Erdős-Falconer distance problem under Hamming metric in vector spaces over finite fields

Zixiang Xu[a] and Gennian Ge[a,*]

[a] School of Mathematics Sciences, Capital Normal University, Beijing 100048, China.

Abstract

For a subset $I \subseteq \mathbb{F}_q^n$, let $\Delta(I)$ be the set of distances determined by the elements of I. The Erdős-Falconer distance problem in $\mathbb{F}_q^n$ asks for a threshold on the cardinality $|I|$ so that $\Delta(I)$ contains a positive proportion of the whole distance set. In this paper, we consider the analogous question under Hamming distance, which is the most important metric in coding theory. When $q \geqslant 4$ is a fixed prime power and n goes to infinity, our main result shows that, for arbitrary positive proportion α, we can find αn distinct Hamming distances in $\Delta(I)$ if $|I| > q^{(1-\beta)\cdot n}$, where β is a positive number depending on α. Unlike using Fourier analytical method as usual, our main tools include the celebrated dependent random choice and some results from additive number theory and coding theory. Hence our bound is much smaller than the previously known bound which was obtained by Fourier analytic machinery.

Key words and phrases: Erdős-Falconer distance problem, Hamming distance, dependent random choice.

AMS subject classifications: 11H71, 52C10.

1 Introduction

The classical Erdős-Falconer conjecture in the Euclidean setting says that if the Hausdorff dimension of a set in $\mathbb{R}^d$ exceeds $\frac{d}{2}$, then the Lebesgue measure of the distance set is positive.

*Corresponding author (e-mail: gnge@zju.edu.cn). Research supported by the National Natural Science Foundation of China under Grant Nos. 11431003, 61571310 and 11971325, Beijing Scholars Program, Beijing Hundreds of Leading Talents Training Project of Science and Technology, and Beijing Municipal Natural Science Foundation.

This conjecture implies that if the size of the set is greater than $q^{\frac{d}{2}}$, then the distance set contains a positive proportion of all the possible distances. The first result on the Falconer distance conjecture [11] showed that if the Hausdorff dimension of a set in $\mathbb{R}^d$, $d \geqslant 2$, is greater than $\frac{d+1}{2}$, then the Lebesgue measure of the Euclidean distance set is positive. From then on, several researches improved this result via different methods, e.g. see [4, 8, 9, 17, 33] and the references therein.

Recently, finite field analogs of classical problems in harmonic analysis, geometry and combinatorics have received much attention because of the relative technical transparency afforded by the discrete setting. In [20], Iosevich and Rudnev investigated the finite field analog of the Erdős-Falconer distance problems and developed the Fourier analytic machinery to study such combinatorial problem. For more literature on the Euclidean distance and related geometric configurations, we refer the readers to [1, 15, 18, 19, 21] and the references therein. Very recently, Yazici [35] considered a similar problem under Hamming distance in $\mathbb{F}_q^n$. Using the Fourier analytic machinery, Yazici proved that if $|I| > \frac{q^{n-1}}{n}\binom{n}{\frac{n}{2}}\binom{\frac{n}{2}}{\frac{n}{4}}$, $4|n$, then the points of I determine a Hamming distance d for every even $0 < d < n-2$. Inspired by Yazici's result, in this paper we further consider the Erdős-Falconer type problem under Hamming distance.

Let $\mathbb{F}_q$ be the finite field of order q, where $q \geqslant 4$ is a prime power. For two elements $\boldsymbol{a} = (a_1, a_2, \ldots, a_n)$ and $\boldsymbol{b} = (b_1, b_2, \ldots, b_n)$ in vector space $\mathbb{F}_q^n$, the Hamming distance between $\boldsymbol{a}$ and $\boldsymbol{b}$ can be defined as $d_H(\boldsymbol{a}, \boldsymbol{b}) = \sum_{i=1}^{n} d_H(a_i, b_i)$, where $d_H(a_i, b_i)$ equals 1 if $a_i \neq b_i$, and 0 otherwise. Let I be a subset of $\mathbb{F}_q^n$ and I can be seen as a q-ary code of length n. In coding theory, we usually consider the problems under the assumption that q is fixed and $n \to \infty$. The question we will be dealing with in this paper is that, how large does the size of I need to be, to guarantee that $\Delta(I)$ contains the positive proportion of n. For convenience, we give the following definition of α-distance q-ary code under Hamming distance.

Definition 1.1. *Let I be a subset of $\mathbb{F}_q^n$, we call I an α-distance code if $\Delta(I)$ contains at least αn distinct Hamming distances. Denote the function $\mathcal{I}(n, q, \alpha)$ as the minimum size of I such that if $|I| \geqslant \mathcal{I}(n, q, \alpha)$ then I must be an α-distance code.*

An Erdős-Falconer type problem under Hamming metric in $\mathbb{F}_q^n$ can be written as follows.

Question 1.2. *For given prime power q, and a real number $0 < \alpha < 1$, determine the value of function $\mathcal{I}(n, q, \alpha)$.*

In this viewpoint, Yazici's result [35] showed that $\mathcal{I}(n, q, \frac{1}{2} - \frac{1}{n}) \leqslant \frac{q^{n-1}}{n}\binom{n}{\frac{n}{2}}\binom{\frac{n}{2}}{\frac{n}{4}}$ with $4|n$. We will focus on the asymptotic behavior of this function assuming q is fixed and $n \to \infty$. Our main result shows that, for arbitrary positive proportion $0 < \alpha < 1$, there exists some positive constant $\beta = \beta(\alpha)$, such that if $|I| > q^{(1-\beta)\cdot n}$ then I must be an α-distance code.

Theorem 1.3. *Let $q \geqslant 4$ be a prime power. For given $0 < \alpha < 1$, there exists a positive constant $\beta = \beta(\alpha) > 0$ such that*

$$q^{H_q(\alpha-\frac{1}{n})\cdot n - o(n)} \leqslant \lim_{n\to\infty} \mathcal{I}(n, q, \alpha) \leqslant q^{(1-\beta)\cdot n},$$

where $H_q(x) = x\log_q(q-1) - x\log_q x - (1-x)\log_q(1-x)$.

The rest of this paper is organized as follows. In Section 2, we introduce some relevant tools and results from coding theory, extremal combinatorics and number theory. The proof of our main result is presented in Section 3. Finally we conclude this paper and pose some open problems in Section 4.

2 Preliminaries

In this section, we briefly introduce three important tools which are useful in the proof of our main result. We first introduce some basic knowledge on coding theory such as Hamming ball and anti-code method. The second one is the celebrated dependent random choice which plays an important role in extremal combinatorics. The third result is from additive number theory.

