# Extended Binary Linear Codes from Legendre Sequences

T. Aaron Gulliver and Matthew G. Parker\*

#### Abstract

A construction based on Legendre sequences is presented for a doubly-extended binary linear code of length 2p+2 and dimension p+1. This code has a double circulant structure. For p=4k+3, we obtain a doubly-even self-dual code. Another construction is given for a class of triply extended rate 1/3 codes of length 3p+3 and dimension p+1. For p=4k+1, these codes are doubly-even self-orthogonal.

# 1 Introduction

A binary [n, K] code C is a K-dimensional vector subspace of  $\mathbb{F}_2^n$ , where  $\mathbb{F}_2$  is the field of two elements. The parameter n is called the length of C. The elements of a code C are called *codewords* and the *weight* of a codeword is the number of non-zero coordinates. Denote the weight of a codeword c as wt(c). The *minimum weight* of C is the smallest weight among all non-zero codewords of C. An [n, K, d] code is an [n, K] code with minimum weight d. Two codes are *equivalent* if one can be obtained from the other by a permutation of coordinates.

<sup>\*</sup>T.A. Gulliver is with the Dept. of Electrical and Computer Engineering, University of Victoria, P.O. Box 3055 STN CSC, Victoria, BC V8W 3P6 Canada. agullive@ece.uvic.ca. Web: http://www.ece.uvic.ca/-agullive/ M.G. Parker is with the Selmer Centre, Inst. for Informatikk, Høyteknologisenteret i Bergen, University of Bergen, Bergen 5020, Norway. E-mail: matthew@ii.uib.no. Web: http://www.ii.uib.no/-matthew/ This work was done while the first author was visiting the Selmer Centre.

The dual code  $C^{\perp}$  of C is defined as  $C^{\perp} = \{x \in \mathbb{F}_2^n | (x, y) = 0 \text{ for all } y \in C\}$  where (x, y) denotes the inner product. A code C is called *self-dual* if  $C = C^{\perp}$ . A self-dual code C is called *doubly-even* or *singly-even* if all codewords have weight  $\equiv 0 \pmod{4}$  or if some codeword has weight  $\equiv 2 \pmod{4}$ , respectively.

Let  $D_p$  and  $D_b$  be codes with generator matrices of the form

$$I_n R$$
 (1)

and

$$I_{n+1}$$
  $0 \quad 1 \quad \cdots \quad 1$   $1 \quad \vdots \quad R'$   $(2)$ 

respectively, where I is the identity matrix of order n and R and R' are  $n \times n$  circulant matrices. The codes  $D_p$  and  $D_b$  are called *pure double circulant* and bordered double circulant, respectively. The two families are collectively called double circulant codes. Many of the known self-dual codes are double circulant [2, 3, 5, 6, 9].

It was shown in [14] that the minimum weight d of a doubly-even self-dual code of length n is bounded by  $d \leq 4[n/24]+4$ . We call a doubly-even self-dual code meeting this upper bound extremal. The largest possible minimum weights of doubly-even self-dual codes of lengths up to 72 are given in [2, Table I]. This work was revised and extended to lengths up to 96 in [3, Table V]. We say that a doubly-even self-dual code with the largest possible minimum weight given in [2, Table I], [3, Table V] is extremal. Many extremal self-dual codes are double circulant [2, 3, 5, 6, 7, 9].

In this paper we employ a Legendre sequence [16] of length p, p an odd prime, to build a circulant matrix which is then used to construct a bordered double circulant code of length n=2p+2 and dimension K=p+1. We show that these codes have good distance, in particular when 2 is a quadratic nonresidue, mod p. For p=4k+3, we show that these codes are self-dual. Another construction based on these sequences is used to obtain a class of triply extended rate 1/3 codes of length 3p+3 and dimension p+1. For p=4k+1, these codes are doubly-even self-orthogonal.

### 2 The Construction

### 2.1 Legendre Sequences

Let a be a primitive integer root, mod p, where p is an odd prime. Let  $\mathcal{A} = \{a^{2i}\}$  be the set of even powers of a, mod p, and  $\mathcal{B} = \{a^{2i+1}\}$  be the set of odd powers of a, mod p.

