Published in final edited form as:

Finite Fields Appl. 2020 February ; 62: . doi:10.1016/j.ffa.2019.101607.

Moment subset sums over finite fields

Tim Lai[1], **Alicia Marino**[2], **Angela Robinson**[3], **Daqing Wan**[4]

[1]Indiana University, Bloomington

[2]University of Hartford

[3]National Institute of Standards and Technology

[4]University of California, Irvine

Abstract

The k-subset sum problem over finite fields is a classical NP-complete problem. Motivated by coding theory applications, a more complex problem is the higher m-th moment k-subset sum problem over finite fields. We show that there is a deterministic polynomial time algorithm for the m-th moment k-subset sum problem over finite fields for each fixed m when the evaluation set is the image set of a monomial or Dickson polynomial of any degree n. In the classical case $m = 1$, this recovers previous results of Nguyen-Wang (the case $m = 1$, $p > 2$) [24] and the results of Choe-Choe (the case $m = 1$, $p = 2$) [3].

1. Introduction

One of the most puzzling problems in theoretical computer science, originally posed in 1971, is to determine whether P = NP [5]. That is, to determine whether the complexity class of problems which can be solved in deterministic polynomial time is equivalent to the class of problems whose solutions, if any, can be verified in deterministic polynomial time. For a comprehensive survey on this topic, see Widgerson's forthcoming monograph [25].

All NP-complete problems are equivalent to each other under polynomial time reduction. One approach to proving that P = NP is to find an NP-complete problem and prove (or disprove) that it is deterministically solvable in polynomial time. We choose the k-subset sum problem over finite fields [6], which is a classical NP-complete problem. Although this problem is out of reach, our aim of this paper is to explore deterministic polynomial time algorithms to this and similar variations of this problem in various interesting special cases.

Let p be a prime, $q = p^s$ for some integer $s > 0$, and $\mathbb{F}_q$ the finite field of q elements. Given a subset $D = \{x_1, \ldots, x_d\} \subset \mathbb{F}_q$ and $b \in \mathbb{F}_q$, let

$$N(D, b) = \#\left\{S \subseteq D : \sum_{x \in S} x = b\right\}.$$

The dense input size of D is $d \log q$, since one can simply list all the d elements of D in $\mathbb{F}_q$ where each takes $\log q$ space. The decision subset sum problem (SSP) over finite fields

asks if given D and b, can one determine whether $N(D, b) > 0$ in polynomial time in terms of the dense input size $d \log q$? If $N(D, b) > 0$, then there exists at least one collection $S \subseteq D$ of elements which sum to b. This solution, S, can be checked by addition of $|S| \leq d$ elements of size $\log q$, thus SSP $\in$ NP for every fixed p. When $p = 2$, it is a linear algebra problem and thus SSP $\in$ P. It is known SSP is NP-complete for each fixed $p > 2$.

Motivated by numerous applications, a more precise version of the SSP is to determine whether there exists a subset $S \subseteq D$ of given size k whose elements sum to b given a set D and target b as above. The decision version of this k-subset sum problem (k-SSP) is as follows. Given a subset $D = \{x_1, \ldots, x_d\} \subset \mathbb{F}_q, k \in \{1, \ldots, d\}$ and $b \in \mathbb{F}_q$, for

$$N_k(D, b) = \#\left\{S \subseteq D: \sum_{x \in S} x = b, \left|S\right| = k\right\},$$

determine whether $N_k(D, b) > 0$. The decision k-SSP problem is NP-hard for every fixed p, including the more difficult case $p = 2$ which is the main result in [23] determining that computing the minimum distance of binary codes is NP-hard. In general, the complexity of the k-SSP problem depends on the relationship between d and the modulus q. When $q = \mathcal{O}(\text{poly}(d))$, dynamic programming solves the problem in polynomial time [9, 20]. The trivial exhaustive search algorithm shows that $k - \text{SSP} \in \text{P}$ when $q = \mathcal{O}(\log \log q)$. It is known that k-SSP is NP-hard when $d = (\log q)^c$ for constant $c > 0$, see [15, 9]. An explicit formula for $N_k(D, b)$ was presented for the case of $D = \mathbb{F}_q$ [16].

In coding theory, k-SSP arises from computing the minimum distance of a linear code and the deep hole problem for Reed-Solomon codes. The set D is called the *evaluation set* as it is exactly the evaluation set of the corresponding Reed-Solomon code. If one moves further to consider the harder problem of computing the error distance of a received word (namely, maximal likelihood decoding) in Reed-Solomon codes, one is naturally lead to the following higher moment k-subset sum problem. More formally, given a subset $D = \{x_1 \ldots x_d\} \subset \mathbb{F}_q, k \in \{1 \ldots d\}, m \in \mathbb{N}$, and $\boldsymbol{b} = (b_1, \ldots, b_m) \in \mathbb{F}_q^m$, determine whether

$$N_k(D, \boldsymbol{b}, m) = \#\left\{S \subseteq D: \sum_{y \in S} y^j = b_j, 1 \leq j \leq m, \left|S\right| = k\right\},$$

is positive. This problem is known as the m-th moment k-SSP and its complexity has been studied recently. It is proven to be NP-hard for general D if $m \leq 3$ [10] or smaller than $\mathcal{O}(\log \log \log q)$ [11]. An explicit combinatorial formula for $N_k(D, \boldsymbol{b}, m)$ is obtained in [21] when $m = 2$ and $D = \mathbb{F}_q$.

All the problems and results above are based on a model where we use the dense input $\{D, b\}$ of size $\mathcal{O}(d \log q)$ by listing all the d elements of D. Though improved solutions to the decision k-SSP with such dense input are desired, one may also consider an *algebraic input* model wherein D is the set of images under some polynomial map applied to field elements. That is, for some monic polynomial $g(x) \in \mathbb{F}_q[x]$ of degree n,

$$D = g(\mathbb{F}_q) = \{g(a) : a \in \mathbb{F}_q\}.$$

In this situation, the algebraic input size would be $n \log q$ since it is enough to write down the n coefficients of the input polynomial $g(x)$. A fundamental problem is to ask if the k-SSP and the m-th moment k-SSP can be solved in deterministic polynomial time in terms of the algebraic input size $n \log q$. This appears more difficult as it is not even clear if the problem is in NP because both k and the set size $d = |D| \geq q/n$ can already be exponential in terms of the algebraic input size $n \log q$. No complexity result is yet known for the algebraic model.

The last author conjectured that k-SSP can be solved in deterministic polynomial time in algebraic input size $n \log q$ if the order of the Galois group G_g of $g(x) - t$ over $\mathbb{F}_q(t)$ is bounded by a polynomial in $n \log q$. The last condition is trivially satisfied if

$$n = O(\log \log q / \log \log \log q)$$

since then $|G_g| \leq n!$ is bounded by a polynomial in $\log q$. This condition is also satisfied when $g(x)$ is a monomial or Dickson polynomial of any degree n. Note that this conjecture is highly non-trivial, as it is not even clear whether the problem is in NP since we are using the algebraic (sparse) input size and $d \geq q/n$ is exponential in $n \log q$ for $n = O(\log \log q)$. Thus, we cannot write down all the elements of D as listing all elements of D already takes exponential time. In a sense, our set D is given as a black-box.

As a supporting evidence, this conjecture has been proved to be true in the special case when the evaluation set D is the image of the monomial x^n or Dickson polynomials of degree n: see [24] for the case $p > 2$ and [3] for the case $p = 2$. The aim of the present paper is to extend these results from $m = 1$ (k-SSP) to the higher m-moment k-SSP for each fixed m. Namely, our main result is

Theorem 1.

Let the evaluation set D be the image set of a monomial or a Dickson polynomial of degree n over $\mathbb{F}_q$. There is a deterministic algorithm which for any given $m \in \mathbb{N}$, $b \in \mathbb{F}_q^m$ and integer $k \geq 0$, decides if $N_k(D, b, m) > 0$ in time $(n \log q)^{C_m}$, where C_m is a constant depending only on m. In particular, this is a polynomial time algorithm in the algebraic input size $n \log q$ for each fixed m.

To prove the above theorem, we will need to combine all the techniques available so far: dynamic programming for large $n > q^{\epsilon}$, Kayal's algorithm [13] for constant k, Brun sieve for medium k, the Li-Wan sieve for large k and $p > 2$, and the recent Choe-Choe argument [3] for large k and $p = 2$. In addition, we need to employ the Weil bound to prove a crucial new partial character sum estimate.

2. Background

One important tool in our proof is character sum estimates. Let $\psi: \mathbb{F}_q \to \mathbb{C}$ be an additive character. We know from character theory that for a nontrivial character ψ we have $\sum_{x \in \mathbb{F}_q} \psi(x) = 0$. However, in the case of the trivial character, the sum is the size of the finite field.

