A Construction of Binary Linear Codes from Boolean Functions

Cunsheng Ding

Department of Computer Science and Engineering, The Hong Kong University of Science and Technology, Clear Water Bay, Kowloon, Hong Kong

Abstract

Boolean functions have important applications in cryptography and coding theory. Two famous classes of binary codes derived from Boolean functions are the Reed-Muller codes and Kerdock codes. In the past two decades, a lot of progress on the study of applications of Boolean functions in coding theory has been made. Two generic constructions of binary linear codes with Boolean functions have been well investigated in the literature. The objective of this paper is twofold. The first is to provide a survey on recent results, and the other is to propose open problems on one of the two generic constructions of binary linear codes with Boolean functions. These open problems are expected to stimulate further research on binary linear codes from Boolean functions.

Keywords: Almost bent functions, bent functions, difference sets, linear codes, semibent functions, o-polynomials.

1. Introduction

Let p be a prime and let $q = p^m$ for some positive integer m. An [n, k, d] code \mathcal{C} over GF(p) is a k-dimensional subspace of $GF(p)^n$ with minimum (Hamming) distance d. Let A_i denote the number of codewords with Hamming weight i in a code \mathcal{C} of length n. The weight enumerator of \mathcal{C} is defined by $1 + A_1z + A_2z^2 + \cdots + A_nz^n$. The sequence $(1, A_1, A_2, \cdots, A_n)$ is called the weight distribution of the code \mathcal{C} . A code \mathcal{C} is said to be a t-weight code if the number of nonzero A_i in the sequence (A_1, A_2, \cdots, A_n) is equal to t.

Boolean functions are functions from $GF(2^m)$ or $GF(2)^m$ to GF(2). They are important building blocks for certain types of stream ciphers, and can also be employed to construct binary codes. Two famous families of binary codes are the Reed-Muller codes [64, 60] and Kerdock codes [10, 11, 47]. In the literature two generic constructions of binary linear codes from Boolean functions have been well investigated. A lot of progress on the study of one of the two constructions has been made in the past decade. The objective of this paper is twofold. The first one is to provide a survey on recent development on this construction, and the other is to propose open problems on this generic constructions of binary linear codes with Boolean functions. These open problems are expected to stimulate further research on binary linear codes from Boolean functions.

Preprint submitted to Elsevier

C. Ding's research was supported by The Hong Kong Research Grants Council, under Proj. No. 16300415.

Email address: cding@ust.hk (Cunsheng Ding)

2. Mathematical foundations

2.1. Difference sets

For convenience later, we define the *difference function* of a subset D of an abelian group (A, +) as

$$diff_D(x) = |D \cap (D+x)|,\tag{1}$$

where $D + x = \{y + x : y \in D\}.$

A subset D of size k in an abelian group (A, +) with order v is called a (v, k, λ) difference set in (A, +) if the difference function $\mathrm{diff}_D(x) = \lambda$ for every nonzero $x \in A$. A difference set D in (A, +) is called *cyclic* if the abelian group A is so.

Difference sets could be employed to construct linear codes in different ways. The reader is referred to [27, 28] for detailed information. Some of the codes presented in this survey paper are also defined by difference sets.

2.2. *Group characters in* GF(q)

An *additive character* of GF(q) is a nonzero function χ from GF(q) to the set of nonzero complex numbers such that $\chi(x+y) = \chi(x)\chi(y)$ for any pair $(x,y) \in GF(q)^2$. For each $b \in GF(q)$, the function

$$\chi_b(c) = \varepsilon_p^{\text{Tr}(bc)} \quad \text{for all } c \in \text{GF}(q)$$
 (2)

defines an additive character of GF(q), where and whereafter $\varepsilon_p = e^{2\pi\sqrt{-1}/p}$ is a primitive complex pth root of unity and Tr is the absolute trace function. When b=0, $\chi_0(c)=1$ for all $c\in GF(q)$, and is called the *trivial additive character* of GF(q). The character χ_1 in (2) is called the *canonical additive character* of GF(q). It is known that every additive character of GF(q) can be written as $\chi_b(x) = \chi_1(bx)$ [51, Theorem 5.7].

2.3. Special types of polynomials over GF(q)

It is well known that every function from GF(q) to GF(q) can be expressed as a polynomial over GF(q). A polynomial $f \in GF(q)[x]$ is called a *permutation polynomial* if the associated polynomial function $f : a \mapsto f(a)$ from GF(q) to GF(q) is a permutation of GF(q).

Dickson polynomials of the first kind over GF(q) are defined by

$$D_h(x,a) = \sum_{i=0}^{\lfloor \frac{h}{2} \rfloor} \frac{h}{h-i} \binom{h-i}{i} (-a)^i x^{h-2i}, \tag{3}$$

where $a \in GF(q)$ and h is called the *order* of the polynomial. Some of the linear codes that will be presented in this paper are defined by Dickson permutation polynomials of order 5 over $GF(2^m)$.

A polynomial $f \in GF(q)[x]$ is said to be *e-to-1* if the equation f(x) = b over GF(q) has either e solutions $x \in GF(q)$ or no solution for every $b \in GF(q)$, where $e \ge 1$ is an integer, and e divides q. By definition, permutation polynomials are 1-to-1. In this survey paper, we need e-to-1 polynomials over $GF(2^m)$ for the construction of binary linear codes.

2.4. Boolean functions and their expressions

A function f from $GF(2^m)$ or $GF(2)^m$ to GF(2) is called a *Boolean function*. A function f from $GF(2^m)$ to GF(2) is called *linear* if f(x+y)=f(x)+f(y) for all $(x,y)\in GF(2^m)^2$. A function f from $GF(2^m)$ to GF(2) is called *affine* if f or f-1 is linear.

The Walsh transform of $f: GF(2^m) \to GF(2)$ is defined by

$$\hat{f}(w) = \sum_{x \in GF(2^m)} (-1)^{f(x) + Tr(wx)}$$
(4)

where $w \in GF(2^m)$. The Walsh spectrum of f is the following multiset

$$\left\{ \left\{ \hat{f}(w) : w \in \mathrm{GF}(2^m) \right\} \right\}.$$

Let f be a Boolean function from $GF(2^m)$ to GF(2). The *support* of f is defined to be

$$D_f = \{ x \in GF(2^m) : f(x) = 1 \} \subseteq GF(2^m).$$
 (5)

Clearly, $f \mapsto D_f$ is a one-to-one correspondence between the set of Boolean functions from $GF(2^m)$ to GF(2) and the power set of $GF(2^m)$.

3. The first generic construction of linear codes from functions

Let f be any polynomial from GF(q) to GF(q), where $q = p^m$. A code over GF(p) is defined by

$$C(f) = \{ \mathbf{c} = (\operatorname{Tr}(af(x) + bx))_{x \in \operatorname{GF}(q)} : a \in \operatorname{GF}(q), \ b \in \operatorname{GF}(q) \},$$

where Tr is the absolute trace function. Its length is q, and its dimension is at most 2m and is equal to 2m in many cases. The dual of C(f) has dimension at least q-2m.

Let f be any polynomial from GF(q) to GF(q) such that f(0)=0. A code over GF(p) is defined by

$$C^*(f) = \{ \mathbf{c} = (\text{Tr}(af(x) + bx))_{x \in GF(q)^*} : a \in GF(q), b \in GF(q) \}.$$

Its length is q-1, and its dimension is at most 2m and is equal to 2m in many cases. The dual of $C^*(f)$ has dimension at least q-1-2m.

This is a generic construction of linear codes, which has a long history and its importance is supported by Delsarte's Theorem [24]. It gives a coding-theory characterisation of APN monomials, almost bent functions, and semibent functions (see, for examples, [13], [8] and [44]) when q = 2. We will not deal with this construction in this paper.

4. The second generic construction of linear codes from functions

In this section, we present the second generic construction of linear codes over GF(p) with any subset D of $GF(p^m)$, and introduce basic results about the linear codes. In Section 5, we will consider specific families of binary linear codes from Boolean functions obtained with this generic construction.

4.1. The description of the construction of linear codes

Let $D = \{d_1, d_2, \dots, d_n\} \subseteq GF(q)$, where again $q = p^m$. Recall that Tr denotes the trace function from GF(q) onto GF(p) throughout this paper. We define a linear code of length n over GF(p) by

$$C_D = \{ (\operatorname{Tr}(xd_1), \operatorname{Tr}(xd_2), \dots, \operatorname{Tr}(xd_n)) : x \in \operatorname{GF}(q) \},$$
(6)

and call D the *defining set* of this code C_D . By definition, the dimension of the code C_D is at most m.

This construction is generic in the sense that many classes of known codes could be produced by selecting the defining set $D \subseteq GF(q)$ properly. This construction technique was employed in [28], [29], [30], [33] and other papers for obtaining linear codes with a few weights. If the set D is properly chosen, the code C_D may have good or optimal parameters. Otherwise, the code C_D could have bad parameters.

4.2. The weights in the linear codes C_D

It is convenient to define for each $x \in GF(q)$,

$$\mathbf{c}_{x} = (\operatorname{Tr}(xd_{1}), \operatorname{Tr}(xd_{2}), \dots, \operatorname{Tr}(xd_{n})). \tag{7}$$

The Hamming weight wt(\mathbf{c}_x) of \mathbf{c}_x is $n - N_x(0)$, where

$$N_x(0) = |\{1 \le i \le n : \text{Tr}(xd_i) = 0\}|$$

for each $x \in GF(q)$.