**Definition 1.** The binary Legendre sequence, s, of length p (see e.g. [1, 10]), satisfies

$$s = (s_0, s_1, \dots, s_{p-1}) \mid s_0 = 0, s_t = 1 \text{ if } t \in A, s_t = 0 \text{ if } t \in B.$$

We have chosen in this case to assign  $s_0 = 0$ , but we retain the possibility to assign 0 or 1 to  $s_0$ .

**Definition 2.** The alternative Legendre sequence  $\tilde{s}$ , has  $\tilde{s}_0 = 1$ , and  $\tilde{s}_t = s_t$  if  $t \neq 0$ .

Define  $u = (u_0, u_1, \dots, u_{p-1})$  as the cyclic autocorrelation of s with

$$u_j = \sum_{t=0}^{p-1} (-1)^{s_t - s_{t+j}},$$

where the index of s is taken mod p. Similarly, define  $\tilde{\mathbf{u}}$  as the cyclic autocorrelation of  $\tilde{\mathbf{s}}$ . The following properties of s and  $\tilde{\mathbf{s}}$  are well-known

# Lemma 1. /16/

$$u_0 = \tilde{u}_0 = p,$$
  
 $u_j, \tilde{u}_j = -1,$   $j \neq 0, p = 4k + 3,$   
 $u_j, \tilde{u}_j \in \{1, -3\},$   $j \neq 0, p = 4k + 1,$   
 $u_j + \tilde{u}_j = -2,$   $j \neq 0.$ 

In the sequel we make particular use of the property that  $u_j + \tilde{u}_j = -2$  when  $j \neq 0$  or p to construct, for all odd primes p, a double circulant code of length 2p. We illustrate the code construction by means of an example.

### 2.2 Example

Consider the length p=5 Legendre sequence  $\mathbf{s}=01001$ , where  $s_t=1$  for  $t\in\mathcal{A}=\{1,4\}$  and  $s_t=0$  for  $t\in\mathcal{B}=\{2,3\}$ . The alternative Legendre sequence is  $\tilde{\mathbf{s}}=11001$ . It follows that  $\mathbf{u}=5,-3,1,1,-3$  and  $\tilde{\mathbf{u}}=5,1,-3,-3,1$ , and therefore  $\mathbf{u}+\tilde{\mathbf{u}}=10,-2,-2,-2,-2$ . This suggests that appropriate bordering of the concatenation of the circulant matrices formed by  $\mathbf{s}$  and  $\tilde{\mathbf{s}}$  by two additional columns could give a matrix with orthogonal rows, and this proves to be the case for p=4k+3.

For the example above, concatenating the circulant matrices formed from the Legendre and alternative Legendre sequences gives

 $\mathbf{D'} = \begin{array}{c} 01001|11001\\ 10100|11100\\ \mathbf{D'} = 01010|01110\\ 00101|00111\\ 10010|10011 \end{array}$ 

This is a double circulant generator matrix for a [10,5,3] binary linear code (D' always generates a cyclic code). The above matrix can be bordered by the all-ones and all-zeroes columns, and then the all-ones row resulting in

<b>D</b> =	11 11111 11111
	10 01001 11001
	10 10100 11100
	10 01010 01110
	10 00101 00111
	10 10010 10011

D can be transformed into a bordered double circulant generator matrix for a [12, 6, 4] optimal binary linear code, as will be shown later.

We generalise this construction to any length p Legendre sequence in the next section.

# 2.3 The Doubly-Extended Legendre Code Construction

Let  $q = s|\tilde{s}$ .

#### Lemma 2.

$$wt(\mathbf{q}) = p.$$

*Proof.* From the definition of s, wt(s) = (p-1)/2 and therefore wt( $\tilde{s}$ ) = (p-1)/2+1. Thus wt(q) = 2(p-1)/2+1=p.

Define  $\rho = (\rho_0, \rho_1, \dots, \rho_{2p-1})$  as the cyclic autocorrelation of q, where

$$\rho_j = \sum_{t=0}^{2p-1} (-1)^{q_t - q_{t+j}},$$

and the index of q is taken mod 2p.

#### Lemma 3.

$$\rho_j = -2, \qquad 0 < j < 2p, \ j \neq p.$$

*Proof.* Follows immediately from Lemma 1 as  $\rho_j = u_j + \tilde{u}_j$ .