Let $G = \mathbb{F}_q$ and let $\widehat{G}$ be the set of all additive characters for $\mathbb{F}_q$. Then we have the following equality

$$\sum_{\psi \in \widehat{G}} \psi(x) = \begin{cases} q \text{ if } x = \mathbf{0} \\ 0 \text{ if } x \neq \mathbf{0} \end{cases}.$$

Definition

(Dickson Polynomial). Let n be a positive integer and $a \in \mathbb{F}_q$. The Dickson polynomial of degree n is defined as

$$D_n(x, a) = \sum_{i=0}^{\lfloor n/2 \rfloor} \frac{n}{n-i} \binom{n-i}{i} (-a)^i x^{n-2i}.$$

If $n = pn_1$ is divisible by p, one checks that $D_{pn_1}(x, a) = D_{n_1}(x, a)^p$. Thus, we can assume that n is not divisible by p.

Note that for $a = 0$, $D_n(x, 0) = x^n$, so we see that Dickson polynomials are generalizations of monomials. Of particular use to us is the size of the image of these polynomials, also known as the *value set*. A simple fact for the monomial $D_n(x, 0) = x^n$ is that

$$\left| D_n(\mathbb{F}_q^\times, 0) \right| = \begin{cases} q-1 & \gcd(n, q-1) = 1 \\ \frac{1}{\ell}(q-1) & \gcd(n, q-1) = \ell \end{cases}$$

In the first case, the map is 1 to 1; in the latter case, the map is ℓ to 1. It turns out an analogous preimage-counting statement holds when $a \neq 0$. Chou, Mullen, and Wassermann in [4] used a character sum argument to calculate the following.

Notations.

For $b, c, d \in \mathbb{Z}$, Let $b^c \| d$ denote that b^c fully divides d so that $b^{c+1} \nmid d$.

Theorem 2.

Let $n \geq 2$ and $a \in \mathbb{F}_q^*$. If q is even, then $\left| D_n^{-1}(D_n(x_0, a)) \right| =$

$$\begin{cases} \gcd(n, q-1) & if\ condition\ A\ holds \\ \gcd(n, q+1) & if\ condition\ B\ holds \\ \frac{\gcd(n, q-1) + \gcd(n, q+1)}{2} & D_n(x_0, a) = 0, \end{cases}$$

where 'condition A' holds if $x^2 + x_0 x + a$ *is reducible over* $\mathbb{F}_q$ *and* $D_n(x_0, a) \neq 0$; *'condition B' holds if* $x^2 + x_0 x + a$ *is irreducible over* $\mathbb{F}_q$ *and* $D_n(x_0, a) \neq 0$.

If q *is odd, let* η *be the quadratic character of* $\mathbb{F}_q$. *If* $2^r \| (q^2 - 1)$ *then* $|D_n^{-1}(D_n(x_0, a))| =$

$$\begin{cases} \gcd(n, q-1) & if\ \eta(x_0^2 - 4a) = 1\ and\ D_n(x_0, a) \neq \pm 2a^{n/2} \\ \gcd(n, q+1) & if\ \eta(x_0^2 - 4a) = -1\ and\ D_n(x_0, a) \neq \pm 2a^{n/2} \\ \frac{\gcd(n, q-1)}{2} & if\ \eta(x_0^2 - 4a) = 1\ and\ condition\ C\ holds \\ \frac{\gcd(n, q+1)}{2} & if\ \eta(x_0^2 - 4a) = -1\ and\ condition\ C\ holds \\ \frac{\gcd(n, q-1) + \gcd(n, q+1)}{2} & otherwise, \end{cases}$$

where 'condition C' holds if

$$2^t \| n\ with\ 1 \leq t \leq r-1, \eta(a) = -1,\ and\ D_n(x_0, a) = \pm 2a^{n/2}$$

or

$$2^t \| n\ with\ 1 \leq t \leq r-2, \eta(a) = 1,\ and\ D_n(x_0, a) = -2a^{n/2}.$$

They also showed an explicit formula for the size of the value set of $D_n(x, a)$, denoted $|V_{D_n(x,a)}|$. We state their result in the odd q case.

Theorem 3.

Let $a \in \mathbb{F}_q^*$. *If* $2^r \| (q^2 - 1)$ *and* η *is the quadratic character on* $\mathbb{F}_q$ *when* q *is odd, then*

$$|V_{D_n(x,a)}| = \frac{q-1}{2\gcd(n, q-1)} + \frac{q+1}{2\gcd(n, q+1)} + \delta$$

where

$$\delta = \begin{cases} 1 & if\ q\ is\ odd,\ 2^r - 1 \| n\ and\ \eta(a) = -1 \\ \frac{1}{2} & if\ q\ is\ odd,\ 2^t \| n\ with\ 1 \leq t \leq r-2 \\ 0 & otherwise. \end{cases}$$

As a consequence, for Dickson polynomials of degree n, the value set cardinality $d = |D|$ can be computed in polynomial time in $n \log q$. Note that for a general polynomial

$g(x) \in \mathbb{F}_q[x]$ of degree n, computing the image size $|g(\mathbb{F}_q)|$ is a difficult problem, and there is no known polynomial time algorithm in terms of the algebraic input size $n \log q$, see [2] for complexity results and p-adic algorithm.

Weil's Character Sum Bound

The following classical case of the Weil bound is well known. We shall give a more general form later.

Theorem 4.—*(Weil Bound) Let* $f(x) \in \mathbb{F}_q[x]$ *be a polynomial of degree* m, *where* $(p, m) = 1$ *and* ψ *a non-trivial additive character of* $\mathbb{F}_q$. *Then*

$$\left| \sum_{x \in \mathbb{F}_q} \psi(f(x)) \right| \leq (m-1)\sqrt{q}.$$

For our purposes it will be important to have a good estimate for certain incomplete character sums, where the sum is not summing over the full field $\mathbb{F}_q$, but over the image set D of another polynomial $g(x)$. This is not available yet for general $g(x)$, but can be proved for monomials and Dickson polynomials. The monomial case is straightforward.

Proposition 1.—*Let* $f(x) \in \mathbb{F}_q[x]$ *be a polynomial of degree* m *such that* $p \nmid m$. Let $D = \{x^n : x \in \mathbb{F}_q\}$ *where* $(n+1)^2 \leq q$. *Then*

$$\left| \sum_{x \in D} \psi(f(x)) \right| \leq m\sqrt{q}.$$

Proof. Without loss of generality, we can assume that $n \mid (q-1)$. Let $D^\times = \{x^n : x \in \mathbb{F}_q^\times\}$. Using the Weil bound above,

$$\begin{aligned}
\left| \sum_{x \in D} \psi(f(x)) \right| &= \left| \psi(f(0)) + \sum_{x \in D^\times} \psi(f(x)) \right| \\
&= \left| \psi(f(0)) + \frac{1}{n} \sum_{x \in \mathbb{F}_q^\times} \psi(f(x^n)) \right| \\
&= \left| \psi(f(0)) + \frac{1}{n} \sum_{x \in \mathbb{F}_q} \psi(f(x^n)) - \frac{1}{n}\psi(f(0)) \right| \\
&\leq 1 + \frac{1}{n}(mn-1)\sqrt{q} + \frac{1}{n} \\
&= 1 + m\sqrt{q} - \frac{\sqrt{q}-1}{n}.
\end{aligned}$$

If $(n+1)^2 \leq q$ then we conclude

$$\left|\sum_{x \in D} \psi(f(x))\right| \leq m\sqrt{q}.$$

When D is the image of Dickson polynomials, the corresponding character sum estimate is harder. We need the following version of Weil's bound, which is the case $d = 1$ of Theorem 5.6 in [8].

Theorem 5.—*Let $f_i(t)$ $(1 \leq i \leq n)$ be polynomials in $\mathbb{F}_q[t]$, let $f_{n+1}(t)$ be a rational function in $\mathbb{F}_q(t)$, let D_1 be the degree of the highest square free divisor of $\prod_{i=1}^{n} f_i(t)$, let*

$$D_2 = \begin{cases} 0 & \deg(f_{n+1}) \leq 0 \\ \deg(f_{n+1}) & \deg(f_{n+1}) > 0, \end{cases}$$

let D_3 be the degree of the denominator of f_{n+1}, and let D_4 be the degree of the highest square free divisor of the denominator of $f_{n+1}(t)$ which is relatively prime to $\prod_{i=1}^{n} f_i(t)$.

