

VC-DIMENSION OF SALEM SETS OVER FINITE FIELDS

MOUSTAPHA DIALLO AND BRIAN MCDONALD

ABSTRACT. The VC-dimension, introduced by Vapnik and Chervonenkis in 1968 in the context of learning theory, has in recent years provided a rich source of problems in combinatorial geometry. Given $E \subseteq \mathbb{F}_q^d$ or $E \subseteq \mathbb{R}^d$, finding lower bounds on the VC-dimension of hypothesis classes defined by geometric objects such as spheres and hyperplanes is equivalent to constructing appropriate geometric configurations in E . The complexity of these configurations increases exponentially with the VC-dimension.

These questions are related to the Erdős distance problem and the Falconer problem when considering a hypothesis class defined by spheres. In particular, the Erdős distance problem over finite fields is equivalent to showing that the VC-dimension of translates of a sphere of radius t is at least one for all nonzero $t \in \mathbb{F}_q$. In this paper, we show that many of the existing techniques for distance problems over finite fields can be extended to a much broader context, not relying on the specific geometry of circles and spheres. We provide a unified framework which allows us to simultaneously study highly structured sets such as algebraic curves, as well as random sets.

CONTENTS

1. Introduction	1
1.1. Distance problems	1
1.2. VC-dimension and learning theory	3
1.3. Discrete Fourier analysis preliminaries	5
1.4. Results	5
2. Examples of Salem sets over finite fields	7
2.1. Explicit examples	7
2.2. Random sets	14
3. Proof of Theorem 1.20	16
4. Symmetrized Parabola	23
4.1. Euclidean space	24
4.2. Finite fields	29
References	31

1. INTRODUCTION

1.1. Distance problems. In this paper, we study various generalizations of the Erdős distance problem over finite fields. For a set of n points $E \subseteq \mathbb{R}^2$, Erdős investigated the size of the distance set

$$\Delta(E) := \{|x - y| : x, y \in E\}.$$

The size of the distance set is trivially bounded above,

$$|\Delta(E)| \leq \binom{n}{2}.$$

Here and throughout, we denote the cardinality of a finite set S by $|S|$. This trivial upper bound is sharp, since it is not hard to find examples of sets $E \subseteq \mathbb{R}^2$ of arbitrary size such that all distances are distinct. Lower bounds are much more interesting. Erdős [9] showed in 1946 that $|\Delta(E)| \gtrsim \sqrt{n}$, and conjectured that a square lattice is the optimal configuration to minimize the size of the distance set, i.e.

$$|\Delta(E)| \gtrsim \frac{n}{\sqrt{\log n}}.$$

Here, and throughout, we use asymptotic notation $f(n) \lesssim g(n)$ interchangeably with $f(n) = O(g(n))$.

After almost 70 years and several iterations of improvement on the optimal bound, Guth and Katz [13] essentially resolved the conjecture, up to a factor of $\sqrt{\log n}$.

Theorem 1.1 (Guth-Katz). *For $E \subseteq \mathbb{R}^2$, $|E| = n$,*

$$|\Delta(E)| \gtrsim \frac{n}{\log n}.$$

Iosevich and Rudnev [20] studied a related problem over finite fields in 2007. In the finite setting, in order for these asymptotic questions to have meaning, we generally consider $\mathbb{F}_q$, a finite field with $q = p^r$ elements, as $q \rightarrow \infty$. Of course, it is not immediately clear what is meant by distances over finite fields in the first place.

Definition 1.2. We take the distance between $x, y \in \mathbb{F}_q^d$ to be

$$||x - y|| := x_1^2 + \cdots + x_d^2,$$

Remark 1.3. This distance is an element of $\mathbb{F}_q$ and not a real number. This does not yield a norm or a metric space, yet it shares many properties with the usual Euclidean sphere algebraically and geometrically.

We notice that the distance set can never be larger than the field itself, so a reasonable formulation of the Erdős distance problem over finite fields is to ask how large $E \subseteq \mathbb{F}_q^d$ must be (or $E \subseteq \mathbb{F}_q^d$) to ensure that $\Delta(E) = \mathbb{F}_q$ recovers the entire field?

Theorem 1.4 (Iosevich-Rudnev). *If $E \subseteq \mathbb{F}_q^d$, $|E| \geq cq^{\frac{d+1}{2}}$, c sufficiently large, it follows that*

$$\Delta(E) = \mathbb{F}_q.$$

Indeed, they also demonstrated a more precise quantitative formulation of this result, counting the number of pairs $(x, y) \in E$ with $||x - y|| = t$ for each $t \neq 0$. We will make use of parts of this argument later, see for example Theorem 2.1.

Definition 1.5. Let $\mathcal{G}_t(E)$ be the distance- t graph on $E \subseteq \mathbb{F}_q^d$, i.e. there is an edge $x \sim y$ whenever $||x - y|| = t$.

Thus, the above discussion can be formulated in terms of counting edges in $\mathcal{G}_t(E)$, or equivalently counting embeddings of the graph with two vertices and one edge into $\mathcal{G}_t(E)$. It is natural to extend this study to counting embeddings $G \rightarrow \mathcal{G}_t(E)$ for other graphs G . Indeed, this was done for paths of arbitrary length by Bennett, Chapman, Covert, Hart, Iosevich, and Pakianathan [5], and was done for cycles of arbitrary length and for trees by the second listed author in joint work with Iosevich and Jardine [17]. The tools used to handle sparse graphs like paths and cycles are somewhat different from the tools used for dense graphs, e.g. the complete graph K_n . Iosevich and Parshall were able to formulate a very general framework which can apply to arbitrary graphs G , with bounds depending on the number of vertices and the maximum vertex degree. In general, taking an approach which applies to a wide family of graphs generally results in weaker bounds than taking an approach specifically tailored to a narrower family of graphs.

Another generalization of the Erdős distance problem over finite fields is to consider the generalized distance set $\Delta_G(E)$ for a given graph G .

Definition 1.6. Given a graph G with vertex set $V = \{v_1, \dots, v_n\}$ and edge set $\mathcal{E}$, and a set $E \subseteq \mathbb{F}_q^d$, let

$$\Delta_G(E) := \{(t_{ij})_{ij \in \mathcal{E}} : \exists x^1, \dots, x^n \in E, \|x^i - x^j\| = t_{ij} \text{ whenever } v_i \sim v_j\}$$

This has been studied in the context $G = K_n$ by Bennett, Hart, Iosevich, Pakianathan, and Rudnev [6], and in the context when G is a tree by [17]. The approaches are quite different, as trees are maximally flexible and complete graphs are maximally rigid. The second listed author, with Aksoy and Iosevich [1], developed an interpolation scheme utilizing Fourier restriction to handle cases of graphs with both rigid and flexible components. This technique was extended by the SMALL REU 2024 [7] to a wider class of graphs.

1.2. VC-dimension and learning theory. One family of graphs which has been increasingly studied in the context of distance problems in the last few years arises from the notion of VC-dimension, developed by Vapnik and Chervonenkis [29] in 1968 in the context of machine learning theory, particularly the paradigm of PAC learning (probably approximately correct). We will introduce the central ideas here, for a more thorough introduction to the subject see for example [28].

Definition 1.7. Let X be a set and $\mathcal{H}$ collection of functions from X to $\{0, 1\}$. We say that $\mathcal{H}$ shatters a finite set $C \subset X$ if the restriction of $\mathcal{H}$ to C yields every function from C to $\{0, 1\}$.

Definition 1.8. Let X and $\mathcal{H}$ be defined as above. We say that a non-negative integer n is the VC-Dimension of $\mathcal{H}$ if there exists a $C \subset X$ of size n such that $\mathcal{H}$ shatters C , and no subset of size $n + 1$ is shattered by $\mathcal{H}$.

Remark 1.9. Shattering is a monotone condition, i.e. if $C' \subseteq C$ and $\mathcal{H}$ shatters C , then $\mathcal{H}$ shatters C' as well. Moreover if $\mathcal{H} \subseteq \mathcal{H}'$ and $\mathcal{H}$ shatters C , then $\mathcal{H}'$ shatters C as well.

Definition 1.10. Let G_n be the graph which represents shattering n points, i.e. G_n is the graph with vertex set

$$\{x^1, \dots, x^n\} \cup \{y^I : I \subseteq [n]\},$$

with an edge $x^i \sim y^I$ whenever $i \in I$. See for example G_3 in Figure 1.

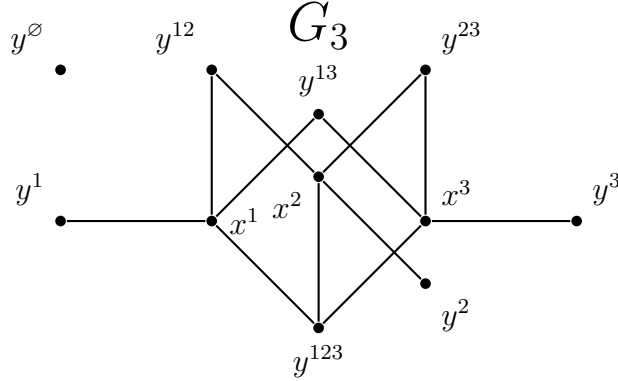


FIGURE 1. The graph G_3 represents shattering 3 points.

Haussler and Welzl [15] demonstrated examples of set systems whose VC-dimension is determined by their geometry in 1987. In 1995 Haussler [14] showed that assuming bounded VC-dimension of a subset of the Boolean cube yields improved bounds for sphere packing with respect to the hamming distance, compared to the result of Dudley [8] in 1978. More recently, in 2019 and 2021 Fox, Pach, and Suk [11, 12] showed that the Schur-Erdős conjecture, as well as a slightly weaker version of the Erdős-Hajnal conjecture, both hold for graphs with bounded VC-dimension. In 2025 Nguyen, Scott, and Seymour [25] strengthened this and showed that the Erdős-Hajnal conjecture holds for graphs of bounded VC-dimension.

In recent years there have also been a handful of results on VC-dimension of hypothesis classes defined over finite fields, in the context of distance sets, hyperplanes, and quadratic residues, and pseudo-random graphs [2, 3, 10, 19, 24, 26, 27]. These results generally take the form of defining a hypothesis class $\mathcal{H}(E)$ depending on a set $E \subseteq \mathbb{F}_q^d$, and determining thresholds on the size of E to guarantee that $\mathcal{H}(E)$ has the same VC-dimension as the ambient space $\mathbb{F}_q^d$. Recently the second listed author, in joint work with Iosevich, Magyar, and A. McDonald [18], extended these techniques to fractals in Euclidean space as well. In this context one seeks a threshold for the Hausdorff dimension of $E \subseteq \mathbb{R}^d$, rather than cardinality.

Our new contribution here is as follows: The condition that $\|x - y\| = t$ can be reformulated as $x - y \in S$, where S is the circle of radius t . The central question we seek to answer here is how the analysis changes by considering different sets S . In this investigation, we need to understand how the techniques of these other papers depend on specific properties of circles and spheres. In particular, we find that most of these results boil down to having good bounds on the Fourier transform of a sphere, understanding intersection properties of spheres, circles, hyperplanes, etc., and having precise estimates for the number of points on a sphere. Theorem 1.20 demonstrates that a many of these results can be derived purely from these properties. We are then able to extend the constructions of [10] not only to a larger class of algebraic curves and surfaces, but also to random sets as in Theorem 1.22.

On the other hand, many of the results in this area rely on symmetry of spheres in an essential way by making use of the action of the orthogonal group. Because of this, we explicitly cannot extend the techniques of [1, 6, 7, 23] to this context.

1.3. Discrete Fourier analysis preliminaries. We will need some basic facts about Fourier analysis over finite fields, and here we record our normalization conventions.

Definition 1.11. Let χ be a nontrivial additive character on $\mathbb{F}_q$, so that every additive character over $\mathbb{F}_q^d$ can be represented in the form

$$\chi_m(x) = \chi(m \cdot x).$$

The choice of nontrivial $\chi \in \widehat{\mathbb{F}_q}$ is arbitrary, but will remain fixed throughout.

