Cyclic Low-Density MDS Array Codes [1]

Yuval Cassuto
Department of Electrical Engineering
California Institute of Technology
Pasadena, CA 91125, U.S.A.
Email: ycassuto@paradise.caltech.edu

Jehoshua Bruck
Department of Electrical Engineering
California Institute of Technology
Pasadena, CA 91125, U.S.A.
Email: bruck@paradise.caltech.edu

***Abstract*— We construct two infinite families of low density MDS array codes which are also cyclic. One of these families includes the first such sub-family with redundancy parameter $r > 2$. The two constructions have different algebraic formulations, though they both have the same indirect structure. First MDS codes that are not cyclic are constructed and then by applying a certain mapping to their parity check matrices, non-equivalent cyclic codes with the same distance and density properties are obtained. Using the same proof techniques, a third infinite family of quasi-cyclic codes can be constructed.**

I. Introduction

MDS (maximum distance separable) codes over large symbol alphabets are ubiquitous in data storage applications. Being MDS, they offer the maximum protection against device failures for a given amount of redundancy. Array codes are one type of such codes that is very useful to dynamic high-speed storage applications as they enjoy low-complexity decoding algorithms over small fields as well as low update complexity when small changes are applied to the stored content. That is in contrast to the family of Reed-Solomon codes that in general has none of these favorable properties. Examples of constructions that yield array codes with these properties can be found in [1],[2],[3],[4],[5] and [6] (partial list). In this paper we wish to propose codes of this type that are also cyclic. As an example we examine the following code defined on a 2×6 array

a_0	a_1	a_2	a_3	a_4	a_5
a_2+a_3+a_4	a_3+a_4+a_5	a_4+a_5+a_0	a_5+a_0+a_1	a_0+a_1+a_2	a_1+a_2+a_3

This code has 6 information bits $a_0, \ldots, a_5$ all of which can be recovered from *any* set of 3 columns that in total have 6 bits. Hence the code is MDS. However, the focus of this paper is a different property of this sample code; its cyclicity. To convince oneself that the code is cyclic, we observe that all the indices in a column can be obtained by adding one (modulo 6) to the indices in the column to its (cyclic) left. Thus any shift of the information bits row results in an identical shift in the parity bits row.

Beyond theoretical interest, cyclic array codes whose other properties match those of their best-known non-cyclic counterparts, offer significant practical advantages. Using cyclic array codes for high-speed storage applications can reduce the implementation cost of the codes. This improved cost-performance is thanks to savings in time and space resources needed for encoding and decoding and also by allowing many useful operations on the array to be carried out using simple regular circuits. Another advantage is having a uniform design for the individual storage units that implement the code array. Examples of benefits of cyclic array codes are provided in section III. The previously only known family of low-density MDS cyclic codes is the code with redundancy $r = 2$ that was proposed in [6]. In this paper we present two families of cyclic MDS array codes. The first, given in section IV, is defined on arrays with dimensions $\frac{p-1}{r} \times (p-1)$, where r is the code redundancy and p is a prime. It includes the first family of cyclic MDS array codes with $r > 2$. Its construction builds upon a known construction for non-cyclic codes (first appeared in [1], later extended in [7], and in [5]) by first shortening the codes and then explicitly providing a class of mappings of parity check matrix locations from the original shortened codes to the cyclic codes. The second construction, given in section V, proposes a new non-cyclic code family on arrays with dimensions $(p-1) \times (p-1)$ and $r = 2$, then proves its MDS property and similarly shows how to map the codes to cyclic codes. The first part of that construction (the non-cyclic part) can be alternately presented using graph theoretic construction tools: perfect 1-factorizations of complete bi-partite graphs, and thus it is a generalization of the method proposed in [5] that uses factorizations of complete *uni*-partite graphs. Combining the proof techniques from the two codes constructed in this paper, a third family of lowest density *quasi*-cyclic MDS codes can be constructed. These codes have dimensions $(p-1) \times 2(p-1)$ and $r = 2$. Due to space limitations, we omit the presentation of these codes to allow a greater focus on the proof techniques. To make the constructions clearer, an example for each one is provided following its formal description.

