

LITTLESTONE AND VC-DIMENSION OF FAMILIES OF ZERO SETS

VINCENT GUINGONA, ALEXEI KOLESNIKOV, JULIE NIERWINSKI,
AND RICHARD SOUCY

ABSTRACT. We prove that, for any d linearly independent functions from some set into a d -dimensional vector space over any field, the family of zero sets of all non-trivial linear combination of these functions has VC-dimension and Littlestone dimension $d - 1$. Additionally, we characterize when such families are maximal of VC-dimension $d - 1$ and give a sufficient condition for when they are maximal of Littlestone dimension $d - 1$.

1. INTRODUCTION

Complexity, in the sense of machine learning theory, of the sets of positivity of linear combinations of real-valued functions is fairly well understood. If $f : \mathbb{R}^k \rightarrow \mathbb{R}$ is a function, the *set of positivity of f* is the set $\{\bar{x} \in \mathbb{R}^k : f(\bar{x}) > 0\}$. If $\{f_1, \dots, f_d\}$ is a set of linearly independent functions from \mathbb{R}^k to \mathbb{R} , then the set system of sets of positivity of all the linear combinations of $\{f_1, \dots, f_d\}$ has the VC dimension d (see [4] or [5, Theorem A]). If X is a finite subset of the domain of the functions, by Sauer–Shelah lemma, the size of the induced set system on X is bounded above by $\sum_{i=0}^d \binom{|X|}{i}$ (we abbreviate this sum as $\binom{|X|}{\leq d}$).

There are known sufficient conditions on functions guaranteeing that for suitable finite sets X , the induced set systems have maximal size $\binom{|X|}{\leq d}$. Floyd shows in [6] that a condition on linear dimension of the family of functions relative to the set X and a condition on the number of zeros are sufficient for maximality. Johnson provides analytic sufficient conditions in [7] for maximality of the set system. It is worth pointing out that set systems given by the sets of positivity generally have infinite Littlestone dimension.

Date: September 8, 2021.

This work was partially supported by National Security Agency (NSA) grant H98230-20-1-0005, which funds a Research Experiences for Undergraduates program at Towson University. This work also received financial support from the Towson University Jess and Mildred Fisher College of Science and Mathematics.

We study set systems of zero-sets of non-trivial linear combinations of functions from a set to a field. One expects to see a lower combinatorial complexity of this set system, and indeed we show that both Littlestone dimension and VC-dimension of zero-sets of linear combinations of d linearly independent functions is $d - 1$. Somewhat surprisingly, the system of zero-sets can be maximal (we formally define the notion in Definition 2.8). We obtain a characterization of maximality of the set system.

Bhaskar has studied a shatter function related to Littlestone dimension (called “thicket dimension” in [1]). Bhaskar showed that the size of the set system that well-labels the leaves of the tree of depth n (this is a natural notion for the shatter function for Littlestone dimension) is bounded above by $\binom{n}{\leq d}$, where d is the Littlestone dimension. It is natural to ask for examples of maximal families for this shatter function.

In this paper, we establish the following results: Let $\{f_1, \dots, f_d\}$ be a set of functions from a set X to a vector space F^d , where F is a field and d is a positive integer, and let \mathcal{C} be the set of all zero-sets of non-trivial linear combinations of f_1, \dots, f_d .

Theorem 3.10: If $\{f_1, \dots, f_d\}$ is linearly independent (in the vector space of all functions from X to F), then \mathcal{C} has VC-dimension and Littlestone dimension $d - 1$.

Theorem 5.6: \mathcal{C} is maximum of VC-dimension $d - 1$ if and only if the image of (f_1, \dots, f_d) is not contained in the union of finitely many proper subspaces of F^d .

Corollary 4.5: If the image of (f_1, \dots, f_d) is not contained in the union of finitely many proper subspaces of F^d , then \mathcal{C} is maximum of Littlestone dimension $d - 1$.

2. PRELIMINARIES

For this paper, let \mathbb{N} denote the set of natural numbers, which for us is the set of all non-negative integers (including 0). Let \mathbb{Z}^+ denote the set of positive integers (i.e., $\mathbb{Z}^+ = \mathbb{N} \setminus \{0\}$).

For $n, k \in \mathbb{N}$ with $k \leq n$, let

$$\binom{n}{k} = \frac{n!}{k!(n-k)!} \text{ and } \binom{n}{\leq k} = \sum_{i=0}^k \binom{n}{i}.$$

For all sets X and Y , let ${}^X Y$ denote the set of all functions from X to Y . For any set X and $f : X \rightarrow \mathbb{R}$, let $\text{supp}(f) = \{a \in X : f(a) \neq 0\}$. For all sets X and all $n \in \mathbb{N}$, let $\binom{X}{n}$ denote the set of all subsets of X

of size n . Similarly, let

$$\binom{X}{\leq n} = \bigcup_{i \leq n} \binom{X}{i} \text{ and } \binom{X}{< n} = \bigcup_{i < n} \binom{X}{i}.$$

For all $n \in \mathbb{N}$, let

$$[n] = \{k \in \mathbb{N} : 0 \leq k < n\}.$$

(Note that $[0] = \emptyset$.) Then, we have that, for all $0 \leq k \leq n$,

$$\left| \binom{[n]}{k} \right| = \binom{n}{k} \text{ and } \left| \binom{[n]}{\leq k} \right| = \binom{n}{\leq k}.$$

Definition 2.1. Let X be a set, $\mathcal{C} \subseteq \mathcal{P}(X)$, and $d \in \mathbb{N}$. A (X, \mathcal{C}) -labeled tree of depth d is a function T from $\bigcup_{k \leq d} [k][2]$ to $X \cup \mathcal{C}$ such that

- (1) for all $0 \leq k < d$ and $\sigma \in [k][2]$, $T(\sigma) \in X$; and
- (2) for all $\tau \in [d][2]$, $T(\tau) \in \mathcal{C}$

For $0 \leq k \leq d$ and $\sigma \in [k][2]$, we call $(\sigma, T(\sigma))$ a *node* if $k < d$ and a *leaf* if $k = d$. A leaf $(\tau, T(\tau))$ is *well-labeled* if, for all $0 \leq k < d$,

$$T(\tau|_{[k]}) \in T(\tau) \iff \tau(k) = 1.$$

We say that T is *well-labeled* if all leaves of T are well-labeled. We say that T of depth d is *level-balanced* if, for all $k < d$ and for all $\sigma_0, \sigma_1 \in [k][2]$, $T(\sigma_0) = T(\sigma_1)$.

Definition 2.2. Let X be a set and $\mathcal{C} \subseteq \mathcal{P}(X)$.