It is easily seen that for any $D = \{d_1, d_2, \dots, d_n\} \subseteq GF(q)$ we have

$$pN_x(0) = \sum_{i=1}^n \sum_{y \in GF(p)} e^{2\pi\sqrt{-1}y Tr(xd_i)/p} = \sum_{i=1}^n \sum_{y \in GF(p)} \chi_1(yxd_i) = n + \sum_{y \in GF(p)^*} \chi_1(yxD)$$

where χ_1 is the canonical additive character of GF(q), aD denotes the set $\{ad: d \in D\}$, and $\chi_1(S) := \sum_{x \in S} \chi_1(x)$ for any subset S of GF(q). Hence,

$$wt(\mathbf{c}_{x}) = n - N_{x}(0) = \frac{(p-1)n - \sum_{y \in GF(p)^{*}} \chi_{1}(yxD)}{p}.$$
 (8)

4.3. Differences between the first and second generic constructions

The second generic construction of this section is different from the first generic construction of Section 3 in the following aspects:

- While the length of the codes in the first generic construction in Section 3 is either q or q-1, that of the codes in the second generic construction could be any integer between 1 and q, depending on the underlying defining set D.
- While the dimension of the codes in the first construction in Section 3 is usually 2m, that of the codes in the second construction is usually m and is at most m.

5. Binary codes from the preimage $f^{-1}(b)$ of Boolean functions f

Let f be a function from $GF(2^m)$ to GF(2), and let D be any subset of the preimage $f^{-1}(b)$ for any $b \in GF(2)$. In general, it is very hard to determine the parameters of the code \mathcal{C}_D . Recall the support D_f of f defined in (5). Let $n_f = |D_f|$. In this section, we deal with the binary code \mathcal{C}_{D_f} with length n_f and dimension at most m, and will focus on the weight distribution of the code \mathcal{C}_{D_f} for several classes of Boolean functions f.

The following theorem plays a major role in this section whose proof can be found in [28].

Theorem 1. Let f be a function from $GF(2^m)$ to GF(2), and let D_f be the support of f. If $2n_f + \hat{f}(w) \neq 0$ for all $w \in GF(2^m)^*$, then C_{D_f} is a binary linear code with length n_f and dimension m, and its weight distribution is given by the following multiset:

$$\left\{ \left\{ \frac{2n_f + \hat{f}(w)}{4} : w \in GF(2^m)^* \right\} \right\} \cup \left\{ \{0\} \right\}.$$
(9)

Theorem 1 establishes a connection between the set of Boolean functions f such that $2n_f + \hat{f}(w) \neq 0$ for all $w \in \mathrm{GF}(2^m)^*$ and a class of binary linear codes. The determination of the weight distribution of the binary linear code \mathcal{C}_{D_f} is equivalent to that of the Walsh spectrum of the Boolean function f satisfying $2n_f + \hat{f}(w) \neq 0$ for all $w \in \mathrm{GF}(2^m)^*$. When the Boolean function f is selected properly, the code \mathcal{C}_{D_f} has only a few weights and may have good parameters. We will demonstrate this in the remainder of this section.

We point out that Theorem 1 can be generalized into the following whose proof is the same as that of Theorem 1.

Theorem 2. Let f be a function from $GF(2^m)$ to GF(2), and let D_f be the support of f. Let e_w denote the multiplicity of the element $\frac{2n_f + \hat{f}(w)}{4}$ and e the multiplicity of 0 in the following multiset of (9). Then C_{D_f} is a binary linear code with length n_f and dimension $m - \log_2 e$, and the weight distribution of the code is given by

$$\frac{2n_f + \hat{f}(w)}{4}$$
 with frequency $\frac{e_w}{e}$

for all $\frac{2n_f + \hat{f}(w)}{4}$ in the multiset of (9).

5.1. Linear codes from bent functions

A function from $GF(2^m)$ to GF(2) is called *bent* if $|\hat{f}(w)| = 2^{m/2}$ for every $w \in GF(2^m)$. Bent functions exist only for even m [65].

It is well known that a function f from $GF(2^m)$ to GF(2) is bent if and only if D_f is a difference set in $(GF(2^m), +)$ with the following parameters

$$(2^m, 2^{m-1} \pm 2^{(m-2)/2}, 2^{m-2} \pm 2^{(m-2)/2}).$$
 (10)

Let f be bent. Then by definition $\hat{f}(0) = \pm 2^{m/2}$. It then follows that

$$n_f = |D_f| = 2^{m-1} \pm 2^{(m-2)/2}$$
 (11)

As a corollary of Theorem 1, we have the following [28].

Table 1: The weight distribution of the codes of Corollary 3

Weight w	Multiplicity A_w
0	1
$\frac{n_f}{2} - 2^{\frac{m-4}{2}}$	$\frac{2^m - 1 - n_f 2^{-\frac{m-2}{2}}}{2}$
$\frac{n_f}{2} + 2^{\frac{m-4}{2}}$	$\frac{2^m - 1 + n_f 2^{-\frac{m-2}{2}}}{2}$

Corollary 3. Let f be a Boolean function from $GF(2^m)$ to GF(2) with f(0) = 0, where $m \ge 4$ and is even. Then C_{D_f} is an $[n_f, m, (n_f - 2^{(m-2)/2})/2]$ two-weight binary code with the weight distribution in Table 1, where n_f is defined in (11), if and only if f is bent.

There are many constructions of bent functions and thus Hadamard difference sets. We refer the reader to [4], [58], [59], the book chapter [12] and the references therein for details. Any bent function can be plugged into Corollary 3 to obtain a two-weight binary linear code.

The construction of binary codes with bent functions above can be generalized as follows.

Theorem 4. Let D be a $(2^m, n, \lambda)$ difference sets in $(GF(2^m), +)$. Then C_D is a two-weight binary code with parameters [n, m] and weight enumerator

$$1 + \frac{(2^m - 1)\sqrt{n - \lambda} - n}{2\sqrt{n - \lambda}}z^{\frac{n - \sqrt{n - \lambda}}{2}} + \frac{(2^m - 1)\sqrt{n - \lambda} + n}{2\sqrt{n - \lambda}}z^{\frac{n + \sqrt{n - \lambda}}{2}}.$$

5.2. Linear codes from semibent functions

Let *m* be odd. Then there is no bent Boolean function on $GF(2^m)$. A function *f* from $GF(2^m)$ to GF(2) is called *semibent* if $\hat{f}(w) \in \{0, \pm 2^{(m+1)/2}\}$ for every $w \in GF(2^m)$.

Let f be a semibent function from $GF(2^m)$ to GF(2). It then follows from the definition of semibent functions that

$$n_f = |D_f| = \begin{cases} 2^{m-1} - 2^{(m-1)/2} & \text{if } \hat{f}(0) = 2^{(m+1)/2}, \\ 2^{m-1} + 2^{(m-1)/2} & \text{if } \hat{f}(0) = -2^{(m+1)/2}, \\ 2^{m-1} & \text{if } \hat{f}(0) = 0. \end{cases}$$
(12)

Table 2: The weight distribution of the codes of Corollary 5

Weight w	Multiplicity A _w
0	1
$\frac{n_f - 2^{(m-1)/2}}{2n_f}$	$n_f(2^m - n_f)2^{-m} - n_f2^{-(m+1)/2}$
$n_f + 2^{(m-1)/2}$	$2^m - 1 - n_f(2^m - n_f)2^{-(m-1)}$
$\frac{n_f+2}{2}$	$n_f(2^m - n_f)2^{-m} + n_f2^{-(m+1)/2}$

As a corollary of Theorem 1, we have the following [28].

Corollary 5. Let f be a Boolean function from $GF(2^m)$ to GF(2) with f(0) = 0, where m is odd. Then C_{D_f} is an $[n_f, m, (n_f - 2^{(m-1)/2})/2]$ three-weight binary code with the weight distribution in Table 2, where n_f is defined in (12), if and only if f is semibent.

There are a lot of constructions of semibent functions from $GF(2^m)$ to GF(2). We refer the reader to [16, 21, 37, 54, 55, 57] for detailed constructions. All semibent functions can be plugged into Corollary 5 to obtain three-weight binary linear codes.

5.3. Linear codes from almost bent functions

For any function g from $GF(2^m)$ to $GF(2^m)$, we define

$$\lambda_g(a,b) = \sum_{x \in GF(2^m)} (-1)^{Tr(ag(x)+bx)}, \ a, b \in GF(2^m).$$

A function g from $GF(2^m)$ to $GF(2^m)$ is called *almost bent* (AB) if $\lambda_g(a,b) = 0$, or $\pm 2^{(m+1)/2}$ for every pair (a,b) with $a \neq 0$. By definition, almost bent functions over $GF(2^m)$ exist only for odd m. Specific almost bent functions are available in [5, 12].

odd m. Specific almost bent functions are available in [5, 12]. By definition, $\lambda_g(1,0) \in \{0, \pm 2^{(m+1)/2}\}$ for any almost bent function g on $GF(2^m)$. It is straightforward to deduce the following lemma.