Define  $\mathbf{w} = (w_0, w_1, \dots, w_{p-1})$  as the  $\{0, 1\}$ -cyclic autocorrelation of  $\mathbf{q}$ , where

$$w_j = \sum_{t=0}^{2p-1} q_t q_{t+j},$$

and the index of q is taken mod 2p. Note that this is a shortened version of the complete autocorrelation as we are only concerned with the first p elements.

#### Theorem 1.

$$w_j = 2k + 1,$$
  $p = 4k + 3,$   $0 < j < p,$   
=  $2k,$   $p = 4k + 1,$   $0 < j < p.$ 

*Proof.* We can alternatively define  $w_j$  by  $w_j = |\{t|q_t = q_{t+j} = 1, 0 \le t < 2p\}|$ . Define the set  $\mathbf{A} = \{t|q_t \ne q_{t+j}, 0 \le t < 2p\}$ .

Consider the set of bit pairs  $\{(q_t, q_{t+j})\}$ ,  $0 \le t < 2p$ . We have that  $w_j = |\{t|(q_t, q_{t+j}) = (1, 1)\}|$ , and  $\operatorname{wt}(q) = |\{t|(q_t, q_{t+j}) = (1, 0)\}| = |\{t|(q_t, q_{t+j}) = (0, 1)\}|$ . It follows that  $2 \times \operatorname{wt}(q) = |\{t|(q_t, q_{t+j}) = (1, 0)\}| + |\{t|(q_t, q_{t+j}) = (1, 0)\}|$ 

 $|\{t|(q_t,q_{t+j})=(0,1)\}|=|\{t|(q_t,q_{t+j})=(1,0) \text{ or } (0,1)\}|=|\mathbf{A}|.$  Therefore it follows that

$$\operatorname{wt}(\mathbf{q}) = |\{t|q_t = 1\}| = w_j + \frac{|\mathbf{A}|}{2}.$$
 (3)

Lemma 3 implies that  $|\mathbf{A}| = p + 1$  which, together with Lemma 2 and (3), gives  $w_j = \frac{p-1}{2}$ , and the theorem follows.

Let  $d_i$  be the *i*th row of D'. An immediate corollary of Theorem 1 is Corollary 1.

$$wt(\mathbf{d}_i + \mathbf{d}_j) = p + 1.$$

Let s be a length p Legendre sequence, where p is a prime integer, and  $\tilde{\mathbf{S}}$  and  $\tilde{\mathbf{S}}$  be the  $p \times p$  circulant matrices with s and  $\tilde{\mathbf{s}}$  as their first rows, respectively. Then

$$\mathbf{D}' = \mathbf{S}|\mathbf{\tilde{S}}$$

is a length 2p double circulant binary linear code of dimension p. Let 1 be the  $p \times 1$  all-ones vector and 0 be the  $p \times 1$  all-zeroes vector. Then

$$\mathbf{D} = \begin{array}{cc} \mathbf{1}\mathbf{1}\mathbf{1}^T & |\mathbf{1}^T| \\ \mathbf{1}\mathbf{0}\mathbf{S} & |\tilde{\mathbf{S}}| \end{array}$$

is a length 2p+2 bordered double circulant binary linear code of dimension p+2.

**Theorem 2.** The code with generator matrix **D** for p = 4k + 3 is a doubly-even self-dual code.

Proof. Since 4|2p+2 when p is an odd prime, the first row of  $\mathbf{D}$  has weight a multiple of 4. The rows of  $\mathbf{S}$  have weight (p-1)/2 and the rows of  $\tilde{\mathbf{S}}$  have weight (p+1)/2. Adding these together gives 2p/2=p. The all-ones column adds weight 1 to each row, so all rows of  $\mathbf{D}$  have weight p+1. From Corollary 1, the weight of the sum of any two rows of  $\mathbf{D}'$  is even, and this also holds for the rows of  $\mathbf{D}$ , so the rows are orthogonal. When p=4k+3, p+1=4k+4 so the weight of all rows is divisible by 4. Therefore from [12], the code is doubly-even self-dual.

It is obvious that the minimum distance of the code generated by **D** is upperbounded by p+1.