- (1) The *Littlestone dimension* of \mathcal{C} , denoted $\text{Ldim}(\mathcal{C})$, is the largest natural number d such that there exists a well-labeled (X, \mathcal{C}) -labeled tree of depth d . If there exists a well-labeled (X, \mathcal{C}) -labeled tree of depth d for all natural numbers d , we say \mathcal{C} has *infinite Littlestone dimension*, denoted $\text{Ldim}(\mathcal{C}) = \infty$. If there exists no well-labeled (X, \mathcal{C}) -labeled tree of depth 0, we say $\text{Ldim}(\mathcal{C}) = -\infty$.
- (2) The *Littlestone shatter function* of \mathcal{C} , denoted $\rho_{\mathcal{C}}$, is the function from \mathbb{N} to \mathbb{N} given by, for all $n \in \mathbb{N}$, the maximum number of well-labeled leaves of a (X, \mathcal{C}) -labeled tree T of depth n .
- (3) The *Littlestone density* of \mathcal{C} , denoted $\text{Lden}(\mathcal{C})$, is the infimum over all positive $\ell \in \mathbb{R}$ such that there exists $K \in \mathbb{R}$ such that, for all $n \geq 1$, $\rho_{\mathcal{C}}(n) \leq Kn^{\ell}$.
- (4) If we restrict our attention to only level-balanced trees T in the above definitions, we define instead the *VC-dimension* of \mathcal{C} , denoted $\text{VCdim}(\mathcal{C})$, the *VC-shatter function* of \mathcal{C} , denoted $\pi_{\mathcal{C}}$, and the *VC-density* of \mathcal{C} , denoted $\text{VCden}(\mathcal{C})$, respectively.

Remark 2.3. For any set X and $\mathcal{C} \subseteq \mathcal{P}(X)$, we clearly have

- (1) $\text{VCdim}(\mathcal{C}) \leq \text{Ldim}(\mathcal{C})$,
- (2) $\pi_{\mathcal{C}}(n) \leq \rho_{\mathcal{C}}(n)$ for all $n \in \mathbb{N}$, and
- (3) $\text{VCden}(\mathcal{C}) \leq \text{Lden}(\mathcal{C})$.

The above definitions for the VC-dimension and the VC-shatter function are equivalent to the usual definitions given in terms of shattering. In particular, Lemma 2.5 below can be found, for example, in [3].

Definition 2.4. Let X be a set, $Y \subseteq X$, and $\mathcal{C} \subseteq \mathcal{P}(X)$. Let

$$\mathcal{C}|_Y = \{A \cap Y : A \in \mathcal{C}\},$$

we call $\mathcal{C}|_Y$ an *induced set system on Y* . We say that \mathcal{C} *shatters* Y if $\mathcal{C}|_Y = \mathcal{P}(Y)$.

Lemma 2.5. *Let X be a set and $\mathcal{C} \subseteq \mathcal{P}(X)$.*

- (1) *the VC-dimension of \mathcal{C} is the largest $d \in \mathbb{N}$ such that there exists $Y \in \binom{X}{d}$ with $|\mathcal{C}|_Y| = 2^d$.*
- (2) *the VC-shatter function of \mathcal{C} is given by, for all $n \in \mathbb{N}$,*

$$\pi_{\mathcal{C}}(n) = \max \left\{ |\mathcal{C}|_Y| : Y \in \binom{X}{n} \right\}.$$

Lemma 2.6 (Sauer–Shelah Lemma). *For any set X and $\mathcal{C} \subseteq \mathcal{P}(X)$, if \mathcal{C} has VC-dimension d , then, for all $n \geq d$,*

$$\pi_{\mathcal{C}}(n) \leq \binom{n}{\leq d}.$$

An analogous lemma holds in the Littlestone case.

Lemma 2.7 ([1]). *For any set X and $\mathcal{C} \subseteq \mathcal{P}(X)$, if \mathcal{C} has Littlestone dimension d , then, for all $n \geq d$,*

$$\rho_{\mathcal{C}}(n) \leq \binom{n}{\leq d}.$$

This motivates the following two definitions:

Definition 2.8. Let X be a set and $\mathcal{C} \subseteq \mathcal{P}(X)$.

- (1) For $d \in \mathbb{N}$, we say \mathcal{C} is *maximal of VC-dimension d* if, for all $n \geq d$,

$$\pi_{\mathcal{C}}(n) = \binom{n}{\leq d}.$$

- (2) For $d \in \mathbb{N}$, we say \mathcal{C} is *maximal of Littlestone dimension d* if, for all $n \geq d$,

$$\rho_{\mathcal{C}}(n) = \binom{n}{\leq d}.$$

Remark 2.9. For any set X and $\mathcal{C} \subseteq \mathcal{P}(X)$, by the Sauer-Shelah Lemma,

$$\text{VCden}(\mathcal{C}) \leq \text{VCdim}(\mathcal{C}).$$

Similarly, by Lemma 2.7,

$$\text{Lden}(\mathcal{C}) \leq \text{Ldim}(\mathcal{C})$$

Moreover, if \mathcal{C} is maximal of VC-dimension d , then

$$\text{VCden}(\mathcal{C}) = \text{VCdim}(\mathcal{C}) = d$$

and, if \mathcal{C} is maximal of Littlestone dimension d , then

$$\text{Lden}(\mathcal{C}) = \text{Ldim}(\mathcal{C}) = d.$$

If \mathcal{C} is maximal of VC-dimension d and $\text{Ldim}(\mathcal{C}) = d$, then \mathcal{C} is maximal of Littlestone dimension d as well. This is because, in this case, for all $n \geq d$,

$$\binom{n}{\leq d} = \pi_{\mathcal{C}}(n) \leq \rho_{\mathcal{C}}(n) \leq \binom{n}{\leq d}.$$

3. VC-DIMENSION AND LITTLESTONE DIMENSION OF ZERO SETS.

Let X be a set and F a field. The family of functions from X to F has a natural structure of a vector space over F with the pointwise addition operation. The notion of linear independence below is taken in the sense of that vector space. In particular, if $X = F = \mathbb{F}_3$, then the set of functions $\{x, x^3\}$ is linearly dependent, because x and x^3 define the same function from \mathbb{F}_3 to \mathbb{F}_3 .

Definition 3.1. Let F be a field, $d \in \mathbb{Z}^+$, and $\bar{a}, \bar{b} \in F^d$. Let $\bar{a} \cdot \bar{b}$ denote the usual dot product of a and b . That is,

$$(\bar{a}_0, \dots, \bar{a}_{d-1}) \cdot (\bar{b}_0, \dots, \bar{b}_{d-1}) = \sum_{i=0}^{d-1} \bar{a}_i \bar{b}_i$$

for all $\bar{a}_0, \dots, \bar{a}_{d-1}, \bar{b}_0, \dots, \bar{b}_{d-1} \in F$. We say that $\bar{a}, \bar{b} \in F^d$ are *orthogonal* if $\bar{a} \cdot \bar{b} = 0$.

We note that the term *orthogonal* is used as a short-hand; the vector space F^d with the dot product may not be an inner product space.