Lemma 6. For any almost bent function g from $GF(2^m)$ to $GF(2^m)$, define f = Tr(g). Then we have

$$n_f = |D_{\text{Tr}(g)}| = \begin{cases} 2^{m-1} + 2^{(m-1)/2} & \text{if } \lambda_g(1,0) = -2^{(m+1)/2}, \\ 2^{m-1} - 2^{(m-1)/2} & \text{if } \lambda_g(1,0) = 2^{(m+1)/2}, \\ 2^{m-1} & \text{if } \lambda_g(1,0) = 0. \end{cases}$$
(13)

As a corollary of Theorem 1, we have the following [28]

Corollary 7. Let g be an almost bent function from $GF(2^m)$ to $GF(2^m)$ with Tr(g(0)) = 0, where m is odd. Define f = Tr(g). Then C_{D_f} is an $[n_f, m, (n_f - 2^{(m-1)/2})/2]$ three-weight binary code with the weight distribution in Table 2, where n_f is given in (13).

The following is a list of almost bent functions $g(x) = x^d$ on $GF(2^m)$ for odd m:

- 1. $d = 2^h + 1$, where gcd(m, h) = 1 is odd [41].
- 2. $d = 2^{2h} 2^h + 1$, where $h \ge 2$ and gcd(m, h) = 1 is odd [46].
- 3. $d = 2^{(m-1)/2} + 3$ [46].
- 4. $d = 2^{(m-1)/2} + 2^{(m-1)/4} 1$, where $m \equiv 1 \pmod{4}$ [44, 45].
- 5. $d = 2^{(m-1)/2} + 2^{(3m-1)/4} 1$, where $m \equiv 3 \pmod{4}$ [44, 45].

This list of almost bent monomials g(x) are permutation polynomials on $GF(2^m)$. Hence, the length of the code C_{D_f} is equal to 2^{m-1} , and the weight distribution of the code is given in Table 11.

5.4. Linear codes from quadratic Boolean functions

Let

$$f(x) = \text{Tr}_{2^m/2} \left(\sum_{i=0}^{\lfloor m/2 \rfloor} f_i x^{2^i + 1} \right)$$
 (14)

be a quadratic Boolean function from $GF(2^m)$ to GF(2), where $f_i \in GF(2^m)$. The rank of f, denoted by r_f , is defined to be the codimension of the GF(2)-vector space

$$V_f = \{ x \in GF(2^m) : f(x+z) - f(z) - f(z) = 0 \ \forall \ z \in GF(2^m) \}.$$

Table 3: The Walsh spectrum of quadratic Boolean functions

$\hat{f}(w)$	the number of w's
0	$2^{m}-2^{r_{f}}$
$2^{m-r_f/2}$	$2^{r_f-1} + 2^{(r_f-2)/2}$
$-2^{m-r_f/2}$	$2^{r_f-1} - 2^{(r_f-2)/2}$

The Walsh spectrum of f is known [9] and given in Table 3.

Let D_f be the support of f. By definition, we have

$$n_f = |D_f| = 2^{m-1} - \frac{\hat{f}(0)}{2} = \begin{cases} 2^{m-1} & \text{if } \hat{f}(0) = 0, \\ 2^{m-1} - 2^{m-1-r_f/2} & \text{if } \hat{f}(0) = 2^{m-r_f/2}, \\ 2^{m-1} + 2^{m-1-r_f/2} & \text{if } \hat{f}(0) = -2^{m-r_f/2}. \end{cases}$$
(15)

The following theorem then follows from Theorem 1 and Table 3.

Theorem 8. Let f be a quadratic Boolean function of the form in (14) such that $r_f > 2$. Then C_{D_f} is a binary code with length n_f given in (15), dimension m, and the weight distribution in Table 4, where

$$(\varepsilon_{1}, \varepsilon_{1}, \varepsilon_{3}) = \begin{cases} (1,0,0) & \text{if } \hat{f}(0) = 0, \\ (0,1,0) & \text{if } \hat{f}(0) = 2^{m-1-r_{f}/2}, \\ (0,0,1) & \text{if } \hat{f}(0) = -2^{m-1-r_{f}/2}. \end{cases}$$

$$(16)$$

Table 4: The weight distribution of the code C_{D_f} in Theorem 8

Weight w	A_w
0	1
$\frac{n_f}{2}$	$2^m-2^{r_f}-\varepsilon_1$
$\frac{n_f + 2^{m^2 - 1 - r_f/2}}{2}$	$2^{r_f-1} + 2^{(r_f-2)/2} - \varepsilon_2$
$\frac{n_f - 2^{m-1 - r_f/2}}{2}$	$2^{r_f-1}-2^{(r_f-2)/2}-\varepsilon_3$

Note that the code C_{D_f} in Theorem 8 defined by any quadratic Boolean function f is different from any subcode of the second-order Reed-muller code, due to the difference in their lengths. The weight distributions of the two codes are also different.

5.5. Some binary codes C_{D_f} with three weights

Theorem 9. Let $m \ge 4$ be even. Then the code C_{D_f} has parameters $[2^{m-1}, m, 2^{m-2} - 2^{(m-2)/2}]$ and the weight distribution of Table 6, where e = 2, for $f(x) = \text{Tr}(x^d)$ for the following d:

- 1. $d = 2^h + 1$, where gcd(m,h) is odd and $1 \le h \le m/2$ [41].
- 2. $d = 2^{2h} 2^h + 1$, where gcd(m,h) is odd and $1 \le h \le m/2$ [46].
- 3. $d = 2^{m/2} + 2^{(m+2)/4} + 1$, where $m \equiv 2 \pmod{4}$ [23].

Table 5: Boolean functions with three-valued Walsh spectrum

$\hat{f}(w)$	the number of w's
0	$2^{m}-2^{m-e}$
$2^{(m+e)/2}$	$2^{m-e-1} + 2^{(m-e-2)/2}$
$-2^{(m+e)/2}$	$2^{m-e-1} - 2^{(m-e-2)/2}$

4.
$$d = 2^{(m+2)/2} + 3$$
, where $m \equiv 2 \pmod{4}$ [23].

Proof. It can be verified that $gcd(d, 2^m - 1) = 1$ for all the d listed above. Hence $n_f = |D_f| = 2^{m-1}$. The Walsh spectrum of the functions f above is given in Table 5 according to the references given in this theorem. The desired conclusions on the parameters and the weight distribution of the code \mathcal{C}_{D_f} then follow from Theorem 1.

Table 6: The weight distribution of some three-weight codes

Weight w	Multiplicity A _w
0	1
2^{m-2}	$2^m - 2^{m-e} - 1$
$2^{m-2} + 2^{(m+e-4)/2}$	$2^{m-e-1} + 2^{(m-e-2)/2}$
$2^{m-2} - 2^{(m+e-4)/2}$	$2^{m-e-1} - 2^{(m-e-2)/2}$

5.6. Binary codes C_{D_f} with four weights

Table 7: Boolean functions with four-valued Walsh spectrum: Case I

$\hat{f}(w)$	the number of w's
$-2^{m/2}$	$(2^m - 2^{m/2})/3$
0	$2^{m-1} - 2^{(m-2)/2}$
$2^{m/2}$	$2^{m/2}$
$2^{(m+2)/2}$	$(2^{m-1}-2^{(m-2)/2})/3$

The code C_{D_f} has four weights and its weight distribution is known when $f(x) = \text{Tr}(x^d)$ and d is given in the following list.

- When $d = 2^{(m+2)/2} 1$ and $m \equiv 0 \pmod{4}$, the code C_{D_f} has length 2^{m-1} and dimension m, and the weight distribution of C_{D_f} is deduced from Theorem 1 and Table 7, where h = 1 [61].
- When $d=2^{(m+2)/2}-1$ and $m\equiv 2\pmod 4$, the code \mathcal{C}_{D_f} has length $2^{m-1}-2^{m/2}$ and dimension m, and the weight distribution of \mathcal{C}_{D_f} is deduced from Theorem 1 and Table 9 [61]. Note that in this case, $\gcd(d,2^m-1)=3$.

Table 8: Boolean functions with four-valued Walsh spectrum: Case 2

	$\hat{f}(w)$	the number of w's
_	$-2^{m/2}$	$2^{m-1} - 2^{(3m-4)/4}$
	0	$2^{3m/4} - 2^{m/4}$
	$2^{m/2}$	$2^{m-1} - 2^{(3m-4)/4}$
2	$2^{3m/4}$	$2^{m/4}$

Table 9: Boolean functions with four-valued Walsh spectrum: Case 3

$\hat{f}(w)$	the number of w's
$-2^{m/2}$	$(2^m - 2^{m/2} - 2)/3$
0	$2^{m-1} - 2^{(m-2)/2} + 2$
$2^{m/2}$	$2^{m/2}-2$
$2^{(m+2)/2}$	$(2^{m-1}-2^{(m-2)/2}+2)/3$

- When $d = (2^{m/2} + 1)(2^{m/4} 1) + 2$ and $m \equiv 0 \pmod{4}$, the code C_{D_f} has length 2^{m-1} and dimension m, and the weight distribution of C_{D_f} is deduced from Theorem 1 and Table 8 [61].
- When $d = \frac{2^{(m+2)h/2}-1}{2^h-1}$ and $m \equiv 0 \pmod{4}$, where $1 \leq h < m$ and $\gcd(h,m) = 1$, the code \mathcal{C}_{D_f} has length 2^{m-1} and dimension m, and the weight distribution of \mathcal{C}_{D_f} is deduced from Theorem 1 and Table 7, where h = 1 [35].
- When $d = \frac{2^{(m+2)h/2}-1}{2^h-1}$ and $m \equiv 2 \pmod 4$, where $1 \le h < m$ and $\gcd(h,m) = 1$, the code \mathcal{C}_{D_f} has length $2^{m-1} 2^{m/2}$ and dimension m, and the weight distribution of \mathcal{C}_{D_f} is deduced from Theorem 1 and Table 9 [61]. Note that in this case, $\gcd(d, 2^m 1) = 3$.
- When $d = \frac{2^m + 2^{h+1} 2^{m/2+1} 1}{2^h 1}$, where 2h divides m/2 and $m \equiv 0 \pmod{4}$, the code C_{D_f} has length 2^{m-1} and dimension m, and the weight distribution of C_{D_f} is deduced from Theorem 1 and Table 7 [43].
- When $d = (2^{m/2} 1)s + 1$ with $s = 2^h(2^h \pm 1)^{-1} \pmod{2^{m/2} + 1}$, where $e_2(h) < e_2(m/2)$ and $e_2(h)$ denotes the highest power of 2 dividing h, the parameters and the weight distribution of the code C_{D_f} can be deduced from Theorem 1 and the results in [36].
- Let d be any integer such that $1 \le d \le 2^m 2$ and $d(2^\ell + 1) \equiv 2^h \pmod{2^m 1}$ for some positive integers ℓ and h. Then the parameters and the weight distribution of the code C_{D_f} can be deduced from Theorem 1 and the results in [42].

