On weak ϵ-nets and the Radon number

Shay Moran[*]   Amir Yehudayoff[†]

Abstract

We show that the Radon number characterizes the existence of weak nets in separable convexity spaces (an abstraction of the Euclidean notion of convexity). The construction of weak nets when the Radon number is finite is based on Helly's property and on metric properties of VC classes. The lower bound on the size of weak nets when the Radon number is large relies on the chromatic number of the Kneser graph. As an application, we prove an amplification result for weak ϵ-nets.

1 Introduction

Weak and strong ϵ-nets were defined by Haussler and Welzl as a tool for fast processing of geometric range queries [20]. They have been consequently studied in many areas, including computational geometry, combinatorics, and machine learning, and they were used in many algorithmic applications, including range searching and geometric optimization.

An ϵ-net is a set that pierces all large sets in a given family of sets. Formally, let μ be a probability distribution over a domain X and let[1] $C \subseteq 2^X$ be a family of sets. A subset S of X is called a *weak ϵ-net* for C over μ if $S \cap c \neq \emptyset$ for every $c \in C$ with $\mu(c) \geq \epsilon$. A subset S is called a *strong ϵ-net* if in addition S is contained in the support of μ. We say that C has weak/strong ϵ-nets of size $\beta = \beta(C, \epsilon)$ if for every distribution μ there is a weak/strong ϵ-net for C over μ of size at most β (we stress that β may depend on ϵ, but not on μ).

To illustrate the difference between weak and strong nets, consider the uniform distribution on n points on the unit circle in the plane $X = \mathbb{R}^2$, and the family C to be all convex hulls of subsets of these n points. Any strong ϵ-net must contain

[1]Here and below we assume that all sets considered are measurable.

at least $(1 - \epsilon)n$ points, but there are weak ϵ-nets of size $O(\alpha(\epsilon)/\epsilon)$, where $\alpha(\cdot)$ is the inverse Ackermann function [5]; using points inside the unit disc allows to use significantly less points. In general, weak nets may be much smaller than strong nets.

The main question we address is under what conditions do weak nets exist. This question for strong ϵ-nets is fairly well understood; the fundamental theorem of statistical learning, which shows that the VC dimension characterizes PAC learnability, also shows that the VC dimension characterizes the existence of strong nets (see [20, 29] and references within). We show that the Radon number characterizes the existence of weak nets in a pretty general setting (viz. separable convexity spaces).

Weak nets

Weak ϵ-nets were mostly studied in the context of discrete and convex geometry, where they are related to several deep phenomena. For example, Alon and Kleitman [6] used weak nets in their famous solution of the (p, q)-conjecture by Hadwiger and Debrunner [16].

Barany, Furedi, and Lovasz [8] showed that convex sets in the plane admit weak nets, and Alon, Barany, Furedi, and Kleitman [3] established a bound of $O(1/\epsilon^2)$ on their size. Recently Rubin improved this bound to $O(1/\epsilon^{3/2+\gamma})$, where $\gamma > 0$ is arbitrarily small [28]. Alon, Barany, Furedi, and Kleitman [3] extended the existence of weak nets for convex sets to all dimensions d; there are weak ϵ-nets of size at most roughly $(1/\epsilon)^{d+1}$. Their proof relies on several results from convex geometry, like Tverberg's theorem and the colorful Caratheodory theorem. Chazelle, Edelsbrunner, Grigni, Guibas, Sharir, and Welzl [11] later improved the bound to at most roughly $(1/\epsilon)^d$. Overall, there are at least three different constructions of weak nets for convex sets. Matousek [26] showed that any weak ϵ-net for convex sets in $\mathbb{R}^d$ must contain at least $\Omega(\exp(\sqrt{d/2}))$ points for $\epsilon \leq 1/50$. Later, Alon [2] proved the first lower bound that is superlinear in $1/\epsilon$; this was later improved to $\Omega\big((1/\epsilon) \log^d(1/\epsilon)\big)$ by Bukh, Matousek, and Nivasch [10]. Ezra [14] constructed weak ϵ-nets for the more restricted class of axis-parallel boxes in $\mathbb{R}^d$. Alon, Kalai, Matousek, and Meshulam [4] defined weak nets in an abstract setting, and asked about combinatorial conditions that yield their existence.

We identify that the existence of weak nets follows from a basic combinatorial property of convex sets, Radon's theorem:

Theorem 1.1 (Radon). *Any set of $d + 2$ points in $\mathbb{R}^d$ can be partitioned into two disjoint subsets whose convex hulls intersect.*

We show that this property alone is sufficient and necessary for the existence of weak nets in a general setting, which we describe next.

It is worth mentioning that Radon's theorem also plays a central role in the context of strong nets. Indeed, it implies that the VC dimension of half-spaces in $\mathbb{R}^d$ is at most $d + 1$, which consequently bounds the VC dimension of many geometrically defined classes.

Convexity spaces

We consider an abstraction of Euclidean convexity that originated in a paper by Levi [24], and defined in the form presented here by Kay and Womble [21]. For a thorough introduction to this subject see the survey by Danzer, Grunbaum, and Klee [13] or the more recent book by van de Vel [30].

A *convexity space* is a pair (X, C) where $C \subseteq 2^X$ is a family of subsets that satisfies[2]:

- $\emptyset, X \in C$.

- C is closed under intersections[3]: $\cap C' \in C$ for every $C' \subseteq C$.

