Combinatorics, Probability and Computing (1992) 1, 189–200

Point Selections and Weak ε-Nets for Convex Hulls

NOGA ALON[†], IMRE BÁRÁNY[‡], ZOLTÁN FÜREDI[•] and DANIEL J. KLEITMAN[§]

[†] Department of Mathematics, Raymond and Beverly Sackler Faculty of Exact Sciences, Tel Aviv University, Tel Aviv, Israel, and Bellcore, Morristown, NJ 07960, USA.
e-mail: noga@math.tau.ac.il

[‡] Cowles Foundation, Yale University, New Haven, CT 06520, USA, and Courant Institute, New York University, New York, NY 10009, USA.
e-mail: h2923bar@ella.hu

[•] Mathematical Institute of the Hungarian Academy of Sciences POB 127, 1364 Budapest, Hungary, and Department of Mathematics, University of Illinois, Urbana, IL 61801, USA.
e-mail: zoltan@math.uiuc.edu

[§] Department of Mathematics, MIT, Cambridge, MA 02139, USA.
e-mail: djk@math.mit.edu

Received 30 October 1991; revised 7 September 1992

One of our results: let X be a finite set on the plane, $0 < \varepsilon < 1$, then there exists a set F (a weak ε-net) of size at most $7/\varepsilon^2$ such that every convex set containing at least $\varepsilon|X|$ elements of X intersects F. Note that the size of F is independent of the size of X.

1. Introduction

This paper is about weak ε-nets, point selections, convex hulls, and related topics. To explain what they mean, we start with the assumption that $d \geq 2$ and

$$X \subset \mathbb{R}^d \text{ is a set of } n \text{ points in general position.} \tag{1}$$

[†] Research supported by a United States Israel BSF grant.
[‡] On leave from the Mathematical Institute of the Hungarian Academy of Sciences, POB 127, 1364 Budapest, Hungary. Supported by the Program in Discrete Mathematics and its Applications at Yale, NSF grant 8901484, and Hungarian National Science Foundation grant No. 1812.
[•] This research was done while the author visited the Department of Mathematics at Massachusetts Institute of Technology. Supported in part by Hungarian National Science Foundation grant No. 1812.
[§] Supported in part by NSF grants DMS–86–06225 and AFOSR–0271.

We only assume general position to simplify the presentation: all of our results can be extended to any finite set X using an appropriate limit procedure (and suitable extensions of the definitions). Write $\binom{X}{d+1}$ for the set of all $(d+1)$-tuples of X. Since the points of X live in $\mathbb{R}^d$, these $(d+1)$-tuples can, and will, be called simplices when we consider their convex hull. This should not cause any confusion.

We deal with the following three problems. Given a (large) set of simplices $\mathscr{H} \subset \binom{X}{d+1}$, find a point that is contained in the maximum possible number of simplices. We call this the *point selection problem* (Section 2, proof in Section 6). In the *hitting set problem* (Sections 5 and 7), we shall look for a small set E meeting "almost all" simplices in $\binom{X}{d+1}$. Finally, in the *weak ε-net problem* (Sections 4 and 8–10), given a set X and $0 < \varepsilon < 1$, we look for a small set F such that any convex region C with $|X \cap C| \geq \varepsilon |X|$ contains a point of F. We shall find several upper bounds for $\min |F|$, together with polynomial algorithms for finding a small set F.

2. Piercing many simplices by one point

A family $\mathscr{H}$ is called *pierceable* if there exists a point common to int conv(S) for every $S \in \mathscr{H}$. We have the following *Point Selection Theorem*.

Theorem 2.1. *Given $d \geq 2$, there exists a constant $s = s_d$ such that any family $\mathscr{H} \subset \binom{X}{d+1}$, with $|\mathscr{H}| = p\binom{n}{d+1}$, contains a pierceable subfamily $\mathscr{H}'$ such that*

$$|\mathscr{H}'| \gg_d p^s \binom{n}{d+1}.$$

Here, and in what follows, we are using Vinogradov's notation. For two functions f and g, $f \gg g$ means that there are two absolute constants $c_1 > 0$ and $c_2 \in \mathbb{R}$ such that $f \geq c_1 g + c_2$ for all values of the parameters. Similarly, $f \gg_d g$ means that there are constants $c_1(d) > 0$ and $c_2(d)$ such that $f \geq c_1(d) g + c_2(d)$ for all values of the parameters.