II. Definitions

A *linear array code* $(\mathcal{C})_F$ of dimensions $b \times n$ over $F = F_q$ is a linear subspace of the vector space F^{nb}. The dual code $(\mathcal{C})_F^{\perp}$ is the null-space of $(\mathcal{C})_F$ over F. To define the minimum distance of an array code we regard it as a code over the alphabet F^b, where F^b denotes length b vectors over F. Then the minimum distance is simply the minimum Hamming distance of the length n code over F^b. Note that though the

[1]This work was supported in part by the Caltech Lee Center for Advanced Networking and by NSF grant ANI-0322475

code symbols can be regarded as elements in the finite field F_{q^b}, we do not assume linearity over this field. $(\mathcal{C})_F$ can be specified by either its Parity-check matrix H of size $N_p \times nb$ or its Generator matrix G of size $(nb - N_p) \times nb$. A Parity-check (or Generator) matrix is called *systematic* if it has N_p (or $nb - N_p$) not necessarily adjacent columns that when stacked together form the identity matrix I_{N_p} (or I_{nb-N_p}), respectively. Given a systematic H matrix or G matrix (one can be easily obtained from the other), the nb symbols of the $b \times n$ array can be partitioned into N_p parity symbols and $nb - N_p$ information symbols. Define the *density* of the code as the average number of non-zeros in a row of G, $\frac{N(G)}{nb-N_p}$, where $N(M)$ is the number of non-zeros in a matrix M. When H is systematic an alternative expression for the density is $1 + \frac{N(H)-N_p}{nb-N_p}$. We call a code $(\mathcal{C})_F$ *lowest density* if its density equals its minimum distance (the minimum distance is an obvious lower bound on the density [4]). We call a family of codes *low-density* if the density of the codes is $O(1)$. If $b|N_p$ and the minimum distance d equals $\frac{N_p}{b}+1$ then the code is called maximum distance separable (MDS) with redundancy $r = \frac{N_p}{b}$. Throughout the paper $[s,t]$ denotes the set $\{x \in \mathbb{Z} : \ s \le x \le t\}$. To simplify the presentation of the constructions in the paper, we introduce another structure that defines a code when, as is the situation here, the parity check matrix has elements in $\{0,1\}$. Given a parity check matrix H, define the *index array* A_H to be a $b \times n$ array of subsets of $[0, N_p - 1]$. Index arrays can be similarly defined for generator matrices as well, but these do not appear in this paper. h_l denotes the l^{th} column of H and $h_l(x)$ the x^{th} element of h_l, $x \in [0, N_p - 1]$. The set in location i, j of A_H contains the elements $\{x : \ h_{i+bj}(x) = 1\}$. If H is systematic, A_H has N_p subsets of size 1. Note that A_H has the same dimensions as the code array and its sets specify the encodings of the bits of $(\mathcal{C})_F^{\perp}$. As an example we take a $(n = 6, b = 3, N_p = 6)$ systematic code and provide in figure 1 a generator matrix G and a parity check matrix H with its index array A_H.

$$G = \left[\begin{array}{ccc|ccc|ccc|ccc|ccc|ccc}
0&1&0&0&0&0&0&0&0&0&0&0&1&0&0&1&0&0\\
0&0&1&1&0&0&0&0&0&1&0&0&0&0&0&0&0&0\\
1&0&0&0&1&0&0&0&0&0&0&0&0&0&0&1&0&0\\
0&0&0&0&0&1&1&0&0&0&0&0&1&0&0&0&0&0\\
1&0&0&1&0&0&0&1&0&0&0&0&0&0&0&0&0&0\\
0&0&0&0&0&0&0&0&1&1&0&0&0&0&0&1&0&0\\
0&0&0&1&0&0&1&0&0&0&1&0&0&0&0&0&0&0\\
1&0&0&0&0&0&0&0&0&0&0&1&1&0&0&0&0&0\\
0&0&0&0&0&0&1&0&0&1&0&0&0&1&0&0&0&0\\
0&0&0&1&0&0&0&0&0&0&0&0&0&0&1&1&0&0\\
0&0&0&0&0&0&0&0&0&1&0&0&1&0&0&0&1&0\\
1&0&0&0&0&0&1&0&0&0&0&0&0&0&0&0&0&1
\end{array}\right] \quad \begin{matrix}\uparrow\\ nb-N_p \\ \downarrow\end{matrix}$$