The members of C are called *convex sets*. The *convex-hull* of a set $Y \subseteq X$, denoted by $conv(Y) = conv_C(Y)$, is the intersection of all convex sets $c \in C$ that contain Y. A convex set $b \in C$ is called a *half-space* if its complement is also convex.

We next define the notion of separability, which is an abstraction of the hyperplane separation theorem (and the more general Hahn-Banach theorem). The convexity space (X, C) is *separable*, if for every $c \in C$ and $x \in X \setminus c$ there exists a half-space $b \in C$ so that $c \subseteq b$ and $x \notin b$. It can be verified that (X, C) is separable if and only if every convex set $c \in C$ is the intersection of all half-spaces containing it. This form of separability, as well as other forms, have been extensively studied (e.g. [15, 18, 17, 22, 12]).

Convexity spaces appear in many contexts in mathematics. For instance, the family of closed subsets in a topological space, the subgroups of a given groups, and the subrings of a ring are all examples[4] of convexity spaces. They are also closely related to the notion of π-systems in probability theory. In Section 1.1 below we discuss a few examples of convexity spaces that arise in algebra and combinatorics.

Main results

The combinatorial property that characterizes the existence of weak nets is the Radon number [24, 21], which is an abstraction of Radon's theorem: we say that C *Radon-shatters* a set $Y \subseteq X$ if for every partition of Y into two parts Y_1, Y_2 it

---

[2]Note that van de Vel [30] also requires that the union of an ascending chain of convex sets is convex.

[3]We use the standard notation $\cap C' = \bigcap_{c \in C'} c$.

[4]One should sometimes add the empty set in order to satisfy all axioms of a convexity space.

holds that $conv(Y_1) \cap conv(Y_2) = \emptyset$. The *Radon number* of (X, C) is the minimum number r such that C does not Radon-shatter any set of size r. Radon's theorem states that the Radon number of the space of convex sets in $\mathbb{R}^d$ is at most $d + 2$.

Theorem 1.2. *Let (X, C) be a finite separable convexity space.*

1. *If the Radon number of (X, C) is at most r then it has weak ϵ-nets of size at most $(120r^2/\epsilon)^{4r^2 \ln(1/\epsilon)}$ for every $\epsilon > 0$.*

2. *If the Radon number of (X, C) is more than r then there is a distribution μ over X such that every $\frac{1}{4}$-net for C over μ has size at least $r/2$.*

We often refer to a construction of small weak ϵ-nets as an *upper bound*, and to a proof that no small weak ϵ-nets exists as a *lower bound*. The upper bound in 1 above is quantitively worse than the one for Euclidean convex sets [11], but holds in a more general setting. We do not know what is the optimal bound in this generality.

A possible interpretation of the upper bound is that the existence of weak nets is not directly related to "geometric" properties of the underlying space (as in the various constructions surveyed above). This is somewhat surprising: consider a family C and a distribution μ such that C has no strong ϵ-net with respect to μ; in order to construct a weak ϵ-net we should have a mechanism that suggests "good points" outside the support of μ. In Euclidean geometry there are such natural choices, like "center of mass". We notice that the Radon number provides such a mechanism (see Section 1.2).

We next discuss conditions that are equivalent to the existence of small ϵ-nets. It is convenient to present these equivalences for infinite spaces. Quantitative variants of these statements apply to finite spaces as well.

We use the following standard notion of compactness; a family C is *compact* if for every $C' \subseteq C$ so that $\cap C' = \emptyset$ there is a finite $C'' \subseteq C'$ so that $\cap C'' = \emptyset$. This condition is satisfied e.g. by closed sets in a compact topological space.

Corollary 1.3 (Equivalences)**.** *The following are equivalent for a compact separable convexity space (X, C):*

1. *(X, C) has a finite Radon number.*

2. *(X, C) has weak ϵ-nets of finite size for some $0 < \epsilon < 1/2$.*

3. *(X, C) has weak ϵ-nets of finite size for every $\epsilon > 0$.*

The proof of Corollary 1.3 appears in Section 4.1. It shows that the role of the Radon number in the existence of weak nets is similar to the role of the VC dimension in the existence of strong nets, at least for separable compact convexity spaces.

4

In Section 1.4, we provide an example showing that the compactness assumption in Corollary 1.3 is necessary. We do not know if the separability assumption is necessary.

The implication $2 \Rightarrow 3$ in Corollary 1.3 is an amplification statement for the parameter ϵ in weak nets. In Section 4.2 we give an example showing that the threshold $1/2$ in item 2 is sharp for amplification of weak nets:

Example 1.4. *There is a compact separable convexity space that has weak ϵ-nets of finite size for every $\epsilon > 1/2$ but has no weak ϵ-nets of finite size for $\epsilon < \frac{1}{2}$.*

This demonstrates an interesting difference with strong ϵ-nets, for which there is no such threshold: if C has strong ϵ-nets for some $\epsilon < 1$ then it has finite VC dimension, which implies the existence of strong ϵ-net for all $\epsilon > 0$.

Organization

In Section 1.1 we provide some examples of convexity spaces. In Section 1.2 we outline the construction that leads to the upper bound in Theorem 1.2, in Section 1.3 we outline the lower bound in Theorem 1.2, and in Section 1.4 we provide some examples that demonstrate the necessity of some of our assumptions.

In Section 2 we prove the upper bound, and in Section 3 we prove the lower bound. In Section 4 we prove the characterizations (Corollary 1.3), and in Section 5 we discuss an extension to convexity spaces that are not necessarily separable or compact (like bounded convex sets in $\mathbb{R}^d$). Finally, in Section 6 we conclude the paper and offer some directions for future research.

