Title
Higher Moments Subset Sum Problem over Finite Fields

Author
Nguyen, Jennifer

Publication Date
2019

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA,
IRVINE

Higher Moments Subset Sum Problem over Finite Fields

DISSERTATION

submitted in partial satisfaction of the requirements
for the degree of

DOCTOR OF PHILOSOPHY

in Mathematics

by

Jennifer Nguyen

Dissertation Committee:
Professor Daqing Wan, Chair
Professor Alice Silverberg
Professor Nathan Kaplan

2019

TABLE OF CONTENTS

ACKNOWLEDGMENTS

First, I would like to express my deepest gratitude to my advisor, Professor Daqing Wan, for his guidance and support over the years. His suggestions and advice have been invaluable.

Next, I would like to thank my committee members, Professor Alice Silverberg and Professor Nathan Kaplan, for their help throughout this process. Their comments and feedback are greatly appreciated.

Lastly, I would like to thank my parents and my sister. I would not be here without them.

CURRICULUM VITAE

Jennifer Nguyen

EDUCATION

Ph.D. in Mathematics **2019**
University of California, Irvine

M.S. in Mathematics **2015**
University of California, Irvine

B.S. in Mathematics **2012**
University of California, San Diego

ABSTRACT OF THE DISSERTATION

Higher Moments Subset Sum Problem over Finite Fields

By

Jennifer Nguyen

Doctor of Philosophy in Mathematics

University of California, Irvine, 2019

Professor Daqing Wan, Chair

Let $\mathbb{F}_q$ be a finite field and let $D \subseteq \mathbb{F}_q$. Let m be a positive integer and let k be an integer such that $1 \leq k \leq |D|$. For $b = (b_1, \ldots, b_m) \in \mathbb{F}_q^m$, let $N_m(k, b)$ denote the number of subsets $S \subseteq D$ with cardinality k such that for $i = 1, \ldots, m$, $\sum_{a \in S} a^i = b_i$. The Moments Subset Sum Problem is to determine if $N_m(k, b) > 0$. There are many results for when $m = 1$, but not much is known about the higher moments. In this dissertation, we obtain a formula for $N_m(k, b)$ when $m = 2$ and conditions on the solvability of the Moments Subset Sum Problem by using the Li-Wan sieve and properties of character sums and Gauss sums.

Chapter 1

Introduction

1.1 The Moments Subset Sum Problem

Let $\mathbb{F}_q$ be a finite field with cardinality q and characteristic p and let $D \subseteq \mathbb{F}_q$. Let m be a positive integer and let k be an integer such that $1 \leq k \leq |D|$. For $b = (b_1, \ldots, b_m) \in \mathbb{F}_q^m$, let $N_m(k, b)$ denote the number of subsets $S \subseteq D$ with cardinality k such that for $i = 1, \ldots, m$,

$$\sum_{a \in S} a^i = b_i.$$

Understanding the number $N_m(k, b)$ is the Moments Subset Sum Problem.

Definition 1.1. ***(Moments Subset Sum Problem [6])*** *Determine if* $N_m(k, b) > 0$.

Let j be an integer such that $1 \leq j \leq \lfloor \frac{m}{p} \rfloor$. Then, $1 \leq pj \leq m$ and

$$\begin{aligned} b_{pj} &= \sum_{a \in S} a^{pj} \\ &= \left(\sum_{a \in S} a^j \right)^p \\ &= (b_j)^p. \end{aligned}$$

Therefore, we may assume, without loss of generality, that for all $1 \leq j \leq \lfloor \frac{m}{p} \rfloor$, $b_{pj} = b_j^p$ and focus on $i = 1, \ldots, m$ such that $p \nmid i$.

We may also assume, without loss of generality, that $k \leq \frac{|D|}{2}$ because of the symmetry

$$N_m(k, b) = N_m \left(|D| - k, \left(\left(\sum_{a \in D} a \right) - b_1, \ldots, \left(\sum_{a \in D} a^m \right) - b_m \right) \right).$$

If $m = 1$, this problem becomes the decision version of the k-Subset Sum Problem and it has been shown that, for general D, this is NP-complete [5]. This version arises in several applications in many different fields. For example, in cryptography, Merkle and Hellman [17] presented a public key cryptosystem based on a variation of the k-Subset Sum Problem. It was one of the earliest public key cryptosystems, though it has since been broken [18].

If $m = 2$ or $m = 3$, then Gandikota, Ghazi, and Grigorescu [6] proved that the Moments Subset Sum Problem is NP-hard. They also proved [7] that there exists $c > 0$ such that if $1 \leq m \leq c \frac{\log n}{\log \log n}$, then the Moments Subset Sum Problem is NP-hard for prime fields of size $2^{polynomial(n)}$. The higher moments of this problem can be found in coding theory, where solving the Moments Subset Sum Problem helps to answer the Deep Hole Problem and to decode received words under certain conditions [9, 10, 13, 14, 21].

The main difficulty of this problem comes from the subset D. Since there are no restrictions on the choice of D, D might lack any algebraic structure. If D is a special subset of $\mathbb{F}_q$, then it is possible to obtain an exact value or an asymptotic formula for $N_m(k,b)$. In Chapter 2, we will review previous work on the Moments Subset Sum Problem for special subsets D, as well as state our main results. Next, in Chapter 3, we will introduce the tools we will use in our proofs that are in Chapter 4. Lastly, we will apply our results to the Deep Hole Problem in Chapter 5.

1.2 Notation

In order to clarify some notations that we will be using in this dissertation, we will define them here.

Definition 1.2. *A permutation τ in the symmetric group S_k is of cycle type $(c_1, \cdots, c_k)$ if τ has exactly c_i cycles of length i.*

Let $N(c_1, \cdots, c_k)$ be the number of permutations S_k of cycle type $(c_1, \cdots, c_k)$. Then

$$N(c_1, \cdots, c_k) = \frac{k!}{1^{c_1}c_1!2^{c_2}c_2! \cdots k^{c_k}c_k!}.$$

Definition 1.3. *Let η be the quadratic character of $\mathbb{F}_q$ and let $\psi_1 = e^{\frac{2\pi i}{p}Tr_{\mathbb{F}_q/\mathbb{F}_p}(x)}$. The character ψ_1 is called the canonical additive character of $\mathbb{F}_q$.*

Definition 1.4. *Let k be a positive integer. Then, for any number x,*

$$(x)_k \coloneqq x(x-1)\dots(x-k+1).$$

Chapter 2

Previous Results

2.1 For $m = 1$

If $D = \mathbb{F}_q$ or $= \mathbb{F}_q^*$, Li and Wan [10] obtained explicit formulas for $N_1(k, b)$.

Theorem 2.1. *(Li, Wan [10])*

(1) *When $D = \mathbb{F}_q$, if $p \nmid k$, then for all $b \in \mathbb{F}_q$,*

$$N_1(k,b) = \frac{1}{q}\binom{q}{k}.$$

If $p \mid k$ and $b = 0$, then

$$N_1(k,0) = \frac{1}{q}\binom{q}{k} + (-1)^{k+\frac{k}{p}}\left(\frac{q-1}{q}\right)\binom{q/p}{k/p}.$$

If $p \mid k$ and $b \neq 0$, then

$$N_1(k,b) = \frac{1}{q}\binom{q}{k} + (-1)^{k+\frac{k}{p}}\left(\frac{-1}{q}\right)\binom{q/p}{k/p}.$$

(2) When $D = \mathbb{F}_q^*$, if $b = 0$, then

$$N_1(k, 0) = \frac{1}{q}\binom{q-1}{k} + (-1)^{k+\lfloor \frac{k}{p} \rfloor}\left(\frac{q-1}{q}\right)\binom{q/p-1}{\lfloor k/p \rfloor}.$$

If $b \neq 0$, then

$$N_1(k, b) = \frac{1}{q}\binom{q-1}{k} + (-1)^{k+\lfloor \frac{k}{p} \rfloor}\left(\frac{-1}{q}\right)\binom{q/p-1}{\lfloor k/p \rfloor}.$$

To solve the decision version of the k-Subset Sum Problem, we only need conditions on p, k, and b to determine when $N_1(k, b) > 0$. In addition to simplifying Theorem 2.1, Li and Wan [10] were able to find good asymptotic formulas for when $\mathbb{F}_q - D$ is small.

Theorem 2.2. *(Li, Wan [10])*

(1) Let $D = \mathbb{F}_q$. If $p > 2$, then for $0 < k < q$, $N_1(k, b) > 0$ for all $b \in \mathbb{F}_q$. If $p = 2$, then for $2 < k < q - 2$, $N_1(k, b) > 0$ for all $b \in \mathbb{F}_q$.

(2) Let $|D| = q - 1 > 4$. If $p > 2$, then for $1 < k < q - 2$, $N_1(k, b) > 0$ for all $b \in \mathbb{F}_q$. If $p = 2$, then for $2 < k < q - 3$, $N_1(k, b) > 0$ for all $b \in \mathbb{F}_q$.

(3) Let $q > p$ and $c \geq 2$. Let $|D| = q - c$. If

$$\frac{q-c}{2}(1-w) \leq k \leq \frac{q-c}{2}(1+w),$$

then $N_D(k, b) > 0$ for all $b \in \mathbb{F}_q$, where $0 < w \leq 1$ is an explicit constant.

(4) Let $q = p$. If $|D| \geq k + \frac{p-1}{k}$, then $N_1(k, b) > 0$ for all $b \in \mathbb{F}_p$.

Another interesting choice of D would be when D is a subgroup of $\mathbb{F}_q$. If D is a multiplicative subgroup of $\mathbb{F}_q^*$ of index d, the k-Subset Sum Problem becomes much harder because it is nonlinear. When D is a subgroup of index 2, Wang, Wang, and Zhou [22] were able to find an explicit formula for $N_1(k,b)$.

Theorem 2.3. *(Wang, Wang, and Zhou [22]) Let $b \in \mathbb{F}_q^*$.*

Define

$$A_{k,b}(u,v,w) := \sum_{\substack{0\le c_i\le k\\ \sum ic_i=k}} N(c_1,\ldots,c_k)t_1^{c_1}\ldots t_k^{c_k},$$

where

$$t_i = \begin{cases} u, & \text{if } p\nmid i, \eta(i)=\eta(b)\\ v, & \text{if } p\nmid i, \eta(i)=-\eta(b)\\ w, & \text{if } p\mid i \end{cases}.$$

Let t be the integer such that $q=p^t$ and let $D=\{x^2 \mid x\in\mathbb{F}_q^*\}$.