#### 2.4 Reduced Echelon Form

It is often desirable to have a code in systematic or reduced echelon form

I|P

where I is the  $p \times p$  identity matrix. The double circulant form of our construction should then be converted to the form (1). To achieve this, it is necessary that S or  $\tilde{S}$  be invertible. This in turn implies that s or  $\tilde{s}$ , when viewed as polynomials, s(x) or  $\tilde{s}(x)$ , should be invertible, mod  $x^p + 1$ , mod 2. It turns out that, for  $p = 8k \pm 1$ , s(x) and  $\tilde{s}(x)$  are never invertible, for p = 8k + 3 s(x) is always invertible, and for p = 8k - 3  $\tilde{s}(x)$  is always invertible. These conditions reflect the fact that 2 is a quadratic residue for  $p = 8k \pm 1$  and a quadratic nonresidue for  $p = 8k \pm 3$ . Therefore a row echelon form for the doubly-extended Legendre code, D, with the identity in the first p+2 or last p+2 columns, can only be achieved when  $p = 8k \pm 3$ , i.e. neither columns 0 to p+1, or columns p+2 to 2p+3 are information sets). Let  $\overline{s(x)}$  denote that every coefficient of s(x) is negated. Then, when  $p = 8k \pm 3$ , it can be shown that

$$\tilde{s}(x)^{-1} = \tilde{s}(x)^2 = \overline{\tilde{s}(x)} \mod x^p + 1, \mod 2, \quad p = 8k - 3$$
 $s(x)^{-1} = s(x)^2 = \overline{\tilde{s}(x)} \mod x^p + 1, \mod 2, \quad p = 8k + 3$ 
 $\tilde{s}(x)^{-1}s(x) = \overline{\tilde{s}(x)} \mod x^p + 1, \mod 2, \quad p = 8k - 3$ 
 $s(x)^{-1}\tilde{s}(x) = \overline{s(x)} \mod x^p + 1, \mod 2, \quad p = 8k + 3.$ 

Therefore, when 2 is a quadratic nonresidue, mod p, we obtain a  $p \times p$  circulant matrix,  $\mathbf{P}$ , whose first row is the negation of  $\tilde{\mathbf{s}}$  for p = 8k - 3, and the negation of  $\mathbf{s}$  for p = 8k + 3. In this case, we obtain a double circulant code having the first row of the circulant matrix as defined above. When p = 8k + 3, the codes (bordered or pure) are equivalent to those given in [15, 13, 8, 11].

### 2.4.1 Example

For p=5,  $\tilde{s}=11001$  and  $\tilde{s}(x)=x^4+x+1$  has multiplicative order 3 mod  $x^5+1 \pmod{2}$ . Moreover  $\tilde{s}(x)^{-1}=x^3+x^2+1$ . Thus

$$\tilde{\mathbf{S}} = \begin{array}{cccc} 11001 & & 10110 \\ 11100 & & 01011 \\ \tilde{\mathbf{S}} = & 01110 & & \text{and } \tilde{\mathbf{S}}^{-1} = & 10101 \\ 00111 & & & 11010 \\ 10011 & & & 01101 \end{array}$$

since  $\tilde{\mathbf{S}}^{-1}\tilde{\mathbf{S}} = \mathbf{I}$ . Thus

$$\tilde{\mathbf{S}}^{-1}\mathbf{D}'=\mathbf{P}|\mathbf{I}$$

where

$$\mathbf{P} = \begin{array}{c} 00110 \\ 00011 \\ 10001 \\ 11000 \\ 01100 \end{array}$$

since

$$\tilde{s}(x)^{-1}s(x) = (x^3 + x^2 + 1)(x^4 + x) \mod x^5 + 1 = x^3 + x^2$$

The generator matrix then has the form

$$\mathbf{G} = \begin{array}{c} 100000|011111\\ 010000|100011\\ 001000|110001\\ 000100|111000\\ 000010|101100\\ 000001|100110 \end{array}$$

This is a bordered double circulant generator matrix for a [12, 6, 4] binary linear code.