1.1 Some convexity spaces

We now present a few examples of "non Euclidean" convexity spaces. These examples will be used later on to show that some of our theorems/lemmas are tight in the sense that each premise is necessary. More examples can be found in the book [30].

Example 1 (power set). Let X be a set. Perhaps the simplest convexity space is $(X, 2^X)$. Here every convex set is a half-space and therefore this space is separable. When X is finite, the Radon number of this space is $|X| + 1$. When X is infinite, the Radon number is ∞.

Example 2 (subgroups). Let G be a group with identity e. The space $(G \setminus \{e\}, \{H \setminus \{e\} : H \leq G\})$ of all subgroups of G (with the identity removed) is a convexity space. Here, $Y \subseteq G$ is Radon-shattered, if every two disjoint subsets of Y generate groups whose intersection is $\{e\}$.

Example 3 (cylinders). Let $X = \{0,1\}^n$, and let the C be the family of *cylinders*: a set $c \subseteq \{0,1\}^n$ is called a cylinder if there exists $Y \subseteq [n]$, and $v \in \{0,1\}^Y$ such

that
$$c = \big\{u \in \{0,1\}^n : u|_Y = v\big\}.$$

The size of Y is called the co-dimension of the cylinder. This is a separable convexity space. Its half-spaces are the cylinders with co-dimension 1, and its Radon number is $\Theta(\log n)$.

Example 4 (subtrees). Let $T = (V, E)$ be a finite tree. Consider the convexity space (V, C), where

$$C = \big\{U \subseteq V : \text{the induced subgraph on } U \text{ is connected}\big\}.$$

It is a separable convexity space and its Radon number is at most 4. Theorem 1.2 hence implies the existence of weak nets of size depending only on ϵ (in this case there are elementary constructions of ϵ-nets of size $O(1/\epsilon)$). This example is a special case of *geodesic convexity* in metric spaces (see [30]).

Example 5 (convex lattice sets). Consider the space $(\mathbb{Z}^d, C)$, where C is the family of *convex lattice sets* in $\mathbb{R}^d$; these are sets of the form $K \cap \mathbb{Z}^d$ for some convex $K \subseteq \mathbb{R}^d$.

This is a separable convexity space. Indeed, let $c = K \cap \mathbb{Z}^d \in C$ and $x \in \mathbb{Z}^d \setminus c$. Since $x \notin c$ it follows that $x \notin K$. and therefore there is a half-space in $\mathbb{R}^d$ separating x and K. This half-space induces a half-space in $(\mathbb{Z}^d, C)$.

Onn [27] proved that the Radon number of this space is at most $O(d2^d)$ and at least 2^d. Our results imply that weak ϵ-nets exist in this case as well.

Note that the family of half-spaces has VC dimension $d + 1$, which is much smaller than the Radon number. Thus, Theorem 1.6 (which we state in the next subsection) gives better bounds on the size of the ϵ-net than Theorem 1.2.

Example 6 (linear extensions of posets). Let Ω be a set. For a partial order P on Ω, let $c(P) \subseteq X$ denote the set of all linear orders that extend P. Fix a partial order P_0, and consider the family C of all sets of the form $c(P)$, where P is a partial order that extends P_0. The space $\big(c(P_0), C\big)$ is a separable convexity space whose half-spaces correspond to partial orders defined by taking two P_0-incomparable elements $x, y \in \Omega$ and extending P_0 by setting $x < y$.

1.2 Upper bound

From Radon to Helly and VC

Let (X, C) be a separable convexity space and let B denote the family of half-spaces of C (the notation B is chosen to reflect that B generates all convex sets in C by taking intersection, and hence can be seen as a *basis*).

We first observe that the Radon number is an upper bound on the Helly number and the VC dimension of B. The *Helly number* of a family B is the minimum

number h such that every finite[5] $B' \subseteq B$ with $\cap B' = \emptyset$ contains a subfamily B'' with at most h sets such that already the intersection of the sets in B'' is empty. Helly theorem states that the Helly number of half-spaces in $\mathbb{R}^d$ is at most $d + 1$. The *VC dimension* of $B \subseteq 2^X$ is the supremum over v for which there exists $Y \subseteq X$ of size $|Y| = v$ so that for every $Z \subseteq Y$ there is a member of the family that contains Z and is disjoint from $Y \setminus Z$.

Lemma 1.5. *Let B the family of half-spaces of a separable convexity space C. If the Radon number of C is r then the VC dimension and the Helly number of B are less than r.*

The bound on the Helly number follows from a result by Levi [24], and the bound on the VC dimension is straightforward. The proof appears in Appendix A.

Lemma 1.5 reduces the construction of weak nets for separable convexity spaces to the following, more general construction.

Theorem 1.6. *Let X be a set, let $B \subseteq 2^X$ be a compact family with VC dimension v and Helly number h, and let $B^\cap$ denote the family generated by taking arbitrary intersections of members of B. Then $B^\cap$ has weak ϵ-nets of size at most $\beta < (120h^2/\epsilon)^{4hv\ln(1/\epsilon)}$ for every $\epsilon > 0$.*

We next give an overview of the proof of Theorem 1.6; the complete proof appears in Section 2.