All these cases of d above are derived from the cross-correlation of a binary maximum-length sequence with its d-decimation version.

5.7. Other binary codes C_{D_f} with at most five weights

The code C_{D_f} has at most five weights for the following f:

Table 10: The weight distribution of the codes of Theorem 10

weight w	Multiplicity A _w
0	1
$\frac{q-\sqrt{q}}{4} + \frac{q+\sqrt{q}}{4r^m}(r^m - 2r + 2)$	$\frac{(q-1)(r^m-r+1)}{r^m}$
$\frac{q+\sqrt{q}}{4} + \frac{q+\sqrt{q}}{4r^m}(r^m - 2r + 2)$	$\frac{(q-1)(r-1)}{r^m}$

- When $f(x) = \text{Tr}(x^{2^{m/2}+3})$, where $m \ge 6$ and is even, C_{D_f} is a five-weight code with length 2^{m-1} and dimension m, and its weight distribution can be derived from [34].
- When $f(x) = \text{Tr}(ax^{(2^m-1)/3})$ with $\text{Tr}_4^{2^m}(a) \neq 0$, where m is even, C_{D_f} is a two-weight code with length $(2^{m+2}-4)/6$ and dimension m, and its weight distribution can be derived from [48].
- When $f(x) = \text{Tr}(\lambda x^{2^{m/2}+1}) + \text{Tr}(x)\text{Tr}(\mu x^{2^{m/2}-1})$, where m is even, $\mu \in \text{GF}(2^{m/2})^*$, and $\lambda \in \text{GF}(2^m)$ with $\lambda + \lambda^{2^m} = 1$, \mathcal{C}_{D_f} is a five-weight code [3].
- When $f(x) = (1 + \text{Tr}(x))\text{Tr}(\lambda x^{2^{m/2} + 1}) + \text{Tr}(x)\text{Tr}(\mu x^{2^{m/2} 1})$, where m is even, $\mu \in \text{GF}(2^{\frac{m}{2}})^*$, and $\lambda \in \text{GF}(2^m)$ with $\lambda + \lambda^{2^m} = 1$, C_{D_f} is a five-weight code [3].

Some Boolean functions f documented in [63] gives also binary linear code C_{D_f} with five weights.

5.8. A class of two-weight codes from the preimage of a type of Boolean functions

Let m be a positive integer and let r be a prime such that 2 is a primitive root modulo r^m . Let $q = 2^{\phi(r^m)}$, where ϕ is the Euler function. Define

$$D = \left\{ x \in \mathrm{GF}(q)^* : \mathrm{Tr}\left(x^{\frac{q-1}{r^m}}\right) = 0 \right\}. \tag{17}$$

The following theorem was proved in [69].

Theorem 10. Let $r^m \ge 9$ and let D be defined in (17). Then the set C_D of (6) is a binary code with length $(q-1)(r^m-r+1)/r^m$, dimension $(r-1)r^{m-1}$ and the weight distribution in Table 10.

5.9. Binary codes from Boolean functions whose supports are relative difference sets

Let (A, +) be an abelian group of order $m\ell$ and (N, +) a subgroup of A of order ℓ . A n-subset D of A is called an (m, ℓ, n, λ) relative difference set, if the multiset $\{\{d_1 - d_2 : d_1, d_2 \in D, d_1 \neq d_2\}\}$ does not contain all elements in N, but every element in $A \setminus N$ exactly λ times.

It well known in combinatorics that $|\chi(D)|^2 \in \{n, n - \lambda \ell\}$ for any nontrivial group character χ . Hence, any relative difference set D in $(GF(2^m), +)$ defines a binary code C_D with at most the following four weights:

$$\frac{n\pm\sqrt{n}}{2}$$
, $\frac{n\pm\sqrt{n-\lambda\ell}}{2}$.

Obviously, D is the support of a Boolean function on $GF(2^m)$.

6. Binary codes from the images of certain functions on $GF(2^m)$

Let f(x) be a function from $GF(2^m)$ to $GF(2^m)$. We define

$$D(f) = \{f(x) : x \in GF(2^m)\} \text{ and } D(f)^* = \{f(x) : x \in GF(2^m)\} \setminus \{0\}.$$

In this section, we consider the code $C_{D(f)}$. In general, it is difficult to determine the length $n_f := |D(f)|$ of this code, not to mention its weight distribution. However, in certain special cases, the parameters and the weight distribution of $C_{D(f)}$ can be settled.

If $0 \notin D(f)$, then the two codes $C_{D(f)}$ and $C_{D(f)^*}$ are the same. Otherwise, the length of the code $C_{D(f)^*}$ is one less than that of the code $C_{D(f)}$, but the two codes have the same weight distribution. Thus, we will not give information about the codes $C_{D(f)^*}$ in this section.

Let D be any subset of $GF(2^m)$. The *characteristic function*, denoted by $f_D(x)$, of D is defined by

$$f_D(x) = \begin{cases} 1 & \text{if } x \in D, \\ 0 & \text{otherwise.} \end{cases}$$

Hence, the Boolean function f_D has support D. Thus, the code $C_{D(f)}$ is in fact defined by the support of the characteristic function (Boolean function) of the set D(f). Therefore, the construction method of this section is actually equivalent to that of Section 5.

6.1. The codes $C_{D(f)}$ from o-polynomials on $GF(2^m)$

A permutation polynomial f on $GF(2^m)$ is called an *o-polynomial* if f(0) = 0, and for each $s \in GF(2^m)$,

$$f_s(x) = (f(x+s) + f(s))x^{2^m - 2}$$
(18)

is also a permutation polynomial. O-polynomial can be used to construct hyperovals in finite geometry.

In the original definition of o-polynomials, it is required that f(1) = 1. However, this is not essential, as one can always normalise f(x) by using $f(1)^{-1}f(x)$ due to that $f(1) \neq 0$.

In this section, we consider binary codes $C_{D(f)}$, where f is defined by an o-polynomial in some way.

6.1.1. O-polynomials and their binary codes $C_{D(f_n)}$

For any permutation polynomial f(x) over $GF(2^m)$, we define $\overline{f}(x) = xf(x^{2^m-2})$, and use f^{-1} to denote the compositional inverse of f, i.e., $f^{-1}(f(x)) = x$ for all $x \in GF(2^m)$.

The following two theorems introduce basic properties of o-polynomials whose proofs can be found in references about hyperovals. .

Theorem 11. Let f be an o-polynomial on $GF(2^m)$. Then the following statements hold:

- 1. f^{-1} is also an o-polynomial;
- 2. $f(x^{2^j})^{2^{m-j}}$ is also an o-polynomial for any $1 \le j \le m-1$;
- 3. \overline{f} is also an o-polynomial; and
- 4. f(x+1) + f(1) is also an o-polynomial.

Theorem 12. Let x^k be an o-polynomial on $GF(2^m)$. Then every polynomial in

$$\left\{x^{\frac{1}{k}}, x^{1-k}, x^{\frac{1}{1-k}}, x^{\frac{k}{k-1}}, x^{\frac{k-1}{k}}\right\}$$

is also an o-polynomial, where 1/k denotes the multiplicative inverse of k modulo $2^m - 1$.

The following property of o-polynomials plays an important role in our construction of binary linear codes with o-polynomials.

Theorem 13 ([15]). A polynomial f from $GF(2^m)$ to $GF(2^m)$ with f(0) = 0 is an o-polynomial if and only if $f_u := f(x) + ux$ is 2-to-1 for every $u \in GF(2^m)^*$.

Let f be any o-polynomial over $GF(2^m)$. Define

$$f_u(x) = f(x) + ux$$

where $u \in GF(2^m)^*$. It follows from Theorem 13 that f_u is 2-to-1 for every $u \in GF(2^m)^*$. In the rest of Section 6.1, we consider the codes $C_{D(f_u)}$ defined by o-polynomials. By Theorem 13, the o-polynomial property of f guarantees that the length of the code $C_{D(f_u)}$ is equal to 2^{m-1} for any $u \in GF(2^m)^*$. The dimension of $C_{D(f_u)}$ usually equals m, but may be less than m. The minimum weight and the weight distribution of $C_{D(f_u)}$ cannot be determined by the o-polynomial property alone, and differ from case to case.

6.1.2. Binary codes from the translation o-polynomials

The translation o-polynomials are described in the following theorem [66].

