Learning Spheres and Chains in $\mathbb{F}_q^d$

Ryan Jeong, Wyatt Milgrim

Joint work with Ruben Ascoli, Livia Betti, Justin Cheigh, Xuyan Liu, Brian McDonald, Francisco Romero, and Santiago Velazquez; Advisors: Alex Iosevich and Steven J. Miller

SMALL REU 2022, Williams College

Young Mathematicians Conference

August 14 2022

Preliminaries

Preliminaries

- We let $\mathbb{F}_q^d$ denote the d-dimensional vector space over the finite field $\mathbb{F}_q$.

Preliminaries

- We let $\mathbb{F}_q^d$ denote the d-dimensional vector space over the finite field $\mathbb{F}_q$.
- We denote $||x|| = x_1^2 + x_2^2 ... \in \mathbb{F}_q$, where $x = (x_1, ..., x_d) \in \mathbb{F}_q^d$. (This is not technically a norm but behaves like one for our purposes.)

Preliminaries

- We let $\mathbb{F}_q^d$ denote the d-dimensional vector space over the finite field $\mathbb{F}_q$.
- We denote $||x|| = x_1^2 + x_2^2 ... \in \mathbb{F}_q$, where $x = (x_1, ..., x_d) \in \mathbb{F}_q^d$. (This is not technically a norm but behaves like one for our purposes.)
- For the remainder of the talk we let t be a fixed, nonzero element of the finite field $\mathbb{F}_q$ (i.e. this is not a variable).

VC-Dimension

VC-Dimension

Fix a set E and a collection of functions $\mathcal{H}$ from E to $\{0,1\}$.

VC-Dimension

Fix a set E and a collection of functions $\mathcal{H}$ from E to $\{0, 1\}$.

Definition

Say $\mathcal{H}$ *shatters* a finite subset $A \subset E$ if restricting functions in $\mathcal{H}$ to A yields all $2^{|A|}$ functions from A to $\{0, 1\}$.

VC-Dimension

Fix a set E and a collection of functions $\mathcal{H}$ from E to $\{0, 1\}$.

Definition

Say $\mathcal{H}$ *shatters* a finite subset $A \subset E$ if restricting functions in $\mathcal{H}$ to A yields all $2^{|A|}$ functions from A to $\{0, 1\}$.

In other words, functions in the collection $\mathcal{H}$ realize all possible behaviors on the subset $A \subset E$.

VC-Dimension

Fix a set E and a collection of functions $\mathcal{H}$ from E to $\{0,1\}$.

Definition

Say $\mathcal{H}$ *shatters* a finite subset $A \subset E$ if restricting functions in $\mathcal{H}$ to A yields all $2^{|A|}$ functions from A to $\{0,1\}$.

In other words, functions in the collection $\mathcal{H}$ realize all possible behaviors on the subset $A \subset E$.

Definition (Vapnik and Chervonenkis, 1968)

The *Vapnik–Chervonenkis dimension* (VCD) of $\mathcal{H}$ is the cardinality of the largest subset $A \subset E$ that is shattered by $\mathcal{H}$.

VC-Dimension

Fix a set E and a collection of functions $\mathcal{H}$ from E to $\{0,1\}$.

Definition

Say $\mathcal{H}$ *shatters* a finite subset $A \subset E$ if restricting functions in $\mathcal{H}$ to A yields all $2^{|A|}$ functions from A to $\{0,1\}$.

In other words, functions in the collection $\mathcal{H}$ realize all possible behaviors on the subset $A \subset E$.

Definition (Vapnik and Chervonenkis, 1968)

The *Vapnik–Chervonenkis dimension* (VCD) of $\mathcal{H}$ is the cardinality of the largest subset $A \subset E$ that is shattered by $\mathcal{H}$.

Explicitly, $VCD(\mathcal{H}) = n$ if there exists $A \subset E$, $|A| = n$ such that A is shattered by $\mathcal{H}$, but no such subset of size $n + 1$.

