Journal of Machine Learning Research 20 (2019) 1-8 Submitted 10/18; Published 3/19

Tight Lower Bounds on the VC-dimension of Geometric Set Systems

Mónika Csikós MONIKA.CSIKOS@ESIEE.FR
Nabil H. Mustafa MUSTAFAN@ESIEE.FR
Université Paris-Est, LIGM, Equipe A3SI, ESIEE Paris,
Cité Descartes - 2 boulevard Blaise Pascal,
93162 Noisy-le-Grand Cedex, France.

Andrey Kupavskii KUPAVSKII@YANDEX.RU
Laboratory of Advanced Combinatorics and Networks Applications,
Moscow Institute of Physics and Technology,
9, Institutskiy per., Dolgoprudniy, 141701, Russia and
Mathematical Institute, University of Oxford,
Woodstock Rd, Oxford OX2 6GG, UK

Editor: Alexander Rakhlin

Abstract

The VC-dimension of a set system is a way to capture its complexity and has been a key parameter studied extensively in machine learning and geometry communities. In this paper, we resolve two longstanding open problems on bounding the VC-dimension of two fundamental set systems: k-fold unions/intersections of half-spaces and the simplices set system. Among other implications, it settles an open question in machine learning that was first studied in the foundational paper of Blumer et al. (1989) as well as by Eisenstat and Angluin (2007) and Johnson (2008).

Keywords: VC-dimension, union of concepts, intersection of concepts, combinatorial problems, PAC learning

1. Introduction

Let $(X, \mathcal{R})$ be a *set system*, where X is a set of elements and $\mathcal{R}$ is a set of subsets of X. In the theory of learning, the elements of $\mathcal{R}$ are also called *concepts*, and $\mathcal{R}$ is called a *concept class* on X. For any integer $k \geq 2$, define the *k-fold union* of $\mathcal{R}$ as the set system induced on X by the ranges

$$\mathcal{R}^{k\cup} = \{R_1 \cup \cdots \cup R_k \colon R_1, \ldots, R_k \in \mathcal{R}\}.$$

Similarly, one can define the *k-fold intersection* of $\mathcal{R}$, denoted by $\mathcal{R}^{k\cap}$, as the set system consisting of all subsets derived from the common intersection of at most k sets of $\mathcal{R}$. Note that as the subsets $R_1, \ldots, R_k$ need not necessarily be distinct, we have $\mathcal{R} \subseteq \mathcal{R}^{k\cup}$ and $\mathcal{R} \subseteq \mathcal{R}^{k\cap}$. Analogously, the *$k$-fold symmetric difference* of $\mathcal{R}$ is defined as $\mathcal{R}^{k\oplus} = \{R_1 \oplus \cdots \oplus R_k \colon R_1, \ldots, R_k \in \mathcal{R}\}$, where $R_1 \oplus \cdots \oplus R_k$ is the set of those elements that are contained in an odd number of $R_1, \ldots, R_k$.

One of the fundamental measures of complexity of a set system is its *Vapnik-Chervonenkis dimension*, or in short, *VC-dimension*. Given a set system $(X, \mathcal{R})$, for any set $Y \subseteq X$, define the *projection* of $\mathcal{R}$ onto Y as

$$\mathcal{R}|_Y = \{Y \cap R \colon R \in \mathcal{R}\}.$$

We say that $\mathcal{R}$ *shatters* Y if $|\mathcal{R}|_Y| = 2^{|Y|}$; in other words, any subset of Y can be derived as the intersection of Y with an element of $\mathcal{R}$. The *VC-dimension* of $\mathcal{R}$, denoted by VC-dim$(\mathcal{R})$, is the size of the largest subset of X that can be shattered by $\mathcal{R}$. Originally introduced in statistical learning by

Vapnik and Chervonenkis (1971), it has turned out to be a key parameter in several areas, including learning theory, combinatorics, and computational geometry.

1.1. Learning Theory.

In learning theory, the VC-dimension of a concept class measures the difficulty of learning a concept of the class. The foundational paper of Blumer et al. (1989) states that "the essential condition for distribution-free learnability is finiteness of the Vapnik-Chervonenkis dimension". Among their results, they prove the following theorem.