(1) If either $p\equiv 1 \pmod 4$ or $p\equiv 3 \pmod 4$ and t is even, then

$$N_1(k,0) = \frac{1}{q}\binom{\frac{q-1}{2}}{k} - (-1)^k\frac{q-1}{2qk!}\Bigg[A_{k,1}\left(\frac{1-\sqrt{q}}{2},\frac{1+\sqrt{q}}{2},\frac{1-q}{2}\right) + A_{k,1}\left(\frac{1+\sqrt{q}}{2},\frac{1-\sqrt{q}}{2},\frac{1-q}{2}\right)\Bigg]$$

and

$$N_1(k,b) = \frac{1}{q}\binom{\frac{q-1}{2}}{k} - \frac{(-1)^k}{2qk!}\Bigg[(1-\sqrt{q})A_{k,b}\left(\frac{1-\sqrt{q}}{2},\frac{1+\sqrt{q}}{2},\frac{1-q}{2}\right) + (1+\sqrt{q})A_{k,b}\left(\frac{1+\sqrt{q}}{2},\frac{1-\sqrt{q}}{2},\frac{1-q}{2}\right)\Bigg]$$

for $b\in\mathbb{F}_q^*$.

(2) If $p \equiv 3 \pmod 4$ *and* t *is odd, then*

$$N_1(k,0) = \frac{1}{q}\binom{\frac{q-1}{2}}{k} - (-1)^k \frac{q-1}{2qk!}\left[A_{k,1}\left(\frac{1-\sqrt{q}i}{2}, \frac{1+\sqrt{q}i}{2}, \frac{1-q}{2}\right) + A_{k,1}\left(\frac{1+\sqrt{q}i}{2}, \frac{1-\sqrt{q}i}{2}, \frac{1-q}{2}\right)\right]$$

and

$$N_1(k,b) = \frac{1}{q}\binom{\frac{q-1}{2}}{k} - \frac{(-1)^k}{2qk!}\left[(1+\sqrt{q}i)A_{k,b}\left(\frac{1-\sqrt{q}i}{2}, \frac{1+\sqrt{q}i}{2}, \frac{1-q}{2}\right) + (1-\sqrt{q}i)A_{k,b}\left(\frac{1+\sqrt{q}i}{2}, \frac{1-\sqrt{q}i}{2}, \frac{1-q}{2}\right)\right]$$

for $b \in \mathbb{F}_q^*$*, where* $i = \sqrt{-1}$.

For general d, Zhu and Wan [23] provided an asymptotic formula.

Theorem 2.4. *(Zhu, Wan [23]) Let* D *be a multiplicative subgroup of* $\mathbb{F}_q^*$ *with index* d*. Let* $p > 2$*. There is an effectively computable absolute constant* $0 < c < 1$ *such that if* $d < c\sqrt{q}$ *and* $6 \ln q < k \leq \frac{q-1}{2d} = \frac{|D|}{2}$*, then* $N_1(k,b) > 0$ *for all* $b \in \mathbb{F}_q$*.*

Moving further away from an algebraic structure, Keti and Wan [9] studied the case when D is the image of a Dickson polynomial of degree d,

$$D_d(x,a) = \left(\frac{x+\sqrt{x^2+4a}}{2}\right)^d + \left(\frac{x-\sqrt{x^2+4a}}{2}\right)^d,$$

where $a \in \mathbb{F}_q$. Dickson polynomials are like generalized monomials because when $a = 0$, then $D_d(x,0) = x^d$.

Theorem 2.5. *(Keti, Wan [9]) Let* $D = \{D_d(x,a) \mid x \in \mathbb{F}_q\}$ *for* $a \in \mathbb{F}_q^*$*. There exist* *computable constants* $c_1, c_2 > 0$ *such that if the conditions*

$$\frac{d+1}{2}\sqrt{q} < c_1|D| \text{ and } \log_2 q \leq k < c_2|D|$$

are satisfied, then $N_1(k,b) > 0$ *for all* $b \in \mathbb{F}_q$.

If D is a more general subset of $\mathbb{F}_q$, then Wang and Nguyen [21] were able to answer the k-Subset Sum Problem, relying on character sums over D.

Theorem 2.6. *(Wang, Nguyen [21]) Let* $\mathbb{F}_q$ *be the finite field, where* p *is an odd prime. Let* $D \subseteq \mathbb{F}_q$. *If* $q \geq 227584$, $|D| \geq 36\ln^2 q$, *and for all nontrivial additive characters* $\psi : (\mathbb{F}_q, +) \to \mathbb{C}^*$,

$$\left|\sum_{x\in D}\psi(x)\right| \leq \frac{1}{\sqrt[3]{2q}}|D|,$$

then $N_1(k,b) > 0$ *for all* $b \in \mathbb{F}_q$ *and* $3 \leq k \leq \frac{|D|}{2}$.

Using Theorem 2.6, Wang and Nguyen [21] were able to conclude the following conditions for when D is the multiplicative subgroup of $\mathbb{F}_q^*$ with index d and when D is the image of a Dickson polynomial of degree d.

Corollary 2.1. *(Wang, Nguyen [21]) Let* p *be an odd prime and* $D = \{x^d \mid x \in \mathbb{F}_q^*\}$. *If* $d < 0.8\sqrt[6]{q}$, *then for all* $b \in \mathbb{F}_q$ *and* $3 \leq k \leq \dfrac{|D|}{2}$, $N_1(k,b) > 0$.

Corollary 2.2. *(Wang, Nguyen [21]) Let* p *is an odd prime,* $a \in \mathbb{F}_q^*$ *and* $D = \{D_d(x,a) | x \in \mathbb{F}_q\}$. *If*

$$q\left[\frac{1}{\gcd(d,q-1)} + \frac{1}{\gcd(d,q+1)}\right] \geq 72\ln^2 q + 1 \text{ and}$$
$$d+1 \leq 0.39 \cdot \sqrt[6]{q}\left[\frac{1}{\gcd(d,q-1)} + \frac{1}{\gcd(d,q+1)}\right],$$

then for all $b \in \mathbb{F}_q$ *and* $3 \leq k \leq \dfrac{|D|}{2}$, $N_1(k,b) > 0$.

One of the most important consequences of Corollaries 2.1 and 2.2 is the following theorem.

Theorem 2.7. *(Wang, Nguyen [21]) If $p > 2$ and D is the image of the monomial or a Dickson polynomial of degree d, then the k-Subset Sum problem can be solved in deterministic polynomial time in $d \log q$.*

Wang and Nguyen focused on the case when $p > 2$. If $p = 2$, Choe and Choe [4] found a similar theorem.

Theorem 2.8. *(Choe, Choe [4]) Let $q = 2^t$, where $t \geq 11$ and let $D \subseteq \mathbb{F}_q$ such that $|D| > \max\{5q^{\frac{2}{3}}, (3.05t)^2\}$. If*

$$\left|\sum_{x \in D} \psi(x)\right| \leq \frac{1}{\sqrt[3]{2q}}|D|$$

for all nontrivial additive characters ψ of $\mathbb{F}_q$, then $N_1(k,b) > 0$ whenever $b \in \mathbb{F}_q$ and $3 \leq k \leq \frac{|D|}{2}$.

Choe and Choe [4] also applied their theorem for when D is the multiplicative subgroup of $\mathbb{F}_q^*$ with index d and when D is the image of a Dickson polynomial of degree d.

Corollary 2.3. *(Choe, Choe [4]) Let $q = 2^t$, where $t \geq 13$ and let D be the subgroup of $\mathbb{F}_q^*$ with index d. If $d \leq \frac{1}{\sqrt[3]{2}}\sqrt[6]{q}$, then $N_1(k,b) > 0$ whenever $b \in \mathbb{F}_q$ and $3 \leq k \leq \frac{|D|}{2}$.*

Corollary 2.4. *(Choe, Choe [4]) Let $q = 2^t$, where $t \geq 11$, $a \in \mathbb{F}_q^*$ and $D = \{D_d(x,a) | x \in \mathbb{F}_q\}$. If $d \leq \frac{1}{\sqrt[3]{16}}\sqrt[6]{q}$, then $N_1(k,b) > 0$ whenever $b \in \mathbb{F}_q$ and $3 \leq k \leq \frac{|D|}{2}$.*

Similar to the case of $p > 2$, Corollaries 2.3 and 2.4 imply the following theorem.

Theorem 2.9. *(Choe, Choe [4]) If $p = 2$ and D is the image of the monomial or a Dickson polynomial of degree d, then the k-Subset Sum problem can be solved in deterministic polynomial time in $d \log q$.*

2.2 For general m

The k-Subset Sum Problem has been studied extensively, but not much is known about $N_m(k,b)$ for $m > 1$. The work from Li and Wan [11] implies the following asymptotic formula for general m when $D = \mathbb{F}_q$.

Theorem 2.10. *Let $D = \mathbb{F}_q$. For any $\epsilon > 0$, there is a constant $c_\epsilon > 0$ s.t. if $m < \epsilon k^{1/2}$ and $4\epsilon^2 \ln^2 q < k \leq c_\epsilon q$, then $N_m(k,b) > 0$ for all $b \in \mathbb{F}_q^m$.*

2.3 New results for $m = 2$

In this dissertation, we obtain a formula for $N_2(k,(0,0))$ when $D = \mathbb{F}_q$.

Theorem 2.11. *Let $D = \mathbb{F}_q$, where p is an odd prime, and let t be the integer such that $q = p^t$.*

For $0 \leq c_i \leq k$, $i = 1, \ldots, k$, let $s = \sum_{p \nmid i} c_i$ and $r = \sum_{p \mid i} c_i$.