VC-Dimension: Example

VC-Dimension: Example

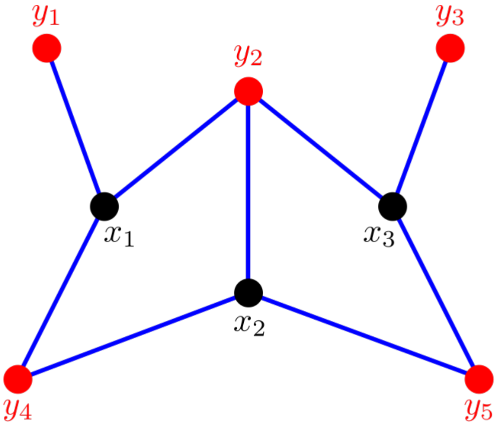


Figure: $E = \{x_1, x_2, x_3\}$ and classifiers in $\mathcal{H}$ correspond to spheres of some fixed radius centered at points y_i. We can shatter $A = \{x_1, x_2\} \subset E$: centers $\{y_1, y_3, y_4, y_5\}$ yield all four possible behaviors. Here, $\text{VCD}(\mathcal{H}) = 2$, since there is no y_i adjacent to strictly x_1 and x_3.

Learning Problems: Motivation

Learning Problems: Motivation

VC-dimension has historically been of central importance in *statistical and computational learning theory*.

Learning Problems: Motivation

VC-dimension has historically been of central importance in *statistical and computational learning theory*.

It is a key object of study in proofs yielding bounds on the number of data points needed to efficiently learn a collection of classifiers.

Learning Problems: Motivation

VC-dimension has historically been of central importance in *statistical and computational learning theory*.

It is a key object of study in proofs yielding bounds on the number of data points needed to efficiently learn a collection of classifiers.

Fundamental Theorem of Statistical Learning:

Theorem (Vapnik, 1999)

The collection of classifiers $\mathcal{H}$ is PAC-learnable iff $VCD(\mathcal{H}) < \infty$.

Learning Problems: Motivation

VC-dimension has historically been of central importance in *statistical and computational learning theory*.

It is a key object of study in proofs yielding bounds on the number of data points needed to efficiently learn a collection of classifiers.

Fundamental Theorem of Statistical Learning:

Theorem (Vapnik, 1999)

The collection of classifiers $\mathcal{H}$ is PAC-learnable iff $VCD(\mathcal{H}) < \infty$.

Motivated by these seminal results, we can study *learning natural families of classifiers over finite fields.*

Prior Work: Learning Spheres in $\mathbb{F}_q^2$

Prior Work: Learning Spheres in $\mathbb{F}_q^2$

Spheres in $\mathbb{F}_q^2$: Fix $t \neq 0$. For $E \subset \mathbb{F}_q^2$, define the class of functions $\mathcal{H}_t^2(E) = \{h_y : y \in E\}$, where $h_y : E \to \{0,1\}$ is the indicator function for the sphere of radius t centered at y:

$$h_y(x) = \begin{cases} 1 & ||y-x|| = t \\ 0 & ||y-x|| \neq t. \end{cases}$$

Prior Work: Learning Spheres in $\mathbb{F}_q^2$

Spheres in $\mathbb{F}_q^2$: Fix $t \neq 0$. For $E \subset \mathbb{F}_q^2$, define the class of functions $\mathcal{H}_t^2(E) = \{h_y : y \in E\}$, where $h_y : E \to \{0,1\}$ is the indicator function for the sphere of radius t centered at y:

$$h_y(x) = \begin{cases} 1 & ||y-x|| = t \\ 0 & ||y-x|| \neq t. \end{cases}$$

Theorem (Fitzpatrick, Iosevich, McDonald, and Wyman, 2021)

If $|E| \geq Cq^{15/8}$ for C large, then $VCD\left(\mathcal{H}_t^2(E)\right) = 3$.

Prior Work: Learning Spheres in $\mathbb{F}_q^2$

Spheres in $\mathbb{F}_q^2$: Fix $t \neq 0$. For $E \subset \mathbb{F}_q^2$, define the class of functions $\mathcal{H}_t^2(E) = \{h_y : y \in E\}$, where $h_y : E \to \{0,1\}$ is the indicator function for the sphere of radius t centered at y:

$$h_y(x) = \begin{cases} 1 & ||y-x|| = t \\ 0 & ||y-x|| \neq t. \end{cases}$$

Theorem (Fitzpatrick, Iosevich, McDonald, and Wyman, 2021)

If $|E| \geq Cq^{15/8}$ for C large, then $VCD\left(\mathcal{H}_t^2(E)\right) = 3$.