2.1 Coding theory

2.1.1 Hamming Ball

The fundamental problem in coding theory asks that for given minimum distance d and code length n, how large of size can a q-ary code be? The direct theoretical bounds are so-called Gilbert-Varshamov bound and sphere-packing bound, both of which depend on the volume of Hamming ball. Recall that the Hamming ball of radius w in $\mathbb{F}_q^n$ is the set $B_q(n, w)$ of all q-ary words of length n and Hamming weight at most w. Then the volume of Hamming ball $B_q(n, pn)$ is $Vol_q(n, pn) = \sum_{i=0}^{pn} \binom{n}{i}(q-1)^i$. We will introduce the lower bound and the upper bound on $Vol_q(n, pn)$ as follows.

Proposition 2.1. *Let $q \geqslant 3$ be an integer and $0 \leqslant p \leqslant 1 - \frac{1}{q}$ be a real number. Then for large enough n, we have*

$$q^{H_q(p)\cdot n - o(n)} \leqslant Vol_q(n, pn) \leqslant q^{H_q(p)\cdot n},$$

where $H_q(x) = x\log_q(q-1) - x\log_q x - (1-x)\log_q(1-x)$.

Proof. First we prove $Vol_q(n, pn) \leqslant q^{H_q(p)\cdot n}$ and consider the following sequence of relations:

$$\begin{aligned}
1 =& (p + (1-p))^n \\
=& \sum_{i=0}^{n} \binom{n}{i} p^i (1-p)^{n-i} \\
\geqslant& \sum_{i=0}^{pn} \binom{n}{i} p^i (1-p)^{n-i} \\
=& \sum_{i=0}^{pn} \binom{n}{i} (q-1)^i (1-p)^n (\frac{p}{(q-1)(1-p)})^i \\
\geqslant& \sum_{i=0}^{pn} \binom{n}{i} (q-1)^i (1-p)^n (\frac{p}{(q-1)(1-p)})^{pn} \\
=& \sum_{i=0}^{pn} \binom{n}{i} (q-1)^i (\frac{p}{q-1})^{pn} (1-p)^{(1-p)n}.
\end{aligned}$$

In the above, the first inequality follows by dropping some terms from the summation and the second inequality follows from the fact that $\frac{p}{(q-1)(1-p)} \leqslant 1$ as $q \geqslant 3$, $p \leqslant 1 - \frac{1}{q}$ and $pn \geqslant 1$. Since

$$q^{H_q(p)\cdot n} = (\frac{p}{q-1})^{pn} (1-p)^{(1-p)n},$$

the last expression implies that

$$Vol_q(n, pn) \cdot q^{-H_q(p)\cdot n} \leqslant 1.$$

To prove $q^{H_q(p)\cdot n - o(n)} \leqslant Vol_q(n, pn)$, we need the following Stirling's approximation:

$$\sqrt{2\pi n}(\frac{n}{e})^n \leqslant n! \leqslant \sqrt{2\pi n}(\frac{n}{e})^n e^{\lambda(n)},$$

where $\lambda(n) = \frac{1}{12n}$. Since

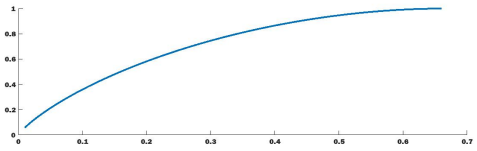
$$\begin{aligned}
\binom{n}{pn} =& \frac{n!}{(pn)!(n-pn)!} \\
\geqslant& \frac{(\frac{n}{e})^n}{(\frac{pn}{e})^{pn} (\frac{(1-p)n}{e})^{(1-p)n}} \cdot \frac{e^{-\lambda(pn)-\lambda(n-pn)}}{\sqrt{2\pi p(n-pn)}} \\
=& \frac{\ell(n)}{p^{pn}(1-p)^{(1-p)n}},
\end{aligned}$$

where $\ell(n) = \frac{e^{-\lambda(pn)-\lambda(n-pn)}}{\sqrt{2\pi p(n-pn)}}$. Now we have

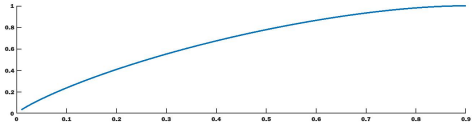
$$\begin{aligned} Vol_q(n, pn) \geqslant & \binom{n}{pn}(q-1)^{pn} \\ > & \frac{(q-1)^{pn}}{p^{pn}(1-p)^{(1-p)n}} \cdot \ell(n) \\ \geqslant & q^{H_q(n)\cdot n + \log_q(\ell(n))}. \end{aligned}$$

The proof is finished. □

Remark 2.2. The function $H_q(p)$ has several good properties. For example, $H_q(p) < 1$ for any integer $q \geqslant 3$ and the real number $0 \leqslant p \leqslant 1 - \frac{1}{q}$. Moreover, this function is monotonically increasing in p while it is monotonically decreasing in q. In order to have a more intuitive understanding of this function, let us show some examples in the figures as follows.



(a) $H_3(p)$ with $0 < p < \frac{2}{3}$



(b) $H_{10}(p)$ with $0 < p < \frac{9}{10}$

Remark 2.3. Consider the Hamming ball $B_q(n, \lceil \frac{\alpha n-1}{2} \rceil)$ with radius $\lceil \frac{\alpha n-1}{2} \rceil$. It is easy to see $|\Delta(B_q(n, \lceil \frac{\alpha n-1}{2} \rceil))| \leqslant \alpha n - 1$. Combining with Proposition 2.1, we obtain the lower bound in Theorem 1.3.

2.1.2 Anti-code method

To attack the fundamental problem in coding theory, many interesting theoretical bounds such as Plotkin bound, Griesmer bound and Johnson bound were proposed via different ideas. For more information on theoretical bounds in coding theory, we refer the readers to the textbook [29] and the references therein. In analogy to the definition of a code, a subset $A \subseteq \mathbb{F}_q^n$ is an anti-code with maximal distance d, if any two of its elements are at distance at most d. The anti-code method was usually used to construct codes that attain the Griesmer bound with equality. In particular, the exact expressions of maximal anti-code can be used to derive better upper bounds on the fundamental problem. Recently, many other variations on different spaces and metrics have created a wealth of anti-codes (see, e.g. [10, 23, 26, 27]).

Denote $C_r(t)$ as the set

$$C_r(t) = \{\boldsymbol{c} \in \mathbb{F}_q^n : |\{i : 1 \leqslant i \leqslant t+2r, c_i = 1\}| \geqslant t+r\}.$$

A result from Frankl and Tokushige [12] showed that $C_r(t)$ is the maximal non-binary anti-code with maximal distance $n-t$ under certain conditions.

