A Class of Optimal Cyclic Symbol-pair Codes from Combinatorial Designs

Jie Yan¹, Xiaofeng Xu² and Chengmin Wang²

¹School of Science, Jiangnan University, Wuxi, China
²Department of Mathematics, Taizhou University, Taizhou, China

Abstract. Motivated by the application of high-density data storage technologies, symbol-pair codes were defined by Cassuto and Blaum recently. In this coding model, every two consecutive symbols are read together to address channels with high write resolution but low read resolution, so that individual symbols cannot be read off due to physical limitations. Since such kind of code was introduced, a lot of work has been carried out by many researchers, including upper bounds, constructions, decoding et.al. In this work, we continue the research on the construction of symbol-pair codes. We will present one special construction for cyclic symbol-pair codes from the perspective of combinatorial designs.

1. Introduction
Symbol-pair read channels were recently studied by Cassuto and Blaum [1], in which the outputs of the read process are pairs of consecutive symbols instead of individual symbols. This kind of channels has been set up and studied deeply because of their application of high-density data storage technologies [1,2]. In this paradigm, a relevant pair-distance metric was set up to establish necessary and sufficient conditions for the pair error model. Also the relation between Hamming distance and pair-distance has also been studied. Some constructions of symbol-pair codes are presented and the decoding method is discussed.

Soon later, Chee et.al. [4] established the Singleton type upper bound of symbol-pair code and its corresponding optimal codes are called maximum distance separable (MDS) symbol-pair codes. Since then, MDS symbol-pair codes have attracted a lot of researchers’ interest. A lot of optimal such codes are constructed via different kind of techniques from algebraic, geometry, graph and so on [4-6,12,13,15]. One generalization of the symbol-pair model was studied in [19] for the b-symbol read channel where the assumption is that every b > 2 consecutive symbols are read together. This model was further deeply studied for MDS symbol-pair codes in [5], [13].

Decoding is an important content in the research of coding theory and the study of symbol-pair code is no exception. Several decoding techniques on symbol-pair codes have been studied deeply by some researchers [9,10,14,16-18]. For example, syndrome decoding is developed in [9], and linear programming decoding is studied in [10]. Algebraic decoding for BCH type symbol-pair codes and decoding for cyclic symbol-pair codes are developed in [17] and [16,18] respectively.

As another core problem of coding theory, the research on upper bounds of symbol-pair codes has also been developed in several aspects. For example, the Gilbert-Varshamov type bound is established in [1] and the Singleton type bound is established in [4]. Recently in [7,8], two other special upper bounds are established, one is called John type bound and the other is called linear programming type bound. With these two bounds, some new optimal symbol-pair codes are also presented. Some more work on bounds has been focused in [19, 20].
We know that a lot of work has been carried out from different aspects. In this paper, we continue the research on the construction of symbol-pair codes. We will present one special construction for cyclic symbol-pair codes from combinatorial design. It is known that combinatorial design is an ancient subject with rich content and special skills. As an important means, it has been widely used in coding. By using a kind of combinatorial designs called Mendelsohn designs, we obtain a class of optimal cyclic symbol-pair codes.

This paper is organized as follows. In Section 2, we introduce some basic notations and definitions for later use. In Section 3, we will present a kind of combinatorial design, called Mendelsohn designs. In Section 4, a class of optimal cyclic symbol-pair codes constructed from Mendelsohn designs is displayed. Section 5 finally comes to a conclusion.

2. Preliminaries

In this section, some basic notations of symbol-pair codes are introduced.

