < j \leq t \\ x_i \frac{r-\alpha}{n-t} & i \leq t < j \\ \frac{(r-\alpha)(r-\alpha-1)}{(n-t)(n-t-1)} & t < i < j. \end{cases} \tag{4.2}$$

This implies that the algorithm can compute β in quadratic time directly via Eq. (4.1). Similarly, the algorithm computes β_0 and β_1. Observe that

$$\beta = Z(x_1, \ldots, x_t) = \frac{r - \alpha}{n - t}\beta_1 + \frac{n - t - (r - \alpha)}{n - t}\beta_0.$$

Namely, β is a convex combination of β_0 and β_1. Thus, if $\beta_0 \leq \beta$ then the algorithm sets $x_{t+1} = 0$, and otherwise the algorithm sets $x_{t+1} = 1$.

The algorithm now performs n such assignment steps, for $t = 0, \ldots, n-1$, to compute an assignment of $x_1, \ldots, x_n$ such that $Z(x_1, \ldots, x_n) \leq \mathbb{E}[\|Y\|^2]$. Overall, this leads to a $O(dn^3)$ time algorithm. Specifically, the algorithm outputs a set $R \subseteq P$ of size r, such that $R = \{p_i \mid x_i = 1, i = 1, \ldots, n\}$. Observe that $Z(x_1, \ldots, x_n) = \|r\overline{\mathsf{m}}_R\|^2 \leq \mathbb{E}[\|Y\|^2]$. Thus, by Eq. (2.3) and Lemma 1, we have

$$\|\overline{\mathsf{m}}_R - \overline{\mathsf{m}}_P\| = \|\overline{\mathsf{m}}_R\| \leq \sqrt{\frac{\mathbb{E}[\|Y\|^2]}{r^2}} \leq \sqrt{\frac{r\nabla^2}{r^2}} = \frac{\nabla}{\sqrt{r}} \leq \frac{\operatorname{diam}(P)}{\sqrt{2r}}. \qquad ◀$$

With some care, the running time of the algorithm of Lemma 15 can be improved to $O(dn)$ time, but the details are tedious, and we delegate the proof of the following lemma to the full version.

▶ **Lemma 16.** *Let P be a set of n points in $\mathbb{R}^d$. Then, for any integer $r \geq 1$, one can compute, in $O(dn)$ deterministic time, a subset $R \subset P$ of size r, such that $\|\overline{\mathsf{m}}_P - \overline{\mathsf{m}}_R\| \leq \nabla(P)/\sqrt{r} \leq \operatorname{diam}(P)/\sqrt{2r}$.*

4.2 Derandomizing the halving scheme

The algorithm of Lemma 9 can be similarly derandomized.

▶ **Lemma 17.** *Let $U = \{u_1, \ldots, u_{2n}\}$ be a set of $2n$ points in $\mathbb{R}^d$ with $\Delta = \mathrm{diam}(U)$. One can partition U, in deterministic $O(dn)$ time, into two equal size sets P and Q, such that $\left\|\overline{\mathsf{m}}_P - \overline{\mathsf{m}}_Q\right\| \leq \Delta/\sqrt{n}$.*

Proof. We follow Lemma 9. To this end, let $v_i = u_{2i-1} - u_{2i}$, for $i = 1, \ldots, n$. Let $Y = \sum_{i=1}^{n} \frac{X_i v_i}{n}$, where $X_i \in \{-1, +1\}$. Next, consider the quantity

$$Z(x_1, \ldots, x_t) = \mathbb{E}\Big[\|Y\|^2 \;\Big|\; \mathcal{E}\Big], \qquad \mathcal{E} \equiv (X_1 = x_1, \ldots, X_t = x_t),$$

where the expectation is over the random choices of $X_{t+1}, \ldots, X_n$. By Eq. (2.5), we have $Z(x_1, \ldots, x_t) = \frac{1}{n^2}\sum_{i=1}^{n} v_i^2 + \frac{2}{n^2}\sum_{i<j} \mathbb{E}\big[X_i X_j v_i v_j \;\big|\; \mathcal{E}\big]$. The latter term is