Theorem 14. Trans(x) = x^{2^h} is an o-polynomial on GF(2^m), where gcd(h, m) = 1.

The following is a list of known properties of translation o-polynomials.

- 1. Trans⁻¹ $(x) = x^{2^{m-h}}$ and
- 2. $\overline{\text{Trans}}(x) = xf(x^{2^m 2}) = x^{2^m 2^{m h}}$.

The proof of the following theorem is straightforward.

Theorem 15. Let $f(x) = x^{2^h}$, where gcd(h, m) = 1. Then for any $u \in GF(2^m)^*$, the code $C_{D(f_u)}$ has parameters $[2^{m-1}, m-1, 2^{m-2}]$ and is a one-weight code.

The codes $C_{D(f_u)}$ in Theorem 15 have the same parameters as a subcode of the first order binary Reed-Muller code.

6.1.3. Binary codes from the Segre and Glynn o-polynomials

The following theorem describes a class of o-polynomials, which are an extension of the original Segre o-polynomials.

Theorem 16 ([32]). Let m be odd. Then $\operatorname{Segre}_a(x) = x^6 + ax^4 + a^2x^2$ is an o-polynomial on $\operatorname{GF}(2^m)$ for every $a \in \operatorname{GF}(2^m)$.

Proof. The conclusion follows from

$$Segre_a(x) = (x + \sqrt{a})^6 + \sqrt{a}^3.$$

We have the following remarks on this family of o-polynomials.

- 1. Segre₀(x) = x⁶ is the original Segre o-polynomial [67, 68]. So this is an extended family.
- 2. Segre_a(x) = $xD_5(x,a) = a^2D_5(x^{2^m-2},a^{2^m-2})x^7$, where $D_5(x,a) = x^5 + ax^3 + a^2x$, which is the Dickson polynomial of the first kind of order 5.
- 3. $\overline{\text{Segre}}_a = D_5(x^{2^m-2}, a) = a^2 x^{2^m-2} + a x^{2^m-4} + x^{2^m-6}$
- 4. Segre_a⁻¹(x) = $(x + \sqrt{a^3})^{\frac{5 \times 2^{m-1} 2}{3}} + \sqrt{a}$.

Theorem 17 ([32]). *Let m be odd. Then*

$$\overline{\text{Segre}}_{1}^{-1}(x) = \left(D_{\frac{3 \times 2^{2m} - 2}{5}}(x, 1)\right)^{2^{m} - 2}.$$
(19)

Glynn discovered two families of o-polynomials [40]. The first is described as follows.

Theorem 18. Let m be odd. Then $Glynni(x) = x^{3 \times 2^{(m+1)/2} + 4}$ is an o-polynomial.

Conjecture 19. Let $m \ge 3$ be odd, and let $f(x) = x^{3 \times 2^{(m+1)/2} + 4}$ be the Glynn o-polynomial. When $m \in \{5,7\}$, $C_{D(f_u)}$ is a $[2^{m-1}, m]$ code with the weight distribution of Table 11. When $m \ge 9$, $C_{D(f_u)}$ is a $[2^{m-1}, m]$ code with five weights.

An extension of the second family of o-polynomials discovered by Glynn is documented in the following theorem.

Theorem 20 ([32]). *Let m be odd. Then*

$$\text{Glynnii}_{a}(x) = \begin{cases} x^{2^{(m+1)/2} + 2^{(3m+1)/4}} + ax^{2^{(m+1)/2}} + (ax)^{2^{(3m+1)/4}} & \text{if } m \equiv 1 \pmod{4}, \\ x^{2^{(m+1)/2} + 2^{(m+1)/4}} + ax^{2^{(m+1)/2}} + (ax)^{2^{(m+1)/4}} & \text{if } m \equiv 3 \pmod{4}. \end{cases}$$

is an o-polynomial for all $a \in GF(q)$.

Proof. Let $m \equiv 1 \pmod{4}$. Then

Glynnii_a(x) =
$$(x + a^{(m-1)/4})^{2^{(m+1)/2} + 2^{(3m+1)/4}} + a^{2^{(m+1)/2} + 2^{(3m+1)/4}}$$

The desired conclusion for the case $m \equiv 1 \pmod{4}$ can be similarly proved.

Note that $Glynnii_0(x)$ is the original Glynn o-polynomial. So this is an extended family. For some applications, the extended family may be useful.

For certain quadratic Boolean functions f, the code $C_{D(f_u)}$ has good parameters and its weight distribution is known. The following result is an extended version of a result proved in [28].

Theorem 21. Let $\rho = 2^i + 2^j$ and define

$$f_u(x) = x^{\rho} + ux, u \in GF(2^m)^*.$$

If $f_u(x)$ is 2-to-1 on $GF(2^m)$ and $gcd(2^{\kappa}+1,2^m-1)=1$, where $\kappa=j-i$, then the binary code $C_{D(f_u)}$ has parameters $[2^{m-1},m,2^{m-2}-2^{(m-3)/2}]$ and the weight distribution of Table 11 for any $u \in GF(2^m)^*$.

The following ρ satisfies the conditions of Theorem 21:

Table 11: The weight distribution of the codes of Theorem 21

Weight w	Multiplicity A_w
0	1
$2^{m-2} - 2^{(m-3)/2}$	$2^{m-2} + 2^{(m-3)/2}$
2^{m-2}	$2^{m-1}-1$
$2^{m-2} + 2^{(m-3)/2}$	$2^{m-2} - 2^{(m-3)/2}$

- $\rho = 6$ (Segre case).
- $\rho = 2^{\sigma} + 2^{\pi}$ with $\sigma = (m+1)/2$ and $4\pi \equiv 1 \mod m$ (Glynn I case).

Theorem 22. Let $f(x) = xD_5(x,a)$, where $a \in GF(2^m)$, and let m be odd. Then the code $C_{D(f_n)}$ has parameters $[2^{m-1}, m]$ and the weight distribution of Table 11 for any $u \in GF(2^m)^*$.

Proof. It is easily verified that $f(x) = (x + \sqrt{a})^6 + \sqrt{a^3}$. The desired conclusions then follow from Theorem 21 in the Segre case, i.e., $\rho = 6$.

Theorem 23. Let $F(x) = x^{(5 \times 2^{m-1} - 2)/3}$, and let m be odd. Then code $C_{D(f_u)}$ has parameters $[2^{m-1}, m]$ and the weight distribution of Table 11 for any $u \in GF(2^m)$.

Proof. Note that the multiplicative inverse of 6 modulo $2^m - 1$ is equal to $(5 \times 2^{m-1} - 2)/3$. The desired conclusion then follows from Theorem 21.

6.1.4. Binary codes from the Cherowitzo o-polynomials

The following describes another conjectured class of o-polynomials.

Conjecture 24 ([32]). *Let m be odd and e* = (m+1)/2. *Then*

Cherowitzo_a
$$(x) = x^{2^e} + ax^{2^e+2} + a^{2^e+2}x^{3\times 2^e+4}$$

is an o-polynomial on $GF(2^m)$ for every $a \in GF(2^m)$.

We have the following remarks on this family.

- 1. Cherowitzo $_1(x)$ is the original Cherowitzo o-polynomial [17, 18]. So this is an extended
- 2. No proof of the o-polynomial property is known in the literature.

 3. Cherowitzo(x) = $x^{2^m-2^e} + ax^{2^m-2^e-2} + a^{2^e+2}x^{2^m-3\times 2^e-4}$.
- 4. Carlet and Mesnager showed that Cherowitzo₁⁻¹ $(x) = x(x^{2^e+1} + x^3 + x)^{2^{e-1}-1}$.

We can prove the following.

Theorem 25 ([32]).

Cherowitzo_a⁻¹(x) =
$$x(ax^{2^e+1} + a^{2^e}x^3 + x)^{2^{e-1}-1}$$
.

Theorem 26 ([32]).

$$\overline{\text{Cherowitzo}} = (ax^{2^m - 2^e - 2} + a^{2^e}x^{2^m - 4} + x^{2^m - 2})^{2^{e - 1} - 1}.$$

Conjecture 27. Let m be odd, and let

$$f(x) = b^{2^{(m+1)/2} + 2} x^{2^{(m+1)/2}} + b^{2^{(m+1)/2} + 1} x^{2^{(m+1)/2} + 2} + x^{3 \times 2^{(m+1)/2} + 4},$$

where $b \in GF(2^m)$. If $m \in \{5,7\}$, $C_{D(f_u)}$ is a $[2^{m-1}, m]$ code with at most five weights for every $u \in GF(2^m)^*$. If $m \ge 9$, $C_{D(f_u)}$ is a five-weight code with length 2^{m-1} and dimension m for every $u \in GF(2^m)^*$.

6.1.5. Binary codes from the Payne o-polynomials

The following documents a conjectured family of o-trinomials.

Conjecture 28 ([32]). Let m be odd. Then $\operatorname{Payne}_a(x) = x^{\frac{5}{6}} + ax^{\frac{3}{6}} + a^2x^{\frac{1}{6}}$ is an a-polynomial on $\operatorname{GF}(2^m)$ for every $a \in \operatorname{GF}(2^m)$.

We have the following remarks on this family.