If $\bar{f} = (f_0, \dots, f_{d-1})$ is a tuple of functions from X to F and $\bar{a} = (a_0, \dots, a_{d-1})$ is a tuple of scalars from F , it is convenient to write the linear combination $a_0 f_0 + \dots + a_{d-1} f_{d-1}$ as the dot product $\bar{a} \cdot \bar{f}$.

Definition 3.2. Let X be a set, F a field, $d \in \mathbb{Z}^+$, $\bar{a} \in F^d$, and $\bar{f} : X \rightarrow F^d$. Define the *zero set* of \bar{f} and \bar{a} by

$$Z_{\bar{f}, \bar{a}} = \{c \in X : \bar{a} \cdot \bar{f}(c) = 0\}.$$

Define

$$\mathcal{C}_{\bar{f}} = \{Z_{\bar{f}, \bar{a}} : \bar{a} \in F^d \setminus \{\bar{0}\}\}.$$

Example 3.3. If $X = \mathbb{R}^2$, $F = \mathbb{R}$, $d = 6$, and

$$\bar{f}(x, y) = (x^2, xy, y^2, x, y, 1),$$

then $\mathcal{C}_{\bar{f}}$ is the set of all conic sections in \mathbb{R}^2 .

Lemma 3.4. Let X be a set, F a field, $d \in \mathbb{Z}^+$, and $\bar{f} : X \rightarrow F^d$. The following are equivalent:

- (1) $\{f_0, f_1, \dots, f_{d-1}\}$ is a linearly independent subset of the F -vector space ${}^X F$;
- (2) for all $\bar{a} \in F^d \setminus \{\bar{0}\}$, there exists $c \in X$ such that $\bar{a} \cdot \bar{f}(c) \neq 0$; and
- (3) $\bar{f}(X)$ is not contained in a proper subspace of F^d .

Proof. (1) \Rightarrow (2): If $\{f_0, f_1, \dots, f_{d-1}\}$ is a linearly independent set, then, for all $\bar{a} \in F^d$, $\sum_{k < d} a_k f_k$ is identically zero if and only if $\bar{a} = \bar{0}$. Therefore, for any non-zero $\bar{a} \in F^d$, $\sum_{k < d} a_k f_k$ is not identically zero, so there exists $c \in X$ such that $\bar{a} \cdot \bar{f}(c) \neq 0$.

(2) \Rightarrow (3): Suppose that (3) fails. So $\bar{f}(X)$ is contained in a proper subspace V of F^d . Take $\bar{a} \in F^d$ non-zero and orthogonal to V . Then, for all $c \in X$, $\bar{a} \cdot \bar{f}(c) = 0$. Hence, (2) fails.

(3) \Rightarrow (1): Suppose that (1) fails. So, there exists $\bar{a} \in F^d \setminus \{\bar{0}\}$ such that $\sum_{k < d} a_k f_k$ is identically zero. Therefore, $\bar{f}(X)$ is contained in the subspace of F^d orthogonal to \bar{a} . \square

Definition 3.5. Let X be a set, F a field, $d \in \mathbb{Z}^+$, and $\bar{f} : X \rightarrow F^d$. If \bar{f} satisfies any of the conditions of Lemma 3.4, we say that \bar{f} is *linearly independent*.

In Proposition 3.6 below, we show that the Littlestone dimension of $\mathcal{C}_{\bar{f}}$ is bounded above by $d - 1$. This is proven by contradiction: from a well-labeled $(X, \mathcal{C}_{\bar{f}})$ -labeled tree of depth d , we create a $d \times d$ matrix that has both full rank and less than full rank.

Proposition 3.6. Let X be a set, F a field, $d \in \mathbb{Z}^+$, and $\bar{f} : X \rightarrow F^d$. Then, $\text{Ldim}(\mathcal{C}_{\bar{f}}) < d$.

Proof. For each natural number k , let 1_k denote the function from $[k]$ to \mathbb{N} that is constantly 1. Assume that T is a well-labeled $(X, \mathcal{C}_{\bar{f}})$ -labeled tree of depth d . First, we will show that the set

$$V = \{\bar{f}(T(1_k)) : k \in [d]\} \subseteq F^d$$

is linearly dependent.

Since $T(1_d) \in \mathcal{C}_{\bar{f}}$, there exists $\bar{a} \in F^d \setminus \{\bar{0}\}$ such that $T(1_d) = Z_{\bar{f}, \bar{a}}$. On the other hand, since T is well-labeled, $T(1_k) \in T(1_d)$ for all $k < d$. Therefore, for all $k < d$,

$$\bar{a} \cdot \bar{f}(T(1_k)) = 0.$$

Therefore, we have

$$\begin{bmatrix} f_0(T(1_0)) & \cdots & f_{d-1}(T(1_0)) \\ \vdots & \ddots & \vdots \\ f_0(T(1_{d-1})) & \cdots & f_{d-1}(T(1_{d-1})) \end{bmatrix} \begin{bmatrix} a_0 \\ \vdots \\ a_{d-1} \end{bmatrix} = \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix}.$$

Since $\bar{a} \neq \bar{0}$, we conclude that V is linearly dependent.

Since V is linearly dependent, there exists $0 \leq k < d$ and $\bar{b} \in F^k$ such that

$$\bar{f}(T(1_k)) = \sum_{j=0}^{k-1} b_j \bar{f}(T(1_j)).$$

Consider $\tau : [d] \rightarrow [2]$ given by

$$\tau(j) = \begin{cases} 1 & \text{if } j < k \\ 0 & \text{if } k \leq j < d \end{cases}.$$

Then, for all $j < k$, $T(1_j) \in T(\tau)$ and $T(1_k) \notin T(\tau)$. Choose $\bar{a} \in F^d \setminus \{\bar{0}\}$ such that $T(\tau) = Z_{\bar{f}, \bar{a}}$. Then, for all $j < k$,

$$\bar{a} \cdot \bar{f}(T(1_j)) = 0 \text{ and } \bar{a} \cdot \bar{f}(T(1_k)) \neq 0.$$

However,

$$\begin{aligned} \bar{a} \cdot \bar{f}(T(1_k)) &= \bar{a} \cdot \left(\sum_{j=0}^{k-1} b_j \bar{f}(T(1_j)) \right) = \\ &= \sum_{j=0}^{k-1} b_j (\bar{a} \cdot \bar{f}(T(1_j))) = 0. \end{aligned}$$

This is a contradiction.

Therefore, no such well-labeled tree T exists. That is, $\text{Ldim}(\mathcal{C}_{\bar{f}}) < d$. \square

Next, in Theorem 3.10 below we establish that, if \bar{f} is linearly independent, then the VC-dimension and Littlestone dimension of $\mathcal{C}_{\bar{f}}$ are both exactly $d - 1$. This is done through Lemma 3.7, where we create a new tuple of functions in the span of \bar{f} that acts as an indicator function on a subset of X of size d . Then, in Proposition 3.8, we use this auxiliary tuple of functions to produce a large subset of X that is shattered by $\mathcal{C}_{\bar{f}}$.