$$\longleftarrow \quad nb \quad \longrightarrow$$

$$H = \left[\begin{array}{ccc|ccc|ccc|ccc|ccc|ccc}
1&0&0&0&1&0&0&1&0&0&0&1&0&0&0&0&0&1\\
0&0&1&1&0&0&0&1&0&0&1&0&0&0&1&0&0&0\\
0&0&0&0&0&1&1&0&0&0&1&0&0&1&0&0&0&1\\
0&0&1&0&0&0&0&0&1&1&0&0&0&1&0&0&1&0\\
0&1&0&0&0&1&0&0&0&0&0&1&1&0&0&0&1&0\\
0&1&0&0&1&0&0&0&1&0&0&0&0&0&1&1&0&0
\end{array}\right] \quad \begin{matrix}\uparrow\\ N_p \\ \downarrow\end{matrix}$$

$$\longleftarrow \quad n \quad \longrightarrow$$

$$A_H = \left[\begin{array}{c|c|c|c|c|c}
0 & 1 & 2 & 3 & 4 & 5\\
4,5 & 5,0 & 0,1 & 1,2 & 2,3 & 3,4\\
1,3 & 2,4 & 3,5 & 4,0 & 5,1 & 0,2
\end{array}\right] \quad \begin{matrix}\uparrow\\ b \\ \downarrow\end{matrix}$$

Fig. 1. G,H and A_H for a sample $n = 6, b = 3, N_p = 6$ code

III. Cyclic array codes

The codes we hereafter construct are codes of length n over F^b which are cyclic but not linear. In this section we wish to discuss such codes in general, providing conditions for a code to be cyclic and highlighting the potential benefits of using them. One way to characterize cyclic array codes is as cyclic group codes over the direct-product group of the additive group of F. Another is to view them as length nb linear b-*quasi-cyclic* codes. For the most part, the latter view will prove more useful since the constructions below are not explicit group theoretic ones. In fact, the description of array codes using index arrays we chose here was used in [8] to describe quasi-cyclic code constructions. We start off with the basic definition of cyclic codes.

Definition 1: The code $\mathcal{C}$ over F^b is *cyclic* if $s = (s_0, s_1, \ldots, s_{n-1}) \in \mathcal{C} \ \Rightarrow \ s' = (s_{n-1}, s_0, \ldots, s_{n-2}) \in \mathcal{C}$ and $s_i \in F^b$.

The following proposition is a straight forward generalization of a folklore fact about linear cyclic codes. Let S be the matrix that cyclically right-shifts the columns of H, b times. If H is viewed as concatenation of n matrices of size $N_p \times b$, $H = [H_0 \mid H_1 \mid \cdots \mid H_{n-1}]$ then $HS = [H_{n-1} \mid H_0 \mid \cdots \mid H_{n-2}]$.

Proposition 1: 1) The code $(\mathcal{C})_F$ with a parity check matrix H is cyclic if and only if there exists an invertible $N_p \times N_p$ matrix L such that $HS = LH$.

2) The code $(\mathcal{C})_F$ with a generator matrix G is cyclic if and only if there exists an invertible $(nb-N_p) \times (nb-N_p)$ matrix L' such that $GS = L'G$.

Corollary 1: If $(\mathcal{C})_F$ is cyclic, so is $(\mathcal{C})_F^{\perp}$.

While proposition 1 is not unique to array codes, for a class of array codes that are soon defined, a stronger necessary and sufficient condition for cyclicity can be proved. A parity check matrix of a systematic array code is called *regular* if $n|N_p$ and the subset $J_l \subset [1, b]$ of the columns of H_l that represent parity bits is the same for every l. In figure 1 H is regular since the first column of every $N_p \times b$ sub-matrix corresponds to a parity bit. The vector that has one in location i and zeros else where is denoted by e_i.

Theorem 1: A code $(\mathcal{C})_F$ with a regular parity check matrix H is cyclic if and only if $HS = L_s H$ where L_s is the $N_p \times N_p$

permutation matrix that performs $\frac{N_p}{n}$ downward cyclic shift of the rows of H.

Corollary 2: A code given as a regular index array A_H is cyclic if and only if adding $\frac{N_p}{n}$ modulo N_p to the elements of the sets of A_H yields a cyclic shift of A_H.