Let $\chi_i : \mathbb{F}_q^* \to \mathbb{C}^*(1 \leq i \leq n)$ be multiplicative characters of $\mathbb{F}_q$, and let $\psi = \psi_p \circ Tr_{\mathbb{F}_{q^l}/\mathbb{F}_p}$ for a non-trivial additive character $\psi_p : \mathbb{F}_p \to \mathbb{C}^*$ of $\mathbb{F}_q$. Extend χ_i to $\mathbb{F}_q$ by setting $\chi_i(0) = 0$. Suppose that $f_{n+1}(t)$ is not of the form $r(t)^p - r(t) + c$ in $\mathbb{F}_q(t)$. Then, we have

$$\left|\sum_{a \in \mathbb{F}_q,\, f_{n+1}(a) \neq \infty} \chi_1(f_1(a)) \cdots \chi_n(f_n(a)) \psi(f_{n+1}(a))\right| \leq (D_1 + D_2 + D_3 + D_4 - 1)\sqrt{q},$$

where the sum is taken over those $a \in \mathbb{F}_q$ such that $f_{n+1}(a)$ is well-defined.

As a consequence, we derive the following character sum bounds.

Corollary 1.—Let $\psi_{\mathrm{Tr}} = \psi_p \circ \mathrm{Tr}_{\mathbb{F}_{q^l}/\mathbb{F}_p}$ be the canonical additive character, $\psi : \mathbb{F}_q \to \mathbb{C}^*$ any non-trivial additive character of $\mathbb{F}_q$, and $\eta : \mathbb{F}_q^* \to \mathbb{C}^*$ the quadratic character if q is odd. Let $f(x)$ be a polynomial in $\mathbb{F}_q[x]$ of degree m not divisible by p.

1. For all q, we have

$$\left|\sum_{x \in \mathbb{F}_q} \psi(f(D_n(x, a)))\right| \leq (mn - 1)\sqrt{q}.$$

2. If q is odd, then

$$\left|\sum_{x \in \mathbb{F}_q} \eta\left(x^2 - 4a\right)\psi(f(D_n(x, a)))\right| \leq (mn + 1)\sqrt{q}.$$

3. If q is even, then

$$\left|\sum_{x\in\mathbb{F}_q^*}\psi_{\mathrm{Tr}}\left(f(D_n(x,a))+a/x^2\right)\right| = \left|\sum_{x\in\mathbb{F}_q^*}\psi_{\mathrm{Tr}}\left(f(D_n(x,a))+a^{q/2}/x\right)\right| \le (mn+1)\sqrt{q}.$$

Note that none of the polynomials in place of $f_{n+1}(x)$ are of the form $r(t)^2 - r(t)+c$. This is clear if n is also not divisible by p. If n is divisible by can be reduced to the case when n is not divisible by p using the identity $D_{pn_1}(x,a) = D_{n_1}(x,a)^p$.

The following lemma is the key character sum estimate we need. The proof follows the method used in [14], where the case $m = 1$ is treated.

Lemma 1.—*Let $f(x)$ be a polynomial in $\mathbb{F}_q[x]$ of degree m not divisible by p. Let $D = \{D_n(x,a) \mid x \in \mathbb{F}_q\}$, for $a \in \mathbb{F}_q^*$. If $\psi:(\mathbb{F}_q, +) \to \mathbb{C}^*$ is a non-trivial additive character, then the following estimates hold:*

$$\left|\sum_{x\in D}\psi(f(x))\right| \le (mn+1)\sqrt{q}.$$

Proof. The sum can be rewritten in the following way:

$$S_f := \sum_{y\in D}\psi(f(y)) = \sum_{x\in\mathbb{F}_q}\psi(f(D_n(x,a)))\frac{1}{N_x},$$

where $N_x = \left|D_n^{-1}(D_n(x,a))\right|$ is size of the preimage of the value $D_n(x,a)$.

When q is even:

By Theorem 2, N_x can be quantified. Let $\mathrm{Tr}:\mathbb{F}_q \to \mathbb{F}_2$ denote the absolute trace. Using the fact that $z^2 + xz + a$ is reducible over $\mathbb{F}_q$ if and only if $\mathrm{Tr}(a/x^2) = 0$, we obtain

$$\begin{aligned} S_f = &\sum_{\substack{x\in\mathbb{F}_q^* \\ \mathrm{Tr}\left(a/x^2\right)=0}} \frac{1}{\gcd(n,q-1)}\psi(f(D_n(x,a))) \\ &+ \sum_{\substack{x\in\mathbb{F}_q^* \\ \mathrm{Tr}\left(a/x^2\right)=1}} \frac{1}{\gcd(n,q+1)}\psi(f(D_n(x,a))) \\ &+ \frac{1}{\gcd(n,q-1)}\psi(f(D_n(0,a))) + O(1), \end{aligned}$$

where $O(1)$ is a constant of size at most 1, which we accept by dropping the $D_n(x,a) = 0$ case. Denote $\psi_1:\mathbb{F}_2 \to \mathbb{C}^*$ as the order two additive character and $\psi_{\mathrm{Tr}} = \psi_1 \circ \mathrm{Tr}$, which is an additive character from $\mathbb{F}_q \to \mathbb{C}^*$. Simplifying and rearranging gives

$$
\begin{aligned}
S_f &= \frac{1}{2\gcd(n,q-1)} \sum_{x \in \mathbb{F}_q^*} \psi(f(D_n(x,a)))\left(1+\psi_{\mathrm{Tr}}\left(a/x^2\right)\right) \\
&+ \frac{1}{2\gcd(n,q+1)} \sum_{x \in \mathbb{F}_q^*} \psi(f(D_n(x,a)))\left(1-\psi_{\mathrm{Tr}}\left(a/x^2\right)\right) \\
&+ \frac{1}{\gcd(n,q-1)}\psi(D_n(0,a)) + O(1) \\
&= \left(\frac{1}{2\gcd(n,q-1)} + \frac{1}{2\gcd(n,q+1)}\right) \sum_{x \in \mathbb{F}_q^*} \psi(f(D_n(x,a))) \\
&+ \left(\frac{1}{2\gcd(n,q-1)} - \frac{1}{2\gcd(n,q+1)}\right) \sum_{x \in \mathbb{F}_q^*} \psi(f(D_n(x,a)))\psi_{\mathrm{Tr}}\left(a/x^2\right) \\
&+ \frac{1}{\gcd(n,q-1)}\psi(f(D_n(0,a))) + O(1).
\end{aligned}
$$

We add and subtract $\left(\frac{1}{2\gcd(n,q-1)} + \frac{1}{2\gcd(n,q+1)}\right)\psi(D_n(0,a))$ to complete the first sum:

$$
\begin{aligned}
&= \left(\frac{1}{2\gcd(n,q-1)} + \frac{1}{2\gcd(n,q+1)}\right) \sum_{x \in \mathbb{F}_q} \psi(f(D_n(x,a))) \\
&+ \left(\frac{1}{2\gcd(n,q-1)} - \frac{1}{2\gcd(n,q+1)}\right) \sum_{x \in \mathbb{F}_q^*} \psi(f(D_n(x,a)))\psi_{\mathrm{Tr}}\left(a/x^2\right) \\
&+ \left(\frac{1}{2\gcd(n,q-1)} - \frac{1}{2\gcd(n,q+1)}\right)\psi(f(D_n(0,a))) + O(1).
\end{aligned}
$$

In order to estimate the sum in second term, take $b \in \mathbb{F}_q^*$ so that $\psi(x) = \psi_{\mathrm{Tr}}(bx)$. Then,

$$
\sum_{x \in \mathbb{F}_q^*} \psi(f(D_n(x,a)))\psi_{\mathrm{Tr}}\left(a/x^2\right) = \sum_{x \in \mathbb{F}_q^*} \psi_{\mathrm{Tr}}\left(bf(D_n(x,a)) + a/x^2\right).
$$

Applying the bounds in Corollary 1 with f replaced by bf,

$$
\begin{aligned}
\left|\sum_{y \in D} \psi(f(y))\right| &\le \left(\frac{1}{2\gcd(n,q-1)} + \frac{1}{2\gcd(n,q+1)}\right)(mn-1)\sqrt{q} \\
&+ \left|\frac{1}{2\gcd(n,q-1)} - \frac{1}{2\gcd(n,q+1)}\right|(mn+1)\sqrt{q} + 2 \\
&\le (mn+1)\sqrt{q}.
\end{aligned}
$$

When q is odd:

We use Theorem 2 again to calculate N_x. Let η be the quadratic character of $\mathbb{F}_q$. Then,

$$
\begin{aligned}
S_f = &\sum_{\substack{x \in \mathbb{F}_q \\ \eta\left(x^2-4a\right)=1}} \frac{1}{\gcd(n,q-1)}\psi(f(D_n(x,a))) \\
&+ \sum_{\substack{x \in F_q \\ \eta\left(x^2-4a\right)=-1}} \frac{1}{\gcd(n,q+1)}\psi(f(D_n(x,a))) + O(1).
\end{aligned}
$$