Lemma 3.7. *Let X be a set, F a field, $d \in \mathbb{Z}^+$, and $\bar{f} : X \rightarrow F^d$ linearly independent. Then, there exists $\bar{c} \in X^d$ and $\bar{g} : X \rightarrow F^d$ such that $g_k \in \text{span}\{f_0, \dots, f_{d-1}\}$ for all $k < d$ and, for all $i, j < d$,*

$$g_j(c_i) = \begin{cases} 1 & \text{if } j = i \\ 0 & \text{otherwise} \end{cases}.$$

Proof. By induction on d . If $d = 1$, choose any $c_0 \in X$ where $f_0(c_0) \neq 0$ and set $g_0(x) = (f_0(c_0))^{-1} f_0(x)$.

Fix $d > 1$ and suppose that $\bar{f} : X \rightarrow F^d$ is linearly independent. By the induction hypothesis, there exists $\bar{c} \in X^{d-1}$ and $\bar{g}' : X \rightarrow F^{d-1}$ such that $g'_k \in \text{span}\{f_0, \dots, f_{d-2}\}$ for all $k < d - 1$ and, for all $i, j < d - 1$,

$$g'_j(c_i) = \begin{cases} 1 & \text{if } j = i \\ 0 & \text{otherwise} \end{cases}.$$

Define

$$g'_{d-1}(x) = f_{d-1}(x) - \sum_{i=0}^{d-2} f_{d-1}(c_i) g'_i(x).$$

Then, $g'_{d-1}(c_i) = 0$ for all $i < d - 1$. On the other hand, since g'_{d-1} is a non-trivial linear combination of $\{f_0, \dots, f_{d-1}\}$, which is linearly independent, g'_{d-1} is non-zero. Thus, there exists $c_{d-1} \in X$ such that $g'_{d-1}(c_{d-1}) \neq 0$. Finally, set

$$g_{d-1}(x) = (g'_{d-1}(c_{d-1}))^{-1} g'_{d-1}(x).$$

and set

$$g_i(x) = g'_i(x) - g'_i(c_{d-1}) g_{d-1}(x)$$

for $i < d - 1$. Check that this gives the desired conclusion. \square

Proposition 3.8. *Let X be a set, F a field, $d \in \mathbb{Z}^+$, and $\bar{f} : X \rightarrow F^d$ linearly independent. Then, $\text{VCdim}(\mathcal{C}_{\bar{f}}) \geq d - 1$.*

Proof. By Lemma 3.7, there exists $\bar{c} \in X^d$ and $\bar{g} : X \rightarrow F^d$ such that $g_k \in \text{span}\{f_0, \dots, f_{d-1}\}$ for all $k < d$ and, for all $i, j < d$,

$$g_j(c_i) = \begin{cases} 1 & \text{if } j = i \\ 0 & \text{otherwise} \end{cases}.$$

For any non-empty $S \subseteq [d]$, let

$$g_S(x) = \sum_{i \in S} g_i(x).$$

Then, for any non-empty $S \subseteq [d]$ and any $i < d$, note that

$$g_S(c_i) = 0 \iff i \notin S.$$

On the other hand, since $g_k \in \text{span}\{f_0, \dots, f_{d-1}\}$ and g_k is not identically zero, for each non-empty $S \subseteq [d]$, there exists $\bar{a}_S \in F^d \setminus \{0\}$ such that $g_S = \bar{a}_S \cdot \bar{f}$. Therefore, for any non-empty $S \subseteq [d]$ and any $i < d$,

$$c_i \in Z_{f, a_s} \iff i \notin S.$$

Thus, $\mathcal{C}_{\bar{f}}$ shatters $\{c_0, \dots, c_{d-2}\}$. So the VC-dimension of $\mathcal{C}_{\bar{f}}$ is at least $d - 1$. \square

Remark 3.9. We see from the above proof that, if $\bar{f} : X \rightarrow F^d$ is linearly independent, then $\pi_{\mathcal{C}_{\bar{f}}}(d) \geq 2^d - 1$. Using the next theorem together with the Sauer–Shelah Lemma, we conclude that $\pi_{\mathcal{C}_{\bar{f}}}(d) = 2^d - 1$.

Theorem 3.10. *Let X be a set, F a field, $d \in \mathbb{Z}^+$, and $f : X \rightarrow F^d$ linearly independent. Then,*

$$\text{VCdim}(\mathcal{C}_{\bar{f}}) = \text{Ldim}(\mathcal{C}_{\bar{f}}) = d - 1.$$

Proof. By Proposition 3.6, Remark 2.3, and Proposition 3.8, we obtain

$$d - 1 \leq \text{VCdim}(\mathcal{C}_{\bar{f}}) \leq \text{Ldim}(\mathcal{C}_{\bar{f}}) < d.$$

The conclusion follows. \square

4. VC-DENSITY AND LITTLESTONE DENSITY OF THE SYSTEM OF ZERO SETS.

Now that we have established the VC-dimension and Littlestone dimension of $\mathcal{C}_{\bar{f}}$, we turn our attention to analyzing the VC-density and Littlestone density of $\mathcal{C}_{\bar{f}}$.

Proposition 4.1. *Let X be a set, F a field, $d \in \mathbb{Z}^+$, and $\bar{f} : X \rightarrow F^d$. Suppose that there exist 1-dimensional subspaces $V_i \subseteq F^d$ for $i < k$ and suppose that $\bar{f}(X) \subseteq \bigcup_{i < k} V_i$. Then, $|\mathcal{C}_{\bar{f}}| \leq 2^k$. In particular, $\mathcal{C}_{\bar{f}}$ has VC-density 0 and Littlestone density 0.*

Proof. Without loss of generality, we may assume that $V_i \neq V_j$ for all $i \neq j$. Therefore, since V_i and V_j are 1-dimensional subspaces, $V_i \cap V_j = \{\bar{0}\}$. For each $i < k$, let

$$S_i = \bar{f}^{-1}(V_i \setminus \{\bar{0}\})$$

and let $S_k = \bar{f}^{-1}(\{\bar{0}\})$. For each $i < k$, consider $\mathcal{C}_{\bar{f}}|_{S_i}$. By assumption, for all $c \in S_i$, $\bar{f}(c) \in V_i \setminus \{\bar{0}\}$. Thus, for all $\bar{a} \in F^d$ and $c \in S_i$, $\bar{a} \cdot \bar{f}(c) = 0$ if and only if \bar{a} is orthogonal to V_i . Since this is independent of the choice of $c \in S_i$, we see that $Z_{\bar{f}, \bar{a}} \cap S_i = \emptyset$ or $Z_{\bar{f}, \bar{a}} \cap S_i = S_i$. Therefore,

$$\mathcal{C}_{\bar{f}}|_{S_i} = \{\emptyset, S_i\}.$$

Similarly, one can check that $\mathcal{C}_{\bar{f}}|_{S_k} = \{S_k\}$. Therefore,

$$\mathcal{C}_{\bar{f}} \subseteq \left\{ \bigcup_{i \in I} S_i \cup S_k : I \subseteq [k] \right\}.$$

Thus, $|\mathcal{C}_{\bar{f}}| \leq 2^k$. □

Remark 4.2. Let X be a set, F a field, $d \in \mathbb{Z}^+$, and $\bar{f} : X \rightarrow F^d$. If F is a finite field, then $\mathcal{C}_{\bar{f}}$ is finite (since $F^d \setminus \{\bar{0}\}$ is finite). Therefore, $\mathcal{C}_{\bar{f}}$ has VC-density 0 and Littlestone density 0.