Proposition 1.12 (Orthogonality of characters). *Let δ_0 be the Dirac delta. Then,*

$$\sum_{m \in \mathbb{F}_q^d} \chi(m \cdot x) = q^d \delta_0(x) = \begin{cases} q^d & : x = 0 \\ 0 & : x \neq 0 \end{cases}$$

Definition 1.13. For a function $f : \mathbb{F}_q^d \rightarrow \mathbb{C}$, the Fourier transform of f is defined by

$$\widehat{f}(m) = q^{-d} \sum_{x \in \mathbb{F}_q^d} \chi(-m \cdot x) f(x).$$

Proposition 1.14 (Fourier inversion). *We have the inversion formula*

$$f(x) = \sum_{m \in \mathbb{F}_q^d} \chi(m \cdot x) \widehat{f}(m).$$

Proposition 1.15 (Plancharel identity).

$$\sum_{x \in \mathbb{F}_q^d} |f(x)|^2 = q^d \sum_{m \in \mathbb{F}_q^d} |\widehat{f}(m)|^2.$$

1.4. Results.

Definition 1.16. For $E \subseteq \mathbb{F}_q^d$, consider the hypothesis class $\mathcal{H}_S(E)$, with domain $X = E$, defined by

$$\mathcal{H}_S(E) := \{h_y : y \in E\}, \quad h_y(x) := \begin{cases} 1 & x - y \in S \\ 0 & x - y \notin S \end{cases}$$

Definition 1.17. We say that $S \subseteq \mathbb{F}_q^d$ is a γ -Salem set if for all $m \neq 0$,

$$|\widehat{S}(m)| \leq q^{-d} (\log q)^\gamma |S|^{\frac{1}{2}}.$$

In the case $\gamma = 0$, we simply call it a Salem set.

Remark 1.18. We will need this notion of a γ -Salem set for the case when S is a randomly chosen set. For all the other results we could simply work with Salem sets and not keep track of this logarithmic factor. However, it is convenient to take a unified approach which handles both at the same time, and one can always set $\gamma = 0$ and ignore the logarithmic factor in settings when it is not needed.

Definition 1.19. Given $E, S \subseteq \mathbb{F}_q^d$, let $\mathcal{G}_S(E)$ be the directed graph with vertex set E , and with an edge $x \sim y$ whenever $x - y \in S$. In the case when S is symmetric, i.e. $S = -S$, we may consider $\mathcal{G}_S(E)$ to be an undirected graph.

Theorem 1.20. Suppose $S \subset \mathbb{F}_q^d$ is a γ -Salem set, symmetric in the sense $S = -S$, $|S| = Kq^{d-1}(1 + O(q^{-\alpha}))$ for a constant K , $\alpha > \frac{d-1}{2}$, and $|S \cap (S - x)| \leq 2q^\beta$, $\beta < \frac{d-1}{2}$, for all nonzero $x \in \mathbb{F}_q^d$. If $E \subset \mathbb{F}_q^d$ with $|E| \geq cq^{\frac{7d+1}{8}}$, for sufficiently large constant c . Then the VC-dimension of $\mathcal{H}_S^2(E)$ is at least 3.

Corollary 1.21. If S and E satisfy the conditions of Theorem 1.20, and $\beta = 0$, then we can conclude that

$$VCdim(\mathcal{H}_S(E)) = 3.$$

Proof. we need to check that it is impossible to shatter 4 points. This is straightforward, since if we have some shattering configuration

$$\{x^i : 1 \leq i \leq 4\} \cup \{y^I : I \subseteq [4]\} \subseteq E,$$

such that $x^i - y^I \in S$ if and only if $i \in I$, we see that

$$\{x^1, x^2, x^3\} \subseteq (y^{1234} + S) \cap (y^{123} + S),$$

contradicting that this intersection can have at most 2 points. $\square$

In Section 2, we will demonstrate some applications of Theorem 1.20. In particular, we generalize [10] to hypothesis classes over $\mathbb{F}_q^2$ defined by arbitrary quadratic curves, as opposed to circles. The same framework allows us to consider random sets.

Theorem 1.22. Let S be chosen uniformly at random from the subsets of $\mathbb{F}_q^2$ of size q . Let $T = S \cup (-S)$. Then, with probability $1 - o(1)$,

$$VCdim(\mathcal{H}_S(E)) \geq 2 \quad \text{and} \quad VCdim(\mathcal{H}_T(E)) \geq 3.$$

Theorem 1.23. Let $E \subseteq \mathbb{F}_q^2$, $|E| \geq cq^{\frac{15}{8}}$ for c sufficiently large. Let

$$f(x, y) = Ax^2 + Bxy + Cy^2 + Dx + Ey + F,$$

and let Z_f be the zero set of f . Suppose that

$$\det \begin{pmatrix} A & B/2 & D/2 \\ B/2 & C & E/2 \\ D/2 & E/2 & F \end{pmatrix} \neq 0.$$

If

$$\det \begin{pmatrix} A & B/2 \\ B/2 & C \end{pmatrix} \neq 0,$$

Then there exists a translation S of Z_f such that

$$2 \leq VCdim(\mathcal{H}_{Z_f}(E)) \leq VCdim(\mathcal{H}_S(E)) = 3.$$

Otherwise, if the quadratic part is degenerate, the VC-dimension is equal to 2 for any translation. In this case,

$$VCdim(\mathcal{H}_T(E)) \geq 3,$$

where $T = S \cup (-S)$.

Remark 1.24. In the setting of Theorem 1.23, when the quadratic part of f is non-degenerate, we do not know in general whether $VCdim(\mathcal{H}_{Z_f}(E))$ is equal to 2 or 3. We suspect that it might always be equal to 3, but the techniques we develop in Section 3 depend on symmetry in an essential way, and so this is left as an open question.

Finally, we have one more application of Theorem 1.20.

Theorem 1.25. *Let $E \subseteq \mathbb{F}_q^2$, $|E| \geq cq^{\frac{15}{8}}$ for c sufficiently large. Consider a degree n polynomial $f \in \mathbb{F}_q[x]$, $p \nmid n$, and let $\Gamma_f := \{(x, f(x)) : x \in \mathbb{F}_q\}$ be the graph of the function. Then $VCdim(\mathcal{H}_{\Gamma_f}(E)) \geq 2$, and if f is an odd function then $VCdim(\mathcal{H}_{\Gamma_f}(E)) \geq 3$.*

2. EXAMPLES OF SALEM SETS OVER FINITE FIELDS

2.1. Explicit examples. Two explicit examples of Salem sets in $\mathbb{F}_q^d$ are given by [20].

Theorem 2.1 (Iosevich-Rudnev). *The sphere of radius t ,*

$$\mathcal{S}_t = \{x \in \mathbb{F}_q^d : x_1^2 + \cdots + x_d^2 = t\},$$

and the paraboloid

$$\mathcal{P} = \{(x_1, x_2, \dots, x_{d-1}, x_1^2 + \cdots + x_{d-1}^2) : x_1, \dots, x_{d-1} \in \mathbb{F}_q\}$$

are both Salem sets.

We will provide a few more examples and discuss how they relate to Theorem 1.20.

Lemma 2.2. *Suppose that $S \subseteq \mathbb{F}_q^d$ is obtained from $S_0 \subseteq \mathbb{F}_q^d$ via an invertible linear transformation, i.e.*

$$S(x) = S_0(Tx).$$

Then S is a γ -Salem set if and only if S_0 is.

Proof. We will assume that S_0 is a Salem set and show that S must be as well, noting that the other direction follows by symmetry. We can compute directly, where T^{-t} is the inverse transpose of T and $m \neq 0$,

$$\begin{aligned} \widehat{S}(m) &= q^{-d} \sum_{x \in \mathbb{F}_q^d} \chi(-m \cdot x) S_0(Tx) = q^{-d} \sum_x \chi(-m \cdot (T^{-1}x)) S_0(x) \\ &= q^{-d} \sum_x \chi(-x \cdot (T^{-t}m)) S_0(x) = \widehat{S_0}(T^{-t}m) \end{aligned}$$

□

Lemma 2.3. *Suppose that $S \subseteq \mathbb{F}_q^d$ is obtained from $S_0 \subseteq \mathbb{F}_q^d$ via a translation, i.e. $S(x) = S_0(x + v)$ for some vector $v \in \mathbb{F}_q^d$. Then S is a γ -Salem set if and only if S_0 is.*

Proof. This is another straightforward calculation:

$$\begin{aligned}\widehat{S}(m) &= q^{-d} \sum_x \chi(-m \cdot x) S_0(x + v) = q^{-d} \sum_x \chi(-m \cdot (x - v)) S_0(x) \\ &= \chi(m \cdot v) \widehat{S}_0(m),\end{aligned}$$

and so $|\widehat{S}(m)| = |\widehat{S}_0(m)|$. □

Lemma 2.4. *Any degree two polynomial $f \in \mathbb{F}_q[x, y]$, $f \notin \mathbb{F}_q[x] \cup \mathbb{F}_q[y]$, q odd, can be reduced to one of the following forms via an invertible linear change of variables and a translation:*

- (1) $f(x, y) = y - x^2$
- (2) $f(x, y) = ax^2 + by^2 + c$

Moreover, it takes the first form if the quadratic part is a non-degenerate quadratic form, and takes the second form if it is degenerate. If f takes the second form and is a smooth curve, then $c \neq 0$.

Proof. Let

$$f(x, y) = Ax^2 + Bxy + Cy^2 + Dx + Ey + F.$$

Consider the quadratic form $Q(x, y) = Ax^2 + Bxy + Cy^2$, which corresponds to the symmetric matrix

$$M = \begin{pmatrix} A & \frac{B}{2} \\ \frac{B}{2} & C \end{pmatrix}$$

in the sense that for any column vector $v = \begin{pmatrix} x \\ y \end{pmatrix}$,

$$Q(x, y) = Q(v) = v^t M v.$$

This is where we have used the assumption q is odd, so that $\frac{1}{2} \in \mathbb{F}_q$. M is a real symmetric matrix and hence is diagonalizable. Therefore there is an orthogonal transformation U such that $U^t M U$ is a diagonal matrix. This means that we may assume without loss of generality that $B = 0$, since otherwise it can be eliminated via the change of coordinates

$$v' = \begin{pmatrix} x' \\ y' \end{pmatrix} = U^t \begin{pmatrix} x \\ y \end{pmatrix} = U^t v.$$

Let Q' be the quadratic form corresponding to the matrix $U^t M U$. Under this transformation,

$$Q'(v') = (v')^t (U^t M U) v' = (U v')^t M (U v') = v^t M v = Q(v).$$

If M is invertible, let

$$w = -\frac{1}{2} M^{-1} \begin{pmatrix} D \\ E \end{pmatrix},$$

so that

$$f(x, y) = f(v) = v^t M v + Dx + Ey + F$$

$$\begin{aligned}
&= (v - w)^t M(v - w) + w^t Mv + v^t Mw - w^t Mw + \begin{pmatrix} D \\ E \end{pmatrix}^t (v - w) + \begin{pmatrix} D \\ E \end{pmatrix}^t w + F \\
&= Q(v - w) + \left(2Mw + \begin{pmatrix} D \\ E \end{pmatrix} \right)^t \cdot (v - w) + w^t Mw + \begin{pmatrix} D \\ E \end{pmatrix}^t w + F \\
&= Q(v - w) + F',
\end{aligned}$$

where

$$F' = w^t Mw + \begin{pmatrix} D \\ E \end{pmatrix}^t w + F.$$

Since F' is a constant vector, not depending on the variables x and y , we see that translation by w eliminates linear terms assuming M is invertible. This combined with the previous discussion of diagonalizing M proves the Lemma in the case when M is invertible.

