ON THE SUBSET SUM PROBLEM OVER FINITE FIELDS

JIYOU LI AND DAQING WAN

Abstract. The subset sum problem over finite fields is a well known **NP**-complete problem. It arises naturally from decoding generalized Reed-Solomon codes. In this paper, we study the number of solutions of the subset sum problem from a mathematicial point of view. In several interesting cases, we obtain explicit or asymptotic formulas for the solution number. As a consequence, we get some information on the decoding problem of Reed-Solomon codes.

1. Introduction

Let $\mathbf{F}_q$ be a finite field of characteristic p. Let $D \subseteq \mathbf{F}_q$ be a subset of cardinality $|D| = n > 0$. Let $1 \le m \le k \le n$ be integers. Given m elements $b_1, \cdots, b_m$ in $\mathbf{F}_q$. Let $V_{b,k}$ denote the affine variety in $\mathbf{A}^k$ defined by the following system of equations

$$\sum_{i=1}^{k} X_i = b_1,$$

$$\sum_{1 \le i_1 < i_2 \le k} X_{i_1} X_{i_2} = b_2,$$

$$\cdots,$$

$$\sum_{1 \le i_1 < i_2 < \cdots < i_m \le k} X_{i_1} \cdots X_{i_k} = b_m,$$

$$X_i - X_j \neq 0 \ (i \neq j).$$

A fundamental problem arising from decoding Reed-Solomon codes is to determine for any given $b = (b_1, \cdots, b_m) \in \mathbf{F}_q^m$, if the variety $V_{b,k}$ has an $\mathbf{F}_q$-rational point with all $x_i \in D$, see section 5 for more detail. This problem is apparently difficult due to several parameters of different nature involved. The high degree of the variety naturally introduces a substantial algebraic difficulty, but this can at least be overcome in some cases when D is the full field $\mathbf{F}_q$ and m is small, using the Weil bound. The requirement that the x_i's are distinct leads to a significant combinatorial difficulty. From computational point of view, a more substantial difficulty is caused by the flexibility of the subset D of $\mathbf{F}_q$. In fact, even in the case $m = 1$ and so the algebraic difficulty disappear, the problem is known to be **NP**-complete. In this case, we are reduced to the well known subset sum problem over $D \subseteq \mathbf{F}_q$, that is, to determine for a given $b \in \mathbf{F}_q$, if there is a non-empty subset $\{x_1, x_2, \cdots, x_k\} \subseteq D$ such that

$$x_1 + x_2 + \cdots + x_k = b. \tag{1.1}$$

This research is supported by the National Natural Science Foundation of China(10331030).

This subset sum problem is known to be **NP**-complete. Given integer $1 \le k \le n$, and $b \in \mathbf{F}_q$, a more precise problem is to determine

$$N(k,b,D) = \#\{\{x_1, x_2, \cdots, x_k\} \subseteq D \mid x_1 + x_2 + \cdots + x_k = b\},$$

the number of k-element subsets of D whose sum is b. The decision version of the above subset sum problem is then to determine if $N(k,b,D) > 0$ for some k, that is, if

$$N(b,D) := \sum_{k=1}^{n} N(k,b,D) > 0.$$

In this paper, we study the approximation version of the above subset sum problem for each k from a mathematical point of view, that is, we try to approximate the solution number $N(k,b,D)$. Intuitively, the problem is easier if D is close to be the full field $\mathbf{F}_q$, i.e., when $q-n$ is small. Indeed, we obtain an asymptotic formula for $N(k,b,D)$ when $q-n$ is small. Heuristically, $N(k,b,D)$ should be approximately $\frac{1}{q}\binom{n}{k}$. The question is about the error term. We have

Theorem 1.1. *Let $p < q$, that is, $\mathbf{F}_q$ is not a prime field. Let $D \subseteq \mathbf{F}_q$ be a subset of cardinality n. For any $1 \le k \le n \le q-2$, any $b \in \mathbf{F}_q$, we have the inequality*

$$\left| N(k,b,D) - \frac{1}{q}\binom{n}{k} \right| \le \frac{q-p}{q} \binom{k+q-n-2}{q-n-2} \binom{q/p-2}{\lfloor k/p \rfloor}.$$

Furthermore, let $D = \mathbf{F}_q \backslash \{a_1, \cdots, a_{q-n}\}$ with $a_1 = 0$, and if $b, a_2, \cdots, a_{q-n}$ are linearly independent over $\mathbf{F}_p$, then we have the improved estimate

$$\left| N(k,b,D) - \frac{1}{q}\binom{n}{k} \right| \le \max_{0 \le j \le k} \frac{p}{q} \cdot \binom{k+q-n-2-j}{q-n-2} \binom{q/p-1}{\lfloor j/p \rfloor}.$$

This theorem implies

$$N(b,D) > 1/q\big(2^n - p \cdot q^{q-n-2} 2^{q/p-2}\big).$$

It follows that there is a constant C such that if $q-n \le C \cdot (\frac{q \ln 2}{\ln q})$ then $N(b,D) > 0$ for large $q > p$.

The above theorem gives a sufficient condition on k for which $N(k,b,D) > 0$ for all b when c is small. Our proof shows that the problem also becomes somewhat easier when the characteristic p is small compared to q.

The above theorem assumes that $n \le q-2$. In the remaining case $n \ge q-2$, that is, $n \in \{q-2, q-1, q\}$, the situation is nicer and we obtain explicit formulas for $N(k,b,D)$. Here we state the results for $q-n \le 1$ and thus we can take $D = \mathbf{F}_q$ or $\mathbf{F}_q^*$.

Theorem 1.2. *If $p \nmid k$, then*

$$N(k,b,\mathbf{F}_q) = \frac{1}{q}\binom{q}{k}.$$

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

$$N(k,0,\mathbf{F}_q) = \frac{1}{q}\binom{q}{k} + (-1)^{k+\frac{k}{p}} \frac{q-1}{q} \binom{q/p}{k/p}.$$

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

$$N(k,b,\mathbf{F}_q) = \frac{1}{q}\binom{q}{k} + (-1)^{k+\frac{k}{p}} \frac{-1}{q} \binom{q/p}{k/p}.$$

If $b = 0$, then

$$N(k, 0, \mathbf{F}_q^*) = \frac{1}{q}\binom{q-1}{k} + (-1)^{k+\lfloor k/p \rfloor}\frac{q-1}{q}\binom{q/p-1}{\lfloor k/p \rfloor}.$$

If $b \neq 0$, then

$$N(k, b, \mathbf{F}_q^*) = \frac{1}{q}\binom{q-1}{k} + (-1)^{k+\lfloor k/p \rfloor}\frac{-1}{q}\binom{q/p-1}{\lfloor k/p \rfloor}.$$

The explicit formula for the case $q - n = 2$ is given in Theorem 3.5. Corollary 3.6 shows that the estimate in Theorem 1.1 is sharp for $q - n = 2$. Applications to coding theory are given in section 5. The case $p = q$ is discussed in section 4.

Notations. Let $(x)_0 = 1$ and $(x)_k = x(x-1)\cdots(x-k+1)$ for $k \in \mathbb{Z}^+ = \{1, 2, 3, \cdots\}$. For $k \in \mathbb{N} = \{0, 1, 2, \cdots\}$ define the binomial coefficient $\binom{x}{k} = \frac{(x)_k}{k!}$. For a real number a we denote $\lfloor a \rfloor$ to be the largest integer not greater than a.

2. Proof of Theorem 1.2

When D equals $q-1$, it suffices to consider $N(k, b, \mathbf{F}_q^*)$ by a simple linear substitution. Let $M(k, b, D)$ denote the number of ordered tuples $(x_1, x_2, \cdots, x_k)$ satisfying equation (1.1). Then

$$M(k, b, D) = k!N(k, b, D)$$

is the number of solutions of the equation

$$x_1 + \cdots + x_k = b, x_i \in D, x_i \neq x_j \ (i \neq j). \tag{2.0}$$

It suffices to determine $M(k, b, D)$. We use a pure combinatorial method to find recursive relations among the values of $M(k, b, \mathbf{F}_q)$ and $M(k, b, \mathbf{F}_q^*)$.