In Proposition 4.3 below, we establish that, if $\bar{f}(X)$ is not contained in a finite union of proper subspaces of F^d , then $\mathcal{C}_{\bar{f}}$ is maximal of VC-dimension $d - 1$. This is done by establishing an infinite sequence of elements from X that have d -wise linearly independent images under \bar{f} . Then, we use linear independence to select every subset of this sequence of size at most $d - 1$ with some $Z_{\bar{f}, \bar{b}}$, establishing the maximality of the VC-dimension.

Proposition 4.3. *Let X be a set, F a field, $d \in \mathbb{N}$ with $d \geq 2$, and $\bar{f} : X \rightarrow F^d$. If $\bar{f}(X)$ is not contained in a finite union of proper subspaces of F^d , then $\mathcal{C}_{\bar{f}}$ is maximal of VC-dimension $d - 1$. In particular, $\text{VCden}(\mathcal{C}_{\bar{f}}) = d - 1$.*

Proof. First, create a sequence $(c_i)_{i \in \mathbb{N}}$ with $c_i \in X$ for all $i \in \mathbb{N}$ such that, for all $I \in \binom{\mathbb{N}}{\leq d}$,

$$(1) \quad \{\bar{f}(c_i) : i \in I\} \text{ is linearly independent.}$$

We do this recursively as follows:

First, choose $c_0 \in X$ such that $\bar{f}(c_0) \neq \bar{0}$ (which can be done, since $\bar{f}(X) \not\subseteq \{\bar{0}\}$). Now, assume that c_0, \dots, c_{n-1} have been constructed so

that, for all $I \in \binom{[n]}{\leq d}$, (1) holds. Consider the set

$$U = \bigcup_{I \in \binom{[n]}{< d}} \text{span} \{ \bar{f}(c_i) : i \in I \}.$$

This is a union of finitely many proper subspaces of F^d , so, by the assumption, there exists $c_n \in X$ such that $\bar{f}(c_n) \notin U$. It is easy to check that, for all $I \in \binom{[n+1]}{\leq d}$, (1) holds. This concludes our construction.

For all $I \in \binom{[n]}{d-1}$, consider the homogeneous system

$$\begin{bmatrix} f_0(c_{i_0}) & \cdots & f_{d-1}(c_{i_0}) \\ \vdots & \ddots & \vdots \\ f_0(c_{i_{d-2}}) & \cdots & f_{d-1}(c_{i_{d-2}}) \end{bmatrix} \begin{bmatrix} x_0 \\ \vdots \\ x_{d-1} \end{bmatrix} = \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix},$$

where $i_0 < \cdots < i_{d-2}$ enumerate I . Since the matrix has rank $d-1$, this has a non-trivial solution $\bar{b}_I \in F^d$. That is, for all $i \in I$, $\bar{b}_I \cdot \bar{f}(c_i) = 0$. On the other hand, for any $j \in \mathbb{N} \setminus I$, $\{\bar{f}(c_i) : i \in I \cup \{j\}\}$ is linearly independent, so we have that

$$\begin{bmatrix} f_0(c_{i_0}) & \cdots & f_{d-1}(c_{i_0}) \\ \vdots & \ddots & \vdots \\ f_0(c_{i_{d-2}}) & \cdots & f_{d-1}(c_{i_{d-2}}) \\ f_0(c_j) & \cdots & f_{d-1}(c_j) \end{bmatrix} \bar{b}_I \neq \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 0 \end{bmatrix}.$$

(This holds since $\bar{b}_I \neq 0$ and this matrix has rank d .) Therefore, we see that $\bar{b}_I \cdot f(c_j) \neq 0$. In other words, we have constructed, for each $I \in \binom{[n]}{d-1}$, $\bar{b}_I \in F^d$ such that, for all $i \in \mathbb{N}$,

$$\bar{b}_I \cdot \bar{f}(c_i) = 0 \iff i \in I.$$

In other words,

$$c_i \in Z_{\bar{f}, \bar{b}_I} \iff i \in I.$$

For any $n \geq d$, for any $I \in \binom{[n]}{< d}$, set $I' = I \cup \{n, n+1, \dots, n+(d-|I|-2)\}$ and we have, for all $i \in [n]$,

$$c_i \in Z_{\bar{f}, \bar{b}_{I'}} \iff i \in I.$$

Thus,

$$|\mathcal{C}_{\bar{f}}|_{\{c_i : i \in [n]\}} \geq \binom{n}{< d}.$$

Thus, since $\text{VCdim}(\mathcal{C}_{\bar{f}}) = d-1$, $\mathcal{C}_{\bar{f}}$ is maximal of VC-dimension $d-1$. \square

Unfortunately, we do not necessarily get an upper bound on the VC-density of $\mathcal{C}_{\bar{f}}$, even if $\bar{f}(X)$ is contained in a union of proper subspaces of F^d .

Example 4.4. Fix $d \in \mathbb{N}$ with $d \geq 3$ and F an infinite field. For each $i < d$, let $\bar{e}_i \in F^d$ denote the i th standard basis vector. For each $i < d - 1$, let

$$L_i = \text{span} \{ \bar{e}_0, \bar{e}_{i+1} \}.$$

Let $X = \bigcup_{i=0}^{d-2} L_i$ and let $\bar{f} : X \rightarrow F^d$ be the identity embedding. Then, $\text{VCden}(\mathcal{C}_{\bar{f}}) = d - 1$, even though $\bar{f}(X)$ is contained in the union of 2-dimensional subspaces of F^d .