If M is not invertible, Then there is some nonzero vector u such that $Mu = 0$. Without loss of generality assume that $u = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$, since otherwise we may apply an orthogonal transformation mapping $\begin{pmatrix} 0 \\ 1 \end{pmatrix}$ to u . This change of basis does not affect whether the linear transformation corresponding to M is invertible. If $M \begin{pmatrix} 0 \\ 1 \end{pmatrix} = 0$, then $B = C = 0$, and so

$$f(x, y) = Ax^2 + Dx + Ey + F.$$

Since we assumed $f \notin \mathbb{F}_q[x]$, we know that $E \neq 0$, so we may express this in the standard quadratic form

$$y = ax^2 + bx + c.$$

Then, via translation, we get $y = x^2$. □

We still need to show that Z_f satisfies the hypotheses of Theorem 1.20. We will need the well-known Weil bound for Kloosterman sums. We will also later use another form of the Weil bounds [30], so we will introduce both here. See, for example, Iwaniec and Kowalski [21] for an overview of the various forms of Weil bounds for character sums over finite fields.

Theorem 2.5 (Weil, 1948). *For $a, b \in \mathbb{F}_q^*$,*

$$\left| \sum_{j \in \mathbb{F}_q} \chi(aj + bj^{-1}) \right| \leq 2\sqrt{q}.$$

Moreover, if $f \in \mathbb{F}_q[x]$ is a polynomial of degree n , f not of the form $c + h^p - h$ for any polynomial $h \in \mathbb{F}_q[x]$, then

$$\left| \sum_{j \in \mathbb{F}_q} \chi(f(j)) \right| \leq (n - 1)\sqrt{q}.$$

Lemma 2.6. *Let f be a degree 2 polynomial $f \in \mathbb{F}_q[x, y]$, $f \notin \mathbb{F}_q[x] \cup \mathbb{F}_q[y]$, q odd. As long as the reduction from Lemma 2.4 does not yield the equation $ax^2 + by^2 = 0$, it follows that Z_f is a Salem set with $|Z_f|$ equal to $q - 1$, q , or $q + 1$.*

Proof. A parabola $y = x^2$ clearly has exactly q solutions. We know from [20] that the parabola is a Salem set. It remains to consider the case

$$Z_f = S_{a,b} := \{(x, y) : ax^2 + by^2 = 1\}.$$

Viewing this as the projective curve $aX^2 + bY^2 = Z^2$, there are exactly $q + 1$ points. This is because for a fixed point P on the curve, every line through P intersects the projective curve in exactly one other point. This yields a bijection between the curve and a projective line, which has $q + 1$ points. To find the number of points on the curve in affine space, observe that in projective space the line at infinity intersects the curve in at most 2 points, and so the number of points on the curve in affine space is one of $q - 1$, q , or $q + 1$.

We still need a pointwise bound on the Fourier transform $|\widehat{S}_{a,b}(m)|$ for $m \neq 0$.

$$\begin{aligned} \widehat{S}_{a,b}(m) &= q^{-2} \sum_{x \in \mathbb{F}_q^2} \chi(-m \cdot x) S_{a,b}(x) = q^{-3} \sum_{x \in \mathbb{F}_q^2} \sum_{j \in \mathbb{F}_q} \chi(-m \cdot x) \chi(j(ax_1^2 + bx_2^2 - 1)) \\ &= q^{-3} \sum_{j \in \mathbb{F}_q^*} \chi(-j) \sum_{x \in \mathbb{F}_q^2} \chi(jax_1^2 + jbx_2^2 - m \cdot x) \end{aligned}$$

This last step is where we have used the fact $m \neq 0$, for this ensures that the $j = 0$ term vanishes. For this inner sum, we have

$$\sum_x \chi(jax_1^2 + jbx_2^2 - m \cdot x) = \left(\sum_{x_1} \chi(jax_1^2 - m_1 x_1) \right) \left(\sum_{x_2} \chi(jbx_2^2 - m_2 x_2) \right),$$

where

$$\sum_{x_1 \in \mathbb{F}_q} \chi(jax_1^2 - m_1 x_1) = \chi\left(-\frac{m_1^2}{4ja}\right) \sum_{x_1} \chi\left(ja\left(x_1 - \frac{m_1}{2ja}\right)^2\right),$$

and

$$\sum_{x_1} \chi\left(ja\left(x_1 - \frac{m_1}{2ja}\right)^2\right) = \sum_{x_1} \chi(jax_1^2)$$

is a Gauss sum, for which it is well known for $k \neq 0$,

$$g(k) := \sum_{x \in \mathbb{F}_q} \chi(kx^2) = \varepsilon_q \eta(k) \sqrt{q},$$

where $\varepsilon_q \in \{\pm 1, \pm i\}$ depends on the field but not on k , and

$$\eta(k) = \begin{cases} 1 & : k \neq 0 \text{ is a square in } \mathbb{F}_q \\ -1 & : k \neq 0 \text{ is not a square in } \mathbb{F}_q \\ 0 & : k = 0 \end{cases}$$

Therefore, putting this all together we obtain

$$\begin{aligned}
\widehat{S}_{a,b}(m) &= q^{-3} \sum_{j \in \mathbb{F}_q^*} \chi(-j) \chi \left(\frac{1}{j} \left(-\frac{m_1^2}{4a} - \frac{m_2^2}{4b} \right) \right) g(ja)g(jb) \\
&= q^{-2} \varepsilon_q^2 \sum_{j \in \mathbb{F}_q^*} \chi(-j) \chi \left(\frac{1}{j} \left(-\frac{m_1^2}{4a} - \frac{m_2^2}{4b} \right) \right) \eta(j^2 ab) \\
&= q^{-2} \varepsilon_q^2 \eta(ab) \sum_{j \in \mathbb{F}_q^*} \chi(-j) \chi \left(\frac{1}{j} \left(-\frac{m_1^2}{4a} - \frac{m_2^2}{4b} \right) \right).
\end{aligned}$$

Finally, by Theorem 2.5 we see that

$$|\widehat{S}_{a,b}(m)| \leq 2q^{-\frac{3}{2}}$$

as long as $\frac{m_1^2}{4a} + \frac{m_2^2}{4b} \neq 0$. If $\frac{m_1^2}{4a} + \frac{m_2^2}{4b} = 0$, $m \neq 0$, then we can compute exactly

$$|\widehat{S}_{a,b}(m)| = q^{-2} < q^{-\frac{3}{2}}.$$

□

We are almost ready to prove Theorem 1.23, we just need one more Lemma. We will defer the proof of Lemma 2.7 until Section 3, after we develop some more Fourier analytic machinery.

Lemma 2.7. *Suppose $S \subseteq \mathbb{F}_q^d$ is a γ -Salem set, $|S| = Kq^{d-1}(1 + O(q^{-\alpha}))$ for a constant K and for $\alpha > \frac{d-1}{2}$, and $|S \cap (S - x)| \leq q^\beta$, $\beta < \frac{d-1}{2}$, for all nonzero $x \in \mathbb{F}_q^d$. If $E \subseteq \mathbb{F}_q^d$, $|E| \geq cq^{\frac{d+1}{2}}$ for c sufficiently large, then*

$$VCdim(\mathcal{H}_S(E)) \geq 2.$$

This corresponds to the configuration in Figure 2.

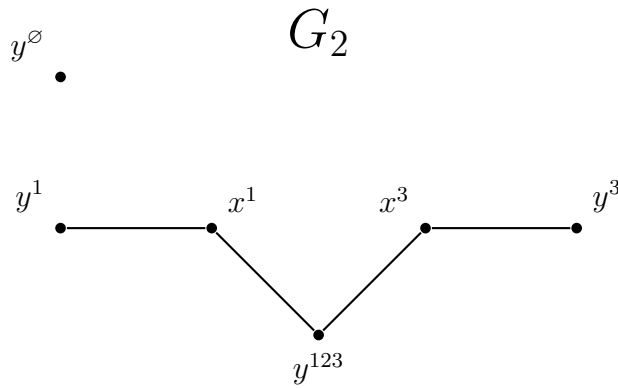


FIGURE 2. The graph G_2 represents shattering 2 points.

We now have all the tools we need to prove Theorem 1.23.

proof of Theorem 1.23. This is a direct application of Theorem 1.20, Lemma 2.4, and Lemma 2.6. Since we require

$$\det \begin{pmatrix} A & B/2 & D/2 \\ B/2 & C & E/2 \\ D/2 & E/2 & F \end{pmatrix} \neq 0,$$

it follows that the curve is smooth. To see this, homogenize $f(x, y)$ to obtain

$$F(X, Y, Z) = Ax^2 + Bx + Cy^2 + Dxz + Eyz + Fz^2,$$

and note that the system of equations

$$0 = \nabla F(X, Y, Z) = \langle 2Ax + By + Dz, Bx + 2Cy + Ez, Dx + Ey + 2Fz \rangle$$

has a nontrivial solution if and only if the above matrix is invertible. If

$$\det \begin{pmatrix} A & B/2 \\ B/2 & C \end{pmatrix} \neq 0,$$

then by Lemma 2.4 we may assume without loss of generality that $f(x, y) = ax^2 + by^2 + c$. Since the curve is smooth, we know $c \neq 0$. We claim that for any vector $(h, k) \in \mathbb{F}_q^2$,

$$f(x, y) = f(x + h, y + k) = 0$$

has at most 2 solutions. To see this, note that

$$f(x + h, y + k) - f(x, y) = 0$$

is a line, the empty set, or the whole space in the case when $(h, k) = (0, 0)$. Therefore,

$$|Z_f \cap (Z_f - (h, k))| \leq 2$$

whenever $(h, k) \neq (0, 0)$, since a line intersects a quadratic curve in at most 2 points. Clearly it is possible to find a translation S of Z_f such that $S = -S$. This is achieved by completing the square to eliminate linear terms, as in Lemma 2.4. Therefore, S satisfies all conditions of Corollary 1.21, and so we can conclude that

$$\text{VCdim}(\mathcal{H}_S(E)) = 3.$$

To see that

$$\text{VCdim} \mathcal{H}_{Z_f}(E) \leq 3,$$

note that the argument in Corollary 1.21 that it is impossible to shatter 4 points still holds here, even though Z_f does not necessarily satisfy all hypotheses of Theorem 1.20.