The first point selection theorem is due to Boros and Füredi [7]. They show that for $d = 2$, the family $\binom{X}{3}$ contains a pierceable subfamily of size $(2/9)\binom{n}{3}$. This is extended for any dimension in Bárány [3], where it is proved that $\binom{X}{d+1}$ contains a pierceable subfamily of size at least

$$(d+1)\binom{\lfloor n/(d+1) \rfloor}{d+1} = (1 - o(1))\frac{1}{(d+1)^d}\binom{n}{d+1}. \tag{2}$$

The term "point selection" comes from Aronov et al. [2]. They prove, again when $d = 2$, that any family $\mathscr{H} \subset \binom{X}{3}$ of size $n^{3-\alpha}$ contains a pierceable subfamily of size

$$2^{-9}\frac{n^{3-3\alpha}}{\log^5 n}. \tag{3}$$

Thus, $s = s_2 = 3 + \delta$ will do (for any positive δ) in the point selection theorem in the range $p = n^{-\alpha}$, $\alpha > 0$. Here we prove that, in general, one can take

$$s_d = (4d+1)^{d+1}. \tag{4}$$

3. The multicoloured Tverberg theorem

Our point selection theorem will follow from a nice recent result of Živaljević and Vrećica [19], which was also conjectured in [4]. The result is a "multicoloured" version of Tverberg's theorem [16]. One form of the latter says that any set of $(d+1)t$ points in $\mathbb{R}^d$ can be partitioned into t sets, $S_1, \ldots, S_t$, each of cardinality $d+1$, so that

$$\bigcap_{i=1}^{t} \operatorname{conv}(S_i) \neq \emptyset.$$

In the multicoloured version, the $(d+1)t$ points come in $d+1$ classes $C_1, \ldots, C_{d+1}$, or *colours*, each of cardinality t, and one wants to find "many" pairwise disjoint sets $S_1, \ldots, S_r$, each of cardinality $d+1$, such that every S_i is multicoloured (i.e., $|S_i \cap C_j| = 1$ for every i and j) and

$$\bigcap_{i=1}^{r} \operatorname{conv}(S_i) \neq \emptyset.$$

The question is how large $t = T(r,d)$ must be in order to ensure the existence of such sets $S_1, \ldots, S_r$. In the planar case, one can take $T(r,2) = r$ (see [5] or [12]) and this is clearly best possible. Using tools of algebraic topology, Živaljević and Vrećica [19] show that

$$T(r,d) \leq 2p(r) - 1, \tag{5}$$

where $p(r)$ is the smallest prime that is not smaller than r. It is well known that $p(r) \leq 2r-1$, whence $T(r,d) \leq 4r-3$. We will see later that this is where the value in (4) comes from.

4. Piercing all large convex sets

A set $F \subset \mathbb{R}^d$ is called a *weak ε-net* for X if, for every $Y \subset X$ with $|Y| \geq \varepsilon n$, the intersection $F \cap \operatorname{conv}(Y)$ is nonempty. At a DIMACS workshop in 1990, E. Welzl [17] asked whether there exists a weak ε-net for X whose size depends only on ε and d. This had been proved true in the planar case in [4] before Welzl posed his question; however, the bound $O(\varepsilon^{-1026})$ given in [4] is enormous compared to the bound in the following *weak ε-net theorem.*

Theorem 4.1. *For any $X \subset \mathbb{R}^d$ there exists a weak ε-net F with*

$$|F| \ll_d \varepsilon^{-(d+1)(1-1/s)}.$$

Here, s is the constant s_d of the point selection theorem. In the planar case, (3) gives $s_2 = 3 + \delta$, i.e., a weak ε-net of size $O(\varepsilon^{-(2+\delta')})$ for any positive δ'. We present here a separate argument for the planar case.

Theorem 4.2. *For any $X \subset \mathbb{R}^2$ there exists a weak ε-net of size $7\varepsilon^{-2}$.*

The proof works in any dimension but gives $O(\varepsilon^{-2^{d-1}})$ for the size of the weak ε-net; for $d > 2$, this bound is worse than the bound in Theorem 4.1. Also, in Section 10 we give an algorithm with running time $O(n \log(1/\varepsilon))$, which, for a planar set, yields an ε-net of size $O(\varepsilon^{-4.818\ldots})$.

Define

$$f_d(\varepsilon) = \max\min\{|F| : F \text{ is a weak } \varepsilon\text{-net for } X\},$$

where the maximum is taken over all X satisfying (1). It is clear that $f_d(\varepsilon) \geq 1/\varepsilon$, so

$$\frac{1}{\varepsilon} \leq f_d(\varepsilon) \ll_d \varepsilon^{-(d+1)(1-1/s_d)}.$$

It is not known whether $\varepsilon f_d(\varepsilon)$ is bounded when ε tends to 0.

Weak ε-nets and the discrepancy of triangles. Consider the case when X is a set of n points chosen randomly, independently, and uniformly from the unit square. When ε is fixed and n is large, every triangle of area ε (and contained in the unit square) will contain about εn points of X. Using this one can show that there is a weak ε-net F for X of size $O((1/\varepsilon)\log(1/\varepsilon))$. On the other hand, finding a lower bound for $|F|$ leads to the following old problem of Danzer (see [6] page 285) about irregularities of distributions. How many points are needed to hit every triangle of area ε contained in the unit square?