The term $O(1)$ is a constant of size at most 2, which we accept by dropping the complicated 'condition C' and 'otherwise' cases. Simplifying and rearranging gives

$$\begin{aligned}
&= \frac{1}{2\gcd(n, q-1)} \sum_{x \in \mathbb{F}_q} \psi(f(D_n(x,a)))\left(1 + \eta\left(x^2 - 4a\right)\right) \\
&+ \frac{1}{2\gcd(n, q+1)} \sum_{x \in \mathbb{F}_q} \psi(f(D_n(x,a)))\left(1 - \eta\left(x^2 - 4a\right)\right) + O(1) \\
&= \left(\frac{1}{2\gcd(n, q-1)} + \frac{1}{2\gcd(n, q+1)}\right) \sum_{x \in \mathbb{F}_q} \psi(f(D_n(x,a))) \\
&+ \left(\frac{1}{2\gcd(n, q-1)} - \frac{1}{2\gcd(n, q+1)}\right) \sum_{x \in \mathbb{F}_q} \psi(f(D_n(x,a)))\eta\left(x^2 - 4a\right) + O(1).
\end{aligned}$$

Again applying the bounds in Corollary 1,

$$\begin{aligned}
\left|\sum_{x \in D} \psi(f(x))\right| &\le \left(\frac{1}{2\gcd(n, q-1)} + \frac{1}{2\gcd(n, q+1)}\right)(mn-1)\sqrt{q} \\
&+ \left|\frac{1}{2\gcd(n, q-1)} - \frac{1}{2\gcd(n, q+1)}\right|(mn+1)\sqrt{q} + 2 \\
&\le (mn+1)\sqrt{q},
\end{aligned}$$

which was to be shown.

3. k-MSS(m)

We are now ready to consider the m-th moment k-subset sum problem, called k-MSS(m) in short. Let m be a fixed positive integer, and $g(x) \in \mathbb{F}_q[x]$ a polynomial of degree n with $1 \le n \le q-1$. Let $D = g(F_q)$ and $\boldsymbol{b} = (b_1, b_2, \ldots, b_m) \in \mathbb{F}_q^m$. Since we are working in characteristic p, we have that

$$\left(x_1^i + \ldots + x_k^i\right)^p = x_1^{ip} + \ldots + x_k^{ip}.$$

Thus if $b_i^p \ne b_{ip}$ for some $ip \le m$, there will be no solutions for k-MSS(m). We may and will assume without loss of generality that $b_i^p = b_{ip}$ for all $ip \le m$ in the remainder of this paper. Under this assumption, the j-th power equation in the k-MSS(m) can and will be dropped for all j divisible by p. We introduce the moment subset sum problem over subsets of size k with the value

$$N_k(D, \boldsymbol{b}, m) = \#\left\{S \subseteq D : |S| = k, \sum_{y \in S} y^j = b_j, 1 \le j \le m, p \nmid j\right\}. \tag{1}$$

Thus, from now on, the index j is not divisible by p.

Determining whether $N_k(D, \boldsymbol{b}, m) > 0$ for given $\{D, \boldsymbol{b}\}$ is the decision version of the k-MSS(m) problem. As indicated before, we shall use the algebraic input size $n \log q$.

A closely related number is the following integer

$$M_k(D, \boldsymbol{b}, m) = \#\left\{(x_1, \ldots, x_k) \in D^k : \sum_{i=1}^{k} x_i^j = b_j, \; x_{i_1} \neq x_{i_2}, \forall 1 \leq i_1 < i_2 \leq k, p \nmid j\right\}.$$

It is clear that $M_k(D, \boldsymbol{b}, m) = k! N_k(D, \boldsymbol{b}, m)$. We deduce

Theorem 6.

$M_k(D, \boldsymbol{b}, m) > 0$ *if and only if* $N_k(D, \boldsymbol{b}, m) > 0$.

Our problem is then reduced to deciding if $M_k(D, \boldsymbol{b}, m) > 0$. We can reduce this further by assuming from duality that $k \leq \frac{|D|}{2}$. The strategy to solve this new problem is to combine all established strategy for the original subset sum problem and apply the character sum estimate from the previous section. We shall divide k into three different ranges (constant size, medium size, and large size) and use different methods for each range. The main idea is to use algorithms to solve boundary cases of parameters and to use mathematics to prove that there is a solution when the parameters are in the interior.

If $n > q^{\epsilon}$ for constant $\epsilon > 0$, then q is polynomial in $n \log q$, we can list all elements of D and use the dynamic programming algorithm to solve the moment subset sum problem in polynomial time. In the rest of the paper, we can and will assume that $n > q^{\epsilon}$ for whatever positive constant ϵ we like.

k-MSS(m) for constant size k

The main result that we depend on in this case is due to Kayal's solvability algorithm for polynomial systems over $\mathbb{F}_q$ [13], which we summarize in this context below. Let $f_1, \ldots, f_m \in \mathbb{F}_q[x_1, \ldots, x_n]$, where d is the maximum degree of all the polynomials. Let $X = V(f_1, \ldots, f_m)$ be the vanishing locus of the polynomials. Then the result of Kayal [13] states the following.

Theorem 7.—*The decision problem of* $\# X(\mathbb{F}_q) > 0$ *can be solved in time* $\left(d^{n^{cn}} m \log q\right)^{O(1)}$ for some constant $c > 0$.

Most of the conditions in our k-MSS(m) are polynomial equations, with the exception of the condition that the individual elements be distinct. However, we can easily consider this as a polynomial equation at the cost of additional variables. Recall that $D = \{g(x) : x \in \mathbb{F}_q, g(x) \in \mathbb{F}_q[x]\}$ for a polynomial g such that $\deg(g) = n$. For the context of the k-MSS(m) problem, we are deciding if the variety determined by the vanishing locus of

$$f_j(x_1, \ldots, x_k) := \left(\sum_{i=1}^{k} g(x_i)^j\right) - b_j, \; 1 \leq j \leq m, p \nmid j$$

and the additional polynomial

$$\left(\prod_{i_1 \neq i_2} (g(x_{i_1}) - g(x_{i_2}))\right) x_{k+1} - 1$$

have any $\mathbb{F}_q$-rational points. Each f_j has degree at most mn while the latter polynomial has degree $n\binom{k}{2} + 1$.

Now, we assume $k \leq 3m + 1$. Then, $n\binom{k}{2} + 1 \leq 9nm^2$ and so all the polynomials have degrees bounded by $9nm^2$. Kayal's theorem then states that the decision problem can be solved in time which is bounded by a polynomial in

$$\left(9nm^2\right)^{(k+1)^{O((k+1))}} \log q = \left(9nm^2\right)^{(3m+2)^{O((3m+2))}} \log q.$$

This is $(n \log q)^{O(1)}$ if m is a constant. Thus, we have proved the following

Theorem 8.—*Let* $D = \{g(x): x \in \mathbb{F}_q\}$, *where* $g(x) \in \mathbb{F}_q[x]$ *is any polynomial of degree* n. *Let* m *be a fixed positive integer. Assume* $k \leq 3m + 1$. *Then k-MSS(m) can be solved in time* $(n \log q)^{O(1)}$.

The condition $k \leq 3m + 1$ is all we need. It can be replaced by any bound $k \leq C$, where C is a positive constant.

k-MSS(m) for medium k

We now consider the moment k-subset sum problem for medium-sized values of k. Fix $m \in \mathbb{N}$ and $\boldsymbol{b} = (b_1, \ldots, b_m) \in \mathbb{F}_q^m$. Let $m_p = |\{j: 1 \leq j \leq m, p \nmid j\}| = m - \left\lfloor \frac{m}{p} \right\rfloor$. Recall

$$M_k(D, \boldsymbol{b}, m) = \left| \left\{ (x_1, \ldots, x_k) \in D^k : \sum_{i=1}^{k} x_i^j - b_j = 0, \; x_{i_1} \neq x_{i_2} \text{ for } i_1 \neq i_2, \; 1 \leq j \leq m, p \nmid j \right\} \right|.$$

and

$$M_k(D, \boldsymbol{b}, m) = k! \cdot N_k(D, \boldsymbol{b}, m),$$

where

$$N_k(D, \boldsymbol{b}, m) = \left| \left\{ S \subseteq D : |S| = k, \sum_{y \in S} y^j = b_j, 1 \leq j \leq m, p \nmid j \right\} \right|.$$

We wish to decide when $M_k(D, \boldsymbol{b}, m) > 0$. The following theorem solves this problem in the medium k case if certain character sum estimate is satisfied.

Theorem 9.—*Let $D = g(\mathbb{F}_q)$ where $g \in \mathbb{F}_q[x]$ with $deg(g) = n$. Let ψ be a non-trivial additive character of $\mathbb{F}_q$. Assume for all $f \in \mathbb{F}_q[x]$ of degree at most m with $p \nmid deg(f)$, we have*

$$\left| \sum_{x \in D} \psi(f(x)) \right| \leq (mn + 1)\sqrt{q}.$$

Then $M_k(D, \boldsymbol{b}, m) > 0$ if $2n(mn + 1) < q^{\frac{1}{6}}$ and $3m_p + 1 < k < q^{\frac{5}{12}}$.