- Three points determine a circle in $\mathbb{F}_q^2$, so VCD $\left(\mathcal{H}_t^2(E)\right) \leq 3$. More generally, VCD $\left(\mathcal{H}_t^d(E)\right) \leq d+1$.

Prior Work: Learning Spheres in $\mathbb{F}_q^2$

Spheres in $\mathbb{F}_q^2$: Fix $t \neq 0$. For $E \subset \mathbb{F}_q^2$, define the class of functions $\mathcal{H}_t^2(E) = \{h_y : y \in E\}$, where $h_y : E \to \{0,1\}$ is the indicator function for the sphere of radius t centered at y:

$$h_y(x) = \begin{cases} 1 & ||y-x|| = t \\ 0 & ||y-x|| \neq t. \end{cases}$$

Theorem (Fitzpatrick, Iosevich, McDonald, and Wyman, 2021)

If $|E| \geq Cq^{15/8}$ for C large, then $VCD\left(\mathcal{H}_t^2(E)\right) = 3$.

- Three points determine a circle in $\mathbb{F}_q^2$, so VCD $\left(\mathcal{H}_t^2(E)\right) \leq 3$. More generally, VCD $\left(\mathcal{H}_t^d(E)\right) \leq d+1$.
- However, the theorem does not generalize easily to higher dimensions (i.e. the task of learning spheres $\mathbb{F}_q^d$ with $d \geq 3$).

New Classifiers: Learning 2-Chains in $\mathbb{F}_q^d$

New Classifiers: Learning 2-Chains in $\mathbb{F}_q^d$

2-Chains in $\mathbb{F}_q^d$: Fix $t \neq 0$, $d \geq 3$, and $E \subset \mathbb{F}_q^d$. Define the collection of functions $\mathcal{H}_t^d(E) = \{h_{y,z} : y, z \in E, y \neq z\}$, where $h_{y,z} : E \to \{0,1\}$ is the indicator function for the intersection of spheres of radius t centered at y and z:

$$h_{y,z}(x) = \begin{cases} 1 & \text{if } ||y-x|| = ||z-x|| = t \\ 0 & \text{otherwise.} \end{cases}$$

New Classifiers: Learning 2-Chains in $\mathbb{F}_q^d$

2-Chains in $\mathbb{F}_q^d$: Fix $t \neq 0$, $d \geq 3$, and $E \subset \mathbb{F}_q^d$. Define the collection of functions $\mathcal{H}_t^d(E) = \{h_{y,z} : y, z \in E, y \neq z\}$, where $h_{y,z} : E \to \{0, 1\}$ is the indicator function for the intersection of spheres of radius t centered at y and z:

$$h_{y,z}(x) = \begin{cases} 1 & \text{if } ||y - x|| = ||z - x|| = t \\ 0 & \text{otherwise.} \end{cases}$$

Theorem (SMALL 2022++)

If $|E| \geq \begin{cases} Cq^{\frac{7}{3}} & d = 3 \\ Cq^{d-\frac{1}{d-1}} & d \geq 4 \end{cases}$

Where C depends on d but not q, then $VCD\left(\mathcal{H}_t^d(E)\right) = d$.

New Classifiers: Learning 2-Chains in $\mathbb{F}_q^d$

2-Chains in $\mathbb{F}_q^d$: Fix $t \neq 0$, $d \geq 3$, and $E \subset \mathbb{F}_q^d$. Define the collection of functions $\mathcal{H}_t^d(E) = \{h_{y,z} : y, z \in E, y \neq z\}$, where $h_{y,z} : E \to \{0,1\}$ is the indicator function for the intersection of spheres of radius t centered at y and z:

$$h_{y,z}(x) = \begin{cases} 1 & \text{if } ||y - x|| = ||z - x|| = t \\ 0 & \text{otherwise.} \end{cases}$$

Theorem (SMALL 2022++)

If $|E| \geq \begin{cases} Cq^{\frac{7}{3}} & d = 3 \\ Cq^{d - \frac{1}{d-1}} & d \geq 4 \end{cases}$

Where C depends on d but not q, then $VCD\left(\mathcal{H}_t^d(E)\right) = d$.