Lemma 2.4 ([12]). *Let $q \geqslant 3$, and set $r := \lfloor \frac{t-1}{q-2} \rfloor$, then $C_r(t)$ is the maximal anti-code with maximal distance $n-t$ for $n \geqslant t+2r$. Moreover, we have*

$$|C_r(t)| = q^{n-t-2r} \sum_{i=0}^{r} \binom{t+2r}{i} (q-1)^i.$$

Using the previous result, we can obtain the following consequence.

Corollary 2.5. *Let $q \geqslant 3$ and $\frac{1}{2} < \gamma < 1$. Set $t := (\frac{1-\alpha}{2}) \cdot n$ and $r := \lfloor \frac{t-1}{q-2} \rfloor$. If $n \geqslant t+2r$, then we have*

$$|C_r(t)| \leqslant q^{(1-(1-H_q(\gamma))(\frac{1-\alpha}{2}))\cdot n}.$$

Proof. By Proposition 2.1, we have $|C_r(t)| \leqslant q^{n-(1-H_q(\frac{r}{t+2r}))(t+2r)}$. The result follows since $t+2r \geqslant (\frac{1-\alpha}{2}) \cdot n$ and $H_q(p)$ is monotonically increasing in p. □

Remark 2.6. In our main result, we set $q \geqslant 4$ because when $q = 2$, the maximal anti-code is Hamming ball exactly. On the other hand, when $q = 2$, the code $\mathcal{C} = \{\boldsymbol{c} : \sum_{i=1}^{n} c_i \equiv 0 \mod 2\}$ has size $\Omega(2^n)$ but $\Delta(\mathcal{C})$ just contains about $\frac{n}{2}$ distinct distances. Moreover, when $q = 3$, if $\alpha < \frac{1}{3}$, then $t+2r$ may be larger than n. Hence we will take advantage of Corollary 2.5 in the proof of Lemma 3.7 under the assumption that $q \geqslant 4$.

2.2 Dependent random choice

Early versions of the dependent random choice lemma were proved and applied by various researchers, starting with Gowers [13], who gave a new proof of Szemerédi's theorem for arithmetic progressions of length four. From then on, there were several striking applications of dependent random choice to extremal graph theory, Ramsey theory, additive combinatorics, and combinatorial geometry. For more information, we refer the readers to the survey [7] and the references therein. We state an asymmetric version of the dependent random choice lemma and give a simple proof as follows.

Lemma 2.7. *Let $H = (A \cup B, E)$ be a bipartite graph with $|A| = n$, $|B| = m$, and $e(H) = Cnm$. For a given positive integer t, there will be a subset $A' \subseteq A$ with $|A'| \geqslant \frac{C^t(n+1)}{2}$ such that every pair of vertices in A' have at least $Cmn^{-\frac{1}{t}}$ common neighbors.*

Proof. Pick a set T of t vertices from B uniformly at random with repetition. By linearity of expectation, we have

$$\mathbb{E}[|N(T)|] = \sum_{a\in A}(\frac{N_B(a)}{m})^t \geqslant n^{t-1}(\sum_{a\in A}(\frac{N_B(a)}{m}))^t = n^{t-1}(Cn)^t = C^t n,$$

here we take advantage of the mean inequality: for any $0 < p \leqslant q$, we have $\sum\limits_{i=1}^{n} x_i^p \leqslant n^{1-\frac{p}{q}}(\sum\limits_{i=1}^{n} x_i^q)^{\frac{p}{q}}$. Let Y denote the random variable counting the number of pairs $(a_1, a_2) \subseteq N(T)$ with fewer than $Cmn^{-\frac{1}{t}}$ common neighbors. For a given such pair (a_1, a_2), the probability that it is a subset of $N(T)$ is $(\frac{|N_B(a_1)\cap N_B(a_2)|}{m})^t$. Since there are at most $\binom{n}{2}$ pairs for which $|N_B(a_1) \cap N_B(a_2)| < Cmn^{-\frac{1}{t}}$, it follows that

$$\mathbb{E}[Y] < \binom{n}{2}(\frac{|N_B(a_1) \cap N_B(a_2)|}{m})^t = \frac{C^t(n-1)}{2}.$$

Using the linearity of expectation again,

$$\mathbb{E}[|N(T)| - Y] \geqslant C^t n - \frac{C^t(n-1)}{2} = \frac{C^t(n+1)}{2}.$$

Hence there is a choice of T such that $|N(T)| - Y \geqslant \frac{C^t(n+1)}{2}$. Delete one vertex from each pair of vertices with fewer than $Cmn^{-\frac{1}{t}}$ common neighbors. Let A' be the remaining subset of A, and it is easy to check that A' has at least $\frac{C^t(n+1)}{2}$ vertices and each pair of vertices in A' have at least $Cmn^{-\frac{1}{t}}$ common neighbors. □

Remark 2.8. In this version of dependent random choice lemma, we do not require H to be a dense graph. For instance, we can choose $C = \Theta(n^c)$ instead of a constant, where c is a negative number. In this case, we just need to verify that $\frac{C^t(n+1)}{2} \geqslant 2$ and $Cmn^{-\frac{1}{t}} \geqslant 2$.

2.3 Additive number theory

Many classical problems in additive number theory revolve around the study of sum sets for specific sets A and B. For example, let $\mathbb{N}^2 = \{1, 4, 9, 16, \cdots\}$ be the set of square numbers, then there is a famous theorem of Lagrange that $4\mathbb{N}^2 = \mathbb{N}$, that is, every natural number is the sum of four squares. Let $\mathbb{P} = \{2, 3, 5, 7, \cdots\}$ be the set of prime number, the infamous Goldbach conjecture asserts that $2\mathbb{P}$ contains every even integer greater than 2, but this conjecture remains far from resolution. In order to solve the Goldbach conjecture, there is a famous theorem of Vinogradov [30], which states that $(2 \cdot \mathbb{N} + 1) \setminus 3\mathbb{P}$ is finite, i.e. every sufficiently large odd number is the sum of three primes. For more information on additive number theory, we

refer the readers to the great textbooks [25] and [28]. Very recently, using the transference principle introduced by Green [14], Matomäki, Maynard and Shao [24] showed an extension of Vinogradov's theorem as follows.

Theorem 2.9 ([24]). *Let $\theta > \frac{11}{20}$. Every sufficiently large odd integer n can be written as a sum of three primes $n = p_1 + p_2 + p_3$ with $|p_i - \frac{n}{3}| \leqslant n^\theta$ for $i \in \{1, 2, 3\}$.*

Since the best known result on primes in short intervals due to Baker, Harman and Pintz [3] showed the existence of primes in intervals $[x, x + x^{0.525}]$, we can obtain a consequence for large even integer as follows.

