

# Deep Holes in Reed-Solomon Codes Based on Dickson Polynomials

Matt Ketı, Daqing Wan

*Department of Mathematics, University of California Irvine, CA 92697-3875, USA*

---

## Abstract

For an  $[n, k]$  Reed-Solomon code  $\mathcal{C}$ , it can be shown that any received word  $r$  lies a distance at most  $n - k$  from  $\mathcal{C}$ , denoted  $d(r, \mathcal{C}) \leq n - k$ . Any word  $r$  meeting the equality is called a deep hole. Guruswami and Vardy (2005) showed that for a specific class of codes, determining whether or not a word is a deep hole is NP-hard. They suggested passingly that it may be easier when the evaluation set of  $\mathcal{C}$  is large or structured. Following this idea, we study the case where the evaluation set is the image of a Dickson polynomial, whose values appear with a special uniformity. To find families of received words that are not deep holes, we reduce to a subset sum problem (or equivalently, a Dickson polynomial-variation of Waring's problem) and find solution conditions by applying an argument using estimates on character sums indexed over the evaluation set.

*Keywords:* character sum, deep hole, Dickson polynomial, Reed-Solomon code, subset sum, Waring's problem

---

## 1. Introduction

Reed-Solomon error-correcting codes are used routinely in technological applications when there is a risk for transmitted data to be lost or corrupt. The classical set-up fixes a finite field  $\mathbb{F}_q$ , message block length  $k$ , and subset of  $\mathbb{F}_q$  of size  $n > k$  denoted  $D = \{x_1, x_2, \dots, x_n\}$ .  $D$  is often referred to as the evaluation set with typical choices  $D = \mathbb{F}_q$  or  $\mathbb{F}_q^*$ . A message  $(m_0, m_1, \dots, m_{k-1})$  is represented by the polynomial  $m(x) = m_0 + m_1x + \dots + m_{k-1}x^{k-1}$ . The message is encoded by calculating  $(m(x_1), m(x_2), \dots, m(x_n))$ , called a codeword. The set of all possible encoded messages is defined as the codebook and is denoted by  $\mathcal{C}$ .

Let  $a = (a_1, a_2, \dots, a_n)$  and  $b = (b_1, b_2, \dots, b_n)$  be two words. Define the Hamming distance  $d(a, b)$  as the number of coordinates in which  $a$  and  $b$  differ. The distance between a word  $u$  and the codebook  $\mathcal{C}$  is defined as  $d(u, \mathcal{C}) =$

---

*Email addresses:* [mketi@uci.edu](mailto:mketi@uci.edu) (Matt Ketı), [dwan@math.uci.edu](mailto:dwan@math.uci.edu) (Daqing Wan)

$\min_{v \in \mathcal{C}} d(u, v)$ . It is well known that for Reed-Solomon codes,  $d(u, \mathcal{C}) \leq n - k$  for any received word  $u$ . In studying the error-correcting capacity for Reed-Solomon codes, Guruswami and Vardy in [8] found that for a special family of codes with a small evaluation set, determining whether or not  $d(u, \mathcal{C}) = n - k$  for a given  $u$  is NP-hard. They called any word  $u$  satisfying the equality a deep hole and suggested in passing that finding deep holes might be easier when the evaluation set is large. We will further investigate the problem of finding deep holes.

## 2. Overview of Previous Work

One way to measure  $d(u, \mathcal{C})$  is to run Lagrange Interpolation on the word  $u = (u_1, \dots, u_n)$  to get a fitted polynomial  $u(x)$  satisfying  $u(x_i) = u_i$  for all  $1 \leq i \leq n$ . Then, if  $\deg u(x) \leq k - 1$ , then  $u$  is a codeword and  $d(u, \mathcal{C}) = 0$ . Otherwise,  $k \leq \deg u(x) \leq n - 1$ , and Li and Wan in [13] gave the bound

$$n - \deg u(x) \leq d(u, \mathcal{C}) \leq n - k$$

which shows that if  $\deg u(x) = k$ , then  $u$  is automatically a deep hole. Many of the results toward the deep hole problem are geared toward examining families of words by degree.