Note that we have the bound VCD $\left(\mathcal{H}_t^d(E)\right) \leq d$, since $d+1$ points determine a sphere in $\mathbb{F}_q^d$.

Definitions

Definitions

Definition

Fix $E \subset \mathbb{F}_q^d$, and let $A \subset E$. Call $x \in E$ a *pole* of A if $||x - a|| = t$ whenever $a \in A$, and let $\text{Pole}(A) \subset E$ denote the set of poles of A.

Definitions

Definition

Fix $E \subset \mathbb{F}_q^d$, and let $A \subset E$. Call $x \in E$ a *pole* of A if $||x - a|| = t$ whenever $a \in A$, and let $\text{Pole}(A) \subset E$ denote the set of poles of A.

Definition

The $(d+2)$-tuple $P = (y, z, x^1, \ldots, x^d) \in (\mathbb{F}_q^d)^{d+2}$ is an *d-prism* if for all $i \leq n$, $||x^i - y|| = ||x^i - z|| = t$.

Definitions

Definition

Fix $E \subset \mathbb{F}_q^d$, and let $A \subset E$. Call $x \in E$ a *pole* of A if $||x - a|| = t$ whenever $a \in A$, and let $\text{Pole}(A) \subset E$ denote the set of poles of A.

Definition

The $(d+2)$-tuple $P = (y, z, x^1, \ldots, x^d) \in (\mathbb{F}_q^d)^{d+2}$ is an *d-prism* if for all $i \leq n$, $||x^i - y|| = ||x^i - z|| = t$.

- The *tail* $\mathcal{T}(P)$ of P is the set $\{y, z\}$.

Definitions

Definition

Fix $E \subset \mathbb{F}_q^d$, and let $A \subset E$. Call $x \in E$ a *pole* of A if $||x - a|| = t$ whenever $a \in A$, and let $\mathrm{Pole}(A) \subset E$ denote the set of poles of A.

Definition

The $(d+2)$-tuple $P = (y, z, x^1, \ldots, x^d) \in (\mathbb{F}_q^d)^{d+2}$ is an *d-prism* if for all $i \leq n$, $||x^i - y|| = ||x^i - z|| = t$.

- The *tail* $\mathcal{T}(P)$ of P is the set $\{y, z\}$.
- The *center* $\mathcal{C}(P)$ of P is the set $\{x^1, x^2, \ldots, x^d\}$.

Definitions

Definition

Fix $E \subset \mathbb{F}_q^d$, and let $A \subset E$. Call $x \in E$ a *pole* of A if $||x - a|| = t$ whenever $a \in A$, and let $\text{Pole}(A) \subset E$ denote the set of poles of A.

Definition

The $(d+2)$-tuple $P = (y, z, x^1, \ldots, x^d) \in (\mathbb{F}_q^d)^{d+2}$ is an *d-prism* if for all $i \leq n$, $||x^i - y|| = ||x^i - z|| = t$.

- The *tail* $\mathcal{T}(P)$ of P is the set $\{y, z\}$.
- The *center* $\mathcal{C}(P)$ of P is the set $\{x^1, x^2, \ldots, x^d\}$.

We say a prism is *non-degenerate* if all of its points are distinct.

Prisms: Distance Graph

Prisms: Distance Graph

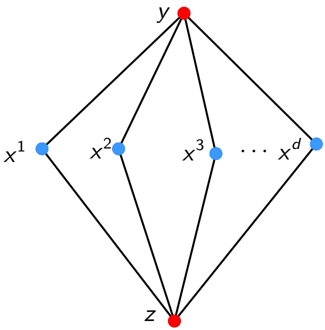


Figure: A nondegenerate d-prism $P = (y, z, x^1, \ldots, x^d)$ with tail $\mathcal{T}(P) = \{y, z\}$ and center $\mathcal{C}(P) = \{x^1, \ldots, x^d\}$. A d-prism can be thought of as possessing the subgraph $K_{2,d}$, with vertices in the edge $\{v, w\}$ edge satisfying $||v - w|| = t$.

