RC25218 (C1110-004) October 11, 2011
Computer Science

IBM Research Report

C-Codes: Cyclic Lowest-Density MDS Array Codes Constructed Using Starters for RAID 6

Mingqiang Li
IBM Research Division
China Research Laboratory
Building 19, Zhouguancun Software Park
8 Dongbeiwang West Road, Haidian District
Beijing, 100193
P.R.China

Jiwu Shu
Department of Computer Science and Technology
Tsinghua University
Qinghuayuan No. 1, Haidian District
Beijing 100084
P. R. China

C-Codes: Cyclic Lowest-Density MDS Array Codes Constructed Using Starters for RAID 6

Mingqiang Li and Jiwu Shu

***Abstract*—The distance-3 cyclic lowest-density MDS array code (called the C-Code) is a good candidate for RAID 6 because of its optimal storage efficiency, optimal update complexity, optimal length, and cyclic symmetry. In this paper, the underlying connections between C-Codes (or quasi-C-Codes) and starters in group theory are revealed. It is shown that each C-Code (or quasi-C-Code) of length $2n$ can be constructed using an even starter (or even multi-starter) in $(Z_{2n}, +)$. It is also shown that each C-Code (or quasi-C-Code) has a twin C-Code (or quasi-C-Code). Then, four infinite families (three of which are new) of C-Codes of length $p-1$ are constructed, where p is a prime. Besides the family of length $p-1$, C-Codes for some sporadic even lengths are also presented. Even so, there are still some even lengths (such as 8) for which C-Codes do not exist. To cover this limitation, two infinite families (one of which is new) of quasi-C-Codes of length $2(p-1)$ are constructed for these even lengths.**

***Index Terms*—RAID 6, array codes, starters, perfect one-factorization.**

I. Introduction

ARRAY CODES [2] are a class of linear codes whose information and parity bits are placed in a two-dimensional (or multidimensional) array rather than a one-dimensional vector. A common property of array codes is that they are implemented based on only simple eXclusive-OR (XOR) operations. This is an attractive advantage in contrast to the family of Reed-Solomon codes [3]–[5] whose encoding and decoding processes use complex finite-field operations. Thus, array codes are ubiquitous in data storage applications.

Among all kinds of array codes, cyclic lowest-density Maximum-Distance Separable (MDS) array codes [6] are regarded as the optimal ones for data storage applications because they have all the following properties:

1) they are MDS codes, which attain the Singleton bound [7] and thus have optimal *storage efficiency* (i.e. the ratio of user data to the total of user data plus redundancy data);
2) their *update complexity* (defined as the average number of parity bits affected by a change of a single information bit) achieves the minimum update complexity that MDS codes can have; and
3) their regularity in the form of cyclic symmetry makes their implementation simpler and potentially less costly.

Fault tolerance is an important concern in the design of disk-based storage systems [8]. As today's storage systems grow in size and complexity, they are increasingly confronted with disk failures [9], [10] together with latent sector errors [11]. Then, RAID 5 [12], which has been widely used in modern storage systems to recover one disk failure, cannot provide sufficient reliability guarantee. This results in the demand of RAID 6 [12], [13], which can tolerate two disk failures.

TABLE I
Comparison among some representative MDS array codes for RAID 6.

Array Code	Optimal Update Complexity	Optimal Length[a]	Cyclic Symmetry
EVENODD [15], RDP [16], Liberation [17]	No	—	No
X-Code [18]	Yes	No	No
B-Code[b] [14]	Yes	Yes	No
C-Code	Yes	Yes	Yes

[a]Considered only for lowest-density MDS codes (see [14]).
[b]Including ZZS Code [19] and P-Code [20].

RAID 6 is designed based on a distance-3 MDS linear code. In applications with many small writes, such as the On-Line Transaction Processing (OLTP) application, the distance-3 cyclic lowest-density MDS array code (called the C-Code) defined in Section II-A can be regarded as a good candidate for RAID 6 (see Table I). Under this background, we will make a systematic study on the C-Code in this paper.

Here, the distance-3 cyclic lowest-density MDS array code is one particular kind of the B-Code described in [14]. Its additional feature is the regularity in the form of cyclic symmetry. It is thus called the C-Code in this paper.

Definition 1.1 ([21]): A *one-factorization* of a graph is a partitioning of the set of its edges into subsets such that each subset is a graph of degree one. Here, each subset is called a *one-factor*. A *perfect one-factorization* (or *P1F*) is a particular one-factorization in which the union of any pair of one-factors is a Hamiltonian cycle.

(*Remark:* A Hamiltonian cycle is a cycle in an undirected graph, which visits each vertex exactly once and also returns

This work was supported by the National Science Foundation for Distinguished Young Scholars of China (Grant No. 60925006) and the National High Technology Research and Development Program of China (Grant No. 2009AA01A403). The material in this paper was presented in part at the IEEE International Symposium on Information Theory, Austin, TX, June 2010 (see [1]).

Mingqiang Li is with the IBM China Research Laboratory, Diamond Building 19-A, Zhongguancun Software Park, Dongbeiwang West Road No. 8, ShangDi, Haidian District, Beijing 100193, P.R.China (e-mail: mingqiangli@cn.ibm.com). The main part of this work was finished while Mingqiang Li was a Ph.D. student in the Department of Computer Science and Technology, Tsinghua University.
Jiwu Shu is with the Department of Computer Science and Technology, Tsinghua University, Qinghuayuan No. 1, Haidian District, Beijing 100084, P.R.China (e-mail:shujw@tsinghua.edu.cn).

to the starting vertex.)

The C-Code can be described using a graph approach proposed in [14] (see Section II-B). It will be shown in Section II-C that the constructions of the C-Code of length $2n$ (denoted by $\mathbb{C}_{2n}$) are equivalent to bipyramidal P1Fs of a $2n$-regular graph on $2n+2$ vertices. In the literature of graph theory, we noticed that several known P1Fs [22]–[28] were constructed using *starters* [29] in group theory. Inspired by this, we immediately raise a question: *Which kind of starter can be used to construct the C-Code?* In Section III, we will show that each $\mathbb{C}_{2n}$ instance can be constructed using an *even starter* [30] in $(Z_{2n}, +)$. The necessary and sufficient condition is that the even starter can induce a bipyramidal P1F of a $2n$-regular graph on $2n+2$ vertices. Then, we will obtained C-Codes for some sporadic even lengths listed as follows:

$$4, 6, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 50.$$

Among them, 14, 20, 24, 26, 32, 34, 38, and 50 are not covered by the family of length $p-1$ presented in [6] (where p is a prime). We will also show that a C-Code exists for most but not all of even lengths (one exception we found is 8) and that there are often more than one C-Code instances for a given even length (each C-Code instance always has a *twin C-Code instance*). Then, someone may wonder if there exists an infinite family of C-Codes. In Section IV, we will construct four infinite families of $\mathbb{C}_{p-1}$ instances (which cover the family of $\mathbb{C}_{p-1}$ instances constructed in [6]) from two infinite families of even starters in $\left(Z_p^*, \times\right)$. We will also conclude that non-cyclic B-Codes of length $p-1$ constructed in [20], [14], and [19] can always be transformed to $\mathbb{C}_{p-1}$ instances.

Besides, we noticed that there is no C-Code for some even lengths, such as 8. Then, someone may ask: *Can we construct quasi-C-Codes (which partially hold cyclic symmetry [6]) for these even lengths?* In Section V, we will introduce a concept of *even multi-starters* and then discuss how to construct quasi-C-Codes using even multi-starters (with an example of length 8). We will also show that each quasi-C-Code instance has a *twin quasi-C-Code instance*. Similarly, in Section VI, we will present an infinite family of even 2-starter in $\left(Z_{2(p-1)}, +\right)$ and then construct two infinite families of quasi-C-Code instances of length $2(p-1)$ (which cover the family of quasi-C-Code instances of length $2(p-1)$ constructed in [6]) using this family of even 2-starter. We will also conclude that non-cyclic B-Codes of length $2(p-1)$ constructed in [14] can always be transformed to quasi-C-Code instances of length $2(p-1)$.

Two work very related to this paper are [14] and [6]. Unlike the work in [14], which studied the constructions of B-Codes, the work in this paper focuses on the constructions of C-Codes (or quasi-C-Codes), which cannot be obtained directly from the constructions of B-Codes. Besides, in [6], although Cassuto and Bruck constructed one infinite family of $\mathbb{C}_{p-1}$ instances and one infinite family of quasi-C-Code instances of length $2(p-1)$, they did not study the general constructions of C-Codes (or quasi-C-Codes). In contrast, this paper carries out a systematic study on the constructions of C-Codes by revealing the underlying connections between C-Codes (or quasi-C-Codes) and starters in group theory.

We begin this paper with an introduction of the C-Code in the next section.