We still need to consider the case when

$$\det \begin{pmatrix} A & B/2 \\ B/2 & C \end{pmatrix} = 0,$$

and so Z_f is a parabola. In this case, there is no hope to achieve $S = -S$ for any translation of Z_f . Indeed, we can see immediately that it is impossible to shatter 3 points, because $|S \cap (S - x)| \leq 1$ for all $x \neq 0$. Therefore we may make an argument similar to Corollary 1.21. If

$$\{x^1, x^2, x^3\} \cup \{y^I : I \subseteq [3]\}$$

is a shattering configuration, we get a contradiction based on the fact that

$$\{x^1, x^2\} \subseteq (y^{123} + S) \cap (y^{12} \cap S).$$

Finally, we know by Lemma 2.7 that

$$\text{VCdim}(\mathcal{H}_S(E)) = 2.$$

In this case, we obtain the result for the symmetrized parabola $P = S \cup (-S)$ via Corollary 2.14 later in the section. □

Lemma 2.8. *Consider a degree $n \geq 2$ polynomial $f \in \mathbb{F}_q[x]$, and let $\Gamma_f := \{(t, f(t)) : t \in \mathbb{F}_q\}$ be the graph of the function. Γ_f is a Salem set with exactly q elements, and $|\Gamma_f \cap (\Gamma_f - x)| \leq n - 1$ for all $x \in \mathbb{F}_q^2$.*

Proof. We immediately get $|\Gamma_f| = q$ and $|\Gamma_f \cap (\Gamma_f - x)| \leq n - 1$ by the fact that f is a polynomial of degree n . Let

$$f(t) = a_0 + a_1 t + \cdots + a_n t^n.$$

To see that it is a Salem set, we compute for $m \neq 0$,

$$\begin{aligned} \widehat{\Gamma}_f(m) &= q^{-2} \sum_{x \in \mathbb{F}_q^2} \chi(-m \cdot x) \Gamma_f(x) = q^{-3} \sum_{x \in \mathbb{F}_q^2} \sum_{r \in \mathbb{F}_q} \chi(-m \cdot x) \chi(r(f(x_1) - x_2)) \\ &= q^{-3} \sum_r \left(\sum_{x_2} \chi(-(r + m_2)x_2) \right) \left(\sum_{x_1} \chi(rf(x_1) - m_1 x_1) \right) \end{aligned}$$

Here, $g(t) = rf(t) - m_1t$ is a degree n polynomial which is not of the form $c + h^p - h$ for any polynomial h , since the degree is not divisible by p . Therefore, by Theorem 2.5 we have

$$\widehat{\Gamma}_f(m) \lesssim q^{-\frac{5}{2}} \sum_r \sum_{x_2} \chi(-(r + m_2)x_2) = q^{-\frac{3}{2}}$$

□

Corollary 2.9. *The conclusion of Theorem 1.25 follows immediately from Lemma 2.7, Lemma 2.8 and Theorem 1.20, since $\Gamma_f = -\Gamma_f$ if f is an odd function.*

2.2. Random sets. For the remainder of Section 2, we will consider the situation where $S \subseteq \mathbb{F}_q^2$ is randomly chosen, uniformly from the set of subsets of $\mathbb{F}_q^2$ of size q . Throughout the section, for sets S determined by such a random process, we will refer to them simply as random subsets of $\mathbb{F}_q^2$. We will make use of the following result of Hayes [16], which answers a question of Babai [4].

Theorem 2.10 (Hayes). *Let $\varepsilon > 0$. Let G be a finite abelian group of order n . Let $k \leq n$, and let $m = \min(k, n - k)$. For all but an $O(n^{-\varepsilon})$ fraction of subsets $S \subseteq G$ such that $|S| = k$,*

$$\Phi(S) := \max_{\chi} \left| \sum_{x \in G} \chi(x) f(x) \right| < 2\sqrt{2(1 + \varepsilon)m \log n},$$

where this max is take over nontrivial characters of G .

Corollary 2.11. *Translating this to our notation in the case $G = \langle \mathbb{F}_q^2, + \rangle$, $S \subseteq \mathbb{F}_q^d$ chosen randomly with $|S| = q$, we see that with probability $1 - o(1)$, for all $m \neq 0$,*

$$|\widehat{S}(m)| \lesssim q^{-\frac{3}{2}} \sqrt{\log q}.$$

In other words, S is a $\frac{1}{2}$ -Salem set.

Lemma 2.12. *If $S \subseteq \mathbb{F}_q^2$ is a random subset, then S satisfies the hypotheses of Theorem 1.20, with the exception of the symmetry property $S = -S$, with probability $1 - o(1)$ as $q \rightarrow \infty$. In other words, with high probability, we have*

- (1) $|S| = q(1 + O(q^{-\alpha}))$, $\alpha > \frac{1}{2}$.
- (2) S is a γ -Salem set with $\gamma = \frac{1}{2}$.
- (3) $|S \cap (S - x)| \lesssim q^\beta$, $\beta < \frac{d-1}{2}$ for all nonzero $x \in \mathbb{F}_q^d$.

Proof. (1) follows immediately from the definition of S , since $|S| = q$ exactly. (2) follows immediately from Theorem 2.10. It remains to prove (3). Fix $x \in \mathbb{F}_q^2$, and let

$$\Omega^x = |S \cap (S - x)| = \sum_{y \in \mathbb{F}_q^2} S(y)S(x + y) = \sum_{y \in \mathbb{F}_q^2} \Omega_y^x,$$

where $\Omega_y^x = 1$ in the event that $y \in S \cap (S - x)$. We see that

$$\mathbb{E}[\Omega_y^x] = P(\Omega_y^x = 1) = \frac{1}{q(q+1)},$$

and so

$$\mathbb{E}[\Omega^x] = \sum_y \mathbb{E}[\Omega_y^x] = \frac{q}{q+1} \leq 1.$$

Therefore by Markov's inequality,

$$P(\Omega^x > q^\beta) \leq q^{-\beta} = o(1).$$

□

Lemma 2.13. *Suppose $T = S \cup (-S)$, and S is a set, not necessarily random, which satisfies the three properties of Lemma 2.12. Then as long as $|S \cap (-S)| = O(q^{1-\alpha})$, T satisfies properties (1) and (2), with the exception that we gain a factor of 2 in the first property, i.e.*

$$|T| = 2q(1 + O(q^{-\alpha})).$$

Proof. Property (1) is trivial. Property (2) follows by direct calculation. Let

$$R(x) = S(x)S(-x),$$

so that $T(x) = S(x) + S(-x) - R(x)$. Then,

$$|\widehat{T}(m)| \lesssim q^{-\frac{3}{2}} \sqrt{\log \gamma} + |\widehat{R}(m)|,$$

where

$$|\widehat{R}(m)| \leq \frac{|R|}{q^2} = O(q^{-1-\alpha}).$$

□

We are now ready to prove Theorem 1.22.

Proof of Theorem 1.22. This follows immediately from Theorem 1.20, Lemma 2.12, and Lemma 2.13, as long as we can prove that the intersection $T \cap (T - x)$ is small. But this follows immediately by writing

$$T \cap (T - x) = (S \cap (S - x)) \cup (S \cap (-S - x)) \cup (-S \cap (S - x)) \cup (-S \cap (-S - x)).$$

The argument that showed $S \cap (S - x)$ is small with high probability applies to these other sets just the same. $\square$

Corollary 2.14. *The symmetrized parabola $P = \{(x, x^2)\} \cup \{(x, -x^2)\}$ satisfies the hypotheses of Theorem 1.20, and therefore the conclusion of Theorem 1.23 follows.*

Proof. Since the symmetrized parabola $P = P^+ \cup P^-$ satisfies the finite intersection property

$$|P \cap (P - x)| \leq 6,$$

and P^+ is a Salem set with q elements, we see that P satisfies the hypotheses of Theorem 1.20. $\square$

3. PROOF OF THEOREM 1.20

Lemma 3.1. *Suppose that $S \subseteq \mathbb{F}_q^d$,*

$$|\widehat{S}(m)| \lesssim q^{-\frac{d+1}{2}} (\log q)^\gamma,$$

and $|S| = Kq^{d-1}(1 + O(q^{-\alpha}))$ for some constant K , $\alpha > \frac{d-1}{2}$. Then it follows that

$$\nu_S(E) := |\{(x, y) \in E^2 : x - y \in S\}| = K \frac{|E|^2}{q} + \mathcal{E},$$

where

$$|\mathcal{E}| \lesssim q^{\frac{d-1}{2}} (\log q)^\gamma |E|$$

In particular, if $|E| \geq cq^{\frac{d+1}{2}} (\log q)^\gamma$ for a sufficiently large constant c , it follows that

$$\nu_S(E) \geq K \frac{|E|^2}{2q}$$

Proof. We generalize the method of Iosevich and Rudnev [20], who obtained this result in the special case when S is a sphere of radius $t \neq 0$.

$$\sum_{x, y \in \mathbb{F}_q^d} E(x)E(y)S(x - y) = \sum_{x, y, m \in \mathbb{F}_q^d} E(x)E(y)\chi(m \cdot (x - y))\widehat{S}(m) = q^{2d} \sum_m \overline{\widehat{E}(m)} \widehat{E}(m) \widehat{S}(m)$$

$$= \frac{|E|^2|S|}{q^d} + q^{2d} \sum_{m \neq 0} |\widehat{E}(m)|^2 \widehat{S}(m) = K \frac{|E|^2}{q} (1 + O(q^{-\alpha})) + q^{2d} \sum_{m \neq 0} |\widehat{E}(m)|^2 \widehat{S}(m)$$

The $O(q^{-\alpha})$ term is irrelevant, since for any $E \subseteq \mathbb{F}_q^2$, we have $\frac{|E|^2}{q^{1+\alpha}} \leq q^{\frac{d-1}{2}} |E|$. To bound the terms with $m \neq 0$, we use the pointwise bound $|\widehat{S}(m)| \lesssim q^{-\frac{d+1}{2}}$, followed by Cauchy-Schwarz and Plancharel.

$$\begin{aligned} q^{2d} \sum_{m \neq 0} |\widehat{E}(m)|^2 \widehat{S}(m) &\lesssim q^{2d} q^{-\frac{d+1}{2}} (\log q)^\gamma \sum_{m \in \mathbb{F}_q^2} |\widehat{E}(m)|^2 \\ &= q^d q^{-\frac{d+1}{2}} (\log q)^\gamma \left(\sum_x |E(x)|^2 \right) = q^{\frac{d-1}{2}} (\log q)^\gamma |E| \end{aligned}$$

□

We can use this to demonstrate an abundance of tuples (x^1, x^2, x^3, x^4) with vertices contained in $E \subseteq \mathbb{F}_q^d$, satisfying $x^1 - x^2 = x^3 - x^4 \in S$ and $x^1 - x^3 = x^2 - x^4 \in S$. This generalizes the method of [10], which handles the special case where S is a circle and so we would be constructing rhombi.

Lemma 3.2. *Suppose that $E \subseteq \mathbb{F}_q^d$, $|E| \geq cq^{\frac{3d+1}{4}} (\log q)^{\gamma/2}$ for a sufficiently large constant c . Let $S \subseteq \mathbb{F}_q^d$ be a set satisfying the conditions of Lemma 3.1. Fix a non-zero vector $v \in \mathbb{F}_q^d$. Then there exist distinct $x^1, x^2, x^3, x^4 \in E$ such that*

$$x^1 - x^2 = x^3 - x^4 \in S, \quad x^1 - x^3 = x^2 - x^4 \in S.$$

and $x^i - x^j$ is not equal to $\pm v$ for any i, j .

Proof. Observe that less than half of the pairs $(x, y) \in E^2$ satisfying $x - y \in S$ happen to also satisfy $x - y = \pm v$. To see this,

$$|\{(x, y) \in E^2 : x - y = \pm v\}| \leq 2|E| < K \frac{|E|^2}{4q} < \frac{1}{2} |\{(x, y) \in E^2 : x - y \in S\}|.$$

Since $E \geq cq^{\frac{3d+1}{4}} (\log q)^{\gamma/2} \geq cq^{\frac{d+1}{2}} (\log q)^\gamma$, it follows from Lemma 3.1 that

$$\sum_{x, y \in \mathbb{F}_q^d} E(x)E(y)S(x - y) \geq K \frac{|E|^2}{2q},$$

and therefore by the pigeonhole principle there must exist some $u \in S$, $u \neq \pm v$ so that

$$|E \cap (E - u)| = \sum_{\substack{x, y \in \mathbb{F}_q^d \\ x - y = u}} E(x)E(y) \geq \frac{|E|^2}{4q^d}$$

Let $E_u = E \cap (E - u)$, and since

$$|E_u| \geq \frac{|E|^2}{4q^d} \geq \frac{1}{4}c^2q^{\frac{d+1}{2}}(\log q)^\gamma,$$

we can apply Lemma 3.1 once again to find that

$$\sum_{x, y \in \mathbb{F}_q^2} E_u(x)E_u(y)S(x - y) \geq K \frac{|E_u|^2}{2q}.$$

Therefore, since

$$|\{(x, y) \in E_u^2 : x - y \in \{\pm u, \pm v, \pm u \pm v\}\}| \leq 8|E_u| < K \frac{|E_u|^2}{2q} \leq |\{(x, y) \in E_u^2 : x - y \in S\}|,$$

we find that there must exist some pair $(x, y) \in E_u^2$ such that $x - y \in S$ but $x - y \notin \{\pm u, \pm v\}$, since the second inequality above is strict. Let $x^1 = x + u$, $x^2 = y + u$, $x^3 = x$, $x^4 = y$, and this completes the proof. $\square$

In the course of proving Lemma 3.2, we have also obtained the following fact:

Corollary 3.3. *If $|E| \geq cq^{\frac{d+1}{2}}(\log q)^\gamma$, c sufficiently large, it follows that there exists $v \in S$ satisfying*

$$|E_v| \geq K \frac{|E|^2}{4q^d}.$$

This fact, along with Lemma 3.2, allows us to construct a cube graph which contains most of the features necessary for shattering 3 points. If $|E| \geq Cq^{\frac{7d+1}{8}}(\log q)^{\gamma/4}$, then

$$|E_v| \geq \frac{|E|^2}{4q^d} \geq \frac{1}{4}c^2q^{\frac{3d+1}{4}}(\log q)^{\gamma/2},$$

and so we can apply Lemma 3.2 to the set E_v instead of E , using the vector v chosen by Corollary 3.3.