For a positive integer \( n, \) \( \mathbb{Z}_n \) denote the ring \( \mathbb{Z}/n\mathbb{Z} \). Let \( \Sigma \) be a set of symbol alphabet, each element in \( \Sigma \) is called a symbol. Let \( \Sigma^n \) be the set of all \( n \)-length sequences over \( \Sigma \). The coordinates of \( x \in \Sigma^n \) are indexed by elements of \( \mathbb{Z}_n \), so that \( x \) is written as \((x_0, x_1, \ldots, x_{n-1})\). A pair-vector over \( \Sigma \) is a vector in \( (\Sigma \times \Sigma)^n \). For any \( x = (x_0, x_1, \ldots, x_{n-1}) \in \Sigma^n \), the symbol-pair read vector of \( x \) is the pair-vector (over \( \Sigma \))

\[
\pi(x) = ((x_0, x_1), (x_1, x_2), \ldots, (x_{n-2}, x_{n-1}), (x_{n-1}, x_0)).
\]

Obviously, each vector \( x \in \Sigma^n \) has a unique symbol-pair read vector \( \pi(x) \in (\Sigma \times \Sigma)^n \). But the opposite is not true, as any element of \( (\Sigma \times \Sigma)^n \) is not necessarily a pair-vector.

For any two vectors \( x, y \in \Sigma^n \), denote by \( d_H(x, y) \) the Hamming distance between \( x \) and \( y \). The pair-distance between \( x \) and \( y \) is defined as

\[
d_p(x, y) = d_H(\pi(x), \pi(y)).
\]

**Example 1** Let \( x = (0, 1, 1, 0, 1, 0) \), \( y = (0, 1, 1, 1, 1, 0) \). Then we have the following fact.

1. \( d_p(x, y) = 3 \);
2. \( \pi(x) = ((0, 1), (1, 1), (1, 0), (0, 1), (1, 0), (0, 0)) \);
3. \( \pi(y) = ((0, 1), (1, 0), (1, 1), (1, 1), (1, 0), (0, 0)) \);
4. \( d_p(x, y) = 5 \).

For a code \( C \), we denote its minimum Hamming distance by \( d_H(C) \). An \((q \text{-ary}) \) code of length \( n \) is a nonempty set \( C \subseteq \Sigma^n \). Define the minimum pair-distance of \( C \) as

\[
d_p(C) = \min\{d_p(x, y) | x, y \in C, x \neq y \}.
\]

A code \( C \) of length \( n \) over \( \Sigma \) is called an \((n, M, d_p)\), symbol-pair code if its size is \( M \), the minimum pair-distance is \( d_p \), and \(|\Sigma| = q \). A code \( C \) of length \( n \) is cyclic provided that for each codeword \( c = (c_1, c_2, c_3, \ldots, c_n) \) in \( C \), the vector \( e = (c_2, c_3, \ldots, c_n, c_1) \), obtained from \( c \) by the cyclic shift of coordinates \( i \mapsto i+1 \) (mod \( n \)), is also in \( C \).

**Example 2** Let \( C = \{c_1, c_2, c_3, c_4\} \) where

\[
c_1 = (0, 0, 0, 0);
\]
\[
c_2 = (0, 1, 1, 0);
\]
\[
c_3 = (1, 0, 1, 0);
\]
\[
c_4 = (0, 0, 1, 1).
\]

It is readily to check that the Hamming distance of \( C \) is \( d_H(C) = 2 \), while the minimum pair-distance of \( C \) is \( d_p(C) = 3 \). So \( C \) is a \((4, 4, 3)\)-symbol-pair code.
Let \( q, n, d_p \) be integers. Let \( A_q(n, d_p) \) be the maximal size of a \( q \)-ary symbol-pair code of length \( n \) with pair-distance \( d_p \). A \( q \)-ary \( (n, d_p) \) symbol-pair code with size \( A_q(n, d_p) \) is called optimal.

3. A Kind of Combinatorial Designs - Mendelsohn Designs

In this section, we introduce a kind of combinatorial designs, called Mendelsohn designs, which can be used to produce optimal cyclic symbol-pair codes.

The ordered \( k \)-tuple \( (x_1, x_2, \ldots, x_k) \) cyclically contains the ordered pairs \( (x_0, x_1), (x_1, x_2), \ldots, (x_k, x_0) \), and no others.