II. An Introduction of the C-Code

A. Definition and Structure

The C-Code is one particular kind of the B-Code described in [14]. Its additional feature is the regularity in the form of cyclic symmetry, and its algebraic definition is given as follows:

Definition 2.1: Let

$$\mathbf{H}_{2n} = (H_0 \ H_1 \ \cdots \ H_{2n-1})$$

be a binary matrix, where

$$H_k = (h_{i,j})_{2n\times n}$$

is a binary submatrix of size $2n \times n$, for $0 \le i \le 2n-1$, $0 \le j \le n-1$, and $0 \le k \le 2n-1$. Suppose $\mathbf{H}_{2n}$ meets the following four conditions:

1) for $k = 0, 1, \cdots, 2n-1$, the last column of H_k is the same as the k-th column of a binary $2n \times 2n$ identity matrix;
2) for $k = 0, 1, \cdots, 2n-1$,
$$H_k = E_{2n}^k \times H_0, \tag{1}$$
where E_{2n} is a binary *elemental cyclic matrix* defined as
$$E_{2n} = \begin{pmatrix} \overrightarrow{0} & 1 \\ I_{2n-1} & \overrightarrow{0}^T \end{pmatrix},$$
where I_{2n-1} is a binary $(2n-1)\times(2n-1)$ identity matrix, $\overrightarrow{0}$ is a binary $1\times(2n-1)$ vector of 0's, and $\overrightarrow{0}^T$ is a binary $(2n-1)\times 1$ vector of 0's;
3) the *weight* (i.e. the number of 1's) of each row of $\mathbf{H}_{2n}$ is $2n-1$; and
4) for any m and k (where $0 \le m < k \le 2n-1$), the square matrix $(H_m \ H_k)$ is nonsingular.

If a code's parity-check matrix is $\mathbf{H}_{2n}$, the code is then called the C-Code of length $2n$, denoted by $\mathbb{C}_{2n}$.

In the above definition, it should be noted that the length of the C-Code is always an even number. This is guaranteed by the MDS property of the B-Code [14].

Take $\mathbb{C}_4$ for example. The parity-check matrix for a $\mathbb{C}_4$ instance is as follows:

$$\mathbf{H}_4 = \left(\begin{array}{cc|cc|cc|cc} 0 & 1 & 0 & 0 & 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 & 0 & 0 & 1 \end{array}\right).$$

It can be easily checked that

$$\left\{I_{2n}, E_{2n}, E_{2n}^2, \cdots, E_{2n}^{2n-1}\right\}$$

forms a cyclic group with binary matrix multiplication. We also have

$$E_{2n}^{2n} = I_{2n}. \tag{2}$$

Thus, the code defined in Definition 2.1 has the property of cyclic symmetry.

In addition, it can also be deduced from the results in [14] that the C-Code defined in Definition 2.1 has all the following optimal properties:

1) it is Maximum-Distance Separable (MDS);
2) its update complexity is 2, which is the minimum update complexity that MDS codes of distance 3 can have; and
3) it achieves the maximum length that MDS codes with optimal update complexity can have.

The structure of the C-Code of length $2n$ evolves from that of the B-Code described in [14]. This kind of array code has dimensions $n \times 2n$, i.e. n rows and $2n$ columns, where n is an integer not smaller than 2. It was proved in [19] that this size has optimal length. The first $n-1$ rows are information rows, and the last row is a parity row. In other words, the bits in the first $n-1$ rows are information bits, while those in the last row are parity bits. Because of the optimal update complexity, each information bit contributes to the calculation of (or is protected by) exactly 2 parity bits contained in other columns. Moreover, any two information bits do not contribute to the calculation of the same pair of parity bits.

Take the foregoing $\mathbb{C}_4$ instance for example. It has 2 rows and 4 columns. Its array representation is given as follows:

$d_{1,2}$	$d_{2,3}$	$d_{3,0}$	$d_{0,1}$
p_0	p_1	p_2	p_3

,

where $d_{i,j}$ ($0 \le i \ne j \le 3$) represents a information bit that contributes to the calculation of (or is protected by) 2 parity bits p_i and p_j. Then, take the parity bit p_0 for example. It can be calculated by $p_0 = d_{3,0} + d_{0,1}$.

B. Graph Description

In the C-Code, each information bit contributes to the calculation of (or is protected by) exactly 2 parity bits contained in other columns. Moreover, any two information bits do not contribute to the same pair of parity bits. Thus, a graph approach [14] can be used to describe the C-Code.

In the graph description of the C-Code, each parity bit is represented by a vertex, and each information bit that contributes to the calculation of 2 parity bits is represented by the edge that connects the two corresponding vertices. Then, a C-Code of length $2n$ can be described by a $(2n-2)$-regular graph G on $2n$ vertices. We label the $2n$ vertices with integers from 0 to $2n-1$ such that the i-th vertex ($i = 0, 1, \cdots, 2n-1$) represents the parity bit contained in the i-th column of the C-Code. Then, for $i = 0, 1, \cdots, 2n-1$, the i-th column of the C-Code can be represented by a set of $n-1$ edges, i.e.

$$C_i = \{\{x_{i,1}, y_{i,1}\}, \{x_{i,2}, y_{i,2}\}, \cdots, \{x_{i,n-1}, y_{i,n-1}\}\},$$

where $\{x_{i,j}, y_{i,j}\}$ ($j = 1, 2, \cdots, n-1$) is an edge corresponding to an information bit contained in the i-th column. According to the cyclic symmetry of the C-Code, for $i = 0, 1, \cdots, 2n-1$, we have

$$C_i = \{\{x + i \bmod 2n, y + i \bmod 2n\} : \{x, y\} \in C_0\}. \quad (3)$$

Thus, in this paper, we sometimes use C_0 to simply represent a C-Code.

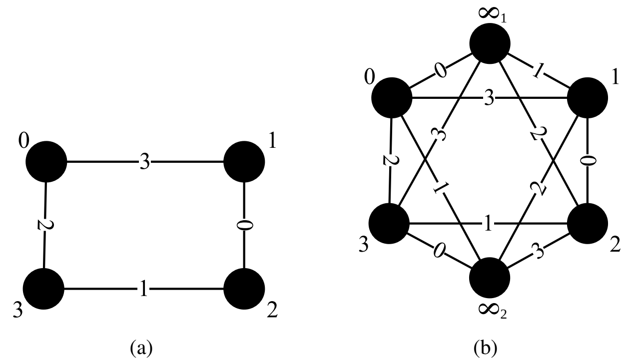

Fig. 1. Constructing (a) a $\mathbb{C}_4$ instance from (b) a bipyramidal P1F of a 4-regular graph on 6 vertices.

Take the foregoing $\mathbb{C}_4$ instance for example. Figure 1(a) shows its graph representation. The corresponding graph G is a 2-regular graph on a set of 4 vertices $\{0, 1, 2, 3\}$. The four columns of the code can be represented by

$$\begin{cases} C_0 = \{\{1, 2\}\}; \\ C_1 = \{\{2, 3\}\}; \\ C_2 = \{\{3, 0\}\}; \\ C_3 = \{\{0, 1\}\}. \end{cases}$$

It is clear that $\{C_0, C_1, C_2, C_3\}$ meets Equation (3). Then, this $\mathbb{C}_4$ instance can be represented simply by $C_0 = \{\{1, 2\}\}$.

Recall the C-Code of length $2n$ can recover the erasure of any two columns. This is guaranteed by the fourth condition of Definition 2.1. In the graph description, this condition is equivalent to the following one:

Condition 2.1: For any m and k (where $0 \le m < k \le 2n-1$), the subgraph

$$G^* = (\{0, 1, \cdots, 2n-1\}, C_m \cup C_k)$$

does not contain a cycle or a path whose terminal vertices are the two vertices m and k.

The above condition is explained by contradiction as follows:

We first consider the first opposite case where G^* contains a cycle of length r. In such a cycle, suppose the r edges are $e_1, e_2, \cdots, e_r$. As we know, in the corresponding square matrix mentioned in the fourth condition of Definition 2.1, the column vector corresponding to each edge is a vector of weight 2, whose two 1's are in the two rows corresponding to the two vertices of the edge. Then, in the square matrix, the binary sum of the r column vectors corresponding to $e_1, e_2, \cdots, e_r$ is a zero vertical vector, which conflicts with the nonsingular property of the square matrix. Thus, G^* should not contain a cycle.

We then consider the second opposite case where G^* contains a path of length r' whose terminal vertices are the two vertices m and k. In such a path, suppose the r' edges are ${e'}_1, {e'}_2, \cdots, {e'}_{r'}$. As we know, in the corresponding square matrix mentioned in the fourth condition of Definition 2.1, the column vector corresponding to the terminal vertex m (or k) is a vector of weight 1, whose only 1 is in the row corresponding to the vertex m (or k). Then, in the square matrix, the binary sum of the $r'+2$ column vectors corresponding to the two terminal vertices m and k and the r' edges $e'_1, e'_2, \cdots, e'_{r'}$ is a zero vertical vector, which conflicts with the nonsingular property of the square matrix. Thus, G^* should not contain a path whose terminal vertices are the two vertices m and k.

C. The Equivalence Between C-Code Constructions and Bipyramidal P1Fs

Definition 2.2: For a $2n$-regular graph on $2n+2$ vertices $0, 1, \cdots, 2n-1, \infty_1, \infty_2$ (where the two vertices ∞_1 and ∞_2 are not adjacent to each other), a *bipyramidal one-factorization* is a one-factorization consisting of $2n$ factors $F_0, F_1, \cdots, F_{2n-1}$, which are defined as

$$F_i = \left\{ \{\sigma^i(x), \sigma^i(y)\} : \{x, y\} \in F_0 \right\} \tag{4}$$

for $i = 0, 1, \cdots, 2n-1$, where

$$\sigma = (0\ 1\ \cdots\ 2n-1)(\infty_1)(\infty_2)$$

is a permutation represented by a product of disjoint cycles.

In the above definition, if the one-factorization is perfect, it is then called a *bipyramidal P1F* of a $2n$-regular graph on $2n+2$ vertices.

