Partial MDS (PMDS) and Sector-Disk (SD) Codes that Tolerate the Erasure of Two Random Sectors

Mario Blaum[*], **James S. Plank**[†], **Moshe Schwartz**[‡], and **Eitan Yaakobi**[§]
[*]IBM Research Division, Almaden Research Center, San Jose, CA 95120, USA
[†]Dept. of Electrical Engineering and Computer Science, Univ. of Tennessee, Knoxville, TN 37996, USA
[‡]Department of Electrical and Computer Engineering, Ben-Gurion University, Beer Sheva 8410501, Israel
[§] Computer Science Department, Technion – Israel Institute of Technology, Haifa 32000, Israel
mblaum@us.ibm.com, plank@cs.utk.edu, schwartz@ee.bgu.ac.il, yaakobi@cs.technion.ac.il

***Abstract*—Partial MDS (PMDS) codes are erasure codes combining local (row) correction with global additional correction of entries, while Sector-Disk (SD) codes are erasure codes that address the mixed failure mode of current RAID systems. It has been an open problem to construct general codes that have the PMDS and the SD properties, and previous work has relied on Monte-Carlo searches. In this paper, we present a general construction that addresses the case of any number of failed disks and in addition, two erased sectors. The construction requires a modest field size. This result generalizes previous constructions extending RAID 5 and RAID 6.**

I. Introduction

Consider an $r \times n$ array whose entries are elements in a finite field $GF(2^w)$ [9] (in general, we could consider a field $GF(p^w)$, p a prime number, but for simplicity, we constrain ourselves to binary fields). The array may correspond to a stripe on a disk system, where elements co-located in the same column reside on the same disk, or the elements may correspond to disk or SSD blocks on a large storage system. Normally, these arrays are protected using the well known architectures known as Redundant Arrays of Independent Disks (RAID) [5]. Recent work has explored two types of erasure codes that extend RAID and are tailored for these scenarios: Partial-MDS (PMDS) codes and Sector-Disk (SD) codes [2], [3], [11], [12].

Both follow the same methodology — m entire columns of elements are devoted to coding, and each row composes an $[n, n-m, m+1]$ MDS code. In the remaining $n-m$ columns of the array, s more elements are also devoted to coding. The erasure protection that they provide differentiates PMDS and SD codes. SD codes tolerate the erasure of any m columns of elements, plus any additional s elements in the array. PMDS codes tolerate a broader class of erasures — any m elements per row, plus any additional s elements.

As their name implies, SD codes address the combination of disk and sector failures that occurs in modern disk systems. Column failures occur when entire disks break, and sector failures can accumulate over time, typically unnoticed until an entire disk breaks, and the failed sector is required for recovery. PMDS codes are maximally recoverable for codes laid out in the manner described above [2]. Maximally recoverable codes have been applied to cloud storage systems where each element resides on a different storage node [7]. The rows of the array correspond to collections of storage nodes that can decode together with good performance, while the extra s elements allow the system to tolerate broader classes of failures.

We label the codes with $(m;s)$, and illustrate the difference between PMDS and SD in Figure 1. The figure depicts five failure scenarios in a 4×5 array, encoded with a $(1;2)$ code, where erased elements are shaded in gray. The left scenario may be tolerated by both PMDS and SD codes, since each row is an $[5,4,2]$ MDS code. The second two scenarios are also tolerated by both PMDS and SD codes, because four erasures are co-located in the same column. The last of these is an important case, as it is not tolerated by RAID-6, even though RAID-6 devotes two full columns to coding. The two right scenarios are PMDS only, as there is one erasure per row, plus two additional erasures.

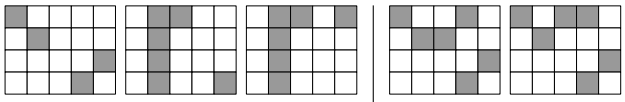

Fig. 1. Five failure scenarios on a 4×5 array of elements.

The challenge of the current work is to define PMDS and SD codes for general parameters. The case of $(m;1)$ PMDS codes was solved in [2]. In this paper, we address the case of $(m;2)$ PMDS and SD codes.

We begin with a formal definition of the two codes.

Definition 1.1: Let $\mathcal{C}$ be a linear $[rn, r(n-m)-s]$ code over a field such that when codewords are taken row-wise as $r \times n$ arrays, each row belongs in an $[n, n-m, m+1]$ MDS code. In other words, each row can correct m erasures independently.