When X is the vertex set of a regular n-gon in the plane, there is a weak ε-net of size $O((1/\varepsilon)2^{\log^*(1/\varepsilon)})$, where $\log^* m$ denotes the function defined by the recursion $\log^*(2^x) = 1 + \log^* x$ and $\log^* 1 = 0$. This is a result of Capoyleas [8].

We do not know how large the smallest weak ε-net is for a set of n distinct points on the moment's curve $\{(t, t^2, \ldots, t^d) : -\infty < t < \infty\}$ in $\mathbb{R}^d$.

A generalization of Helly's theorem. In [1] the above Theorem 4.1 is combined with some additional tools to prove the following Helly-type result, solving an old problem of Hadwiger and Debrunner.

Theorem. [1] *For every $p \geq q \geq d+1$, there is a (finite) $c = c(p, q, d+1)$ such that the following holds: for every family $\mathcal{K}$ of compact convex sets in $\mathbb{R}^d$ with the property that among any p members of the family some q are pierceable, there is a set F of at most c points in $\mathbb{R}^d$ so that every member of $\mathcal{K}$ contains at least one point of F.*

An easy consequence of Theorem 4.1 is the following result.

Proposition 4.1. *For every $\eta > 0$ and for every integer d, there is a $c = c(\eta, d)$ such that for every probability measure μ on $\mathbb{R}^d$ there is a set F of at most c points in $\mathbb{R}^d$ so that every compact, convex set C of measure $\mu(C) \geq \eta$ contains at least one point of F.*

Let us sketch a proof of this result. By a usual compactness argument, it is enough to prove the proposition for any finite family $\{C_1, \ldots, C_N\}$ of compact convex sets, with $\mu(C_i) \geq \eta$ for every i. Choose points $x_1, \ldots, x_n$ randomly, independently, and according to the distribution μ. Set $X = \{x_1, \ldots, x_n\}$. A straightforward application of the large deviation theorem of Chernoff [9] says that the probability that

$|X \cap C_i| \le \frac{\eta}{2}|X|$ is very small. This shows that, with positive probability, for large enough n we have

$$|X \cap C_i| \ge \frac{\eta}{2}|X| \text{ for every } i = 1, \dots, N.$$

Fix such an X and let F be a weak ε-net for X, where $\varepsilon = \eta/2$. Then, clearly, F intersects every C_i and is of size $O(\eta^{(d+1)(1-1/s)})$, completing the proof.

Another way of proving the proposition is to use the theorem establishing the Hadwiger-Debrunner conjecture. Namely, one can show easily that the family of all convex, compact sets whose measure is at least η satisfies the conditions of that theorem, with $p = \lceil d/\eta \rceil + 1$ and $q = d + 1$. This gives that $c(\eta, d) \le c(p, q, d + 1)$. In fact, the first argument given above gives a better bound on $c(\eta, d)$.

5. Piercing most of the simplices by many points

It turns out that the point selection theorem is closely related to some other results, which we now describe. We say that a set E *misses* $S \in \binom{X}{d+1}$ if $E \cap \operatorname{int} \operatorname{conv}(S) = \emptyset$. (Here, again, X is assumed to satisfy condition (1).) The following *hitting set theorem* asserts the existence of a "small" set E that misses only "few" members of $\binom{X}{d+1}$.

Theorem 5.1. *For every $\eta > 0$ and $X \subset \mathbb{R}^d$, there exists a set $E \subset \mathbb{R}^d$ that misses at most $\eta\binom{n}{d+1}$ simplices of X and has size*

$$|E| \ll_d \eta^{1-s},$$

where s is the constant s_d in the point selection theorem.

In fact we shall show that the hitting set and the point selection theorems are equivalent.

Observe that η may depend on $n = |X|$; for instance, one may take $\eta = n^{-1/s}$, which gives a set E of size $O(n^{1-(1/s)})$ missing at most $O(n^{d+1-1/s})$ simplices of X. This special case of Theorem 4 was proved in [4] for $d = 2$ with $s = 343$.

We emphasize again that the point selection, the hitting set, and the multicoloured Tverberg theorems are equivalent. In fact, the multicoloured Tverberg theorem with $r = d + 1$ implies the point selection theorem with $s = s_d = (T(d + 1, d))^{d+1}$, and the latter implies the multicoloured Tverberg theorem with $T(r, d) \ll_d r$. The equivalence of the point selection and the hitting set theorem is stronger, since it carries over to the exponent $s = s_d$. It would be interesting to know the smallest possible exponent s_d.