Then, we present the following theorem:

Theorem 2.1: The constructions of the C-Code of length $2n$ are equivalent to bipyramidal P1Fs of a $2n$-regular graph on $2n+2$ vertices. Suppose $\mathbb{F}$ is a bipyramidal P1F of a $2n$-regular graph on $2n+2$ vertices $0, 1, \cdots, 2n-1, \infty_1, \infty_2$ (where the two vertices ∞_1 and ∞_2 are not adjacent to each other), in which F_0 is the one-factor that contains the edge $\{0, \infty_1\}$. Then, the first column of the corresponding C-Code $\mathbb{C}_{2n}$ is

$$C_0 = F_0 \setminus \{\{0, \infty_1\}, \{r, \infty_2\}\}, \tag{5}$$

where r is the vertex that is adjacent to the vertex ∞_2 in F_0.

Proof: See Appendix A. ■

III. Constructing a C-Code Using Even Starters

We first give the definition of an even starter [30] in $(Z_{2n}, +)$.

Definition 3.1: An *even starter* S_E in $(Z_{2n}, +)$ is a set of $n-1$ pairs of non-zero elements, i.e.

$$S_E = \{\{x_1, y_1\}, \{x_2, y_2\}, \cdots, \{x_{n-1}, y_{n-1}\}\},$$

such that every non-zero element except n occurs in

$$\Delta = \{x-y, y-x : \{x, y\} \in S_E\}.$$

Its *twin even starter* S_E^τ is defined as

$$S_E^\tau = \{\{x-r, y-r\} : \{x, y\} \in S_E\}, \tag{6}$$

where r is the one and only non-zero element that does not occur in S_E.

Take $S_E = \{\{1, 2\}, \{3, 5\}\}$ in $(Z_6, +)$ for example. For every non-zero element except 3, we have

$$\begin{cases} 1 = 2-1 \bmod 6; \\ 2 = 5-3 \bmod 6; \\ 4 = 3-5 \bmod 6; \\ 5 = 1-2 \bmod 6. \end{cases}$$

Thus, S_E is an even starter in $(Z_6, +)$. Its twin even starter is $S_E^\tau = \{\{3, 4\}, \{5, 1\}\}$.

Then, we present the following theorem:

Theorem 3.1: For a $2n$-regular graph on $2n+2$ vertices $0, 1, \cdots, 2n-1, \infty_1, \infty_2$ (where the two vertices ∞_1 and ∞_2 are not adjacent to each other), suppose a one-factorization $\mathbb{F}$ is a bipyramidal one-factorization in which F_0 is the one-factor that contains the edge $\{0, \infty_1\}$. Let

$$S = F_0 \setminus \{\{0, \infty_1\}, \{r, \infty_2\}\}, \tag{7}$$

where r is the vertex that is adjacent to the vertex ∞_2 in F_0. Then, S is an even starter in $(Z_{2n}, +)$.

Proof: See Appendix B. ■

According to Theorem 2.1 in Section II-C, we can further make the following conclusion:

Theorem 3.2: In a C-Code $\mathbb{C}_{2n}$, the first column C_0 is always an even starter in $(Z_{2n}, +)$. An even starter S_E in $(Z_{2n}, +)$ can be used to construct a C-Code of length $2n$ if and only if S_E can induce a bipyramidal P1F of a $2n$-regular graph on $2n+2$ vertices.

At the same time, we can deduce the following conclusion:

Theorem 3.3: If a $\mathbb{C}_{2n}$ instance can be constructed using an even starter S_E in $(Z_{2n}, +)$, another $\mathbb{C}_{2n}$ instance can also be constructed using the twin even starter S_E^τ. They are called *twin* $\mathbb{C}_{2n}$ *instances*.

The above conclusion can be easily understood because twin even starters S_E and S_E^τ induce the same bipyramidal one-factorization.

Now, we discuss how to construct a C-Code $\mathbb{C}_{2n}$ using even starters in $(Z_{2n}, +)$. The steps to construct $\mathbb{C}_{2n}$ are to first find an even starter S_E in $(Z_{2n}, +)$ and then check whether the bipyramidal one-factorization $\mathbb{F}$ induced by S_E is a P1F of a $2n$-regular graph on a set of $2n+2$ vertices $0, 1, \cdots, 2n-1, \infty_1, \infty_2$ (where the two vertices ∞_1 and ∞_2 are not adjacent to each other). Here, let

$$\mathbb{F} = \{F_0, F_1, \cdots, F_{2n-1}\}.$$

According to the cyclic symmetry of $\mathbb{F}$, it is clear that if $F_0 \cup F_i$ is a Hamiltonian cycle for all i from 1 to n, $\mathbb{F}$ is then a P1F. Thus, only n rather than $\binom{2n}{2}$ subgraphs need to be checked in determining whether $\mathbb{F}$ is a P1F. According to Theorem 3.2, if $\mathbb{F}$ is a P1F, a $\mathbb{C}_{2n}$ instance

$$C_0 = S_E \tag{8}$$

can be constructed; otherwise we try other even starters in $(Z_{2n}, +)$ until a C-Code is constructed, or all even starters in $(Z_{2n}, +)$ have been checked.

For example, the foregoing even starter $S_E = \{\{1, 2\}, \{3, 5\}\}$ in $(Z_6, +)$ induces a P1F of a 6-regular graph on 8 vertices. Thus, a $\mathbb{C}_6$ instance illustrated as follows can be constructed using S_E:

$d_{1,2}$	$d_{2,3}$	$d_{3,4}$	$d_{4,5}$	$d_{5,0}$	$d_{0,1}$
$d_{3,5}$	$d_{4,0}$	$d_{5,1}$	$d_{0,2}$	$d_{1,3}$	$d_{2,4}$
p_0	p_1	p_2	p_3	p_4	p_5

At the same time, we can construct the twin $\mathbb{C}_6$ instance illustrated as follows using the twin even starter $S_E^{\tau} = \{\{3,4\},\{5,1\}\}$:

$d_{3,4}$	$d_{4,5}$	$d_{5,0}$	$d_{0,1}$	$d_{1,2}$	$d_{2,3}$
$d_{5,1}$	$d_{0,2}$	$d_{1,3}$	$d_{2,4}$	$d_{3,5}$	$d_{4,0}$
p_0	p_1	p_2	p_3	p_4	p_5

Finally, in the literature of graph theory, some bipyramidal P1Fs of a complete graph on $2n+2$ vertices (denoted by K_{2n+2}), which are induced by even starters in $(Z_{2n},+)$, have been found for the following values of $2n$: 4, 6, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 34, 38, and 50 [22], [24], [25], [27]. Here, a bipyramidal P1F $\mathbb{F}$ of K_{2n+2} is induced by an even starter S_E in $(Z_{2n},+)$ as follows:

$$\mathbb{F} = \{F_0, F_1, \cdots, F_{2n-1}, F_{2n}\},$$

where $\{F_0, F_1, \cdots, F_{2n-1}\}$ is a bipyramidal P1F of a $2n$-regular graph on $2n+2$ vertices induced by S_E, and

$$F_{2n} = \{\{0,n\},\{1,n+1\},\cdots,\{n-1,2n-1\},\{\infty_1,\infty_2\}\}.$$

Thus, C-Codes for these values of $2n$ can be constructed using the corresponding even starters presented in [24], [22], [25], and [27]. For example, a $\mathbb{C}_{50}$ instance can be constructed using the even starter presented in [27]:

$$\begin{aligned} C_0 = \{&\{2,29\},\{3,35\},\{4,16\},\{5,33\},\{6,43\},\{7,15\},\\ &\{8,19\},\{9,30\},\{10,41\},\{11,46\},\{12,17\},\\ &\{13,20\},\{14,28\},\{18,38\},\{21,27\},\{22,23\},\\ &\{24,48\},\{25,34\},\{26,36\},\{31,47\},\{32,49\},\\ &\{37,39\},\{40,44\},\{42,45\}\}. \end{aligned}$$

Then, C-Codes for lengths 14, 20, 24, 26, 34, 38, and 50, which are not covered by the family of length $p-1$ presented in [6] (where p is a prime), can be constructed here.

It should be noted that an exhaustive search showed that a C-Code exists for most but not all of even lengths. For example, a C-Code exists for every even length from 4 to 36 except 8 (see Table II). Here, the length 32 is also not covered by the family of length $p-1$ presented in [6].

It should also be noted that there are often more than one C-Code instances for a given even length. For example, besides the $\mathbb{C}_{34}$ instance listed in Table II, another $\mathbb{C}_{34}$ instance can be constructed using the even starter presented in [24]:

$$\begin{aligned} C_0 = \{&\{1,2\},\{3,5\},\{4,24\},\{6,9\},\{7,22\},\{8,18\},\\ &\{10,17\},\{12,25\},\{13,21\},\{14,23\},\{15,31\},\\ &\{16,28\},\{19,30\},\{20,26\},\{27,32\},\{29,33\}\}. \end{aligned}$$

These two $\mathbb{C}_{34}$ instances are not twin instances. In addition, even for a length 6 covered by the family of length $p-1$ presented in [6], besides the instance $C_0 = \{\{1,3\},\{4,5\}\}$ constructed in [6], we can find another instance in Table II. These two $\mathbb{C}_6$ instances are also not twin instances. Furthermore, as will be shown in the next section, there exist four families of $\mathbb{C}_{p-1}$ instances. Besides, Table III gives the number of C-Codes for even lengths from 4 to 30. These results are derived from [22]. From this table, we can see that for most but not all of even lengths, the number of C-Codes increases with the length. We can also observe that the number of C-Codes for each length is always an even number. The reason is that each C-Code instance always has a twin instance (see Theorem 3.3).