Lemma 2.1. *For $b \neq 0$ and D being $\mathbf{F}_q$ or $\mathbf{F}_q^*$, we have $M(k, b, D) = M(k, 1, D)$.*

Proof. There is a one to one map sending the solution $\{x_1, x_2, \cdots, x_k\}$ of (2.0) to the solution $\{x_1 b^{-1}, x_2 b^{-1}, \cdots, x_k b^{-1}\}$ of (2.0) with $b = 1$. □

Lemma 2.2.

$$M(k, 1, \mathbf{F}_q) = M(k, 1, \mathbf{F}_q^*) + kM(k-1, 1, \mathbf{F}_q^*), \tag{2.1}$$

$$M(k, 0, \mathbf{F}_q) = M(k, 0, \mathbf{F}_q^*) + kM(k-1, 0, \mathbf{F}_q^*), \tag{2.2}$$

$$(q)_k = (q-1)M(k, 1, \mathbf{F}_q) + M(k, 0, \mathbf{F}_q), \tag{2.3}$$

$$(q-1)_k = (q-1)M(k, 1, \mathbf{F}_q^*) + M(k, 0, \mathbf{F}_q^*). \tag{2.4}$$

Proof. Fix an element $c \in \mathbf{F}_q$. The solutions of (2.0) in $\mathbf{F}_q$ can be divided into two classes depending on whether c occurs. By a linear substitution, the number of solutions of (2.0) in $\mathbf{F}_q$ not including c equals $M(k, b - ck, \mathbf{F}_q^*)$. And the number of solutions of (2.0) in $\mathbf{F}_q$ including c equals $kM(k-1, b-ck, \mathbf{F}_q^*)$. Hence we have

$$M(k, b, \mathbf{F}_q) = M(k, b - ck, \mathbf{F}_q^*) + kM(k-1, b-ck, \mathbf{F}_q^*). \tag{2.5}$$

Choose $b = 1, c = 0$ we get (2.1) and choose $b = 0, c = 0$ we obtain (2.2). Note that $(q)_k$ is the number of k-permutations of $\mathbf{F}_q$, and $(q-1)_k$ is the number of k-permutations of $\mathbf{F}_q^*$. Thus, both (2.3) and (2.4) follows. □

The next step is to find more relations between $M(k, b, \mathbf{F}_q)$ and $M(k, b, \mathbf{F}_q^*)$.

Lemma 2.3. *If $p \nmid k$, we have $M(k,b,\mathbf{F}_q) = M(k,0,\mathbf{F}_q)$ for all $b \in \mathbf{F}_q$ and hence*

$$M(k,b,\mathbf{F}_q) = \frac{1}{q}(q)_k.$$

If $p \mid k$, we have $M(k,b,\mathbf{F}_q) = qM(k-1,b,\mathbf{F}_q^*)$ for all $b \in \mathbf{F}_q$.

Proof. **Case 1:** Since $p \nmid k$, we can take $c = k^{-1}b$ in (2.5) and get the relation

$$M(k,b,\mathbf{F}_q) = M(k,0,\mathbf{F}_q^*) + kM(k-1,0,\mathbf{F}_q^*).$$

The right side is just $M(k,0,\mathbf{F}_q)$ by (2.2).

Case 2 : In this case, $p \mid k$. Then $M(k,b,\mathbf{F}_q)$ equals the number of ordered solutions of the following system of equations:

$$\begin{cases} x_1 + x_2 + \cdots + x_k = b, \\ x_1 - x_2 = y_2, \\ \cdots\cdots \\ x_1 - x_k = y_k, \\ y_i \in \mathbf{F}_q^*, \quad y_i \neq y_j, \ 2 \le i < j \le k. \end{cases}$$

Regarding $x_1, x_2, \cdots, x_k$ as variables it is easy to check that the p-rank (the rank of a matrix over the prime field $\mathbf{F}_p$) of the coefficient matrix of the above system of equations equals $k-1$. The system has solutions if and only if $\sum_{i=2}^{k} y_i = b$ and $y_i \in \mathbf{F}_q^*$ being distinct. Furthermore, since the p-rank of the above system is $k-1$, when $y_2, y_3, \cdots, y_k$ and x_1 are given then $x_2, x_3, \cdots, x_k$ will be uniquely determined. This means the number of the solutions of above linear system of equations equals to q times the number of ordered solutions of the following equation:

$$\begin{cases} y_2 + y_3 + \cdots + y_k = b, \\ y_i \in \mathbf{F}_q^*, \quad y_i \neq y_j, \ 2 \le i < j \le k. \end{cases}$$

This number of solutions of the above equation is just $M(k-1,b,\mathbf{F}_q^*)$ and hence $M(k,b,\mathbf{F}_q) = qM(k-1,b,\mathbf{F}_q^*)$. □

We have obtained several relations from Lemma 2.2 and Lemma 2.3. To determine $M(k,b,\mathbf{F}_q)$, it is now sufficient to know $M(k,0,\mathbf{F}_q^*)$. Define for $k > 0$,

$$d_k = M(k,1,\mathbf{F}_q^*) - M(k,0,\mathbf{F}_q^*).$$

Then by (2.4) we have

$$qM(k,0,\mathbf{F}_q^*) = (q-1)_k - (q-1)d_k. \tag{2.6}$$

Heuristically, $M(k,0,\mathbf{F}_q^*)$ should be approximately $\frac{1}{q}(q-1)_k$. To obtain the explicit value of $M(k,0,\mathbf{F}_q^*)$, we only need to know d_k. For convenience we set $d_0 = -1$.

Lemma 2.4. *If d_k is defined as above, then*

$$d_k = \begin{cases} -1, & k = 0; \\ 1, & k = 1; \\ -kd_{k-1}, & p \nmid k, \ 2 \le k \le q-1 \ ; \\ (q-k)d_{k-1}, & p \mid k, \ 2 \le k \le q-1 \ . \end{cases}$$

Proof. One checks that $d_1 = M(1,1,\mathbf{F}_q^*) - M(1,0,\mathbf{F}_q^*) = 1 - 0 = 1$. When $p \nmid k$, by Lemma 2.3 we have $M(k,1,\mathbf{F}_q) = M(k,0,\mathbf{F}_q)$. This together with Lemma 2.2 implies

$$M(k,1,\mathbf{F}_q^*) - M(k,0,\mathbf{F}_q^*) = k(M(k-1,0,\mathbf{F}_q^*) - M(k-1,1,\mathbf{F}_q^*)).$$

Namely, $d_k = -kd_{k-1}$. When $p \mid k$, using Lemma 2.3 we have

$$M(k,1,\mathbf{F}_q) - M(k,0,\mathbf{F}_q) = q(M(k-1,1,\mathbf{F}_q^*) - M(k-1,0,\mathbf{F}_q^*)) = qd_{k-1}.$$

By Lemma 2.2, the left side is $d_k + kd_{k-1}$. Thus, $d_k = (q-k)d_{k-1}$. □

Corollary 2.5.

$$d_k = -(-1)^{k+\lfloor \frac{k}{p} \rfloor} k! \binom{\frac{q}{p} - 1}{\lfloor \frac{k}{p} \rfloor}$$

Proof. One checks $d_0 = -1$ and $d_1 = 1$ are consistent with the above formula for $k \leq 1$. Let $k \geq 2$ and write $k = np + m$ with $0 \leq m < p$. By Lemma 2.4,

$$\begin{aligned} \frac{d_k}{k!} &= (-1)^{n(p-1)+m+1} \prod_{i=1}^{n} \frac{(q-ip)}{ip} \\ &= (-1)^{n(p-1)+m+1} \frac{\prod_{i=1}^{n} (\frac{q}{p} - i)}{n!} \\ &= -(-1)^{k+n} \binom{\frac{q}{p} - 1}{n}. \end{aligned}$$

□

We are now ready to prove the explicit formula for $M(k,b,D)$ when $|D|$ equals q or $q-1$.

Theorem 2.6. *Let $\mathbf{F}_q$ be a finite field of characteristic p. Let D be a subset of $\mathbf{F}_q$. Let $M(k,b,D)$ be the number of solutions of (2.0). Then we have*

$$\begin{aligned} M(k,b,\mathbf{F}_q^*) &= \frac{(q-1)_k - v(b)d_k}{q}; \\ M(k,b,\mathbf{F}_q) &= \frac{(q)_k - v(b)(d_k + kd_{k-1})}{q}; \end{aligned}$$

where $v(b) = -1$ if $b \neq 0$ and $v(b) = q-1$ if $b = 0$.

Proof. If $b = 0$, by (2.6), we obtain

$$qM(k,0,\mathbf{F}_q^*) = (q-1)_k - (q-1)d_k.$$

If $b \neq 0$, then

$$qM(k,b,\mathbf{F}_q^*) = qM(k,1,\mathbf{F}_q^*) = qd_k + qM(k,0,\mathbf{F}_q^*) = (q-1)_k + d_k.$$

The formula for $M(k,b,\mathbf{F}_q^*)$ holds.

If $p \nmid k$, then $d_k + kd_{k-1} = 0$ and the formula for $M(k,b,\mathbf{F}_q)$ holds by Lemma 2.3.

If $p \mid k$, then $d_k + kd_{k-1} = qd_{k-1}$. By Lemma 2.3 and the above formula for $M(k,b,\mathbf{F}_q^*)$, we deduce

$$M(k,b,\mathbf{F}_q) = qM(k-1,b,\mathbf{F}_q^*) = (q-1)_{k-1} - v(b)d_{k-1}.$$

The formula for $M(k,b,\mathbf{F}_q)$ holds. □

Now we turn to deciding when the solution number $N(k,b,\mathbf{F}_q^*)>0$. A sequence $\{a_0,a_1,\cdots,a_n\}$ is **unimodal** if there exits indices k,r with $1\le k,r\le n$ such that

$$a_0\le a_1\le\cdots a_{k-1}\le a_k=a_{k+1}=\cdots=a_{k+r}\ge a_{k+r+1}\ge\cdots\ge a_n.$$

The sequence $\{a_0,a_1,\cdots,a_n\}$ is called symmetric if $a_i=a_{n-i}$ for $0\le i<n$.