Prisms: Distance Graph

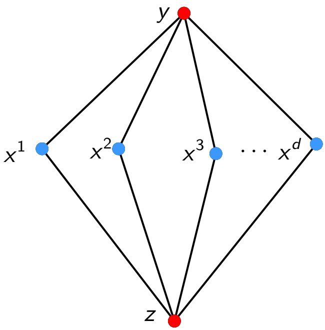


Figure: A nondegenerate d-prism $P = (y, z, x^1, \ldots, x^d)$ with tail $\mathcal{T}(P) = \{y, z\}$ and center $\mathcal{C}(P) = \{x^1, \ldots, x^d\}$. A d-prism can be thought of as possessing the subgraph $K_{2,d}$, with vertices in the edge $\{v, w\}$ edge satisfying $||v - w|| = t$.

Henceforth, the term prism refers to a nondegenerate d-prism.

Using Prisms to Shatter Sets

Using Prisms to Shatter Sets

Why are prisms the right object to study?

Using Prisms to Shatter Sets

Why are prisms the right object to study?

Observation

Using Prisms to Shatter Sets

Why are prisms the right object to study?

Observation

- *Any set* $A \subset E$ *with* $|A| = d$ *shattered by* $\mathcal{H}_t^d(E)$ *is necessarily the center of some prism* P*.*

Using Prisms to Shatter Sets

Why are prisms the right object to study?

Observation

- *Any set* $A \subset E$ *with* $|A| = d$ *shattered by* $\mathcal{H}_t^d(E)$ *is necessarily the center of some prism* P*.*
- *To shatter the center* $\mathcal{C}(P)$ *of a prism* P*, for every subset* $A \subset \mathcal{C}(P)$*, we must find a point distance* t *away only from the points in* A*.*

Proof Strategy

Definition (SMALL 2022++)

Fix a prism $P = (\mathcal{T}, \mathcal{C})$ with center $\mathcal{C}$. A subset $A \subset \mathcal{C}$ with $|A| \geq 2$ is P-bad if every $x \in \text{Pole}(A)$ is also in $\text{Pole}(y)$ for some $y \in \mathcal{C} \setminus A$.

Goal: Find a prism with no bad sets.

Proof Strategy

Definition (SMALL 2022++)

Fix a prism $P = (\mathcal{T}, \mathcal{C})$ with center $\mathcal{C}$. A subset $A \subset \mathcal{C}$ with $|A| \geq 2$ is P-bad if every $x \in \text{Pole}(A)$ is also in $\text{Pole}(y)$ for some $y \in \mathcal{C} \setminus A$.

Goal: Find a prism with no bad sets.

- Count the number of prisms, denoted $N_d(E)$.

Proof Strategy

Definition (SMALL 2022++)

Fix a prism $P = (\mathcal{T}, \mathcal{C})$ with center $\mathcal{C}$. A subset $A \subset \mathcal{C}$ with $|A| \geq 2$ is P-bad if every $x \in \text{Pole}(A)$ is also in $\text{Pole}(y)$ for some $y \in \mathcal{C} \setminus A$.

Goal: Find a prism with no bad sets.

- Count the number of prisms, denoted $N_d(E)$.
- Count the number of prisms that can contain a bad k-set by counting the number of prisms a given k-set can be bad in.

Proof Strategy

Definition (SMALL 2022++)

Fix a prism $P = (\mathcal{T}, \mathcal{C})$ with center $\mathcal{C}$. A subset $A \subset \mathcal{C}$ with $|A| \geq 2$ is P-bad if every $x \in \text{Pole}(A)$ is also in $\text{Pole}(y)$ for some $y \in \mathcal{C} \setminus A$.

Goal: Find a prism with no bad sets.

- Count the number of prisms, denoted $N_d(E)$.
- Count the number of prisms that can contain a bad k-set by counting the number of prisms a given k-set can be bad in.
- Show that, if $|E|$ is large, there must be some prism for which no subset of its center is bad.

Prisms, Pairs, and 2-Paths

Definition

A 2-path in E is a set $(x_1, x_2, x_3) \in E^3$ such that $||x_1 - x_2|| = ||x_2 - x_3|| = t$.