TABLE III
THE NUMBER (#) OF C-CODES FOR EVEN LENGTHS FROM 4 TO 30.

Length	4	6	8	10	12	14	16
#	2	4	0	16	24	12	80
Length	18	20	22	24	26	28	30
#	120	272	440	576	2016	4992	11104

IV. FOUR INFINITE FAMILIES OF $\mathbb{C}_{p-1}$ INSTANCES

In the previous section, we have obtained C-Codes for some sporadic even lengths listed as follows:

$$4,6,10,12,14,16,18,20,22,24,26,28,30,32,34,36,38,50.$$

Then, someone may wonder if there exists an infinite family of C-Codes. A positive answer to this question will be given in this section. Exactly speaking, this section will construct four infinite families of $\mathbb{C}_{p-1}$ instances.

We first give the extended definition of an event starter as follows:

Definition 4.1 ([30]): Let $(A_{2n},\circ)$ be an abelian group (written multiplicatively) of order $2n$ with identity e and unique element a^* of order 2 (i.e. $a^* \circ a^* = e$). An *even starter* $\widehat{S}_E$ in $(A_{2n},\circ)$ is a set of $n-1$ pairs of non-identity elements of A_{2n}, i.e.

$$\widehat{S}_E = \{\{x_1,y_1\},\{x_2,y_2\},\cdots,\{x_{n-1},y_{n-1}\}\},$$

such that every non-identity element of A_{2n} except a^* occurs in

$$\left\{x^{-1}\circ y, x\circ y^{-1} : \{x,y\} \in \widehat{S}_E\right\}.$$

Let $\left(Z_p^*,\times\right)$ be a multiplicative group of congruence classes modulo p. For $p=7$,

$$Z_7^* = \{1,2,\cdots,6\}.$$

It is clear that $(Z_7^*,\times)$ is an abelian group of order 6 with identity 1 and unique element 6 of order 2. Take $\widehat{S}_E = \{\{2,6\},\{3,5\}\}$ in $(Z_7^*,\times)$ for example. For every non-identity element of Z_p^* except 6, we have

$$\begin{cases} 2 = 3\times 5^{-1} \bmod 7; \\ 3 = 2^{-1}\times 6 \bmod 7; \\ 4 = 3^{-1}\times 5 \bmod 7; \\ 5 = 2\times 6^{-1} \bmod 7. \end{cases}$$

Thus, $\widehat{S}_E$ is an even starter in $(Z_7^*,\times)$.

An even starter $\widehat{S}_E$ in $(A_{2n},\circ)$ induces a bipyramidal one-factorization of a $2n$-regular graph on $2n+2$ vertices as follows. Label the $2n+2$ vertices with the elements of A_{2n}

TABLE II
SOME EXAMPLES OF C-CODES FOR EVEN LENGTHS FROM 4 TO 36.

Length	C_0 (i.e. S_E)
4	$\{\{1,2\}\}$
6	$\{\{1,2\},\{3,5\}\}$
8	(Not Exist!)
10	$\{\{1,2\},\{3,5\},\{4,8\},\{6,9\}\}$
12	$\{\{1,10\},\{2,6\},\{3,5\},\{4,9\},\{7,8\}\}$
14	$\{\{1,2\},\{3,11\},\{4,6\},\{5,9\},\{7,10\},\{8,13\}\}$
16	$\{\{1,2\},\{3,13\},\{4,15\},\{5,14\},\{6,8\},\{7,11\},\{9,12\}\}$
18	$\{\{1,2\},\{3,7\},\{4,11\},\{5,15\},\{6,9\},\{8,13\},\{10,16\},\{12,14\}\}$
20	$\{\{1,2\},\{3,5\},\{4,17\},\{6,14\},\{7,18\},\{8,13\},\{9,12\},\{10,16\},\{11,15\}\}$
22	$\{\{1,2\},\{3,6\},\{4,12\},\{5,9\},\{7,13\},\{8,21\},\{10,20\},\{11,18\},\{14,19\},\{15,17\}\}$
24	$\{\{1,2\},\{3,5\},\{4,21\},\{6,11\},\{7,20\},\{8,12\},\{9,19\},\{10,16\},\{13,22\},\{14,17\},\{15,23\}\}$
26	$\{\{1,2\},\{3,6\},\{4,25\},\{5,19\},\{7,14\},\{8,24\},\{9,11\},\{10,18\},\{12,23\},\{13,22\},\{15,21\},\{16,20\}\}$
28	$\{\{1,2\},\{3,6\},\{4,25\},\{5,21\},\{7,11\},\{8,16\},\{9,18\},\{10,27\},\{12,22\},\{13,26\},\{14,20\},\{15,17\},\{19,24\}\}$
30	$\{\{1,2\},\{3,5\},\{4,9\},\{6,25\},\{7,13\},\{8,21\},\{10,24\},\{11,29\},\{12,16\},\{14,23\},\{15,22\},\{17,20\},\{18,28\},$ $\{19,27\}\}$
32	$\{\{1,2\},\{3,5\},\{4,8\},\{6,27\},\{7,24\},\{9,21\},\{10,19\},\{11,29\},\{12,31\},\{13,18\},\{14,17\},\{15,25\},\{16,22\},$ $\{20,28\},\{23,30\}\}$
34	$\{\{1,2\},\{3,5\},\{4,10\},\{6,25\},\{7,14\},\{8,32\},\{9,18\},\{11,22\},\{12,20\},\{13,26\},\{15,33\},\{16,30\},\{17,21\},$ $\{19,31\},\{23,28\},\{24,27\}\}$
36	$\{\{1,2\},\{3,5\},\{4,8\},\{6,11\},\{7,20\},\{9,18\},\{10,34\},\{12,26\},\{13,28\},\{14,33\},\{15,35\},\{16,22\},\{17,25\},$ $\{19,29\},\{21,32\},\{23,30\},\{24,27\}\}$

and two infinity elements ∞_1 and ∞_2 such that there is no edge between the following pairs of vertices: $\{\infty_1,\infty_2\}$ and all $\{a, a\circ a^*\}$ for $a\in A_{2n}$. Let

$$\widetilde{S}_E=\widehat{S}_E\cup\{\{e,\infty_1\},\{r,\infty_2\}\}, \tag{9}$$

where r is the non-identity element that does not appear in $\widehat{S}_E$. For all $a\in A_{2n}$, define $a\circ\infty_1=\infty_1$ and $a\circ\infty_2=\infty_2$. The corresponding bipyramidal one-factorization $\mathbb{F}$ is then given by

$$\mathbb{F}=\left\{a\circ\widetilde{S}_E : a\in A_{2n}\right\}, \tag{10}$$

where

$$a\circ\widetilde{S}_E=\left\{\{a\circ x, a\circ y\}:\{x,y\}\in\widetilde{S}_E\right\}.$$

Here, if the bipyramidal one-factorization induced by $\widehat{S}_E$ in $(A_{2n},\circ)$ is a P1F, a non-cyclic B-Code of length $2n$, in which the a-th column ($a\in A_{2n}$) is $a\circ\widehat{S}_E$, can be constructed [14].

Take the foregoing even starter $\widehat{S}_E=\{\{2,6\},\{3,5\}\}$ in $(Z_7^*,\times)$ for example. Since the corresponding bipyramidal one-factorization is a P1F, a non-cyclic B-Code of length 6 constructed using $\widehat{S}_E$ is illustrated as follows:

$d_{2,6}$	$d_{4,5}$	$d_{6,4}$	$d_{1,3}$	$d_{3,2}$	$d_{5,1}$
$d_{3,5}$	$d_{6,3}$	$d_{2,1}$	$d_{5,6}$	$d_{1,4}$	$d_{4,2}$
p_1	p_2	p_3	p_4	p_5	p_6

This code is the same as the P-Code of length 6 constructed in [20] (in fact, P-Code is just one family of the B-Code of length $p-1$, and its code structure was originally derived from [19]).

We now consider the case where $(A_{2n},\circ)$ is a cyclic group of which g is a generator. Then, in the non-cyclic B-Code of length $2n$ constructed using $\widehat{S}_E$ in $(A_{2n},\circ)$, the (g^i)-th column ($i=0,1,\cdots,2n-1$) can be expressed as $g^i\circ\widehat{S}_E$. For $i=0,1,\cdots,2n-1$, replace g^i with i and then relabel the (g^i)-th column with i. Reorder all the $2n$ columns in order according to their new labels. Then, a C-Code of length $2n$ is obtained.