Proof. Since F is infinite, there exists $g : \mathbb{N} \rightarrow F \setminus \{0\}$ an injective function. For each $i < d - 1$ and $j \in \mathbb{N}$, define

$$\bar{c}_{i,j} = \bar{e}_0 + g(j)\bar{e}_{i+1}.$$

For each $j_0, \dots, j_{d-2} \in \mathbb{N}$, let

$$\bar{b}_{j_0, \dots, j_{d-2}} = \left(\prod_{k < d-1} g(j_k) \right) \bar{e}_0 + \sum_{\ell < d-1} \left(- \prod_{k < d-1, k \neq \ell} g(j_k) \right) \bar{e}_{\ell+1}.$$

Then, for all $i < d - 1$, $j \in \mathbb{N}$, and $j_0, \dots, j_{d-2} \in \mathbb{N}$,

$$\begin{aligned} \bar{b}_{j_0, \dots, j_{d-2}} \cdot \bar{f}(c_{i,j}) &= \bar{b}_{j_0, \dots, j_{d-2}} \cdot \bar{c}_{i,j} = \\ & \prod_{k < d-1} g(j_k) - g(j) \prod_{k < d-1, k \neq i} g(j_k) = \\ & (g(j_i) - g(j)) \prod_{k < d-1, k \neq i} g(j_k). \end{aligned}$$

Therefore, $\bar{c}_{i,j} \in Z_{\bar{f}, \bar{b}_{j_0, \dots, j_{d-2}}}$ if and only if $g(j) = g(j_i)$ if and only if $j = j_i$. Thus, for any $n \in \mathbb{N}$,

$$|\mathcal{C}_{\bar{f}}|_{\{\bar{c}_{i,j} : i < d-1, j < n\}} \geq n^{d-1}.$$

Thus, $\text{VCden}(\mathcal{C}_{\bar{f}}) \geq d - 1$. However, by Theorem 3.10 and Remark 2.9, we conclude that $\text{VCden}(\mathcal{C}_{\bar{f}}) = d - 1$. \square

Corollary 4.5. *Let X be a set, F a field, $d \in \mathbb{N}$ with $d \geq 2$, and $\bar{f} : X \rightarrow F^d$. If $\bar{f}(X)$ is not contained in a finite union of proper subspaces of F^d , then $\mathcal{C}_{\bar{f}}$ is maximal of Littlestone dimension $d - 1$. In particular, $\text{Lden}(\mathcal{C}_{\bar{f}}) = d - 1$.*

Proof. By Proposition 4.3, $\mathcal{C}_{\bar{f}}$ is maximal of VC dimension $d - 1$. By Theorem 3.10, $\mathcal{C}_{\bar{f}}$ has Littlestone dimension $d - 1$. By Remark 2.9, $\mathcal{C}_{\bar{f}}$ is maximal of Littlestone dimension $d - 1$. \square

5. CHARACTERIZATION OF MAXIMALITY OF SYSTEMS OF ZERO SETS.

In this section, we prove the converse to Proposition 4.3.

Definition 5.1. Let F be a field, $d \in \mathbb{Z}^+$, $S \subseteq F^d$ finite, and $\mathcal{C} \subseteq \mathcal{P}(S)$. We say that \mathcal{C} is *span injective* if

- (1) for all $A \in \mathcal{C}$, $\text{span}(A) \neq F^d$, and
- (2) for all $A, B \in \mathcal{C}$, if $\text{span}(A) = \text{span}(B)$, then $A = B$.

In other words, \mathcal{C} is span injective if span is an injective function from \mathcal{C} to the set of all proper subsets of F^d .

Lemma 5.2. Let X be a set, F a field, $d \in \mathbb{Z}^+$, and $\bar{f} : X \rightarrow F^d$. Then, the set

$$\{\bar{f}(Z_{\bar{f}, \bar{a}}) : \bar{a} \in F^d \setminus \{\bar{0}\}\}$$

is span injective.

Proof. For all $\bar{a} \in F^d \setminus \{\bar{0}\}$, for all $c \in Z_{\bar{f}, \bar{a}}$, $\bar{a} \cdot \bar{f}(c) = 0$, so every element of $\bar{f}(Z_{\bar{f}, \bar{a}})$ is orthogonal to \bar{a} . Hence, $\text{span}(\bar{f}(Z_{\bar{f}, \bar{a}})) \neq F^d$. Suppose that $\bar{a}, \bar{b} \in F^d \setminus \{\bar{0}\}$ and $\text{span}(\bar{f}(Z_{\bar{f}, \bar{a}})) = \text{span}(\bar{f}(Z_{\bar{f}, \bar{b}}))$. Then, as \bar{a} is orthogonal to $\bar{f}(Z_{\bar{f}, \bar{a}})$, it is orthogonal to $\text{span}(\bar{f}(Z_{\bar{f}, \bar{a}}))$, hence it is orthogonal to $\text{span}(\bar{f}(Z_{\bar{f}, \bar{b}}))$, so \bar{a} is orthogonal to $\bar{f}(Z_{\bar{f}, \bar{b}})$. Thus, for any $c \in Z_{\bar{f}, \bar{b}}$, $\bar{a} \cdot \bar{f}(c) = 0$, so $c \in Z_{\bar{f}, \bar{a}}$. By symmetry, we obtain that $Z_{\bar{f}, \bar{a}} = Z_{\bar{f}, \bar{b}}$. \square

Lemma 5.3. Let F be a field, $d \in \mathbb{Z}^+$, $S \subseteq F^d$ finite, and $\mathcal{C} \subseteq \mathcal{P}(S)$. If \mathcal{C} is span injective, then there exists $\mathcal{C}' \subseteq \mathcal{P}(S)$ such that

- (1) $|\mathcal{C}'| = |\mathcal{C}|$,
- (2) \mathcal{C}' is span injective,
- (3) the elements of \mathcal{C}' are linearly independent, and
- (4) $\mathcal{C}' \subseteq \binom{S}{< d}$.

Proof. For each $A \in \mathcal{C}$, let $I_A \subseteq A$ be minimal such that $\text{span}(I_A) = \text{span}(A)$ and let

$$\mathcal{C}' = \{I_A : A \in \mathcal{C}\}.$$

Clearly \mathcal{C}' is span injective. Fix $A \in \mathcal{C}$. Since I_A is chosen to be minimal, I_A is linearly independent. Since \mathcal{C} is span injective, $\text{span}(I_A) = \text{span}(A) \neq F^d$, so $|I_A| < d$. Thus,

$$\mathcal{C}' \subseteq \binom{S}{< d}.$$

Fix $A, B \in \mathcal{C}$ and suppose that $I_A = I_B$. Then,

$$\text{span}(A) = \text{span}(I_A) = \text{span}(I_B) = \text{span}(B).$$

Since \mathcal{C} is span injective, $A = B$. Therefore, $A \mapsto I_A$ is a bijection from \mathcal{C} to \mathcal{C}' , so $|\mathcal{C}| = |\mathcal{C}'|$. \square

Lemma 5.4. *Let F be a field, $d \in \mathbb{Z}^+$, $S \subseteq F^d$ finite, and $\mathcal{C} \in \mathcal{P}(S)$. If $|S| > k(d-1)$, S is contained in the union of k proper subsets of F^d , and \mathcal{C} is span injective, then*

$$|\mathcal{C}| < \binom{|S|}{< d}.$$

Proof. Let \mathcal{C}' be given as in Lemma 5.3. Since $|\mathcal{C}| = |\mathcal{C}'|$ and $\mathcal{C}' \subseteq \binom{S}{< d}$, it suffices to show that there exists a $(d-1)$ -element subset of S not in \mathcal{C}' .