A. Merits of cyclic codes

One dimensional linear cyclic codes are known to provide great advantages such as succinct representations and efficient encoding and decoding. Cyclic array codes carry similar advantages when used to protect stored data. A cyclic array code can be specified by providing only the $b - \frac{N_p}{n}$ subsets of the set $[0, N_p - 1]$, corresponding to the first column of the code array. This makes searching for cyclic codes a more tractable operation (the results of section IV provide an example where restricting a search to cyclic codes may not compromise the properties of the codes found). As for decoding, a cyclic code can provide a factor n savings in memory used to decode erasures. A generic (and sometimes the best known) way to decode erasures in array codes is storing a decoding matrix for every combination of r column erasures. This matrix is the inverse of the concatenation of the r sub-matrices (of size $N_p \times b$) of H that correspond to the erased columns. With a cyclic code every n such erasure combinations have cyclically equivalent decoding sub-matrices and need only a single stored decoding matrix. Many other operations can be carried out using simple circuits when the code is cyclic. As examples we can take syndrome computation for error decoding and update operations such as bit,row or column updates. Using cyclic codes can also prove practically appealing for storage applications since its symmetry allows the storage devices to have identical designs, compared to, in general, a specialized design for each unit depending on its index in the code word.

IV. Cyclic lowest-density MDS codes with $n = p - 1, b = \frac{p-1}{r}$

Let r be a divisor of $p - 1$, and p an odd prime. Let α be an element in F_p of order r and β be an element in F_p of order $p - 1$. α and β define a partition of F_p to cosets of its multiplicative subgroup of order r plus a set that contains only the zero element. Except for the zero set, all sets are of cardinality r and there are $\frac{p-1}{r}$ such sets.

$$C_{-1} = \{0\} \qquad C_i = \{\beta^i, \beta^i\alpha, \ldots, \beta^i\alpha^{r-1}\} \quad (1)$$

where $0 \leq i < \frac{p-1}{r}$. The sets C_i are used in [4] and [7] to construct (non-cyclic) lowest density MDS codes with redundancy r. The construction therein is a generalization of the $r = 2$ construction of [1]. In [7], this construction was proved to provide lowest density MDS codes for a wide range of parameters. When F has characteristic 2, MDS codes are obtained for $r = 3$ and $r = 4$ whenever 2 is primitive in F_p. For larger characteristics, codes with additional r values were shown to be MDS. For completeness we present a construction for codes which are shortened versions of the codes therein. Later in the section we show that for every code constructed in this manner, there exists a class of mappings from locations in the parity check matrix of the code to those of a different (non equivalent) code that is cyclic. Under these mappings, the new cyclic codes inherit the distance and density properties from [1] and [7], so the proposed codes enjoy the cyclicity property while not compromising the optimality of their ancestors. Better readability in mind and with a slight abuse of notation, operations on sets denote element-wise operations on the content of the sets. Specifically, if $\langle x + l \rangle_z$ is used to denote $x + l \pmod z$, then $\langle S + l \rangle_z$ denotes the set that is obtained by adding l to the elements of S modulo z; also denote $\langle S \rangle_z \triangleq \langle S + 0 \rangle_z$. Similarly, permutations and arithmetic operations on sets represent the corresponding operations on their elements. For every $0 \leq m < p - 1$ define $I_m = \{i : \forall x \in \langle C_i + m \rangle_p, 0 \leq x < p - 1\}$. It is obvious that for every m, $|I_m| = \frac{p-1}{r}$ since for every translation m of the sets C_i only one set contains the element $p - 1$. Denote the j^{th} element of I_m by $I_m(j)$, where indices in I_m are ordered lexicographically. The code $\mathcal{C}$ is defined via an index array A_H. The set at location $(j, m) \in [0, \frac{p-1}{r} - 1] \times [0, p - 2]$, in A_H is $\langle C_i + m \rangle_p$, $i = I_m(j)$. The code $\mathcal{C}$ is a shortened version of the code constructed in [4],[7]. Note that because of the restriction $i \in I_m$, Theorem 1 implies that $\mathcal{C}$ is *not* a cyclic code. To define the cyclic code $\mathcal{C}_\circ$ ($\circ$ for cyclic), we derive alternative constructing sets D_i from C_i, as described below. The permutation $\psi : [0, p-2] \to [0, p-2]$ is defined to be $\psi(x) = \beta^x - 1 \pmod p$. We also define the inverse of ψ, $\psi^{-1}(y) = \log_\beta(y+1)$. For $i \in I_0$, define $D_i = \psi^{-1}(C_i)$. $\mathcal{C}_\circ$ is similarly defined through its index array $A_{H_\circ}$ as follows. The j, l set of $A_{H_\circ}$, $(j, l) \in [0, \frac{p-1}{r} - 1] \times [0, p-2]$, is $\langle D_i + l \rangle_{p-1}$, and now $i = I_0(j)$ (cf. $i = I_m(j)$ in the definition of $\mathcal{C}$). The fact that for every l translations of the same sets D_i are taken, together with Corollary 2 proves the following proposition.