Take the foregoing non-cyclic B-Code of length 6 for example. Since $(Z_7^*,\times)$ is a cyclic group of which 3 is a generator, the code can then be represented by

$d_{3^2,3^3}$	$d_{3^4,3^5}$	$d_{3^3,3^4}$	$d_{3^0,3^1}$	$d_{3^1,3^2}$	$d_{3^5,3^0}$
$d_{3^1,3^5}$	$d_{3^3,3^1}$	$d_{3^2,3^0}$	$d_{3^5,3^3}$	$d_{3^0,3^4}$	$d_{3^4,3^2}$
p_{3^0}	p_{3^2}	p_{3^1}	p_{3^4}	p_{3^5}	p_{3^3}

In the above representation, replace 3^i with i for $i=0,1,\cdots,5$ and then reorder all the 6 columns in order according to their new labels. We can obtain a $\mathbb{C}_6$ instance as follows:

$d_{2,3}$	$d_{3,4}$	$d_{4,5}$	$d_{5,0}$	$d_{0,1}$	$d_{1,2}$
$d_{1,5}$	$d_{2,0}$	$d_{3,1}$	$d_{4,2}$	$d_{5,3}$	$d_{0,4}$
p_0	p_1	p_2	p_3	p_4	p_5

In a cyclic group (written multiplicatively) of which g is a generator, for $x=g^i$, define $\log_g(x)=i$. Then, we can easily make the following conclusion:

Theorem 4.1: If $(A_{2n},\circ)$ is a cyclic group of which g is a generator, a non-cyclic B-Code of length $2n$ constructed using an even starter $\widehat{S}_E$ in $(A_{2n},\circ)$ can always be transformed to a $\mathbb{C}_{2n}$ instance

$$C_0=\left\{\{\log_g(x),\log_g(y)\}:\{x,y\}\in\widehat{S}_E\right\}. \tag{11}$$

At the same time, we can get the twin $\mathbb{C}_{2n}$ instance

$$C_0^\tau=\{\{x^*-r^*,y^*-r^*\}:\{x^*,y^*\}\in C_0\}, \tag{12}$$

where r^* is the one and only non-zero element of $(Z_{2n},+)$ that does not occur in C_0.

Specially, we consider even starters in $\left(Z_p^*,\times\right)$, where p is a prime. It is well-known that when p is a prime, $\left(Z_p^*,\times\right)$ is a cyclic group in which

$$Z_p^*=\{1,2,\cdots,p-1\}.$$

Thus, we can make the following conclusion:

Theorem 4.2: A non-cyclic B-Code of length $p-1$ constructed using an even starter in $\left(Z_p^*, \times\right)$ can always be transformed to a C-Code of length $p-1$.

Finally, in $\left(Z_p^*, \times\right)$, there exist two infinite families of even starters [28] as follows:

$$\widehat{S}_E^A = \left\{\{x, y\} : x, y \in Z_p^* \setminus \{1, 2^{-1}\}, x + y = 1\right\}, \quad (13)$$

and

$$\widehat{S}_E^B = \big\{\{x, y\} : x, y \in Z_p^* \setminus \{1, 2^{-1}, 2, p-1\}, \\ x + y = 1\} \bigcup \left\{\{2^{-1}, p-1\}\right\}. \quad (14)$$

It was proved in [28] that $\widehat{S}_E^A$ and $\widehat{S}_E^B$ can induce two families of non-isomorphic bipyramidal P1Fs of K_{p+1}, respectively. Note that the P1F induced by $\widehat{S}_E^A$ is isomorphic to the well-known *patterned P1F* (induced by the well-known *patterned starter* in $(Z_p, +)$) [31], which has been used to construct the family of non-cyclic B-Codes of length $p-1$ in [20], [14], and [19]. Thus, two families of non-cyclic B-Codes of length $p-1$ can be constructed using $\widehat{S}_E^A$ and $\widehat{S}_E^B$, respectively.

Suppose g is a generator of $\left(Z_p^*, \times\right)$. Then, two families of $\mathbb{C}_{p-1}$ instances

$$C_0^A = \left\{\left\{\log_g(x), \log_g(y)\right\} : x, y \in Z_p^* \setminus \{1, 2^{-1}\}, x + y = 1\right\} \quad (15)$$

and

$$C_0^B = \big\{\left\{\log_g(x), \log_g(y)\right\} : x, y \in Z_p^* \setminus \{1, 2^{-1}, 2, p-1\}, \\ x + y = 1\} \bigcup \left\{\{2^{-1}, p-1\}\right\} \quad (16)$$

and their twin instances can be constructed. Therefore, there exist four families of $\mathbb{C}_{p-1}$ instances.

Take $(Z_7^*, \times)$ for example. We then have $\widehat{S}_E^A = \{\{2, 6\}, \{3, 5\}\}$ and $\widehat{S}_E^A = \{\{3, 5\}, \{4, 6\}\}$. In $(Z_7^*, \times)$, pick $g = 3$. From $\widehat{S}_E^A$, we obtain a $\mathbb{C}_6$ instance

$$C_0^A = \{\{2, 3\}, \{1, 5\}\}$$

and its twin instance

$$\left(C_0^A\right)^\tau = \{\{4, 5\}, \{3, 1\}\}.$$

Here, the twin instance is the same as the instance constructed in [6]. Also, from $\widehat{S}_E^B$, we obtain a $\mathbb{C}_6$ instance

$$C_0^B = \{\{1, 5\}, \{4, 3\}\}$$

and its twin instance

$$\left(C_0^B\right)^\tau = \{\{5, 3\}, \{2, 1\}\}.$$

Thus, we construct four $\mathbb{C}_6$ instances.

From the above results, we can make two observations as follows:

1) Non-cyclic B-Codes of length $p-1$ constructed in [20], [14], and [19] can always be transformed to $\mathbb{C}_{p-1}$ instances; and
2) The family of $\mathbb{C}_{p-1}$ instances constructed in [6] can also be obtained from $\widehat{S}_E^A$.

V. Constructing a Quasi-C-Code Using Even Multi-Starters

As mentioned in Section III, our exhaustive search showed that there is no C-Code for some even lengths, such as 8. Then, someone may ask: *Can we construct quasi-C-Codes (which partially hold cyclic symmetry [6]) for these even lengths?*

In this section, we will introduce a concept of even multi-starters and then discuss how to construct quasi-C-Codes using even multi-starters.

An even κ-starter in $(Z_{2n}, +)$ (where $\kappa|2n$) is defined as follows:

Definition 5.1: An even *κ-starter* S^κ in $(Z_{2n}, +)$ (where $\kappa|2n$) is a set

$$S^\kappa = \{S_0, S_1, \cdots, S_{\kappa-1}\},$$

where S_i $(i = 0, 1, \cdots, \kappa-1)$ is a set of $n-1$ pairs of non-i elements of Z_{2n}, such that every integer from 1 to $n-1$ occurs κ times as a difference of a pair of S^κ. Its *twin even κ-starter* $(S^\kappa)^\tau$ is defined as

$$(S^\kappa)^\tau = \left\{S'_{r_i \bmod \kappa} = S_i - \kappa\left\lfloor\frac{r_i}{\kappa}\right\rfloor : i = 0, 1, \cdots, \kappa-1\right\}, \quad (17)$$

where r_i is the non-i element that does not appear in S_i, and

$$S_i - \kappa\left\lfloor\frac{r_i}{\kappa}\right\rfloor = \left\{\left\{x - \kappa\left\lfloor\frac{r_i}{\kappa}\right\rfloor, y - \kappa\left\lfloor\frac{r_i}{\kappa}\right\rfloor\right\} : \{x, y\} \in S_i\right\}.$$

Take $S^2 = \{S_0, S_1\}$ in $(Z_8, +)$ for example, where $S_0 = \{\{1, 2\}, \{3, 5\}, \{4, 6\}\}$, and $S_1 = \{\{0, 3\}, \{2, 7\}, \{4, 5\}\}$. For every integer from 1 to 3, we have

$$\begin{cases} 1 = 2 - 1 = 5 - 4 \bmod 8; \\ 2 = 5 - 3 = 6 - 4 \bmod 8; \\ 3 = 3 - 0 = 2 - 7 \bmod 8. \end{cases}$$

Thus, S^2 is an even 2-starter in $(Z_8, +)$. Its twin even 2-starter is $(S^2)^\tau = \{S'_0, S'_1\}$, where $S'_0 = \{\{2, 5\}, \{4, 1\}, \{6, 7\}\}$, and $S'_1 = \{\{3, 4\}, \{5, 7\}, \{6, 0\}\}$.

An even κ-starter

$$S^\kappa = \{S_0, S_1, \cdots, S_{\kappa-1}\}$$

in $(Z_{2n}, +)$ (where $\kappa|2n$) induces a one-factorization of a $2n$-regular graph on $2n+2$ vertices as follows. Label these $2n+2$ vertices with the elements of Z_{2n} and two infinity elements ∞_1 and ∞_2 such that there is no edge between the following pairs of vertices: $\{\infty_1, \infty_2\}$ and all $\{i, i+n\}$ for $i = 0, 1, \cdots, n-1$. For every $z \in Z_{2n}$, define $z + \infty_1 = \infty_1$ and $z + \infty_2 = \infty_2$. For $i = 0, 1, \cdots, \kappa-1$, let

$$\widetilde{S}_i = S_i \cup \{\{i, \infty_1\}, \{r_i, \infty_2\}\}, \quad (18)$$

where r_i is the non-i element that does not appear in S_i. The corresponding one-factorization $\mathbb{F}^\kappa$ is then given by

$$\mathbb{F}^\kappa = \bigg\{\kappa\tau + \widetilde{S}_0, \kappa\tau + \widetilde{S}_1, \cdots, \kappa\tau + \widetilde{S}_{\kappa-1} : \\ \tau = 0, 1, \cdots, \frac{2n}{\kappa} - 1\bigg\}, \quad (19)$$

where

$$\kappa\tau+\widetilde{S}_i=\left\{\{\kappa\tau+x,\kappa\tau+y\}:\{x,y\}\in\widetilde{S}_i\right\}$$

for $i=0,1,\cdots,\kappa-1$. Such a one-factorization is called a *κ-quasi-bipyramidal one-factorization*.