Remark on halving planes. As observed in [2] and [4], the point selection theorem (or the hitting set theorem) implies the following upper bound on the number $H_d(X)$ of halving hyperplanes a set $X \subset \mathbb{R}^d$ can have:

$$H_d(X) \ll_d n^{d-1/s_{d-1}}.$$

The simplest way of proving this bound is to use the fact that no line meets more than $\binom{n}{d-1}$ halving simplices. (This was proved in [14] for the planar case, but the argument goes through in $\mathbb{R}^d$ without difficulty.) Then the projection of X, and of the halving simplices of X, to $\mathbb{R}^{d-1}$ gives rise to a family $\mathscr{H}$ in $\mathbb{R}^{d-1}$ (on n points) so that no point is contained in more than $\binom{n}{d-1}$ simplices of $\mathscr{H}$. By the point selection theorem, $\mathscr{H}$ has a pierceable subfamily of size $\gg_d p^{s_{d-1}}\binom{n}{d}$, where $|\mathscr{H}| = p\binom{n}{d}$. So we get $p \ll_d n^{-1/s_{d-1}}$, as required.

6. Proof of the point selection theorem

Here we prove Theorem 2.1. The method is similar to that of [4]. First, we define $V = V(X)$, the *set of crossings* determined by d distinct hyperplanes through the points of X. To this end, let $Q_1, \ldots, Q_d$ be pairwise disjoint d-tuples from X. Their *crossing* is defined as the point of intersection of the hyperplanes aff$(Q_1), \ldots,$ aff(Q_d). Here we assume that X is in a general position, so that any crossing is a well-defined, unique point. To this end, condition (1) can be understood as saying that the coordinates of X are in algebraically independent position. Clearly,

$$|V| = \frac{1}{d!}\binom{n}{d}\binom{n-d}{d}\cdots\binom{n-d(d-1)}{d},$$

so that $n^{d^2} \ll_d |V| \ll_d n^{d^2}$.

Second, we need a theorem of Erdős and Simonovits [11], which is implicit in Erdős [10] as well.

Theorem. [11] *For all positive integers d and t, there exists a positive constant $b = b(d,t)$ such that the following holds: if $\mathscr{H}$ is an arbitrary $(d+1)$-graph on n vertices and $p\binom{n}{d+1}$ edges, where $n^{-t^{-d}} \ll_d p \leq 1$, then $\mathscr{H}$ contains at least*

$$bp^{t^{d+1}}n^{(d+1)t}$$

copies of $K(t,\ldots,t)$, the complete $(d+1)$-partite $(d+1)$-graph with t vertices in each of its $d+1$ vertex classes.

Proof of Theorem 2.1. Consider the family $\mathscr{H} \subset \binom{X}{d+1}$. Then the Erdős-Simonovits theorem implies that $\mathscr{H}$ contains at least

$$bp^{t^{d+1}}n^{(d+1)t}$$

copies of $K(t,\ldots,t)$, provided $n^{-t^{-d}} \ll_d p \leq 1$. Now choose

$$t = T(d+1,d) \leq 2p(d+1) - 1 \leq 4d + 1 \tag{6}$$

from the multicoloured Tverberg theorem of Živaljević and Vrećica, and consider a copy of $K(t,\ldots,t)$ in $\mathscr{H}$. This consists of $d+1$ pairwise disjoint sets $C_1, \ldots, C_{d+1} \subset X$, each of size t, such that, for any $x_1 \in C_1, \ldots, x_{d+1} \in C_{d+1}$, the $(d+1)$-tuple $\{x_1, \ldots, x_{d+1}\}$ belongs to $\mathscr{H}$. By the multicoloured Tverberg theorem, there are $d+1$ pairwise disjoint $(d+1)$-tuples $S_1, \ldots, S_{d+1}$ such that $\bigcap_{i=1}^{d+1}$ conv(S_i) is nonempty. The general position of X implies that $\bigcap_{i=1}^{d+1}$ conv(S_i) is a polytope P with nonempty interior.

The following simple geometric argument shows that there is an S_j and there are subsets $Q_i \subset S_i$ ($i = 1, \ldots, d+1$, $i \neq j$) with $|Q_i| = d$ such that the crossing of the Q_i's lies in int conv(S_j). Consider a facet F of the polytope P. It lies in a (uniquely determined) facet of a (uniquely determined) simplex $\operatorname{conv}(S_i)$, say, $\operatorname{conv}(S_1)$. Thus F lies in $\operatorname{aff}(Q_1)$ for a (uniquely determined) $Q_1 \subset S_1$, where $|Q_1| = d$. Then the $(d-1)$-dimensional polytope $P_1 = \operatorname{aff}(Q_1) \cap \bigcap_{i=2}^{d+1} \operatorname{conv}(S_i)$ has nonvoid $(d-1)$-dimensional interior. So it has a facet F_1 that lies in a facet of one of the simplices $\operatorname{conv}(S_i)$ $(i \geq 2)$, say $\operatorname{conv}(S_2)$. Thus F_1 lies in a hyperplane $\operatorname{aff}(Q_2)$ for a (uniquely determined) d-tuple $Q_2 \subset S_2$. Then the $(d-2)$-dimensional polytope $P_2 = \operatorname{aff}(Q_1) \cap \operatorname{aff}(Q_2) \cap \bigcap_{i=3}^{d+1} \operatorname{conv}(S_i)$ has nonvoid $(d-2)$-dimensional interior, and so on. We end up with a zero-dimensional polytope, i.e., a singleton

$$\{v\} = \operatorname{aff}(Q_1) \cap \ldots \cap \operatorname{aff}(Q_d) \cap \operatorname{conv}(S_{d+1}).$$

Then v is a crossing in the interior of $\operatorname{conv}(S_{d+1})$.