Proposition 2: The code $\mathcal{C}_\circ$ is cyclic.

Theorem 2: $\mathcal{C}_\circ$ and $\mathcal{C}$ have the same redundancy, minimum distance and density.

proof: We show that $\mathcal{C}$ can be obtained from $\mathcal{C}_\circ$ (and also vice versa) by a distance-preserving bijection between rows and columns of $H_\circ$ to those of H (in array codes column permutations of the parity check matrix, in general, do not preserve distance). To refer to an element x in the set at location (j, l) in an index array A_M we use the tuple (x, j, l, M). The aforementioned bijection is given by showing that A_H is obtained from $A_{H_\circ}$ by mapping $(x, j, l, H_\circ) \leftrightarrow (\psi(x), j', m, H)$. The mapping $x \leftrightarrow \psi(x)$ represents permuting the rows of the parity check matrix and the mapping $(j, l) \leftrightarrow (j', m)$ represents permuting columns of the parity check matrix. The mapping $(j, l) \leftrightarrow (j', m)$ has a special property that it only reorders columns of the index array and reorders sets *within* its columns. Hence the mapping $(x, j, l, H_\circ) \leftrightarrow (\psi(x), j', m, H)$ preserves both the minimum distance and density of the code. More concretely, we need to show that for every $l \in [0, p-2]$ there exists an $m \in [0, p-2]$ such that every $i = I_0(j)$ has a corresponding $t = I_m(j')$ that together satisfy

$$\psi[\langle D_i + l \rangle_{p-1}] = \langle C_t + m \rangle_p$$

Since $\langle D_{-1}+l\rangle_{p-1}$ consists of the single element l and $\langle C_{-1}+m\rangle_p$ consists of the single element m, the integers l and m have to satisfy $m=\psi(l)$. Then we rewrite the above condition as

$$\psi[\langle D_i+l\rangle_{p-1}]=\langle C_t+\psi(l)\rangle_p$$

$$\psi[\langle D_i+l\rangle_{p-1}]=\psi[\langle\psi^{-1}[C_i]+l\rangle_{p-1}]=\langle\beta^{\log_\beta(\langle C_i+1\rangle_{p-1})+l}-1\rangle_p$$
$$=\langle\beta^l C_i+\beta^l-1\rangle_p=\langle C_{\langle i+l\rangle_{\frac{p-1}{r}}}+\psi(l)\rangle_p$$

□

A. Example: cyclic MDS code with p=7, n=6, b=3, r=2

In F_7 pick $\alpha=6$, $\beta=3$ that satisfy $\mathrm{ord}(\alpha)=r=2$, $\mathrm{ord}(\beta)=p-1=6$. These α and β define the following partition of F_7 into sets C_i according to (1).

$$C_{-1}=\{0\},\ C_0=\{1,6\},\ C_1=\{3,4\},\ C_2=\{2,5\}$$

Taking the sets $\langle C_i+m\rangle_7$ to be the sets of A_H in column m, leaving out the particular set in that column that contains the element 6, we get

$$A_H=\begin{array}{|c|c|c|c|c|c|}\hline 0&1&2&3&4&5\\\hline 3,4&0,2&1,3&4,2&5,3&2,1\\\hline 2,5&5,4&4,0&1,5&0,1&3,0\\\hline\end{array}$$

The permutations ψ and ψ^{-1} written explicitly are $[0,1,2,3,4,5]\overset{\psi}{\to}[0,2,1,5,3,4]$ and $[0,1,2,3,4,5]\overset{\psi^{-1}}{\to}[0,2,1,4,5,3]$. ψ^{-1} acting on the array A_H yields

$$\psi^{-1}(A_H)=\begin{array}{|c|c|c|c|c|c|}\hline 0&2&1&4&5&3\\\hline 4,5&0,1&2,4&5,1&3,4&1,2\\\hline 1,3&3,5&5,0&2,3&0,2&4,0\\\hline\end{array}$$

which after reordering of columns and sets within columns results in the cyclic code generated by $\langle D_i+l\rangle_6$, $i\in I_0=\{-1,1,2\}$.