Definition 3.4. Let Q_r be the r -dimensional cube graph, i.e. the graph with vertex set $V = \{0, 1\}^r$, with an edge

$$(a_1, \dots, a_r) \sim (b_1, \dots, b_r)$$

if and only if $a_i \neq b_i$ for exactly one index i .

Corollary 3.5. *Suppose S is symmetric, i.e. $S = -S$, so that $\mathcal{G}_S(E)$ is an undirected graph. If $E \subseteq \mathbb{F}_q^d$, $|E| \geq cq^{\frac{7d+1}{8}}(\log q)^{\gamma/4}$, c sufficiently large, then there is a subgraph of $\mathcal{G}_S(E)$ isomorphic to the graph obtained from Q_3 by deleting one vertex.*

Proof. As mentioned, we combine Lemma 3.2 and Corollary 3.3. With v, x^1, x^2, x^3, x^4 defined as above, the proof is the picture in Figure 3. The fact that Lemma 3.2 ensures $x^i - x^j$ is not equal to $\pm v$ for any (i, j) , and that $u \neq \pm v$, ensures that all the vertices are actually distinct. □

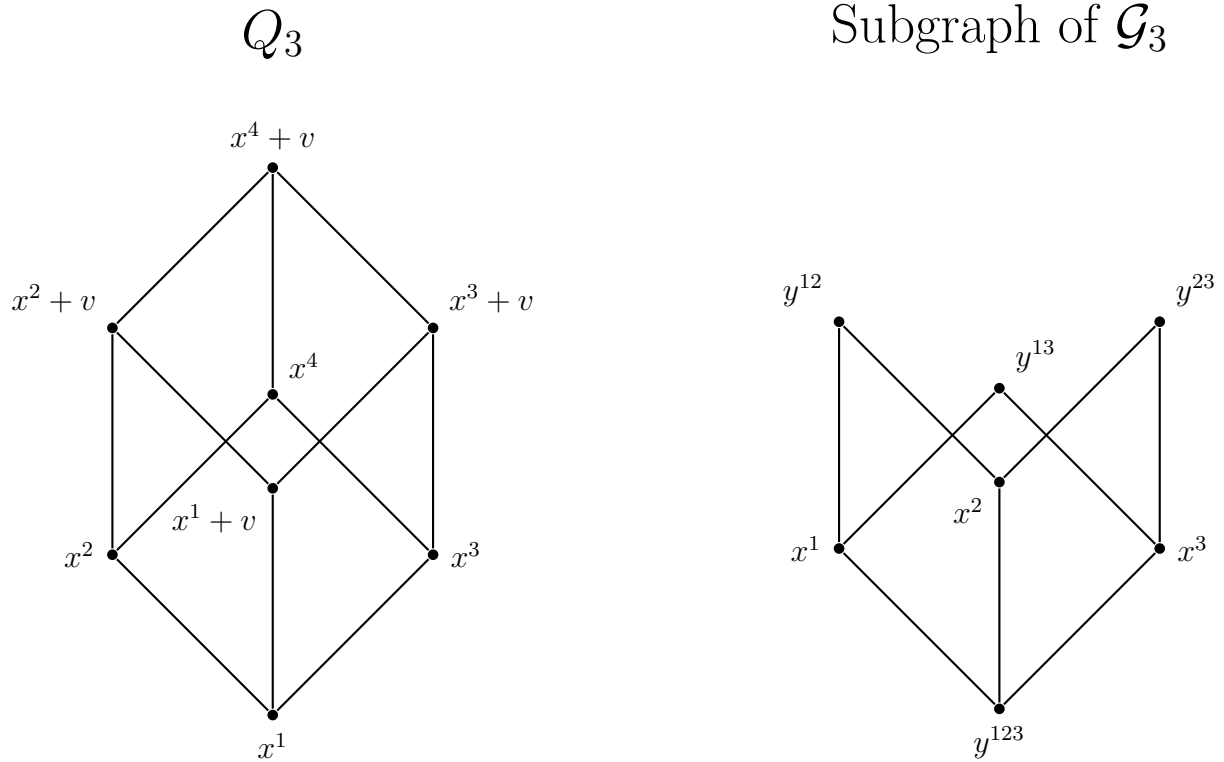


FIGURE 3. Left: The cube graph Q_3 , as constructed by Lemma 3.2 and Corollary 3.5. Right: Re-labeling the vertices to agree with Figure 1, and deleting the vertex $x^4 + v$ which is unnecessary for the proof of Theorem 1.20.

We are almost ready to prove Theorem 1.20, but first we need Lemma 3.6, which will allow us to estimate the number of triples $(x, y, z) \in E^3$ with $x - y, z - y \in S$, as well as Lemma 3.7, which will allow us to assume without loss of generality that each of the vertices constructed above has a positive proportion of the average number of neighbors for vertices in $\mathcal{G}_S(E)$.

Lemma 3.6. *Let $f, g : \mathbb{F}_q^d \rightarrow \mathbb{R}^{\geq 0}$, and let $S \subseteq \mathbb{F}_q^d$ satisfy $|\widehat{S}(m)| \lesssim q^{-\frac{d+1}{2}}(\log q)^\gamma$ for $m \neq 0$, and $|S| = Kq^{d-1}(1 + O(q^{-\alpha}))$. It follows that*

$$\sum_{x,y \in \mathbb{F}_q^d} f(x)g(y)S(x-y) = \frac{K}{q} \|f\|_{L^1} \|g\|_{L^1} (1 + O(q^{-\alpha})) + O\left(q^{\frac{d-1}{2}} (\log q)^\gamma \|f\|_{L^2} \|g\|_{L^2}\right).$$

Proof.

$$\begin{aligned} \sum_{x,y \in \mathbb{F}_q^d} f(x)g(y)S(x-y) &= \sum_{x,y,m \in \mathbb{F}_q^d} f(x)g(y)\chi(m \cdot (x-y))\widehat{S}(m) = q^{2d} \sum_m \overline{\widehat{f}(m)} \widehat{g}(m) \widehat{S}(m) \\ &= \frac{K}{q} \|f\|_{L^1} \|g\|_{L^1} (1 + O(q^{-\alpha})) + \mathcal{R}, \end{aligned}$$

where

$$\begin{aligned} \mathcal{R} &= q^{2d} \sum_{m \neq 0} \overline{\widehat{f}(m)} \widehat{g}(m) \widehat{S}(m) \lesssim q^{\frac{3d-1}{2}} (\log q)^\gamma \sum_{m \in \mathbb{F}_q^2} |\widehat{f}(m)| |\widehat{g}(m)| \leq q^{\frac{3d-1}{2}} (\log q)^\gamma \|\widehat{f}\|_{L^2} \|\widehat{g}\|_{L^2} \\ &= q^{\frac{d-1}{2}} (\log q)^\gamma \|f\|_{L^2} \|g\|_{L^2} \end{aligned}$$

□

Lemma 3.7. *Let $E \subseteq \mathbb{F}_q^d$, $|E| \geq cq^{\frac{d+1}{2}} (\log q)^\gamma$ for c sufficiently large. Suppose that S satisfies the conditions of Lemma 3.1, i.e. $|\widehat{S}(m)| \lesssim q^{-\frac{d+1}{2}} (\log q)^\gamma$ for $m \neq 0$ and $|S| = Kq^{d-1}(1 + O(q^{-\alpha}))$. Then whenever $0 < M \lesssim q^\beta$, $\beta < \frac{d-1}{2}$, there exists a subset $E_M \subseteq E$, $|E_M| \gtrsim |E|$, so that*

$$E * S(x) > M$$

for all $x \in E_M$.

Proof. Summing over $x \in E$, we have

$$\sum_{x \in E} E * S(x) = \sum_{x,y \in \mathbb{F}_q^d} E(x)E(y)S(x-y) \geq K \frac{|E|^2}{2q}.$$

Let

$$E_M = \{x \in E : E * S(x) > M\}.$$

Then, since $M \leq \frac{K}{4}|E|q^{-1}$ for q sufficiently large, it follows that

$$\sum_{x \in E} E * S(x) \leq M|E \setminus E_M| + \sum_{x \in E_M} E * S(x) \leq K \frac{|E|^2}{4q} + \sum_{x \in E_M} E * S(x).$$

Therefore,

$$\sum_{x \in E_M} E * S(x) \geq K \frac{|E|^2}{4q}.$$

By Cauchy-Schwarz,

$$K^2 \frac{|E|^4}{16q^2} \leq \left(\sum_{x \in E_M} E * S(x) \right)^2 \leq |E_M| \sum_{x \in E} (E * S(x))^2.$$

We can bound this last sum using Lemma 3.6 with $f(x) = E(x)E * S(x)$, $g(y) = E(y)$. We have

$$\begin{aligned} \sum_{x \in E} (E * S(x))^2 &= \sum_{x, y \in \mathbb{F}_q^d} E(x)E * S(x)E(y)S(x-y) = \sum_{x, y} f(x)g(y)S(x-y) \\ &= \frac{K}{q} \|f\|_{L^1} \|g\|_{L^1} (1 + O(q^{-\alpha})) + O\left(q^{\frac{d-1}{2}} (\log q)^\gamma \|f\|_{L^2} \|g\|_{L^2}\right) \end{aligned}$$

Clearly, $\|g\|_{L^1} = |E|$ and $\|g\|_{L^2} = |E|^{\frac{1}{2}}$. Moreover,

$$\|f\|_{L^1} = \sum_x E(x)E * S(x) = \sum_{x, y} E(x)E(y)S(x-y) \lesssim K \frac{|E|^2}{q},$$

and

$$\|f\|_{L^2}^2 = \sum_{x \in E} (E * S(x))^2$$

is the original sum we were trying to bound. We see that

$$\frac{K}{q} \|f\|_{L^1} \|g\|_{L^1} (1 + O(q^{-\alpha})) \lesssim K \frac{|E|^3}{q^2}.$$

If this is the larger of the two terms, we are done. Otherwise, we have

$$\sum_{x \in E} (E * S(x))^2 = \|f\|_{L^2}^2 \lesssim q^{\frac{d-1}{2}} (\log q)^\gamma \|f\|_{L^2} \|g\|_{L^2},$$

and so

$$\|f\|_{L^2} \lesssim q^{\frac{d-1}{2}} (\log q)^\gamma |E|^{\frac{1}{2}}.$$

Finally, we obtain

$$\sum_{x \in E} (E * S(x))^2 \lesssim q^{d-1} (\log q)^{2\gamma} |E| \lesssim \frac{|E|^3}{q^2},$$

since $|E| \geq cq^{\frac{d+1}{2}} (\log q)^\gamma$. Putting this all together, we get the desired conclusion that $|E_M| \gtrsim |E|$. □

We are now ready to prove Theorem 1.20.