# 3 The Double Circulant Codes

The most well-known case is p=11 as the [24, 12, 8] Golay code is obtained. Note that p=7 is the first case where both S and  $\tilde{S}$  are singular, but in this case we obtain an extremal code. Table 1 shows the Hamming distances for the first 40 codes ( $n \le 180$ ) constructed from **D**. The extremal codes are denoted by a '\*'. For large n, it was not possible to find the minimum distance, so in these cases bounds are given. Of particular interest is when  $p = 8k \pm 1$ , since in these cases it is not possible to obtain a bordered double circulant code. Such primes are marked in table 1 with a '#'.

Table 1: Hamming Distances for the Doubly-Extended Double Circulant Codes

p	d	p	d	p	d	p	d	p	d
3	4*	29	12	61	20	101	20 - 30	139	20 – 44
5	4	31#	8	67	24*	103#	20	149	18 – 50
7#	4*	37	12	71#	12	107	20 – 36	151#	20
11	8*	41#	10	73#	14	109	20 – 36	157	16 – 52
13	8	43	16*	79#	16	113#	16	163	16 – 56
17#	6	47#	12	83	24	127#	20	167#	16 – 24
19	8*	53	20	89#	18	131	20 – 44	173	16 - 62
23#	8	59	20	97#	16	137#	18 - 22	179	16 – 60

From Table 1 one observes that, in general, the codes for  $p=8k\pm 1$  have lower minimum Hamming distance than those for  $p=8k\pm 3$ . A lower bound on the minimum Hamming distance of the unextended form of the codes (given by  $\mathbf{D}'$ ), when  $p=8k\pm 3$ , can be obtained from the lower bound on Hamming distance for double circulant codes [11]

$$d \geq \frac{2(p+\sqrt{p})}{\sqrt{p}+3}.$$

The corresponding bound for **D** (when  $p = 8k \pm 3$ ) is

$$d \geq \frac{2p+3\sqrt{p}+3}{\sqrt{p}+3}.$$

However, the bounding technique of [11] cannot easily adapted to the case when  $p = 8k \pm 1$ . This is because in this case  $n(x)^i = n(x)$  and  $q(x)^i = q(x)$ 

for all i where q(x) are the quadratic residues and n(x) are the quadratic nonresidues.

# 4 A Construction for Rate 1/3 Codes

Now consider the [3p, p, d] codes with generator matrices

$$\mathbf{E}' = \mathbf{I}|\mathbf{D}' = \mathbf{I}|\mathbf{S}|\tilde{\mathbf{S}}.$$

These can be extended to [3p+3,p+1,d] codes with generator matrices

$$\mathbf{E} = \begin{array}{cc|c} \mathbf{0} & \mathbf{0}^T \\ \mathbf{1} & \mathbf{I} \end{array} \mathbf{D} = \begin{array}{cc|c} \mathbf{0} & \mathbf{0}^T & \mathbf{1} & \mathbf{1} & \mathbf{1}^T & \mathbf{1}^T \\ \mathbf{1} & \mathbf{I} & \mathbf{1} & \mathbf{S} & \mathbf{\tilde{S}} \end{array}$$

**Theorem 3.** The code with generator matrix  $\mathbf{E}$  for p = 4k+1 is a doubly-even self-orthogonal code.

**Proof.** Since 4|2p+2 when p is an odd prime, the first row of E has weight a multiple of 4. The rows of S have weight (p+1)/2 and the rows of S have weight (p-1)/2. Adding these together gives 2p/2=p. The remaining columns in E add 3 to the weight of each of these rows, so they have weight p+3. From Corollary 1, the weight of the sum of any two rows of D' is even, so the rows are orthogonal. The inner product of the first row of D' with any other row is 1, therefore the first column makes the first row of E orthogonal to the others. When p=4k+1, p+3=4k+4 so the weight of all rows is divisible by 4. Therefore from [12], the code is doubly-even self-orthogonal.

Deleting the first row and 3 columns in E we obtain the following.