### 2.1. For $D = \mathbb{F}_q$

The choice of  $D = \mathbb{F}_q$  is referred to as a standard Reed-Solomon code. Cheng and Murray in [3] searched for deep holes in this context, and conjectured that the only deep holes were those satisfying  $\deg u(x) = k$ . More precisely,

**Conjecture** (Cheng-Murray). *All deep holes for standard Reed-Solomon codes are those words  $u$  satisfying  $\deg u(x) = k$ .*

They weren't able to prove this, but they were able to reduce the problem to finding a rational point on an algebraic hypersurface to derive the first result on deep holes for Reed-Solomon codes over prime field  $\mathbb{F}_p$ :

**Theorem 1** (Cheng-Murray). *Let  $p$  be a prime and  $1 < k < p^{1/4-\varepsilon}$  be a positive integer. Consider the standard Reed-Solomon code  $\mathcal{C}_p$ , and let  $u$  be a received word and  $u(x)$  be its interpolated polynomial. If the degree of  $u(x)$  satisfies*

$$k < \deg u(x) < k + p^{3/13-\varepsilon}$$

*then  $u$  is not a deep hole.*

In one of the newest papers, Cheng, Li, and Zhuang [4] were able to resolve in some cases the conjecture over  $\mathbb{F}_p$  using the concept of deep hole trees.

**Theorem 2** (Cheng-Li-Zhuang). *Let  $p > 2$  be a prime number,  $k \geq \frac{p-1}{2}$ ,  $D = \{\alpha_1, \alpha_2, \dots, \alpha_n\}$  with  $k < n \leq p$ . The only deep holes of  $\mathcal{C}_p$  are generated by functions which are equivalent to the following:*

$$f(x) = x^k, \quad f_\delta(x) = \frac{1}{x - \delta}$$

where  $\delta \in \mathbb{F}_p \setminus D$ . Here, two functions  $f(x)$  and  $g(x)$  are equivalent if and only if there exists  $a \in \mathbb{F}_p^*$  and  $h(x)$  with degree less than  $k$  such that

$$g(x) = af(x) + h(x)$$

Over a general finite field  $\mathbb{F}_q$ , they also showed

**Theorem 3.** *Given a finite field  $\mathbb{F}_q$  with characteristic  $p > 2$ , if  $k + 1 \leq p$  or  $3 \leq q - p + 1 \leq k + 1 \leq q - 2$ , then the Cheng-Murray conjecture is true.*

Further evidence in favour of the conjecture has been provided by Li-Wan [13], Liao [11], Zhu-Wan [20], using polynomial congruences and estimates of character sums, Cafure-Matera-Privitelli [1] using algebraic geometry, and Li-Zhu [14] by explicitly counting solutions to certain polynomial equations. To list the one by Li-Zhu,

**Theorem 4** (Li-Zhu). *Let  $\mathcal{C}_q$  be a standard Reed-Solomon code, and  $u$  a received word represented by the polynomial  $u(x) = x^{k+1} - bx^k + v(x)$ , where  $\deg(v(x)) \leq k - 1$ . Then*

$$d(u, \mathcal{C}_q) = \begin{cases} q - k & b = 0, p = 2, k = 1 \text{ or } q - 3 \\ q - k - 1 & \text{otherwise} \end{cases}$$

They also proved a similar distance formula for words of the form  $u(x) = x^{k+2} - bx^{k+1} + cx^k + v(x)$ , showing that such words are not deep holes either.

2.2. For  $D = \mathbb{F}_q^*$

The choice of  $D = \mathbb{F}_q^*$  is referred to as a primitive Reed-Solomon code. Wu and Hong studied this in [18] and showed

**Theorem 5** (Wu-Hong). *Let  $\mathcal{C}$  be a primitive Reed-Solomon code over  $\mathbb{F}_q$  with  $q \geq 4$  and  $2 \leq k \leq q - 2$ . Then polynomials of the form  $u(x) = ax^{q-2} + v(x)$  with  $a \neq 0$ , where  $\deg v(x) \leq k - 1$ , represent deep holes for  $\mathcal{C}$ .*

(This was later generalised to any  $D \neq \mathbb{F}_q$  by Wu and Hong [19], and separately by Zhang, Fu, and Liao [17].) With this new family of deep holes, they revised the Cheng-Murray conjecture:

**Conjecture** (Wu-Hong). *All deep holes for primitive Reed-Solomon codes are those words  $u$  represented by  $u(x) = ax^k + v(x)$  or  $ax^{q-2} + v(x)$  where  $\deg(v(x)) \leq k - 1$ .*

Cheng, Li, and Zhuang in the previously referenced [4] were able to resolve this under some conditions.