For an even κ-starter

$$S^\kappa=\{S_0,S_1,\cdots,S_{\kappa-1}\}$$

in $(Z_{2n},+)$ (where $\kappa|2n$), if the κ-quasi-bipyramidal one-factorization $\mathbb{F}^\kappa$ induced by S^κ is a P1F of a $2n$-regular graph on $2n+2$ vertices, a κ-quasi-C-Code of length $2n$ (denoted by $\mathbb{C}^\kappa_{2n}$), in which the i-th column ($i=0,1,\cdots,2n-1$) is

$$C_i=\left\{\left\{x+\kappa\left\lfloor\frac{i}{\kappa}\right\rfloor \bmod 2n, y+\kappa\left\lfloor\frac{i}{\kappa}\right\rfloor \bmod 2n\right\}:\{x,y\}\in S_{i \bmod \kappa}\right\}, \quad (20)$$

can be constructed using S^κ. It can be easily checked that in a $\mathbb{C}^\kappa_{2n}$ instance, for $i=0,1,\cdots,\kappa-1$, each group of $\frac{2n}{\kappa}$ columns

$$C_{i+\kappa\times0},C_{i+\kappa\times1},\cdots,C_{i+\kappa\times(\frac{2n}{\kappa}-1)}$$

hold cyclic symmetry.

Similar to Theorem 3.3 in Section III, we give the following theorem:

Theorem 5.1: If a $\mathbb{C}^\kappa_{2n}$ instance can be constructed using an even κ-starter S^κ in $(Z_{2n},+)$, another $\mathbb{C}^\kappa_{2n}$ instance can also be constructed using the twin even κ-starter $(S^\kappa)^\tau$. They are called *twin $\mathbb{C}^\kappa_{2n}$ instances*.

The above conclusion can also be easily understood because twin even κ-starters S^κ and $(S^\kappa)^\tau$ induce the same κ-quasi-bipyramidal one-factorization.

For example, the foregoing even 2-starter $S^2=\{S_0,S_1\}$ (where $S_0=\{\{1,2\},\{3,5\},\{4,6\}\}$, and $S_1=\{\{0,3\},\{2,7\},\{4,5\}\}$) in $(Z_8,+)$ induces a 2-quasi-bipyramidal P1F of a 8-regular graph on 10 vertices. Thus, a $\mathbb{C}^2_8$ instance illustrated as follows can be constructed using S^2:

$d_{1,2}$	$d_{0,3}$	$d_{3,4}$	$d_{2,5}$	$d_{5,6}$	$d_{4,7}$	$d_{7,0}$	$d_{6,1}$
$d_{3,5}$	$d_{2,7}$	$d_{5,7}$	$d_{4,1}$	$d_{7,1}$	$d_{6,3}$	$d_{1,3}$	$d_{0,5}$
$d_{4,6}$	$d_{4,5}$	$d_{6,0}$	$d_{6,7}$	$d_{0,2}$	$d_{0,1}$	$d_{2,4}$	$d_{2,3}$
p_0	p_1	p_2	p_3	p_4	p_5	p_6	p_7

.

At the same time, we can construct the twin $\mathbb{C}^2_8$ instance illustrated as follows using the twin even 2-starter $(S^2)^\tau=\{S'_0,S'_1\}$, where $S'_0=\{\{2,5\},\{4,1\},\{6,7\}\}$, and $S'_1=\{\{3,4\},\{5,7\},\{6,0\}\}$:

$d_{2,5}$	$d_{3,4}$	$d_{4,7}$	$d_{5,6}$	$d_{6,1}$	$d_{7,0}$	$d_{0,3}$	$d_{1,2}$
$d_{4,1}$	$d_{5,7}$	$d_{6,3}$	$d_{7,1}$	$d_{0,5}$	$d_{1,3}$	$d_{2,7}$	$d_{3,5}$
$d_{6,7}$	$d_{6,0}$	$d_{0,1}$	$d_{0,2}$	$d_{2,3}$	$d_{2,4}$	$d_{4,5}$	$d_{4,6}$
p_0	p_1	p_2	p_3	p_4	p_5	p_6	p_7

.

VI. Two Infinite Families of $\mathbb{C}^2_{2(p-1)}$ Instances

In this section, we will construct two infinite families of $\mathbb{C}^2_{2(p-1)}$ instances. We start with an infinite family of even 2-starter S^2 and its twin even 2-starter $\left(S^2\right)^\tau$ in $\left(Z_{2(p-1)},+\right)$.

Suppose g is a generator of $\left(Z^*_p,\times\right)$. In $\left(Z^*_p,\times\right)$, for $x=g^i$, define $\log_g(x)=i$. Then, $S^2=\{S_0,S_1\}$ in $\left(Z_{2(p-1)},+\right)$ is constructed as follows:

$$S_0=\left\{\left\{2\log_g(x),2\log_g(y)+1\right\}:x\in Z^*_p\setminus\{1\}, y\in Z^*_p\setminus\{p-1\},x-y=1\right\}, \quad (21)$$

and

$$\begin{aligned}S_1=&\left\{\{2x+1,2y+1\}:\{x,y\}\in C^A_0\right\}\bigcup\\&\left\{\{2x,2y\}:\{x,y\}\in C^A_0\right\}\bigcup\\&\{\{2r,2r+1\}\},\end{aligned} \quad (22)$$

where C^A_0 is defined in Equation (15) in Section IV, and r is the one and only non-zero element of $(Z_{p-1},+)$ that does not occur in C^A_0. Its twin even 2-starter in $\left(Z_{2(p-1)},+\right)$ is $\left(S^2\right)^\tau=\{S'_0,S'_1\}$, where

$$S'_0=S_1, \quad (23)$$

and

$$S'_1=\left\{\left\{2\log_g(x)+1,2\log_g(y)\right\}:x\in Z^*_p\setminus\{1\}, y\in Z^*_p\setminus\{p-1\},x-y=1\right\}. \quad (24)$$

Take $(Z_8,+)$ for example. Pick $g=2$ in $(Z^*_5,\times)$. We then have $S^2=\{S_0,S_1\}$, where $S_0=\{\{2,1\},\{6,3\},\{4,7\}\}$, and $S_1=\{\{2,4\},\{3,5\},\{6,7\}\}$. Its twin even 2-starter in $(Z_8,+)$ is $\left(S^2\right)^\tau=\{S'_0,S'_1\}$, where $S'_0=\{\{2,4\},\{3,5\},\{6,7\}\}$, and $S'_1=\{\{3,0\},\{7,2\},\{5,6\}\}$.

It can be verified that this family of even 2-starter S^2 and its twin even 2-starter $\left(S^2\right)^\tau$ in $\left(Z_{2(p-1)},+\right)$ can induce the same 2-quasi-bipyramidal P1F of K_{2p}, which is isomorphic to the well-known P1F GA_{2p} of K_{2p} [31]. Thus, two families of $\mathbb{C}^2_{2(p-1)}$ instances can be constructed using S^2 and $\left(S^2\right)^\tau$, respectively.

Here, the family of $\mathbb{C}^2_{2(p-1)}$ instances constructed using S^2 can be shown to be the same as those constructed in [6]. Besides, it was proved in [32] that the P1F GN_{2p} of K_{2p}, which was adopted in [14] to construct the family of non-cyclic B-Codes of length $2(p-1)$, is also isomorphic to GA_{2p}. Thus, we can make two observations as follows:

1) Non-cyclic B-Codes of length $2(p-1)$ constructed in [14] can always be transformed to $\mathbb{C}^2_{2(p-1)}$ instances; and
2) The family of $\mathbb{C}^2_{2(p-1)}$ instances constructed in [6] can also be constructed using S^2.

VII. Conclusions and Remarks

This paper investigated the underlying connections between distance-3 cyclic (or quasi-cyclic) lowest-density MDS array codes and starters in group theory. Some interesting new results listed as follows were obtained:

1) Each cyclic code of length $2n$ can be constructed using an even starter in $(Z_{2n}, +)$ (see Section III), while each quasi-cyclic code of length $2n$ can be constructed using an even multi-starter in $(Z_{2n}, +)$ (see Section V);
2) Each cyclic (or quasi-cyclic) code has a twin cyclic (or quasi-cyclic) code (see Sections III and V);
3) A cyclic code exists for most but not all of even lengths (one exception is 8) (see Section III);
4) Four infinite families of cyclic codes of length $p-1$ (which cover the family of cyclic codes of length $p-1$ constructed in [6]) were constructed from two infinite families of even starters in $\left(Z_p^*, \times\right)$ (where p is a prime) in Section IV;
5) Besides the family of length $p-1$, cyclic codes for some sporadic even lengths listed as follows were obtained in Section III:

$$14, 20, 24, 26, 32, 34, 38, 50;$$

6) Two infinite families of quasi-cyclic codes of length $2(p-1)$ (which cover the family of quasi-cyclic codes of length $2(p-1)$ constructed in [6]) were constructed using an infinite family of even 2-starter in $\left(Z_{2(p-1)}, +\right)$ in Section VI; and
7) Non-cyclic B-Codes of length $p-1$ constructed in [20], [14], and [19] can always be transformed to cyclic codes (see Section IV), while non-cyclic B-Codes of length $2(p-1)$ constructed in [14] can always be transformed to quasi-cyclic codes (see Section VI).