Outline of construction

The construction of weak nets underlying Theorem 1.6 is short and simple. We want to pierce all convex sets $c \in C$ such that $\mu(c) \geq \epsilon$. We distinguish between two cases. The simpler case is when c can be written as the intersection of half-spaces, each of which has μ-measure more than $1 - 1/h$. By Helly's property, it follows that there is a single point that pierces all such c's. In the complementary case, when c can not be written in this way, we use Haussler's packing lemma [19] and show that there is a small collection $A \subseteq B$ such that conditioning μ on a single $a \in A$ increases the measure of c by a factor of at least $1 + 1/(2h)$. The size of A can be bounded from above in terms of the VC dimension of B. So, constructing a set that pierces all c's with measure at least ϵ is reduced to constructing a bounded number of nets for larger density $\epsilon' \geq (1 + 1/(2h))\epsilon$.

To summarize, the Helly number yields the theorem for ϵ close to 1, and the VC dimension allows to keep increasing the density until the Helly number becomes relevant.

Going back to the discussion after Theorem 1.2 concerning the mechanism that suggests "good points" we see that this mechanism is based on the Helly property

---

[5]The finiteness assumption can be removed when B is compact.

(roughly speaking, the Helly property is a mechanism that given a collection of sets outputs a point).

1.3 Lower bound

The following lemma is a slight generalization of the lower bound in Theorem 1.2 (we do not assume separability or compactness, and replace $1/4$ by any $\epsilon > 0$).

Lemma 1.7. *Let (X, C) be a convexity space. If the Radon number of C is greater than $r > 0$ then there is a distribution μ on X so that every weak ϵ-net for C over μ has size at least $(1 - 2\epsilon)r$.*

This gives a non-trivial lower bound as long as $\epsilon < 1/2$. Example 1.4 shows that this is sharp in the sense that when $\epsilon > 1/2$ there is no lower bound that tends to infinity with the Radon number.

The proof of Lemma 1.7, as well as a finer distribution dependent lower bound, appear in Section 3. The proof is essentially by reduction to the chromatic number of Kneser graphs.

1.4 The necessity of assumptions

Assumptions in Theorem 1.6

We first show that if either of the assumptions of having bounded VC dimension or of having bounded Helly number is removed then Theorem 1.6 ceases to hold.

To see why bounded Helly number is necessary, let X be any finite set and set $B = \{X\} \cup \{X \setminus \{x\} : x \in X\}$. The VC dimension of B is 1, but any subset of X can be represented as an intersection of members of B. Thus, $B^{\cap} = 2^X$, which does not have weak ϵ-nets of size which is independent of $|X|$.

To see why bounded VC dimension is necessary, consider the convexity space (X, C) of cylinders (Example 3 in Section 1.1). Here, $X = \{0, 1\}^n$, C is the family of cylinders, and the family of half-spaces B consists of cylinders with co-dimension 1: $B = B_0 \cup B_1$ with

$$B_t = \big\{\{x \in X : x_i = t\} : i \in [n]\big\}.$$

The Helly number of B is 2, since an intersection of half-spaces is empty if and only if two complementing cylinders with co-dimension 1 participate in it.

The following claim gives a lower bound on weak $\frac{1}{4}$-nets for C over the uniform distribution μ over X.

Claim 1.8. *Every weak $(1/4)$-net for C over μ has size at least $\log n$.*

Proof. $S \subseteq X$ pierces every cylinder with measure $1/4$ only if for every $i \neq j$ in $[n]$ there is $x \in S$ with $x_i = 0$ and $x_j = 1$. Now, consider the mapping from $[n]$ to $\{0, 1\}^S$, which maps $i \in [n]$ to $(x_i)_{x \in S}$. By the above, this mapping is one-to-one, and in particular $2^{|S|} = \big|\{0, 1\}^S\big| \geq n$. $\qquad\square$

Assumptions in Corollary 1.3

We now describe an example showing that the compactness assumption in Corollary 1.3 is necessary (a related example appears in [9, 31] in the context of strong ϵ-nets). Let $X = \omega_1$ be the first uncountable ordinal, and C be the family of all intervals in the well-ordering of X (a set $I \subseteq X$ is an interval if whenever $a, b \in I$ and $a \leq x \leq b$ then also $x \in I$). The space (X, C) is separable with Radon number 3, but is not compact.

We claim that it does not have finite weak ϵ-nets, even for $\epsilon = 1$. Indeed, let μ be the probability distribution defined over the σ-algebra generated by countable subsets of X and assigns every countable subset of X measure 0. Every interval is either countable or has a countable complement, and is therefore measurable.

We now claim that there is no finite $S \subseteq X$ that pierces all intervals of measure 1. Indeed, let S be finite, and let m be the maximum element in S. The interval $\{x \in X : x > m\}$ has measure 1 but is not pierced by S.

2 Proof of upper bound

Here we construct weak ϵ-nets when the Helly number and the VC dimension of the half-spaces are bounded (Theorem 1.6). The property of VC classes that we use is the following packing lemma due to Haussler [19].

Theorem 2.1 (Haussler). *Let $B \subseteq 2^X$ be a class of VC dimension v. For every distribution μ on X and for every $\delta > 0$, there is $A \subseteq B$ of size $|A| \leq (4e^2/\delta)^v$ such that for every $b \in B$ there is $a \in A$ with $\mu(a \Delta b) \leq \delta$.*

Haussler's stated the lemma in a dual way; the number of disjoint balls of a given radius in a VC class is small. Haussler's proof is elaborate, but a weaker bound can be proved fairly easily. Indeed, consider a finite set A so that $\mu(a \Delta a') > \delta$ for all $a \neq a'$ in A. Let $x_1, \ldots, x_m$ be m independent samples from μ for $m \geq 2 \log(|A|)/\delta$. Let $Y = \{x_1, \ldots, x_m\}$, and let $A|_Y = \{a \cap Y : a \in A\}$. On one hand, the Sauer-Shelah-Perles lemma implies that $A|_Y$ is small: $|A|_Y| \leq (em/v)^v$. On the other hand, by the union bound, $|A|_Y| = |A|$ with positive probability. This implies that A is small.