Since S is contained in the union of k proper subsets of F^d and $|S| > k(d-1)$, by the pigeonhole principle, there exists a proper subspace L of F^d that contains at least d elements of S . Suppose that there exists $A, B \in \binom{S \cap L}{d-1}$ such that $A, B \in \mathcal{C}'$. Since A and B are linearly independent of cardinality $d-1$ and L has dimension less than d , $\text{span}(A) = \text{span}(B) = L$. Since \mathcal{C}' is span injective, $A = B$. That is, only one element of $\binom{S \cap L}{d-1}$ belongs to \mathcal{C}' .

Therefore,

$$|\mathcal{C}| = |\mathcal{C}'| < \binom{|S|}{< d}.$$

\square

We are ready to establish the converse to Proposition 4.3.

Proposition 5.5. *Let X be a set, F a field, $d \in \mathbb{N}$ with $d \geq 2$, and $\bar{f} : X \rightarrow F^d$. If $\bar{f}(X)$ is contained in a finite union of proper subspaces of F^d , then $\mathcal{C}_{\bar{f}}$ is not maximal of VC-dimension $d-1$.*

Proof. Let $n = k(d-1) + 1$ and fix some $X_0 \in \binom{X}{n}$. Consider $\bar{f}|_{X_0} : X_0 \rightarrow F^d$ and, for each $\bar{a} \in F^d \setminus \{\bar{0}\}$, the corresponding $Z_{\bar{f}|_{X_0}, \bar{a}} \in \mathcal{P}(X_0)$. By Lemma 5.2, the set

$$\mathcal{C} = \{\bar{f}|_{X_0}(Z_{\bar{f}|_{X_0}, \bar{a}}) : \bar{a} \in F^d \setminus \{\bar{0}\}\}$$

is span injective. By Lemma 5.4,

$$|\mathcal{C}| < \binom{n}{< d}.$$

It is clear that, for all $\bar{a} \in F^d \setminus \{\bar{0}\}$,

$$Z_{\bar{f}, \bar{a}} \cap X_0 = Z_{\bar{f}|_{X_0}, \bar{a}}.$$

Thus, $\mathcal{C}_{\bar{f}}|_{X_0} = \mathcal{C}_{\bar{f}|_{X_0}}$. Therefore,

$$|\mathcal{C}_{\bar{f}}|_{X_0}| = |\mathcal{C}_{\bar{f}|_{X_0}}| \leq |\mathcal{C}| < \binom{n}{< d}.$$

Thus, $\mathcal{C}_{\bar{f}}$ is not maximal of VC-dimension $d - 1$. \square

We thus obtain the following result.

Theorem 5.6. *Let X be a set, F a field, $d \in \mathbb{N}$ with $d \geq 2$, and $\bar{f} : X \rightarrow F^d$. Then, the following are equivalent:*

- (1) $\mathcal{C}_{\bar{f}}$ is maximal of VC-dimension $d - 1$;
- (2) $\bar{f}(X)$ is not contained in a finite union of proper subspaces of F^d .

Proof. (1) \Rightarrow (2) follows from the contrapositive of Proposition 5.5 and (2) \Rightarrow (1) follows from Proposition 4.3. \square

It turns out that the conditions in Theorem 5.6 are not equivalent to $\mathcal{C}_{\bar{f}}$ being maximal Littlestone dimension $d - 1$.

Example 5.7. Consider the function $\bar{f} : X \rightarrow F^d$ in Example 4.4. That is, F is an infinite field, $d \geq 3$,

$$X = \bigcup_{i=0}^{d-2} \text{span}\{\bar{e}_0, \bar{e}_{i+1}\},$$

and $\bar{f} : X \rightarrow F^d$ is the identity embedding. Then, $\mathcal{C}_{\bar{f}}$ is maximal of Littlestone dimension $d - 1$, even though $\bar{f}(X)$ is contained in a union of 2-dimensional subspaces of F^d .

Proof. Define $\bar{c}_{i,j}$ for $i \in [d - 1]$ and $j \in \mathbb{N}$ and define $\bar{b}_{j_0, \dots, j_{d-2}}$ for $j_0, \dots, j_{d-2} \in \mathbb{N}$ as in Example 4.4. In particular, $\bar{c}_{i,j} \in X$ for all $i \in [d - 1]$ and $j \in \mathbb{N}$ and

$$\bar{c}_{i,j} \in Z_{\bar{f}, \bar{b}_{j_0, \dots, j_{d-2}}} \iff j = j_i$$

for all $i \in [d - 1]$, $j, j_0, \dots, j_{d-2} \in \mathbb{N}$. Fix $n \in \mathbb{N}$ and we define a $(X, \mathcal{C}_{\bar{f}})$ -labeled tree T of depth n with $\binom{n}{< d}$ well-labeled leaves, showing that $\mathcal{C}_{\bar{f}}$ is maximal of Littlestone dimension $d - 1$.

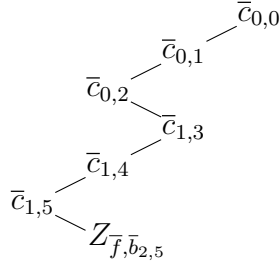
For each $k \in [n]$ and $\sigma \in [k][2]$, let $i = |\text{supp}(\sigma)|$ and let

$$T(\sigma) = \begin{cases} \bar{c}_{i,k} & \text{if } i < d - 1 \\ \bar{c}_{0,n} & \text{otherwise} \end{cases}$$

For $\tau \in [n][2]$, let $j_0 < j_1 < \cdots < j_{\ell-1}$ enumerate the set $\text{supp}(\sigma)$. If $\ell < d-1$, let $j_\ell = j_{\ell+1} = \cdots = j_{d-2} = n$. Then, let

$$T(\tau) = Z_{\bar{f}, \bar{b}_{j_0, \dots, j_{d-2}}}.$$

We claim that, for each $\tau \in [n][2]$, if $|\text{supp}(\tau)| < d$, then $(\tau, T(\tau))$ is well-labeled. This exhibits $\binom{n}{<d}$ well-labeled leaves of T .



Fix $\tau \in [n][2]$ with $|\text{supp}(\tau)| < d$. Say $j_0 < \cdots < j_{\ell-1}$ enumerates $\text{supp}(\tau)$. Clearly $\ell \leq d-1$. Set $j_\ell = j_{\ell+1} = \cdots = j_{d-2} = n$ and $j_{-1} = -1$. Fix $k \in [n]$. If $j_{i-1} < k \leq j_i$ for some $i \in [d-1]$, then

$$|\text{supp}(\tau|_{[k]})| = |\{j_0, \dots, j_{i-1}\}| = i.$$

Therefore, $T(\tau|_{[k]}) = \bar{c}_{i,k}$. Moreover,

$$T(\tau) = Z_{\bar{f}, \bar{b}_{j_0, \dots, j_{d-2}}}.$$

By construction, $T(\tau|_{[k]}) \in T(\tau)$ if and only if $k = j_i$ if and only if $\tau(k) = 1$. Finally, if $j_{d-2} < k < n$, then $|\text{supp}(\tau|_{[k]})| = d-1$, so $T(\tau|_{[k]}) = \bar{c}_{0,n}$. Thus, since $j_0 < n$, $T(\tau|_{[k]}) \notin T(\tau)$. Moreover, $\tau(k) = 0$.