$$A_{H_\circ}=\begin{array}{|c|c|c|c|c|c|}\hline 0&1&2&3&4&5\\\hline 4,5&5,0&0,1&1,2&2,3&3,4\\\hline 1,3&2,4&3,5&4,0&5,1&0,2\\\hline\end{array}$$

V. Cyclic low-density MDS codes with $n=p-1, b=p-1, r=2$

Let p be an odd prime. Define the *ordered pairs* C_i, $i\in[0,p-2]$ to be

$$C_i=(a_i,c_{p-1-i}) \tag{2}$$

For an ordered pair $P=(a_j,c_m)$ define $P^{(l)}=(a_j,c_{\langle m+l\rangle_p})$. The pairs $C_i^{(l)}$ are first used to define a non-systematic, non-cyclic MDS code $\mathcal{B}$. To define the code $\mathcal{B}$, we again use the index array A_H, but now, for convenience of presentation, the set $[0,N_p-1]=\{0,1,\ldots,2(p-1)-1\}$ is mapped to the set $\{a_0,c_0,a_1,c_1,\ldots,a_{p-2},c_{p-2}\}$. The set at location $(0,l)$, $l\in[0,p-2]$, in A_H is $\{a_l\}$ and the set at location $(j,l)\in[1,p-2]\times[0,p-2]$, is $\{C^{(l)}_{\langle j+l\rangle_{p-1}}\}$. Note that since the index of c in $C_i^{(l)}$ is incremented modulo p while the index of a is incremented modulo $p-1$, $\mathcal{B}$ is not cyclic. To prove that $\mathcal{B}$ is MDS (has minimum column weight 3) we show that for every l_1,l_2, the matrix that is obtained by the juxtaposition $[H_{l_1}\mid H_{l_2}]$ is nonsingular (hence there are no weight 2 words in the code). As a reminder, H_{l_i} is the $N_p\times b=2(p-1)\times(p-1)$ matrix that corresponds to column l_i of the index array A_H, in the way described in section II. To do that we resort to some additional definitions.

Let E be a $m\times m$ matrix with entries $e_{ij}\in\{0,1\}$. Define a *singleton* column to be a column with a single 1. Two rows i_1,i_2 are called $2-$*connected* if there is a column with exactly two 1s in locations i_1,i_2. Define a *protruding* row to be a row with a 1 in a singleton column. A row is called *noble*, for reasons that will soon become clear, if it is protruding or if it is $2-$connected to another noble row. It is not hard to see that if a set of row vectors in E is linearly dependent it does *not* include any noble rows. Using the above, the following lemma is fairly easy to prove.

Lemma 1: If in a matrix E all of the rows are noble, then E is nonsingular.

Now take $E=[H_{l_1}\mid H_{l_2}]$. E has two singleton columns (and corresponding two protruding rows) defined by the sets $\{a_{l_1}\},\{a_{l_2}\}$, and $2(p-2)$ $2-$connections defined by the sets $\{C_i^{(l_1)}\},\{C_i^{(l_2)}\}$. We want to show that all the rows of E are noble by arguing that each row $a_0,c_0,a_1,c_1,\ldots,a_{p-2},c_{p-2}$ is connected to a protruding row by a chain of $2-$connections. Given the two columns $l_1=l,l_2=l+r$, $1\le r<p-1$, the chains of noble rows starting from a_l,a_{l+r} are, respectively

$$a_l\to c_{r-1}\to a_{l-r}\to c_{2r-1}\to a_{l-2r}\to\cdots\to c_{tr-1}$$

$$a_{l+r}\to c_{-r-1}\to a_{l+2r}\to c_{-2r-1}\to a_{l+3r}\to\cdots\to c_{-sr-1}$$

This follows since by construction (2), in columns l and $l+r$ the a and c indices sum to $l-1$ and $l+r-1$, respectively. If $s+t<p$ then the chains are disjoint and neither include repetitions. Otherwise there would be $s',t':s'+t'<p$ such that $t'r-1=-s'r-1\pmod p$ or $l-t'r=l+s'r\pmod p$ and both are impossible since $\gcd(r,p)=1$. Therefore, the first and second chains necessarily terminate when $tr-1=l\pmod p$ and $-sr-1=l+r\pmod p$ respectively. These correspond to c_l and c_{l+r} that are absent from columns l and $l+r$ respectively. To show that all $2(p-1)$ row indices appear in the chains we subtract the two equations and get $(t+s+1)r=0\pmod p$ and thus $t+s=p-1$. We conclude that all rows are noble and $E=[H_{l_1}\mid H_{l_2}]$ is non-singular for every l_1,l_2. □