The first condition $2n(mn + 1) < q^{\frac{1}{6}}$ is already satisfied, since we assumed that $n < q^{\epsilon}$ and m is a constant. The second condition $3m_p + 1 < k < q^{\frac{5}{12}}$ gives the medium range of k.

Towards this goal, we define

$$R = \left| \left\{ (x_1, \ldots, x_k) \in D^k : \sum_{i=1}^{k} x_i^j - b_j = 0, 1 \leq j \leq m, p \nmid j \right\} \right|.$$

We say that $\boldsymbol{x} = (x_1, \ldots, x_k) \in \mathbb{F}_q^k$ *is a solution* if $\boldsymbol{x}$ satisfies the conditions of R. Note that R counts solutions allowing for those with repeated entries, while $M_k(D, \boldsymbol{b}, m)$ strictly counts solutions with distinct entries. We define a new number to compute the size of R with the added condition that the first two entries of x are equal. Let

$$R_{12} = \left| \left\{ (x_1, \ldots, x_k) \in D^k : 2x_2^j + \sum_{i=3}^{k} x_i^j - b_j = 0, 1 \leq j \leq m, p \nmid j \right\} \right|.$$

Then the Brun sieve tells us that

$$M_k(D, \boldsymbol{b}, m) \geq R - \sum_{1 \leq i_1 < i_2 \leq k} R_{i_1 i_2} = R - \binom{k}{2} R_{12}.$$

In order to rewrite R and R_{12} and obtain bounds for them we use the theory of characters.

Let ψ be a non-trivial additive character of $\mathbb{F}_q$. Recall that we have the following summation:

$$\sum_{c \in \mathbb{F}_q} \psi(cx) = \begin{cases} q \text{ if } x = 0 \\ 0 \text{ if } x \neq 0 \end{cases}$$

We would like to take advantage of this character sum equation and have it evaluate solutions positively and evaluate non-solutions to zero. Thus we have the following identity.

$$\prod_{j=1,\, p\nmid j}^{m}\left(\sum_{c\in\mathbb{F}_q}\psi\left(c\left(\sum_{i=1}^{k}x_i^j-b_j\right)\right)\right)=\begin{cases}q^{m_p} & \text{if } x \text{ is a solution}\\ 0 & \text{if } x \text{ is not a solution}\end{cases}$$

With this in mind, we can rewrite R as below

$$\begin{aligned}R&=\frac{1}{q^{m_p}}\sum_{x\in D^k}\prod_{j=1,\, p\nmid j}^{m}\sum_{c\in\mathbb{F}_q}\psi\left(c\left(\sum_{i=1}^{k}x_i^j-b_j\right)\right)\\&=\frac{1}{q^{m_p}}\sum_{\boldsymbol{x}\in D^k}\sum_{\boldsymbol{c}\in\mathbb{F}_q^{m_p}}\prod_{j=1,\, p\nmid j}^{m}\psi\left(c_j\left(\sum_{i=1}^{k}x_i^j-b_j\right)\right)\\&=\frac{1}{q^{m_p}}\sum_{\boldsymbol{c}\in\mathbb{F}_q^{m_p}}\sum_{\boldsymbol{x}\in D^k}\prod_{j=1,\, p\nmid j}^{m}\psi\left(c_j\left(\sum_{i=1}^{k}x_i^j-b_j\right)\right)\\&=\frac{1}{q^{m_p}}\sum_{\boldsymbol{c}\in\mathbb{F}_q^{m_p}}\sum_{\boldsymbol{x}\in D^k}\psi\left(\sum_{j=1,\, p\nmid j}^{m}c_j\left(\sum_{i=1}^{k}x_i^j-b_j\right)\right)\end{aligned}$$

By separating the contribution of the trivial term, we obtain the following.

$$\begin{aligned}R&=\frac{1}{q^{m_p}}\sum_{\boldsymbol{x}\in D^k}\psi(0)+\frac{1}{q^{m_p}}\sum_{0\neq\boldsymbol{c}\in\mathbb{F}_q^{m_p}}\sum_{\boldsymbol{x}\in D^k}\psi\left(\sum_{j=1,\, p\nmid j}^{m}c_j\left(\sum_{i=1}^{k}x_i^j-b_j\right)\right)\\&=\frac{|D|^k}{q^{m_p}}+\frac{1}{q^{m_p}}\sum_{0\neq\boldsymbol{c}\in\mathbb{F}_q^{m_p}}S_c,\end{aligned}$$

where

$$S_c=\sum_{\boldsymbol{x}\in D^k}\psi\left(\sum_{j=1,\, p\nmid j}^{m}c_j\left(\sum_{i=1}^{k}x_i^j-b_j\right)\right).$$

Define

$$f(x)=\sum_{j=1,\, p\nmid j}^{m}c_jx^j\in\mathbb{F}_q[x].$$

Note that the degree of f is not divisible by p and at most m if $c\neq 0$. We now want to find an upper bound for S_c. Notice that

$$\psi\left(\sum_{j=1,\,p\nmid j}^{m} c_j\left(\sum_{i=1}^{k} x_i^j - b_j\right)\right) = \psi\left(\sum_{i=1}^{k}\sum_{j=1,\,p\nmid j}^{m} c_j x_i^j - \sum_{j=1,\,p\nmid j}^{m} c_j b_j\right)$$
$$= \psi(f(x_1))\cdots\psi(f(x_k))\psi\left(-\sum_{j=1,\,p\nmid j}^{m} c_j b_j\right)$$
$$= A\cdot\prod_{i=1}^{k}\psi(f(x_i)).$$

Here, $A = \psi\left(-\sum_{j=1,\,p\nmid j}^{m} c_j b_j\right)$ and so $|A| = \prod_{j=1,\,p\nmid j}^{m}|\psi(-c_j b_j)| = 1$. Thus

$$|S_c| = \left|\sum_{\boldsymbol{x}\in D^k}\prod_{i=1}^{k}\psi(f(x_i))\right| = \left(\left|\sum_{x\in D}\psi(f(x))\right|\right)^k.$$

By our assumptions, $|S_c| \le (mn+1)^k(\sqrt{q})^k$. It follows that

$$\left|R - \frac{|D|^k}{q^{m_p}}\right| = \frac{1}{q^{m_p}}\sum_{0\neq \boldsymbol{c}\in\mathbb{F}_q^{m_p}}|S_c| \le \frac{q^{m_p}-1}{q^{m_p}}(mn+1)^k q^{\frac{k}{2}} < (mn+1)^k q^{\frac{k}{2}}.$$

Remark.—Igor Shparlinski kindly informed us that the average trick in [22] can be used to improve the above coefficient $(mn+1)^k$ to $(mn+1)k-2$. The idea is to apply the character sum estimate only to the first $(k-2)$-th power in $|S_c|$, and then compute the remaining quadratic moment over c, resulting in a saving of the factor $(mn+1)^2$. This type of improvement is theoretically interesting, but would not significantly improve the lower bound condition $3m_p + 1 < k$ in our theorem, which is enough for our algorithmic purpose of this paper.

Now we can rewrite R_{12} in a similar way.

$$R_{12} = \frac{1}{q^{m_p}}\sum_{\boldsymbol{x}\in D^{k-1}}\prod_{j=1,\,p\nmid j}^{m}\sum_{c\in\mathbb{F}_q}\psi\left(c\left(2x_1^j + \sum_{i=3}^{k} x_i^j - b_j\right)\right)$$
$$= \frac{1}{q^{m_p}}\sum_{\boldsymbol{x}\in D^{k-1}}\sum_{\boldsymbol{c}\in\mathbb{F}_q^{m_p}}\prod_{j=1,\,p\nmid j}^{m}\psi\left(c_j\left(2x_1^j + \sum_{i=3}^{k} x_i^j - b_j\right)\right)$$
$$= \frac{1}{q^{m_p}}\sum_{\boldsymbol{c}\in\mathbb{F}_g^{m_p}}\sum_{\boldsymbol{x}\in D^{k-1}}\prod_{j=1,\,p\nmid j}^{m}\psi\left(c_j\left(2x_1^j + \sum_{i=3}^{k} x_i^j - b_j\right)\right)$$
$$= \frac{1}{q^{m_p}}\sum_{\boldsymbol{c}\in\mathbb{F}_q^{m_p}}\sum_{\boldsymbol{x}\in D^{k-1}}\psi\left(\sum_{j=1,\,p\nmid j}^{m} c_j\left(2x_1^j + \sum_{i=3}^{k} x_i^j - b_j\right)\right)$$