TABLE IV
THE EXISTENCE OF DISTANCE-3 CYCLIC (OR QUASI-CYCLIC) LOWEST-DENSITY MDS ARRAY CODES FOR EVEN LENGTHS FROM 4 TO 58.

Length	4	6	8	10	12	14	16
Cyclic	√	√	×	√	√	√	√
Quasi-Cyclic			√		√		

Length	18	20	22	24	26	28	30
Cyclic	√	√	√	√	√	√	√
Quasi-Cyclic		√		√			

Length	32	34	36	38	40	42	44
Cyclic	√	√	√	√	√	√	?
Quasi-Cyclic	√		√				√

Length	46	48	50	52	54	56	58
Cyclic	√	?	√	√	?	?	√
Quasi-Cyclic						√	

√: existence; ×: inexistence; ?: unknown.

According to the above results, we can obtain Table IV. From this table, we can see that for even lengths from 4 to 58, there are one length 8, for which the cyclic code does not exist, and four lengths 44, 48, 54, and 56, for which cyclic codes are still unknown. Luckily, quasi-cyclic codes for lengths 8, 44, and 56 can be constructed in Section VI. Then, the constructions of cyclic (or quasi-cyclic) codes for the rest two lengths 48 and 54 are left as open problems. Here, two points deserve future researchers' attention:

1) Non-cyclic B-Codes of length 48 can be constructed using P1Fs of K_{50} found in [26]. However, these P1Fs were induced by starters in $(Z_{49}, +)$. Whether these non-cyclic B-Codes can be transformed to cyclic codes is left as an open problem.
2) Since 2009, when a P1F of K_{52} was found (see [27]), K_{56} has been the smallest complete graph for which a P1F has not been known. The construction of a P1F of K_{56} is left as an open problem in graph theory. Consequently, the construction of a B-Code of length 54 is still unknown.

APPENDIX A
PROOF OF THEOREM 2.1

We prove this theorem by two algorithms. Here, note that the basic idea comes from the work of [14], and a similar proof was given in [33].

We now first propose Algorithm 1 to construct a $\mathbb{C}_{2n}$ instance from a bipyramidal P1F of a $2n$-regular graph on $2n+2$ vertices. In this algorithm, it is clear that $\{C_0, C_1, \cdots, C_{2n-1}\}$ meets Equation (3) in Section II-B. It can also be proved as follows that $\{C_0, C_1, \cdots, C_{2n-1}\}$ meets Condition 2.1 in Section II-B. Thus, a corresponding $\mathbb{C}_{2n}$ instance is constructed by Algorithm 1.

Algorithm 1 Constructing a $\mathbb{C}_{2n}$ instance from a bipyramidal P1F of a $2n$-regular graph on $2n+2$ vertices.

(S1) Choose arbitrary pair of vertices that are not adjacent to each other in the regular graph and label them with ∞_1 and ∞_2. Then, label the other $2n$ vertices of the regular graph with integers from 0 to $2n-1$.
(S2) If a bipyramidal P1F exists for the regular graph, then let F_i denote the one-factor that contains the edge $\{i, \infty_1\}$, where $i = 0, 1, \ldots, 2n-1$.
(S3) In each F_i, delete the two edges that are incident to the two vertices ∞_1 and ∞_2. Then, delete the two vertices ∞_1 and ∞_2 in the graph. For $i = 0, 1, \ldots, 2n-1$, let $C_i = F_i \setminus \{\{i, \infty_1\}, \{r_i, \infty_2\}\}$, where r_i is the vertex that is adjacent to the vertex ∞_2 originally in F_i, and label all the edges in C_i with i.

According to Definition 1.1 in Section I, in a P1F, for any pair of one-factors F_{i_1} and F_{i_2}, the union of them forms a Hamiltonian cycle. Then, in the union of F_{i_1} and F_{i_2}, after we delete all the edges that are incident to the two vertices ∞_1 and ∞_2, no cycle can exist. In addition, there also does not exist a path whose terminal vertices are the two vertices i_1 and i_2, otherwise the union of the path and the two edges $\{i_1, \infty_1\}$ (contained in F_{i_1}) and $\{i_2, \infty_1\}$ (contained in F_{i_2}) can form a cycle that does not visit the vertex ∞_2, which conflicts with the fact that the union of the two one-factors F_{i_1} and F_{i_2} forms a Hamiltonian cycle. Thus, $\{C_0, C_1, \cdots, C_{2n-1}\}$ in Algorithm 1 meets Condition 2.1 in Section II-B.

To make Algorithm 1 more easily understood, we give an example of constructing the $\mathbb{C}_4$ instance in Section II-A from a bipyramidal P1F of a 4-regular graph on 6 vertices in Figure 1.

Then, the next natural question is: Can we get a bipyramidal P1F of a $2n$-regular graph on a set of $2n+2$ vertices from a

$\mathbb{C}_{2n}$ instance? A positive answer to this question will be given by Algorithm 2.

Algorithm 2 Constructing a bipyramidal P1F of a $2n$-regular graph on $2n+2$ vertices from a $\mathbb{C}_{2n}$ instance.

(S1) If a $\mathbb{C}_{2n}$ instance exists, describe the code using the graph representation mentioned in Section II-B and let C_i represent the i-th column of the code in the graph description, where $i = 0, 1, \dots, 2n-1$.

(S2) Add two vertices ∞_1 and ∞_2 to the $(2n-2)$-regular graph G of vertices $0, 1, \dots, 2n-1$.

(S3) For $i = 0, 1, \dots, 2n-1$, add two edges $\{i, \infty_1\}$ and $\{r_i, \infty_2\}$ to C_i, where r_i is an integer from 0 to $2n-1$ such that the expanded set $\widetilde{C}_i$ is a one-factor of the expanded graph $\widetilde{G}$ of vertices $0, 1, \cdots, 2n-1, \infty_1, \infty_2$.

In Algorithm 2, for $i = 0, 1, \cdots, 2n-1$, $\widetilde{C}_i$ has the following form:

$$\widetilde{C}_i = C_i \cup \{\{i, \infty_1\}, \{r_i, \infty_2\}\}. \tag{25}$$

It is clear that the new graph $\widetilde{G}$ is a $2n$-regular graph on $2n+2$ vertices.

For $i = 0, 1, \cdots, 2n-1$, define $\infty_1 + i = \infty_1$ and $\infty_2 + i = \infty_2$. Then, for $i = 0, 1, \cdots, 2n-1$, we have

$$\widetilde{C}_i = \left\{\{x + i \bmod 2n, y + i \bmod 2n\} : \{x, y\} \in \widetilde{C}_0\right\}. \tag{26}$$

Consequently,

$$\mathbb{F} = \left\{\widetilde{C}_0, \widetilde{C}_1, \cdots, \widetilde{C}_{2n-1}\right\} \tag{27}$$

is a bipyramidal one-factorization of a $2n$-regular graph on $2n+2$ vertices.

Take the $\mathbb{C}_4$ instance in Section II-A for example. Figure 1(b) shows the corresponding expanded graph $\widetilde{G}$, which is a 4-regular graph on a set of 6 vertices $\{0, 1, 2, 3, \infty_1, \infty_2\}$. The four corresponding expanded sets are

$$\begin{cases} \widetilde{C}_0 = \{\{1,2\}, \{0, \infty_1\}, \{3, \infty_2\}\}; \\ \widetilde{C}_1 = \{\{2,3\}, \{1, \infty_1\}, \{0, \infty_2\}\}; \\ \widetilde{C}_2 = \{\{3,0\}, \{2, \infty_1\}, \{1, \infty_2\}\}; \\ \widetilde{C}_3 = \{\{0,1\}, \{3, \infty_1\}, \{2, \infty_2\}\}. \end{cases}$$

It is clear that

$$\mathbb{F} = \left\{\widetilde{C}_0, \widetilde{C}_1, \widetilde{C}_2, \widetilde{C}_3\right\}$$

is a bipyramidal one-factorization of a 4-regular graph on 6 vertices.

We then prove that the bipyramidal one-factorization obtained in Algorithm 2 is perfect as follows:

From Condition 2.1 in Section II-B, we can deduce that for any m and k (where $0 \le m < k \le 2n-1$), the subgraph

$$G^* = (\{0, 1, \cdots, 2n-1\}, C_m \cup C_k)$$

can be in one of the following two forms:

1) G^* consists of an isolated vertex m (or k) and a path of length $2n-2$ one of whose terminal vertices is the other vertex k (or m); or
2) G^* consists of two paths that satisfy: i) the sum of their length is $2n-2$, and ii) one of the terminal vertices of each path is the vertex m or k.

Then, in the one-factorization constructed in Algorithm 2, the union of any pair of one-factors forms a Hamiltonian cycle. Thus, according to Definition 1.1 in Section I, the one-factorization obtained in Algorithm 2 is a bipyramidal P1F of a $2n$-regular graph on $2n+2$ vertices.

Appendix B
Proof of Theorem 3.1

It is clear that S consists of $n-1$ pairs of non-zero elements in $(Z_{2n}, +)$. We then prove by contradiction that every non-zero element except n occurs in Δ.