Corollary 2.7. *For any $b\in\mathbf{F}_q$, both the sequence $N(k,b,\mathbf{F}_q)$ $(1\le k\le q)$ and the sequence $N(k,b,\mathbf{F}_q^*)$ $(1\le k\le q-1)$ are unimodal and symmetric.*

Proof. The symmetric part can be verified using Theorem 2.6. A simpler way is to use the relation

$$\sum_{a\in\mathbf{F}_q}a=\sum_{a\in\mathbf{F}_q^*}a=0.$$

To prove the unimodal property for $N(k,b,\mathbf{F}_q^*)$, by the symmetry it is sufficient to consider the case $k\le\frac{q-1}{2}$. Then, by Theorem 1.2, we deduce

$$q(N(k,0,\mathbf{F}_q^*)-N(k-1,0,\mathbf{F}_q^*))\ge(\binom{q-1}{k}-\binom{q-1}{k-1})-(q-1)(\binom{\frac{q}{p}-1}{\lfloor\frac{k}{p}\rfloor}-\binom{\frac{q}{p}-1}{\lfloor\frac{k-1}{p}\rfloor}).$$

If $p\nmid k$, then $\lfloor\frac{k}{p}\rfloor=\lfloor\frac{k-1}{p}\rfloor$ and the right side is clearly positive. If $p\mid k$, then the right side is

$$\begin{aligned}&=\frac{q-2k}{k}\binom{q-1}{k-1}-(q-1)\frac{\frac{q}{p}-\frac{2k}{p}}{\frac{k}{p}}\binom{\frac{q}{p}-1}{\frac{k}{p}-1}\\&=\frac{q-2k}{k}\left[\binom{q-1}{k-1}-(q-1)\binom{\frac{q}{p}-1}{\frac{k}{p}-1}\right]\\&\ge 0.\end{aligned}$$

The last inequality can be verified from the following Vandermonde's convolution

$$\binom{q-1}{k-1}=\sum_{i=0}^{\frac{q}{p}-1}\binom{\frac{q}{p}-1}{i}\binom{q-\frac{q}{p}}{k-1-i}>\binom{\frac{q}{p}-1}{\frac{k}{p}-1}\binom{q-\frac{q}{p}}{k-\frac{k}{p}}.$$

And so $N(k,0,\mathbf{F}_q^*)$ is unimodal. The proof for the unimodal property of $N(k,b,\mathbf{F}_q)$ is similar. This completes the proof.

□

Corollary 2.8. *Let $|D|=q-1>4$. If p is an odd prime then for $1<k<q-2$ the equation* (1.1) *always has a solution. If $p=2$, then for $2<k<q-3$ the equation* (1.1) *always has a solution.*

Proof. For any $a\in\mathbf{F}_q$ we have $N(k,b,\mathbf{F}_q\backslash\{a\})=N(k,b-ka,\mathbf{F}_q^*)$. Hence by Lemma 2.1 it is sufficent to consider $N(k,1,\mathbf{F}_q^*)$ and $N(k,0,\mathbf{F}_q^*)$. When p is odd and $k=2$ using Theorem 2.6 $N(2,0,\mathbf{F}_q^*)=\frac{1}{q}(\binom{q-1}{2}+(q-1))=\frac{q-1}{2}>0$, and $N(2,1,\mathbf{F}_q^*)=\frac{1}{q}(\binom{q-1}{2}-1)=\frac{q-3}{2}>0$. Then, by the unimodality of $N(k,1,\mathbf{F}_q^*)$ and $N(k,0,\mathbf{F}_q^*)$, for $1<k<q-2$, $N(k,b,\mathbf{F}_q\backslash\{a\})$ must be positive.

When $p=2$ and $k=3$, from Theorem 2.6, $N(3,0,\mathbf{F}_q^*)=\frac{1}{q}(\binom{q-1}{3}+(q-1)(\frac{q}{2}-1))=\frac{(q-1)(q-2)}{6}>0$ and $N(3,1,\mathbf{F}_q^*)=\frac{1}{q}(\binom{q-1}{3}-(\frac{q}{2}-1))=\frac{(q-2)(q-4)}{6}>0$. By the unimodality and symmetry we complete the proof. □

Corollary 2.9. *Let $D = \mathbf{F}_q$. If p is an odd prime then the equation (1.1) always has a solution if and only if $0 < k < q$. If $p = 2$, then for $2 < k < q-2$ the equation (1.1) always has a solution.*

Proof. It is straightforward from Corollary 2.8 and Theorem 2.6. □

3. The Case $|D| = q - c$

Let $D = \mathbf{F}_q \backslash \{a_1, a_2, \cdots a_c\}$ where $a_1, a_2, \cdots a_c$ are distinct elements in $\mathbf{F}_q$. In this section we first obtain the explicit formula of the solution number $N(k, b, D)$ for $c = 2$ and then give a general formula for $c > 2$. The solution number $N(k, b, \mathbf{F}_q \backslash \{a_1, a_2, \cdots a_c\})$ are closely related to the $\mathbf{F}_p$-linear relations among the set $\{a_1, \cdots a_c\}$ which we will see in Theorem 3.8. After that we obtain some simple bounds and we get a sufficient condition of k ensuring $N(k, b, D) > 0$. It can help us to investigate the subset sum problem over finite fields.

The following equality is an useful formula for the summand of sign-alternating binomial coefficients. It can be easily proved by comparing the coefficients of x^m in both sides of $(1-x)^{-1}(1-x)^r = (1-x)^{r-1}$. For other proofs and more details we refer to [3].

Lemma 3.1.

$$\sum_{k \le m} (-1)^k \binom{r}{k} = (-1)^m \binom{r-1}{m}. \tag{3.0}$$

Let $< k >_p$ denote the least non-negative residue of k modulo p. The following equality is the summand of the modular alternating binomial coefficients.

Lemma 3.2. *For any positive integers a, k we have*

$$\sum_{j=0}^{k} -(-1)^{\lfloor j/p \rfloor} \binom{a}{\lfloor j/p \rfloor} = -p(-1)^{\lfloor k/p \rfloor} \binom{a-1}{\lfloor k/p \rfloor} + (p - 1 - < k >_p)(-1)^{\lfloor k/p \rfloor} \binom{a}{\lfloor k/p \rfloor}$$

and so

$$\left| \sum_{j=0}^{k} -(-1)^{\lfloor j/p \rfloor} \binom{a}{\lfloor j/p \rfloor} \right| \le p \binom{a}{\lfloor k/p \rfloor}. \tag{3.1}$$

Proof. Let $j = n_j p + m_j$ with $0 \le m_j < p$. By using (3.0) we have

$$\begin{aligned}
&\sum_{j=0}^{k} -(-1)^{\lfloor \frac{j}{p} \rfloor} \binom{a}{\lfloor j/p \rfloor} \\
&= -p \sum_{n_j=0}^{n_k} (-1)^{n_j} \binom{a}{n_j} + (p - 1 - < k >_p)(-1)^{n_k} \binom{a}{n_k} \\
&= -p(-1)^{\lfloor k/p \rfloor} \binom{a-1}{\lfloor k/p \rfloor} + (p - 1 - < k >_p)(-1)^{\lfloor k/p \rfloor} \binom{a}{\lfloor k/p \rfloor}.
\end{aligned}$$

The inequality is obvious by noting the alternating signs before the two binomial coefficients. □

We omit the proofs of the following two lemmas for the first is direct from Lemma 3.2 and the proof of the second is similar to that of Lemma 3.2.

Lemma 3.3. *Define* $R_k^1 = (-1)^k d_k/k! = -(-1)^{\lfloor k/p \rfloor}\binom{q/p-1}{\lfloor k/p \rfloor}$ *and* $R_k^2 = \sum_{j=0}^{k} R_j^1$. *Let* $< k >_p$ *denote the least non-negative residue of* k *modulo* p. *Then we have*

$$R_k^2 = -p(-1)^{\lfloor k/p \rfloor}\binom{q/p-2}{\lfloor k/p \rfloor} + (p-1-< k >_p)(-1)^{\lfloor k/p \rfloor}\binom{q/p-1}{\lfloor k/p \rfloor}. \tag{3.2}$$

Lemma 3.4. *Let* $R_k^1 = -(-1)^{\lfloor k/p \rfloor}\binom{q/p-1}{\lfloor k/p \rfloor}$. *Let* $< k >_p$ *denote the least non-negative residue of* k *modulo* p. *Let* $b \in \mathbf{F}_p$. *Define* $\delta_{b,k} = 1$ *if* $< b >_p$ *is greater than* $< k >_p$ *and* $\delta_{b,k} = 0$ *otherwise. Then we have*

$$M(k,b) \triangleq \sum_{\substack{0 \le i \le k \\ i \equiv b(\mod p)}} R_i^1 = -(-1)^{\lfloor k/p \rfloor}\binom{q/p-2}{\lfloor k/p \rfloor} + \delta_{b,k}(-1)^{\lfloor k/p \rfloor}\binom{q/p-1}{\lfloor k/p \rfloor}. \tag{3.3}$$

For $b \notin \mathbf{F}_p$ we define $M(k,b) = 0$ for any nonnegative integer k. Note $M(k,b) \le \binom{q/p-2}{k}$. In the following theorem we give the accurate formula of $N(k,b,D)$ when $D = \mathbf{F}_q \backslash \{a_1, a_2\}$. Note that we can always assume $a_1 = 0$ and $a_2 = 1$ by a linear substitution.