Proof of Theorem 1.20. To begin, using the labeling conventions from Figure 3, we simply ignore the vertex $x^4 + v$, and relabel the rest of the vertices to agree with the labelling of Figure 1. It only remains to find points $y^1, y^2, y^3, y^\emptyset$ with the desired connections. $y^\emptyset$ is trivial, since it is not connected to any other vertices of $\mathcal{G}_3$, so we need only find a point in E which is not connected to any of the previously constructed points. But there are finitely many such points (7 to be precise) and each has $O(q)$ neighbors in $\mathcal{G}_S(E)$. Since $E \geq cq^{\frac{7d+1}{8}}$, clearly we can pick a vertex which is not connected to any of these 7 points in $\mathcal{G}_S(E)$.

Next we need to find $y^1, y^2, y^3 \in E$ so that $x^i - y^j \in S$ if and only if $i = j$. Let E_M be as in Lemma 3.7, and note that we may assume without loss of generality that $x^1, x^2, x^3 \in E_M$. This is because $|E_M| \gtrsim |E|$, so E_M satisfies the same assumptions as E , and thus all of these constructions apply to E_M as well as E .

Thus, we have guaranteed that each of x^1, x^2, x^3 have more than M neighbors in $\mathcal{G}_S(E)$. Therefore we can choose an appropriate y^1 as long as M is large enough to ensure that x^1 has a neighbor which avoids the 7 points already in the diagram, and also avoids any points which are neighbors of x^2 or x^3 . But, since $|S \cap (S - x)| \leq q^\beta$ for all $x \in \mathbb{F}_q^2$, there are only $7 + 2q^\beta$ points which must be avoided. Put $M = 7 + 2q^\beta \lesssim q^\beta$, and thus we can find y^1, y^2, y^3 with the desired connections in $\mathcal{G}_S(E)$. This completes the proof. □

In addition, we now have the Fourier analytic machinery needed to prove Lemma 2.7.

Proof of Lemma 2.7. We need to find

$$\{x^1, x^2\} \cup \{y^{12}, y^1, y^2, y^\emptyset\} \subseteq E$$

such that $x^i \in y^I$ if and only if $i \in I$. It is trivial to find $y^\emptyset$ once the other points are constructed, since we have to avoid $O(q)$ points which is significantly less than the size of E . Therefore, we really just need to achieve the rest of the configuration as displayed in Figure 2.

In the proof of Lemma 3.7, we use Lemma 3.6 to find that

$$\sum_{x \in E} (E * S(x))^2 \lesssim \frac{|E|^3}{q^2}$$

as long as $|E| \geq cq^{\frac{d+1}{2}}(\log q)^\gamma$ for c sufficiently large. In fact, that same calculation yields the more precise estimate

$$\sum_{x \in E} (E * S(x))^2 = K^2 \frac{|E|^3}{q^2} (1 + O(q^{-\alpha})) + O(q^{d-1}(\log q)^{2\gamma}|E|),$$

where $O(q^{d-1}(\log q)^{2\gamma}|E|)$ is a small error term. But the sum $\sum_{x \in E} (E * S(x))^2$ is counting the number of tuples $(x^1, x^2, y^{12}) \in E^3$ with $x^1 - y^{12}, x^2 - y^{12} \in S$. Therefore, we may apply Lemma 3.7 once again to ensure that $(x^1, x^2, y^{12}) \in E_M \subseteq E$, $M > q^\beta$. This means that the number of points $y^1 \in E$ with $x^1 - y^1 \in S$ is greater than the number of points in the intersection

$$|(x^1 - S) \cap (x^2 - S)| \leq q^\beta,$$

and therefore we may choose a point $y^1 \in E$ satisfying the shattering configuration property that $x^1 - y^1 \in S$ and $x^1 - y^2 \notin S$. We may similarly choose an appropriate y^2 , and this completes the proof of the Lemma. $\square$

4. SYMMETRIZED PARABOLA

When first considering the question of how the results of [10] generalize to $\mathcal{H}_S(E)$ where S is a general quadratic curve as opposed to a circle, it did not take long to realize that the argument does not apply to a parabola because of a lack of symmetry. In attempts to fix this by considering a symmetrized parabola, we then lose the necessary intersection properties, since a symmetrized parabola may intersect a translation of itself up to 6 times over a finite field (only 4 times in Euclidean space), which is more than the 2 points of intersection obtained by translating a circle. In this section we demonstrate how the situation over finite fields is fundamentally different from Euclidean space, in contrast with the situation for circles and spheres, in which the VC-dimension is the same in $\mathbb{F}_q^d$ and in $\mathbb{R}^d$. We will show that one can only shatter 3 points with the symmetrized parabola over Euclidean space, while it is possible to shatter 4 points over (some) finite fields.

This leads to a very natural follow-up question: In [18], the VC-dimension of a hypothesis class of spheres in fractal sets in $\mathbb{R}^d$ is studied. In particular, they show that whenever the Hausdorff dimension of $E \subseteq \mathbb{R}^d$, $d \geq 3$, is above a certain threshold, it follows that $\text{VCdim}(\mathcal{H}_{S_t}(E)) \geq 3$, where $\mathcal{H}_{S_t}(E)$ is the hypothesis class defined by translations of the sphere S_t of radius t intersected with E . The methods do not provide a result in $d = 2$, and indeed there is a fundamental obstruction due to Maga's construction [22] of full-dimensional sets avoiding parallelograms. There may be no hope to extend the results of [18] to $d = 2$, but what about for the symmetrized parabola, or other curves besides the circle? It is plausible that one could obtain a positive result in this case, but simply applying the techniques from [18] is not sufficient. For now we leave this as an open problem.

Here we will discuss the VC-dimension of a hypothesis class defined by a symmetrized parabola, and compare the finite field situation with Euclidean space.

Definition 4.1. Let $P = P^+ \cup P^-$ be the symmetrized parabola, where $P^+ = \{(x, x^2)\}$ and $P^- = -P^+$. Let L, R be the left and right halves of the symmetrized parabola respectively, and let

$$L^+ = P^+ \cap L, \quad L^- = P^- \cap L, \quad R^+ = P^+ \cap R, \quad R^- = P^- \cap R$$

Remark 4.2. The definition for P^+ and P^- can be interpreted over any ring, but we will consider the cases $\mathbb{R}^2$ and $\mathbb{F}_q^2$ here. The definition for $L, R, L^\pm, R^\pm$ can be interpreted over any ordered ring. In particular this excludes $\mathbb{F}_q$.

4.1. Euclidean space.

Definition 4.3. Consider the symmetrized parabola in Euclidean space. Let $\mathcal{H}_P(\mathbb{R}^2)$ denote the hypothesis class consisting of indicator functions h_y of sets of the form $y + P$ for $y \in \mathbb{R}^2$.

Proposition 4.4. $\mathcal{H}_P(\mathbb{R}^2)$ has VC-dimension equal to 3.

First we will show that no set of 4 points, $\{x^1, x^2, x^3, x^4\}$, is shattered by $\mathcal{H}_P(\mathbb{R}^2)$. For sake of contradiction, assume that we have a shattering configuration

$$\{x^1, x^2, x^3, x^4\} \cup \{y^I : I \subseteq [4]\}$$

such that $x^i \in y^I$ if and only if $i \in I$. Without loss of generality, assume $y^{1234} = 0$. We observe that there are several cases based on the position of the points $x^1, \dots, x^4$ with respect to the 4 quadrants of the plane.

Lemma 4.5. Consider the points $X = (x, x^2) \in L^+$ and $Y = (y, -y^2) \in L^-$, so that $x, y < 0$. If there exists $P_0 \in \mathbb{R}^2$ such that $X \in P_0 + L^+$ and $Y \in P_0 + L^-$, then $P_0 = (0, 0)$

Proof. Let $(a, b) = (y - x, y^2 - x^2)$, and the Lemma is equivalent to showing that

$$\begin{pmatrix} x + a \\ x^2 + b \end{pmatrix} = \begin{pmatrix} y \\ -y^2 \end{pmatrix}$$

has at most one solution (x, y) with $x, y < 0$. But this is the intersection of the circle $x^2 + y^2 = -b$ and the line $y = x + a$, which is at most 2 points. One of these points has $x > 0$, so there is only one solution. $\square$

Lemma 4.6. Let $D = L^- \cup R^+$. Then, for any two points P and Q , there are at most 2 solutions $v \in \mathbb{R}^2$ satisfying $P, Q \in (D + v)$. The same holds for $-D$ by symmetry.

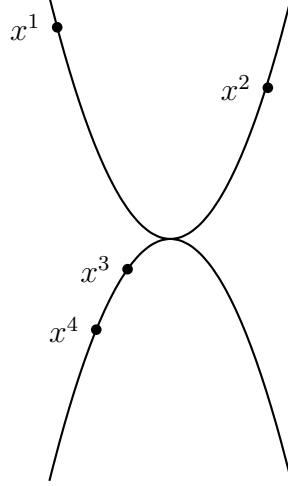


FIGURE 4. Four points placed on the Symmetrized Parabola, used to illustrate the different geometric configurations examined in the casework that follows

Proof. Let

$$f(x) = \begin{cases} x^2 & : x \geq 0 \\ -x^2 & : x < 0 \end{cases}$$

We are looking for solutions to $g(x) := f(x+h) - f(x) = k$, where $Q - P = (h, k)$. There is a one-to-one correspondence between solutions to this equation and translations satisfying $P, Q \in (D + v)$, by taking $v = P - (x_0, f(x_0))$. We compute the derivative

$$g'(x) = \begin{cases} 2h & : x \geq 0 \\ -2h & : x \leq -h \\ 4x + 2h & : -h < x < 0 \end{cases}$$

Therefore, g is decreasing to the left of $x = -\frac{h}{2}$ and increasing to the right of $x = -\frac{h}{2}$, so we conclude that there are at most 2 solutions to the equation $g(x) = k$. $\square$

Lemma 4.7. *Assume that $\{x^1, \dots, x^4\}$ are shattered, with the notation from above for x^i and y^I , where $x^i = (x_1^i, x_2^i) \in \mathbb{R}^2$. Then $x_2^1, x_2^2, x_2^3, x_2^4$ are distinct.*

Proof. Suppose, for contradiction, that there exists distinct points $x^i \neq x^j$ such that $x_2^i = x_2^j$. Without loss of generality, assume $x^i, x^j \in y^{1234} + P^+$, since they must either be both in the top branch or both in the bottom branch. Since $(y^{1234} + P^+) \cap (y^{123} + P^+)$ has at most one point, we see that x^i and x^j cannot both lie in $y^{123} + P^+$. Moreover, since they share the same height, it is impossible for one to lie on $y^{123} + P^+$ and the other on $y^{123} + P^-$. It follows that $x^i, x^j \in y^{123} + P^-$. By a similar argument, we must also have $x^i, x^j \in y^{124} + P^-$. This implies $y^{123} = y^{124}$, a contradiction. $\square$

Lemma 4.8. *Assume that $\{x^1, \dots, x^4\}$ are shattered, with the notation from above for x^i and y^I , where $x^i = (x_1^i, x_2^i) \in \mathbb{R}^2$. Further assume that the points are indexed so that their heights are in strictly decreasing order, i.e. $x_2^1 > x_2^2 > x_2^3 > x_2^4$. Then x^1, x^2, x^3, x^4 cannot all lie in P^+ .*

Proof. Suppose $x^1, \dots, x^4 \in P^+$. It follows that $x^3, x^4 \in y^{134} + P^-$, since $x^1 \in P^+ \cap (y^{134} + P^+)$, and the intersection contains at most one point. For similar reasons $x^3, x^4 \in y^{234} + P^-$. This implies $y^{134} = y^{234}$, a contradiction. $\square$

Lemma 4.9. *Assume that $\{x^1, \dots, x^4\}$ are shattered, with the notation from above for x^i and y^I , where $x^i = (x_1^i, x_2^i) \in \mathbb{R}^2$. Further assume that the points are indexed so that their heights are in strictly decreasing order, i.e. $x_2^1 > x_2^2 > x_2^3 > x_2^4$. Then x^3 cannot lie in P^+ .*

Proof. Suppose $x^3 \in P^+$. It follows that $x^1, x^2 \in P^+$. Moreover, following lemma 4.8 $x^4 \in P^-$. By the same reasoning as before, x^3 must simultaneously satisfy

$$(4.1) \quad x^3 \in y^{134} + P^- \quad \text{and} \quad x^3 \in y^{234} + P^-$$

However, clearly $x^4 \in (y^{134} + P^-) \cap (y^{234} + P^-)$ as well, forcing $y^{134} = y^{234}$, a contradiction. $\square$

Lemma 4.10. *Suppose that*

$$\{x^1, x^2, x^3, x^4\} \cup \{y^I : I \subseteq [4]\}$$

is a shattering configuration with $x^i = (x_1^i, x_2^i)$, and $x_2^1 > x_2^2 > x_2^3 > x_2^4$. Then, we must have $x^1, x^2 \in L^+$, $x^3 \in L^-$, $x^4 \in R^-$.