Theorem A (Blumer et al. 1989) *Let $(X, \mathcal{R})$ be a set system and k be any positive integer. Then*

$$\text{VC-dim}\left(\mathcal{R}^{k\cup}\right) = O\Big(\text{VC-dim}\left(\mathcal{R}\right) \cdot k \log k\Big),$$

$$\text{VC-dim}\left(\mathcal{R}^{k\cap}\right) = O\Big(\text{VC-dim}\left(\mathcal{R}\right) \cdot k \log k\Big).$$

Moreover, there are set systems such that $\text{VC-dim}\left(\mathcal{R}^{k\cup}\right) = \Omega\left(\text{VC-dim}\left(\mathcal{R}\right) \cdot k\right)$ *and* $\text{VC-dim}\left(\mathcal{R}^{k\cap}\right) = \Omega\left(\text{VC-dim}\left(\mathcal{R}\right) \cdot k\right)$*.*

Remark. The upper bound holds in a more general setting: for any fixed set-theoretic expression $F(R_1, R_2, \ldots, R_k)$ (consisting of operations of set union, intersection, and difference) and range set $\mathcal{R}^{k*} := \{F(R_1, \ldots, R_k) : R_1, \ldots, R_k \in \mathcal{R}\}$, we have $\text{VC-dim}(\mathcal{R}^{k*}) = O(\text{VC-dim}\left(\mathcal{R}\right) \cdot k \log k)$ (see eg. Matoušek, 2002, chap. 10). In particular, $\text{VC-dim}\left(\mathcal{R}^{k\oplus}\right) = O\left(\text{VC-dim}\left(\mathcal{R}\right) \cdot k \log k\right)$.

They also considered the question of whether the upper bounds of Theorem A are tight in the most basic geometric case when $X \subseteq \mathbb{R}^d$ is a set of points and $\mathcal{R}$ is the projection of the family of all half-spaces of $\mathbb{R}^d$ onto X. They proved that the VC-dimension of the k-fold union of half-spaces in two dimensions is exactly $2k+1$. For general dimensions $d \geq 3$, they upper-bound the VC-dimension of the k-fold union of half-spaces by $O(d \cdot k \log k)$. This follows from Theorem A together with the fact that the VC-dimension of the set system induced by half-spaces in $\mathbb{R}^d$ is $d+1$. The same upper bound holds for the VC-dimension of the k-fold intersection of half-spaces in $\mathbb{R}^d$. Later Dobkin and Gunopulos (1995) showed that the VC-dimension of the k-fold union of half-spaces in $\mathbb{R}^3$ is upper-bounded by $4k$.

Eisenstat and Angluin (2007) proved, by giving a probabilistic construction of an abstract set system, that the upper bound of Theorem A is asymptotically tight if $\text{VC-dim}\left(\mathcal{R}\right) \geq 5$ and that for $\text{VC-dim}\left(\mathcal{R}\right) = 1$, an upper bound of k holds which is tight. A few years later, Eisenstat (2009) filled the gap by showing that there exists a set system $(X, \mathcal{R})$ of VC-dimension 2 such that $\text{VC-dim}\left(\mathcal{R}^{k\cup}\right) = \Omega\left(\text{VC-dim}\left(\mathcal{R}\right) \cdot k \log k\right)$.

For $d \geq 4$, the current best upper-bound for the k-fold union and the k-fold intersection of half-spaces in $\mathbb{R}^d$ is still the one given by Theorem A almost 30 years ago, while the lower-bound has remained $\Omega\left(\text{VC-dim}(\mathcal{R}) \cdot k\right)$. We refer the reader to the PhD thesis of Johnson (2008) for a summary of the bounds on VC-dimensions of these basic combinatorial and geometric set systems.

1.2. Computational Geometry.

The following set system is fundamental in computational geometry. Given a set $\mathcal{H}$ of hyperplanes in $\mathbb{R}^d$, define

$$\Delta\left(\mathcal{H}\right) = \Big\{\mathcal{H}' \subseteq \mathcal{H} \colon \exists \text{ an open } d\text{-dimensional simplex } S \text{ in } \mathbb{R}^d \text{ such that } H \in \mathcal{H} \text{ intersects } S \text{ if and only if } H \in \mathcal{H}'\Big\}.$$

Its importance derives from the fact that it is the set system underlying the construction of *cuttings* via random sampling (see Chazelle and Friedman 1990 as well as the recent survey of Mustafa and Varadarajan 2017, to appear). Cuttings are the key tool for fast point-location algorithms and were studied in detail recently by Ezra et al. (2017). They provided the best bounds so far for the VC-dimension of $\Delta(\mathcal{H})$.