Corollary 2.10. *Let $\theta > \frac{11}{20}$. Every sufficiently large even integer n can be written as a sum of four primes $n = p_1 + p_2 + p_3 + p_4$ with $|p_i - \frac{n}{4}| \leqslant n^\theta$ for $i \in \{1, 2, 3, 4\}$.*

3 Proof of the main result

Our proof is under the assumption that $q \geqslant 4$ is a prime power and n goes to infinity.

3.1 The existence of small prime distances

The first step of our proof is to show the existence of small prime distances in $\Delta(I)$. We need the following result in [2].

Lemma 3.1 ([2]). *Let p be a prime and $I \subseteq \mathbb{F}_q^n$. If*

$$|I| > \sum_{i=0}^{s} (q-1)^i \binom{n}{i},$$

and $s \geqslant p - 1$, then there exists some element $d \in \Delta(I)$ such that $d \equiv 0 \mod p$.

Then we show how to find prime distances in certain interval.

Lemma 3.2. *For given real numbers $0 < \alpha < 1$ and $\frac{1}{2} < \gamma < 1$, let I be a subset of $\mathbb{F}_q^n$ with $|I| > q^{(1-\frac{(1-\alpha)(1-H_q(\gamma))}{2})\cdot n}$. If p is a prime number with $\frac{(1-\alpha)\cdot n}{2} < p < \gamma \cdot n$, then $\Delta(I)$ contains p.*

Proof. Suppose $|I| > q^{(1-\frac{(1-\alpha)(1-H_q(\gamma))}{2})\cdot n}$, solving the following inequality on x,

$$\frac{n-x}{2} < p < \gamma \cdot (n - x),$$

we obtain that

$$n - 2p < x < n - \frac{p}{\gamma} < (1 - \frac{1-\alpha}{2\gamma}) \cdot n.$$

Now we choose a positive integer x from interval $[n-2p, n-\frac{p}{\gamma}]$. Consider the subset I_{n-x} of I such that every element in I_{n-x} has the same first x coordinates. It is easy to see that

$$|I_{n-x}| > q^{(1-\frac{(1-\alpha)(1-H_q(\gamma))}{2})\cdot n-x}.$$

Since $\gamma < 1$ and $x < (1-\frac{1-\alpha}{2\gamma})\cdot n < (1-\frac{1-\alpha}{2})\cdot n$, we have that

$$(1-H_q(\gamma))\cdot(n-x) > (\frac{(1-H_q(\gamma)(1-\alpha))}{2})\cdot n.$$

Hence

$$q^{(1-\frac{(1-\alpha)(1-H_q(\gamma))}{2})\cdot n-x} \geqslant q^{H_q(\gamma)\cdot(n-x)} \geqslant \sum_{i=0}^{\gamma(n-x)} (q-1)^i \binom{n-x}{i}.$$

We see that $p < \gamma\cdot(n-x)$, by Lemma 3.1, there exists some element $d \in \Delta(I_{n-x})$ such that $d \equiv 0 \mod p$. Since $p > \frac{n-x}{2}$, then we conclude $\Delta(I_{n-x})$ contains the prime p. Now we have found the prime distance p in $\Delta(I)$ since $\Delta(I_{n-x}) \subseteq \Delta(I)$. □

3.2 Find the small distances

The previous result guarantees the existence of small prime distances in $\Delta(I)$. Next we will combine the operation from coding theory and the result from number theory to find the small distances in $\Delta(I)$.

First we define the direct sum of two elements $\boldsymbol{c}_1, \boldsymbol{c}_2$ in $\mathbb{F}_q^{n_1}$ and $\mathbb{F}_q^{n_2}$, respectively.

Definition 3.3. *For given two disjoint subspaces $\mathbb{F}_q^{n_1}$ and $\mathbb{F}_q^{n_2}$, we define the direct sum $\boldsymbol{c}_1 \oplus \boldsymbol{c}_2$ of $\boldsymbol{c}_1 \in \mathbb{F}_q^{n_1}$ and $\boldsymbol{c}_2 \in \mathbb{F}_q^{n_2}$ as the element in $\mathbb{F}_q^{n_1+n_2}$. More precisely, if we view $\boldsymbol{c}_1$ and $\boldsymbol{c}_2$ as vectors in corresponding spaces, then the i-th coordinate of $\boldsymbol{c}_1 \oplus \boldsymbol{c}_2$ is*

$$(\boldsymbol{c}_1 \oplus \boldsymbol{c}_2)_i = \begin{cases} (\boldsymbol{c}_1)_i, & \text{if } i \in [1, n_1] \\ (\boldsymbol{c}_2)_i, & \text{otherwise.} \end{cases}$$

Moreover, define the direct sum of $I_1 \subseteq \mathbb{F}_q^{n_1}$ and $I_2 \subseteq \mathbb{F}_q^{n_2}$ as

$$I_1 \oplus I_2 = \{\boldsymbol{c_1} \oplus \boldsymbol{c_2} : \boldsymbol{c_1} \in I_1, \boldsymbol{c_2} \in I_2\}.$$

Lemma 3.4. *For given integers n_1, n_2, d_1 and d_2, let $I_1 \in \mathbb{F}_q^{n_1}$ and $I_2 \in \mathbb{F}_q^{n_2}$. Let δ_1, δ_2 and t satisfy the following conditions*

- *if $|I_1| > \delta_1 q^{n_1}$, then $d_1 \in \Delta(I_1)$,*
- *if $|I_2| > \delta_2 q^{n_2}$, then $d_2 \in \Delta(I_2)$,*

- $2\delta_1\delta_2^{-t} \geqslant q^{n_1}$.

Then for any subset $I \subseteq \mathbb{F}_q^{n_1+n_2}$ with $|I| > (2\delta_1)^{\frac{1}{t}} q^{n_1+n_2}$, $\Delta(I)$ contains distance $d_1 + d_2$.