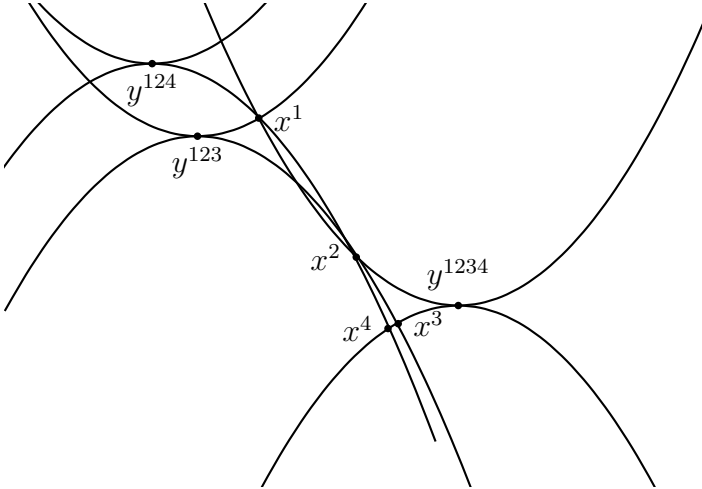


FIGURE 5. Geometric configuration used in proof of Lemma 4.10, Case 1, where the translates $y^{123} + P$ and $y^{124} + P$ force x^3, x^4 to lie to the right of x^1, x^2 .

Proof. Suppose $x^3 \in P^-$. By Lemma 4.9, we know $x^1, x^2 \in P^+$. We divide into cases based on the number of points in L^+ and L^- .

Case 1: $x^1, x^2 \in L^+$ and $x^3, x^4 \in L^-$. Since $(y^{1234} + P^+) \cap (y^{123} + P^+)$ contains at most one point, x^1 and x^2 cannot both lie on $y^{123} + P^+$. Thus, x^2 must lie on $y^{123} + P^-$. If $x^2 \in y^{123} + L^-$, then so is x^3 . This also forces $x^1 \in y^{123} + L^+$. By Lemma 4.5, since $x^1 \in L^+$ and $x^3 \in L^-$, we cannot also have $x^1 \in y^{123} + L^+$ and $x^3 \in y^{123} + L^-$. Therefore, we conclude that $x^2 \in y^{123} + R^-$, and therefore $x^3 \in y^{123} + R^-$ as well.

A symmetric argument applied to y^{124} gives $x^2, x^4 \in y^{124} + R^-$. Hence, x^3 and x^4 both lie to the right of x^1, x^2 . However, by reflecting the entire picture across the horizontal axis we see that the same argument could have been applied to show that x^3 and x^4 both lie to the left of x^1 and x^2 . This is a contradiction. Therefore, it is impossible to have $x^1, x^2 \in L^+$ and $x^3, x^4 \in L^-$.

Case 2: $x^1, x^2 \in L^+$ and $x^3, x^4 \in R^-$. Since $(y^{1234} + P^+) \cap (y^{123} + P^+)$ contains at most one point, x^1 and x^2 cannot both lie on $y^{123} + P^+$. Hence, at least one of them must lie on $y^{123} + P^-$. Furthermore, in order to have $x^3 \in y^{123} + P$, we must have $x^1, x^2 \in y^{123} + L$, since otherwise we would have y^{123} lying entirely to the left of P , in which case $y^{123} + P$ would not intersect R^- . It is impossible for $x^1 \in y^{123} + L^-$ for then, x^2 could not be hit. Thus, we must have $x^1 \in y^{123} + L^+$ and $x^2 \in y^{123} + L^-$. A symmetric argument with y^{124} gives $x^1 \in y^{124} + L^+$ and $x^2 \in y^{124} + L^-$. However, by Lemma 4.5 this implies $y^{123} = y^{124}$, a contradiction.

Case 3: $x^1, x^2 \in L^+$, $x^3 \in R^-$, $x^4 \in L^-$. Now, since $(y^{1234} + P^-) \cap (y^{234} + P^-)$ contains at most one point, x^3 and x^4 cannot both lie on $y^{234} + P^-$, and so $x^3 \in y^{234} + P^+$, and so is x^2 . We see that $x^2 \in y^{234} + L^+$, since otherwise y^{234} would lie entirely to the left of P , contradicting that $x^3 \in y^{234} + P$. We see that $x^3 \in y^{234} + R^+$ for similar reasons. When we consider the position of x^4 , we know by Lemma 4.5 it cannot be in $y^{234} + L^-$ since we already have $x^2 \in L^+$ and $x^4 \in L^-$. We cannot have $x^4 \in y^{234} + R$, since then y^{234} would lie entirely to the left of P . Thus, we conclude that $x^4 \in y^{234} + L^+$. The same argument applies to y^{134} , and so $\{x^1, x^3, x^4\} \subseteq y^{134} + P^+$. This yields a contradiction, since $(y^{134} + P^+) \cap (y^{234} + P^+)$ has at most one point and cannot contain both x^3 and x^4 .

Case 4: $x^1 \in L^+$, $x^2 \in R^+$, $x^3 \in L^-$, $x^4 \in R^-$. Since $(y^{1234} + P^+) \cap (y^{123} + P^+)$ contains at most one point, x^1 and x^2 cannot both lie on $y^{123} + P^+$. Therefore, $x^2, x^3 \in y^{123} + P^-$. Moreover, x^1 is also in $y^{123} + P^-$ because it cannot be in $y^{123} + L^+$ because of Lemma 4.5 applied to $\{x^1, x^3\} \subseteq L$, and it cannot be in $y^{123} + R$ because then y^{123} would lie entirely to the left of P , contradicting that $x^2 \in y^{123} + P$. Now considering y^{124} , note that x^1 and x^2 cannot both lie on $y^{124} + P^+$, since $x^1, x^2 \in P^+$, and also cannot both lie on $y^{124} + P^-$, since $x^1, x^2 \in y^{123} + P^-$. Therefore, $x^1 \in y^{124} + P^+$ and $x^2 \in y^{124} + P^-$. We also see that $x^2, x^4 \in y^{124} + R^-$, since if either were in a left branch then y^{124} would lie entirely to the right of P , contradicting that $x^1 \in y^{124} + P$. For similar reasons, $x^1 \in y^{124} + L^+$.

Now we consider y^{234} . Since $\{x^3, x^4\} \subseteq P^-$, they cannot both be contained in $y^{234} + P^-$, so we see that $x^3 \in y^{234} + P^+$ and hence so is x^2 . Once again we conclude that in fact $x^3 \in y^{234} + L^+$, $x^2 \in y^{234} + R^+$, and $x^4 \in y^{234} + R$, because y^{234} cannot lie either entirely to the left or entirely to the right of P . Using Lemma 4.5, we see that $x^4 \in y^{234} + R^+$, since if it were in $y^{234} + R^-$ this would contradict the fact that $x^2 \in R^+$ and $x^4 \in R^-$.

Finally, consider y^{134} . Since $\{x^3, x^4\} \subseteq P^-$, they cannot both be contained in $y^{134} + P^-$, so we see that $x^1, x^3 \in y^{134} + P^+$. Thus, once again they must be in $y^{134} + L^+$ to avoid y^{134} lying entirely to the left of P . Similarly, x^3 and x^4 cannot both lie in $y^{134} + P^+$, since they already both lie in $y^{234} + P^+$. Hence, $x^4 \in y^{134} + P^-$, and in fact $x^4 \in y^{134} + R^-$ to avoid y^{134} lying entirely to the right of P . By Lemma 4.6, we have a contradiction.

Case 5: $x^1 \in L^+$, $x^2 \in R^+$, $x^3 \in R^-$, and $x^4 \in L^-$. Note that $x^2, x^4 \in y^{124} + P^-$, since $\{x^1, x^2\} \subseteq P^+$ so they cannot share another upward parabola, and x^4 is below x^2 . Moreover, $x^2 \in y^{124} + R^-$ to avoid y^{124} being to the right of P . Similarly, $x^4 \in y^{124} + L^-$

and $x^1 \in y^{124} + L$. We claim $x^1 \in y^{124} + L^-$. If not, then by Lemma 4.5 we would have $y^{124} = y^{1234}$.

Consider y^{134} , observing that $x^1, x^3 \in y^{134} + P^+$. This is because $x^3, x^4 \in y^{1234} + P^-$. In particular, $x^3 \in y^{134} + R^+$ because otherwise y^{134} would lie entirely to the right of P . Similarly, we conclude that $x^1 \in y^{134} + L^+$ and $x^4 \in y^{134} + L$. We claim $x^4 \in y^{134} + L^+$. If not, then by Lemma 4.5 we would have $y^{134} = y^{1234}$.

For y^{234} , note that $x^2, x^3 \in y^{234} + P^+$, since $x^3, x^4 \in P^-$. Moreover, $x^3 \in y^{234} + R^+$, to avoid y^{234} lying entirely to the right of P . For similar reasons we conclude $x^2 \in y^{234} + R^+$ and $x^4 \in y^{234} + L$. We claim that $x^4 \in y^{234} + L^-$. This is because $x^3, x^4 \in y^{134} + P^+$.

Finally, we consider y^{123} , observing that $x^2, x^3 \in y^{123} + P^-$, since $\{x^1, x^2\} \subseteq P^+$. Note that $x^2 \in y^{123} + R^-$, since otherwise y^{123} would lie entirely to the right of P . Similarly, $x^3 \in y^{123} + R^-$. This yields a contradiction, as $x^2, x^3 \in y^{123} + R^-$ implies that x^3 is to the right of x^2 , while $x^2, x^3 \in y^{234} + R^+$ implies that x^2 is to the right of x^3 . $\square$

Lemma 4.11. *Suppose that*

$$\{x^1, x^2, x^3, x^4\} \cup \{y^I : I \subseteq [4]\}$$

is a shattering configuration, and assume $x_2^1 > x_2^2 > x_2^3 > x_2^4$, $y^{1234} = (0, 0)$. It is not possible to have

$$x^1, x^2 \in L^+, \quad x^3 \in L^-, \quad x^4 \in R^-.$$

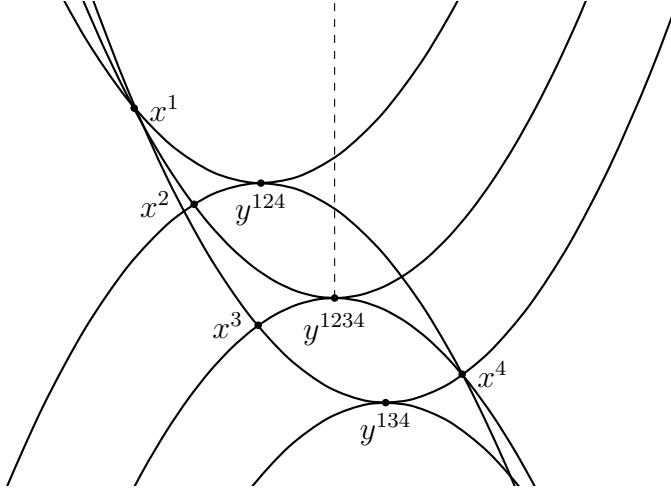


FIGURE 6. Geometric configuration used in proof of Lemma 4.11. The positioning $x^1, x^2 \in L^+$, $x^3 \in L^-$, and $x^4 \in R^-$ forces the placement of y^{124} and y^{134} to agree with this diagram.