Lemma B (Ezra et al. 2017) *For $d \geq 9$, we have*

$$d\,(d+1) \leq \text{VC-dim}\,(\Delta\,(\mathcal{H})) \leq 5 \cdot d^2 \log d.$$

2. Our Results

For some time now, it has generally been expected that VC-dim $\left(\mathcal{R}^{k\cup}\right) =$ VC-dim $\left(\mathcal{R}^{k\cap}\right) = O\,(dk)$ for the k-fold unions and intersections of half-spaces. This upper-bound indeed holds for a related notion: the *primal shattering dimension* of the k-fold unions and intersections of half-spaces is $O(dk)$. In fact, as it was pointed out by Bachem (2018), several papers in learning theory assume the same for VC-dimension. Likewise for computational geometry literature: for example, the coreset size bounds in the constructions of Feldman and Langberg (2011), Balcan et al. (2013), and Lucic et al. (2016) would require an additional $\log k$ factor in the coreset size—if the upper-bound of Theorem A was tight for the k-fold intersection of half-spaces.

In this paper, we resolve the question of VC-dimension for the above two set systems. Our proofs are short and we make an effort to keep them self-contained.

We show an optimal lower-bound on the VC-dimension of the k-fold union, the k-fold intersection, and the k-fold symmetric difference of half-spaces in $\mathbb{R}^d$, matching the $O(d \cdot k \log k)$ upper bound of Theorem A, and thus settling affirmatively one of the main open questions studied by Eisenstat and Angluin (2007), Johnson (2008), and Eisenstat (2009).

Theorem 1 *Let k be a given positive integer and $d \geq 4$ be an integer. Then there exists a set $\mathcal{P}$ of points in $\mathbb{R}^d$ such that the set system $\mathcal{R}$ induced on $\mathcal{P}$ by half-spaces satisfies*

$$\text{a)}\quad \text{VC-dim}\left(\mathcal{R}^{k\cup}\right) = \Omega\Big(\text{VC-dim}(\mathcal{R}) \cdot k \log k\Big) = \Omega\Big(d \cdot k \log k\Big),$$

$$\text{b)}\quad \text{VC-dim}\left(\mathcal{R}^{k\cap}\right) = \Omega\Big(\text{VC-dim}(\mathcal{R}) \cdot k \log k\Big) = \Omega\Big(d \cdot k \log k\Big),$$

$$\text{c)}\quad \text{VC-dim}\left(\mathcal{R}^{k\oplus}\right) = \Omega\Big(\text{VC-dim}(\mathcal{R}) \cdot k \log k\Big) = \Omega\Big(d \cdot k \log k\Big).$$

Remark 1. This statement also provides a non-probabilistic proof of the lower-bound of Eisenstat and Angluin (2007).

Remark 2. Note that a set of points is shattered by (the k-fold union of) closed half-spaces if and only if it is shattered by (the k-fold union of) open half-spaces. Thus Theorem 1 also holds for open half-spaces.

Remark 3. Observe that if $\overline{\mathcal{R}} := \{\mathbb{R}^d \setminus R : \ R \in \mathcal{R}\}$, then VC-dim$(\overline{\mathcal{R}}) =$ VC-dim$(\mathcal{R})$ and

$$\text{VC-dim}\left(\mathcal{R}^{k\cap}\right) = \text{VC-dim}\left(\overline{\mathcal{R}^{k\cap}}\right) = \text{VC-dim}\left(\overline{\mathcal{R}}^{k\cup}\right).$$

holds by the De Morgan laws. Since for half-spaces $\overline{\mathcal{R}} = \mathcal{R}$, part a) of Theorem 1 implies part b).

Using Theorem 1, we show an asymptotically optimal bound on the VC-dimension of $\Delta(\mathcal{H})$, improving the bound of Ezra et al. (2017) and resolving a question that was studied in the computational geometry community starting in the 1980s.