A. Cyclic code

Let β be a primitive element in F_p. The permutation $\psi:[0,p-2]\to[0,p-2]$ is once again defined to be $\psi(x)=\beta^x-1\pmod p$. The inverse permutation ψ^{-1} is then $\psi^{-1}(y)=\log_\beta(y+1)$. For notational convenience we use $\phi(a_i)$,$\phi(c_i)$ to denote $a_{\phi(i)}$, $c_{\phi(i)}$ respectively, where ϕ is an arbitrary permutation and also a_i+l,c_i+l for a_{i+l},c_{i+l} respectively. Define the ordered pairs D_j, $j\in[1,p-2]$ to be

$$D_j=\left(\psi^{-1}(a_j),\psi^{-1}(c_{p-1-j})\right) \tag{3}$$

The code $\mathcal{B}_\circ$ has an index array $A_{H_\circ}$ whose set in location $(0,l)$, $l \in [0,p-2]$ is $\{a_l\}$ and the set in location $(j,l) \in [1,p-2]\times[0,p-2]$ is $\langle D_j+l\rangle_{p-1}$. To use Corollary 2 to prove the cyclicity of $\mathcal{B}_\circ$, we map $\{a_0,c_0,a_1,c_1,\ldots,a_{p-2},c_{p-2}\}$ back to $[0,2p-3]$ and observe that adding 2 modulo $2(p-1)$ to the sets mapped from $\langle D_j+l\rangle_{p-1}$ yields a cyclic shift of $A_{H_\circ}$.

Theorem 3: $\mathcal{B}_\circ$ and $\mathcal{B}$ have the same redundancy, minimum distance and density.

proof: We show, similarly to the proof of Theorem 2, that $\mathcal{B}$ can be obtained from $\mathcal{B}_\circ$ (and also vice versa) by a distance-preserving bijection between rows and columns of $H_\circ$ and those of H. Specifically, we prove the claim that for every $j \in [1,p-2]$ there exists a $i \in [0,p-2]$ such that $\psi(\langle D_j+l\rangle_{p-1}) = C_i^{(\psi(l))}$. Note that as mentioned before, the a indices and c indices of $C_i^{(m)}$ sum to $m-1 \pmod p$. The proof simplifies thanks to the following two observations. First, the elements of $\langle D_j+l\rangle_{p-1}$ are distinct for different j. Second, there are exactly $p-2$ pairs $s,t \in [0,p-2]$ that give $s+t = m-1 \pmod p$ for every $m \in [0,p-2]$ (this is not true for $m=p-1$ where we have $p-1$ such pairs). Consequently, proving that the a and c indices of $\psi(\langle D_j+l\rangle_{p-1})$, sum to $\psi(l)-1$ for each j establishes that these sets are indeed $C_i^{(\psi(l))}$. So proving the following suffices.

$$\psi(\psi^{-1}(j)+l)+\psi(\psi^{-1}(p-1-j)+l) = \psi(l)-1 \pmod p$$

Proving the above is then straight forward

$$\psi(\psi^{-1}(j)+l)+\psi(\psi^{-1}(p-1-j)+l) =$$

$$= \beta^{\log_\beta(j+1)+l} - 1 + \beta^{\log_\beta(p-1-j+1)+l} - 1 =$$

$$= \beta^l(j+1+p-j) - 2 = \beta^l - 1 - 1 = \psi(l) - 1$$

□

B. Systematic parity-check matrix

The parity check matrix obtained for $\mathcal{B}_\circ$ (and similarly for $\mathcal{B}$) earlier in the section is not systematic. The bits c_i do not appear in it as singleton columns. Non-systematic parity check matrices are undesirable since they do not offer the simple encoding allowed by systematic ones. Moreover, when the parity check matrix is systematic, one can easily use it to extract the density of the code. We now derive a systematic, cyclic parity check matrix from the non-systematic cyclic one of section V-A by setting $c_i = c_i - a_i$ for $i \in [0,p-2]$. Since the a and c indices of $C_i^{(m)}$ sum to $m-1 \in \{p-1\}\cup[0,p-3]$, for every m there exists a pair $C_i^{(m)}$ of the form $C_i^{(m)} = (a_{(m-1)/2}, c_{(m-1)/2})$. The sets D_i are obtained from C_i by permutation so for each m, one of the sets $\langle D_j+m\rangle_{p-1}$ is of the form $(a_{\psi^{-1}((m-1)/2)}, c_{\psi^{-1}((m-1)/2)})$. Thus the transformation $c_i = c_i - a_i$ makes the parity check matrix systematic. The number of ones in each column of the non-systematic part of the modified parity check matrix is 3, therefore the density is 4.