If $p \equiv 1 \pmod 4$, then

$$N_2(k,(0,0)) = \frac{1}{k!q^2}\Bigg[(q)_k + (q-1)\sum_{\substack{0\leq c_i \leq k \\ \sum ic_i = k \\ s=0}} (-1)^{k-\sum c_i} N(c_1,\cdots,c_k) q^r$$

$$+ q(q-1)\sum_{\substack{0\leq c_i \leq k \\ \sum ic_i = k \\ s\equiv 0 (mod\ 2p)}} (-1)^{k-\sum c_i} N(c_1,\cdots,c_k)\eta\left(\prod_{i=1}^{k} i^{c_i}\right) q^{\frac{2r+s}{2}}$$

$$+ (-1)^{t-1}(q-1)\sqrt{q}\sum_{\substack{0\leq c_i \leq k \\ \sum ic_i = k \\ s\not\equiv 0 (mod\ p) \\ s\equiv 0 (mod\ 2)}} (-1)^{k-\sum c_i} N(c_1,\cdots,c_k)\eta\left(-s\prod_{i=1}^{k} i^{c_i}\right) q^{\frac{2r+s}{2}}\Bigg]$$

and if $p \equiv 3 \pmod 4$, *then*

$$N_2(k,(0,0)) = \frac{1}{k!q^2}\Bigg[(q)_k + (q-1)\sum_{\substack{0\le c_i\le k\\ \sum ic_i=k\\ s=0}} (-1)^{k-\sum c_i} N(c_1,\cdots,c_k)q^r$$

$$+\, q(q-1)\sum_{\substack{0\le c_i\le k\\ \sum ic_i=k\\ s\equiv 0(mod\ 2p)}} (-1)^{k-\sum c_i+\frac{st}{2}} N(c_1,\cdots,c_k)\eta\left(\prod_{i=1}^{k} i^{c_i}\right) q^{\frac{2r+s}{2}}$$

$$+\, (-1)^{\frac{3t}{2}-1}(q-1)\sqrt{q}\sum_{\substack{0\le c_i\le k\\ \sum ic_i=k\\ s\not\equiv 0(mod\ p)\\ s\equiv 0(mod\ 2)}} (-1)^{k-\sum c_i+\frac{st}{2}} N(c_1,\cdots,c_k)\eta\left(-s\prod_{i=1}^{k} i^{c_i}\right) q^{\frac{2r+s}{2}}\Bigg].$$

In order to answer the Moments Subset Sum Problem under certain conditions, we have the following corollary.

Corollary 2.5. *Let* $q = p$, *where* p *is an odd prime, and let* $D = \mathbb{F}_q$. *For a positive constant* c *such that* $0 < c < \frac{1}{2}$, *if* $\frac{-2\log(q)}{\log(2c)} \le k \le c\sqrt{q}$, *then* $N_2(k,(0,0)) > 0$.

Note that since $0 < c < \frac{1}{2}$, $\log(2c) < 0$ and therefore, $\frac{-2\log(q)}{\log(2c)} > 0$. Also, note that the lower bound for k in Corollary 2.5 is an improvement on the lower bound for k in Theorem 2.10.

We prove Theorem 2.11 in Section 4.2 and Corollary 2.5 in Section 4.3 by utilizing a sieve by Li and Wan [11], properties of characters sums, and the Gauss sum over $\mathbb{F}_q$.

Chapter 3

Tools

3.1 Li-Wan Sieve

To solve the Moments Subset Sum Problem, we need to count vectors with distinct coordinates.

Let D be a finite set. Let $D^k = D \times D \times \cdots \times D$ $(k \in \mathbb{N}^+)$ be the Cartesian product of k copies of D and let X be a subset of D^k. We are interested in the number of elements in X with distinct coordinates, i.e., the cardinality of the set

$$\overline{X} = \{(x_1, \cdots, x_k) \in X \mid x_i \neq x_j \text{ for } \forall\ i \neq j\}.$$

Let $X_{ij} = \{(x_1, \ldots, x_k) \in X \mid x_i = x_j\}$. Then, by the Inclusion-Exclusion Principle,

$$|\overline{X}| = |X| - \sum_{1 \leq i < j \leq k} |X_{ij}| + \sum_{\substack{1 \leq i < j \leq k \\ 1 \leq s < t \leq k}} |X_{ij} \bigcap X_{st}| - \ldots + (-1)^{\binom{k}{2}} \left| \bigcap_{1 \leq i < j \leq k} X_{ij} \right|.$$

This equation has $2^{\binom{k}{2}}$ terms. When k is relatively large, the total error term may be greater than the main term. To avoid this, another method to counting these vectors is to use a sieve proposed by Li and Wan [11], which we will introduce here.

Let S_k be the symmetric group on $\{1, \dots, k\}$. For a given permutation $\tau \in S_k$, we can write it as the product of disjoint cycles, i.e., $\tau = (i_1, \cdots, i_{a_1})(j_1, \cdots, j_{a_2}) \cdots (l_1, \cdots, l_{a_s})$, where $a_i \geq 1, 1 \leq i \leq s$. The group S_k acts on D^k by permuting its coordinates, that is

$$\tau \circ (x_1, \cdots, x_k) = (x_{\tau(1)}, \cdots, x_{\tau(k)}).$$

Definition 3.1. *The set X is called symmetric if it is invariant under the action of S_k, i.e., for any $x \in X$ and any $\tau \in S_k$, $\tau \circ x \in X$.*

Let $f(x_1, x_2, \cdots, x_k)$ be a complex valued function defined over X, and denote

$$F = \sum_{x \in \overline{X}} f(x_1, x_2, \cdots, x_k).$$

In order to illustrate the sieve, we define for $\tau = (i_1, \cdots, i_{a_1})(j_1, \cdots, j_{a_2}) \cdots (l_1, \cdots, l_{a_s})$,

$$X_\tau = \{(x_1, \cdots, x_k) \in X \mid x_{i_1} = \cdots = x_{ia_1}, \cdots, x_{l_1} = \cdots = x_{l_{a_s}}\}.$$

Similarly, we can define

$$F_\tau = \sum_{x \in X_\tau} f(x_1, x_2, \cdots, x_k).$$

Definition 3.2. *A complex-valued function f defined on X is called normal on X if X is symmetric, and for any two conjugate elements τ and τ' in S_k, we have*

$$\sum_{x \in X_\tau} f(x_1, x_2, \cdots, x_k) = \sum_{x \in X_{\tau'}} f(x_1, x_2, \cdots, x_k).$$

Theorem 3.1. *(Li, Wan [11]) If f is normal on X, then we have*

$$F = \sum_{\substack{0 \le c_i \le k \\ \sum i c_i = k}} (-1)^{k - \sum c_i} N(c_1, \cdots, c_k) F_\tau,$$

where $\tau \in S_k$ is of cycle type $(c_1, \cdots, c_k)$.

3.2 Gauss Sums

In our proof, we will also need a few properties about Gauss sums.

Definition 3.3. *Let χ be a multiplicative character of $\mathbb{F}_q$ and let ψ be an additive character of $\mathbb{F}_q$. Suppose that we extend χ to the whole field $\mathbb{F}_q$ by defining*

$$\chi(0) = \begin{cases} 1, & \text{if } \chi \text{ is the trivial character} \\ 0, & \text{otherwise} \end{cases}.$$

The Gauss sum $G(\chi, \psi)$ is defined by

$$G(\chi, \psi) = \sum_{x \in \mathbb{F}_q} \chi(x)\psi(x).$$

Theorem 3.2. *[16] Let t be the integer such that $q = p^t$. Then,*

$$G(\eta, \psi_1) = \begin{cases} (-1)^{t-1}\sqrt{q}, & \text{if } p \equiv 1 \pmod 4 \\ (-1)^{\frac{3t}{2}-1}\sqrt{q}, & \text{if } p \equiv 3 \pmod 4 \end{cases}.$$

3.3 Some Combinatorial Formulas

We will also need another combinatorial tool in our proof.

Definition 3.4. *Let* $C_k(t_1, \dots, t_k)$ *be the generating function*

$$C_k(t_1, \dots, t_k) := \sum_{\sum ic_i = k} N(c_1, \dots, c_k) t_1^{c_1} \dots t_k^{c_k}.$$

Lemma 3.1. *(Li, Wan [12]) Let* a, b *be two nonnegative real numbers such that* $b \geq a$ *and let be* p *be a prime number. If* $t_i = a$ *for* $p \nmid i$, $t_i = b$ *for* $p \mid i$, *then*

$$\begin{aligned} C_k(t_1, \dots, t_k) &= C_k(\overbrace{a, \cdots, a}^{p-1}, b, \overbrace{a, \cdots, a}^{p-1}, b, \dots) \\ &\leq \left(a + k + \frac{b-a}{p} - 1\right)_k. \end{aligned}$$

In the special case when $a = 0$ and $b = q$, we have the exact value of the generating function.

Lemma 3.2.

$$C_k(\overbrace{0, \cdots, 0}^{p-1}, q, \overbrace{0, \cdots, 0}^{p-1}, q, \dots) = \begin{cases} k!\binom{\frac{q+k}{p} - 1}{\frac{k}{p}}, & \text{if } p \mid k \\ 0, & \text{if } p \nmid k \end{cases}.$$

Proof. By definition of the generating function,

$$\begin{aligned} C_k(t_1, \dots, t_k) &= \sum_{\sum ic_i = k} \frac{k!}{1^{c_1} c_1! 2^{c_2} c_2! \cdots k^{c_k} c_k!} t_1^{c_1} \dots t_k^{c_k} \\ &= \sum_{\sum ic_i = k} \frac{k!}{c_1! c_2! \cdots c_k!} \left(\frac{t_1}{1}\right)^{c_1} \left(\frac{t_2}{2}\right)^{c_2} \cdots \left(\frac{t_k}{k}\right)^{c_k} \end{aligned}$$

Therefore, we have the following exponential generating function,

$$\sum_{k\geq 0} C_k(t_1,\ldots,t_k)\frac{u^k}{k!} = e^{ut_1+u^2\cdot\frac{t_2}{2}+u^3\cdot\frac{t_3}{3}+\cdots}.$$

If $t_i = 0$ for $p \nmid i$, $t_i = q$ for $p \mid i$, then we have

$$\begin{aligned}
\sum_{k\geq 0} C_k(\overbrace{0,\cdots,0}^{p-1},q,\overbrace{0,\cdots,0}^{p-1},q,\ldots)\frac{u^k}{k!} &= e^{u^p\cdot\frac{q}{p}+u^{2p}\cdot\frac{q}{2p}+u^{3p}\cdot\frac{q}{3p}+\cdots}\\
&= e^{\frac{q}{p}\left(u^p+\frac{u^{2p}}{2}+\frac{u^{3p}}{3}+\cdots\right)}\\
&= e^{-\frac{q}{p}\log(1-u^p)}\\
&= \frac{1}{(1-u^p)^{\frac{q}{p}}}\\
&= \sum_{i\geq 0}\binom{\frac{q}{p}+i-1}{i}u^{pi}.
\end{aligned}$$

Thus, $C_k(\overbrace{0,\cdots,0}^{p-1},q,\overbrace{0,\cdots,0}^{p-1},q,\ldots)$ is the coefficient of the term $\frac{u^k}{k!}$ in the sum $\sum_{i\geq 0}\binom{\frac{q}{p}+i-1}{i}u^{pi}$.