Now we give a lower bound for the number of pairs (S, v) with $S \in \mathscr{H}$, $v \in V$, and $v \in \operatorname{int} \operatorname{conv}(S)$. Such a pair can be identified with the $(d+1)$-tuple of sets $(S, Q_1, \ldots, Q_d)$. As we have seen, every $K(t, \ldots, t)$ contains such a $(d+1)$-tuple with

$$\bigcap_{i=1}^{d} \operatorname{aff}(Q_i) \subset \operatorname{int} \operatorname{conv}(S).$$

A given $(d+1)$-tuple $(S, Q_1, \ldots, Q_d)$ can appear in at most

$$Bn^{t(d+1)-(d^2+d+1)}$$

copies of $K(t, \ldots, t)$. Consequently

$$\left|\{(S, v) \in \mathscr{H} \times V : v \in \operatorname{int} \operatorname{conv}(S)\}\right| \geq \frac{\text{number of copies of } K(t, \ldots, t)}{Bn^{t(d+1)-(d^2+d+1)}} \gg_d p^{t^{d+1}} n^{d^2+d+1}.$$

This shows that there is a crossing v in at least

$$\frac{p^{t^{d+1}} n^{d^2+d+1}}{|V|} \gg_d p^{t^{d+1}} \binom{n}{d+1}$$

simplices of $\mathscr{H}$. Let $\mathscr{H}'$ be the set of those $(d+1)$-tuples of $\mathscr{H}$ whose convex hull contains v. Then, indeed, $\mathscr{H}'$ is pierceable and

$$|\mathscr{H}'| \gg_d p^{t^{d+1}} \binom{n}{d+1}. \tag{7}$$

Here t comes from (6). In the hypergraph theorem we needed $p \gg_d n^{-t^{-d}}$, but (7) holds trivially if this condition is violated, since then

$$p^{t^{d+1}} \binom{n}{d+1} \ll_d 1.$$

□

Remark (1). We deduced the point selection theorem from the multicoloured Tverberg theorem of Živaljević and Vrećica. Now we show, in turn, that the latter follows from the point selection theorem. To see this, take $d+1$ sets $C_1, \ldots, C_{d+1}$ in $\mathbb{R}^d$, each of cardinality t, and set $X = \bigcup_{i=1}^{d+1} C_i$. Define $\mathscr{H}$ to be the complete $(d+1)$-partite $(d+1)$-graph with

$(d+1)$-partition $C_1, \ldots, C_{d+1}$. Then $n = t(d+1)$ and $|\mathscr{H}| = t^{d+1} \gg_d \binom{n}{d+1}$. By the point selection theorem, $\mathscr{H}$ has a pierceable subfamily $\mathscr{H}'$ of size

$$|\mathscr{H}'| \gg_d \binom{n}{d+1} \gg_d t^{d+1}.$$

Consider the largest integer r for which there exist pairwise disjoint $(d+1)$-tuples $S_1, \ldots, S_r$ in $\mathscr{H}'$. Then any other $S \in \mathscr{H}'$ intersects $\bigcup_{i=1}^r S_i$, and the number of such $(d+1)$-tuples is at most $(d+1)rt^d$. Since here we counted $S_1, \ldots, S_r$ as well, we get

$$(d+1)rt^d \geq |\mathscr{H}'| \gg_d t^{d+1},$$

which shows that, indeed, $T(r,d) \leq t \ll_d r$.

Remark (2). We mention further that the proof method of the point selection theorem cannot give a selection exponent s_d smaller than $(d+1)^{(d+1)}$. This is because $T(r,d) \geq r$ implies that $t \geq d+1$ in (6).

7. Proof of the hitting set theorem

Proof of Theorem 5.1. We are going to use a greedy algorithm to produce the hitting set E.

Start with $\mathscr{H} = \binom{X}{d+1}$ and $E = \emptyset$. The algorithm proceeds by choosing a maximal cardinality pierceable subfamily $\mathscr{H}' \subset \mathscr{H}$, together with a point $v \in \bigcap\{\text{int conv}(S) : S \in \mathscr{H}'\}$. Then set $\mathscr{H} = \mathscr{H} \setminus \mathscr{H}'$ and $E = E \cup \{v\}$. We stop as soon as $|\mathscr{H}| \leq \eta\binom{n}{d+1}$. We claim that when the algorithm stops

$$|E| \ll_d \eta^{1-s}.$$

Assume the algorithm produced the sequence of families $\binom{X}{d+1} = \mathscr{H}_0 \supset \mathscr{H}_1 \supset \ldots \supset \mathscr{H}_m$. Denote by k_i the index where

$$|\mathscr{H}_{k_i}| \geq 2^{-i}\binom{n}{d+1} \text{ and } |\mathscr{H}_{k_i+1}| < 2^{-i}\binom{n}{d+1}.$$