$$\begin{aligned}\sum_{i<j} \mathbb{E}\big[X_i X_j v_i v_j \;\big|\; \mathcal{E}\big] &= \sum_{i<j: i,j \leq t} x_i x_j v_i v_j + \sum_{i<j: i \leq t < j} \mathbb{E}\big[x_i X_j v_i v_j\big] + \sum_{i<j: t<i,j} \mathbb{E}\big[X_i X_j v_i v_j\big] \\ &= \sum_{i<j\leq t} x_i x_j v_i v_j,\end{aligned}$$

as $\mathbb{E}[X_i] = \mathbb{E}\big[X_i X_j\big] = 0$. Thus, $Z(x_1, \ldots, x_t) = \frac{1}{n^2}\sum_{i=1}^{n} v_i^2 + \frac{2}{n^2}\sum_{i<j\leq t} x_i x_j v_i v_j$. The key observation is that

$$Z(x_1, \ldots, x_t) = \frac{Z(x_1, \ldots, x_t, -1) + Z(x_1, \ldots, x_t, +1)}{2}.$$

Our goal is to compute the assignment of $x_1, \ldots, x_n$ that minimizes Z. Observe that

$$D_t = Z(x_1, \ldots, x_t, +1) - Z(x_1, \ldots, x_t) = \frac{2}{n^2}\Big(\sum_{i<t+1} x_i v_i\Big) v_{t+1}.$$

If $D_t \leq 0$, then the algorithm sets $x_{t+1} = +1$, otherwise the algorithm sets $x_{t+1} = -1$. The algorithm has to repeat this process for $t = 1, \ldots, n$, and naively, each step takes $O(dn)$ time. Observe that if the algorithm maintains the quantity $V_t = \sum_{i=1}^{t} x_i v_i$, then D_t can be computed in $O(d)$ time. This determines the value of x_{t+1}, and the value of $V_{t+1} = V_t + x_{t+1} v_{t+1}$ can be maintained in $O(d)$ time. As each iteration takes $O(d)$ time, the algorithm overall takes $O(dn)$ time. By the end of this process, the algorithm will have computed an assignment $x_1, \ldots, x_n$, with an associated partition of U into P and Q. By Eq. (2.5), we have $\left\|\overline{\mathsf{m}}_P - \overline{\mathsf{m}}_Q\right\|^2 \leq \mathbb{E}\big[\|Y\|^2\big] \leq \Delta^2/n$. ◀

4.3 A deterministic approximate Tverberg partition

▶ **Lemma 18.** *Let P be a set of n points in $\mathbb{R}^d$, and $\delta \in (0, 1/4)$ be a parameter. Then, one can compute, in $O(nd \log n)$ deterministic time, a partition of P into sets $P_1, \ldots, P_k$, and a ball $\mathcal{b}$, such that*

(i) $\forall i \ |P_i| \leq 8/\delta^2$,
(ii) $\forall i \ \mathsf{CH}(P_i) \cap \mathcal{b} \neq \emptyset$,
(iii) $\mathrm{radius}(\mathcal{b}) \leq \delta\,\mathrm{diam}(P)$, *and*
(iv) $k \geq n\delta^2/8$.

Proof. Assume for the time being that n is a power of 2. As done for discrepancy, we halve the current point set, and then continue doing this recursively (on both resulting sets), using the algorithm of Lemma 17 at each stage. Conceptually, this is done in a binary tree fashion, and doing this for i levels breaks the point set into 2^i sets. Let ℓ_i be an upper bound on the distance of the centroid of a set in the ith level from the centroid of its parent. By Lemma 17, we have $2\ell_i \leq \Delta/\sqrt{n/2^i}$ (where $i = 1$ in the top level). Thus, repeating this process for t levels, we have that the distance of any centroid at the leaves to the global centroid is bounded by

$$\begin{aligned} L_t = \sum_{i=1}^{t} \ell_i \leq \sum_{i=1}^{t} \frac{\Delta}{2\sqrt{n/2^i}} = \frac{\Delta}{\sqrt{2n}} \sum_{i=0}^{t-1} \sqrt{2^i} = \frac{\Delta}{\sqrt{2n}} \left(\frac{2^{t/2}-1}{\sqrt{2}-1} \right) \\ \leq \frac{5\Delta}{2\sqrt{2n}} 2^{t/2} = \frac{5\Delta}{2\sqrt{2}} \sqrt{\frac{1}{n/2^t}}. \end{aligned} \tag{4.3}$$