If $p \mid k$, then the term u^k appears when $i = \frac{k}{p}$. Thus,

$$C_k(\overbrace{0,\cdots,0}^{p-1},q,\overbrace{0,\cdots,0}^{p-1},q,\ldots) = k!\binom{\frac{q+k}{p}-1}{\frac{k}{p}}.$$

If $p \nmid k$, then the term u^k does not appear in the sum and

$$C_k(\overbrace{0,\cdots,0}^{p-1},q,\overbrace{0,\cdots,0}^{p-1},q,\ldots) = 0.$$

□

To combine Lemma 3.2 into one estimate, we have the following corollary.

Corollary 3.1.

$$C_k(\overbrace{0,\cdots,0}^{p-1},q,\overbrace{0,\cdots,0}^{p-1},q,\ldots) \le k!\binom{\frac{q+k}{p}-1}{\frac{k}{p}}$$

Chapter 4

Approaching the case of $m = 2$

4.1 Redefine the problem

Let $D = \mathbb{F}_q$, where p is an odd prime, and $m = 2$. Let $N_2(k,(0,0))$ be the number of unordered k-tuples $x = (x_1, \ldots, x_k)$ with distinct $x_i \in \mathbb{F}_q$ such that

$$\begin{cases} x_1^2 + \ldots + x_k^2 = 0 \\ x_1 + \ldots + x_k = 0 \end{cases} \tag{4.1}$$

Let $\widetilde{N}_2(k,(0,0))$ be the number of unordered k-tuples $x = (x_1, \ldots, x_k)$ with distinct $x_i \in \mathbb{F}_q$ such that

$$\begin{cases} \sum_{1 \le i < j \le k} x_i x_j = 0 \\ x_1 + \ldots + x_k = 0 \end{cases} \tag{4.2}$$

Lemma 4.1. $N_2(k,(0,0)) = \widetilde{N}_2(k,(0,0))$.

Proof. By squaring $(x_1 + \ldots + x_k)$ and rearranging the terms, we have

$$x_1^2 + \ldots + x_k^2 = (x_1 + \ldots + x_k)^2 - 2 \sum_{1 \le i < j \le k} x_i x_j.$$

Then, $(x_1, \ldots, x_k)$ is a solution to the system of equations (4.1) if and only if it is a solution to the system of equations (4.2). □

Let $x = (x_1, \ldots, x_k)$ be a solution to the system of equations (4.2). Then, for some $g(y) \in y^3\mathbb{F}_q[y]$,

$$\begin{aligned}\prod_{i=1}^{k}(1 + x_i y) &= 1 + (x_1 + \ldots + x_k)y + \left(\sum_{1 \le i < j \le k} x_i x_j\right) y^2 + g(y) \\ &= 1 + g(y) \\ &\equiv 1 \pmod{y^3}.\end{aligned}$$

Therefore, we can redefine $N_2(k, (0,0))$ as the following.

Definition 4.1. *The number $N_2(k, (0,0))$ is the number of unordered k-tuples $x = (x_1, \ldots, x_k)$ with distinct $x_i \in \mathbb{F}_q$ such that*

$$\prod_{i=1}^{k}(1 + x_i y) \equiv 1 \pmod{y^3},$$

i.e.,

$$N_2(k, (0,0)) = \left| \left\{ \{x_1, \ldots, x_k\} \subseteq \mathbb{F}_q \mid \prod_{i=1}(1 + x_i y) \equiv 1 \pmod{y^3}, x_i \neq x_j, \text{ for } i \neq j \right\} \right|.$$

Let $G = ((1 + yF_q[y])/(1 + y^3\mathbb{F}_q[y]))^*$. Then, all the multiplicative characters χ of G are given such that for $a \in F_q$, $\chi(1 + ay) = \psi_1(\sigma_2 a^2 + \sigma_1 a)$, where $\sigma_1, \sigma_2 \in \mathbb{F}_q$ [20].

The q^2 characters of G are parametrized precisely by the q^2 pairs $(\sigma_1, \sigma_2) \in F_q^2$. For each $\sigma = (\sigma_1, \sigma_2) \in F_q^2$, let $\chi_\sigma(1 + ay) = \psi_1(\sigma_2 a^2 + \sigma_1 a)$. Note that if $\sigma = (0, 0)$, then χ_σ has order p.

Lemma 4.2. *Let χ_σ be a multiplicative character of G. Then,*

$$\sum_{a \in \mathbb{F}_q} \chi_\sigma(1 + ay) = \begin{cases} q, & \text{if } \sigma_1 = 0, \sigma_2 = 0 \\ 0, & \text{if } \sigma_1 \neq 0, \sigma_2 = 0 \\ \psi_1\left(\dfrac{-\sigma_1^2}{4\sigma_2^2}\right)\eta(\sigma_2)G(\eta, \psi_1), & \text{if } \sigma_2 \neq 0 \end{cases} \cdot$$

Proof. (1) If $\sigma_1 = \sigma_2 = 0$, then for all $a \in \mathbb{F}_q$, $\chi_\sigma(1 + ay) = 1$ and

$$\sum_{a \in \mathbb{F}_q} \chi_\sigma(1 + ay) = q.$$

(2) If $\sigma_1 \neq 0, \sigma_2 = 0$, then

$$\begin{aligned} \sum_{a \in \mathbb{F}_q} \chi_\sigma(1 + ay) &= \sum_{a \in \mathbb{F}_q} \psi_1(\sigma_1 a) \\ &= \sum_{a \in \mathbb{F}_q} \psi_1(a) \\ &= 0. \end{aligned}$$

(3) If $\sigma_2 \neq 0$, then

$$\begin{aligned}
\sum_{a\in\mathbb{F}_q} \chi_\sigma(1+ay) &= \sum_{a\in\mathbb{F}_q} \psi_1(\sigma_2 a^2 + \sigma_1 a) \\
&= \sum_{a\in\mathbb{F}_q} \psi_1\left(\sigma_2\left(a+\frac{\sigma_1}{2\sigma_2}\right)^2 - \frac{\sigma_1^2}{4\sigma_2^2}\right) \\
&= \psi_1\left(\frac{-\sigma_1^2}{4\sigma_2^2}\right)\sum_{a\in\mathbb{F}_q}\psi_1(\sigma_2 a^2) \\
&= \psi_1\left(\frac{-\sigma_1^2}{4\sigma_2^2}\right)\sum_{x\in\mathbb{F}_q}\psi_1(\sigma_2 x)(1+\eta(x)) \\
&= \psi_1\left(\frac{-\sigma_1^2}{4\sigma_2^2}\right)\sum_{x\in\mathbb{F}_q}\psi_1(\sigma_2 x)\eta(x) \\
&= \psi_1\left(\frac{-\sigma_1^2}{4\sigma_2^2}\right)\eta(\sigma_2^{-1})\sum_{x\in\mathbb{F}_q}\psi_1(\sigma_2 x)\eta(\sigma_2 x) \\
&= \psi_1\left(\frac{-\sigma_1^2}{4\sigma_2^2}\right)\eta(\sigma_2)G(\eta,\psi_1).
\end{aligned}$$

□

4.2 Proof of Theorem 2.11

Instead of focusing on the number $N_2(k,(0,0))$, we will be looking at the following number.

Definition 4.2. *Let $M_2(k,(0,0))$ be the number of ordered k-tuples $x = (x_1,\dots,x_k)$ with distinct $x_i \in \mathbb{F}_q$ such that*

$$\prod_{i=1}^{k}(1+x_i y) \equiv 1 \pmod{y^3},$$

i.e.,

$$M_2(k,(0,0)) = \left|\left\{(x_1,\dots,x_k)\in\mathbb{F}_q^k \,\middle|\, \prod_{i=1}^{k}(1+x_i y)\equiv 1 \pmod{y^3}, x_i \neq x_j,\ \textit{for } i\neq j\right\}\right|.$$

Let $\widehat{G}$ be the group of multiplicative characters χ_σ of G. Based on the properties of character sums,

$$M_2(k,(0,0)) = \frac{1}{q^2} \sum_{\substack{x_i \in \mathbb{F}_q \\ x_i \text{ distinct}}} \sum_{\chi_\sigma \in \widehat{G}} \chi_\sigma \left(\prod_{i=1}^{k} (1 + x_i y) \right). \tag{4.3}$$

Let $X = \mathbb{F}_q^k$,

$$\overline{X} = \{(x_1, \ldots, x_k) \in X \mid x_i \neq x_j, \text{ for } i \neq j\},$$

and for $\tau = (i_1, \cdots, i_{a_1})(j_1, \cdots, j_{a_2}) \cdots (l_1, \cdots, l_{a_s}) \in S_k$,

$$X_\tau = \{(x_1, \cdots, x_k) \in X \mid x_{i_1} = \cdots = x_{ia_1}, \cdots, x_{l_1} = \cdots = x_{l_{a_s}}\}.$$

For $\chi_\sigma \in \widehat{G}$, define $f_{\chi_\sigma}(x) = f_{\chi_\sigma}(x_1, \ldots, x_k) = \chi_\sigma \left(\prod_{i=1}^{k} (1 + x_i y) \right)$ and define

$$F_{\chi_\sigma} = \sum_{\substack{x_i \in \mathbb{F}_q \\ x_i \text{ distinct}}} \chi_\sigma \left(\prod_{i=1}^{k} (1 + x_i y) \right) = \sum_{x \in \overline{X}} f_{\chi_\sigma}(x).$$