It may happen that $k_i = k_{i+1}$, but that will not matter. We know that

$$|\mathscr{H}_{k_i+1} \setminus \mathscr{H}_{k_{i+1}+1}| \leq |\mathscr{H}_{k_i+1}| < 2^{-i}\binom{n}{d+1}.$$

We also know from the point selection theorem that, for $j \geq k_{i+1}$, we have

$$|\mathscr{H}_j \setminus \mathscr{H}_{j+1}| \gg_d (2^{-(i+1)})^s\binom{n}{d+1},$$

since the deleted subfamily $\mathscr{H}_j \setminus \mathscr{H}_{j+1}$ was of maximum cardinality. This shows that

$$k_{i+1} - k_i \ll_d \frac{2^{-i}\binom{n}{d+1}}{2^{-(i+1)s}\binom{n}{d+1}} = 2^{i(s-1)+s}.$$

https://doi.org/10.1017/S0963548300000225 Published online by Cambridge University Press

Since we stop as soon as $2^{-i} \leq \eta$, i.e., $i \geq \lceil \log 1/\eta \rceil$, we get that the basic step of the algorithm is carried out

$$m \leq \sum_{i=0}^{\lceil \log 1/\eta \rceil} (k_{i+1} - k_i) \ll_d \sum_{i=0}^{\lceil \log 1/\eta \rceil} 2^{i(s-1)+s} \ll_d \eta^{1-s}$$

times. This proves the claim. □

Remark. The hitting set theorem implies the point selection theorem. Indeed, let $\mathscr{H} \subset \binom{X}{d+1}$ with $|\mathscr{H}| = p\binom{n}{d+1}$. Set $\eta = p/2$ and let E be a set of cardinality $O(\eta^{-(s-1)})$ missing at most $\eta\binom{n}{d+1}$ simplices of X. Define

$$\mathscr{H}_1 = \{S \in \mathscr{H} : E \cap \operatorname{int} \operatorname{conv} S \neq \emptyset\}.$$

Clearly, $|\mathscr{H}_1| \geq \frac{p}{2}\binom{n}{d+1}$. Since E meets every simplex in $\mathscr{H}_1$, there is a point $v \in E$ that is contained in at least

$$\frac{|\mathscr{H}_1|}{|E|} \gg_d p^s \binom{n}{d+1}$$

simplices of $\mathscr{H}_1$.

8. Weak ε-nets for convex sets in $\mathbb{R}^d$

Proof of Theorem 4.1. First we give a simple algorithm producing a weak ε-net F, of size $\ll_d \varepsilon^{-(d+1)}$.

Start with $F = \emptyset$. Check whether there is a set $Y \subset X$, $|Y| \geq \varepsilon n$ such that F misses all simplices of $\binom{Y}{d+1}$. If there is no such Y, stop. In this case F is a weak ε-net for X. If there is a Y like that, choose a point v common to at least

$$(1 - o(\mathrm{i}))(d+1)^{-d}\binom{|Y|}{d+1}$$

simplices from Y. Such a point exists by (2). Set $F = F \cup \{v\}$.

In each step of the algorithm, the number of missed simplices decreases by at least

$$(1-o(1))(d+1)^{-d}\binom{|Y|}{d+1} \geq (1-o(1))(d+1)^{-d}\binom{\varepsilon n}{d+1} \gg_d \binom{n}{d+1}\varepsilon^{d+1}.$$

Therefore, the algorithm terminates after at most

$$\frac{\binom{n}{d+1}}{\varepsilon^{d+1}\binom{n}{d+1}} \ll_d \varepsilon^{-(d+1)}$$

steps, showing that $|F| \ll_d \varepsilon^{-(d+1)}$.

To get the sharper estimate in the theorem, we apply the previous algorithm, but instead of starting with $F = \emptyset$, we start with $F = E$, where E comes from Theorem 5.1, i.e., E misses at most $\eta\binom{n}{d+1}$ simplices of X and $|E| \ll_d \eta^{1-s}$.

https://doi.org/10.1017/S0963548300000225 Published online by Cambridge University Press

This time the algorithm terminates after

$$(1+o(1))\frac{\eta\binom{n}{d+1}}{(d+1)^{-d}\binom{\varepsilon n}{d+1}} \ll_d \eta\varepsilon^{-(d+1)}$$

steps, producing a weak ε-net F of size $\ll_d |E| + \eta\varepsilon^{-(d+1)} \ll_d \eta^{1-s} + \eta\varepsilon^{-(d+1)}$. The right choice for η is $\varepsilon^{(d+1)/s}$, which gives

$$|F| \ll_d \varepsilon^{-(d+1)+(d+1)/s}.$$

□

9. Weak ε-nets in the plane

Here we prove Theorem 4.2 by an inductive procedure. Let us start with some remarks and definitions.