Theorem 2 *Let $d \geq 4$ be a given integer. Then there exists a set $\mathcal{H}$ of $\Omega\left(d^2 \log d\right)$ hyperplanes in $\mathbb{R}^d$ such that $\Delta(\mathcal{H}) = 2^{\mathcal{H}}$, that is, we have*

$$\text{VC-dim}\left(\Delta\left(\mathcal{H}\right)\right) = \Omega\left(d^2 \log d\right).$$

Remark. In fact, we prove a more general result bounding the VC-dimension of the set system induced by intersection of hyperplanes with k-dimensional simplices in $\mathbb{R}^d$. See Section 4 for details.

Organization. Section 3 contains the proof of Theorem 1 and Section 4 contains the proof of Theorem 2.

3. Proof of Theorem 1.

We prove the theorem for d even, starting from $d = 4$. The asymptotic lower-bound for odd values of d follows from the one in $\mathbb{R}^{d-1}$. The starting point of the proof is the following lemma.

Lemma 3 (Kupavskii et al., 2016, Lemma 2) *Let $n, d' \geq 2$ be integers. Then there exists a set $\mathcal{B}_{n,d'}$ of axis-parallel boxes in $\mathbb{R}^{d'}$, with $|\mathcal{B}_{n,d'}| = (d'-1)(n+1)2^{n-2}$, such that for any subset $\mathcal{S} \subseteq \mathcal{B}_{n,d'}$, there exists a set $\mathcal{Q}$ of points in $\mathbb{R}^{d'}$ such that $|\mathcal{Q}| = 2^{n-1}$ and*

(i) $|\mathcal{Q} \cap B| = 1$ for any $B \in \mathcal{B}_{n,d'} \setminus \mathcal{S}$, and

(ii) $\mathcal{Q} \cap B = \emptyset$ for any $B \in \mathcal{S}$.

Remark. In Kupavskii et al. (2016) this lemma is stated in a weaker form, however the above stronger statement is implicit in their proof.

Let $d' = \left\lfloor \frac{d}{2} \right\rfloor$ and apply Lemma 3 with $n = \lfloor \log k \rfloor + 1$ in $\mathbb{R}^{d'}$ to get a set $\mathcal{B} = \mathcal{B}_{n,d'}$ of boxes in $\mathbb{R}^{d'}$. By translation, we can assume that all coordinates of points lying in each box in $\mathcal{B}$ are positive. Now we will construct the following mappings:

$$\text{set of boxes } \mathcal{B} \text{ in } \mathbb{R}^{d/2} \xrightarrow{\ \pi\ } \text{points in } \mathbb{R}^d,$$

$$\text{points in } \mathbb{R}^{d/2} \xrightarrow{\ \beta\ } \text{boxes in } \mathbb{R}^d \xrightarrow{\ \gamma\ } \text{half-spaces in } \mathbb{R}^d.$$

We will then prove the key property of these mappings, that for any $q \in \mathbb{R}^{d/2}$ and $B \in \mathcal{B}$

$$q \in B \quad \overset{(1)}{\Longleftrightarrow} \quad \pi(B) \in \beta(q) \quad \overset{(2)}{\Longleftrightarrow} \quad \pi(B) \in \gamma(\beta(q)).$$

We first define mappings π and β, with $\pi : \mathcal{B} \to \mathbb{R}^d$ and with β mapping points in $\mathbb{R}^{d'}$ to axis-parallel boxes in $\mathbb{R}^d$, such that for any $B \in \mathcal{B}$ and $q \in \mathbb{R}^{d'}$, we have

$$q \in B \iff \pi(B) \in \beta(q). \tag{1}$$

- Let $B \in \mathcal{B}$ be defined as the product of d' intervals:

$$B = [x_1, x'_1] \times [x_2, x'_2] \times \cdots \times [x_{d'}, x'_{d'}], \quad \text{with} \quad x_i, x'_i > 0 \text{ for each } i \in [d'].$$

Then π maps B in $\mathbb{R}^{d'}$ to the following point in $\mathbb{R}^d$ (see Pach and Tardos, 2013)

$$\pi(B) = \left(x_1, \frac{1}{x'_1}, x_2, \frac{1}{x'_2}, \ldots, x_{d'}, \frac{1}{x'_{d'}}\right) \in \mathbb{R}^d,$$

and we let $\pi(\mathcal{B}) := \{\pi(B) \colon B \in \mathcal{B}\}$.