For $\tau \in S_k$, define

$$F_{\tau,\chi_\sigma} = \sum_{x \in X_\tau} \chi_\sigma \left(\prod_{i=1}^{k} (1 + x_i y) \right) = \sum_{x \in X_\tau} f_{\chi_\sigma}(x). \tag{4.4}$$

We can rewrite equation (4.3) as

$$q^2 M_2(k,(0,0)) = \sum_{\chi_\sigma \in \widehat{G}} \sum_{x \in \overline{X}} f_{\chi_\sigma}(x). \tag{4.5}$$

Recalling Definitions 3.1 and 3.2, the set X is symmetric and the function f_{χ_σ} is normal on X. Therefore, by applying Theorem 3.1 to equation (4.5), we have

$$
\begin{aligned}
q^2 M_2(k,(0,0)) &= \sum_{\chi_\sigma \in \widehat{G}} \sum_{\substack{0\le c_i \le k \\ \sum i c_i = k}} (-1)^{k-\sum c_i} N(c_1,\cdots,c_k) F_{\tau,\chi_\sigma} \\
&= (q)_k + \sum_{\substack{\chi_\sigma \in \widehat{G} \\ \chi_\sigma \neq 1}} \sum_{\substack{0\le c_i \le k \\ \sum i c_i = k}} (-1)^{k-\sum c_i} N(c_1,\cdots,c_k) F_{\tau,\chi_\sigma} \\
&= (q)_k + \sum_{\substack{\sigma_1 \neq 0 \\ \sigma_2 = 0}} \sum_{\substack{0\le c_i \le k \\ \sum i c_i = k}} (-1)^{k-\sum c_i} N(c_1,\cdots,c_k) F_{\tau,\chi_\sigma} \\
&\qquad + \sum_{\substack{\sigma_2 \neq 0 \\ \sigma_1 \in \mathbb{F}_q}} \sum_{\substack{0\le c_i \le k \\ \sum i c_i = k}} (-1)^{k-\sum c_i} N(c_1,\cdots,c_k) F_{\tau,\chi_\sigma} \\
&= (q)_k + \sum_{\substack{0\le c_i \le k \\ \sum i c_i = k}} (-1)^{k-\sum c_i} N(c_1,\cdots,c_k) \sum_{\substack{\sigma_1 \neq 0 \\ \sigma_2 = 0}} F_{\tau,\chi_\sigma} \\
&\qquad + \sum_{\substack{0\le c_i \le k \\ \sum i c_i = k}} (-1)^{k-\sum c_i} N(c_1,\cdots,c_k) \sum_{\substack{\sigma_2 \neq 0 \\ \sigma_1 \in \mathbb{F}_q}} F_{\tau,\chi_\sigma}. \qquad (4.6)
\end{aligned}
$$

If τ is of cycle type $(c_1,\cdots,c_k)$, then equation (4.4) becomes

$$
F_{\tau,\chi_\sigma} = \prod_{i=1}^{k} \left(\sum_{a\in\mathbb{F}_q} \chi_\sigma^i (1+ay) \right)^{c_i}.
$$

If $p \mid i$, then

$$
\left(\sum_{a\in\mathbb{F}_q} \chi_\sigma^i (1+ay) \right)^{c_i} = q^{c_i}.
$$

If $p \nmid i$ and $\sigma_2 \neq 0$, then by the third case of Lemma 4.2,

$$\begin{aligned}\left(\sum_{a\in\mathbb{F}_q}\chi_\sigma^i(1+ay)\right)^{c_i} &= \left(\sum_{a\in\mathbb{F}_q}\psi_1^i(\sigma_2 a^2+\sigma_1 a)\right)^{c_i}\\ &= \left(\sum_{a\in\mathbb{F}_q}\psi_1(i\sigma_2 a^2+i\sigma_1 a)\right)^{c_i}\\ &= \left(\psi_1\left(\frac{-\sigma_1^2}{4\sigma_2^2}\right)\eta(i\sigma_2)G(\eta,\psi_1)\right)^{c_i}\\ &= \eta(i^{c_i})\psi_1^{c_i}\left(\frac{-\sigma_1^2}{4\sigma_2^2}\right)\eta^{c_i}(\sigma_2)G^{c_1}(\eta,\psi_1).\end{aligned}$$

Therefore, if $\sigma_2 \neq 0$,

$$F_{\tau,\chi_\sigma} = \eta\left(\prod_{i=1}^{k} i^{c_i}\right) q^r \psi_1^s\left(\frac{-\sigma_1^2}{4\sigma_2^2}\right)\eta^s(\sigma_2)G^s(\eta,\psi_1). \tag{4.7}$$

If $p \nmid i$, $\sigma_1 \neq 0, \sigma_2 = 0$, and $c_i \neq 0$, then by the second case of Lemma 4.2,

$$\begin{aligned}\left(\sum_{a\in\mathbb{F}_q}\chi_\sigma^i(1+ay)\right)^{c_i} &= \left(\sum_{a\in\mathbb{F}_q}\psi_1^i(\sigma_1 a)\right)^{c_i}\\ &= \left(\sum_{a\in\mathbb{F}_q}\psi_1(i\sigma_1 a)\right)^{c_i}\\ &= 0.\end{aligned}$$

If $p \nmid i$, $\sigma_1 \neq 0, \sigma_2 = 0$, and $c_i = 0$, then

$$\left(\sum_{a\in\mathbb{F}_q}\chi_\sigma^i(1+ay)\right)^{c_i} = 1.$$

Let $s = \sum_{p\nmid i} c_i$ and let $r = \sum_{p\mid i} c_i$.

If $\sigma_1 \neq 0, \sigma_2 = 0$, and $s \neq 0$, then there is at least one i such that $p \nmid i$ and $c_i \neq 0$. Thus,

$$F_{\tau,\chi_\sigma} = 0. \tag{4.8}$$

If $s = 0$, then

$$F_{\tau,\chi_\sigma} = q^r. \tag{4.9}$$

Using the values (4.7), (4.8), and (4.9) for F_{τ,χ_σ}, equation (4.6) becomes

$$\begin{aligned}
q^2 M_2(k,(0,0)) &= (q)_k + (q-1) \sum_{\substack{0\le c_i \le k \\ \sum ic_i = k \\ s=0}} (-1)^{k-\sum c_i} N(c_1,\cdots,c_k) q^r \\
&\quad + \sum_{\substack{0\le c_i \le k \\ \sum ic_i = k}} (-1)^{k-\sum c_i} N(c_1,\cdots,c_k) \sum_{\substack{\sigma_2\neq 0 \\ \sigma_1 \in \mathbb{F}_q}} \eta\left(\prod_{i=1}^k i^{c_i}\right) q^r \psi_1^s\left(\frac{-\sigma_1^2}{4\sigma_2^2}\right)\eta^s(\sigma_2) G^s(\eta,\psi_1) \\
&= (q)_k + (q-1) \sum_{\substack{0\le c_i \le k \\ \sum ic_i = k \\ s=0}} (-1)^{k-\sum c_i} N(c_1,\cdots,c_k) q^r \\
&\quad + \sum_{\substack{0\le c_i \le k \\ \sum ic_i = k}} (-1)^{k-\sum c_i} N(c_1,\cdots,c_k) \eta\left(\prod_{i=1}^k i^{c_i}\right) q^r G^s(\eta,\psi_1) \sum_{\substack{\sigma_2\neq 0 \\ \sigma_1 \in \mathbb{F}_q}} \psi_1^s\left(\frac{-\sigma_1^2}{4\sigma_2^2}\right)\eta^s(\sigma_2).
\end{aligned} \tag{4.10}$$

If $s \equiv 0 \pmod{2p}$, then

$$\psi_1^s\left(\frac{-\sigma_1^2}{4\sigma_2^2}\right)\eta^s(\sigma_2) = 1$$

and

$$\sum_{\substack{\sigma_2\neq 0 \\ \sigma_1 \in \mathbb{F}_q}} \psi_1^s\left(\frac{-\sigma_1^2}{4\sigma_2^2}\right)\eta^s(\sigma_2) = q(q-1). \tag{4.11}$$

If $s \not\equiv 0 \pmod{p}$ and $s \equiv 0 \pmod{2}$, then similar to the proof of the third case of Lemma 4.2,

$$
\begin{aligned}
\sum_{\substack{\sigma_2\neq 0\\ \sigma_1\in\mathbb{F}_q}} \psi_1^s\left(\frac{-\sigma_1^2}{4\sigma_2^2}\right)\eta^s(\sigma_2) &= \sum_{\substack{\sigma_2\neq 0\\ \sigma_1\in\mathbb{F}_q}} \psi_1^s\left(\frac{-\sigma_1^2}{4\sigma_2^2}\right)\\
&= \sum_{\substack{\sigma_2\neq 0\\ \sigma_1\in\mathbb{F}_q}} \psi_1\left(\frac{-s\sigma_1^2}{4\sigma_2^2}\right)\\
&= \sum_{\substack{\sigma_2\neq 0\\ \sigma_1\in\mathbb{F}_q}} \psi_1\left(\frac{-s}{4}\sigma_1^2\right)\\
&= (q-1)\sum_{\sigma_1\in\mathbb{F}_q} \psi_1\left(\frac{-s}{4}\sigma_1^2\right)\\
&= (q-1)\eta\left(-\frac{s}{4}\right)G(\eta,\psi_1)\\
&= (q-1)\eta(-s)G(\eta,\psi_1). \qquad (4.12)
\end{aligned}
$$

If $s \not\equiv 0 \pmod 2$, then

$$
\begin{aligned}
\sum_{\substack{\sigma_2\neq 0\\ \sigma_1\in\mathbb{F}_q}} \psi_1^s\left(\frac{-\sigma_1^2}{4\sigma_2^2}\right)\eta^s(\sigma_2) &= \sum_{\substack{\sigma_2\neq 0\\ \sigma_1\in\mathbb{F}_q}} \psi_1\left(\frac{-s\sigma_1^2}{4\sigma_2^2}\right)\eta(\sigma_2)\\
&= \sum_{\sigma_2\neq 0}\eta(\sigma_2)\sum_{\sigma_1\in\mathbb{F}_q} \psi_1\left(\frac{-s\sigma_1^2}{4\sigma_2^2}\right)\\
&= \sum_{\sigma_2\neq 0}\eta(\sigma_2)\sum_{\sigma_1\in\mathbb{F}_q} \psi_1\left(\frac{-s}{4}\left(\frac{\sigma_1}{\sigma_2}\right)^2\right)\\
&= \left(\sum_{\sigma_2\neq 0}\eta(\sigma_2)\right)\left(\sum_{\sigma_1\in\mathbb{F}_q} \psi_1\left(\frac{-s}{4}\sigma_1^2\right)\right)\\
&= 0. \qquad (4.13)
\end{aligned}
$$