By separating the contribution of the trivial character, we obtain the following.

$$R_{12} = \frac{1}{q^{m_p}} \sum_{\boldsymbol{x} \in D^{k-1}} \psi(0) + \frac{1}{q^{m_p}} \sum_{0 \neq \boldsymbol{c} \in \mathbb{F}_q^{m_p}} \sum_{\boldsymbol{x} \in D^{k-1}} \psi\left(\sum_{j=1, p \nmid j}^{m} c_j \left(2x_1^j + \sum_{i=3}^{k} x_i^j - b_j \right) \right)$$
$$= \frac{|D|^{k-1}}{q^{m_p}} + \frac{1}{q^{m_p}} \sum_{0 \neq \boldsymbol{c} \in \mathbb{F}_q^{m_p}} S_{\boldsymbol{c}}^{12},$$

where

$$S_{\boldsymbol{c}}^{12} = \sum_{\boldsymbol{x} \in D^{k-1}} \psi\left(\sum_{j=1, p \nmid j}^{m} c_j \left(2x_1^j + \sum_{i=3}^{k} x_i^j - b_j \right) \right).$$

By a similar manipulation in the previous case,

$$S_{\boldsymbol{c}}^{12} = \sum_{\boldsymbol{x} \in D^{k-1}} \psi(2f(x_1))\psi(f(x_3)) \cdots \psi(f(x_k))\psi\left(- \sum_{j=1, p \nmid j}^{m} c_j b_j \right)$$
$$= A \sum_{\boldsymbol{x} \in D^{k-1}} \psi(2f(x_1)) \prod_{i=3}^{k} \psi(f(x_i)).$$

By a rearrangement, we see that

$$\left| S_{\boldsymbol{c}}^{12} \right| = \left| \sum_{\boldsymbol{x} \in D^{k-1}} \psi(2f(x_1)) \left(\prod_{i=3}^{k} \psi(f(x_i)) \right) \right|$$
$$= \left| \left(\sum_{x \in D} \psi(2f(x)) \right) \left(\sum_{x \in D} \psi(f(x)) \right)^{k-2} \right|$$

By our assumptions, if $p > 2$ (and thus $2 \neq 0$),

$$\left| S_{\boldsymbol{c}}^{12} \right| \leq (mn+1)\sqrt{q}(mn+1)^{k-2}(\sqrt{q})^{k-2}$$
$$= (mn+1)^{k-1} q^{\frac{k-1}{2}}.$$

The case $p = 2$ can be handled in a similar way, and one get the alternate bound

$$\left| S_{\boldsymbol{c}}^{12} \right| \leq |D| (mn+1)^{k-2} q^{\frac{k-2}{2}}.$$

We assume that $p > 2$ for simplicity. Now we have that

$$\begin{aligned}\left|R_{12}-\frac{|D|^{k-1}}{q^{m_p}}\right| &= \frac{1}{q^{m_p}}\left|\sum_{\mathbf{0}\neq\mathbf{c}\in\mathbb{F}_q^{m_p}} S_c^{12}\right| \\ &\leq \frac{1}{q^{m_p}}\sum_{\mathbf{0}\neq\mathbf{c}\in\mathbb{F}_q^{m_p}} (mn+1)^{k-1}q^{\frac{k-1}{2}} \\ &= \frac{q^{m_p}-1}{q^{m_p}}(mn+1)^{k-1}q^{\frac{k-1}{2}} \\ &< (mn+1)^{k-1}q^{\frac{k-1}{2}}.\end{aligned}$$

Since we have the following two inequalities,

$$\left|R_{12}-\frac{|D|^{k-1}}{q^{m_P}}\right| < (mn+1)^{k-1}q^{\frac{k-1}{2}}$$

$$\left|R-\frac{|D|^{k}}{q^{m_p}}\right| < (mn+1)^{k}q^{\frac{k}{2}}$$

We see that

$$\frac{|D|^{k}}{q^{m_p}}-(mn+1)^{k}q^{\frac{k}{2}} < R, \text{ and}$$

$$R_{12} < \frac{|D|^{k-1}}{q^{m_p}}+(mn+1)^{k-1}q^{\frac{k-1}{2}}.$$

Then

$$\begin{aligned}R-\binom{k}{2}R_{12} &> \frac{|D|^{k}}{q^{m_p}}-(mn+1)^{k}q^{\frac{k}{2}}-\binom{k}{2}\left(\frac{|D|^{k-1}}{q^{m_p}}+(mn+1)^{k-1}q^{\frac{k-1}{2}}\right) \\ &= \left|D\right|^{k-1}\frac{1}{q^{m_p}}\left(\left|D\right|-\binom{k}{2}\right)-(mn+1)^{k-1}q^{\frac{k-1}{2}}\left((mn+1)\sqrt{q}+\binom{k}{2}\right) \\ &= \frac{1}{q^{m_p}}\left(\left|D\right|^{k-1}\left(\left|D\right|-\binom{k}{2}\right)\right)-(mn+1)^{k-1}q^{\frac{k-1}{2}}\left((mn+1)\sqrt{q}+\binom{k}{2}\right).\end{aligned}$$

We wish to show that $R-\binom{k}{2}R_{12}$ is positive and thus we need to show that

$$\left|D\right|^{k-1}\left(\left|D\right|-\binom{k}{2}\right) \geq q^{m_p}(mn+1)^{k-1}q^{\frac{k-1}{2}}\left((mn+1)\sqrt{q}+\binom{k}{2}\right).$$

However since $\deg(g)=n$ we know that $|D|\geq\frac{q}{n}$. Thus it is enough to show that

$$\left(\frac{q}{n}\right)^{k-1}\left(\frac{q}{n}-\binom{k}{2}\right) \geq q^{m_p}+\frac{k-1}{2}(mn+1)^{k-1}\left((mn+1)\sqrt{q}+\binom{k}{2}\right).$$

Towards this goal, we utilize our assumptions that $2n(mn+1) < q^{\frac{1}{6}}$ and $3m_p + 1 < k < q^{\frac{5}{12}}$. It is enough to prove

$$\left(\frac{q}{n}\right)^{k-1} \geq q^{m_p} + \frac{k-1}{2}(mn+1)^{k-1}, \left(\frac{q}{n} - \binom{k}{2}\right) \geq \left((mn+1)\sqrt{q} + \binom{k}{2}\right).$$

For the first inequality, we have

$$\left(\frac{q}{n}\right)^{k-1} > q^{m_p} + \frac{k-1}{2}(mn+1)^{k-1} \Leftrightarrow q^{k-1-m_p-\frac{k-1}{2}} > (mn+1)^{k-1} n^{k-1}$$
$$\Leftrightarrow q^{\frac{k-1}{2}-m_p} > (mn+1)^{k-1} n^{k-1}.$$

Since, $2n(mn+1) < q^{\frac{1}{6}}$, the right side is bounded by

$$(mn+1)^{k-1} n^{k-1} < (n(mn+1))^{k-1} < q^{\frac{k-1}{6}}.$$

Our problem is now reduced to showing that $q^{\frac{k-1}{6}+m_p} < q^{\frac{k-1}{2}}$. Namely,

$$m_p < \frac{k-1}{2} - \frac{k-1}{6} = \frac{k-1}{3}.$$

This is satisfied since $3m_p + 1 < k$. Thus we have shown that

$$\left(\frac{q}{n}\right)^{k-1} > q^{m_p} + \frac{k-1}{2}(mn+1)^{k-1}. \tag{2}$$

For the second inequality, we need to show that $n(mn+1)\sqrt{q} + 2n\binom{k}{2} < q$. Since $k < q^{\frac{5}{12}}$ and $2n(mn+1) < q^{\frac{1}{6}}$, we know that $k^2 n < q^{\frac{5}{6}} q^{\frac{1}{6}}/2 = q/2$. We deduce that

$$n(mn+1)\sqrt{q} + 2n\binom{k}{2} < \frac{q^{1/6+1/2}}{2} + \frac{q}{2} < q. \tag{3}$$

The theorem is proved.

Corollary 2.—Let $D = \left\{x^d : x \in \mathbb{F}_q\right\}$ or $D = \{D_n(x,a) : x \in \mathbb{F}_q\}$ for $a \in \mathbb{F}_q^\times$. Then $M_k(D,b,m) > 0$ if $2n(mn+1) < q^{\frac{1}{6}}$ and $3m_p + 1 < k < q^{\frac{5}{12}}$.