Observation

Up to ordering a prism corresponds to a choice of pair (y, z) and a choice of d 2-paths between them—(y, z) are the tails and the set of midpoints of these paths is the center (there is some concern regarding degenerate prisms but this is easy to deal with).

Prisms, Pairs, and 2-Paths

Definition

A 2-path in E is a set $(x_1, x_2, x_3) \in E^3$ such that $||x_1 - x_2|| = ||x_2 - x_3|| = t$.

Observation

Up to ordering a prism corresponds to a choice of pair (y, z) and a choice of d 2-paths between them—(y, z) are the tails and the set of midpoints of these paths is the center (there is some concern regarding degenerate prisms but this is easy to deal with).

Theorem (Iosevich et al., 2018)

Let $E \subset \mathbb{F}_q^d$ with $|E| > \frac{4}{\ln(2)} q^{\frac{d+1}{2}}$ and let $\Gamma_2(E)$ be the number of 2-paths in E. Then $\Gamma_2(E) = \frac{|E|^3}{q^2} + \mathcal{D}_2(E)$ where $\mathcal{D}_2(E) \leq C \frac{|E|^2}{q}$.

Prisms, Pairs, and 2-Paths

Definition

A 2-path in E is a set $(x_1, x_2, x_3) \in E^3$ such that $||x_1 - x_2|| = ||x_2 - x_3|| = t$.

Observation

Up to ordering a prism corresponds to a choice of pair (y, z) and a choice of d 2-paths between them—(y, z) are the tails and the set of midpoints of these paths is the center (there is some concern regarding degenerate prisms but this is easy to deal with).

Theorem (Iosevich et al., 2018)

Let $E \subset \mathbb{F}_q^d$ with $|E| > \frac{4}{\ln(2)} q^{\frac{d+1}{2}}$ and let $\Gamma_2(E)$ be the number of 2-paths in E. Then $\Gamma_2(E) = \frac{|E|^3}{q^2} + \mathcal{D}_2(E)$ where $\mathcal{D}_2(E) \leq C \frac{|E|^2}{q}$.

Counting Prisms (sketch)

- Given a pair $(x, y) \in E \times E$ we let $k_{(x,y)}$ be the number of 2-paths in E with x, y as endpoints.

Counting Prisms (sketch)

- Given a pair $(x, y) \in E \times E$ we let $k_{(x,y)}$ be the number of 2-paths in E with x, y as endpoints.
- Assume $|E| > 4/\log(2)q^{\frac{d+1}{2}}$. Then

$$N_d(E) = d! \sum_{(x,y)\in E\times E} \binom{k_{(x,y)}}{d}$$

We also have

$$\sum_{(x,y)\in E\times E} k_{(x,y)} \gtrsim \frac{|E|^3}{q^2}.$$

Counting Prisms (sketch)

- Given a pair $(x, y) \in E \times E$ we let $k_{(x,y)}$ be the number of 2-paths in E with x, y as endpoints.
- Assume $|E| > 4/\log(2)q^{\frac{d+1}{2}}$. Then

$$N_d(E) = d! \sum_{(x,y)\in E\times E} \binom{k_{(x,y)}}{d}$$

  We also have

$$\sum_{(x,y)\in E\times E} k_{(x,y)} \gtrsim \frac{|E|^3}{q^2}.$$

- From here it is straightfoward to apply Hölder's Inequality (proof omitted) to obtain $N_d(E) \geq C\frac{|E|^{d+2}}{q^{2d}}$

Finding a Good Prism

Definition (SMALL 2022++)

In the $d = 3$ case we say that a pair $\{a, b\}$ is P-bad for a prism $P = \{(y, z, a, b, x)\}$ if $\text{Pole}(\{a, b\}) \subset \text{Pole}(\{a, b, x\})$. We say a pair is **bad** if it is P-bad for some prism P.

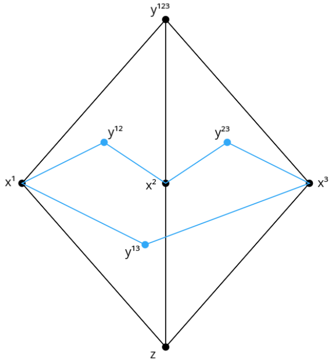
- To shatter 3 points we must find a prism P such that no $A \subset \mathcal{C}(P)$ is P-bad.