- For any point $q = (q_1, q_2, \ldots, q_{d'}) \in \mathbb{R}^{d'}$, define $\beta(q)$ to be the box

$$\beta(q) = [0, q_1] \times \left[0, \frac{1}{q_1}\right] \times \cdots \times [0, q_{d'}] \times \left[0, \frac{1}{q_{d'}}\right] \subset \mathbb{R}^d.$$

Proposition 4 *The mappings π and β satisfy* (1).

Proof Let $q = (q_1, q_2, \ldots, q_{d'})$ be a point in $\mathbb{R}^{d'}$ and $B = [x_1, x'_1] \times [x_2, x'_2] \times \cdots \times [x_{d'}, x'_{d'}]$ be an axis-parallel box in $\mathbb{R}^{d'}$. Then $q \in B$ if and only if $x_i \leq q_i \leq x'_i$ for all $i \in [d']$.

On the other hand, $\pi(B)$ lies in $\beta(q)$ if and only if $0 \leq x_i \leq q_i$ and $0 \leq 1/x'_i \leq 1/q_i$ for each $i \in [d']$—or equivalently, $0 \leq x_i \leq q_i$ and $q_i \leq x'_i$ for each $i \in [d']$.

Note that these two conditions are exactly the same, implying (1). ■

Recall that $\mathcal{B} = \mathcal{B}_{n,d'}$ is the set of boxes provided by Lemma 3 with parameters $d' = \lfloor \frac{d}{2} \rfloor$ and $n = \lfloor \log k \rfloor + 1$. Thus we have

$$|\mathcal{B}| = \left(\left\lfloor \frac{d}{2} \right\rfloor - 1 \right) \lfloor \log k \rfloor 2^{\lfloor \log k \rfloor - 1}.$$

Proposition 5 ("Lifted dual version" of Lemma 3) *$\pi(\mathcal{B})$ is a set of $\left(\lfloor \frac{d}{2} \rfloor - 1\right) \lfloor \log k \rfloor 2^{\lfloor \log k \rfloor - 1}$ points in $\mathbb{R}^d$ such that for any subset $\mathcal{T} \subseteq \pi(\mathcal{B})$, there is a set $\mathcal{Q}^*$ of at most k axis-parallel boxes in $\mathbb{R}^d$ such that*

(i) each point of $\mathcal{T}$ is contained in exactly one box in $\mathcal{Q}^*$, and

(ii) no point of $\pi(\mathcal{B}) \setminus \mathcal{T}$ is contained in any box of $\mathcal{Q}^*$.

In particular, $\pi(\mathcal{B})$ is shattered by the set system induced by the k-fold union of axis-parallel boxes in $\mathbb{R}^d$ and also shattered by the set system induced by the k-fold symmetric difference of axis-parallel boxes in $\mathbb{R}^d$.

Proof Let $\mathcal{S} = \left\{\pi^{-1}(p) \ : \ p \in \pi(\mathcal{B}) \setminus \mathcal{T}\right\}$. By Lemma 3, there is a set $\mathcal{Q}$ of at most k points in $\mathbb{R}^{d'}$ such that (i), (ii) of Lemma 3 hold for $\mathcal{S}$ and $\mathcal{Q}$. Letting $\mathcal{Q}^* = \{\beta(q) : q \in \mathcal{Q}\}$, the claim follows from (1). ■

Next we define the function $\gamma(\cdot)$ mapping boxes in $\mathbb{R}^d$ to half-spaces in $\mathbb{R}^d$ such that for any point $p \in \pi(\mathcal{B})$ and box $B = \beta(q)$, we have

$$p \in B \iff p \in \gamma(B). \tag{2}$$

For every $i \in [d]$, let $0 < \alpha_{i,1} < \alpha_{i,2} < \ldots$ denote the sequence of distinct values of the x_i-coordinates of the elements of $\pi(\mathcal{B})$. Every such sequence has length at most $|\pi(\mathcal{B})|$. By re-scaling the coordinates, we can assume that

$$\text{for each } i \in [d] \text{ and } j \leq |\pi(\mathcal{B})|, \qquad \frac{\alpha_{i,j+1}}{\alpha_{i,j}} > d. \tag{3}$$

Denote the resulting point set by $\mathcal{P}$. Note that scaling along each coordinate does not change incidences with respect to axis-parallel boxes, thus Proposition 5 still holds if we replace $\pi(\mathcal{B})$ by $\mathcal{P}$ and that $|\mathcal{P}| = |\pi(\mathcal{B})| = |\mathcal{B}|$.