Proof of Theorem 1.6. We start by focusing on the set $B_0 = \{b \in B : \mu(b) > 1 - 1/h\}$. By the union bound, every h members of B_0 intersect. Since $B_0 \subseteq B$ is compact with Helly number h, there is a single point $x_0 \in X$ that pierces all sets in B_0.

Let $0 < \epsilon < 1$. We construct the ϵ-net by induction on $N(\epsilon)$, which is defined to be the minimum integer n such that $\epsilon\big(1 + 1/(2h)\big)^n > 1 - 1/h$.

Induction base

If $N(\epsilon) = 0$, define the piercing set $S = S(\mu, \epsilon)$ as

$$S = \{x_0\},$$

where x_0 is the point that pierces all half-spaces in B_0. Indeed, S is an ϵ-net as every $c \in C$ with $\mu(c) \geq \epsilon > 1 - 1/h$ is the intersection of half-spaces from B_0, so x_0 pierces c as well.

Induction step

Let $1 > \epsilon > 0$ such that $N(\epsilon) > 0$. We construct the piercing set $S = S(\mu, \epsilon)$ as follows: Set $\delta = \epsilon/(2h)^2$ and pick some $A \subseteq B$ as in Theorem 2.1. Also set $\epsilon' = (1 + 1/(2h))\epsilon$. Note that $N(\epsilon') = N(\epsilon) - 1$. By induction, for each $a \in A$ with $\mu(a) > 0$, pick a piercing set $S_a = S(\mu|_a, \epsilon')$, where $\mu|_a$ denotes the distribution μ conditioned on a. Finally, let

$$S = \{x_0\} \cup \bigcup_{a \in A : \mu(a) > 0} S_a.$$

It remains to prove that S satisfies the required properties.

$\underline{S \text{ is piercing:}}$ Let $c \in C$ with $\mu(c) \geq \epsilon$. If c is generated[6] by B_0 then x_0 pierces c. Otherwise, there is some $b \in B$ with $\mu(b) \leq 1 - 1/h$ that contains c. Pick $a \in A$ such that $\mu(a \Delta b) \leq \delta$. Since $\mu(b) \geq \epsilon$ and $\mu(a \Delta b) \leq \delta$,

$$\mu(a) \geq \mu(b) - \mu(b \setminus a) \geq \epsilon - \delta > 0,$$

which means that $\mu|_a$ is well-defined. We claim that $S(\mu|_a, \epsilon')$ pierces c. To this end, it suffices to show that $\mu|_a(c) \geq \epsilon'$:

$$
\begin{aligned}
\mu|_a(c) &= \frac{\mu(c \cap a)}{\mu(a)} \\
&= \frac{\mu(c) - \mu(c \cap (b \setminus a))}{\mu(a)} & (c \subseteq b) \\
&\geq \frac{\mu(c) - \mu(b \setminus a)}{\mu(a)} \\
&\geq \frac{\epsilon - \delta}{1 - 1/h + \delta} & (\mu(c) \geq \epsilon, \mu(b) \leq 1 - 1/h, \mu(a \Delta b) \leq \delta) \\
&\geq \frac{\epsilon(1 - 1/(2h)^2)}{1 - 1/(2h)} & (\delta = \epsilon/(2h)^2) \\
&= \epsilon'.
\end{aligned}
$$

---

[6]I.e. c can be presented as an intersection of sets from B_0.

<u>S is small:</u> Let $\beta(n)$ denote the maximum possible size of S for ϵ with $N(\epsilon) = n$. The argument above yields

$$\beta(n) \leq 1 + (4e^2/\delta)^v \cdot \beta(n-1) = 1 + (16e^2h^2/\epsilon)^v \cdot \beta(n-1).$$

Since $\beta(0) = 1$, we get

$$\beta(n) \leq (120h^2/\epsilon)^{vn}.$$

Finally, since $N(\epsilon) \leq 4h\ln(1/\epsilon))$, we get the bound

$$|S(\mu, \epsilon)| \leq (120h^2/\epsilon)^{4hv\ln(1/\epsilon)}.$$

$\square$

3 Proof of lower bound

The proof of Lemma 1.7 is based on the following distribution-dependent lower bound on the size of weak nets:

Lemma 3.1. *Let C be a family of subsets over a domain X and let μ be a distribution on X. For $\epsilon > 0$ define a graph $G = G(\mu, \epsilon)$ whose vertices are the sets $c \in C$ such that $\mu(c) \geq \epsilon$, and two sets are connected by an edge if and only if they are disjoint. Then, every weak ϵ-net for C over μ has size at least the chromatic number of G, which is denoted by $\chi(G)$.*

Proof. Let S be a set that pierces all $c \in C$ with $\mu(c) \geq \epsilon$. Define a coloring of G by assigning to every vertex c an element $x \in S \cap c$. This is a proper coloring of G, since if $\{c, c'\}$ is an edge of G then c and c' are disjoint and therefore can not be pierced by the same element of S. $\square$

Lemma 3.1 is tight whenever the class C has Helly number 2 (like in examples 3 and 4 in Section 1.1). Indeed, consider an optimal coloring of $G(\mu, \epsilon)$. Every color class is an independent set in G (which means that every two sets in it have a non-empty intersection). So, when the Helly number is 2, each color class can be pierced by a single element, and we get a piercing set of size $\chi(G)$.