Proof. For a given subset $I \subseteq \mathbb{F}_q^{n_1+n_2}$ with $|I| > (2\delta_1)^{\frac{1}{t}} q^{n_1+n_2}$, consider the bipartite graph $H = (A \cup B, E)$ with $A = \mathbb{F}_q^{n_1}$ and $B = \mathbb{F}_q^{n_2}$ such that two vertices $\boldsymbol{a} \in A$ and $\boldsymbol{b} \in B$ are adjacent if $a \oplus b \in I$. It is easy to see the number of edges in H is $e(H) > (2\delta_1)^{\frac{1}{t}} q^{n_1+n_2}$. By Lemma 2.7, there exists a subset $A' \subseteq A$ with at least $\delta_1(q^{n_1}+1)$ vertices such that, every pair $(\boldsymbol{a_1}, \boldsymbol{a_2}) \subseteq A'$ has at least $(2\delta_1)^{\frac{1}{t}} q^{n_2 - \frac{n_1}{t}}$ common neighbors. By the property of δ_1, there exists a pair $(\boldsymbol{a_1}, \boldsymbol{a_2}) \subseteq A'$ such that $d_H(\boldsymbol{a_1}, \boldsymbol{a_2}) = d_1$. Fix such a pair $(\boldsymbol{a_1}, \boldsymbol{a_2})$, the set of common neighbors $N_B(\boldsymbol{a_1}, \boldsymbol{a_2})$ satisfies

$$|N_B(\boldsymbol{a_1}, \boldsymbol{a_2})| \geqslant (2\delta_1)^{\frac{1}{t}} q^{n_2 - \frac{n_1}{t}} > \delta_2 q^{n_2}.$$

Hence there exists a pair $(\boldsymbol{b_1}, \boldsymbol{b_2}) \subseteq N_B(\boldsymbol{a_1}, \boldsymbol{a_2})$ such that $d_H(\boldsymbol{b_1}, \boldsymbol{b_2}) = d_2$. Now we have found a pair of elements $(\boldsymbol{a_1} \oplus \boldsymbol{b_1}, \boldsymbol{a_2} \oplus \boldsymbol{b_2}) \subseteq I$ such that $d_H(\boldsymbol{a_1} \oplus \boldsymbol{b_1}, \boldsymbol{a_2} \oplus \boldsymbol{b_2}) = d_1 + d_2$. □

Now we are ready to show the existence of small distances in $\Delta(I)$. First we consider the complicated case that d is even. As we can see from Corollary 2.10, we can only guarantee that a large even natural number can be written as the sum of four almost equal primes.

Lemma 3.5. *Let $\epsilon > 0$ be a sufficiently small real number and $\frac{1}{2} < \gamma < 1$. For a given even integer d with $(\frac{1-\alpha}{2}) \cdot n \leqslant d \leqslant (\gamma - \epsilon) \cdot n$, if a code $I \subseteq \mathbb{F}_q^n$ satisfies $|I| > q^{(1 - \frac{(1-\alpha)^4(1-H_q(\gamma))^4}{3 \cdot 2^{17}}) \cdot n}$, then $d \in \Delta(I)$.*

Proof. As $(\frac{1-\alpha}{2}) \cdot n \leqslant d \leqslant (\gamma - \epsilon) \cdot n$, by Corollary 2.10, we can write the even distance d as

$$d = d_1 + d_2 + d_3 + d_4$$

such that d_i is prime and $|d_i - \frac{d}{4}| \leqslant d^\theta$ with $\theta > \frac{11}{20}$ for $i \in \{1, 2, 3, 4\}$.

On the other hand, we can partition n into four almost equal parts, that is,

$$n = n_1 + n_2 + n_3 + n_4$$

such that $|n_i - \frac{n}{4}| \leqslant 1$ for $i \in \{1, 2, 3, 4\}$.

As $n \to \infty$, we show the direct lower bound and upper bound for d_i as follows.

$$\begin{aligned} d_i \geqslant & \frac{d}{4} - d^\theta \geqslant (\frac{1-\alpha}{8}) \cdot n - d^\theta \geqslant (\frac{1-\alpha}{4}) \cdot n_i, \\ d_i \leqslant & \frac{d}{4} + d^\theta \leqslant (\frac{\gamma - \epsilon}{4}) \cdot n + d^\theta \leqslant \gamma \cdot n_i. \end{aligned}$$

Now we partition $\mathbb{F}_q^n$ into four disjoint parts $\mathbb{F}_q^{n_1}$, $\mathbb{F}_q^{n_2}$, $\mathbb{F}_q^{n_3}$ and $\mathbb{F}_q^{n_4}$. Lemma 3.2 tells that if $I_i \in \mathbb{F}_q^{n_i}$ with $|I_i| > q^{(1-\frac{(1-\alpha)(1-H_q(\gamma))}{2})\cdot n_i}$, then $\Delta(I_i)$ contains the prime number $d_i \in ((\frac{1-\alpha}{2})\cdot n_i, \gamma \cdot n_i)$, for $i \in \{1,2,3,4\}$.

Then we want to find distance $d_1 + d_2$ in $\Delta(I_1 \oplus I_2)$, for convenience, write $I_1 \oplus I_2$ as $I_{1,2}$. We will take advantage of Lemma 3.4. We begin to confirm the following conditions.

- $|I_1| > q^{(1-\frac{(1-\alpha)(1-H_q(\gamma))}{2})\cdot n_1}$,
- $|I_2| > q^{(1-\frac{(1-\alpha)(1-H_q(\gamma))}{2})\cdot n_2}$,
- $2 \cdot q^{-(\frac{(1-\alpha)(1-H_q(\gamma))}{4})\cdot n_1} \cdot q^{(\frac{(1-\alpha)(1-H_q(\gamma))}{4})\cdot n_2 t_1} \geqslant q^{n_1}$.

Solving the third inequality, we can set $t_1 = \lceil \frac{8}{(1-\alpha)(1-H_q(\gamma))} \rceil$. After some easy calculations, we obtain that if $|I_{1,2}| > q^{(1-\frac{(1-\alpha)^2(1-H_q(\gamma))^2}{128})\cdot(n_1+n_2)}$, then $\Delta(I_{1,2})$ contains the distance $d_1 + d_2$. Then we use Lemma 3.4 again, that is, we need to confirm the following conditions.

- $|I_{1,2}| > q^{(1-\frac{(1-\alpha)^2(1-H_q(\gamma))^2}{128})\cdot(n_1+n_2)}$,
- $|I_3| > q^{(1-\frac{(1-\alpha)(1-H_q(\gamma))}{2})\cdot n_3}$,
- $2 \cdot q^{(1-\frac{(1-\alpha)^2(1-H_q(\gamma))^2}{128})\cdot(n_1+n_2)} \cdot q^{(\frac{(1-\alpha)(1-H_q(\gamma))}{4})\cdot n_3 t_2} \geqslant q^{n_1+n_2}$.

We can set $t_2 = \lceil \frac{16}{(1-\alpha)(1-H_q(\gamma))} \rceil$. Denote $I_{1,2,3} = I_1 \oplus I_2 \oplus I_3$, similarly we obtain that if $|I_{1,2,3}| > q^{(1-\frac{(1-\alpha)^3(1-H_q(\gamma))^3}{3\cdot 2^{11}})\cdot(n_1+n_2+n_3)}$, then $\Delta(I_{1,2,3})$ contains the distance $d_1 + d_2 + d_3$. We will use Lemma 3.4 for the last time by checking the following conditions.