**Corollary 2.** The code with generator matrix  $\mathbf{E}'$  for p=4k+1 is a singly-even self-orthogonal code.

### 4.1 Example

Consider as before the length p=5 Legendre sequence. The circulant matrices formed from the Legendre and alternative Legendre sequences

give

 $\mathbf{E'} = \begin{array}{c} 10000|01001|11001 \\ 01000|10100|11100 \\ \mathbf{E'} = 00100|01010|01110 \\ 00010|00101|00111 \\ 00001|10010|10011 \end{array}$ 

This is the generator matrix for a [15, 5, 6] self-orthogonal quasi-cyclic code. This leads to the following extended code

 $\mathbf{E} = \frac{\frac{110|00000|11111|11111}{101|10000|01001|11001}}{\frac{101|01000|10100|11100}{101|00100|01010|01110}} \cdot \frac{110|00010|00101|00111}{101|00001|10010|100111}$ 

**E** is a bordered generator matrix for an [18, 6, 8] optimal self-orthogonal binary linear code.

Table 2 gives the distances of the first few codes generated from  $\mathbf{E}'$ , and Table 3 gives distances and bounds for  $\mathbf{E}$  up to p=151. Note that the code from  $\mathbf{E}'$  has distance 2 less than the corresponding code from  $\mathbf{E}$ . Several of these codes attain the lower bound on the maximum minimum distance for a binary linear code [4]. For large n, it was not possible to find the minimum distance, so in these cases bounds are given.

Table 2: Hamming Distances for Codes Generated Using  $\mathbf{E}'$ 

p	d	p	d
5	6	17	10
7	6	19	14
11	10	29	22
13	10		

Table 3: Hamming Distances for Rate 1/3 Codes Generated by E

p	d	p	d	p	d	p	d	p	d
3	6	29	24	61	20	101	32 - 56	139	24 - 86
5	8	31	16	67	36	103	30 – 40	149	24 – 88
7	8	37	24	71	24	107	30 – 66	151	24 – 40
11	12	41	20	73	28	109	32 – 64		
13	12	43	28	79	32	113	28 - 32	1	
17	12	47	24	83	34 – 48	127	26 – 40		
19	16	53	32	89	32 – 36	131	28 – 78		
23	16	59	30	97	32	137	28 – 44		

# References

- [1] T. Beth, D. Jungnickel, and H. Lenz, *Design Theory*, Cambridge University Press, Cambridge, 2nd edition, 1999.
- [2] J.H. Conway and N.J.A. Sloane, A new upper bound on the minimal distance of self-dual codes, *IEEE Trans. Inform. Theory* 36 (1990), 1319–1333.
- [3] S.T. Dougherty, T.A. Gulliver and M. Harada, Extremal binary selfdual codes, *IEEE Trans. Inform. Theory* 43 (1997), 2036–2047.
- [4] M. Grassl, Bounds on the minimum distance of linear codes, available at http://www.codetables.de/.
- [5] T.A. Gulliver and M. Harada, Weight enumerators of double circulant codes and new extremal self-dual codes, Des. Codes and Cryptogr. 11 (1997), 141–150.
- [6] T.A. Gulliver and M. Harada, Classification of extremal double circulant self-dual codes of lengths 64 to 72, Des. Codes and Cryptogr. 13 (1998), 257–269.

- [7] T.A. Gulliver, M. Harada and J.-L. Kim, Construction of new extremal self-dual codes, *Discrete Math.* 263 (2003), 81–91.
- [8] T.A. Gulliver and N. Senkevitch, On a class of self-dual codes derived from quadratic residues, *IEEE Trans. Inform. Theory* 45 (1999), 701– 702.
- [9] M. Harada, T.A. Gulliver and H. Kaneta, Classification of extremal double circulant self-dual codes of length up to 62, *Discrete Math.* 188 (1998), 127–136.
- [10] T. Helleseth, Legendre sums and codes related to QR codes, Discr. Appl. Math. 35 (1992), 107-113.
- [11] T. Helleseth and J. F. Voloch, Double circulant quadratic residue codes, *IEEE Trans. Inform. Theory* 50 (2004), 2154-2155.
- [12] W.C. Huffman and V. Pless, Fundamentals of Error-Correcting Codes, Cambridge University Press, 2003.
- [13] M. Karlin, New binary coding results by circulants, IEEE Trans. Inform. Theory 15 (1969), 81-92.
- [14] C.L. Mallows and N.J.A. Sloane, An upper bound for self-dual codes, Inform. Control 22 (1973), 188–200, 2001
- [15] E.H. Moore, Double Circulant Codes and Related Algebraic Structures, Ph.D. dissertation, Dartmouth College, 1976.
- [16] M.R. Schroeder, Number Theory in Science and Communication, Springer-Verlag, 1993.