We claim that $\mathcal{P}$ is shattered by the set system induced by the k-fold union of half-spaces in $\mathbb{R}^d$ and also shattered by the set system induced by the k-fold symmetric difference of half-spaces in $\mathbb{R}^d$. To see that, let $\mathcal{P}'$ be any subset of $\mathcal{P}$. Let $\mathcal{Q}^*$ be the set of axis-parallel boxes corresponding to $\mathcal{P}'$ provided by Proposition 5.

For each box $B \in \mathcal{Q}^*$, we can re-scale B if necessary, without changing its intersection with $\mathcal{P}$ so that B is of the form

$$B = [0, b_1] \times [0, b_2] \times \cdots \times [0, b_d],$$

where each b_i is equal to α_{i,j_i}, for a suitable j_i. Now for each box $B \in \mathcal{Q}^*$, we define a half-space $\gamma(B)$ as the set of points $(x_1, \ldots, x_d) \in \mathbb{R}^d$ satisfying

$$\frac{x_1}{b_1} + \frac{x_2}{b_2} + \cdots + \frac{x_d}{b_d} \leq d. \tag{4}$$

Clearly for any point $(x_1, \ldots, x_d)$ contained in the box B, we have $x_i \in [0, b_i]$ and thus each term on the left-hand side of the Equation (4) is at most 1. This implies that $B \subset \gamma(B)$ and so any point in the box B lies in the half-space $\gamma(B)$.

On the other hand, for any point $(x_1, \ldots, x_d) \in \mathcal{P} \setminus B$, there exists an index $i \in [d]$ such that $x_i > b_i$. By (3), we have that $x_i/b_i > d$, and thus $(x_1, \ldots, x_d)$ cannot lie in the half-space $\gamma(B)$.

Consider the set of at most k half-spaces defined as $\mathcal{H} = \{\gamma(B) : B \in \mathcal{Q}^*\}$. Now by Proposition 5 and (2), we have

(i) each point of $\mathcal{P}'$ is contained in exactly one half-space in $\mathcal{H}$, and

(ii) no point of $\mathcal{P} \setminus \mathcal{P}'$ is contained in any half-space of $\mathcal{H}$.

In other words, the union as well as the symmetric difference of the half-spaces in $\mathcal{H}$ contains precisely the set $\mathcal{P}'$. As this is true for any $\mathcal{P}' \subseteq \mathcal{P}$, the k-fold union of half-spaces in $\mathbb{R}^d$ shatters $\mathcal{P}$ and the same holds for the k-fold symmetric difference of half-spaces in $\mathbb{R}^d$. Finally, we have

$$|\mathcal{P}| = |\mathcal{B}| = \left(\left\lfloor \frac{d}{2} \right\rfloor - 1 \right) \lfloor \log k \rfloor 2^{\lfloor \log k \rfloor - 1} = \Omega\left(d \cdot k \log k\right),$$

as desired. ∎

4. Proof of Theorem 2.

Given a set $\mathcal{H}$ of hyperplanes in $\mathbb{R}^d$, define the set system

$$\Delta_k\left(\mathcal{H}\right) = \Big\{ \mathcal{H}' \subseteq \mathcal{H} \colon \exists \text{ an open } k\text{-dimensional simplex } S \text{ in } \mathbb{R}^d \text{ such that } H \in \mathcal{H}' \text{ if and only if } H \text{ intersects } S \Big\}.$$

We prove the following more general theorem from which Theorem 2 follows immediately by setting $k = d$.

Theorem 6 *For any integer $d \geq 4$ and $k \leq d$, there exists a set $\mathcal{H}$ of $\Omega\left(dk \log k\right)$ hyperplanes in $\mathbb{R}^d$ such that $|\Delta_k(\mathcal{H})| = 2^{|\mathcal{H}|}$, that is, we have*

$$\text{VC-dim}\left(\Delta_k\left(\mathcal{H}\right)\right) = \Omega\left(d \cdot k \log k\right).$$