- $|I_{1,2,3}| > q^{(1-\frac{(1-\alpha)^3(1-H_q(\gamma))^3}{3\cdot 2^{11}})\cdot(n_1+n_2+n_3)}$,
- $|I_4| > q^{(1-\frac{(1-\alpha)(1-H_q(\gamma))}{2})\cdot n_4}$,
- $2 \cdot q^{(1-\frac{(1-\alpha)^3(1-H_q(\gamma))^3}{3\cdot 2^{11}})\cdot(n_1+n_2+n_3)} \cdot q^{(\frac{(1-\alpha)(1-H_q(\gamma))}{4})\cdot n_4 t_3} \geqslant q^{n_1+n_2+n_3}$.

We can set $t_3 = \lceil \frac{24}{(1-\alpha)(1-H_q(\gamma))} \rceil$. Finally, we obtain that if $|I| > q^{(1-\frac{(1-\alpha)^4(1-H_q(\gamma))^4}{3\cdot 2^{17}})\cdot n}$, then $d \in \Delta(I)$. The proof is finished.

□

Remark 3.6. When d is odd, we can obtain a similar result via Theorem 2.9. The arguments for odd case are easier and the threshold on cardinality $|I|$ is smaller than that in Lemma 3.5.

3.3 Find the large distances

In the last step, we want to find the remaining distances in our code I, that is, we will show the existence of the distances $d \in [(\gamma-\epsilon)\cdot n, (\frac{1+\alpha}{2})\cdot n]$. First we recall two results we have shown in Lemma 2.4 and Lemma 3.5.

- If $I \in \mathbb{F}_q^n$ satisfies $|I| > q^{(1-f_1(\frac{1-\alpha}{2}))\cdot n}$, then $\Delta(I)$ contains some distance $d \in ((\frac{1+\alpha}{2})\cdot n, n]$, where $f_1(\frac{1-\alpha}{2}) = \frac{(1-H_q(\gamma-\epsilon))\cdot(1-\alpha)}{2}$.
- If $I \in \mathbb{F}_q^n$ satisfies $|I| > q^{(1-f_2(\frac{1-\alpha}{2}))\cdot n}$, then $\Delta(I)$ contains all distances d in interval $[(\frac{1-\alpha}{2})\cdot n, (\gamma-\epsilon)\cdot n]$, where $f_2(\frac{1-\alpha}{2}) = \frac{(1-\alpha)^4(1-H_q(\gamma))^4}{3\cdot 2^{17}}$.

Our goal is to prove the following result.

Lemma 3.7. *Let $c, \epsilon > 0$ be sufficiently small real numbers and $\frac{1}{2} < \gamma < 1$. For a given integer d with $(\gamma-\epsilon)\cdot n \leqslant d \leqslant (\frac{1+\alpha}{2})\cdot n$, if a code $I \subseteq \mathbb{F}_q^n$ satisfies $|I| > q^{(1-\frac{(1-\alpha)\cdot f_1(\frac{(1-\alpha)(\gamma-\epsilon-c)}{4})\cdot f_2(\frac{\gamma-\epsilon}{2}-c)}{32})\cdot n}$, then $d \in \Delta(I)$.*

Proof. Suppose that d is even and the case of odd d is similar. We first partition $\mathbb{F}_q^n$ into two disjoint parts $\mathbb{F}_q^{n_1}$ and $\mathbb{F}_q^{n_2}$, where $n = n_1 + n_2$. Moreover, we need the following conditions on n_1, n_2 and d.

- $|n_1 + \frac{(\gamma-\epsilon)\cdot n_2}{2} - d| \leqslant 1$,
- $\frac{n}{4} \leqslant n_1 \leqslant \frac{(1+\alpha)\cdot n}{2}$.

Then we will show that there exists some distance $\bar{d} \in [(1-\frac{(1-\alpha)\cdot(\gamma-\epsilon-c)}{4})\cdot n_1, n_1]$. Similar as Lemma 3.4, we consider the corresponding bipartite graph $H = (A \cup B, E)$ with $A = \mathbb{F}_q^{n_1}$ and $B = \mathbb{F}_q^{n_2}$ such that $\boldsymbol{a} \in A$ is adjacent to $\boldsymbol{b} \in B$ if $\boldsymbol{a} \oplus \boldsymbol{b} \in I$. It is easy to see $e(H) > q^{-(\frac{(1-\alpha)\cdot f_1(\frac{(1-\alpha)(\gamma-\epsilon-c)}{4})\cdot f_2(\frac{\gamma-\epsilon}{2}-c)}{32})\cdot n}\cdot q^n$. Using Lemma 2.7 and setting $t = \lceil \frac{4}{(1-\alpha)\cdot f_2(\frac{\gamma-\epsilon}{2}-c)} \rceil$, we can find a subset $A' \subseteq A$ with size

$$\begin{aligned}
|A'| \geqslant & \frac{q^{-(\frac{(1-\alpha)\cdot f_1(\frac{(1-\alpha)(\gamma-\epsilon-c)}{4})\cdot f_2(\frac{\gamma-\epsilon}{2}-c)}{32})\cdot nt}\cdot(q^{n_1}+1)}{2} \\
\geqslant & \frac{q^{-\frac{f_1(\frac{(1-\alpha)(\gamma-\epsilon-c)}{4})\cdot n}{8}}\cdot(q^{n_1}+1)}{2} \\
> & q^{-\frac{f_1(\frac{(1-\alpha)(\gamma-\epsilon-c)}{4})\cdot n}{4}}\cdot q^{n_1} \\
\geqslant & q^{(1-f_1(\frac{(1-\alpha)(\gamma-\epsilon-c)}{4}))\cdot n_1},
\end{aligned}$$

where the last inequality is from the assumption $n_1 \geqslant \frac{n}{4}$. Using Lemma 2.4, there is some pair $(\boldsymbol{a_1}, \boldsymbol{a_2}) \subseteq A'$ such that $d_H(\boldsymbol{a_1}, \boldsymbol{a_2}) = \bar{d} \in [(1 - \frac{(1-\alpha)\cdot(\gamma-\epsilon-c)}{4}) \cdot n_1, n_1]$. Moreover, we consider the set $N_B(\boldsymbol{a_1}, \boldsymbol{a_2})$ which consists of all common neighbors in $B = \mathbb{F}_q^{n_2}$. Using Lemma 2.7, we have