Solving for $\frac{5}{2\sqrt{2}}\sqrt{\frac{1}{n/2^t}} \leq \delta$, we get that this holds for $n/2^t \geq 3.2/\delta^2$. We stop our halving procedure once t is large enough such that the preceding inequality no longer holds, implying the stated bound on the size of each set.

If n is not a power of 2 then we apply the above algorithm to the largest subset that has size that is a power of two, and then add the unused points in a round robin fashion to the sets computed. ◀

▶ Remark 19. If instead of keeping both halves, as done by the algorithm of Lemma 18, one throws one of the halves away, and repeats the halving process on the other half, we end up with a single sample. One can repeat this halving process until the "sample" size is $\Theta(1/\delta^2)$. Using the same argument as in Eq. (4.3) to bound the error, we obtain a sample R of size $\Theta(1/\delta^2)$, such that $\|\overline{\mathsf{m}}_R - \overline{\mathsf{m}}_P\| \leq \delta \operatorname{diam}(P)$. The running time is $\sum_i O(dn/2^i) = O(dn)$. Namely, we get a deterministic $O(dn)$ time algorithm that computes a sample with the same guarantees as Lemma 16 – this version is somewhat less flexible and the constants are somewhat worse.

5 Conclusions

Given a data set, archetypal analysis [6] aims to identify a small subset of points such that all (or most) points in the data can be represented as a sparse convex-combination of these "archtypes". Thus, for a sparse convex-combination of points, generating a point can be viewed as an "explanation" of how it is being induced by the data. It is thus natural to ask for as many *independent* explanations as possible for a point – the more such combinations, the more a point "arises" naturally from the data. Thus, an approximate Tverberg partition can be interpreted as stating that high dimensional data has certain points (i.e., the centroid) that are robustly generated by the data.

From a data-analysis point of view, an interesting open question is whether one can do better than the "generic" guarantees provided here. If, for example, a smaller radius centroid ball exists, can it be approximated efficiently? Can a sparser convex-combination of points be computed efficiently?

While these questions in the most general settings seem quite challenging, even solving them in some special cases might be interesting.

In addition, prior works consider other no-dimensional results, such as a no-dimensional version of Helly's theorem [1], and a no-dimensional version of the colorful Tverberg theorem [7]. Our work did not address these problems because of the focus on simplicity, and a possible further direction is to address these variants with extensions of the techniques used here.