Using values (4.11), (4.12), (4.13), equation (4.10) becomes

$$q^2M_2(k,(0,0)) = (q)_k + (q-1)\sum_{\substack{0\le c_i\le k\\ \sum ic_i=k\\ s=0}} (-1)^{k-\sum c_i}N(c_1,\cdots,c_k)q^r$$
$$+ q(q-1)\sum_{\substack{0\le c_i\le k\\ \sum ic_i=k\\ s\equiv 0(mod\ 2p)}} (-1)^{k-\sum c_i}N(c_1,\cdots,c_k)\eta\left(\prod_{i=1}^{k} i^{c_i}\right)q^rG^s(\eta,\psi_1)$$
$$+ (q-1)G(\eta,\psi_1)\sum_{\substack{0\le c_i\le k\\ \sum ic_i=k\\ s\not\equiv 0(mod\ p)\\ s\equiv 0(mod\ 2)}} (-1)^{k-\sum c_i}N(c_1,\cdots,c_k)\eta\left(-s\prod_{i=1}^{k} i^{c_i}\right)q^rG^s(\eta,\psi_1). \tag{4.14}$$

By Theorem 3.2, we have that if $p\equiv 1\ (\text{mod } 4)$, then $G(\eta,\psi_1) = (-1)^{t-1}\sqrt{q}$ and $G^s(\eta,\psi_1) = ((-1)^{t-1}\sqrt{q})^s$. Since s is even, $G^s(\eta,\psi_1) = q^{\frac{s}{2}}$. Therefore, equation (4.14) becomes

$$q^2M_2(k,(0,0)) = (q)_k + (q-1)\sum_{\substack{0\le c_i\le k\\ \sum ic_i=k\\ s=0}} (-1)^{k-\sum c_i}N(c_1,\cdots,c_k)q^r$$
$$+ q(q-1)\sum_{\substack{0\le c_i\le k\\ \sum ic_i=k\\ s\equiv 0(mod\ 2p)}} (-1)^{k-\sum c_i}N(c_1,\cdots,c_k)\eta\left(\prod_{i=1}^{k} i^{c_i}\right)q^{\frac{2r+s}{2}}$$
$$+ (-1)^{t-1}(q-1)\sqrt{q}\sum_{\substack{0\le c_i\le k\\ \sum ic_i=k\\ s\not\equiv 0(mod\ p)\\ s\equiv 0(mod\ 2)}} (-1)^{k-\sum c_i}N(c_1,\cdots,c_k)\eta\left(-s\prod_{i=1}^{k} i^{c_i}\right)q^{\frac{2r+s}{2}}.$$

If $p\equiv 3\ (\text{mod } 4)$, then $G(\eta,\psi_1) = (-1)^{\frac{3t}{2}-1}\sqrt{q}$ and $G^s(\eta,\psi_1) = ((-1)^{\frac{3t}{2}-1}\sqrt{q})^s$. Since s

is even, $G^s(\eta,\psi_1) = (-1)^{\frac{st}{2}} q^{\frac{s}{2}}$. Therefore, equation (4.14) becomes

$$
\begin{aligned}
q^2 M_2(k,(0{,}0)) = (q)_k &+ (q-1) \sum_{\substack{0\le c_i\le k\\ \sum ic_i=k\\ s=0}} (-1)^{k-\sum c_i} N(c_1,\cdots,c_k) q^r \\
&+ q(q-1) \sum_{\substack{0\le c_i\le k\\ \sum ic_i=k\\ s\equiv 0 (mod\ 2p)}} (-1)^{k-\sum c_i+\frac{st}{2}} N(c_1,\cdots,c_k)\eta\left(\prod_{i=1}^{k} i^{c_i}\right) q^{\frac{2r+s}{2}} \\
&+ (-1)^{\frac{3t}{2}-1}(q-1)\sqrt{q} \sum_{\substack{0\le c_i\le k\\ \sum ic_i=k\\ s\not\equiv 0 (mod\ p)\\ s\equiv 0 (mod\ 2)}} (-1)^{k-\sum c_i+\frac{st}{2}} N(c_1,\cdots,c_k)\eta\left(-s\prod_{i=1}^{k} i^{c_i}\right) q^{\frac{2r+s}{2}}.
\end{aligned}
$$

Recalling Definition 4.1 of $N_2(k,(0,0))$ and Definition 4.2 of $M_2(k,(0,0))$, we have $M_2(k,(0,0)) = k!N_2(k,(0,0))$ and $q^2M_2(k,(0,0)) = k!q^2N_2(k,(0,0))$. Therefore, if $p \equiv 1$ (mod 4), then

$$
\begin{aligned}
N_2(k,(0,0)) =& \frac{1}{k!q^2}\Bigg[(q)_k + (q-1) \sum_{\substack{0\le c_i\le k\\ \sum ic_i=k\\ s=0}} (-1)^{k-\sum c_i} N(c_1,\cdots,c_k) q^r \\
&+ q(q-1) \sum_{\substack{0\le c_i\le k\\ \sum ic_i=k\\ s\equiv 0 (mod\ 2p)}} (-1)^{k-\sum c_i} N(c_1,\cdots,c_k)\eta\left(\prod_{i=1}^{k} i^{c_i}\right) q^{\frac{2r+s}{2}} \\
&+ (-1)^{t-1}(q-1)\sqrt{q} \sum_{\substack{0\le c_i\le k\\ \sum ic_i=k\\ s\not\equiv 0 (mod\ p)\\ s\equiv 0 (mod\ 2)}} (-1)^{k-\sum c_i} N(c_1,\cdots,c_k)\eta\left(-s\prod_{i=1}^{k} i^{c_i}\right) q^{\frac{2r+s}{2}}\Bigg]
\end{aligned}
$$

and if $p \equiv 3 \pmod 4$, then

$$N_2(k,(0,0)) = \frac{1}{k!q^2}\Bigg[(q)_k + (q-1)\sum_{\substack{0\le c_i\le k\\ \sum ic_i=k\\ s=0}}(-1)^{k-\sum c_i}N(c_1,\cdots,c_k)q^r$$

$$+\, q(q-1)\sum_{\substack{0\le c_i\le k\\ \sum ic_i=k\\ s\equiv 0 (mod\ 2p)}}(-1)^{k-\sum c_i+\frac{st}{2}}N(c_1,\cdots,c_k)\eta\left(\prod_{i=1}^{k} i^{c_i}\right)q^{\frac{2r+s}{2}}$$

$$+\,(-1)^{\frac{3t}{2}-1}(q-1)\sqrt{q}\sum_{\substack{0\le c_i\le k\\ \sum ic_i=k\\ s\not\equiv 0 (mod\ p)\\ s\equiv 0 (mod\ 2)}}(-1)^{k-\sum c_i+\frac{st}{2}}N(c_1,\cdots,c_k)\eta\left(-s\prod_{i=1}^{k} i^{c_i}\right)q^{\frac{2r+s}{2}}\Bigg].$$

4.3 Proof of Corollary 2.5

Recalling Definition 4.1 of $N_2(k,(0,0))$ and Definition 4.2 of $M_2(k,(0,0))$, to find conditions such that $N_2(k,(0,0)) > 0$, it is enough to find conditions for $M_2(k,(0,0)) > 0$.

From equation (4.10) of the previous proof, we have

$$q^2M_2(k,(0{,}0)) = (q)_k + (q-1)\sum_{\substack{0\le c_i\le k\\ \sum ic_i=k\\ s=0}}(-1)^{k-\sum c_i}N(c_1,\cdots,c_k)q^r$$

$$+\sum_{\substack{0\le c_i\le k\\ \sum ic_i=k}}(-1)^{k-\sum c_i}N(c_1,\cdots,c_k)\eta\left(\prod_{i=1}^{k} i^{c_i}\right)q^rG^s(\eta,\psi_1)\sum_{\substack{\sigma_2\neq 0\\ \sigma_1\in\mathbb{F}_q}}\psi_1^s\left(\frac{-\sigma_1^2}{4\sigma_2^2}\right)\eta^s(\sigma_2). \tag{4.15}$$

If $s=0$, then let $0^s=1$ and if $s\neq 0$, then let $0^s=0$. By rearranging equation (4.15) and taking absolute values, we have

$$|q^2 M_2(k,(0,0)) - (q)_k| \le \left| (q-1) \sum_{\substack{0 \le c_i \le k \\ \sum i c_i = k}} (-1)^{k-\sum c_i} N(c_1, \cdots, c_k) q^r 0^s \right.$$
$$\left. + \sum_{\substack{0 \le c_i \le k \\ \sum i c_i = k}} (-1)^{k-\sum c_i} N(c_1, \cdots, c_k) \eta \left(\prod_{i=1}^{k} i^{c_i} \right) q^r G^s(\eta, \psi_1) \sum_{\substack{\sigma_2 \ne 0 \\ \sigma_1 \in \mathbb{F}_q}} \psi_1^s \left(\frac{-\sigma_1^2}{4\sigma_2^2} \right) \eta^s(\sigma_2) \right|$$
$$\le (q-1) \left| \sum_{\substack{0 \le c_i \le k \\ \sum i c_i = k}} (-1)^{k-\sum c_i} N(c_1, \cdots, c_k) q^r 0^s \right|$$
$$+ \left| \sum_{\substack{0 \le c_i \le k \\ \sum i c_i = k}} (-1)^{k-\sum c_i} N(c_1, \cdots, c_k) \eta \left(\prod_{i=1}^{k} i^{c_i} \right) q^r G^s(\eta, \psi_1) \sum_{\substack{\sigma_2 \ne 0 \\ \sigma_1 \in \mathbb{F}_q}} \psi_1^s \left(\frac{-\sigma_1^2}{4\sigma_2^2} \right) \eta^s(\sigma_2) \right|$$
$$\le (q-1) \sum_{\substack{0 \le c_i \le k \\ \sum i c_i = k}} N(c_1, \cdots, c_k) q^r 0^s + q(q-1) \sum_{\substack{0 \le c_i \le k \\ \sum i c_i = k}} N(c_1, \cdots, c_k) q^r (\sqrt{q})^s. \tag{4.16}$$