- 1. Payne₁(x) is the original Payne o-polynomial [62]. So this is an extended family.
- 2. Payne_a(x) = $xD_5(x^{\frac{1}{6}}, a)$.
- 3. $\overline{\text{Payne}}_{a}^{a}(x) = a^{2^{m}-3} \text{Payne}_{a^{2^{m}-2}}(x)$.
- 4. Note that

$$\frac{1}{6} = \frac{5 \times 2^{m-1} - 2}{3}.$$

We have then

Payne_a(x) =
$$x^{\frac{2^{m-1}+2}{3}} + ax^{2^{m-1}} + a^2x^{\frac{5\times 2^{m-1}-2}{3}}$$
.

Theorem 29 ([32]). *Let m be odd. Then*

Payne₁⁻¹(x) =
$$\left(D_{\frac{3\times 2^{2m}-2}{5}}(x,1)\right)^6$$
 (20)

and $\overline{\text{Payne}}_1^{-1}(x)$ are an o-polynomial.

Proof. Note that the multiplicative inverse of 5 modulo is $\frac{3 \times 2^{2m} - 2}{5}$. The conclusion then follows from the definition of the Payne polynomial and the fact that

$$D_5(x,1)^{-1} = D_{\frac{3\times 2^{2m}-2}{5}}(x,1).$$

Conjecture 30 ([32]). *Let m be odd, and let*

$$f(x) = x^{\frac{5}{6}} + bx^{\frac{3}{6}} + b^2x^{\frac{1}{6}} = D_5\left(x^{\frac{5\times 2^{m-1}-2}{3}}, b\right),$$

where $b \in GF(2^m)$. If $m \ge 7$, $C_{D(f_u)}$ is a three-weight or five-weight code with length 2^{m-1} and dimension m for all $u \in GF(2^m)^*$.

6.1.6. Binary codes from the Subiaco o-polynomials

The Subiaco o-polynomials are given in the following theorem [19].

Theorem 31. Define

Subiaco_a(x) =
$$((a^2(x^4 + x) + a^2(1 + a + a^2)(x^3 + x^2))(x^4 + a^2x^2 + 1)^{2^m - 2} + x^{2^{m-1}},$$

where Tr(1/a) = 1 and $d \notin \text{GF}(4)$ if $m \equiv 2 \mod 4$. Then $\text{Subiaco}_a(x)$ is an o-polynomial on $\text{GF}(2^m)$.

As a corollary of Theorem 31, we have the following.

Corollary 32. Let m be odd. Then

Subiaco₁(x) =
$$(x + x^2 + x^3 + x^4)(x^4 + x^2 + 1)^{2^m - 2} + x^{2^{m-1}}$$
 (21)

is an o-polynomial over $GF(2^m)$.

Experimental data shows that the binary codes $C_{D(f_u)}$ from the Subiaco o-polynomials have many weights and have smaller minimum weights compared with binary codes from other o-polynomials described in the previous subsections. Hence, the binary code $C_{D(f_u)}$ from an o-polynomial could be very good and bad, depending on the specific o-polynomial f.

6.2. Binary codes $C_{D(f)}$ from functions on $GF(2^m)$ of the form f(x) = F(x) + F(x+1) + 1A function F(x) over $GF(2^m)$ is called almost perfect nonlinear (APN) if

$$\max_{a \in GF(2^m)^*} \max_{b \in GF(2^m)} |\{x \in GF(2^m) : F(x+a) - F(x) = b\}| = 2.$$

Let F be any function on $GF(2^m)$. Define

$$f(x) = F(x) + F(x+1) + 1.$$

For certain APN functions F(x) over $GF(2^m)$, it is known that f is 2-to-1.

Conjecture 33. Let $F(x) = x^{2^{(m-1)/2}+3}$ and m be odd. It is known that F is both APN and AB. If $m \in \{5,7\}$, $C_{D(f)}$ is a three-weight code with length 2^{m-1} and dimension m. If $m \ge 9$, $C_{D(f)}$ is a five-weight code with length 2^{m-1} and dimension m.

Conjecture 34. Let $F(x) = x^{2^{2h}-2^h+1}$, and let gcd(h, m) = 1. It is known that F is both APN and AB

When h = 1, $C_{D(f)}$ is a $[2^{m-1}, m-1, 2^{m-2}]$ one-weight code.

When $h \ge 2$ and m is odd, $C_{D(f)}$ is a three-weight or five-weight code with length 2^{m-1} and dimension m.

In particular, when h = 3 and m is odd, $C_{D(f)}$ is a three-weight code with length 2^{m-1} and dimension m for every odd $m \ge 5$ and $m \not\equiv 0 \pmod{3}$. In this case, d = 57 and the weight distribution of the code $C_{D(f)}$ is given in Table 11.

It is known that
$$f(x) = x^{2^{2h}-2^h+1} + (x+1)^{2^{2h}-2^h+1}$$
 is 2^s -to-1, where $s = \gcd(h, m)$ [25].

Theorem 35. Let $F(x) = x^{2^h+1}$, and let gcd(h,m) = 1. Then $C_{D(f)}$ is a one-weight code with parameters $[2^{m-1}, m-1, 2^{m-2}]$.

Proof. The proof is straightforward and omitted, as f(x) is linear.

We have the following comments on other APN monomials.

- a) Let $F(x) = x^{2^m-2}$. Then $C_{D(f)}$ is a binary code with length 2^{m-1} and dimension m, and has at most m weights. The weights are determined by the Kloosterman sums.
- b) For the Niho function $F(x) = x^{2^{(m-1)/2} + 2^{(m-1)/4} 1}$, where $m \equiv 1 \pmod{4}$, the code $C_{D(f)}$ has length 2^{m-1} and dimension m, but many weights.
- c) For the Niho function $F(x) = x^{2^{(m-1)/2} + 2^{(3m-1)/4} 1}$, where $m \equiv 3 \pmod{4}$, the code $C_{D(f)}$ has length 2^{m-1} and dimension m, but many weights.

It would be extremely difficult to determine the weight distribution of the code $C_{D(f)}$ for these three classes of APN monomials.

6.3. Binary linear codes from some trinomials

A lot of constructions of cyclic difference sets in $(GF(2^m)^*, \times)$ with the Singer parameters $(2^m-1, 2^{m-1}, 2^{m-2})$ or $(2^m-1, 2^{m-1}-1, 2^{m-2}-1)$ are proposed in the literature [26, 27]. These difference sets can certainly be plugged into the second generic construction of this paper and obtain binary linear codes with good parameters. But determining the parameters of the binary linear codes may be difficult in general.

There are also a number of conjectured cyclic difference sets in $(GF(2^m)^*, \times)$. They give naturally binary linear codes with this construction. The following is a list of conjectured cyclic difference sets in $(GF(2^m)^*, \times)$ with Singer parameters (see Chapter 4 of [27]).

Conjecture 36. [27, Chapter 4] For any $f \in GF(2^m)[x]$, we define

$$D(f)^* = \{f(x) : x \in GF(2^m)\} \setminus \{0\}.$$

Let $m \ge 5$ be odd. Then $D(f)^*$ is a difference set in $(GF(2^m)^*, \times)$ with Singer parameters $(2^m - 1, 2^{m-1}, 2^{m-2})$ for the following trinomials $f \in GF(2^m)[x]$:

- a) $f(x) = x^{2^m 17} + x^{(2^m + 19)/3} + x$.
- b) $f(x) = x^{2^m 2^{m-4} 1} + x^{2^m (2^{m-2} + 4)/3} + x$.
- c) $f(x) = x^{2^m 3} + x^{2^{(m+3)/2} + 2^{(m+1)/2} + 4} + x$.
- d) $f(x) = x^{2^m 2^{(m-1)/2} 1} + x^{2^{m-1} 2^{(m-1)/2}} + x$.
- e) $f(x) = x^{2^m 2 (2^{m-1} 2^2)/3} + x^{2^m 2^2 (2^m 2^3)/3} + x$.
- f) $f(x) = x^{2^m 2^{(m+1)/2} + 2^{(m-1)/2}} + x^{2^m 2^{(m+1)/2} 1} + x$.
- g) $f(x) = x^{2^m 3(2^{(m+1)/2} 1)} + x^{2^{(m+1)/2} + 2^{(m-1)/2} 2} + x$.
- h) $f(x) = x^{2^m 2^{m-2} 1} + x^{2^{m-1} 2} + x$.
- i) $f(x) = x^{2^m 2^{(m+3)/2} 3} + x^{2^{(m+1)/2} + 2} + x$.
- i) $f(x) = x^{2^m 3(2^{(m-1)/2} + 1)} + x^{2^{m-1} 1} + x$.
- k) $f(x) = x^{2^m-5} + x^6 + x$.

For the linear codes $C_{D(f)^*}$ of the conjectured difference sets $D(f)^*$ in Conjecture 36, we have the following conjectured parameters.

Conjecture 37. Let $m \ge 5$ and let $D(f)^*$ be defined as in Conjecture 36. Then for every f given in Conjecture 36, the binary linear code $C_{D(f)^*}$ has parameters $[2^{m-1}, m, 2^{m-2} - 2^{(m-3)/2}]$ and weight enumerator

$$1 + (2^{m-2} - 2^{(m-3)/2})z^{2^{m-2} - 2^{(m-3)/2}} + (2^{m-1} - 1)z^{2^{m-2}} + (2^{m-2} + 2^{(m-3)/2})z^{2^{m-2} + 2^{(m-3)/2}}.$$

The dual code of $C_{D(f)^*}$ has parameters $[2^{m-1}, 2^{m-1} - m, 3]$.

Conjecture 37 describes binary three-weight codes for the case that m is odd. The next one is about binary three-weight codes for the case that m is even.