**Theorem 6.** *Given a finite field  $\mathbb{F}_q$  with characteristic  $p > 2$ , if  $3 \leq k < \frac{\sqrt{q}+1}{4}$  or  $3 \leq k < \frac{p}{45}$  when  $q = p$  is prime, then the Wu-Hong conjecture is true.*

Li and Zhu [14] and Zhang, Fu, and Liao [17] found families of words of degree  $k+1$  and  $k+2$  that were not deep holes. The latter authors also studied codes over fields of characteristic two, finding yet another family of deep holes.

**Theorem 7** (Zhang-Fu-Liao). *Let  $q > 4$  be a power of 2 and let  $\mathcal{C}$  be a Reed-Solomon code over  $\mathbb{F}_q$  with  $D = \mathbb{F}_q^*$  or  $D = \mathbb{F}_q^*/\{1\}$  and  $k = q-4$ . If  $a \neq 0$ , then polynomials of the form  $u(x) = ax^{q-3} + v(x)$ , where  $\deg v(x) \leq k-1$ , represent deep holes for  $\mathcal{C}$ .*

This example shows that the deep hole problem in characteristic two, the most important setting for applications, may be very complicated.

### 2.3. For other choices of $D$

In a previous paper with Zhu [10], we proved a scheme to find families of words with degree slightly larger than  $k$  that were not deep holes, if a certain character sum depending on  $D$  could be estimated.

**Theorem 8.** *Let  $\mathcal{C}$  be the generalised Reed-Solomon code over  $\mathbb{F}_q$  using the evaluation set  $D$ . Let  $u$  be a received word. Suppose we can write*

$$\left( \frac{w(x_1)}{h(x_1)}, \frac{w(x_2)}{h(x_2)}, \dots, \frac{w(x_{|D|})}{h(x_{|D|})} \right) = u$$

for some  $h(x) \in \mathbb{F}_q[x]$ , with no roots in  $D \cup 0$ , and  $\deg h(x) + k \leq \deg w(x) \leq |D| - 1$ . Let  $m$  be the smallest such degree of  $w(x)$ . Let  $1 \leq r \leq d := m - k \leq |D| - k - 1$ . If the bound

$$\left| \sum_{a \in D} \chi(1 - ax) \right| \leq Kq^{1/2}$$

is true over all nontrivial characters  $\chi : (\mathbb{F}[x]/(\bar{h}(x)))^* \rightarrow \mathbb{C}^*$  with  $\chi(\mathbb{F}_q^*) = 1$  for some  $K \geq d$  and  $\bar{h}(x) = x^{m-k+1}h(1/x)$ , there are positive constants  $c_1$  and  $c_2$  such that if

$$d \leq K < c_1 \frac{|D|}{q^{1/2}}, \left( \frac{d+r}{2} + 1 \right) \log_2 q < k < c_2 |D|$$

then  $d(u, \mathcal{C}) \leq |D| - k - r$ .

We were able to derive a suitable bound taking  $D$  as any subgroup of  $\mathbb{F}_q^*$ :

**Theorem 9.** *Let  $\mathcal{C}$  be the Reed-Solomon code over  $\mathbb{F}_q$  using the evaluation set  $D = (\mathbb{F}_q^*)^{\frac{q-1}{r}}$  of size  $\ell$ . Let  $r \geq 1$  be an integer and  $u$  a received word with interpolated polynomial  $u(x)$  such that  $r \leq d := \deg(u(x)) - k \leq q - 2 - k$ . There are positive constants  $c_1$  and  $c_2$  such that if*

$$d < c_1 \frac{\ell}{q^{1/2}}, \left( \frac{d+r}{2} + 1 \right) \log_2 q < k < c_2 \ell$$

then  $d(u, \mathcal{C}) \leq \ell - k - r$ .

This shows that some progress can be made on classifying deep holes when  $D$  is relatively small.

### 3. New Results

Much of the previous work took Reed-Solomon codes with  $D = \mathbb{F}_q, \mathbb{F}_q^*$ , a subgroup of  $\mathbb{F}_q^*$ , or a very large subset of  $\mathbb{F}_q$ . Any subset  $D$  of  $\mathbb{F}_q$  can be represented as the image of a polynomial  $f(x)$  over  $\mathbb{F}_q$ . Previous investigations take the polynomial  $f(x)$  to be the linear polynomial  $x$  or the monomial  $x^{(q-1)/\ell}$ . For a more general polynomial, the problem becomes much harder, as even computing the cardinality of the image  $f(\mathbb{F}_q)$  is complicated, see [2] for more detail.