Theorem 3.5. *Let* $\mathbf{F}_q$ *be a finite field of characteristic* p. *Let* $a_1 = 0, a_2 = 1 \in \mathbf{F}_q^*$. *Let* $N(k,b,D = \mathbf{F}_q \backslash \{a_1, a_2\})$ *be the number of solutions of (2.0). Then we have*

$$N(k,b,\mathbf{F}_q \backslash \{a_1,a_2\}) = \frac{1}{q}\binom{q-2}{k} + \frac{1}{q}(-1)^k R_k^2 - (-1)^k M(k,k-b), \tag{3.4}$$

where $R_k^2, M(k,b)$ *are defined as in (3.2) and (3.3).*

Proof. Use the simple inclusion-exclusion sieving method by considering whether a_2 appears in the solution of equation (1.1) we have

$$\begin{aligned}
&N(k,b,\mathbf{F}_q \backslash \{a_1,a_2\}) \\
&= N(k,b,\mathbf{F}_q \backslash \{a_1\}) - N(k-1,b-a_2,\mathbf{F}_q \backslash \{a_1,a_2\}) \\
&= N(k,b,\mathbf{F}_q \backslash \{a_1\}) - (N(k-1,b-a_2,\mathbf{F}_q \backslash \{a_1\}) \\
&\quad -N(k-2,b-2a_2,\mathbf{F}_q \backslash \{a_1,a_2\})) \\
&\quad \cdots\cdots \\
&= \sum_{i=0}^{k-1}(-1)^i N(k-i,b-ia_2,\mathbf{F}_q \backslash \{a_1\}) \\
&\quad +(-1)^k N(0,b-ka_2,\mathbf{F}_q \backslash \{a_1,a_2\}).
\end{aligned}$$

It is consistent with the above formula if we define $N(0,b,D)$ to be 1 if and only if $b = 0$ for a nonempty set D. Note $a_1 = 0$ and we have

$$N(k,b,\mathbf{F}_q \backslash \{a_1,a_2\}) = \sum_{i=0}^{k}(-1)^i N(k-i,b-ia_2,\mathbf{F}_q^*).$$

By Theorem 2.6 we have the following formula

$$N(k,b,\mathbf{F}_q^*) = \frac{1}{q}\binom{q-1}{k} - \frac{1}{q}(-1)^k v(b) R_k^1$$

where $R_k^1 = -(-1)^{\lfloor k/p \rfloor}\binom{q/p-1}{\lfloor k/p \rfloor}$, $v(b) = -1$ if $b \neq 0$ and $v(b) = q-1$ if $b = 0$. So

$$N(k,b,\mathbf{F}_q \backslash \{a_1,a_2\})$$

$$= \sum_{i=0}^{k}(-1)^i\Big[\frac{1}{q}\binom{q-1}{k-i} - \frac{1}{q}(-1)^{k-i}v(b-ia_2)R^1_{k-i}\Big].$$
$$= \frac{1}{q}((-1)^k\sum_{k-i=0}^{k}(-1)^{k-i}\binom{q-1}{k-i} - (-1)^k\sum_{k-i=0}^{k}v(b-ia_2)R^1_{k-i}$$
$$= \frac{1}{q}((-1)^k\sum_{j=0}^{k}(-1)^j\binom{q-1}{j} - (-1)^k\sum_{j=0}^{k}v(b-ka_2+ja_2)R^1_j$$
$$= \frac{1}{q}(\binom{q-2}{k} - (-1)^k\sum_{j=0}^{k}v(b-ka_2+ja_2)R^1_j).$$

The last equality is from (3.0). Note $a_2 = 1$ and by the definition of $v(b)$ we have

$$\begin{aligned} &N(k,b,\mathbf{F}_q\backslash\{a_1,a_2\}) \\ &= \frac{1}{q}\binom{q-2}{k} - \frac{1}{q}(-1)^k\sum_{j=0}^{k}v(b-k+j)R^1_j \qquad (3.5)\\ &= \frac{1}{q}\binom{q-2}{k} - \frac{1}{q}(-1)^k\sum_{j=0}^{k}(-1)\cdot R^1_j - \frac{1}{q}(-1)^k\sum_{\substack{0\le j\le k\\ b-k+j=0}} q\cdot R^1_j \\ &= \frac{1}{q}\binom{q-2}{k} + \frac{1}{q}(-1)^k R^2_k - (-1)^k\cdot M(b,k-b). \end{aligned}$$

□

Combining (3.4), (3.2) and (3.3) we obtain the following simple solution number formula compared with the formulas stated in Theorem 1.2.

Corollary 3.6. *If* $< k >_p = p-1$ *and* $b \in \mathbf{F}_p$ *then we have*

$$N(k,b,\mathbf{F}_q\backslash\{0,1\}) = \frac{1}{q}\binom{q-2}{k} + (-1)^{k+\lfloor k/p\rfloor}\frac{q-p}{q}\binom{q/p-2}{\lfloor k/p\rfloor}.$$

Now we turn to deriving a general formula for $c > 2$ using a similar method. For the purpose of further investigation on the solution number $N(k,b,D)$ first we have the following lemma.

Lemma 3.7. *Let* $R^1_k = -(-1)^{\lfloor k/p\rfloor}\binom{q/p-1}{\lfloor k/p\rfloor}$. *For* $c > 1$ *if we define recursively that* $R^c_k = \sum_{j=0}^{k} R^{c-1}_j$, *then we have*

$$R^c_k = -\sum_{j=0}^{k}(-1)^{\lfloor j/p\rfloor}\binom{k+c-2-j}{c-2}\binom{q/p-1}{\lfloor j/p\rfloor}. \qquad (3.6)$$

Proof. When $c = 2$, this formula is consistent with the definition of R^2_k. Assume it is true for some $c \ge 2$, then we have

$$R^{c+1}_k = \sum_{i=0}^{k} R^c_i$$

$$
\begin{aligned}
&= \sum_{i=0}^{k}(-1)\cdot\sum_{j=0}^{i}(-1)^{\lfloor j/p\rfloor}\binom{i+c-2-j}{c-2}\binom{q/p-1}{\lfloor j/p\rfloor}\\
&= -\sum_{j=0}^{k}\sum_{i=j}^{k}(-1)^{\lfloor j/p\rfloor}\binom{i+c-2-j}{c-2}\binom{q/p-1}{\lfloor j/p\rfloor}\\
&= -\sum_{j=0}^{k}(-1)^{\lfloor j/p\rfloor}\sum_{i=j}^{k}\binom{i+c-2-j}{c-2}\binom{q/p-1}{\lfloor j/p\rfloor}\\
&= -\sum_{j=0}^{k}(-1)^{\lfloor j/p\rfloor}\binom{k+c-1-j}{c-1}\binom{q/p-1}{\lfloor j/p\rfloor}.
\end{aligned}
$$

The last equality follows from the following simple binomial coefficient identity

$$\sum_{j\le k}\binom{j+n}{m}=\binom{k+n+1}{m+1}.$$

□

Now we consider the general case when c is a positive integer greater than 2. Note if $k>\frac{q-c}{2}$ then it is equivalent to consider

$$N(q-c-k,-b-\sum_{i=0}^{c}a_i,\mathbf{F}_q\backslash\{a_1,a_2,\cdots,a_c\}).$$

Hence we may always assume $k\le\frac{q-c}{2}$. Recall we have extended the definition of $M(k,b)$ in Lemma 3.3. That is, for $b\notin\mathbf{F}_p$, $M(k,b)=0$ for any integer k. For convenience of obtaining our bounds and more investigation we give two different types of formulas.