Proof Apply Theorem 1 to get a set $\mathcal{P}$ of $\Omega\left(dk \log k\right)$ points in $\mathbb{R}^d$ that is shattered by the k-fold union of open half-spaces. Using point-hyperplane duality (see eg. Matoušek, 2002), map each point $p \in \mathcal{P}$ to a hyperplane $\alpha(p)$ by

$$p = (p_1, \ldots, p_d) \quad \longmapsto \quad \alpha(p) := \big\{ (x_1, \ldots, x_d) : p_1 x_1 + p_2 x_2 + \cdots + p_{d-1} x_{d-1} - x_d = -p_d \big\}.$$

Our desired set $\mathcal{H}$ of hyperplanes will simply be

$$\mathcal{H} = \big\{ \alpha(p) \colon p \in \mathcal{P} \big\}.$$

It is easy to check that the mapping α is injective and thus $|\mathcal{H}| = |\mathcal{P}| = \Omega(dk \log k)$. We claim that $\mathcal{H}$ is shattered by the set system induced by open k-dimensional simplices; in other words, for any $\mathcal{H}' \subseteq \mathcal{H}$, there exists a k-dimensional simplex S such that the interior of S intersects each hyperplane of $\mathcal{H}'$, and no hyperplane of $\mathcal{H} \setminus \mathcal{H}'$.

Fix any $\mathcal{H}' \subseteq \mathcal{H}$ and let $\mathcal{P}' = \alpha^{-1}(\mathcal{H}')$ be the corresponding points of $\mathcal{P}$. Then there exists a set $\mathfrak{H}(\mathcal{P}')$ of k open half-spaces whose union contains all points in $\mathcal{P}'$ and no point in $\mathcal{P} \setminus \mathcal{P}'$. From Equation (4), it follows that each half-space in $\mathfrak{H}(P')$ is of the form

$$\frac{x_1}{b_1} + \frac{x_2}{b_2} + \cdots + \frac{x_d}{b_d} < d,$$

where $b_1, \ldots, b_d$ are positive reals. Map each half-space $H \in \mathfrak{H}(\mathcal{P}')$ to the point $\delta(H)$, given by

$$H = \left\{ (x_1, \ldots, x_d) : \frac{x_1}{b_1} + \frac{x_2}{b_2} + \cdots + \frac{x_d}{b_d} < d \right\} \quad \longmapsto \quad \delta(H) := \left(\frac{b_d}{b_1}, \ldots, \frac{b_d}{b_{d-1}}, d \cdot b_d \right).$$

It is easy to verify that for a point $p \in \mathbb{R}^d$ and the half-space H, we have

$$\begin{aligned}
p \in H \iff & \frac{p_1}{b_1} + \frac{p_2}{b_2} + \cdots + \frac{p_d}{b_d} < d \\
\iff & p_1 \cdot \frac{b_d}{b_1} + p_2 \cdot \frac{b_d}{b_2} + \cdots + p_{d-1} \cdot \frac{b_d}{b_{d-1}} + p_d < d \cdot b_d \\
\iff & \text{point } \left(\frac{b_d}{b_1}, \frac{b_d}{b_2}, \ldots, \frac{b_d}{b_{d-1}}, d \cdot b_d \right) \text{ lies strictly above the hyperplane} \\
& \qquad p_1 x_1 + p_2 x_2 + \cdots + p_{d-1} x_{d-1} - x_d = -p_d \\
\iff & \text{ the point } \delta(H) \text{ lies strictly above the hyperplane } \alpha(p).
\end{aligned} \tag{5}$$

Here we crucially needed the fact that all half-spaces in $\mathfrak{H}(\mathcal{P}')$ are downward facing, that is, each half-space in $\mathfrak{H}(\mathcal{P}')$ contains the origin, which lies below (with respect to the x_d-coordinate) its bounding hyperplane.

Now consider the k open half-spaces in $\mathfrak{H}(\mathcal{P}')$ and let

$$\Delta' = \{\delta(H) \colon H \in \mathfrak{H}(\mathcal{P}')\}$$

be k points in $\mathbb{R}^d$. From the relation (5), it follows that

- As each point $p \in \mathcal{P}'$ lies in *some* half-space $H \in \mathfrak{H}(\mathcal{P}')$, the point $\delta(H)$ lies strictly above the hyperplane $\alpha(p)$ in $\mathcal{H}$—or equivalently, the hyperplane $\alpha(p)$ has at least one of the k points in the set Δ' lying strictly above it.
- For each point $p \in \mathcal{P} \setminus \mathcal{P}'$, all the k points in Δ' lie on or below the hyperplane $\alpha(p) \in \mathcal{H}$.