Here we study the case where  $D$  is slightly more general - it will be the image of a Dickson polynomial over  $\mathbb{F}_q$ , which is defined as follows:

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

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

Note that for  $a = 0$ ,  $D_n(x, 0) = x^n$ , so we see that Dickson polynomials are a sort of generalisation 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 the image of the map  $D_n : \mathbb{F}_q^* \rightarrow \mathbb{F}_q^*$  has size  $q-1$  if  $\gcd(n, q-1) = 1$  and size  $(q-1)/\ell$  if  $\gcd(n, q-1) = \ell$ . In the first case, the map is 1 to 1; in the latter case, the map is  $\ell$  to 1. It turns out an analogous preimage-counting statement holds when  $a \neq 0$ . Chou, Mullen, and Wassermann in [5] used a character sum argument to calculate

**Theorem 10.** *Let  $n \geq 2$  and  $a \in \mathbb{F}_q^*$ . If  $q$  is even, then  $|D_n^{-1}(D_n(x_0, a))| =$*

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

where ‘condition A’ holds if  $x^2 + x_0x + a$  is reducible over  $\mathbb{F}_q$  and  $D_n(x_0, a) \neq 0$ ; ‘condition B’ holds if  $x^2 + x_0x + a$  is irreducible over  $\mathbb{F}_q$  and  $D_n(x_0, a) \neq 0$ .

If  $q$  is odd, let  $\eta$  be the quadratic character of  $\mathbb{F}_q$ . If  $2^r \mid (q^2 - 1)$  then

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

where ‘condition  $C$ ’ holds if

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

or

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

They also showed an explicit formula for the size of the value set of  $D_n(x, a)$ , denoted  $|V_{D_n(x, a)}|$ .

**Theorem 11.** *Let  $a \in \mathbb{F}_q^*$ . If  $2^r || (q^2 - 1)$  and  $\eta$  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

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

These results together with a character sum argument lead us to our main theorem.

### 3.1. Main Theorem

**Theorem 12.** *Let  $\mathcal{C}$  be the Reed-Solomon code over  $\mathbb{F}_q$  with message length  $k$ , using the evaluation set  $D = \{D_n(x, a) \mid x \in \mathbb{F}_q\}$ , for  $a \in \mathbb{F}_q^*$ . Let  $u$  be a received word and  $u(x)$  its interpolated polynomial with  $\deg u(x) = k+1$ . There exist computable positive constants  $c_1$  and  $c_2$  such that if the conditions*

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

*are satisfied, then  $u$  is not a deep hole.*

Theorem 11 shows that  $D$  takes on a variety of sizes, large and small, depending on the parameters. This implies that progress is possible toward the deep hole problem for some small evaluation sets without any obvious algebraic structure to rely on.

### 3.2. Example

Take  $q = 2^{16}$  to consider a Reed-Solomon code over  $\mathbb{F}_{2^{16}}$ . When  $n = 3$ , the Dickson polynomial family is  $D_3(x, a) = x^3 - 3ax = x^3 + ax$ . Theorem 11 gives  $|D| = 43691$ . The conditions to satisfy are

$$640 < 43691c_1 \quad \text{and} \quad 16 \leq k < 43691c_2$$

Fix  $c_1 = .015$ . The proof shows that the largest choice of  $c_2$  is

$$c_2 = 65536^{-\frac{1}{k+1}} - \frac{1}{2} - .015$$

to examine

$$16 \leq k < 43691 \left( 65536^{-\frac{1}{k+1}} - \frac{1}{2} - .015 \right).$$

The supported message sizes are  $16 \leq k \leq 21182$ .

### 3.3. Preliminaries

#### 3.3.1. Weil's Character Sum Bound

Our results rely on the following generalisation of Weil's classical character sum bound, see [7]:

**Theorem 13.** *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  $D_2 = 0$  if  $\deg(f_{n+1}) \leq 0$  and  $D_2 = \deg(f_{n+1})$  if  $\deg(f_{n+1}) > 0$ , 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  $\chi_i : \mathbb{F}_q^* \rightarrow \mathbb{C}^*$  ( $1 \leq i \leq n$ ) be multiplicative characters of  $\mathbb{F}_q$ , and let  $\psi = \psi_p \circ \text{Tr}_{\mathbb{F}_q/\mathbb{F}_p}$  for a non-trivial additive character  $\psi_p : \mathbb{F}_p \rightarrow \mathbb{C}^*$  of  $\mathbb{F}_p$ . Extend  $\chi_i$  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 for any  $m \geq 1$ , we have*

$$\left| \sum_{a \in \mathbb{F}_{q^m}, f_{n+1}(a) \neq \infty} \chi_1(N_{\mathbb{F}_{q^m}/\mathbb{F}_q}(f_1(a))) \cdots \chi_n(N_{\mathbb{F}_{q^m}/\mathbb{F}_q}(f_n(a))) \psi(\text{Tr}_{\mathbb{F}_{q^m}/\mathbb{F}_q}(f_{n+1}(a))) \right| \leq (D_1 + D_2 + D_3 + D_4 - 1)q^{m/2}$$

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

We specify the parameters to obtain various character sum bounds.

**Corollary 1.** *Let  $\psi_{\text{Tr}} = \psi_p \circ \text{Tr}_{\mathbb{F}_q/\mathbb{F}_p}$  be as above,  $\psi : \mathbb{F}_q \rightarrow \mathbb{C}^*$  a non-trivial additive character, and  $\eta : \mathbb{F}_q^* \rightarrow \mathbb{C}^*$  the quadratic character if  $q$  is odd. Set  $m = 1$ .*

1. If  $f_1(x) = D_n(x, a)$  with  $a \neq 0$ :

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

2. If  $q$  is odd,  $f_1(x) = x^2 - 4a$ , and  $f_2(x) = D_n(x, a)$  with  $a \neq 0$ :

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

3. If  $q$  is even and  $f_1(x) = bD_n(x, a) + a/x^2$  with  $a, b \neq 0$ :

$$\left| \sum_{x \in \mathbb{F}_q^*} \psi_{\text{Tr}}(bD_n(x, a) + a/x^2) \right| = \left| \sum_{x \in \mathbb{F}_q^*} \psi_{\text{Tr}}(bD_n(x, a) + a^{q/2}/x) \right| \leq (n+1)\sqrt{q}$$

Note that none of the polynomials in place of  $f_{n+1}(x)$  are of the form  $r(t)^2 - r(t) + c$ . For instance, in part 3, such an  $r$  would have to take the form  $r(t) = d/x + f(x)$ , where  $f(x)$  is a polynomial. Expanding this shows that  $d^2/x^2 + d/x = a^{q/2}/x$ , or  $d = 0$ , which is a contradiction.

**Lemma 1.** Let  $D = \{D_n(x, a) \mid x \in \mathbb{F}_q\}$ , for  $a \in \mathbb{F}_q^*$ . If  $\psi : (\mathbb{F}_q, +) \rightarrow \mathbb{C}^*$  is a non-trivial additive character, then the following estimate holds:

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

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

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

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

**When  $q$  is even:**

By Theorem 10,  $N_x$  can be quantified. Let  $\text{Tr} : \mathbb{F}_q \rightarrow \mathbb{F}_2$  denote the absolute trace. Using the fact that  $z^2 + xz + a$  is reducible over  $\mathbb{F}_q$  if and only if  $\text{Tr}(a/x^2) = 0$ ,

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

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



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

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:

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

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

Then,

$$\sum_{x \in \mathbb{F}_q^*} \psi(D_n(x, a)) \psi_{\text{Tr}}(a/x^2) = \sum_{x \in \mathbb{F}_q^*} \psi_{\text{Tr}}(bD_n(x, a) + a/x^2).$$

Applying the bounds in Corollary 1,

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

**When  $q$  is odd:**

We use Theorem 10 again to calculate  $N_x$ . Let  $\eta$  be the quadratic character of  $\mathbb{F}_q$ .

$$= \sum_{\substack{x \in \mathbb{F}_q \\ \eta(x^2-4a)=1}} \frac{1}{\gcd(n, q-1)} \psi(D_n(x, a)) + \sum_{\substack{x \in \mathbb{F}_q \\ \eta(x^2-4a)=-1}} \frac{1}{\gcd(n, q+1)} \psi(D_n(x, a)) + O(1).$$