Theorem 3.8. *Let $D=\mathbf{F}_q\backslash\{a_1,a_2,\cdots,a_c\}$ and $c\ge 3$, where $a_1=0,a_2=1,a_3,\cdots,a_c$ are distinct elements in the finite field $\mathbf{F}_q$ of characteristic p. Define the integer valued function $v(b)=-1$ if $b\ne 0$ and $v(b)=q-1$ if $b=0$. Then for any $b\in\mathbf{F}_q$, we have the following formulas*

$$
\begin{aligned}
&N(k,b,\mathbf{F}_q\backslash\{a_1,a_2,\cdots,a_c\})\\
&=\frac{1}{q}\binom{q-c}{k}-\frac{1}{q}(-1)^k\cdot\\
&\sum_{i_1=0}^{k}\sum_{i_2=0}^{k-i_1}\cdots\sum_{i_{c-1}=0}^{k-i_1-\cdots-i_{c-2}}v(b-i_1a_c-\cdots-(k-\sum_{j=1}^{c-1}i_j)a_2)R_j^1 \qquad (3.7)\\
&=\frac{1}{q}\binom{q-c}{k}+\frac{1}{q}(-1)^kR_k^c-(-1)^k\cdot\\
&\sum_{i_1=0}^{k}\sum_{i_2=0}^{k-i_1}\cdots\sum_{i_{c-2}=0}^{k-i_1-\cdots-i_{c-3}}M(k-\sum_{j=1}^{c-2}i_j,k-\sum_{j=1}^{c-2}i_j-b+\sum_{j=1}^{c-2}i_ja_{c+1-j}) \qquad (3.8)
\end{aligned}
$$

where R_k^c is defined by (3.6), and $M(k,b)$ is defined by (3.3) if $b\in\mathbf{F}_p$ and otherwise $M(k,b)=0$. Moreover if $a_1=0$, and $b,a_2,\cdots,a_c$ are linear independent over $\mathbf{F}_p$

then we have

$$N(k,b,\mathbf{F}_q\backslash\{a_1,a_2,\cdots,a_c\}) = \frac{1}{q}\binom{q-c}{k} + \frac{1}{q}(-1)^k R_k^c. \tag{3.9}$$

Proof. Using the simple inclusion-exclusion sieving method we have

$$\begin{aligned}
&N(k,b,\mathbf{F}_q\backslash\{a_1,a_2,\cdots,a_c\})\\
&= N(k,b,\mathbf{F}_q\backslash\{a_1,a_2,\cdots,a_{c-1}\})\\
&\quad -N(k-1,b-a_c,\mathbf{F}_q\backslash\{a_1,a_2,\cdots,a_{c-1}\})\\
&\quad \cdots\cdots\\
&= \sum_{i=0}^{k}(-1)^i N(k-i,b-ia_c,\mathbf{F}_q\backslash\{a_1,a_2\cdots,a_{c-1}\}).
\end{aligned}$$

When $c=3$, by the formula given by Theorem 3.5 and note $a_2=1$ we have

$$\begin{aligned}
&N(k,b,\mathbf{F}_q\backslash\{a_1,a_2,a_3\})\\
&= \sum_{i=0}^{k}(-1)^i\left[\frac{1}{q}\binom{q-2}{k-i} + \frac{1}{q}(-1)^{k-i}R_{k-i}^2 - (-1)^{k-i}M(k-i,k-i-(b-ia_3))\right].\\
&= \frac{1}{q}\binom{q-3}{k} + \frac{1}{q}(-1)^k R_k^3 - (-1)^k\sum_{i=0}^{k}M(k-i,k-i-b+ia_3).
\end{aligned}$$

By induction, (3.8) follows for $c\geq 3$. Similarly, (3.7) follows from (3.5) and we omit the proof. If $b,a_2=1,\cdots,a_c$ are linear independent over $\mathbf{F}_p$ then first note $b\notin\mathbf{F}_p$. When $c=2$ by its extended definition we have $M(k,k-b)=0$. When $c>2$, by the independence of $b,a_2=1,\cdots,a_c$ we know $k-\sum_{j=1}^{c-2}i_j - b + \sum_{j=1}^{c-2}i_j a_{c+1-j}\notin\mathbf{F}_p$ for any index $(i_1,i_2,\cdots,i_{c-2})$ in the summation. Hence the summation in (3.8) always vanishes for any c and hence the proof is complete.

□

Now we have obtained the two formulas of the solution number $N(k,b,D)$. It suffices to evaluating R_k^c and the summation in (3.8). Unfortunately the formula given by (3.8) is extremely complicated when c is large. The **NP**-hardness of the subset sum problem indicates the hardness of precisely evaluating the summation in (3.8). In the following corollaries we deduce some bounds by using some combinatorial properties of the binomial coefficients.

Like the proof of Lemma 3.7, simple counting shows there are $\binom{k+c-1}{c-1}$ terms in the summation of (3.7). Note $v(b)\leq q-1$ and $|R_j^1|\leq\binom{q/p-1}{\lfloor k/p\rfloor}$. Hence we obtain the following simple bound directly.

Corollary 3.9.

$$\left|N(k,b,D) - \frac{1}{q}\binom{q-c}{k}\right| \leq \frac{q-1}{q}\binom{k+c-1}{c-1}\binom{q/p-1}{\lfloor k/p\rfloor}. \tag{3.10}$$

We now derive two improved bounds from (3.8) and (3.9). The first one is an improved bound of (3.10).

Lemma 3.10. *Let $p<q$. Let*

$$M_k^c = \sum_{i_1=0}^{k}\sum_{i_2=0}^{k-i_1}\cdots\sum_{i_{c-2}=0}^{k-i_1-\cdots-i_{c-3}} M(k-\sum_{j=1}^{c-2}i_j, b-\sum_{j=1}^{c-2}i_j a_{c+1-j}),$$

then we have

$$qM_k^c - R_k^c \le (q-p)\binom{k+c-2}{c-2}\binom{q/p-2}{\lfloor k/p\rfloor}. \tag{3.11}$$

Proof. By the definition of R_k^c and the proof of Lemma 3.7 we have

$$R_k^c = \sum_{i_1=0}^{k}\sum_{i_2=0}^{k-i_1}\cdots\sum_{i_{c-2}=0}^{k-i_1-\cdots-i_{c-3}} R^2(k-\sum_{j=1}^{c-2} i_j)$$

where $R^2(k) = R_k^2$. By (3.2) and (3.3) when $p < q$ it is easy to check that

$$R_k^2 - qM(k,b) \le (q-p)\binom{q/p-2}{\lfloor k/p\rfloor}$$

for any $b \in \mathbf{F}_q$ and hence (3.11) follows by noting that both the numbers of terms appear in the two summations of R_k^c and M_k^c are $\binom{k+c-2}{c-2}$. □

Corollary 3.11. *Let $D = \mathbf{F}_q\backslash\{a_1, a_2, \cdots, a_c\})$. If $p < q$ then we have*

$$\left|N(k,b,D) - \frac{1}{q}\binom{q-c}{k}\right| \le \frac{q-p}{q}\binom{k+c-2}{c-2}\binom{q/p-2}{\lfloor k/p\rfloor}. \tag{3.12}$$

Proof. By the notation M_k^c in Lemma we rewrite (3.8) as

$$N(k,b,\mathbf{F}_q\backslash\{a_1,a_2,\cdots,a_c\}) = \frac{1}{q}\binom{q-c}{k} + \frac{1}{q}(-1)^k(R_k^c - qM_k^c)$$

and then (3.12) follows from the bound given by (3.11). □

We will do more asymptotic analysis in the Corollary 3.20 by using the bound (3.12). Now we turn to consider the second bound at the special cases when $b, a_2, \cdots, a_c$ are linear independent over $\mathbf{F}_p$. By (3.9) it is sufficient to investigate R_k^c. Unfortunately, even though R_k^c can be written as sum of binomial coefficients, evaluating it precisely seems still nontrivial. To show this, by using some combinatorial identities by (3.6) we can easily have the following equality

$$R_k^c = -\sum_{j=0}^{\lfloor k/p\rfloor-1}(-1)^j\left[\binom{k+c-1-ip}{c-1} - \binom{k+c-1-ip-p}{c-1}\right]\binom{q/p-1}{j} +$$
$$\binom{< k >_p + c-1}{c-1}\binom{q/p-1}{\lfloor k/p\rfloor}. \tag{3.13}$$

It seems R_k^c may have no closed form. When $p \ge 3$ it has been proved that even the following simpler sum $\sum_{j=0}^{n}(-1)^j\binom{k+n-1-ip}{n-1}\binom{n}{j}$ has no closed form which means it cannot be expressed as a fixed number of hypergeometric terms ([6], p. 160). Now we try to give a bound of R_k^c by using some simple combinatorial analysis.

In section 2 we have defined the unimodality of a sequence. A stronger property than unimodality is logarithmic concavity. First recall that a function f on the real line is concave if whenever $x < y$ we have $f((x+y)/2) \ge (f(x)+f(y))/2$. Similarly, a sequence $a_0, a_1 \cdots, a_n$ of positive numbers is **log concave** if $\log a_i$ is a concave function of i which is to say that $(\log a_{i-1} + \log a_{i+1})/2 \le \log a_i$. If we exponentiate both sides of the above, to eliminate all of the logarithms, we find that the sequence is log concave if $a_{i-1}a_{i+1} \le a_i^2$. And by this definition we have the following Lemmas which will be used to evaluate R_k^c.

Lemma 3.12. *If the sequence $\{a_i\}$ is log concave then it must be unimodal.*

Lemma 3.13. *If two sequences $\{a_i\}$ and $\{b_i\}$ are log concave then $\{a_ib_i\}$ is log concave.*

Lemma 3.14. *If the sequence $\{a_i\}$ of nonnegative integers is unimodal, then $\sum_{i=0}^{k}(-1)^i a_i \le \max_{0\le i\le k} a_i$.*

Using the above three Lemmas now we can obtain a bound of R_k^c.