Conjecture 38. Let $f(x) = x + x^{2^{(m-2)/2} + 2^{m-1}} + x^{2^{(m-2)/2} + 2^{m-1} + 1} \in GF(2^m)[x]$, where $m \equiv 2 \mod 4$ and $m \ge 6$. Define

$$D(f) = \{ f(x) : x \in GF(2^m) \}.$$

Then the binary code $C_{D(f)}$ has parameters $[2^{m-1}, m, 2^{m-2} - 2^{(m-2)/2}]$ and weight enumerator

$$1 + (2^{m-3} + 2^{(m-4)/2})z^{2^{m-2} - 2^{(m-2)/2}} + (3 \times 2^{m-2} - 1)z^{2^{m-2}} + (2^{m-3} - 2^{(m-4)/2})z^{2^{m-2} + 2^{(m-2)/2}}.$$

It was conjectured in [27, Chapter 4] that $D(f)^*$ is a difference set in $(GF(2^m)^*, \times)$ with the parameters $(2^m - 1, 2^{m-1} - 1, 2^{m-2} - 1)$.

Table 12: The weight distribution of the codes of Conjecture 39

Weight w	Multiplicity A_w
0	1
$2^{m-2} - 2^{(m-2)/2}$	$2^{(m-2)/2}$
$2^{m-2} - 2^{(m-4)/2}$	$2^{m-1}-2^{m/2}$
2^{m-2}	$2^{m/2} + 2^{(m-2)/2} - 1$
$2^{m-2} + 2^{(m-4)/2}$	$2^{m-1}-2^{m/2}$

Binary four-weight codes may also be produced with difference sets in $(GF(2^m)^*, \times)$ as follows.

Conjecture 39. Let $f(x) = x + x^2 + x^{2^m - 2^{m/2} + 1} \in GF(2^m)[x]$, where $m \ge 4$ and m is even. Define

$$D(f) = \{ f(x) : x \in GF(2^m) \}.$$

Then the binary linear code $C_{D(f)}$ has parameters $[2^{m-1}, m, 2^{m-2} - 2^{(m-2)/2}]$ and the weight distribution of Table 12. It was conjectured in [27, Chapter 4] that $D(f)^*$ is a difference set in $(GF(2^m)^*, \times)$ with parameters $(2^m - 1, 2^{m-1} - 1, 2^{m-2} - 1)$.

The following is an another list of conjectured cyclic difference sets in $(GF(2^m)^*, \times)$ with Singer parameters (see Chapter 4 of [27]).

Conjecture 40. [27, Chapter 4] For any $f \in GF(2^m)[x]$, we define

$$D(f) = \{ f(x(x+1)) : x \in GF(2^m) \}.$$

Let $m \ge 4$. Then $D(f)^*$ is a difference set in $(GF(2^m)^*, \times)$ with Singer parameters $(2^m - 1, 2^{m-1} - 1, 2^{m-2} - 1)$ for the following polynomials $f \in GF(2^m)[x]$:

- 1. $f(x) = x + x^{2^{(m+1)/2}-1} + x^{2^m-2^{(m+1)/2}+1}$, where m is odd.
- 2. $f(x) = x + x^{(2^m+1)/3} + x^{(2^{m+1}-1)/3}$, where m is odd.
- 3. $f(x) = x + x^{2^{(m+2)/2}-1} + x^{2^m-2^{m/2}+1}$, where m is even.

For the linear codes $C_{D(f)^*}$ of the conjectured difference sets $D(f)^*$ in Conjecture 40, we have the following conjectured parameters.

Conjecture 41. Let $m \ge 4$ and let D(f) be defined as in Conjecture 40. Then for every f given in Conjecture 40, the binary linear code $C_{D(f)^*}$ has parameters $[2^{m-1}-1, m-1, 2^{m-2}]$ and is a one-weight code.

To determine the weight distribution of the code $C_{D(f)}$ or $C_{D(f)}^*$ of the conjectured difference sets listed in this section, one does not have to prove the difference set property of the set D(f) or $D(f)^*$.

7. An expansion of the binary codes

Table 13: The weight distribution of the codes of Theorem 42

Weight w	Multiplicity A_w
0	1
$2^{m-2} - 2^{(m-3)/2}$	$2^{m-1} + 2^{(m-1)/2}$
2^{m-2}	$2^{m}-2$
$2^{m-2} + 2^{(m-3)/2}$	$2^{m-1} - 2^{(m-1)/2}$
2^{m-1}	1

Let **1** denote the all-one vector, that is, $(1,1,\ldots,1)$, of any length. The complement of any vector $\mathbf{c} \in \mathrm{GF}(2)^n$ is defined to be $\mathbf{c}+\mathbf{1}$. For any binary code \mathcal{C} , we define

$$\overline{C} = C \cup \{c+1 : c \in C\}.$$

Then \overline{C} is a binary linear code, which has the same length as C. For most of the binary codes C presented in this paper, the dimension of \overline{C} is one more than that of C. In many cases, the weight distribution of \overline{C} can be deduced from that of C. As an example, we have the following.

Theorem 42. Let m be odd and let C be any binary linear code with parameters $[2^{m-1}, m]$ and the weight distribution of Table 11. Then \overline{C} is binary linear code with parameters $[2^{m-1}, m+1]$ and the weight distribution of Table 13.

Example 43. When m = 5, the code \overline{C} of Theorem 42 has parameters [16,6,6] and is optimal. When m = 7, the code \overline{C} of Theorem 42 has parameters [64,8,28] and is almost optimal.

8. Concluding remarks

In this paper, we surveyed binary linear codes from Boolean functions and functions on $GF(2^m)$ obtained from the second generic construction. Our focus was on such binary linear codes with at most five weights. Many one-weight codes, two-weight codes, three-weight codes, four-weight codes are presented in this paper. Some of them are optimal and some are almost optimal. The codes are also quite interesting in the sense that they may have applications in secret sharing [1, 14, 71] and authentication codes [31]. The parameters of some of the binary codes are different from those in [7], [6], [22], [20], [38], [49], [50], [70], and [72].

A number of conjectures were presented in this paper as open problems. All the conjectures on difference sets, o-polynomials and the corresponding binary codes were confirmed for sufficiently many integers *m* by Magma. The reader is warmly invited to attack these open problems.

Finally, we make it clear that this is by no means a survey of all binary linear codes from Boolean functions, but a survey of binary linear codes from Boolean functions from the second generic construction described in Section 4.

References

- [1] R. Anderson, C. Ding, T. Helleseth, T. Kløve, How to build robust shared control systems, Des. Codes Cryptogr. 15(2) (1998) 111–124.
- [2] E. F. Assmus, J. D. Key, Designs and Their Codes, Cambridge Tracts in Mathematics, Vol. 103, Cambridge University Press, Cambridge, 1992.
- [3] X. Cao, L. Hu, Two Boolean functions with five-valued Walsh spectrum and high nonlinearity, preprint 2014.
- [4] L. Budaghyan, C. Carlet, T. Helleseth, A. Kholosha, S. Mesnager, Further results on Niho bent functions, IEEE Trans. Inf. Theory 58(11) (2012) 6979–6985.
- [5] L. Budaghyan, C. Carlet, A. Pott, New classes of almost bent and almost perfect nonlinear polynomials, IEEE Trans. Inf. Theory 52(3) (2006) 11411152.
- [6] A. R. Calderbank, J. M. Goethals, Three-weight codes and association schemes, Philips J. Res. 39 (1984) 143–152.
- [7] A. R. Calderbank, W. M. Kantor, The geometry of two-weight codes, Bull. London Math. Soc. 18 (1986) 97–122.
- [8] A. Canteaut, P. Charpin, H. Dobbertin, Weight divisibility of cyclic codes, highly nonlinear functions on F_{2^m} , and crosscorrelation of maximum-length sequences, SIAM J. Discrete Math. 12(1) (2000) 105–138.
- [9] A. Canteaut, P. Charpin, G. M. Kyureghyan, A new class of monomial bent functions, Finite Fields Appl. 14 (2008) 221–241.
- [10] C. Carlet, A simple description of Kerdock codes, Lect. Notes Computer Science 388 (1989) 202208.
- [11] C. Carlet, The automorphism groups of the Kerdock codes, J. Information and Optimization Sciences 12(3) (1991) 387–400.
- [12] C. Carlet, Boolean functions for cryptography and error-correcting codes, in: Y. Crama and P. Hammer (Eds.), Boolean Models and Methods in Mathematics, Computer Science, and Engineering, Cambridge University Press, Cambridge, 2010, pp. 257–397.
- [13] C. Carlet, P. Charpin, V. Zinoviev, Codes, bent functions and permutations for DES-like cryptosystems, Des. Codes Cryptogr. 15 (1998) 125–156.
- [14] C. Carlet, C. Ding, J. Yuan, Linear codes from perfect nonlinear mappings and their secret sharing schemes, IEEE Trans. Inf. Theory 51(6) (2005) 2089–2102.
- [15] C. Carlet, S. Mesnager, On Dillion's class H of bent functions, Niho bent functions and o-polynomials, J. Combinatorial Theory Series A 118 (2011) 2392–2410.
- [16] C. Carlet, S. Mesnager, On semibent functions, IEEE Trans. Inf. Theory 58(5) (2012) 3287–3292.
- [17] W. Cherowitzo, Hyperovals in Desarguesian planes of even order, Annals of Discrete Mathematics 37 (1988) 87–
- [18] W. Cherowitzo, Hyperovals in Desarguesian planes: an update, Discrete Math. 155 (1996) 31-38.
- [19] W. Cherowitzo, T. Penttila, I. Pinneri, G.F. Royle, Flocks and ovals, Geometriae Dedicata 60 (1996) 17–37.
- [20] S.-T. Choi, J.-Y. Kim, J.-S. No, H. Chung, Weight distribution of some cyclic codes, in: Proc. of the 2012 International Symposium on Information Theory, IEEE Press, 2012, pp. 2911–2913.
- [21] G. Cohen, S. Mesnager, On constructions of semi-bent functions from bent functions, J. Contemporary Mathematics, vol. 625, Discrete Geometry and Algebraic Combinatorics, American Mathematical Society, 2014, pp. 141–154.