Given $k \geq 3$ and a finite set $X \subset \mathbb{R}^2$, let $f(X,k)$ denote the minimal size of a weak ε-net for X, where $\varepsilon = k/|X|$, i.e.,

$$f(X,k) = \min\{|F| : F \subset \mathbb{R}^2, \text{ int conv}(Y) \cap F \neq \emptyset \text{ for every } Y \subset X \text{ with } |Y| \geq k\}.$$

Note that this definition is stronger than the original, as here we require that F intersect the interior of conv(C). Consequently, our Theorem 9.1 below is a little stronger than Theorem 4.2. Let $f(n,k)$ be the maximum of $f(X,k)$, where X satisfies (1) with $d = 2$, i.e., no three points from X lie on a line. Obviously, we have $f(k,k) = 1$, and, more generally, $f(n,k) = 1$ if $n \leq k$. In order to bound $f(n,k)$ for small values of k, we shall need a result of Katchalski and Meir [13] claiming that

$$f(n,3) = 2n-5. \tag{8}$$

Theorem 9.1. *$f(n,k) < 7(n/k)^2$ for all $n \geq k \geq 3$.*

Proof. The function $f(n,k)$ is monotonic in the sense that

$$f(n,k) \geq f(n',k')$$

holds for $n \geq n'$ and $k \leq k'$. Since $7n^2/25 > 2n-5$ for all n, relation (8) implies that the theorem holds for $k = 3$, 4, and 5. From now on we suppose that $n \geq k \geq 6$.

Now let X be an n-set. First, find a line L that bisects X into two parts X_1 and X_2 of almost equal size, i.e., $|X_i| = m_i$ with $|m_1 - m_2| \leq 1$. Next we construct a set V such that V intersects int conv(Y) for every $Y \subset X$, $|Y| \geq k$ that has more than ℓ points on both sides of L. (We shall choose $\ell = \lfloor k/6 \rfloor$ later. For the time being we only need $0 \leq \ell \leq (k/2) - 1$.) To this end consider the intersections of L with the line segments connecting $x_1 \in X_1$ to $x_2 \in X_2$. There are $m_1 m_2$ such intersection points, $u(1), \ldots, u(m_1 m_2)$, indexed consecutively on L. (We may suppose that L is in general position with respect to the lines $z_1 z_2$, i.e., all of these intersections are distinct.) Clearly, for any set $Y \subset X$, $|Y| \geq k$ that has at least $\ell + 1$ points on both sides of L, conv(Y) contains at least

$$h = (\ell+1)(k-\ell-1)$$

of the $u(i)$'s. For V, choose a point from L between $u(h-1)$ and $u(h)$, $u(2h-2)$ and $u(2h-1)$, $u(3h-3)$ and $u(3h-2)$, etc. Then

$$|V| = \left\lceil \frac{m_1 m_2}{h-1} \right\rceil - 1.$$

What are those sets $Y \subset X$, $|Y| \geq k$, whose convex hull contains no point from V? They are the Y that have at most ℓ points either in X_2 or in X_1. But such a Y must have at least $(k-\ell)$ points in X_1 (or in X_2, respectively). So it will be enough to find a weak ε_1-net for X_1, where $\varepsilon_1 = (k-\ell)/m_1$ (and a weak ε_2-net for X_2 with $\varepsilon_2 = (k-\ell)/m_2$). These two sets together with V form a weak ε-net for X. Next we apply the induction hypothesis twice, and obtain

$$f(n,k) \leq f(m_1, k-\ell) + f(m_2, k-\ell) + |V| \leq \frac{7(m_1^2 + m_2^2)}{(k-\ell)^2} + \frac{m_1 m_2}{(\ell+1)(k-\ell-1)-1}. \tag{9}$$

Using the facts that $(m_1^2 + m_2^2) \leq (n^2+1)/2$, $m_1 m_2 \leq n^2/4$ and for $k \geq 6$, $\ell = \lfloor k/6 \rfloor$ one has $(k-\ell)^2 \geq (25/36)k^2$ and $(\ell+1)(k-\ell-1) - 1 \geq (5/36)k^2$, we obtain that the right hand side of (9) is at most $(252/50)(n^2+1)/k^2 + (36/20)(n^2/k^2)$. This is at most $6.98(n/k)^2$ for $n \geq k \geq 6$. □

Remark. Without finding the fine structure (the clusters) of the set X, one cannot obtain a smaller ε-net than $\Omega(1/\varepsilon^2)$. This can be seen from the following example. Let C_1, C_2, ..., $C_{2/\varepsilon}$ be disjoint, small circular discs in the plane such that there is no point P lying in three of the regions $\mathrm{conv}(C_i \cup C_j)$, except if all the three contain the same disk C_i. Put $\varepsilon n/2$ points around the centre of each C_i. Then, every ε-net avoiding $\bigcup C_i$ must have at least $\Omega(1/\varepsilon^2)$ points.