Let ψ be a non-trivial additive character of $\mathbb{F}_q$. We have shown that all $f \in \mathbb{F}_q[x]$ of degree at most m with $p \nmid \deg(f)$,

$$\left|\sum_{x \in D} \psi(f(x))\right| \leq m\sqrt{q}$$

if $D = \left\{x^d : x \in \mathbb{F}_q\right\}$, and

$$\left|\sum_{x \in D} \psi(f(x))\right| \leq (mn+1)\sqrt{q}$$

if $D = \{D_n(x, a) : x \in \mathbb{F}_q\}$. Since $m\sqrt{q} \leq (mn+1)\sqrt{q}$, the character sum condition in Theorem 9 is satisfied. The medium case is proved.

k-MSS(m) for large k

Following established procedures, we use the Li-Wan sieve [17] to analyze large values of k. This method has been used several times [26, 14, 17, 18, 19, 24] and is now standard. So, we will only give an outline and indicate the differences. We begin by discussing the relevant notation and concepts that we will apply in our context. In this section, we assume that D is the image of a monomial or Dickson polynomial of degree n. The relevant character sum estimate is then true.

We use the notation S_k to denote the symmetric group on k letters. For a permutation $\tau \in S_k$, its disjoint cycle decomposition is written as

$$\tau = (a_1 a_2 \cdots a_{m_1})(a_{m_1+1} \cdots a_{m_2}) \cdots (a_{m_{k-1}+1} \cdots a_{m_k}).$$

We shall refer to τ interchangeably with its disjoint cycle decomposition, which we fix beforehand.

Denote by $\overline{X} = \left\{(x_1, \ldots, x_k) \in D^k : x_i \neq x_j, \forall i \neq j\right\}$. For the sake of brevity, we will denote k-tuples from such products by $x = (x_1, ..., x_k)$ when there is no risk of confusion. Let ψ be a fixed non-trivial additive character of $\mathbb{F}_q$. Recall from earlier sections that we are interested in

$$h_c(x_1, \ldots, x_k) = \psi\left(\sum_{j=1,\, p \nmid j}^{m} c_j \left(\sum_{i=1}^{k} x_i^j - b_j\right)\right),$$

where c is not the zero vector. Now define

$$F(c) = \sum_{x \in \overline{X}} h_c(x_1, \ldots, x_k), F_\tau(c) = \sum_{x \in X_\tau} h_c(x_1, \ldots x_k),$$

where X_τ consists of tuples in $\overline{X}$ such that

$$x_{a_1} = \ldots = x_{a_{m_1}}, x_{m_1+1} = \ldots = x_{m_2}, \ldots, x_{m_{k-1}+1} = \ldots = x_{m_k}$$

and so on. Now, let's think of τ as having e_1 cycles of length 1, e_2 cycles of length 2, and so on, up until e_k cycles of length k. Note that $\sum_{i=1}^{k} ie_i = k$. This allows us to express $F_\tau(c)$ as:

$$\begin{aligned} F_\tau(c) &= \sum_{x \in X_\tau} \psi\left(\sum_{j=1,\, p \nmid j}^{m} c_j \left(\sum_{i=1}^{k} i\left(x_{i1}^j + \cdots + x_{ie_i}^j\right) - b_j \right) \right) \\ &= \sum_{\substack{x_{il} \in D \\ 1 \le i \le k \\ 1 \le l \le e_i}} \psi\left(\sum_{j=1,\, p \nmid j}^{m} c_j \left(\sum_{i=1}^{k} \sum_{l=1}^{e_i} i x_{il}^j \right) \right) \psi\left(\sum_{j=1,\, p \nmid j}^{m} - c_j b_j \right) \\ &= \sum_{\substack{x i l \in D \\ 1 \le i \le k \\ 1 \le l \le e_i}} \prod_{i=1}^{k} \psi^i\left(\sum_{p \nmid j,\, l} c_j x_{il}^j \right) \psi\left(\sum_{j=1,\, p \nmid j}^{m} - c_j b_j \right). \end{aligned}$$

Let's consider the inner sum.

$$\sum_{x_{il} \in D} \psi^i\left(\sum_{p \nmid j} c_j x_{il}^j \right) = \sum_{x \in D} \psi^i(f(x))$$

where $f(x) = \sum_{j=1,\, p \nmid j}^{m} c_j x^j$. Hence, if the c_j's are not all zero, we have

$$\left| \sum_{x \in D} \psi\left(\sum_{j=1,\, p \nmid j}^{m} c_j x^j \right) \right| \le (mn+1)\sqrt{q}\,.$$

Now the order of ψ is p so the order of ψ^i is $\frac{p}{(i,p)}$, which is p unless $p \mid i$, in which case it is 1. Therefore,

$$\begin{aligned} |F_\tau(c)| &= \left| \prod_{i=1}^{k} \left(\sum_{x \in D} \psi^i\left(\sum_{j=1,\, p \nmid j}^{m} c_j x^j \right) \right)^{e_i} \psi\left(\sum_{j=1,\, p \nmid j}^{m} - c_j b_j \right) \right| \\ &\le \prod_{\substack{i \\ 1 \le i \le k \\ p \nmid i}} ((mn+1)\sqrt{q})^{e_i} \cdot \prod_{\substack{i \\ 1 \le i \le k \\ p \mid i}} \left| D \right|^{e_i}. \end{aligned}$$

The Li-Wan sieve says that

$$F(c) = \sum_{\sum i e_i = k} (-1)^{k - \sum e_i} N(e_1, \ldots, e_k) F_{e_1, \ldots, e_k}(c),$$

where $N(e_1, ..., e_k)$ denote the number of permutations in S_k with cycle type $(e_1, ..., e_k)$, and $F_{e_1, \ldots, e_k}(c)$ denotes $F_\tau(c)$ for any τ of cycle type $(e_1, ..., e_k)$. Using the above estimates and Lemma 2.1 in [24], one obtains

$$\begin{aligned} \left| F(c) \right| &\leq \sum_{\sum i e_i = k} N(e_1, \ldots, e_k) \prod_{(i,p)=1} ((mn+1)\sqrt{q})^{e_i} \cdot \prod_{p \mid i} \left| D \right|^{e_i} \\ &\leq \left((mn+1)\sqrt{q} + k + \frac{\left| (mn+1)\sqrt{q} - \left| D \right| \right|}{p} - 1 \right)_k \end{aligned}$$

where we define $(x)_k := x(x-1)\cdots(x-k+1)$.

This concludes our discussion of the Li-Wan sieve and the appropriate adaptation to our context. We now return to the framework in the previous sections, with notations as before. Let's see how the above Li-Wan helps. Recall

$$\begin{aligned} M_k(D, b, m) &= \sum_{x \in \mathbf{X}} \frac{1}{q^{m_p}} \sum_{\psi} \sum_{c_j \in \mathbf{F}_q} \psi\left(\sum_{j=1, p \nmid j}^{m} c_j \left(\sum_{i=1}^{k} x_i^j - b_j \right) \right) \\ &= \frac{1}{q^{m_p}} (|D|)_k + \sum_{(\cdots c_j \cdots) \neq 0} \frac{1}{q^{m_p}} F(c). \end{aligned}$$

Therefore,

$$\left| M_k(D, b, m) - \frac{1}{q^{m_p}} (|D|)_k \right| < \left((mn+1)\sqrt{q} + k + \frac{\left| (mn+1)\sqrt{q} - \left| D \right| \right|}{p} - 1 \right)_k \tag{4}$$

$$\leq \left(0.013 \left| D \right| + k + \frac{|D|}{p} \right)_k. \tag{5}$$

This estimate is the analogue of equation (2.3) in [24], resulting from assuming that

$$(mn+1)\sqrt{q} \leq 0.013 \left| D \right|.$$

If further, $6m_p \ln q \leq k \leq \frac{|D|}{2}$, the same argument as in the proof of Theorem 2.3 in [24] shows that $M_k(D, b, m) > 0$. We obtain

Theorem 10.

Let D be the image of a monomial of Dickson polynomial of degree n. Assume that $p > 2$, $(mn+1)\sqrt{q} \le 0.013|D|$, and $6m_p \ln q \le k \le \frac{|D|}{2}$. Then, $M_k(D, b, m) > 0$.

Note that if $p = 2$, the same proof works, but only for k in the shorter range $6m_p \ln q \le k \le \frac{(1-\epsilon)|D|}{2}$. That is, k cannot reach all the way to $|D|/2$ if $p = 2$.

Since $|D| \ge q/n$, the condition $(mn+1)\sqrt{q} \le 0.013|D|$ is satisfied if $n(mn+1)\sqrt{q} \le 0.013q$, which is certainly true since m is fixed and $n < q^{\epsilon}$.

4. Case $p = 2$

Finally, we examine the k-MSS(m) over finite fields of characteristic 2. The result of Kayal used for k-MSS(m) for constant k and our proof for medium-sized k still hold in fields of characteristic 2. Thus Theorem 8 and Theorem 9 hold for $q = p^s$ for all p.