We first consider the first opposite case where n occurs in Δ. Suppose the corresponding pair is $\{x^*, y^*\}$, i.e.

$$y^* - x^* = x^* - y^* = n.$$

Then, we have

$$\{x^* + n, y^* + n\} = \{x^*, y^*\}.$$

$\mathbb{F}$ is a bipyramidal one-factorization in which

$$F_n = \{\{x + n, y + n\} : \{x, y\} \in F_0\}.$$

Consequently, $\{x^*, y^*\}$ is contained in both F_0 and F_n — a contradiction!

Now, under the condition $n \notin \Delta$, we then consider the second opposite case where a non-zero and non-n element does not occur in Δ. Then, according to the pigeonhole principle, there exist two pairs $\{x'_1, y'_1\}$ and $\{x'_2, y'_2\}$, which meet

$$y'_1 - x'_1 = y'_2 - x'_2.$$

Then, we have

$$x'_2 - x'_1 = y'_2 - y'_1.$$

Let

$$k = x'_2 - x'_1.$$

Then, we have

$$\{x'_1 + k, y'_1 + k\} = \{x'_2, y'_2\}.$$

$\mathbb{F}$ is a bipyramidal one-factorization in which

$$F_k = \{\{x + k, y + k\} : \{x, y\} \in F_0\}.$$

Consequently, $\{x'_2, y'_2\}$ is contained in both F_0 and F_k — a contradiction!

Therefore, according to Definition 3.1 in Section III, S is an even starter in $(Z_{2n}, +)$.

REFERENCES

[1] M. Li and J. Shu, "On cyclic lowest density MDS array codes constructed using starters," in *Proceedings of the 2010 IEEE International Symposium on Information Theory (ISIT '10)*, Austin, TX, Jun. 2010, pp. 1315–1319.

[2] M. Blaum, P. Farrell, and H. van Tilborg, "Array codes," in *Handbook of Coding Theory*, V. Pless and W. Huffman, Eds. Amsterdam, The Netherlands: Elsevier Science B.V., 1998, pp. 1805–1909.

[3] I. S. Reed and G. Solomon, "Polynomial codes over certain finite fields," *Journal of the Society for Industrial and Applied Mathematics*, vol. 8, no. 2, pp. 300–304, Jun. 1960.

[4] R. M. Roth and G. Seroussi, "On generator matrices of MDS codes," *IEEE Transactions on Information Theory*, vol. 31, no. 6, pp. 826–830, Nov. 1985.

[5] R. M. Roth and A. Lempel, "On MDS codes via Cauchy matrices," *IEEE Transactions on Information Theory*, vol. 35, no. 6, pp. 1314–1319, Nov. 1989.

[6] Y. Cassuto and J. Bruck, "Cyclic lowest density MDS array codes," *IEEE Transactions on Information Theory*, vol. 55, no. 4, pp. 1721–1729, Apr. 2009.

[7] F. J. MacWilliams and N. J. A. Sloane, *The Theory of Error-Correcting Codes*. New York: North-Holland, 1977.

[8] M. Li, J. Shu, and W. Zheng, "GRID Codes: Strip-based erasure codes with high fault tolerance for storage systems," *ACM Transactions on Storage*, vol. 4, no. 4, pp. 1–22, Jan. 2009.

[9] E. Pinheiro, W.-D. Weber, and L. A. Barroso, "Failure trends in a large disk drive population," in *Proceedings of the 5th USENIX conference on File and Storage Technologies (FAST'07)*, San Jose, CA, Feb. 2007, pp. 17–28.

[10] B. Schroeder and G. A. Gibson, "Disk failures in the real world: What does an MTTF of 1,000,000 hours mean to you?" in *Proceedings of the 5th USENIX conference on File and Storage Technologies (FAST'07)*, San Jose, CA, Feb. 2007, pp. 1–16.

[11] L. N. Bairavasundaram, G. R. Goodson, S. Pasupathy, and J. Schindler, "An analysis of latent sector errors in disk drives," in *Proceedings of the 2007 ACM SIGMETRICS International Conference on Measurement and Modeling of Computer Systems (SIGMETRICS'07)*, San Diego, CA, Jun. 2007, pp. 289–300.

[12] P. M. Chen, E. K. Lee, G. A. Gibson, R. H. Katz, and D. A. Patterson, "RAID: High-performance, reliable secondary storage," *ACM Computing Surveys*, vol. 26, no. 2, pp. 145–185, Jun. 1994.

[13] A. Thomasian and M. Blaum, "Higher reliability redundant disk arrays: Organization, operation, and coding," *ACM Transactions on Storage*, vol. 5, no. 3, pp. 1–59, Nov. 2009.

[14] L. Xu, V. Bohossian, J. Bruck, and D. G. Wagner, "Low-density MDS codes and factors of complete graphs," *IEEE Transactions on Information Theory*, vol. 45, no. 6, pp. 1817–1826, Sep. 1999.

[15] M. Blaum, J. Brady, J. Bruck, and J. Menon, "EVENODD: An efficient scheme for tolerating double disk failures in RAID architectures," *IEEE Transactions on Computers*, vol. 44, no. 2, pp. 192–202, Feb. 1995.

[16] P. Corbett, B. English, A. Goel, T. Grcanac, S. Kleiman, J. Leong, and S. Sankar, "Row-diagonal parity for double disk failure correction," in *Proceedings of the 3rd USENIX Conference on File and Storage Technologies (FAST'04)*, San Francisco, CA, Mar. 2004, pp. 1–14.

[17] J. S. Plank, "The RAID-6 Liberation Codes," in *Proceedings of the 6th USENIX Conference on File and Storage Technologies (FAST'08)*, San Jose, CA, Feb. 2008, pp. 97–110.

[18] L. Xu and J. Bruck, "X-Code: MDS array codes with optimal encoding," *IEEE Transactions on Information Theory*, vol. 45, no. 1, pp. 272–276, Jan. 1999.

[19] G. V. Zaitsev, V. A. Zinov'ev, and N. V. Semakov, "Minimum-check-density codes for correcting bytes of errors, erasures, or defects," *Problems of Information Transmission*, vol. 19, no. 3, pp. 197–204, 1983.

[20] C. Jin, H. Jiang, D. Feng, and L. Tian, "P-Code: A new RAID-6 code with optimal properties," in *Proceedings of the 23rd International Conference on Supercomputing (ICS'09)*, Yorktown Heights, NY, Jun. 2009, pp. 360–369.

[21] W. D. Wallis, *One-Factorizations*. Norwell, MA: Kluwer, 1997.

[22] D. A. Pike and N. Shalaby, "Non-isomorphic perfect one-factorizations from Skolem sequences and starters," *Journal of Combinatorial Mathematics and Combinatorial Computing*, vol. 44, pp. 23–32, 2003.

[23] E. S. Seah and D. R. Stinson, "A perfect one-factorization for K_{36}," *Discrete Mathematics*, vol. 70, no. 2, pp. 199–202, 1988.

[24] M. Kobayashi, H. Awoki, Y. Nakazaki, and G. Nakamura, "A perfect one-factorization of K_{36}," *Graphs and Combinatorics*, vol. 5, pp. 243–244, 1989.

[25] E. S. Seah and D. R. Stinson, "A perfect one-factorization for K_{40}," *Congressus Numerantium*, vol. 68, pp. 211–213, 1989.

[26] E. C. Ihrig, E. S. Seah, and D. R. Stinson, "A perfect one-factorization of K_{50}," *Journal of Combinatorial Mathematics and Combinatorial Computing*, vol. 1, pp. 217–219, 1987.

[27] A. J. Wolfe, "A perfect one-factorization of K_{52}," *Journal of Combinatorial Designs*, vol. 17, no. 2, pp. 190–196, Mar. 2009.

[28] D. Bryant, B. Maenhaut, and I. M. Wanless, "New families of atomic Latin squares and perfect 1-factorisations," *Journal of Combinatorial Theory, Series A*, vol. 113, no. 4, pp. 608–624, May 2006.

[29] J. H. Dinitz, "Starters," in *Handbook of Combinatorial Designs, Second Edition*, C. J. Colbourn and J. H. Dinitz, Eds. Boca Raton, FL: Chapman and Hall/CRC, Nov. 2006, pp. 622–628.

[30] B. A. Anderson, "Sequencings and starters," *Pacific Journal of Mathematics*, vol. 64, no. 1, pp. 17–24, 1976.

[31] ——, "Symmetry groups of some perfect 1-factorizations of complete graphs," *Discrete Mathematics*, vol. 18, no. 3, pp. 227–234, 1977.

[32] M. Kobayashi, "On perfect one-factorization of the complete graph K_{2p}," *Graphs and Combinatorics*, vol. 5, no. 1, pp. 351–353, 1989.

[33] M. Li and J. Shu, "On the equivalence between the B-Code constructions and perfect one-factorizations," in *Proceedings of the 2010 IEEE International Symposium on Information Theory (ISIT 2010)*, Austin, TX, Jun. 2010, pp. 993–996.

Mingqiang Li received his Ph.D. degree in Computer Science from Tsinghua University in July, 2011. He also received a B.S. degree in Mathematics from the University of Electronic Science and Technology of China in July, 2006. He is now a Staff Researcher in the IBM China Research Laboratory. His research interests include storage systems, wireless networking, coding theory, data compression, and distributed computing.

Jiwu Shu received his Ph.D. degree in Computer Science from Nanjing University in 1998. He finished his postdoctoral research at Tsinghua University in 2000. Since then, he has been working as a teacher at Tsinghua University. He is now a professor in the Department of Computer Science and Technology at Tsinghua University. His research interests include storage systems and parallel computing.