The proof Lemma 1.7 thus reduces to a lower bound on the chromatic number of the relevant graph, which in our case contains a copy of the Kneser graph. The *Kneser graph $KG_{n,k}$* is the graph whose vertices correspond to the k-element subsets of a set of n elements, and where two vertices are adjacent if and only if the two corresponding sets are disjoint.

Lovasz [25] proved Kneser's conjecture on the chromatic number of this graph (this proof is considered seminal in the topological method in combinatorics):

11

Theorem 3.2 (Lovasz)**.** *The chromatic number of $KG_{n,k}$ is $n - 2k + 2$.*

We actually do not need the full strength of Lovasz's result. A lower bound of the form $\chi(KG_{n,n/4}) \geq n/10$ suffices for deducing the equivalence between the existence of weak nets and finite Radon number (in fact, even much weaker bounds suffice). Noga Alon informed us that for this range of the parameters there is a short and elementary proof [1]. Since this argument does not appear in the literature and may be useful elsewhere, we include Alon's proof in Section B.

Proof of Lemma 1.7. Since the Radon number is greater than r, it follows that there is a set $Y \subseteq X$ of size r that is Radon-shattered by C. Pick μ to be the uniform distribution over Y. By Lemma 3.1 it suffices to show that $\chi(G(\mu, \epsilon)) \geq (1 - 2\epsilon)r$. This follows by noticing that since Y is Radon-shattered, it follows that the subgraph of $G(\mu, \epsilon)$ induced by the vertices $conv(Z)$, for $Z \subseteq Y$ is of size $\lceil \epsilon r \rceil$, is isomorphic to $KG_{r, \lceil \epsilon r \rceil}$. ◻

4 Proof of equivalences

4.1 The existence of weak nets

Proof of Corollary 1.3. Let (X, C) be a compact separable convexity space.

$1 \Rightarrow 3$. By Lemma 1.5, the family B of half-spaces of C has finite VC dimension and Helly number. Theorem 1.6 now implies that $B^{\cap} = C$ has finite weak ϵ-nets for every $\epsilon > 0$.

$3 \Rightarrow 2$. Obvious.

$2 \Rightarrow 1$. Assume that C has weak ϵ-nets of size $\beta = \beta(\epsilon) < \infty$ for some $\epsilon < 1/2$. By Lemma 1.7, the Radon number of (X, C) is at most $\frac{\beta}{1 - 2\epsilon}$. ◻

4.2 The threshold $1/2$ is sharp

Here we describe Example 1.4. Let $X = \{0, 1\}^{\mathbb{N}}$ be the Cantor space, and let C be the family of all cylinders; recall that $c \subseteq X$ is a cylinder if there exist $Y \subseteq \mathbb{N}$ and $u \in \{0, 1\}^Y$ such that $c = \{v \in X : v|_Y = u\}$. The space (X, C) is a separable convexity space. It is also compact (this follows e.g. from Tychonoff's theorem).

We claim that (X, C) has ϵ-nets of size 1 for every $\epsilon > 1/2$. This follows since for every distribution μ the family $\{c \in C : \mu(c) > 1/2\}$ is intersecting[7], and since C has Helly number 2. Hence, $\cap\{c \in C : \mu(c) > 1/2\} \neq \emptyset$ for all μ, as claimed.

---

[7]A family of sets is *intersecting* if every two members of it intersect.

It remains to show that (X, C) has no finite weak ϵ-nets for $\epsilon < 1/2$. By Corollary 1.3, it suffices to consider the case $\epsilon = 1/4$. Let μ be the Bernoulli measure on the Cantor space; namely the infinite product of uniform measure on $\{0, 1\}$. Now, $S \subseteq X$ is a weak ϵ-net over μ if and only if it intersects every cylinder with co-dimension 2. In particular for every $i \neq j$ in $\mathbb{N}$ there must be $x \in S$ with $x_i = 0$ and $x_j = 1$. By the proof of Claim 1.8 it follows that such an S must be infinite.

5   An extension

Consider the space of bounded closed convex sets in $\mathbb{R}^d$. The corresponding convexity space in not separable nor compact. Nevertheless, our results extend to this space as well.

The following variants of separability and compactness suffice. A convexity space is called *locally-separable* if there is a set $B \subseteq C$ so that $C = B^\cap$ and for every finite $Y \subseteq X$ and for every $b \in B$ there is $\bar{b} \in B$ so that $b \cap Y$ and $\bar{b} \cap Y$ form a partition of Y. Every separable space is locally-separable, but there are convexity spaces which are locally separable but not separable (like the space of bounded convex sets in $\mathbb{R}^d$). A convex set $c \in C$ is called *compact* if the restricted convexity space $(c, \{c' \cap c : c' \in C\})$ is compact. The *Radon number* of c is the Radon number of the space $(c, \{c' \cap c : c' \in C\})$.

Theorem 5.1. *Let (X, C) be a locally separable convexity space such that there exists a chain $c_1 \subseteq c_2 \subseteq \ldots$ of compact convex sets each of which has Radon number at most r such that $\bigcup_i c_i = X$. Then (X, C) has finite weak ϵ-nets for every $\epsilon > 0$.*

Theorem 1.2 and its proof apply for locally-separable and compact convexity spaces. Theorem 5.1 therefore follows by applying it to a c_i in the chain such that $\mu(c_i) \geq 1 - \epsilon/2$.

6   Future research

We showed that for compact separable convexity spaces, the existence of weak nets is equivalent to having a finite Radon number. We now suggest several directions for future research.