$$\begin{aligned}
|N_B(\boldsymbol{a_1}, \boldsymbol{a_2})| \geqslant & q^{-\frac{(1-\alpha)\cdot f_1(\frac{(1-\alpha)(\gamma-\epsilon-c)}{4})\cdot f_2(\frac{\gamma-\epsilon}{2}-c)\cdot n}{32}} \cdot |A|^{-\frac{1}{t}} \cdot |B| \\
> & q^{-\frac{2n_1}{t}+n_2} \\
\geqslant & q^{\frac{-2(1-\alpha)f_2(\frac{\gamma-\epsilon}{2}-c)\cdot n_1}{4}+n_2} \\
> & q^{(1-f_2(\frac{\gamma-\epsilon}{2}-c))\cdot n_2},
\end{aligned}$$

where the last inequality holds since $\frac{n_1}{n_2} = \frac{n_1}{n-n_1} \leqslant \frac{1+\alpha}{1-\alpha} < \frac{2}{1-\alpha}$. Now we have shown that $\Delta(N_B(\boldsymbol{a_1}, \boldsymbol{a_2}))$ contains every distance in $[(\frac{\gamma-\epsilon}{2} - c) \cdot n_2, (\gamma - \epsilon) \cdot n_2]$. Under the assumption that $|n_1 + \frac{(\gamma-\epsilon)\cdot n_2}{2} - d| \leqslant 1$, one can easily check that $d - \bar{d} \in [(\frac{\gamma-\epsilon}{2} - c) \cdot n_2, (\gamma - \epsilon) \cdot n_2]$. Hence there exists $(\boldsymbol{b_1}, \boldsymbol{b_2}) \in N_B(\boldsymbol{a_1}, \boldsymbol{a_2})$ such that $d_H(\boldsymbol{b_1}, \boldsymbol{b_2}) = d - \bar{d}$. Now we have found a pair of elements $(\boldsymbol{a_1} \oplus \boldsymbol{b_1}, \boldsymbol{a_2} \oplus \boldsymbol{b_2}) \subseteq I$ satisfying $d_H(\boldsymbol{a_1} \oplus \boldsymbol{b_1}, \boldsymbol{a_2} \oplus \boldsymbol{b_2}) = d$. □

Proof of Theorem 1.3. Combining Lemma 3.2, Lemma 3.5 and Lemma 3.7 together gives the existence of any distances in $[\frac{(1-\alpha)\cdot n}{2}, \frac{(1+\alpha)\cdot n}{2}]$, and the parameter β in Theorem 1.3 can be taken as

$$\min\{\frac{(1-\alpha)\cdot f_1(\frac{(1-\alpha)(\gamma-\epsilon-c)}{4})\cdot f_2(\frac{\gamma-\epsilon}{2}-c)}{32}, \frac{(1-\alpha)^4\cdot(1-H_q(\gamma))^4}{3\cdot 2^{17}}\},$$

the proof of our main result is finished. □

4 Conclusions and some open problems

In this paper, we consider the Erdős-Falconer type problem under Hamming distance in vector spaces over finite fields. More precisely, under the assumption that q is fixed and n goes to infinity, our main result shows that for arbitrary positive proportion α, we can find αn distinct Hamming distances in code I with $|I| > q^{(1-\beta)\cdot n}$, where $\beta = \beta(\alpha) > 0$. Unlike using the Fourier analytical method in [35], we propose a combinatorial approach, which can overcome the shortcomings of Fourier analytical methods.

In addition to Euclidean metric and Hamming metric, there are many other different metrics that play an important role in coding theory and applications. Moreover, it will be interesting to study the Erdős-Falconer type problem in other spaces rather than finite field. We list some of them for further research. The first example is permutation code, which has been extensively studied due to its potentials in various applications such as DNA storage and flash memory. Let $\pi = (\pi(1), \pi(2), \ldots, \pi(n))$ be a permutation over $[n]$, known as the vector notation of a permutation.

1. **Hamming distance:** We can view permutation code as a special case of multiply constant weight q-ary code (see, e.g. [6, 31]). Using this relationship, we can obtain a similar result as Theorem 1.3, that is, a permutation code $\mathcal{C} \subseteq S_n$ can determine arbitrary positive proportion α of distance set $[n]$ with $|\mathcal{C}| > (n!)^{1-\beta}$, $\beta = \beta(\alpha) > 0$. It will be interesting to improve this bound via special properties of permutation codes.

2. **Kendall's τ-distance:** Given a permutation $\pi = (\pi(1), \pi(2), \ldots, \pi(n))$, an adjacent transposition is an exchange of two adjacent elements $\pi(i), \pi(i+1)$, for some $1 \leqslant i \leqslant n-1$, resulting in the permutation $(\pi(1), \ldots, \pi(i-1), \pi(i+1), \pi(i), \pi(i+2), \ldots, \pi(n))$. Then the Kendall's τ-distance $d_K(\sigma, \pi)$ is the minimum number of adjacent transpositions required to transform one permutation into the other. Several useful results have been proposed in [5, 32], where the maximal anti-code can be determined in certain conditions. However, our method fails in this situation since the idea of direct sum under Kendall's τ-metric does not work.

3. **Block permutation distance:** Given a permutation $\pi = (\pi(1), \pi(2), \ldots, \pi(n))$, denote the characteristic set $A(\pi)$ as $A(\pi) \triangleq \{(\pi(i), \pi(i+1)) : 1 \leqslant i \leqslant n-1\}$. Then the block permutation distance can be represented by the following formula

$$d_B(\sigma, \pi) = |A(\pi) \setminus A(\sigma)|.$$

   The volume of ball in this metric has been estimated (see [34]), but we do not know the structure of maximal anti-codes in general.

We are also interested in the Lee metric (also called the zig-zag metric, the ℓ_1-norm), due to its applications in interleaving schemes and multidimensional burst-error-correction. Moreover, the so-called Golomb-Welch conjecture (see the survey paper [16] and recent progresses [22, 36]) is based on this metric. For any two words $\boldsymbol{u} = (u_1, u_2, \ldots, u_n) \in \mathbb{Z}^n$ and $\boldsymbol{v} = (v_1, v_2, \ldots, v_n) \in \mathbb{Z}^n$, the Lee metric between them is defined as

$$d_L(\boldsymbol{u}, \boldsymbol{v}) = \sum_{i=1}^{n} |u_i - v_i|.$$

However, the situation is worse when we consider the Erdős-Falconer type problem under Lee metric, since we just know how to estimate the volume of Lee ball. More new ideas are needed to deal with this problem.

On the other hand, it seems to be challenging to solve Question 1.2 completely, hence closing the gap between upper bound and lower bound in Theorem 1.3 is of great interest.