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

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

Again applying the bounds in Corollary 1,

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

which was to be shown.  $\square$

### 3.3.2. Li-Wan’s New Sieve

We also state Li-Wan’s new sieve (as in [12, 20]) to estimate a particular type of character sum: let  $D$  be a finite set and  $D^k = D \times D \times \dots \times D$  be the Cartesian product of  $k$  copies of  $D$ . Let  $X$  be a subset of  $D^k$ . Denote

$$\overline{X} = \{(x_1, x_2, \dots, x_k) \in X \mid x_i \neq x_j, i \neq j\}$$

Let  $f(x_1, x_2, \dots, x_k)$  be a complex-valued function defined over  $X$ . Denote

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

Let  $S_k$  be the symmetric group on  $\{1, 2, \dots, k\}$ . Each permutation  $\tau \in S_k$  can be uniquely factorised as a product of disjoint cycles and each fixed point is viewed as a trivial cycle of length 1. Namely,

$$\tau = (i_1 i_2 \dots i_{a_1})(j_1 j_2 \dots j_{a_2}) \cdots (l_1 l_2 \dots l_{a_s})$$

with  $a_i \geq 1$  and  $1 \leq i \leq s$ . Define

$$X_\tau = \{(x_1, x_2, \dots, x_k) \mid x_{i_1} = \dots = x_{i_{a_1}}, x_{j_1} = \dots = x_{j_{a_2}}, \dots, x_{l_1} = \dots = x_{l_{a_s}}\}$$

Similarly define

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

We say that  $\tau$  is of the type  $(c_1, c_2, \dots, c_k)$  if it has exactly  $c_i$  cycles of length  $i$ . Let  $N(c_1, c_2, \dots, c_k)$  be the number of permutations of type  $(c_1, c_2, \dots, c_k)$ . Define

$$C_k(t_1, t_2, \dots, t_k) = \sum_{\sum ic_i = k} N(c_1, c_2, \dots, c_k) t_1^{c_1} t_2^{c_2} \dots t_k^{c_k}$$

Now we have the following combinatorial result:

**Lemma 2.** *Suppose  $q \geq d$ . If  $t_i = q$  for  $d|i$  and  $t_i = s$  for  $d \nmid i$ , then we have*

$$\begin{aligned} C_k(s, \dots, s, q, s, \dots, x, q, \dots) &= k! \sum_{i=0}^{\lfloor k/d \rfloor} \binom{\frac{q-s}{d} + i - 1}{i} \binom{s + k - di - 1}{k - di} \\ &\leq \left( s + k + \frac{q-s}{d} - 1 \right)_k \end{aligned}$$

where  $(x)_k = x(x-1)(x-2) \dots (x-k+1)$ .

Furthermore, we say that  $X$  is symmetric if for any  $x \in X$  and any  $g \in S_k$ , we have  $g \circ x \in X$ . Also, if a complex-valued function  $f$  is defined on  $X$ , we say that it is normal on  $X$  if  $X$  is symmetric and for any two conjugate elements in  $S_k$ ,  $\tau$  and  $\tau'$ , we have

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

Then, we have the result:

**Lemma 3.** *If  $f$  is normal on  $X$ , then*

$$F = \sum_{\sum ic_i = k} (-1)^{k - \sum c_i} N(c_1, c_2, \dots, c_k) F_\tau$$

### 3.3.3. A Rephrasing of Error Distance

A simple argument shows that error distance can be rephrased in the following way:

**Lemma 4.** *Let  $\mathcal{C}$  be a Reed-Solomon code over  $\mathbb{F}_q$  using the evaluation set  $D$ . Let  $u$  be a received word and  $u(x)$  be its interpolated polynomial with  $\deg u(x) = k+1$ . The error distance satisfies  $d(u, \mathcal{C}) \leq |D| - k - 1$  if and only if there exists a subset  $\{x_{i_1}, x_{i_2}, \dots, x_{i_{k+1}}\} \subset D$  such that*

$$u(x) - v(x) = (x - x_{i_1})(x - x_{i_2}) \dots (x - x_{i_{k+1}})$$

for some  $v(x)$  with  $\deg v(x) \leq k - 1$ .