Lemma 3.15.

$$R_k^c \le p\cdot \max_{0\le j\le k}\binom{k+c-2-j}{c-2}\binom{q/p-1}{\lfloor j/p\rfloor}. \tag{3.14}$$

Proof. It is easy to check that both the two sequences $\binom{k+c-2-j}{c-2}$ and $\binom{q/p-1}{\lfloor j/p\rfloor}$ are log concave on j. So the sequence $a_j=\binom{k+c-2-j}{c-2}\binom{q/p-1}{\lfloor j/p\rfloor}$ is also log concave by Lemma 3.13 and hence unimodal by Lemma 3.12. Then we have

$$\begin{aligned} R_k^c &= -\sum_{j=0}^{k}(-1)^{\lfloor j/p\rfloor}a_j \\ &= -\sum_{i=0}^{\lfloor k/p\rfloor}(-1)^i a_{ip} - \cdots - \sum_{i=0}^{\lfloor k/p\rfloor}(-1)^i a_{ip+<k>_p}\cdots - \sum_{i=0}^{\lfloor k/p\rfloor-1}(-1)^i a_{ip+p-1}. \end{aligned}$$

And hence (3.14) follows Lemma 3.14 and the proof is complete. □

From (3.9) and (3.14) we then have the following improved bound.

Corollary 3.16. *Let $D=\mathbf{F}_q\backslash\{a_1,a_2,\cdots,a_c\}$). If $a_1=0$, and $b,a_2,\cdots,a_c$ are linear independent over $\mathbf{F}_p$ then we have the improved bound*

$$\left|N(k,b,D)-\frac{1}{q}\binom{q-c}{k}\right| \le \frac{p}{q}\max_{0\le j\le k}\binom{k+c-2-j}{c-2}\binom{q/p-1}{\lfloor j/p\rfloor}. \tag{3.15}$$

If one obtains better bounds of M_k^c then we can improve the bound given by (3.12). But it seems evaluating M_k^c is much more complicated than evaluating R_k^c. Let

$$I=\{[i_1,i_2,\cdots,i_{c-2}], 0\le i_t\le k-\sum_{j=1}^{t-1}i_j, 1\le t\le c-2:\ \ b-\sum_{j=1}^{c-2}i_j a_{c+1-j}\in\mathbf{F}_p\}.$$

Simple counting shows that $0\le |I|\le \binom{k+c-2}{c-2}$. In the proof of (3.12) we use the upper bound $|I|\le\binom{k+c-2}{c-2}$ and in the proof of (3.15) it is the exact case $|I|=0$. We can improve the above bound if we know more information about the cardinality of I which is determined by the set $b,a_2,\cdots,a_c$. For example, if we know more about the rank of the set $\{b,a_2,\cdots,a_c\}$ we can also improve the bound given by (3.12) but we will omit the details.

Recall the decision version of the above subset sum problem introduced in section 1 is to determine if $\sum_{k=1}^{n}N(k,b,D)>0$ where $n=q-c$. Let $p<q$. Using the bound (3.12) we have

$$\begin{aligned} N(b,D) &= \sum_{k=1}^{n}N(k,b,D) \\ &= \sum_{k=1}^{n}\frac{1}{q}\binom{n}{k}-\sum_{k=1}^{n}\frac{q-p}{q}\binom{k+q-n-2}{q-n-2}\binom{q/p-2}{\lfloor k/p\rfloor} \end{aligned}$$

$$> \sum_{k=1}^{n} \frac{1}{q}\binom{n}{k} - \sum_{k=1}^{n} q^{q-n-2}\binom{q/p-2}{\lfloor k/p \rfloor}$$
$$> \frac{1}{q}\big(2^n - p\cdot q^{q-n-2}2^{q/p-2}\big).$$

The bound $N(b,D) > \frac{1}{q}\big(2^n - p\cdot q^{q-n-2}2^{q/p-2}\big)$ shows when $q-n$ is small compared to q then the answer of the subset problem over $\mathbf{F}_q$ is positive. For example, simple asymptotic analysis shows there is a constant C such that if $q-n = C\cdot(\frac{q\ln 2}{\ln q})$ then the answer of the subset problem is positive when q is large.

Example 3.17. *Choose $p=2, q=128, c=4$ and $k=5$. Let ω be a primitive element in F_{128}. When we choose $D = F_{128} - \{0,\omega,\omega^2,\omega^3\}$ and $b=1$ computations show there are $N = 1759038$ solutions of the equation (1.1) compared with the average number $\frac{1}{q}\binom{q-c}{k} \approx 1758985$. The bound given by (3.15) shows $1758893 < N < 1769076$.*

Our main task is to find the value of k to ensure the solutions number $N(k,b,D)$ to be positive where $D = \mathbf{F}_q\backslash\{a_1,a_2,\cdots,a_c\}$. It is easy to see that when c is a constant and $c << q$ the bound (3.12) is asymptotic to be the real value. In fact it tells us that for $2 < k < q-c-2$ the linear equation (1.1) always have solutions for a enough large extension $\mathbf{F}_q$ of $\mathbf{F}_p$ by some simple asymptotic analysis. Now we shall give a more precise bound of k to ensure $N(k,b,D) > 0$ for given p,q,c.

The (complete) gamma function is defined to be an extension of the factorial to complex and real number arguments. It is related to the factorial by $\Gamma(n) = (n-1)!$ and it has the analytical integral form

$$\Gamma(z) = \int_0^\infty t^{z-1}e^{-t}dt.$$

And the classic binomial coefficient $\binom{n}{m}$ then can be generalized to positive real numbers as

$$\binom{a}{b} = \frac{\Gamma(a+1)}{\Gamma(b+1)\Gamma(a-b+1)}, \quad a \geq b \geq 0.$$

This generalized binomial coefficient still has many like properties of the classic binomial coefficient such as the complementation rule, the Pascal' recursion and the following lemma directly obtained from the following asymptotic expansion for the complex function $\Gamma(z)$

$$\Gamma(z) \sim z^z e^{-z}(2\pi z)^{1/2}\left\{1 + \frac{1}{12z} + \frac{1}{288z^2} - \frac{139}{51840z^3} - \cdots\right\}.$$

Lemma 3.18. *Let n,m be positive real numbers. If $1 \leq m \leq \frac{n}{2}$ then*

$$e^{-\frac{1}{6m}}\frac{1}{\sqrt{2\pi}}\left(\frac{n}{m}\right)^m\left(\frac{n}{n-m}\right)^{n-m}\left(\frac{n}{m(n-m)}\right)^{\frac{1}{2}} \leq \binom{n}{m}$$

$$\leq \frac{1}{\sqrt{2\pi}}\left(\frac{n}{m}\right)^m\left(\frac{n}{n-m}\right)^{n-m}\left(\frac{n}{m(n-m)}\right)^{\frac{1}{2}}. \tag{3.16}$$

We also have the following simple bound of the binomial coefficient by the simple inequality $n! \geq (\frac{n}{e})^n\sqrt{2\pi n}$.

Lemma 3.19. *Let m, n be positive integers with $m \leq n$. Then we have*

$$\binom{n}{m} \leq (\frac{en}{m})^m. \tag{3.17}$$

By using the above inequalities and some basic calculations now we give a sufficient condition to ensure the solution number $N(k, b, D)$ being positive.

Corollary 3.20. *Let $q > p$ and $c \geq 2$. Let $|D| = q - c$. If*

$$\frac{q-c}{2}(1 - \sqrt{2f(c) - 1}) \leq k \leq \frac{q-c}{2}(1 + \sqrt{2f(c) - 1}) \tag{3.18}$$

where

$$f(c) = \left(1.087q(\frac{eq + ec}{2c-4})^{c-2}\right)^{-\frac{p}{(p-1)(q-c)}}$$

if $c \leq 2p$ and

$$f(c) = \left(1.087qe^{\frac{c-2p}{p}}(\frac{eq + ec}{2c-4})^{c-2}\right)^{-\frac{p}{(p-1)(q-c)}}$$

if $c > 2p$, then $N(k, b, D)$ is positive and hence the equation (1.1) always has solutions.