Acknowledgements

The first author would like to thank Zuo Ye for helpful comments.

References

[1] S. D. Adhikari, A. Mukhopadhyay, and M. Ram Murty. The analog of the Erdös distance problem in finite fields. *Int. J. Number Theory*, 13(9):2319–2333, 2017.

[2] L. Babai, H. Snevily, and R. M. Wilson. A new proof of several inequalities on codes and sets. *J. Combin. Theory Ser. A*, 71(1):146–153, 1995.

[3] R. C. Baker, G. Harman, and J. Pintz. The difference between consecutive primes. II. *Proc. London Math. Soc. (3)*, 83(3):532–562, 2001.

[4] J. Bourgain. Hausdorff dimension and distance sets. *Israel J. Math.*, 87(1-3):193–201, 1994.

[5] S. Buzaglo and T. Etzion. Bounds on the size of permutation codes with the Kendall τ-metric. *IEEE Trans. Inform. Theory*, 61(6):3241–3250, 2015.

[6] Y. M. Chee, Z. Cherif, J.-L. Danger, S. Guilley, H. M. Kiah, J.-L. Kim, P. Solé, and X. Zhang. Multiply constant-weight codes and the reliability of loop physically unclonable functions. *IEEE Trans. Inform. Theory*, 60(11):7026–7034, 2014.

[7] D. Covert, A. Iosevich, and J. Pakianathan. Geometric configurations in the ring of integers modulo p^ℓ. *Indiana Univ. Math. J.*, 61(5):1949–1969, 2012.

[8] X. Du and R. Zhang. Sharp L^2 estimates of the Schrödinger maximal function in higher dimensions. *Ann. of Math. (2)*, 189(3):837–861, 2019.

[9] M. B. Erdoğan. A bilinear Fourier extension theorem and applications to the distance set problem. *Int. Math. Res. Not.*, (23):1411–1425, 2005.

[10] T. Etzion, M. Schwartz, and A. Vardy. Optimal tristance anticodes in certain graphs. *J. Combin. Theory Ser. A*, 113(2):189–224, 2006.

[11] K. J. Falconer. On the Hausdorff dimensions of distance sets. *Mathematika*, 32(2):206–212 (1986), 1985.

[12] P. Frankl and N. Tokushige. The Erdos-Ko-Rado theorem for integer sequences. *Combinatorica*, 19(1):55–63, 1999.

[13] W. T. Gowers. A new proof of Szemerédi's theorem for arithmetic progressions of length four. *Geom. Funct. Anal.*, 8(3):529–551, 1998.

[14] B. Green. Roth's theorem in the primes. *Ann. of Math. (2)*, 161(3):1609–1636, 2005.

[15] A. Greenleaf, A. Iosevich, B. Liu, and E. Palsson. A group-theoretic viewpoint on Erdos-Falconer problems and the Mattila integral. *Rev. Mat. Iberoam.*, 31(3):799–810, 2015.

[16] P. Horak and D. Kim. 50 years of the Golomb-Welch conjecture. *IEEE Trans. Inform. Theory*, 64(4, part 2):3048–3061, 2018.

[17] A. Iosevich. What is Falconer's conjecture? *Notices Amer. Math. Soc.*, 66(4):552–555, 2019.

[18] A. Iosevich and D. Koh. The Erdös-Falconer distance problem, exponential sums, and Fourier analytic approach to incidence theorems in vector spaces over finite fields. *SIAM J. Discrete Math.*, 23(1):123–135, 2008/09.

[19] A. Iosevich and E. Palsson. An improved dimensional threshold for the angle problem. *arXiv preprint*, arXiv: 1807.05465, 2018.

[20] A. Iosevich and M. Rudnev. Erdos distance problem in vector spaces over finite fields. *Trans. Amer. Math. Soc.*, 359(12):6127–6142, 2007.

[21] D. Koh and C.-Y. Shen. The generalized Erdös-Falconer distance problems in vector spaces over finite fields. *J. Number Theory*, 132(11):2455–2473, 2012.

[22] K. H. Leung and Y. Zhou. No lattice tiling of $\mathbb{Z}^n$ by Lee sphere of radius 2. *J. Combin. Theory Ser. A, to appear.*

[23] W. J. Martin and X. J. Zhu. Anticodes for the Grassmann and bilinear forms graphs. *Des. Codes Cryptogr.*, 6(1):73–79, 1995.

[24] K. Matomäki, J. Maynard, and X. Shao. Vinogradov's theorem with almost equal summands. *Proc. Lond. Math. Soc. (3)*, 115(2):323–347, 2017.

[25] M. B. Nathanson. *Additive number theory*, volume 164 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 1996. The classical bases.

[26] M. Schwartz and T. Etzion. Codes and anticodes in the Grassman graph. *J. Combin. Theory Ser. A*, 97(1):27–42, 2002.

[27] M. Schwartz and I. Tamo. Optimal permutation anticodes with the infinity norm via permanents of $(0, 1)$-matrices. *J. Combin. Theory Ser. A*, 118(6):1761–1774, 2011.

[28] T. Tao and V. Vu. *Additive combinatorics*, volume 105 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 2006.

[29] J. H. van Lint. *Introduction to coding theory*, volume 86 of *Graduate Texts in Mathematics*. Springer-Verlag, Berlin, third edition, 1999.

[30] I. M. Vinogradov. The method of trigonometrical sums in the theory of numbers. *Trav. Inst. Math. Stekloff*, 23:109, 1947.

[31] X. Wang, H. Wei, C. Shangguan, and G. Ge. New bounds and constructions for multiply constant-weight codes. *IEEE Trans. Inform. Theory*, 62(11):6315–6327, 2016.

[32] X. Wang, Y. Zhang, Y. Yang, and G. Ge. New bounds of permutation codes under Hamming metric and Kendall's τ-metric. *Des. Codes Cryptogr.*, 85(3):533–545, 2017.

[33] T. Wolff. Decay of circular means of Fourier transforms of measures. *Internat. Math. Res. Notices*, (10):547–567, 1999.

[34] S. Yang, C. Schoeny, and L. Dolecek. Theoretical bounds and constructions of codes in the generalized Cayley metric. *IEEE Trans. Inform. Theory*, 65(8):4746–4763, 2019.

[35] E. A. Yazici. Hamming distances in vector spaces over finite fields. *arXiv preprint*, arXiv: 1910.05557, 2019.

[36] T. Zhang and Y. Zhou. On the nonexistence of lattice tilings of $\mathbb{Z}^n$ by Lee spheres. *J. Combin. Theory Ser. A*, 165:225–257, 2019.