10. An efficient algorithm to find weak ε-nets

By applying the results of [15] and [18], one can give an alternative proof of Theorems 4.1 and 4.2 for $d \leq 3$. This proof gives a slightly worse estimate, but has the advantage that it provides an efficient algorithm for constructing the corresponding weak ε-nets. Here is the assertion for the planar case.

Proposition 10.1. *For every set X of n points in the plane and for every $\varepsilon > 0$, there is a weak ε-net of size $O(\varepsilon^{-\log_{4/3} 4})$. Such a net can be found in time $O(n \log(1/\varepsilon))$.*

Proof. Without loss of generality, we may assume that n is a power of 4. By the main result of [15], one can find in time $O(n)$ two intersecting lines l_1 and l_2 so that the number of points in each of the four closed regions to which they partition the plane is at least $n/4$. Let y be the point of intersection of these two lines, and partition X into four pairwise disjoint subsets $X_1, \ldots, X_4$ of cardinality $n/4$ each, where each X_i is completely contained in one of the above closed regions. Observe that if a convex set contains at least one point from each X_i, then it contains y, i.e., $Y = \{y\}$ is a weak 3/4-net for X.

It follows that any convex set that does not contain y misses completely at least one of the sets X_i, and hence, if it contains at least εn points of X, then it contains at least a fraction $(4/3)\varepsilon$ of one of the sets X_i. Therefore, by recursively constructing $(4/3)\varepsilon$-nets in each X_i we conclude that the size $f(\varepsilon)$ of our net satisfies $f(\varepsilon) \leq 1 + 4f(\frac{4}{3}\varepsilon)$ (and $f(\delta) = 1$ for all $\delta \geq 3/4$). This easily gives the bound stated above. The time $t(n, \varepsilon)$ for finding the net in our construction satisfies $t(n, \varepsilon) \leq O(n) + 4t(n/4, \frac{4}{3}\varepsilon)$, which implies that $t(n, \varepsilon) \leq O(n \log (1/\varepsilon))$, completing the proof. □

References

[1] Alon, N. and Kleitman, D. J. (to appear) Piercing convex sets and the Hadwiger-Debrunner (p, q)-problem. *Adv. Math.*

[2] Aronov, B., Chazelle, B., Edelsbrunner, H., Guibas, L. J., Sharir, M. and Wenger, R. (1991) Points and triangles in the plane and halving planes in space. *Discrete Comp. Geom.* **6** 435–442.

[3] Bárány, I. (1982) A generalization of Caratheodory's theorem. *Discrete Math.* **40** 141–152.

[4] Bárány, I., Füredi, Z. and Lovász, L. (1990) On the number of halving planes. *Combinatorica* **10** 175–183.

[5] Bárány, I. and Larman, G. D. (to appear) A coloured version of Tverberg's theorem. *J. London Math. Soc.*

[6] Beck, J. and Chen, W. (1987) Irregularities of distributions. *Cambridge Tracts in Math.* **89**, Cambridge University Press.

[7] Boros, E. and Füredi, Z. (1984) The number of triangles covering the center of an n-set. *Geom. Dedicata* **17** 69–77.

[8] Capoyleas, V. (1991) *An upper bound for weak ε-nets of points in convex position: preliminary version*, (manuscript).

[9] Chernoff, H. (1952) A measure of asymptotic efficiency for tests of an hypothesis based on the sum of observations. *Ann. Math. Statist.* **23** 497–507.

[10] Erdős, P. (1964) On extremal problems of graphs and generalized graphs. *Israel J. Math.* **2** 183–190.

[11] Erdős, P. and Simonovits, M. (1983) Supersaturated graphs and hypergraphs. *Combinatorica* **3** 181–192.

[12] Jaromczyk, J. W. and Światek, G. (1990) The optimal constant for the colored version of Tverberg's theorem. *DIMACS Technical Report 90-6.*

[13] Katchalski, M. and Meir, A. (1988) On empty triangles determined by points in the plane. *Acta Math Hungar.* **51** 323–328.

[14] Lovász, L. (1971) On the number of halving lines. *Annales Univ Sci Budapest, Eötvös* **14** 107–108.

[15] Megiddo, N. (1985) Partitioning with two lines in the plane. *J. Algorithms* **6** 430–433.

[16] Tverberg, H. (1966) A generalization of Radon's theorem. *J. London Math. Soc.* **41** 123–128.

[17] Welzl, E. (1990) Lecture at a DIMACS conference, January 1990.

[18] Yao, F. F. (1983) A 3-space partition and its applications. *Proc. 15th Annual ACM STOC*, ACM Press, New York, 258–263.

[19] Živaljević, R. T. and Vrećica, S. T. (to appear) The coloured Tverberg problem and complexes of injective functions. *J. Comb. Theory, A*