- [22] B. Courteau, J. Wolfmann, On triple-sum-sets and two or three weight codes, Discrete Mathematics 50 (1984) 179–189.
- [23] T. W. Cusick, H. Dobbertin, Some new three-valued crosscorrelation functions for binary m-sequences, IEEE Trans. Inf. Theory 42(4) (1996) 1238–1240.
- [24] P. Delsarte, On subfield subcodes of modified Reed-Solomon codes, IEEE Trans. Inf. Theory 21(5) (1975) 575–576.
- [25] D. Hertel, A note on the Kasami power function, http://eprint.iacr.org/2005/436.
- [26] J. F. Dillon, H. Dobbertin, New cyclic difference sets with Singer parameters, Finite Fields Appl. 10 (2004) 342–389.
- [27] C. Ding, Codes from Difference Sets, World Scientific, Singapore, 2015.
- [28] C. Ding, Linear codes from some 2-designs, IEEE Trans. Inf. Theory 60(6) (2015) 3265–3275.
- [29] C. Ding, J. Luo, H. Niederreiter, Two weight codes punctured from irreducible cyclic codes, in: Y. Li, S. Ling, H. Niederreiter, H. Wang, C. Xing, S. Zhang (Eds.), Proc. of the First International Workshop on Coding Theory and Cryptography, World Scientific, Singapore, 2008, pp. 119–124.
- [30] C. Ding, H. Niederreiter, Cyclotomic linear codes of order 3, IEEE Trans. Inf. Theory 53(6) (2007) 2274–2277.
- [31] C. Ding, X. Wang, A coding theory construction of new systematic authentication codes, Theoretical Computer Science 330 (2005) 81–99.
- [32] C. Ding and P. Yuan, Five constructions of permutation polynomials over $GF(q^2)$, preprint.
- [33] K. Ding and C. Ding, Binary linear codes with three weights, IEEE Communication Letters 18(11) (2014) 1879– 1882
- [34] T. Helleseth, Some results about the cross-correlation between two maximal linear sequences, Discrete Mathematics 16 (1976) 209–232.
- [35] H. Dobbertin, One-to-one highly nonlinear power functions on GF(2ⁿ), Appl. Algebra Eng. Commun. Comput. 9(2) (1998) 139–152.
- [36] H. Dobbertin, P. Felke, T. Helleseth, P. Rosendahl, Niho type cross-correlation functions via Dickson polynomials and Kloosterman sums, IEEE Trans. Inf. Theory 52(2) (2006) 613–627.
- [37] D. Dong, L. Qu, S. Fu, C. Li, New constructions of semi-bent functions in polynomial forms, Mathematical and Computer Modelling 57 (2013) 1139–1147.
- [38] K. Feng, J. Luo, Value distribution of exponential sums from perfect nonlinear functions and their applications, IEEE Trans. Inf. Theory 53 (2007) 3035–3041.
- [39] K. Feng, J. Luo, Weight distribution of some reducible cycliccodes, Finite Fields Appl. 14(2) (2008) 390-409.
- [40] D. G. Glynn, Two new sequences of ovals in finite Desarguesian planes of even order, in: L. R. A. Casse (Ed.), Combinatorial Mathematics X, Lecture Notes in Mathematics 1036, Springer, 1983, pp. 217–229.
- [41] R. Gold, Maximal recursive sequences with 3-valued recursive cross-correlation function, IEEE Trans. Inf. Theory 14 (1968) 154–156.
- [42] T. Helleseth, L. Hu, A. Kholosha, X. Zeng, N. Li, W. Jiang, Period-different m-sequences with at most four-valued cross correlation, IEEE Trans. Inf. Theory 55(7) (2009) 3305–3311.
- [43] T. Helleseth, P. Rosenberg, New pairs of m-sequences with 4-level cross-correlation, Finite Fields Appl. 11(4) (2005) 674–683.
- [44] H. D. L. Hollmann, Q. Xiang, A proof of the Welch and Niho conjectures on cross-correlations of binary msequences, Finite Fields Appl. 7 (2001) 253–286.
- [45] X. D. Hou, A note on the proof of Niho's conjecture, SIAM J. Discrete Math. 18(2) (2004) 313-319.
- [46] T. Kasami, The weight enumerators for several classes of subcodes of the 2nd order binary RM codes, Information and Control 18 (1971) 369–394.
- [47] A. M. Kerdock, A class of low-rate nonlinear codes, Information and Computation 20 (1972) 182–187.
- [48] C. Li, Q. Yue, The Walsh transform of a class of monomial functions and cyclic codes, Cryptogr. Commun. 7(2) (2015) 217–228.
- [49] C. Li, Q. Yue, F. Li, Weight distributions of cyclic codes with respect to pairwise coprime order elements, Finite Fields Appl. 28 (2014) 94–114.
- [50] C. Li, Q. Yue, F. Li, Hamming weights of the duals of cyclic codes with two zeros, *IEEE Trans. Inf. Theory*, vol. 60, no. 7, pp. 3895–3902, 2014.
- [51] R. Lidl, H. Niederreiter, Finite Fields, Cambridge University Press, Cambridge, 1997.
- [52] F. J. MacWilliams, N. J. A. Sloane, The Theory of Error-Correcting Codes, North-Holland Mathematical Library, North-Holland, Amsterdam, 1977.
- [53] A. Maschietti, Difference sets and hyperovals, Des. Codes Cryptogr. 14 (1998) 89–98.
- [54] S. Mesnager, Semibent functions from Dillon and Niho exponents, Kloosterman sums, and Dickson polynomials, IEEE Trans. Inf. Theory 57(11) (2011) 7443–7458.
- [55] S. Mesnager, Semi-bent functions from oval polynomials, In: M. Stam, Ed., Coding and Cryptography, LNCS 8308, Springer-Verlag, Heidelberg, 2013, pp. 1–15.

- [56] S. Mesnager, Bent vectorial functions and linear codes from o-polynomials, Des. Codes Cryptogr., DOI 10.1007/s10623-014-9989-6.
- [57] S. Mesnager, On semi-bent functions and related plateaued functions over the Galois field F_{2ⁿ}, in: Proceedings of Open Problems in Mathematics and Computational Science, LNCS, Springer, pp. 243–273, 2014.
- [58] S. Mesnager, Several new infinite families of bent functions and their duals, IEEE Trans. Inf. Theory 60(7) (2014) 4397–4407.
- [59] S. Mesnager, Bent functions from spreads, J. of the American Mathematical Society, in: Proceedings the 11th International conference on Finite Fields and their Applications Fq11, Volume 632, to appear.
- [60] D. E. Muller, Application of boolean algebra to switching circuit design and to error detection, IRE Transactions on Electronic Computers 3 (1954) 6–12.
- [61] Y. Niho, Multi-Valued Cross-Correlation Functions Between Two Maximal Linear Recursive Sequences, Ph.D. thesis, University of Southern California, Los Angeles , 1972.
- [62] S. E. Payne, A new infinite family of generalized quadrangles, Congressus Numerantium 49 (1985) 115–128.
- [63] L. Qu, Y. Tan, C. Li, On the Walsh spectrum of a family of quadratic APN functions with five terms, Science China Information Sciences 57(2) (2014) 1–7.
- [64] I. S. Reed, A class of multiple-error-correcting codes and the decoding scheme, IEEE Trans. Inf. Theory 4 (1954) 38–49.
- [65] O. S. Rothaus, On bent functions, J. Comb. Theory Ser. A 20 (1976) 300–305.
- [66] B. Segre, Sui k-archi nei piani finiti di caratteristica 2, Revue de Math. Pures Appl. 2 (1957) 289-300.
- [67] B. Segre, Ovali e curvenei piani di Galois di caratteristica due, Atti Accad. Naz. Lincei Rend. (8) 32 (1962) 785–790
- [68] B. Segre, U. Bartocci, Ovali ed alte curve nei piani di Galois di caratteristica due, Acta Arith., 18 (1971) 423-449.
- [69] Q. Wang, K. Ding, and R. Xue, Binary linear codes with two weights, IEEE Communications Letters, DOI 10.1109/LCOMM.2015.2431253.
- [70] Y. Xia, X. Zeng, L. Hu, Further crosscorrelation properties of sequences with the decimation factor $d = (p^n + 1)/(p+1) + (p^n 1)/2$, Appl. Algebra Eng. Commun. Comput. 21 (2010) 329–342.
- [71] J. Yuan, C. Ding, Secret sharing schemes from three classes of linear codes, IEEE Trans. Inf. Theory 52(1) (2006) 206–212.
- [72] Z. Zhou, C. Ding, A class of three-weight cyclic codes, Finite Fields Appl. 25 (2014) 79–93.