Recalling Definition 3.4, inequality (4.16) becomes

$$|q^2 M_2(k,(0,0)) - (q)_k| \le (q-1) C_k(\overbrace{0, \cdots, 0}^{p-1}, q, \overbrace{0, \cdots, 0}^{p-1}, q, \ldots)$$
$$+ q(q-1) C_k(\overbrace{\sqrt{q}, \cdots, \sqrt{q}}^{p-1}, q, \overbrace{\sqrt{q}, \cdots, \sqrt{q}}^{p-1}, q, \ldots). \tag{4.17}$$

Using Lemma 3.1 and Corollary 3.1, inequality (4.17) becomes,

$$|q^2 M_2(k,(0,0)) - (q)_k| \le (q-1) k! \binom{\frac{q+k}{p} - 1}{\frac{k}{p}} + q(q-1) \left(\sqrt{q} + k + \frac{q - \sqrt{q}}{p} - 1 \right)_k .$$

For $M_2(k,(0,0)) > 0$, it is sufficient to have

$$(q)_k > (q-1)k!\binom{\frac{q+k}{p}-1}{\frac{k}{p}} + q(q-1)\left(\sqrt{q}+k+\frac{q-\sqrt{q}}{p}-1\right)_k. \tag{4.18}$$

If $q = p$, then inequality (4.18) becomes

$$(q)_k > (q-1)k! + q(q-1)\left(\sqrt{q}+k-\frac{\sqrt{q}}{q}\right)_k. \tag{4.19}$$

Suppose that

$$k \geq \frac{-2\log(q)}{\log(2c)} \text{ for some constant } 0 < c < \frac{1}{2}.$$

Since $2\log(q) > \log(q^2-\sqrt{q})$ and $-\log(2c) < \log(\sqrt{q}) - \log(c(\sqrt{q}+1))$, we have that

$$\begin{aligned} k &> \frac{\log(q^2-\sqrt{q})}{\log(\sqrt{q}) - \log(c(\sqrt{q}+1))} \\ &= \frac{\log(q^2-\sqrt{q})}{\log\left(\frac{\sqrt{q}}{c(\sqrt{q}+1)}\right)}. \end{aligned} \tag{4.20}$$

Also, since $0 < c < \frac{1}{2}$, we have

$$0 < c < \frac{\sqrt{q}}{\sqrt{q}+1}$$

and therefore,

$$\frac{\sqrt{q}}{c(\sqrt{q}+1)} > 1.$$

Since $\frac{\sqrt{q}}{c(\sqrt{q}+1)} > 1$, $\log\left(\frac{\sqrt{q}}{c(\sqrt{q}+1)}\right) > 0$. Thus, inequality (4.20) becomes

$$\log\left(\frac{\sqrt{q}}{c(\sqrt{q}+1)}\right) > \frac{1}{k}\log(q^2-\sqrt{q}). \tag{4.21}$$

If we raise e to the two quantities in inequality (4.21), we have a new inequality

$$\frac{\sqrt{q}}{c(\sqrt{q}+1)} > (q^2 - \sqrt{q})^{\frac{1}{k}}. \tag{4.22}$$

If we multiply both sides by $\sqrt{q}+1$, inequality (4.22) becomes

$$\frac{\sqrt{q}}{c} > (q^2 - \sqrt{q})^{\frac{1}{k}}(\sqrt{q}+1). \tag{4.23}$$

Suppose that $k \leq c\sqrt{q}$. Then, $\frac{1}{k} \geq \frac{1}{c\sqrt{q}}$ and inequality (4.23) becomes

$$\frac{q}{k} > (q^2 - \sqrt{q})^{\frac{1}{k}}(\sqrt{q}+1). \tag{4.24}$$

If we raise both sides by the kth power, inequality (4.24) becomes

$$\left(\frac{q}{k}\right)^k > (q^2 - \sqrt{q})(\sqrt{q}+1)^k. \tag{4.25}$$

We have that

$$(q-1)(\sqrt{q}+1)^k > q-1.$$

Therefore, inequality (4.25) becomes

$$\left(\frac{q}{k}\right)^k > (q-1) + (q^2 - q)(\sqrt{q}+1)^k \tag{4.26}$$

We have that

$$\frac{(q)_k}{k!} \geq \left(\frac{q}{k}\right)^k \text{ and } (\sqrt{q}+1)^k \geq \frac{\left(\sqrt{q}+k-\frac{\sqrt{q}}{q}\right)_k}{k!}.$$

Thus, inequality (4.26) becomes

$$\frac{(q)_k}{k!} > (q-1) + (q^2 - q)\frac{\left(\sqrt{q} + k - \frac{\sqrt{q}}{q}\right)_k}{k!}. \tag{4.27}$$

Multiplying both sides by $k!$, inequality (4.27) becomes

$$(q)_k > (q-1)k! + q(q-1)\left(\sqrt{q} + k - \frac{\sqrt{q}}{q}\right)_k,$$

which is exactly inequality (4.19).

Thus, if $q = p$, then inequality (4.18) is fulfilled and $M_2(k,(0,0)) > 0$. Therefore, $N_2(k,(0,0)) > 0$.

Chapter 5

Applications to Coding Theory

5.1 Generalized Reed-Solomon Codes and the Deep Hole Problem

When communicating over noisy channels, it is possible that errors can occur. In coding theory, we study codes that can detect and correct these errors. One important class of error-correcting codes is called the generalized Reed-Solomon codes.

Let $\mathbb{F}_q$ be a finite field of cardinality q and characteristic p. Let $D = \{x_1, \ldots, x_n\} \subseteq \mathbb{F}_q$ be an evaluation set and let $1 \leq k \leq n$. A generalized Reed-Solomon code over $\mathbb{F}_q$ with message length n and dimension k [19] is defined as

$$C = \{(f(x_1), \ldots, f(x_n)) \in \mathbb{F}_q^n \mid x_i \in D, f(x) \in \mathbb{F}_q[x], deg(f) \leq k-1\}.$$

The (Hamming) distance between two words $u, v \in \mathbb{F}_q^n$ is

$$d(u, v) = |\{i \mid u_i \neq v_i\}|$$

and the distance from a received word u to the code C is

$$d(u, C) = \min_{v \in C} d(u, v).$$

The covering radius of C is the maximum possible distance from a word in $\mathbb{F}_q^n$ and a word in C.

For a generalized Reed-Solomon code, the covering radius is $n-k$, i.e., $d(u, C) \leq n-k$, for all $u \in \mathbb{F}_q^n$.

Definition 5.1. *A received word u is called a deep hole if $d(u, C) = n - k$.*

Definition 5.2. ***(The Deep Hole Problem)*** *Determine if a given word u is a deep hole.*

It has been shown that for general evaluation sets D, the Deep Hole Problem is NP-complete [8]. If we look at particular received words u, then this problem can be answered. In Section 4.2, we review a few techniques used to answer this problem and in Section 4.3, we apply our results from Section 2.3 to certain received words u.

5.2 Some Previous Results

Let $u = (u_1, \ldots, u_n) \in \mathbb{F}_q^n$ be a received word. Define

$$u(x) := \sum_{i=1}^{n} u_i \frac{\prod_{j \neq i}(x - x_j)}{\prod_{j \neq i}(x_i - x_j)} \in \mathbb{F}_q[x].$$

The polynomial $u(x)$ is the unique polynomial of degree at most $n-1$ such that $u(x_i) = u_i$, for $1 \leq i \leq n$. Define

$$deg(u) \coloneqq deg(u(x)).$$

We have $d(u, C) = 0$ if and only if $deg(u) \leq k-1$. If $k \leq deg(u) \leq n-1$, then Li and Wan [10] proved that there is a connection between $deg(u)$ and $d(u, C)$.

Theorem 5.1. *(Li, Wan [10]) Let $u \in \mathbb{F}_q^n$ be a word such that $k \leq deg(u) \leq n-1$. Then,*

$$n - k \geq d(u, C) \geq n - deg(u).$$

By Theorem 5.1, if $deg(u) = k$, then u is a deep hole. When $D = \mathbb{F}_q$, Cheng and Murray [3] conjectured the following statement.

Conjecture 5.1. *(Cheng, Murray [3]) For the Reed-Solomon code C with $D = \mathbb{F}_q$, where p is an odd prime, a received word u is a deep hole if and only if deg(u)=k.*

This conjecture has not been proven but, there has been some progress [3].

Theorem 5.2. *(Cheng, Murray [3]) Let $q = p$ be a prime and $1 < k < p^{1/4-\epsilon}$. The vector u is not a deep hole of the Reed-Solomon code C with $D = \mathbb{F}_p$ if $k < deg(u) < k + p^{3/13-\epsilon}$.*

In [14], Li and Zhu were able to find the exact the distance or an upper bound on the distance between a received word and the generalized Reed-Solomon code under different conditions. There are many different cases, but a few cases from their work are as follows.

Theorem 5.3. *(Li, Zhu [14]) Let C be a Reed-Solomon code with $D = \mathbb{F}_q$, $k \geq 1, k+2 \leq q-1$, and $u \in \mathbb{F}_q^n$ represented by polynomial $u(x) = x^{k+2} - bx^{k+1} + cx^k + v(x), deg(v) \leq k-1$, then*

(1) If $k+2 = q-1$, then

$$d(u, C_q) = \begin{cases} q-k-2 & \text{if } b^2 = c \\ q-k-1 & \text{if } b^2 \neq c \end{cases}.$$

(2) If $p \neq 2$ and $k+2 \leq q-2$, then if $p \nmid k+2$, we have $d(u, C) \leq q-k-1$.

In the case that $p \mid k+2$, if $b = c = 0$ and $k+2 > \frac{q}{2}+1$, then $d(u, C) \leq q-k-1$.

There have been a variety of methods used to answer the Deep Hole Problem. In [2], Cheng, Li, and Zhuang used deep hole trees to find when Conjecture 5.1 is true.