Proof. For simplicity write $q - c - k$ to be m. Using (3.12) and (3.17) we have

$$\begin{aligned}
&N(k, b, D) \\
&\geq \frac{1}{q}\binom{q-c}{k} - \frac{q-p}{q}\binom{k+c-2}{c-2}\binom{\lfloor \frac{q-2p}{p} \rfloor}{\lfloor \frac{k}{p} \rfloor} \\
&\geq \frac{1}{q}\binom{q-c}{k} - \frac{q-p}{q}\binom{k+c-2}{c-2}\binom{\frac{q-2p}{p}}{\frac{k}{p}} \\
&\geq \frac{1}{q}e^{-\frac{1}{6k}}\frac{1}{\sqrt{2\pi}}\left(\frac{q-c}{k}\right)^k\left(\frac{q-c}{m}\right)^m\left(\frac{q}{km}\right)^{\frac{1}{2}} \\
&-\frac{q-p}{q}\binom{k+c-2}{c-2}\frac{1}{\sqrt{2\pi}}\left(\frac{q-c}{k}\right)^{\frac{k}{p}}\left(\frac{q-c}{m}\right)^{\frac{m}{p}}\left(\frac{q}{km}\right)^{\frac{1}{2}}(\frac{q-2p}{q-c})^{\frac{q-c}{p}} \\
&\geq (2\pi qkm)^{-\frac{1}{2}}\left[e^{-\frac{1}{6k}}\left(\frac{q-c}{k}\right)^k\left(\frac{q-c}{m}\right)^m\right. \\
&\left.-q(\frac{q-2p}{q-c})^{\frac{q-c}{p}}\binom{k+c-2}{c-2}\left(\frac{q-c}{k}\right)^{\frac{k}{p}}\left(\frac{q-c}{m}\right)^{\frac{m}{p}}\right].
\end{aligned}$$

To ensure $N(k, b, D) > 0$ it is sufficient to consider when

$$e^{-\frac{1}{6k}}\left(\frac{q-c}{k}\right)^k\left(\frac{q-c}{m}\right)^m - q(\frac{q-2p}{q-c})^{\frac{q-c}{p}}\binom{k+c-2}{c-2}\left(\frac{q-c}{k}\right)^{\frac{k}{p}}\left(\frac{q-c}{m}\right)^{\frac{m}{p}} > 0$$

and it is reduced to the following inequality

$$\left(\frac{q-c}{k}\right)^k\left(\frac{q-c}{m}\right)^m > \left[e^{\frac{1}{6k}}q(\frac{q-2p}{q-c})^{\frac{q-c}{p}}\binom{k+c-2}{c-2}\right]^{\frac{p}{p-1}}.$$

By the following well known inequality

$$\frac{1}{\frac{1}{a}\alpha + (1-\alpha)\frac{1}{b}} \leq a^\alpha b^{1-\alpha},\ a, b > 0,\ 0 < \alpha < 1$$

we have

$$\left(\frac{q-c}{k}\right)^k\left(\frac{q-c}{m}\right)^m \geq \left[\frac{(q-c)^2}{k^2+m^2}\right]^{q-c}.$$

So it suffices to consider

$$\left[\frac{(q-c)^2}{k^2+m^2}\right]^{q-c} \geq \left[e^{\frac{1}{6k}}q(\frac{q-2p}{q-c})^{\frac{q-c}{p}}\binom{k+c-2}{c-2}\right]^{\frac{p}{p-1}}. \tag{3.19}$$

By (3.17) and assume $k \leq \frac{q-c}{2}$ then we have

$$\binom{k+c-2}{c-2} < \left[\frac{e(k+c-2)}{c-2}\right]^{c-2} < \left[\frac{eq+ec}{2c-4}\right]^{c-2}$$

and assume $k > 1$, so $e^{\frac{1}{6k}} \leq e^{\frac{1}{12}} < 1.087$. If $c \leq 2p$ then $(\frac{q-2p}{q-c})^{\frac{q-c}{p}} \leq 1$ and it is sufficient to consider

$$[\frac{(q-c)^2}{k^2+m^2}]^{q-c} \geq \big(1.087q(\frac{eq+ec}{2c-4})^{c-2}\big)^{\frac{p}{p-1}}$$

and that is

$$\frac{(q-c)^2}{k^2+m^2} \geq \big(1.087q[\frac{eq+ec}{2c-4}]^{c-2}\big)^{\frac{p}{(p-1)(q-c)}}.$$

Denote the reciprocal of the right side of the above inequality to be $f(c)$ then we obtain an quadratic equation of k

$$k^2+(q-c-k)^2 \leq (q-c)^2 f(c).$$

It is easy to solve out

$$\frac{q-c}{2}(1-\sqrt{2f(c)-1}) \leq k \leq \frac{q-c}{2}(1+\sqrt{2f(c)-1})$$

where $f(c) = \big(1.087q(\frac{eq+ec}{2c-4})^{c-2}\big)^{-\frac{p}{(p-1)(q-c)}}$. Note that the condition $k \leq \frac{q-c}{2}$ can be removed.

If $c > 2p$ then by the inequality $(1+1/x)^x < e = 2.2.71828\cdots$ and we have $(\frac{q-2p}{q-c})^{\frac{q-c}{p}} \leq e^{\frac{c-2p}{p}}$ and from (3.19) it suffices to consider

$$[\frac{(q-c)^2}{k^2+m^2}]^{q-c} \geq \big[1.087qe^{\frac{c-2p}{p}}(\frac{eq+ec}{2c-4})^{c-2}\big]^{\frac{p}{p-1}}.$$

And then we have

$$\frac{q-c}{2}(1-\sqrt{2g(c)-1}) \leq k \leq \frac{q-c}{2}(1+\sqrt{2g(c)-1})$$

where $g(c) = \big(1.087qe^{\frac{c-2p}{p}}(\frac{eq+ec}{2c-4})^{c-2}\big)^{-\frac{p}{(p-1)(q-c)}}$. And the proof is complete. □

From the above Corollary we conclude that when c is a constant then if q is large enough then $N(k,b,D) > 0$ for almost all k satisfying $2 < k < q-c$.

4. Some related additive problems

In Corollary 3.11 we obtain a bound when $p < q$. When $q = p$ we have the following bound.

Theorem 4.1. *Let $q = p$. Then we have*

$$\left|N(k,b,D) - \frac{1}{p}\binom{p-c}{k} + \frac{(-1)^k}{p}\binom{k+c-1}{c-1}\right| \le \binom{k+c-2}{c-2}. \tag{4.0}$$

Proof. When $q = p$, by (3.6) we have

$$R_k^c = -\sum_{j=0}^{k}\binom{k+c-2-j}{c-2} = -\binom{k+c-1}{c-1}$$

and note that when p=q, $M(k,b)$ equals 0 or -1 by its definition given in Lemma 3.4. Then by (3.8) we have

$$N(k,b,D) = \frac{\binom{p-c}{k} - (-1)^k\binom{k+c-1}{k}}{p} + (-1)^k M_k^c \tag{4.1}$$

with $0 \le M_k^c \le \binom{k+c-2}{k}$ and hence the proof is complete. □

We shall use the above bound to give a partial result in additive number theory. Given $D \in \mathbf{F}_p$ define

$$2^D = \{a+b : a,b \in D, a \ne b\}.$$

A natural question in additive number theory is how we can obtain any information on the cardinality of 2^D just from the cardinality of D. In particular, when does 2^D equal $\mathbf{F}_p$? Using (4.2) we may obtain something about this problem.

When $k = 2$, by (4.0) of Theorem 4.1 it is easy to check if $|D| \ge \frac{p+3}{2}$ then $N(2,b,D) > 0$ for any $b \in \mathbf{F}_p$. This means $2^D = \mathbf{F}_p$ if $|D| \ge \frac{p+3}{2}$. This agrees with the following famous theorem, conjectured by Erdős and Heilbronn in 1964. Special cases of this conjecture have been proved by various researchers. The full conjecture was proved by Dias da Sila and Hamidoune in 1994, using some tools from linear algebra and the representation theory of the symmetric group. An elementary proof using polynomial method was given by Alon in 1999 [4].

Theorem 4.2. *If p is a prime, and D is a nonempty subset of $\mathbf{F}_p$, then*

$$\left|\{a+b : a,b \in D, a \ne b\}\right| \ge \min\{p, 2|D| - 3\}.$$

The above theorem was generalized by Dias da Sila and Hamidoune [4]:

Theorem 4.3. *If p is a prime, and D is a nonempty subset of $\mathbf{F}_p$. Let k^D denote the set of all sums of k distinct elements of D. Then*

$$|k^D| \ge \min\{p, k|D| - k^2 + 1\}.$$

By Theorem 4.3, if $|D| \ge k + \frac{p-1}{k}$ then $k^D = \mathbf{F}_p$ which means that the set of the sum of all $k-$ subsets of D covers $\mathbf{F}_p$. Using (4.1) we can obtain another bound. If k is even, and if $|D| \ge \frac{p+k}{2}$ then $k^D = \mathbf{F}_p$. When $k = 2$ the two bounds are identical. The second bound is far weaker than the first one when k is large. Though it does give some information about the number of ways to write an elements in $\mathbf{F}_p$ to be the sum of k distinct elements in D. And if we can improve the estimate on M_k^c in (4.1) then better results will follows.

The above bound in Theorem 4.3 is also correct if $\mathbf{F}_p$ is replaced by $\mathbf{F}_q$ (see [5], Page 98). Let us recite it again. If D is a nonempty subset of $\mathbf{F}_q$. Let k^D denote the set of all sums of k distinct elements of D. Then

$$|k^D| \geq \min\{p, k|D| - k^2 + 1\}.$$

But when p is small compared to q and k, this bound is very weak. It would be interesting to get better bound in this case.