Then, by the above discussion, $H \in \mathcal{H}'$ if and only if one of these is true:

1. H intersects the interior of $\texttt{conv}\ (\Delta')$ and so at least one vertex of Δ' lies strictly above H, or
2. H does not intersect $\texttt{conv}\ (\Delta')$, but then *all* vertices of Δ' lie strictly above H.

Finally consider the k-dimensional simplex

$$S = \texttt{conv}\ \left(\Delta' \bigcup (0, \ldots, 0, -\infty) \right).$$

Clearly, a hyperplane $H \in \mathcal{H}$ intersects the interior of S if and only if $H \in \mathcal{H}'$. Note that the point $(0, \ldots, 0, -\infty)$ can be any point $(0, \ldots, 0, t)$ for a small-enough value of $t \in \mathbb{R}$. This concludes the proof. ∎

Acknowledgements. The research of Mónika Csikós and Nabil H. Mustafa was supported by the grant ANR SAGA (JCJC-14-CE25-0016-01). The research of Andrey Kupavskii was partially supported by the EPSRC grant no. EP/N019504/1 and by the Council for the Support of Leading Scientific Schools of the President of the Russian Federation (grant no. N.Sh.-6760.2018.1). We thank Yaqiao Li for his valuable comments, which improved the statement of Theorem 1.

References

O. Bachem. *Sampling for Large-Scale Clustering.* PhD thesis, ETH Zürich, 2018.

M.-F. Balcan, S. Ehrlich, and Y. Liang. Distributed k-means and k-median clustering on general topologies. In *Proceedings of Neural Information Processing Systems (NIPS)*, 2013.

A. Blumer, A. Ehrenfeucht, D. Haussler, and M. K. Warmuth. Learnability and the Vapnik-Chervonenkis dimension. *J. ACM*, 36(4):929–965, October 1989.

B. Chazelle and J. Friedman. A deterministic view of random sampling and its use in geometry. *Combinatorica*, 10(3):229–249, 1990.

D. P. Dobkin and D. Gunopulos. Concept learning with geometric hypotheses. In *Proceedings of Computational Learning Theory (COLT)*, pages 329–336, 1995.

D. Eisenstat. k-fold unions of low-dimensional concept classes. *Information Processing Letters*, 109 (23-24):1232–1234, 2009.

D. Eisenstat and D. Angluin. The VC dimension of k-fold union. *Information Processing Letters*, 101(5):181 – 184, 2007.

E. Ezra, S. Har-Peled, H. Kaplan, and M. Sharir. Decomposing arrangements of hyperplanes: VC-dimension, combinatorial dimension, and point location. *ArXiv e-prints*, December 2017.

D. Feldman and M. Langberg. A unified framework for approximating and clustering data. In *Proceedings of Symposium on the Theory of Computing (STOC)*, pages 569–578, 2011.

H. Johnson. *Definable Families of Finite Vapnik Chernonenkis Dimension.* PhD thesis, University of Maryland, 2008.

A. Kupavskii, N. H. Mustafa, and J. Pach. New lower bounds for epsilon-nets. In *Proceedings of Symposium on Computational Geometry (SoCG)*, pages 54:1–54:16, 2016.

M. Lucic, O. Bachem, and A. Krause. Strong Coresets for Hard and Soft Bregman Clustering with Applications to Exponential Family Mixtures. In *Proceedings of Artificial Intelligence and Statistics (AISTATS)*, 2016.

J. Matoušek. *Lectures in Discrete Geometry.* Springer-Verlag, New York, NY, 2002.

N. H. Mustafa and K. Varadarajan. Epsilon-approximations and epsilon-nets. In J. E. Goodman, J. O'Rourke, and C. D. Tóth, editors, *Handbook of Discrete and Computational Geometry.* CRC Press LLC, 2017, to appear.

J. Pach and G. Tardos. Tight lower bounds for the size of epsilon-nets. *J. Amer. Math. Soc.*, 26(3): 645–658, 2013.

V. N. Vapnik and A. Ya. Chervonenkis. On the uniform convergence of relative frequencies of events to their probabilities. *Theory of Probability and its Applications*, 16(2):264–280, 1971.