One interesting direction is to find a characterization that is valid even more generally (i.e. for families that are not necessarily separable convexity spaces). It is worth noting in this context that the definition of the Radon number can be extended to arbitrary families. One may extend the definition of Radon-shattering as follows: A family $C \subseteq 2^X$ *Radon-shatters* the set $Y \subseteq X$ if for every partition of Y into two parts Y_1, Y_2 there are two disjoint sets $c_1, c_2 \in C$ such that $c_1 \cap Y = Y_1$ and $c_2 \cap Y = Y_2$.

This extension of the Radon number does not characterize the existence of weak nets for arbitrary families. For instance, let $C = \{Y \subseteq [n] : |Y| > n/2\}$. The Radon number is 2 since every two sets in C intersect, but every weak $\frac{1}{2}$-net over the uniform distribution has size at least $n/2 - 1$. However, this family C is not convex (closed under intersection), which is a crucial property in our work. It may also be interesting to fully understand the role of convexity in the context of weak nets.

Alon et al. [4] studied weak nets and the (p, q)-property in an abstract setting and described connections to fractional Helly properties. It may further be interesting to investigate which other combinatorial properties of convex sets apply in more general settings.

An additional question that comes to mind is the dependence of the size of weak ϵ-nets on ϵ. In this direction, Bukh, Matousek and Nivasch proved an $\Omega\left(\frac{1}{\epsilon}\log^{d-1}\frac{1}{\epsilon}\right)$ lower bound on the size of weak ϵ-nets for convex sets in $\mathbb{R}^d$ [10], and recently Rubin [28] proved an upper bound of roughly $O(1/\epsilon^{3/2})$ in $\mathbb{R}^2$ (which improves upon the general bound in $\mathbb{R}^d$ of roughly $O(1/\epsilon^d)$ by [11]). Proving tight bounds on the size of weak ϵ-nets is a central open problem in this area. The general framework developed here may be useful in proving stronger lower bounds.

Acknowledgements

References

[1] N. Alon. Private communication.

[2] N. Alon. A non-linear lower bound for planar epsilon-nets. *Discrete & Computational Geometry*, 47(2):235–244, 2012.

[3] N. Alon, I. Bárány, Z. Füredi, and D. J. Kleitman. Point selections and weak ε-nets for convex hulls. *Combinatorics, Probability and Computing*, 1(03):189–200, 1992.

[4] N. Alon, G. Kalai, J. Matoušek, and R. Meshulam. Transversal numbers for hypergraphs arising in geometry. *Advances in Applied Mathematics*, 29(1):79–101, 2002.

[5] N. Alon, H. Kaplan, G. Nivasch, M. Sharir, and S. Smorodinsky. Weak ϵ-nets and interval chains. *J. ACM*, 55(6):28:1–28:32, 2008.

[6] N. Alon and D. J. Kleitman. Piercing convex sets and the hadwiger-debrunner (p, q)-problem. *Advances in Mathematics*, 96(1):103–112, 1992.

[7] N. Alon and J. H. Spencer. *The probabilistic method.* John Wiley & Sons, Inc., 2000.

[8] I. Bárány, Z. Füredi, and L. Lovász. On the number of halving planes. *Combinatorica*, 10(2):175–183, 1990.

[9] A. Blumer, A. Ehrenfeucht, D. Haussler, and M. K. Warmuth. Learnability and the Vapnik-Chervonenkis dimension. *J. Assoc. Comput. Mach.*, 36(4):929–965, 1989.

[10] B. Bukh, J. Matoušek, and G. Nivasch. Lower bounds for weak epsilon-nets and stair-convexity. *Israel Journal of Mathematics*, 182(1):199–228, 2011.

[11] B. Chazelle, H. Edelsbrunner, M. Grigni, L. Guibas, M. Sharir, and E. Welzl. Improved bounds on weak ε-nets for convex sets. In *STOC*, pages 495–504, 1993.

[12] V. Chepoi. Separation of two convex sets in convexity structures. *Journal of Geometry*, 50(1):30–51, 1994.

[13] L. Danzer, B. Grünbaum, and V. Klee. *Helly's Theorem and Its Relatives.* Proceedings of symposia in pure mathematics: Convexity. American Mathematical Society, 1963.

[14] E. Ezra. A note about weak epsilon-nets for axis-parallel boxes in d-space. *Inf. Process. Lett.*, 110(18-19):835–840, 2010.

[15] B. Grünbaum and T. S. Motzkin. On components in some families of sets. *Proceedings of the American Mathematical Society*, 12(4):607–613, 1961.

[16] H. Hadwiger and H. Debrunner. Über eine variante zum hellyschen satz. *Archiv der Mathematik*, 8(4):309–313, 1957.

[17] P. Hammer. *Semispaces and the Topology of Convexity.* 1961.

[18] P. C. Hammer. Maximal convex sets. *Duke Math. J.*, 22(1):103–106, 03 1955.

[19] D. Haussler. Sphere packing numbers for subsets of the boolean n-cube with bounded vapnik-chervonenkis dimension. *Journal of Combinatorial Theory, Series A*, 69(2):217–232, 1995.

[20] D. Haussler and E. Welzl. Epsilon-nets and simplex range queries. *Discrete & Computational Geometry*, 2:127–151, 1987.

[21] D. Kay and E. W. Womble. Axiomatic convexity theory and relationships between the carathéodory, helly, and radon numbers. *Pacific Journal of Mathematics*, 38(2):471–485, 1971.