Proof. We suppose that such a shattering configuration does satisfy

$$x^1, x^2 \in L^+, \quad x^3 \in L^-, \quad x^4 \in R^-,$$

and seek a contradiction. First, consider y^{124} . Since $x^1, x^2 \in P^+$, they cannot both be in $y^{124} + P^+$. Therefore, $x^2, x^4 \in y^{124} + P^-$. It is easy to see that y^{124} must lie above the

parabola P^+ and to the left of the vertical axis in order for $y^{124} + P^-$ to meet $x^4 \in R^-$. Since $(y^{124} + P^-) \cap L^+$ in only one point, namely x^2 , it follows that $x^1 \in y^{124} + L^+$.

Next, consider y^{134} . Since $x^3, x^4 \in P^-$, they cannot both be in $y^{134} + P^-$, so we conclude that $x^3 \in y^{134} + P^+$. It follows from Lemma 4.6 that $x^4 \in y^{134} + P^+$ as well, since we already have

$$\{x^1, x^4\} \subseteq (-D) \cap (y^{124} - D),$$

and so we cannot have $\{x^1, x^4\} \subseteq y^{134} - D$ as well. This shows that we cannot have $x^4 \in y^{134} + R^-$. We also cannot have $x^4 \in y^{134} + L^-$, since then y^{134} would lie entirely to the right of P , and so this contradicts that $x^1 \in y^{134} + P^+$.

So far, we have justified that the placement of y^{124} and y^{134} agree with Figure 6. Finally, we will find a contradiction by considering y^{234} . $y^{234} + P$ is not included in the diagram precisely because there is no geometric possibility. We see that $x^4 \in y^{234} + P^-$, since if $x^4 \in y^{234} + P^+$ then so is x^3 , and we already have

$$\{x^3, x^4\} \subseteq y^{134} + P^+.$$

Similarly, we see that $x^3 \in y^{234} + P^+$, since we already have

$$\{x^3, x^4\} \subseteq P^-.$$

Therefore $x^2 \in y^{234} + P^+$ as well. Next we claim that y^{234} lies below and to the right of y^{134} . To see this, note that $y^{134} + P^+$ and $y^{234} + P^+$ are both upward facing parabolas going through the point x^3 . Therefore, whichever vertex is farther to the right corresponds to a parabola with a steeper negative slope to the left of x^3 , and hence lies strictly above the other parabola on the interval between x^3 and the point where it intersects the parabola P^+ . Therefore, whichever one is farther to the right will also have its corresponding intersection point with P^+ be farther to the right. Thus, we must be describing the parabola going through x^3 and x^2 , not the one going through x^3 and x^1 , because x^2 lies to the right of x^1 .

This immediately gives a contradiction, since we have shown that y^{234} lies below y^{134} , and

$$x^4 \in (y^{134} + P^+) \cap (y^{234} + P^-)$$

□

Corollary 4.12. *Combining Lemmas 4.7–4.11, we conclude that x^3 cannot lie in either branch P^+ or P^- . Hence, no set of four points in $\mathbb{R}^2$ can be shattered.*

While the intersection properties of parabolas work the same way over any field, many of our arguments in Euclidean space rely in an essential way on one point being farther left than another point, or being above another point. We are using the fact that $\mathbb{R}$ is an ordered field in an essential way, which does not extend to $\mathbb{F}_q$. This is not just an inconvenience in the proof mechanism, it actually changes the VC-dimension.

4.2. Finite fields. Consider $P = P^+ \cup P^-$ over $\mathbb{F}_q$, and Let $\mathcal{H}_P(\mathbb{F}_q^2)$ be the hypothesis class consisting of indicator functions h_y of sets of the form $y + P$, $y \in \mathbb{F}_q^2$. Unlike the real case, $\mathbb{F}_q^2$ can shatter 4 points. The table below provides explicit configurations that achieve this for several fields.

$$\begin{aligned}
\mathbb{F}_{11}^2 &: \{x^1 = (0, 0), x^2 = (1, 2), x^3 = (2, 8), x^4 = (7, 4)\} \\
\mathbb{F}_{17}^2 &: \{x^1 = (0, 0), x^2 = (0, 1), x^3 = (1, 8), x^4 = (12, 13)\} \\
\mathbb{F}_{23}^2 &: \{x^1 = (0, 0), x^2 = (1, 2), x^3 = (10, 17), x^4 = (13, 6)\} \\
\mathbb{F}_{29}^2 &: \{x^1 = (0, 0), x^2 = (0, 2), x^3 = (8, 7), x^4 = (11, 2)\}
\end{aligned}$$

For the field $\mathbb{F}_{11}$, the centers of the corresponding parabolas are:

$$\begin{aligned}
y_1 &= (0, 0) & y_{1,2} &= (9, 4) \\
y_2 &= (0, 3) & y_{1,3} &= (10, 10) \\
y_3 &= (0, 4) & y_{1,4} &= (1, 1) \\
y_4 &= (0, 9) & y_{2,3} &= (0, 1) \\
& & y_{2,4} &= (2, 1) \\
& & y_{3,4} &= (1, 7) \\
& & y_{1,2,3} &= (7, 5) \\
& & y_{1,2,4} &= (5, 8) \\
& & y_{1,3,4} &= (6, 3) \\
& & y_{2,3,4} &= (10, 6) \\
& & y_{1,2,3,4} &= (3, 9)
\end{aligned}$$

We used a computer to find these values, in order to demonstrate the difference between finite fields and the real numbers in this context.

REFERENCES

- [1] E. Aksoy, A. Iosevich, and B. McDonald. Interpolation of point configurations in the discrete plane, (2024). <https://arxiv.org/abs/2408.07010>
- [2] R. Ascoli, L. Betti, J. Cheigh, A. Iosevich, R. Jeong, X. Liu, B. McDonald, W. Milgrim, S. J. Miller, F. Romero Acosta, and S. Velazquez Iannuzzelli. VC-dimension and distance chains in $\mathbb{F}_q^d$. *Korean J. Math.*, 32(1):43–57, (2024)
- [3] R. Ascoli, L. Betti, J. Cheigh, A. Iosevich, R. Jeong, X. Liu, B. McDonald, W. Milgrim, S. J. Miller, F. Romero Acosta, and S. Velazquez Innuzzelli. VC-dimension of hyperplanes over finite fields. *Graphs Combin.*, 41(2):Paper No. 47, 13, (2025)
- [4] L. Babai. The Fourier transform and equations over finite abelian groups, (2002). <http://people.cs.uchicago.edu/~laci/reu02/fourier.pdf>
- [5] M. Bennett, J. Chapman, D. Covert, D. Hart, A. Iosevich, and J. Pakianathan. Long paths in the distance graph over large subsets of vector spaces over finite fields. *J. Korean Math. Soc.*, 53(1):115–126, (2016)
- [6] M. Bennett, D. Hart, A. Iosevich, J. Pakianathan, and M. Rudnev. Group actions and geometric combinatorics in $\mathbb{F}_q^d$. *Forum Math.*, 29(1):91–110, (2017)
- [7] T. Cheek, J. Cooper, P. Gilman, A. Iosevich, K. Jaber, E. Palsson, V. Sharan, J. Shuffelton, and M.-H. Tomé. Congruence classes of simplex structures in finite field vector spaces, (2024). <https://arxiv.org/abs/2408.07912>
- [8] R. M. Dudley. Central Limit Theorems for Empirical Measures. *Ann. Probability*, 6(6):899 – 929, (1978)
- [9] P. Erdős. On sets of distances of n points. *Amer. Math. Monthly*, 53:248–250, (1946)
- [10] D. Fitzpatrick, A. Iosevich, B. McDonald, and E. Wyman. The VC-dimension and point configurations in $\mathbb{F}_q^2$. *Discrete Comput. Geom.*, (2023)
- [11] J. Fox, J. Pach, and A. Suk. Erdős-Hajnal conjecture for graphs with bounded VC-dimension. *Discrete Comput. Geom.*, 61(4):809–829, (2019)
- [12] J. Fox, J. Pach, and A. Suk. Bounded VC-dimension implies the Schur-Erdős conjecture. *Combinatorica*, 41(6):803–813, (2021)
- [13] L. Guth and N. H. Katz. On the Erdős distinct distances problem in the plane. *Ann. of Math. (2)*, 181(1):155–190, (2015)
- [14] D. Haussler. Sphere packing numbers for subsets of the Boolean n -cube with bounded Vapnik-Chervonenkis dimension. *J. Combin. Theory Ser. A*, 69(2):217–232, (1995)
- [15] D. Haussler and E. Welzl. ϵ -nets and simplex range queries. *Discrete Comput. Geom.*, 2(2):127–151, (1987)
- [16] T. Hayes. A large deviation inequality for vector valued martingales, (2001). <https://newtraell.cs.uchicago.edu/research/publications/techreports/TR-2001-24>
- [17] A. Iosevich, G. Jardine, and B. McDonald. Cycles of arbitrary length in distance graphs on $\mathbb{F}_q^d$. *Tr. Mat. Inst. Steklova*, (2021)
- [18] A. Iosevich, A. Magyar, A. McDonald, and B. McDonald. The VC-dimension and point configurations in $\mathbb{R}^d$, (2025). <https://arxiv.org/abs/2510.13984>
- [19] A. Iosevich, B. McDonald, and M. Sun. Dot products in $\mathbb{F}_q^3$ and the Vapnik-Chervonenkis dimension. *Discrete Math.*, 346(1):Paper No. 113096, 9, (2023)
- [20] A. Iosevich and M. Rudnev. Erdős distance problem in vector spaces over finite fields. *Trans. Amer. Math. Soc.*, 359(12):6127–6142, (2007)
- [21] H. Iwaniec and E. Kowalski. *Analytic Number Theory*, volume 53 of *American Mathematical Society Colloquium Publications*. American Mathematical Society, Providence, RI, (2004)
- [22] P. Maga. Full dimensional sets without given patterns. *Real Anal. Exchange*, (2010)
- [23] A. McDonald. Congruence classes of large configurations in vector spaces over finite fields. *Funct. Approx. Comment. Math.*, 62(1):131–141, (2020)
- [24] B. McDonald, A. Sahay, and E. L. Wyman. The VC-dimension of quadratic residues in finite fields. *Discrete Math.*, 348(1):Paper No. 114192, 13, (2025)
- [25] T. Nguyen, A. Scott, and P. Seymour. Induced subgraph density. VI. Bounded VC-dimension. *Adv. Math.*, 482:Paper No. 110601, (2025)

- [26] T. Pham, S. Senger, M. Tait, and V. T. H. Thu. Geometric structures in pseudo-random graphs. *Canad. J. Math.*, 77(3):1041–1071, (2025)
- [27] T. Pham, S. Senger, M. Tait, and N. Thu-Huyen. VC-dimension and pseudo-random graphs. *Discrete Appl. Math.*, 365:231–246, (2025)
- [28] S. Shalev-Schwartz and S. Ben-David. *Understanding Machine Learning: From Theory to Algorithms*. Cambridge University Press, (2014)
- [29] V. N. Vapnik and A. J. Cervonenkis. The uniform convergence of frequencies of the appearance of events to their probabilities. *Dokl. Akad. Nauk SSSR*, 181:781–783, (1968)
- [30] A. Weil. On some exponential sums. *Proc. Nat. Acad. Sci. U.S.A.*, 34:204–207, (1948)

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF GEORGIA
Email address: moustapha.diallo@uga.edu

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF GEORGIA
Email address: brian.mcdonald@uga.edu