#### 4. The Proof

*Proof.* For the received word  $u$ , write the interpolated polynomial as

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

By Lemma 4, if  $u(x)$  is not a deep hole, then there exists a codeword  $v(x)$  of degree  $\leq k-1$  where

$$u(x) - v(x) = (x - x_1) \cdots (x - x_k)(x - x_{k+1})$$

for distinct values of  $x_1, \dots, x_k, x_{k+1}$  in  $D$ . By expanding this product, we see that  $u$  will not be a deep hole if and only if the equation

$$x_1 + \dots + x_k + x_{k+1} = b_1$$

has a solution with distinct coordinates for every  $b_1$  in  $\mathbb{F}_q$ . Let  $N_u$  be the number of solutions to this equation and  $G$  be the group of additive characters  $\psi : \mathbb{F}_q \rightarrow \mathbb{C}^*$ . By the orthogonality of characters,

$$N_u = \frac{1}{q} \sum_{\substack{x_i \in D \\ \text{distinct}}} \sum_{\psi \in G} \psi(x_1 + \dots + x_k + x_{k+1} - b_1)$$

Removing the trivial character and exchanging the sums, we obtain

$$\left| N_u - \frac{1}{q} (|D|)_{k+1} \right| = \left| \frac{1}{q} \sum_{\substack{\psi \in G \\ \psi \neq 1}} \psi(b_1)^{-1} \sum_{\substack{x_i \in D \\ \text{distinct}}} \psi(x_1 + \dots + x_k + x_{k+1}) \right|.$$

Now we can estimate the inner sum of the right hand side using Li-Wan's new sieve. Let  $X = D^{k+1}$ ,  $f(x_1, \dots, x_k, x_{k+1}) = \psi(x_1 + \dots + x_k + x_{k+1}) = \psi(x_1) \cdots \psi(x_k) \psi(x_{k+1})$ , so  $F = \sum_{\mathbf{x} \in X} f(\mathbf{x})$ .  $F$  is symmetric and normal. Then

$$F_\tau = \sum \psi(x_{11}) \cdots \psi(x_{1c_1}) \cdots \psi^{k+1}(x_{(k+1)c_1}) \cdots \psi^{k+1}(x_{(k+1)c_{k+1}})$$

where the sum runs over  $x_{st_s} \in D$ ,  $1 \leq s \leq k+1$ , and  $1 \leq t_s \leq c_s$ . Applying Lemma 1, we obtain

$$\begin{aligned} \left| N_u - \frac{1}{q} (|D|)_{k+1} \right| &\leq \left| \sum_{\substack{x_i \in D \\ \text{distinct}}} \psi(x_1 + \dots + x_k + x_{k+1}) \right| \\ &\leq \sum_{\sum ic_i = k+1} N(c_1, \dots, c_{k+1}) |F_\tau| \\ &\leq C_{k+1} ((n+1)\sqrt{q}, |D|, (n+1)\sqrt{q}, |D|, \dots, (n+1)\sqrt{q}, |D|) \\ &\leq \left( (n+1)\sqrt{q} + (k+1) + \frac{|D| - (n+1)\sqrt{q}}{2} - 1 \right)_{k+1} \\ &= \left( \frac{(n+1)\sqrt{q}}{2} + k + \frac{|D|}{2} \right)_{k+1}. \end{aligned}$$

To guarantee that  $N_u > 0$ , it suffices to have

$$\frac{1}{q}(|D|)_{k+1} > \left( \frac{(n+1)\sqrt{q}}{2} + k + \frac{|D|}{2} \right)_{k+1},$$

which is true if

$$\frac{|D|}{\frac{n+1}{2}\sqrt{q} + k + \frac{|D|}{2}} > q^{\frac{1}{k+1}}.$$

If we take  $\frac{n+1}{2}\sqrt{q} < c_1|D|$  and  $k < c_2|D|$  for positive constants  $c_1$  and  $c_2$ , we calculate the condition

$$q^{-\frac{1}{k+1}} - \frac{1}{2} > c_1 + c_2.$$

Therefore, it is enough to find

$$q^{-\frac{1}{k+1}} > \frac{1}{2}.$$

With some rearrangement, we have

$$k > (\log_2 q) - 1,$$

or more simply,

$$k \geq \log_2 q.$$

□

## 5. Conclusions

We were able to show that words of degree  $k+1$  were not deep holes in certain Reed-Solomon codes based on Dickson polynomials. We reduced to solving a type of restricted subset sum problem. If  $D$  is the value set of  $D_n(x, a)$  over some finite field  $\mathbb{F}_q$ , then this problem essentially asks to find conditions involving  $q$ ,  $n$ , and a parameter  $r$  such that

$$x_1 + \dots + x_r = c$$

always has a solution using distinct  $x_i \in D$ , for any  $c \in \mathbb{F}_q$ . It is interesting to note that this is actually a harder version of Waring's problem for Dickson polynomials, which was originally studied without a distinctness condition by Gomez and Winterhof [9] and improved later by Ostafe and Shparlinski [15]. If we write  $x_i = D_n(u_i, a)$  for suitable  $u_i \in \mathbb{F}_q$ , then our techniques give conditions to guarantee a solution to

$$D_n(u_1, a) + \dots + D_n(u_r, a) = c$$

where we require that each value  $D_n(u_i, a)$  is distinct.