To analyze the case $p = 2$ for large k, we rely on recent work by Choe and Choe [3] which examines the subset sum problem over finite fields of characteristic 2. We adjust the definitions of this work to fit the higher moment subset sum problem over D which are images of monomials or Dickson polynomials. Note that $p = 2$ in this section.

We will prove an analogue of Theorem 2.3 in [3]. Let $D \subseteq \mathbb{F}_q$, $k \le |D|/2$ and $f(x) = \sum_{j=1, p \nmid j}^{m} c_j x^j$, for $c_j \in \mathbb{F}_q$. For a nontrivial additive character ψ of $\mathbb{F}_q$, define

$$S_D(k, \psi, f) = \sum_{\substack{x_1 \in D \\ x_i \text{ distinct}}} \psi(f(x_1) + f(x_2) + \ldots + f(x_k)).$$

Although $S_D(k, \psi, f)$ sums over distinct x_i, there is no assumption that the $f(x_i)$ are distinct. Over finite fields of characteristic 2, however, if $x_i = x_j$, then $f(x_i) = f(x_j)$, and the sum $f(x_i) + f(x_j)$ is equivalent to $2f(x_i) = 0$. It follows that

$$\begin{aligned} S_D(2, \psi, f) &= \sum_{\substack{x_1, x_2 \in D \\ x_1 \ne x_2}} \psi(f(x_1) + f(x_2)) \\ &= \left(\sum_{x \in D} \psi(f(x)) \right)^2 - |D|. \end{aligned}$$

By induction, one derives the following recursive formula for $S_D(k, \psi, f)$ for all $k > 1$, which is the analogue of Lemma 2.1 [3].

Lemma 2.

Let D be a subset of $\mathbb{F}_q$ with more than 3 elements and ψ be a nontrivial additive character of $\mathbb{F}_q$. Then

- $S_D(1, \psi, f) = \sum_{x \in D} \psi(f(x))$,
- $S_D(2, \psi, f) = S_D(1, \psi, f)^2 - |D|$, *and*
- $S_D(k, \psi, f) = S_D(1, \psi, f) S_D(k-1, \psi, f) - (|D| - k + 2)(k-1) S_D(k-2, \psi, f)$, *where* $3 \le k \le |D|$.

This lemma can be applied to prove analogue of Lemma 2.2 [3]. The statement is as follows.

Lemma 3.

Let D *be a subset of* $\mathbb{F}_q$ *with more than 4 elements and* ψ *be a nontrivial additive character of* $\mathbb{F}_q$. *If*

$$\left| \sum_{x \in D} \psi(f(x)) \right| \le \frac{1}{16} |D|,$$

then

$$|S_D(k, \psi, f)| < \left(\frac{9}{16} |D| \right)^k, \text{ for all } k \le \frac{|D|}{2}.$$

From Proposition 1 and Lemma 1, it follows that when D is the image of a polynomial of degree n such that the value set character sum estimate satisfies

$$\left| \sum_{x \in D} \psi(f(x)) \right| < (mn+1)\sqrt{q},$$

then the condition $n(mn+1) < \frac{1}{16}\sqrt{q}$ implies that

$$\left| \sum_{x \in D} \psi(f(x)) \right| < (mn+1)\sqrt{q} < \frac{1}{16} \frac{q}{n} \le \frac{1}{16} |D|.$$

As in the previous section, a standard character sum argument gives the inequality

$$\left| M_k(D, b, m) - \left(\frac{1}{q} \right)^{m_p} (|D|)_k \right| < \max_{c \in \mathbb{F}_q^{mp} - 0} S(k, \psi, f_c), \tag{6}$$

where $f_c = \sum_{j=1, p \nmid j}^{m} c_j x^j$. It follows that

$$\left| M_k(D, b, m) - \left(\frac{1}{q} \right)^{m_p} (|D|)_k \right| < \left(\frac{9}{16} |D| \right)^k. \tag{7}$$

The same argument as in the proof of Theorem 2.3 in [3] shows that if

$$3.05sm_p = 3.05m_p\log_2 q < k \leq |D|/2,$$

then

$$\frac{1}{q^{m_p}}(|D|)_k > \frac{1}{q^{m_p}}\left(\frac{9}{16}|D|\right)^k 2^{sm_p} = \left(\frac{9}{16}|D|\right)^k, \tag{8}$$

Thus, we obtain

Theorem 11.

Let $p = 2$ *and* $n(mn+1) < \frac{1}{16}\sqrt{q}$. *Then* $M_k(D, b, m) > 0$ *for all* $3.05m_p\log_2 q < k \leq |D|/2$.

We conclude that when D is the image of degree n polynomial satisfying the value set character sum estimate in Lemma 1, the m-th moment subset sum problem over D can be solved in deterministic polynomial time in the algebraic input size $n \log q$, for every constant m. In particular, this is true when D is the image of a monomial of Dickson polynomial of degree n.

5 Conclusion

We show that there is a deterministic polynomial time algorithm for the m-th moment k-subset sum problem over finite fields for each fixed m when the evaluation set is the image set of a monomial or Dickson polynomial of any degree n. An open problem is to ask if Theorem 1 can be proved for larger range of m, say, $m = O(\log \log q)$. The difficulty lies in the small k range such as $k \leq 3m + 1$.

Acknowledgements

This work was supported by the Early Career Research Workshop in Coding Theory, Cryptography, and Number Theory held at Clemson University in 2018, under NSF grant DMS-1547399.

References

[1]. Cheng Q. and Murray E. On deciding deep holes of Reed-Solomon codes. In International Conference on Theory and Applications of Models of Computation pages 296–305. Springer, Berlin, Heidelberg, May 2007.

[2]. Cheng Q, Hill J. and Wan D, Counting value sets: algorithms and complexity. ANTS X, Proceedings of the Tenth Algorithmic Number Theory Symposium, The Open Book Series, 1(2013), 235–248.

[3]. Choe H. and Choe C. The k-subset sum problem over fields of characteristic 2. Finite Fields & Applications, 59(2019), 175–184.

[4]. Chou W, Mullen GL, and Wasserman B. On the number of solutions of equations of Dickson polynomials over finite fields. Taiwanese Journal of Mathematics Vol. 12, No. 4, pp. 917–931, July 2008.

[5]. Cook SA The complexity of theorem-proving procedures. Proceedings of the third annual ACM symposium on Theory of computing, ACM, 1971.

[6]. Cormen TH, Leiserson CE, Rivest RL, and Stein C, C. Introduction to algorithms MIT press, 2009.

[7]. Frieze AM On the Lagarias-Odlyzko algorithm for the subset sum problem. SIAM Journal on Computing 15, no. 2 (1986): 536–539.

[8]. Fu L. and Wan D. A class of incomplete character sums. Q. J. Math 65 (2014), no. 4, 1195–1211.

[9]. Galil Z. and Margalit O. An almost linear-time algorithm for the dense subset-sum problem. SIAM Journal on Computing 20, no. 6 (1991): 1157-1189.

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

[11]. Gandikota V, Ghazi B, Grigorescu E. NP-Hardness of Reed–Solomon Decoding, and the Prouhet–Tarry–Escott Problem. SIAM Journal on Computing, 2018.

[12]. Garey MR and Johnson DS Computers and Intractability: A Guide to the Theory of NP-Completeness. W. H. Freeman, 1979.

[13]. Kayal N. Solvability of a system of bivariate polynomial equations over a finite field (extended abstract). In Automata, languages and programming, volume 3580 of Lecture Notes in Comput. Sci, pages 551–562. Springer, Berlin, 2005.

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

[15]. Lagarias JC and Odlyzko AM Solving low-density subset sum problems. Journal of the ACM (JACM) 32, no. 1 (1985): 229–246.

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

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

[18]. Li J. and Wan D. Counting polynomial subset sums. Ramanujan Journal, 47(2018), 67–84.

[19]. Li J. and Wan D. Distance distribution in Reed-Solomon codes. IEEE Transactions on Information Theory, 2019, to appear.

[20]. Lyubashevsky V. On random high density subset sums. Electronic Colloquium on Computational Complexity (ECCC). Vol. 12. No. 007. 2005.

[21]. Nguyen J, Higher moments subset sums over finite fields, Ph.D. Dissertation, UC Irvine, 2019.

[22]. Shparlinski I, Cayley graphs generated by small degree polynomials over finite fields, SIAM J. Discrete Math, Vol 29, 1(2015), 376–381.

[23]. Vardy Alexander. The intractability of computing the minimum distance of a code. IEEE Trans. Inform. Theory, 43(6):1757–1766, 1997.

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

[25]. Wigderson A, Mathematics and Computation, https://www.math.ias.edu/files/Book-online-Aug0619.pdf#page=1

[26]. Zhu G. and Wan D, An asymptotic formula for counting subset sums over subgroups of finite fields. Finite Fields & Applications, 18(2012), No. 1, 192–209.