5. Applications to Reed-Solomon Codes

Let $D = \{x_1, \cdots, x_n\} \subset \mathbf{F}_q$ be a subset of cardinality $|D| = n > 0$. For $1 \leq k \leq n$, the Reed-Solomon code $D_{n,k}$ has the codewords of the form

$$(f(x_1), \cdots, f(x_n)) \in \mathbf{F}_q^n,$$

where f runs over all polynomials in $\mathbf{F}_q[x]$ of degree at most $k-1$. The minimum distance of the Reed-Solomon code is $n-k+1$ because a non-zero polynomial of degree at most $k-1$ has at most $k-1$ zeroes. For $u = (u_1, u_2, \cdots, u_n) \in \mathbf{F}_q^n$, we can associate a unique polynomial $u(x) \in \mathbf{F}_q[x]$ of degree at most $n-1$ such that

$$u(x_i) = u_i,$$

for all $1 \leq i \leq n$. The polynomial $u(x)$ can be computed quickly by solving the above linear system. Explicitly, the polynomial $u(x)$ is given by the Lagrange interpolation formula

$$u(x) = \sum_{i=1}^{n} u_i \frac{\prod_{j \neq i}(x - x_j)}{\prod_{j \neq i}(x_i - x_j)}.$$

Define $d(u)$ to be the degree of the associated polynomial $u(x)$ of u. It is easy to see that u is a codeword if and only if $d(u) \leq k-1$.

For a given $u \in \mathbf{F}_q^n$, define

$$d(u, D_{n,k}) := \min_{v \in D_{n,k}} d(u, v).$$

The maximun likelihood decoding of u is to find a codeword $v \in D_{n,k}$ such that $d(u,v) = d(u, D_{n,k})$. Thus, computing $d(u, D_{n,k})$ is essentially the decision version for the maximun likelihood decoding problem, which is **NP**-complete for general subset $D \subset \mathbf{F}_q$. For standard Reed-Solomon code with $D = \mathbf{F}_q^*$ or $\mathbf{F}_q$, the complexity of the maximun likelihood decoding is unknown to be **NP**-complete. This is an important open problem. It has been shown by Cheng-Wan [2] to be at least as hard as the discrete logarithm problem.

When $d(u) \leq k-1$, then u is a codeword and thus $d(u, D_{n,k}) = 0$. We shall assume that $k \leq d(u) \leq n-1$. The following simple result gives an elementary bound for $d(u, D_{n,k})$.

Theorem 5.1. *Let $u \in \mathbf{F}_q^n$ be a word such that $k \leq d(u) \leq n-1$. Then,*

$$n - k \geq d(u, D_{n,k}) \geq n - d(u).$$

Proof. Let $v = (v(x_1), \cdots, v(x_n))$ be a codeword of $D_{n,k}$, where $v(x)$ is a polynomial in $\mathbf{F}_q[x]$ of degree at most $k-1$. Then,

$$d(u, v) = n - N_D(u(x) - v(x)),$$

where $N_D(u(x) - v(x))$ denotes the number of zeros of the polynomial $u(x) - v(x)$ in D. Thus,

$$d(u, D_{n,k}) = n - \max_{v \in D_{n,k}} N_D(u(x) - v(x).$$

Now $u(x) - v(x)$ is a polynomial of degree equal to $d(u)$. We deduce that

$$N_D(u(x) - v(x)) \leq d(u).$$

It follows that

$$d(u, D_{n,k}) \geq n - d(u).$$

The lower bound is proved. To prove the upper bound, we choose a subset $\{x_1, \cdots, x_k\}$ in D and let $g(x) = (x - x_1) \cdots (x - x_k)$. Write

$$u(x) = g(x)h(x) + v(x),$$

where $v(x) \in \mathbf{F}_q[x]$ has degree at most $k - 1$. Then, clearly, $N_D(u(x) - v(x)) \geq k$. Thus

$$d(u, D_{n,k}) \leq n - k.$$

The theorem is proved.

We call u to be a deep hole if $d(u, D_{n,k}) = n - k$, that is, the upper bound in the equality holds. When $d(u) = k$, the upper bound agrees with the lower bound and thus u must be a deep hole. This gives $(q - 1)q^k$ deep holes. For a general Reed-Solomon code $D_{n,k}$, it is already difficult to determine if a given word u is a deep hole. In the special case that $d(u) = k+1$, the deep hole problem is equivalent to the the subset sum problem over $\mathbf{F}_q$ which is **NP**-complete if $p > 2$.

For the standard Reed-Solomon code, that is, $D = \mathbf{F}_q^*$ and thus $n = q - 1$, there is the following interesting conjecture of Cheng-Murray [1].

Conjecture Let $q = p$. For the standard Reed-Solomon code with $D = \mathbf{F}_p^*$, the set $\{u \in \mathbf{F}_p^n \big| d(u) = k\}$ gives the set of all deep holes.

Using the Weil bound, Cheng and Murray proved that their conjecture is true if p is sufficiently large compared to k.

The deep hole problem is to determine when the upper bound in the above theorem agrees with $d(u, D_{n,k})$. We now examine when the lower bound $n - d(u)$ agrees with $d(u, D_{n,k})$. It turns out that the lower bound agrees with $d(u, D_{n,k})$ much more often. Fix $0 \leq m \leq n - 1 - k$, we call u an ordinary word of degree $k + m$ if $d(u) = k + m$ and $d(u, D_{k,n}) = n - (k + m)$. A basic problem is then to determine for a given word u of degree $k + m$, when u is ordinary of degree $k + m$.

Without loss of generality, we can assume that $u(x)$ is monic. Let

$$u(x) = x^{k+m} - b_1 x^{k+m-1} + \cdots + (-1)^m b_m x^k + \cdots + (-1)^{k+m} b_{k+m}$$

be a monic polynomial in $\mathbf{F}_q[x]$ of degree $k + m$. By definition, $d(u, D_{n,k}) = n-(k+m)$ if and only if there is a polynomial $v(x) \in \mathbf{F}_q[x]$ of degree at most $k-1$ such that

$$u(x) - v(x) = (x - x_1) \cdots (x - x_{k+m}),$$

with $x_i \in D$ being distinct. This is true if and only if the system

$$\sum_{i=1}^{k+m} X_i = b_1,$$

$$\sum_{1\le i_1<i_2\le k+m} X_{i_1}X_{i_2} = b_2,$$
$$\cdots,$$
$$\sum_{1\le i_1<i_2<\cdots<i_m\le k+m} X_{i_1}\cdots X_{i_m} = b_m.$$

has distinct solutions $x_i \in D$. This explains our motivational problem in the introduction section.

When $d(u) = k$, then u is always a deep hole. The next non-trivial case is when $d(u) = k + 1$. Using the bound in Theorem 3.9, we obtain some positive results related to the deep hole problem in the case $d(u) = k + 1$ (i.e., the case $m = 1$) if $q - n$ is small.

Corollary 5.2. *Let $n = q-c$. Let $d(u) = k+1$. If k satisfy the following inequality*

$$\frac{q-c}{2}(1-\sqrt{2f(c)-1}) \le k+1 \le \frac{q-c}{2}(1+\sqrt{2f(c)-1}) \qquad 5.1$$

where $f(c)$ is defined in Corollary 3.20. Then u can not be a deep hole.

Proof. By the above discussion u is not a deep hole if and only if the equation

$$x_1 + x_2 + \cdots + x_{k+1} = b_1$$

always has distinct solutions in D for any $b_1 \in F_q$. Using Corollary 3.20 we obtain the result immediately. □

For fixed p and c when q tends to infinity simple calculations shows that the condition given by (5.1) becomes $\ln q/2 \le k \le q - 2 - \ln q/2$.

By Corollary 2.8 we still have following simple consequence.

Corollary 5.3. *Let $n > q - 2$. Let $d(u) = k + 1$ and $2 < k < q - 3$. If $q > 5$ then u can not be a deep hole.*

In the present paper, we studied the case $m = 1$ and explored some of the combinatorial aspects of the problem. In a future article, we plan to study the case $m > 1$ by combining the ideas of the present papers with algebraic-geometric techniques such as the Weil bound.

References

[1] Q. Cheng and E. Murray, *On deciding deep holes of Reed-Solomon codes*, TAMS 2007, to appear.

[2] Q. Cheng and D. Wan, *On the list and Bounded distance Decodibility of Reed-Solomon Codes*, FOCS (2004), 335-341.

[3] R. Graham, D. Knuth and and D. Patashnik, *Concrete Mathematics: A Foundation for Computer Science*, 2nd ed., Reading MA: Addison-Wesley, 1994.

[4] N. Alon, *Combinatorial Nullstellensatz*, Combin. Probab. Comput, **8** (1999), 7-29.

[5] M. B. Nathanson, *Additive Number Theory, Inverse Problems and the Geometry of Sumsets*, Springer, 1996.

[6] M. Petkovsek, H. S. Wilf and D. Zeilberger, *A=B*, Wellesley, MA:A. K. Peters, 1996.

School of Mathematical Sciences, Peking University, Beijing, P.R. China
E-mail address: joe@math.pku.edu.cn

Department of Mathematics, University of California, Irvine, CA 92697-3875, USA
E-mail address: dwan@math.uci.edu