References

1. Karim A. Adiprasito, Imre Bárány, Nabil H. Mustafa, and Tamás Terpai. Theorems of Carathéodory, Helly, and Tverberg without dimension. *Discrete Comput. Geom.*, 64(2):233–258, 2020. `doi:10.1007/s00454-020-00172-5`.
2. Noga Alon, Imre Bárány, Zoltán Füredi, and Daniel J. Kleitman. Point selections and weak e-nets for convex hulls. *Comb. Probab. Comput.*, 1:189–200, 1992. `doi:10.1017/S0963548300000225`.
3. Imre Bárány and Pablo Soberón. Tverberg's theorem is 50 years old: a survey. *Bull. Amer. Math. Soc. (N.S.)*, 55(4):459–492, 2018. `doi:10.1090/bull/1634`.
4. Timothy M. Chan, Sariel Har-Peled, and Mitchell Jones. Optimal algorithms for geometric centers and depth. *SIAM J. Comput.*, 51(3):627–663, 2022. `doi:10.1137/21M1423324`.
5. Bernard Chazelle. *The Discrepancy Method: Randomness and Complexity*. Cambridge University Press, New York, 2001. URL: `http://www.cs.princeton.edu/~chazelle/book.html`.
6. Yuansi Chen, Julien Mairal, and Zaid Harchaoui. Fast and robust archetypal analysis for representation learning. In *Proceedings of the 2014 IEEE Conference on Computer Vision and Pattern Recognition*, CVPR '14, pages 1478–1485, USA, 2014. IEEE Computer Society. `doi:10.1109/CVPR.2014.192`.
7. Aruni Choudhary and Wolfgang Mulzer. No-dimensional Tverberg theorems and algorithms. *Discrete Comput. Geom.*, 68(4):964–996, 2022. `doi:10.1007/s00454-022-00380-1`.
8. Kenneth L. Clarkson. Coresets, sparse greedy approximation, and the frank-wolfe algorithm. In Shang-Hua Teng, editor, *Proceedings of the Nineteenth Annual ACM-SIAM Symposium on Discrete Algorithms, SODA 2008, San Francisco, California, USA, January 20-22, 2008*, pages 922–931. SIAM, 2008. URL: `http://dl.acm.org/citation.cfm?id=1347082.1347183`.
9. Kenneth L. Clarkson, David Eppstein, Gary L. Miller, Carl Sturtivant, and Shang-Hua Teng. Approximating center points with iterative Radon points. *Int. J. Comput. Geom. Appl.*, 6(3):357–377, 1996. `doi:10.1142/S021819599600023X`.
10. Jesús A. De Loera, Xavier Goaoc, Frédéric Meunier, and Nabil H. Mustafa. The discrete yet ubiquitous theorems of Carathéodory, Helly, Sperner, Tucker, and Tverberg. *Bull. Amer. Math. Soc. (N.S.)*, 56(3):415–511, 2019. `doi:10.1090/bull/1653`.
11. Sariel Har-Peled and Mitchell Jones. Journey to the center of the point set. *ACM Trans. Algorithms*, 17(1):9:1–9:21, 2021. `doi:10.1145/3431285`.
12. Sariel Har-Peled and Eliot Wong Robson. No-dimensional Tverberg partitions revisited. *CoRR*, abs/2306.01678, 2023. `doi:10.48550/arXiv.2306.01678`.
13. Sariel Har-Peled and Timothy Zhou. Improved approximation algorithms for Tverberg partitions. In Petra Mutzel, Rasmus Pagh, and Grzegorz Herman, editors, *Proc. 30th Annu. Euro. Sympos. Alg.* (ESA), volume 204 of *LIPIcs*, pages 51:1–51:15. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2021. `doi:10.4230/LIPIcs.ESA.2021.51`.
14. Eduard Helly. Über Systeme von abgeschlossenen Mengen mit gemeinschaftlichen Punkten. *Monatsh. Math. Phys.*, 37(1):281–302, 1930. `doi:10.1007/BF01696777`.
15. Mary Inaba, Naoki Katoh, and Hiroshi Imai. Applications of weighted Voronoi diagrams and randomization to variance-based k-clustering (extended abstract). In Kurt Mehlhorn, editor, *Proc. 10th Annu. Sympos. Comput. Geom.* (SoCG), pages 332–339. ACM, 1994. `doi:10.1145/177424.178042`.

16 Shreesh Jadhav and Asish Mukhopadhyay. Computing a centerpoint of a finite planar set of points in linear time. *Discrete Comput. Geom.*, 12:291–312, 1994. doi:10.1007/BF02574382.

17 J. Matoušek. *Lectures on Discrete Geometry*, volume 212 of *Grad. Text in Math.* Springer, 2002. doi:10.1007/978-1-4613-0039-7.

18 Jirí Matoušek. *Geometric Discrepancy*, volume 18 of *Algorithms and Combinatorics*. Springer, 1999. doi:10.1007/978-3-642-03942-3.

19 Gary L. Miller and Donald R. Sheehy. Approximate centerpoints with proofs. *Comput. Geom. Theory Appl.*, 43(8):647–654, 2010. doi:10.1016/j.comgeo.2010.04.006.

20 Helge Tverberg and Sinsia Vreica. On generalizations of Radon's theorem and the ham sandwich theorem. *Eur. J. Comb.*, 14(3):259–264, 1993. doi:10.1006/eujc.1993.1029.