Theorem 5.4. *(Cheng, Li, Zhuang [2]) Given a finite field $\mathbb{F}_q$ with characteristic $p > 2$, if $k+1 \leq p$ or $3 \leq q-p+1 \leq k+1 \leq q-2$, then Conjecture 5.1 (The Cheng-Murray conjecture) is true.*

By looking at the existence of certain $\mathbb{F}_q$-rational points of a family of hypersurfaces defined over $\mathbb{F}_q$, Cafure, Matera, and Privitelli [1] found conditions to answer the Deep Hole Problem.

Theorem 5.5. *(Cafure, Matera, Privitelli [1]) Let u be a received word and $u(x)$ be its interpolated polynomial. Suppose $1 \leq deg(u) - k \leq q-1-k$. Assume that*

$$q > \max\{(k+1)^2, 14deg(u)^{2+\epsilon}\} \text{ and } k > \left(\frac{2}{\epsilon}+1\right) deg(u)$$

for some constant $\epsilon > 0$. Then u is not a deep hole.

Li and Wan [13] and Liao [15] used character sums to find other conditions that rely on the degree of u.

Theorem 5.6. *(Li, Wan [13]) Let u be a received word and $u(x)$ be its interpolated polynomial. Suppose $1 \le deg(u) - k \le q-1-k$. If*

$$q > \max\{(k+1)^2, deg(u)^{2+\epsilon}\} \text{ and } k > \left(\frac{2}{\epsilon}+1\right) deg(u) + \frac{8}{\epsilon} + 2$$

for some constant $\epsilon > 0$, then $d(u, C) < q - k$. In other words, u is not a deep hole.

Furthermore, if

$$q > \max\{(k+deg(u))^2, (deg(u)-1)^{2+\epsilon}\} \text{ and } k > \left(\frac{4}{\epsilon}+1\right) deg(u) + \frac{4}{\epsilon} + 2$$

for some constant $\epsilon > 0$, then $d(u, C) = q - (k + deg(u))$.

Theorem 5.7. *(Liao [15]) Let $r \ge 1$ be an integer. Let u be a received word and $u(x)$ be its interpolated polynomial of degree m. If $m \ge k + r$,*

$$q > \max\left\{2\binom{k+r}{2} + (m-k), (m-k)^{2+\epsilon}\right\} \text{ and } k > \frac{1}{1+\epsilon}\left(r + (2+\epsilon)\left(\frac{m}{2}+1\right)\right)$$

for some constant $\epsilon > 0$, then $d(u, C) \le q - k - r$. So u is not a deep hole.

We were able to determine conditions on the nonexistence of deep holes using our results that were based on the Li-Wan sieve and properties of character sums.

5.3 Reduction to the Deep Hole Problem

In order to connect the Deep Hole Problem to the Moments Subset Sum Problem, we need the following theorem.

Theorem 5.8. *(Li, Wan [13]) Let C be a generalized Reed-Solomon code over $\mathbb{F}_q$ with message length n and dimension k with evaluation set D. Let $u \in \mathbb{F}_q^n$ be a word with* *deg*$(u) = k + d$, *where* $k + 1 \leq k + d \leq n - 1$. *Then, the error distance* $d(u, C) \leq n - k - r$ $(1 \leq r \leq d)$ *if and only if there exists a subset* $\{x_1, \ldots, x_{k+r}\} \subseteq D$ *and a monic polynomial* $g(x) \in \mathbb{F}_q[x]$ *of degree* $d - r$ *such that*

$$u(x) - v(x) = (x - x_1) \ldots (x - x_{k+r}) g(x)$$

for some $v(x) \in \mathbb{F}_q[x]$ *with deg*$(v) \leq k - 1$.

As an application of Corollary 2.5, we were able to find when certain received words are not deep holes.

Theorem 5.9. *Let* $q = p$, *where* p *is an odd prime. Let* C *be a generalized Reed-Solomon code with* $D = \mathbb{F}_p$. *Let* u *be a received word with deg*$(u) = k+2$ *such that* $u(x) = x^{k+2} + f(x)$, *where* $f(x) \in \mathbb{F}_q[x]$ *of degree* $< k$. *Then for a positive constant* c *such that* $0 < c < \frac{1}{2}$, *if* $\frac{-2\log(q)}{\log(2c)} - 2 < k \leq c\sqrt{q} - 2$, *then* u *is not a deep hole.*

Proof. Let u be a received word with *deg*$(u) = k + 2$.

By Theorem 5.8, $d(u, C) \leq n - k - 2$ if and only if there exists a subset $\{x_1, \ldots, x_{k+2}\} \subseteq D$ such that

$$u(x) - v(x) = (x - x_1) \ldots (x - x_{k+2}). \tag{5.1}$$

for some $v(x) \in \mathbb{F}_q[x]$ with *deg*$(v) \leq k - 1$.

Subtracting $v(x)$ from both sides, equation (5.1) becomes

$$u(x) = (x - x_1) \ldots (x - x_{k+2}) - v(x). \tag{5.2}$$

Multiplying the linear terms together, equation (5.2) becomes

$$u(x) = x^{k+2} - (x_1 + \ldots + x_{k+2})x^{k+1} + \left(\sum_{1 \le i < j \le k+2} x_i x_j \right) x^{k+1} + \widetilde{v}(x), \tag{5.3}$$

for some $\widetilde{v}(x) \in \mathbb{F}_q[x]$ with $deg(\widetilde{v}) \le k - 1$.

We have

$$u(x) = x^{k+2} + f(x), \tag{5.4}$$

where $f(x) \in \mathbb{F}_q[x]$ of degree $< k$.

Therefore, comparing equations (5.3) and (5.4), we have that $d(u, C) \le n - k - 2$ if and only if there exists a subset $\{x_1, \ldots, x_{k+2}\} \subseteq D$ such that

$$x_1 + \ldots + x_{k+2} = 0$$

and

$$\sum_{1 \le i < j \le k+2} x_i x_j = 0.$$

This is the Moments Subset Sum Problem when $D = \mathbb{F}_q$, $m = 2$, and $b = (0, 0)$. Using Corollary 2.5, if for a positive constant c such that $0 < c < \frac{\sqrt{q}}{\sqrt{q}+1}$, if $\frac{-2\log(q)}{\log(2c)} \le k + 2 \le c\sqrt{q}$, then $d(u, C) \le n - k - 2$. By Definition 5.1, if $d(u, C) \le n - k - 2$, then u is not a deep hole. $\square$

Bibliography

[1] A. Cafure, G. Matera, and M. Privitelli. Singularities of symmetric hypersurfaces and Reed-Solomon codes. *Advances in Mathematics of Communications*, 6(1):69–94, 2012.

[2] Q. Cheng, J. Li, and J. Zhuang. On determining deep holes of generalized Reed-Solomon codes. In *International Symposium on Algorithms and Computation*, pages 100–110. Springer, 2013.

[3] Q. Cheng and E. Murray. On deciding deep holes of Reed-Solomon codes. In *Proceedings of the 4th international conference on Theory and applications of models of computation*, pages 296–305. Springer-Verlag, 2007.

[4] H. Choe and C. Choe. The k-subset sum problem over finite fields of characteristic 2. *Finite Fields and Their Applications*, to appear, 2019.

[5] T. H. Cormen, C. E. Leiserson, R. L. Rivest, and C. Stein. *Introduction to algorithms*. MIT press, 2009.

[6] V. Gandikota, B. Ghazi, and E. Grigorescu. On the NP-hardness of bounded distance decoding of Reed-Solomon codes. In *2015 IEEE International Symposium on Information Theory (ISIT)*, pages 2904–2908. IEEE, 2015.

[7] V. Gandikota, B. Ghazi, and E. Grigorescu. NP-Hardness of Reed-Solomon Decoding, and the Prouhet-Tarry-Escott Problem. *SIAM Journal on Computing*, 47(4):1547–1584, 2018.

[8] V. Guruswami and A. Vardy. Maximum-likelihood decoding of Reed-Solomon codes is NP-hard. In *Proceedings of the sixteenth annual ACM-SIAM symposium on Discrete algorithms*, pages 470–478. Society for Industrial and Applied Mathematics, 2005.

[9] M. Keti and D. Wan. Deep holes in Reed-Solomon codes based on Dickson polynomials. *Finite Fields and Their Applications*, 40:110–125, 2016.

[10] J. Li and D. Wan. On the subset sum problem over finite fields. *Finite Fields and Their Applications*, 14(4):911–929, 2008.

[11] J. Li and D. Wan. A new sieve for distinct coordinate counting. *Science China Mathematics*, 53(9):2351–2362, 2010.

[12] J. Li and D. Wan. Counting polynomial subset sums. *The Ramanujan Journal*, 47(1):67–84, 2018.

[13] Y. Li and D. Wan. On error distance of Reed-Solomon codes. *Science in China Series A: Mathematics*, 51(11):1982–1988, 2008.

[14] Y. Li and G. Zhu. On error distance of received words with fixed degrees to Reed-Solomon code. *arXiv preprint arXiv:1508.02804*, 2015.

[15] Q. Liao. On Reed-Solomon codes. *Chinese Annals of Mathematics, Series B*, 32(1):89–98, 2011.

[16] R. Lidl and H. Niederreiter. *Finite fields*, volume 20. Cambridge university press, 1997.

[17] R. Merkle and M. Hellman. Hiding information and signatures in trapdoor knapsacks. *IEEE transactions on Information Theory*, 24(5):525–530, 1978.

[18] A. Shamir. A polynomial time algorithm for breaking the basic Merkle-Hellman cryptosystem. In *23rd Annual Symposium on Foundations of Computer Science (sfcs 1982)*, pages 145–152. IEEE, 1982.

[19] J. H. van Lint. *Introduction to Coding Theory.* Springer-Verlag, Berlin, Heidelberg, 3rd edition, 1998.

[20] D. Wan. Generators and irreducible polynomials over finite fields. *Mathematics of Computation of the American Mathematical Society*, 66(219):1195–1212, 1997.

[21] W. Wang and J. Nguyen. The k-subset sum problem over finite fields. *Finite Fields and Their Applications*, 51:204–217, 2018.

[22] W. Wang, L. Wang, and H. Zhou. Subset sums of quadratic residues over finite fields. *Finite Fields and Their Applications*, 43:106–122, 2017.

[23] G. Zhu and D. Wan. An asymptotic formula for counting subset sums over subgroups of finite fields. *Finite Fields and Their Applications*, 18(1):192–209, 2012.