Finding a Good Prism

Definition (SMALL 2022++)

In the $d = 3$ case we say that a pair $\{a, b\}$ is P-bad for a prism $P = \{(y, z, a, b, x)\}$ if $\text{Pole}(\{a, b\}) \subset \text{Pole}(\{a, b, x\})$. We say a pair is **bad** if it is P-bad for some prism P.

- To shatter 3 points we must find a prism P such that no $A \subset \mathcal{C}(P)$ is P-bad.



Counting Bad Pairs for $d = 3$ (sketch)

- Three spheres in $\mathbb{F}_q^3$ intersect at at most two points so any bad pair must have exactly two poles.

Counting Bad Pairs for $d = 3$ (sketch)

- Three spheres in $\mathbb{F}_q^3$ intersect at at most two points so any bad pair must have exactly two poles.
- This means that every prism a given pair is bad in must have the same tails.

Counting Bad Pairs for $d = 3$ (sketch)

- Three spheres in $\mathbb{F}_q^3$ intersect at at most two points so any bad pair must have exactly two poles.
- This means that every prism a given pair is bad in must have the same tails.
- There are $\approx q$ choices for the last point of the prism since it must be distance t away from both poles.

Counting Bad Pairs for $d = 3$ (sketch)

- Three spheres in $\mathbb{F}_q^3$ intersect at at most two points so any bad pair must have exactly two poles.
- This means that every prism a given pair is bad in must have the same tails.
- There are $\approx q$ choices for the last point of the prism since it must be distance t away from both poles.
- Finally, observe that there are only $1/2|E|^2$ pairs total, and so at most $C|E|^2 q$ prisms admit a P-bad set. But we just showed there are at least $|E|^5 q^{-6}$ prisms total so we obtain the following.

Counting Bad Pairs for $d = 3$ (sketch)

- Three spheres in $\mathbb{F}_q^3$ intersect at at most two points so any bad pair must have exactly two poles.
- This means that every prism a given pair is bad in must have the same tails.
- There are $\approx q$ choices for the last point of the prism since it must be distance t away from both poles.
- Finally, observe that there are only $1/2|E|^2$ pairs total, and so at most $C|E|^2 q$ prisms admit a P-bad set. But we just showed there are at least $|E|^5 q^{-6}$ prisms total so we obtain the following.

Theorem (SMALL 2022++)

Let $d = 3$ and suppose $|E| \geq Cq^{7/3}$. Then there exists a prism that admits no P-bad pairs, and thus the VC-dimension of $\mathcal{H}_t^3(E)$ equals 3.

Bad Sets in Higher Dimensions (Sketch)

- Our counting argument gets more complicated in higher dimensions since there are no longer only 2 poles for each bad set, and thus there can be more than one choice of tails for prisms a set can be bad in.

Bad Sets in Higher Dimensions (Sketch)

- Our counting argument gets more complicated in higher dimensions since there are no longer only 2 poles for each bad set, and thus there can be more than one choice of tails for prisms a set can be bad in.
- However, if a bad set has additional poles, then we have additional constraints on prisms it can be bad in since each such pole must also be a pole of some other center point in each prism. This counterbalances the additional choices of tails to an extent.

Bad Sets in Higher Dimensions (Sketch)

- Our counting argument gets more complicated in higher dimensions since there are no longer only 2 poles for each bad set, and thus there can be more than one choice of tails for prisms a set can be bad in.
- However, if a bad set has additional poles, then we have additional constraints on prisms it can be bad in since each such pole must also be a pole of some other center point in each prism. This counterbalances the additional choices of tails to an extent.
- However, for the proof to work we need that the maximum number of poles a k-bad set can have is order q^{d-k-1}. This is always the case if we restrict our attention to prisms with affinely independent centers. We were able to show that enough such prisms exist if $|E| > q^{d-\frac{1}{d-1}}$.