For example, an ordered 4-tuple \( (1, 2, 3, 4) \) cyclically contains the following four ordered pairs \( (1, 2), (2, 3), (3, 4), (4, 1) \).

A \((v,k,\lambda)\)-Mendelsohn design, denoted by \( \text{MD}(v,k,\lambda) \), is a set of \( V \) together with a collection \( \Psi \) of ordered \( k \)-tuples (blocks) of distinct elements from \( V \), such that each ordered pair \((x,y)\) with \( x \neq y \) is cyclically contained in exactly \( \lambda \) block.

**Lemma 1** A \((v,k,\lambda)\)-Mendelsohn design should contain \( \frac{\lambda v(v-1)}{k} \) blocks.

**Proof** Let \( V \) be the point set of a \((v,k,\lambda)\)-Mendelsohn design. Since the set \( V \) with \( v \) points should contain \( \lambda v(v-1) \) ordered \( k \)-tuples, and each ordered \( k \)-tuple costs \( k \) ordered pairs, then the conclusion is achieved.

**Example 3** By Lemma 1, a \((6,4,2)\)-Mendelsohn design should contain \( \frac{\lambda v(v-1)}{k} = 15 \) blocks. We display all the blocks as follows.

\[
(1, 2, 3, 4), (1, 2, 5, 3), (1, 6, 3, 2), (1, 4, 2, 6), (1, 6, 2, 5), \\
(1, 4, 5, 2), (1, 5, 6, 3), (1, 3, 6, 4), (1, 3, 4, 5), (1, 5, 4, 6), \\
(2, 3, 5, 4), (2, 4, 3, 6), (2, 6, 5, 3), (2, 4, 6, 5), (3, 5, 6, 4).
\]

Mendelsohn designs are one important object in combinatorial theory and it has been well studied for many years. In addition, Mendelsohn designs with some special properties are also studied, such as perfect Mendelsohn designs, resolvable Mendelsohn designs, and so on. For more details, the reader may refer to [4]. The following theorems provide some existence for Mendelsohn designs.

**Theorem 1** ([4])

1. Let \( p \) be an odd prime, then there exists an \( \text{MD}(p^2, p, 1) \);
2. If \( p = 1 \pmod{n} \), then there exists an \( \text{MD}(p^r, n, 1) \) for all \( r > 1 \).

**Theorem 2** ([4]) The necessary conditions are asymptotically sufficient for the existence of an \( \text{MD}(q, n, 1) \) with \( q = 0, 1 \pmod{n} \).

4. Optimal Cyclic Symbol-pair Codes from Mendelsohn Designs

To bench the optimality of symbol-pair codes, several upper bounds have been established, such as Singleton type bound, Johnson type bound, linear programming type bound, Plotkin type bound et al. Among of these upper bound, the Singleton type bound has been studied deeply. Symbol-pair codes achieved the Singleton type bound are called maximum distance separable (MDS).

**Theorem 3** [3] (Singleton Type Bound) Let \( q \geq 2 \) and \( 2 \leq d_p \leq n \). If \( C \) is an \( (n,d_p) \) symbol-pair code, then \( |C| \leq q^{n-d_p+2} \).

In this work, we use the Singleton type bound as our bench mark. In the following, Mendelsohn designs are going to be utilized to derive a kind of optimal cyclic symbol-pair codes.

**Theorem 4** If there exists an \( \text{MD}(q, n, 1) \), then there exists an optimal cyclic \( (n,d_p) \) symbol-pair code with \( d_p = n \) and \( A_q(n,n) = q^2 \).