In a larger scope, this work was actually motivated by an attempt to use the stronger result from Ketzi-Zhu in [10]. As summarised in the introduction,

it allows one to bound the error distance between a received word and code if the character sum bound

$$\left| \sum_{b \in D} \chi(1 - bx) \right| \leq Kq^{1/2}$$

can be achieved for a suitable  $K$ , for any non-trivial  $\chi$  belonging to a particular character group. We leave this to a future investigation.

## References

- [1] A. Cafure, G. Matera, and M. Privitelli. Singularities of Symmetric Hypersurfaces and Reed-Solomon Codes. *Advances in Mathematics of Communications*, Vol. 6(1), 2012, pp.69-94.
- [2] Q. Cheng, J. Hill and D. Wan, 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] Q. Cheng and E. Murray. On deciding deep holes of Reed-Solomon codes. *Proceedings of TAMC 2007, LNCS 4484*, pp. 296-305
- [4] Q. Cheng, J. Li, and J. Zhuang. On Determining Deep Holes of Generalized Reed-Solomon Codes. in *Algorithms and Computation, LNCS, Volume 8283 (2013)*, 100-110.
- [5] W. Chou, G.L. Mullen, and B. Wasserman. 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.
- [6] Q. Cheng and D. Wan. On the List and Bounded Distance Decodibility of Reed-Solomon Codes. *Proceedings of the 45th Annual IEEE Symposium on Foundations of Computer Science (FOCS'04)*.
- [7] L. Fu and D. Wan. *L*-functions and Character Sums over Finite Fields. Preprint.
- [8] V. Guruswami and A. Vardy. Maximum-Likelihood Decoding of Reed-Solomon Codes is NP-hard. *SODA '05 Proceedings of the sixteenth annual ACM-SIAM symposium on Discrete algorithms: 470-478*. Society for Industrial and Applied Mathematics Philadelphia, PA, USA, 2005.
- [9] D. Gomez and A. Winterhof. Warings Problem in Finite Fields with Dickson Polynomials. *Finite fields: Theory and applications, Contemp. Math.*, vol. 477, Amer. Math. Soc., 2010, 185192.
- [10] M. Ketı and G. Zhu. Computing Error Distance of Reed-Solomon Codes. Preprint.

- [11] Q. Liao. On Reed-Solomon Codes. Chinese Annals of Mathematics, 32B(1), 89-98. Springer-Verlag Berlin Heidelberg, 2011.
- [12] J. Li and D. Wan. A new sieve for distinct coordinate counting. Science China Mathematics, Vol. 53 No.9: 2351-2362. Science China Press and Springer-Verlag Berlin Heidelberg, 2010.
- [13] J. Li and D. Wan. On the subset sum problem over finite fields. Finite Fields and Their Applications, Volume 14, Issue 4, November 2008: 911-929
- [14] Y. Li and G. Zhu. On error distance of received words with fixed degrees to Reed-Solomon code. Preprint.
- [15] A. Ostafe and I.E. Shparlinski. On the Waring Problem with Dickson Polynomials in Finite Fields. Proceedings of the American Mathematical Society, Vol. 139, No. 11, November 2011, Pages 3815-3820.
- [16] D. Wan. Generators and irreducible polynomials over finite fields. Mathematics of Computation, 66, 119-1212 (1997).
- [17] J. Zhang, F. Fu, and Q. Liao. New Deep Holes of Generalized Reed-Solomon Codes. Science China A: Mathematics (Chinese Edition), vol.43, no.7, pp.727-740, 2013.
- [18] R. Wu and S. Hong. On deep holes of standard Reed-Solomon codes. Science China Math. 55 (2012), 2447-2455.
- [19] R. Wu and S. Hong. On deep holes of generalized Reed-Solomon codes. arXiv: 1205.7016.
- [20] G. Zhu and D. Wan. Computing Error Distance of Reed-Solomon Codes. TAMC 2012, LNCS 7287, pp. 214-224, 2012. Springer-Verlag Berlin Heidelberg, 2012.