Bad Sets in Higher Dimensions (Sketch)

- Our counting argument gets more complicated in higher dimensions since there are no longer only 2 poles for each bad set, and thus there can be more than one choice of tails for prisms a set can be bad in.
- However, if a bad set has additional poles, then we have additional constraints on prisms it can be bad in since each such pole must also be a pole of some other center point in each prism. This counterbalances the additional choices of tails to an extent.
- However, for the proof to work we need that the maximum number of poles a k-bad set can have is order q^{d-k-1}. This is always the case if we restrict our attention to prisms with affinely independent centers. We were able to show that enough such prisms exist if $|E| > q^{d-\frac{1}{d-1}}$.
- This is the strongest bound on $|E|$ needed in our proof so we obtain our main result.

Future Work

One direction would be attempting to extend past work on single-sphere classifiers. This direction is very similar to what we studied, but is much harder since the greatest VC dimension possible is $d+1$ instead of d.

Conjecture (Fitzpatrick, Iosevich, McDonald, and Wyman, 2021)

For large subsets $E \subset \mathbb{F}_q^d$, then VCD $\left(\mathcal{H}_t^d(E)\right) = d+1$, where $\mathcal{H}_t^d(E)$ denote sphere classifiers over the subset E.

Future Work

One direction would be attempting to extend past work on single-sphere classifiers. This direction is very similar to what we studied, but is much harder since the greatest VC dimension possible is $d+1$ instead of d.

Conjecture (Fitzpatrick, Iosevich, McDonald, and Wyman, 2021)

For large subsets $E \subset \mathbb{F}_q^d$, then $VCD\left(\mathcal{H}_t^d(E)\right) = d+1$, where $\mathcal{H}_t^d(E)$ denote sphere classifiers over the subset E.

It might also be interesting to pursue the chain direction further.

Problem (SMALL 2022++)

Can we get an analogous result for longer chains? How would such classifiers even be defined?

Acknowledgments

Thank you!

This research was conducted as part of the **2022 SMALL REU program** at Williams College, and was funded by **NSF Grant DMS1947438**, **Harvey Mudd College**, and **Williams College** funds. We thank Steven J. Miller, Alex Iosevich, and our colleagues from the 2022 SMALL REU program for many helpful conversations.

Bibliography I

- M. Bennett, J. Chapman, D. Covert, D. Hart, A. Iosevich, J. Pakianathan, *Long paths in the distance graph over large subsets of vector spaces over finite fields*, Korean Math. Soc. **53** (2016), 115-126.
- M. Bennett, D. Hart, A. Iosevich, J. Pakianathan, M. Rudnev, *Group actions and geometric combinatorics in* $\mathbb{F}_q^d$, Forum Math. **29** (2017), 91-110.
- D. Fitzpatrick, A. Iosevich, B. McDonald, E. Wyman, *The VC-dimension and point configurations in* $\mathbb{F}_q^2$, 2021, https://arxiv.org/abs/2108.13231.
- D. Hart, A. Iosevich, D. Koh, S. Senger, I. Uriarte-Tuero, *Distance graphs in vector spaces over finite fields*, Recent advances in harmonic analysis and applications (2013), 139-160.

Bibliography II

- A. Iosevich, G. Jardine, B. McDonald, *Cycles of arbitrary length in distance graphs on* $\mathbb{F}_q^d$, Tr. Mat. Inst. Steklova 314 (2021), no. 1, 27-43

- A. Iosevich, B. McDonald, M. Sun, *Dot products in* $\mathbb{F}_q^3$ *and the Vapnik-Chervonenkis dimension*, Discrete Math. **346** (2023), Paper No. 113096.

- A. Iosevich, H. Parshall, *Embedding distance graphs in finite field vector spaces*, J. Korean Math. Soc. 56 (2019), no. 6, 1515-1528

- A. Iosevich, M. Rudnev, *Erdos distance problem in vector spaces over finite fields*, Trans. Amer. Math. Soc **359** (2007), 6127-6142.

- M. Kearns, U. Vazirani, *An introduction to computational learning theory*, MIT press, 1994.

Bibliography III

H. Minkowski, *Grundlagen für eine Theorie quadratischen Formen mit ganzahligen Koeffizienten*, Gesammelte Abhandlungen (1911), 3-145.

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