**Proof** An \( \text{MD}(q, n, 1) \) contains \( q(q-1)/n \) blocks by Lemma 1. Let \( \Psi \) be the block set.
For each block $B = (x_0, x_1, \cdots, x_{n-1}) \in \Psi$, define $n$ permutations $\pi_i (i \in \mathbb{Z}_n)$ on the elements of $B$ as follows.

$$\pi_i (B) = (x_i, x_{i+1}, \cdots, x_{i+n-1}).$$

Here the subscripts of $x$ are taken modulo $n$. Moreover, let $c_i = (i, i, \cdots, i) (i \in \mathbb{Z}_n)$.

Define the following block set.

$$\Omega = \left( \bigcup_{i=0}^{n-1} \pi_i (B) \right) \bigcup \left( \bigcup_{j=0}^{n-1} c_i \right).$$

Clearly, the number of blocks in $\Omega$ is $q^2$. Take all the blocks of $\Omega$ as codewords to obtain a code $C$, then $|C| = q^2$. From above construction, we know that $C$ is cyclic.

In addition, by Theorem 3, $|C| \leq q^{n-d_p+2}$. Since $d_p = n$, then $|C| \leq q^{n-d_p+2} = q^2$. Hence, the symbol-pair codes constructed above are optimal, i.e.,

$$A_q (n, n) = q^2$$

**Example 4** A $(4,3,1)$-Mendelsohn design is given by the following four blocks

$$(0, 1, 2), (1, 0, 3), (2, 1, 3), (0, 2, 3).$$

Then we are able to obtain an optimal cyclic $(3,3)$ symbol-pair code with 16 codewords which are list below.

$$c_1 = (0, 1, 2) \quad c_2 = (1, 2, 0) \quad c_3 = (2, 0, 1)$$

$$c_4 = (1, 0, 3) \quad c_5 = (0, 3, 1) \quad c_6 = (3, 1, 0)$$

$$c_7 = (2, 1, 3) \quad c_8 = (1, 3, 2) \quad c_9 = (3, 2, 1)$$

$$c_{10} = (0, 2, 3) \quad c_{11} = (2, 3, 0) \quad c_{12} = (3, 0, 2)$$

$$c_{13} = (0, 0, 0) \quad c_{14} = (1, 1, 1)$$

$$c_{15} = (2, 2, 2) \quad c_{16} = (3, 3, 3)$$

Apply Theorem 4 to Theorem 1 and Theorem 2, we obtain the following optimal cyclic symbol-pair codes.

**Corollary 1**

1. Let $p$ be an odd prime, then there exists an optimal cyclic $(p, p)_{p^2}$ symbol-pair code;

2. If $p = 1 \pmod{n}$, then there exists an optimal cyclic $(n, n)_{q^2}$ symbol-pair code for any $r > 1$.

**Corollary 2** If $q = 0, 1 \pmod{n}$ and $q$ is large enough, then there exists an optimal cyclic $(n, n)_{q^2}$ symbol-pair code.

5. **Conclusion**

Combinatorial design is an ancient subject with rich content and special skills. As an important means, combinatorial design has been widely used in coding. Symbol-pair codes are a kind of code for the application of high-density data storage technologies. In this paper, we present one special construction for cyclic symbol-pair codes from the perspective of combinatorial designs. By using of Mendelsohn designs, we obtain a class of optimal cyclic symbol-pair codes.

6. **Acknowledgments**

The research is supported by NSFC (11471144) Natural Science Foundation of Jiangsu (Grant No. BK20171318), Qing Lan Project of Jiangsu and Overseas Study Program of Jiangsu University,
Corresponding to: Chengmin Wang.