C. Example: cyclic MDS code with p=5, n=4, b=4, r=2

For $p=5$ the sets C_i are

$$C_0 = (a_0,c_4),\ C_1 = (a_1,c_3),\ C_2 = (a_2,c_2),\ C_3 = (a_3,c_1)$$

Note that the indices of a and c in each of the sets C_i sum to $p-1=4$. Taking the set $C^{(l)}_{\langle j+l\rangle_4}$ to be the (j,l) set of A_H,

$$A_H = \begin{array}{|c|c|c|c|}\hline a_0 & a_1 & a_2 & a_3 \\ \hline a_1,c_3 & a_2,c_3 & a_3,c_3 & a_0,c_2 \\ \hline a_2,c_2 & a_3,c_2 & a_0,c_1 & a_1,c_1 \\ \hline a_3,c_1 & a_0,c_0 & a_1,c_0 & a_2,c_0 \\ \hline \end{array}$$

Now pick a primitive element in F_5 $\beta = 2$ and use it to define the permutations ψ and ψ^{-1}, $[0,1,2,3] \stackrel{\psi}{\to} [0,1,3,2]$, $[0,1,2,3] \stackrel{\psi^{-1}}{\to} [0,1,3,2]$. The permutation ψ^{-1} acting on the sets of A_H yields

$$\psi^{-1}(A_H) = \begin{array}{|c|c|c|c|}\hline a_0 & a_1 & a_3 & a_2 \\ \hline a_1,c_2 & a_3,c_2 & a_2,c_2 & a_0,c_3 \\ \hline a_3,c_3 & a_2,c_3 & a_0,c_1 & a_1,c_1 \\ \hline a_2,c_1 & a_0,c_0 & a_1,c_0 & a_3,c_0 \\ \hline \end{array}$$

which after reordering of columns and sets within columns results in the cyclic code generated by $\langle D_j+l\rangle_4$,

$$A_{H_\circ} = \begin{array}{|c|c|c|c|}\hline a_0 & a_1 & a_2 & a_3 \\ \hline a_1,c_2 & a_2,c_3 & a_3,c_0 & a_0,c_1 \\ \hline a_3,c_3 & a_0,c_0 & a_1,c_1 & a_2,c_2 \\ \hline a_2,c_1 & a_3,c_2 & a_0,c_3 & a_1,c_0 \\ \hline \end{array}$$

VI. Conclusion

An important artifact of the two code families presented here is that they take their MDS property from their non-cyclic peers. A direct proof of their MDS property is still missing, and if found, it may enable the construction of new families of cyclic codes. This optimistic view is supported by computer searches that reveal MDS lowest-density cyclic codes with parameters that are not covered by the known families of non-cyclic codes.

References

[1] G.V Zaitsev, V.A Zinov'ev, and N.V Semakov. Minimum-check-density codes for correcting bytes of errors, erasures, or defects. *Problems Inform. Transm.*, 19:197–204, 1981.

[2] 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*, 44(2):192–202, 1995.

[3] M. Blaum, J. Bruck, and A. Vardy. MDS array codes with independent parity symbols. *IEEE-Trans-IT*, 42(2):529–542, 1996.

[4] M. Blaum and R.M Roth. On lowest density MDS codes. *IEEE-Trans-IT*, 45(1):46–59, 1999.

[5] L. Xu, V. Bohossian, J. Bruck, and D.G Wagner. Low-density MDS codes and factors of complete graphs. *IEEE-Trans-IT*, 45(6):1817–1826, 1999.

[6] L. Xu and J. Bruck. X-code: MDS array codes with optimal encoding. *IEEE-Trans-IT*, 45(1):272–276, 1999.

[7] E. Louidor and R.M Roth. Lowest-density MDS codes over extension alphabets. *Technion CS Technical report*, available at http://www.cs.technion.ac.il/users/wwwb/cgi-bin/tr-get.cgi/2005/CS/CS-2005-09.pdf.

[8] R.L Townsend and E.J Weldon Jr. Self-orthogonal quasi-cyclic codes. *IEEE-Trans-IT*, 13(2):183–195, 1967.