In any case, we see that, for all $k \in [n]$, $T(\tau|_{[k]}) \in T(\tau)$ if and only if $\tau(k) = 1$. Therefore, we conclude that $(\tau, T(\tau))$ is well-labeled. \square

6. CONCLUSION

We have determined the VC-dimension and Littlestone dimension of $\mathcal{C}_{\bar{f}}$, as well as given a characterization for when $\mathcal{C}_{\bar{f}}$ is maximal in VC-dimension. We can apply these results, for example, to the set of conic sections in \mathbb{R}^2 via Example 3.3.

Example 6.1. Let $X = \mathbb{R}^2$, $F = \mathbb{R}$, $d = 6$, and

$$\bar{f}(x, y) = (x^2, xy, y^2, x, y, 1).$$

As noted above, $\mathcal{C}_{\bar{f}}$ is the set of all conic sections in \mathbb{R}^2 . For example, we have in the image of \bar{f} ,

$$\begin{aligned}\bar{f}(0, 0) &= (0, 0, 0, 0, 0, 1), \\ \bar{f}(1, 0) &= (1, 0, 0, 1, 0, 1), \\ \bar{f}(0, 1) &= (0, 0, 1, 0, 1, 1), \\ \bar{f}(1, 1) &= (1, 1, 1, 1, 1, 1), \\ \bar{f}(2, 1) &= (4, 2, 1, 2, 1, 1), \\ \bar{f}(1, 2) &= (1, 2, 4, 1, 2, 1).\end{aligned}$$

It is easy to check that these vectors form a basis for \mathbb{R}^6 . By Lemma 3.4, \bar{f} is linearly independent. Thus, by Theorem 3.10, $\mathcal{C}_{\bar{f}}$ has VC-dimension and Littlestone dimension 5. Moreover, $\bar{f}(\mathbb{R}^2)$ is not contained in a finite union of proper subspaces of \mathbb{R}^6 .

Towards a contradiction, suppose that $\bar{f}(\mathbb{R}^2)$ is contained in finitely many proper subspaces of \mathbb{R}^6 , say L_0, \dots, L_{m-1} . For each $i < m$, suppose that $\bar{a}_i \in \mathbb{R}^6$ is orthogonal to L_i . For each $\bar{b} \in \mathbb{R}^2$, $\bar{f}(\bar{b}) \in \bar{f}(\mathbb{R}^2)$, hence there exists $i < m$ such that $\bar{f}(\bar{b}) \in L_i$. That is, $\bar{a}_i \cdot \bar{f}(\bar{b}) = 0$, so $\bar{b} \in Z_{\bar{f}, \bar{a}_i}$. Thus,

$$\mathbb{R}^2 \subseteq \bigcup_{i < m} Z_{\bar{f}, \bar{a}_i}.$$

However, each $Z_{\bar{f}, \bar{a}_i}$ is a conic section. This is a contradiction, since we cannot cover \mathbb{R}^2 with finitely many conic sections.

Therefore, by Theorem 5.6, $\mathcal{C}_{\bar{f}}$ is maximal of VC-dimension 5. By Corollary 4.5, $\mathcal{C}_{\bar{f}}$ is maximal of Littlestone dimension 5. In particular, the set of conic sections in \mathbb{R}^2 has VC-density 5 and Littlestone density 5.

The results of this paper can be applied to more general situations. For example, consider the set of axes-aligned ellipses. We can use Theorem 3.10 to compute the VC-dimension and Littlestone dimension of this class.

Example 6.2. Let \mathcal{C} be the set of axes-aligned ellipses in \mathbb{R}^2 . Formally, let

$$E_{a,b,c,d} = \{(x, y) \in \mathbb{R}^2 : a(x - b)^2 + c(y - d)^2 = 1\}$$

and let \mathcal{C} be the class of all $E_{a,b,c,d}$ where $a, b, c, d \in \mathbb{R}$ and $a, c > 0$. Note that \mathcal{C} is a subclass of the class $\mathcal{C}_{\bar{f}}$, where $\bar{f} : \mathbb{R}^2 \rightarrow \mathbb{R}^5$ given by

$$\bar{f}(x, y) = (x^2, y^2, x, y, 1).$$

It is easy to show that \bar{f} is linearly independent. By Theorem 3.10, $\mathcal{C}_{\bar{f}}$ has VC-dimension and Littlestone dimension 4. Moreover, it is not hard to show that \mathcal{C} has VC-dimension at least 4 (for example, the set $\{(1, 0), (0, 1), (-1, 0), (0, -1)\}$ is shattered by \mathcal{C}). Therefore,

$$4 \leq \text{VCdim}(\mathcal{C}) \leq \text{Ldim}(\mathcal{C}) \leq \text{Ldim}(\mathcal{C}_{\bar{f}}) = 4.$$

Thus, \mathcal{C} has VC-dimension and Littlestone dimension 4. On the other hand, although we can show that $\mathcal{C}_{\bar{f}}$ is maximal of VC-dimension and Littlestone dimension 4, it is more difficult to apply Theorem 5.6 on the class of axes-aligned ellipses.

In general, it may be interesting to examine what happens when we restrict the parameters of $\mathcal{C}_{\bar{f}}$. Under what conditions does the VC-dimension or Littlestone dimension drop below $d - 1$? What can be said about the maximality of VC-dimension or Littlestone dimension in this case?

REFERENCES

- [1] S. Bhaskar, *Thicket Density*, Journal of Symbolic Logic **86** (2021), no. 1, 110–127.
- [2] H. Chase and J. Freitag, *Model Theory and Machine Learning*, The Bulletin of Symbolic Logic **25** (2019), no. 3, 319–332.
- [3] ———, *Model theory and combinatorics of banned sequences*, The Journal of Symbolic Logic (2020), 1–19.
- [4] T. M Cover, *Geometrical and statistical properties of systems of linear inequalities with applications in pattern recognition*, IEEE transactions on electronic computers **3** (1965), 326–334.
- [5] R. M Dudley, *Balls in \mathbb{R}^k do not cut all subsets of $k + 2$ points*, Advances in Mathematics **31** (1979), no. 3, 306–308.
- [6] S. Floyd, *Space-bounded learning and the Vapnik-Chervonenkis dimension*, Proceedings of the second annual workshop on Computational learning theory, 1989, pp. 349–364.
- [7] H. R Johnson, *Some new maximum VC classes*, Information Processing Letters **114** (2014), no. 6, 294–298.

TOWSON UNIVERSITY, 7800 YORK RD., TOWSON, MD, 21252

Email address: vguingona@towson.edu

Email address: akolesnikov@towson.edu