7. References
[1] Cassuto Y and Blaum M 2011 Codes for symbol-pair read channels, *IEEE Transaction on Information Theory*, vol. 57, pp. 8011–8020
[2] Cassuto Y and Litsyn S 2011 Symbol-pair codes: Algebraic constructions and asymptotic bounds, in *Proceedings of the 2011 IEEE International Symposium on Information Theory* (ISIT2011), St. Petersburg, Russia, pp. 2348–2352
[3] Chee Y M, Ji L, Kiah H M, Wang C, and Yin J 2013 Maximum distance separable codes for symbol-pair read channels, *IEEE Transactions on Information Theory*, vol. 59, no. 11, pp. 7259–7267
[4] Colboun C J, Dinitz J 2007 Handbook of Combinatorial Designs (2nd ed), CRC, Baco Raton
[5] Ding B, Ge G, Zhang J, Zhang T, and Zhang Y 2017 New constructions of MDS symbol-pair codes, *Des. Codes Cryp.*, pp. 1–177
[6] Dinh H Q, Nguyen B T, Singh A K, and Sriboonchitta S 2018 On the symbol-pair distance of repeated-root constacyclic codes of prime power lengths, *IEEE Transactions on Information Theory*, vol. 64, no. 4, pp. 2417–2430
[7] Elishco O, Gabrys R, and Yaakobi E 2018 Bounds and constructions of Codes over Symbol-Pair Read Channels, in *Proceedings of the 2018 IEEE International Symposium on Information Theory* (ISIT2018), Vail, Colorado, USA, pp. 2505–2509
[8] Elishco O, Gabrys R, and Yaakobi E 2020 Bounds and constructions of Codes over Symbol-Pair Read Channels, *IEEE Transactions on Information Theory*, vol. 66, no. 3, pp. 1385–1395
[9] Hirotomo M, Takita M, and Morii M 2014 Syndrome decoding of symbol-pair codes, in *Proc. IEEE Inf. Theory Workshop*, Hobart, TAS, Australia, pp. 162–166
[10] Horii S, Matsushima T, and Hirasawam S 2016 Linear programming decoding of binary linear codes for symbol-pair read channel, *IEICE Trans. on Fund. of Elec., Comm. and Comp. Sci.*, vol. E99-A, no. 12, pp. 2170–2178
[11] Huffman W and Pless V 2003 Fundamentals of Error Correcting Codes. *Cambridge University Press*
[12] Kai X, Zhu S, and Li P 2015 A construction of new mds symbol-pair codes, *IEEE Trans. Inform. Theory*, vol. 61, pp. 5828–5834
[13] Li S and Ge G 2017 Constructions of maximum distance separable sym pair codes using cyclic and constacyclic codes, *Des. Codes Cryp.*, vol. 84, pp. 359–371
[14] Morii M, Hirotomo M, and Takita M 2016 Error-trapping decoding for cyclic codes over symbol-pair read channels, *Int. Symp. on Information Theory and Its App.*, pp. 681–685
[15] Sun Z, Zhu S, and Wang L 2017 The symbol-pair distance distribution of a class of repeated-root cyclic codes over fpm, *Cryp. and Comm.*, pp. 116–124
[16] Takita M, Hirotomo M, and Morii M 2015 A decoding algorithm for cyclic codes over symbol-pair read channels, *IEICE Trans. on Fund. of Elec., Comm. and Comp. Sci.*, vol. E98-A, no. 12, pp. 2415–2422
[17] Takita M, Hirotomo M, and Morii M 2016 Algebraic decoding of bch codes over symbol-pair read channels: Cases of two-pair and three-pair error correction, *IEICE Trans. on Fund. of Elec., Comm. and Comp. Sci.*, vol. E99-A, no. 12, pp. 2179–2191
[18] Takita M, Hirotomo M, and Morii M 2017 Error-trapping decoding for cyclic codes over symbol-pair read channels, *IEICE Trans. on Fund. of Elec., Comm. and Comp. Sci.*, vol. E100-A, no. 12, pp. 2578–2584
[19] Yaakobi E, Bruck J, and Siegel P 2016 Constructions and decoding of cyclic codes over b-symbol read channels, *IEEE Transactions on Information Theory*, vol. 62, no. 4, pp. 1541–1551
[20] Zhang H 2017 Improvement on minimum distance of symbol-pair codes, *Cryptography and Coding*, pp. 116-124