[22] V. L. Klee. The structure of semispaces. *Mathematica Scandinavica*, 4(1):54–64, 1956.

[23] D. J. Kleitman. Families of non-disjoint subsets. *J. Combinatorial Theory*, (1):153–155, 1966.

[24] F. W. Levi. On helly's theorem and the axioms of convexity. *J. Indian Math. Soc*, 15:65–76, 1951.

[25] L. Lovász. Kneser's conjecture, chromatic number, and homotopy. *Journal of Combinatorial Theory, Series A*, 25(3):319–324, 1978.

[26] J. Matoušek. A lower bound for weak epsilon-nets in high dimension. *Discrete & Computational Geometry*, 28(1):45–48, 2002.

[27] S. Onn. On the geometry and computational complexity of radon partitions in the integer lattice. *SIAM J. Discrete Math.*, 4(3):436–447, 1991.

[28] N. Rubin. An improved bound for weak epsilon-nets in the plane. In *FOCS*, pages 224–235. IEEE Computer Society, 2018.

[29] S. Shalev-Shwartz and S. Ben-David. *Understanding machine learning: From theory to algorithms*. Cambridge university press, 2014.

[30] M. van de Vel. *Theory of Convex Structures*, volume 50 of *North-Holland mathematical library*. North-Holland, 1993.

[31] R. S. Wenocur and R. M. Dudley. Some special vapnik-chervonenkis classes. *Discrete Mathematics*, 33(3):313–318, 1981.

A Radon, Helly and VC

Here we prove that the Radon number bounds from above both the Helly number and the VC dimension (Lemma 1.5). The proof follows from the following two claims. Levi [24] proved that

Claim A.1. *Let C be a convexity space. If the Radon number of C is r then its Helly number is smaller than r.*

Proof (for completeness). Let $C' \subseteq C$ be a finite family so that $\bigcap_{c \in C'} c = \emptyset$. Let $K \subseteq C'$ be a minimal subfamily so that $\bigcap_{c \in K} c = \emptyset$. Assume towards a contradiction that $|K| \geq r$. Minimality implies that for each $k \in K$ we have $C_k := \bigcap_{c \in K \setminus \{k\}} c \neq \emptyset$. Let $x_k \in C_k$. The x_k's must be distinct (otherwise $\bigcap_{c \in K} c \neq \emptyset$). Thus, there is a partition of $\{x_k : k \in K\}$ to two parts Y_1, Y_2 such that $conv(Y_1) \cap conv(Y_2) \neq \emptyset$. But

$$conv(Y_1) \subseteq \bigcap_{k \in K : x_k \in Y_2} k \quad \text{and} \quad conv(Y_2) \subseteq \bigcap_{k \in K : x_k \in Y_1} k,$$

by construction. This is a contradiction, so $|K| < r$. $\qquad\square$

We observe that

Claim A.2. *Let C be a convexity space and B be its half-spaces. If the Radon number of C is r then the VC dimension of B is smaller than r.*

Proof. Let $Y \subseteq X$ be of size r. The set Y can thus be partitioned to Y_1, Y_2 so that $conv(Y_1) \cap conv(Y_2) \neq \emptyset$. Assume, towards a contradiction, that there is $b \in B$ so that $b \cap Y = Y_1$. Since B consists of half-spaces[8], there is $\bar{b} \in B \subseteq C$ so that $Y_2 = \bar{b} \cap Y$. This implies that $conv(Y_1) \cap conv(Y_2) = \emptyset$, which is a contradiction. Thus, for all $b \in B$ we have $b \cap Y \neq Y_1$ which means that the VC dimension is less than $|Y| = r$. $\qquad\square$

B   The chromatic number of the Kneser graph

Here we prove a lower bound on the chromatic number of the Kneser graph (which is weaker than Lovasz's). We follow an argument of Alon [1], who informed us that a similar argument was independently found by Szemeredi. We focus on the following case, but the argument applies more generally.

Theorem B.1. *For n be divisible by 4 we have $\chi(KG_{n,n/4}) > n/10$.*

The first step in the proof is the following lemma proved by Kleitman [23]. A family $F \subseteq 2^X$ is called intersecting if $f \cap f' \neq \emptyset$ for all $f, f' \in F$.

Lemma B.2. *If $F_1, \ldots, F_s \subset 2^{[n]}$, where each F_i is intersecting, then*

$$\left| \bigcup_{i \in [s]} F_i \right| \leq 2^n - 2^{n-s}.$$

The lemma can proved by induction on s. The case $s = 1$ just says that an intersecting family has size at most 2^{n-1}. The induction step is based on correlation of monotone events (for more details see, e.g. [7]).

---

[8]Here we use a more general definition of half-spaces, as in the definition of locally-separable in Section 5.

Proof of Theorem B.1. Consider a proper coloring of $KG_{n,n/4}$ with s colors. Let $V_1, \ldots, V_s$ be the partition of the vertices to color classes. Each V_i is an intersecting family. Let F_i be the family of sets $u \subseteq [n]$ that contain some set in V_i. Each F_i is also intersecting. By the lemma above,

$$2^n - \left| \bigcup_{i \in [s]} F_i \right| \geq 2^{n-s}.$$

On the other hand, the complement of $\bigcup_{i \in [s]} F_i$ is of size less than $\sum_{k=0}^{n/4} \binom{n}{k} \leq 2^{nH(1/4)}$, where $H(p) = -p \log(p) - (1-p) \log(1-p)$ is the binary entropy function. Hence,

$$s > n(1 - H(1/4)) \geq n/10.$$

$\square$