1) *$\mathcal{C}$ is an $(m;s)$ partial-MDS (PMDS) code if,* for any $(s_1, s_2, \ldots, s_t)$ *such that each* $s_j \geqslant 1$ *and* $\sum_{j=1}^{t} s_j = s$, *and for any* $i_1, i_2, \ldots, i_t$ *such that* $0 \leqslant i_1 < i_2 < \cdots < i_t \leqslant r-1$, *$\mathcal{C}$ can correct up to $s_j + m$ erasures in each row i_j, $1 \leqslant j \leqslant t$, of an array in $\mathcal{C}$.*
2) *$\mathcal{C}$ is an $(m;s)$ sector-disk (SD) code if, for any* $l_1, l_2, \ldots, l_m$ *such that* $0 \leqslant l_1 < l_2 < \cdots < l_m \leqslant n-1$, *for any* $(s_1, s_2, \ldots, s_t)$ *such that each* $s_j \geqslant 1$ *and* $\sum_{j=1}^{t} s_j = s$, *and for any* $i_1, i_2, \ldots, i_t$ *such that* $0 \leqslant i_1 < i_2 < \cdots < i_t \leqslant r-1$, *$\mathcal{C}$ can correct up to $s_j + m$ erasures in each row i_j, $1 \leqslant j \leqslant t$, of an array in $\mathcal{C}$ provided that locations $l_1, l_2, \ldots l_m$ in each of the rows i_j have been erased.*

In the next section we give a general construction for $(m;2)$ PMDS and SD codes. Constructions of (1;2) SD codes were given in [1] and of (2;2) codes in [3]. These constructions

978-1-4799-5186-4/14/$31.00 ©2014 IEEE

are also summarized in [11] and the construction of (3;2) SD codes was verified for all r,n in $GF(2^8)$ and for $r,n \leqslant 24$ in $GF(2^{16})$. Hence, our results generalize those constructions. We note that we can use an MDS code (like a RS code for example) over the entire array. This will work for the purpose of correcting the maximum number of erasures in the array, but it does not guarantee the first property of PMDS or SD codes, namely that each row belongs in an $[n, n-m, m+1]$ MDS code. As such, it would be much more expensive computationally to encode and decode.

From now on, when we say PMDS or SD codes, we refer to $(m;2)$ PMDS or SD codes.

II. Code Construction

Consider the field $GF(2^w)$ and let α be an element in $GF(2^w)$. The (multiplicative) order of α, denoted $\mathcal{O}(\alpha)$, is the minimum $\ell > 0$ such that $\alpha^\ell = 1$. If α is a primitive element [9], then $\mathcal{O}(\alpha) = 2^w - 1$. To each element $\alpha \in GF(2^w)$, there is an associated (irreducible) minimal polynomial [9] that we denote $f_\alpha(x)$.

Let $\alpha \in GF(2^w)$ and $rn \leqslant \mathcal{O}(\alpha)$. We want to construct an SD-code consisting of $r \times n$ arrays over $GF(2^w)$, such that m of the columns correspond to parity (in RAID 5, $m=1$, while in RAID 6, $m=2$). In addition, two extra symbols also correspond to parity. When read row-wise, the codewords belong in an $[rn, r(n-m)-2]$ code over $GF(2^w)$. Specifically, let $\mathcal{C}(r,n,m,2;f_\alpha(x))$ be the $[rn, r(n-m)-2]$ code whose $(mr+2) \times rn$ parity-check matrix is given by

$$H = \left(\begin{array}{c|c|c|c} H_0 & \underline{0} & \dots & \underline{0} \\ \hline \underline{0} & H_0 & \dots & \underline{0} \\ \hline \vdots & \vdots & \ddots & \vdots \\ \hline \underline{0} & \underline{0} & \dots & H_0 \\ \hline H_1 & H_2 & \dots & H_r \end{array}\right) \quad (1)$$

where

$$H_0 = \begin{pmatrix} 1 & 1 & 1 & \dots & 1 \\ 1 & \alpha & \alpha^2 & \dots & \alpha^{n-1} \\ 1 & \alpha^2 & \alpha^4 & \dots & \alpha^{2(n-1)} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & \alpha^{m-1} & \alpha^{2(m-1)} & \dots & \alpha^{(m-1)(n-1)} \end{pmatrix} \quad (2)$$

and, for $1 \leqslant j \leqslant r$,

$$H_j = \begin{pmatrix} 1 & \alpha^m & \alpha^{2m} & \dots & \alpha^{m(n-1)} \\ \alpha^{(j-1)n} & \alpha^{(j-1)n-1} & \alpha^{(j-1)n-2} & \dots & \alpha^{(j-1)n-(n-1)} \end{pmatrix}. \quad (3)$$

We will show under which conditions codes $\mathcal{C}(r,n,m,2;f_\alpha(x))$ are PMDS or SD. Unless stated otherwise, for simplicity, let us denote $\mathcal{C}(r,n,m,2;f_\alpha(x))$ by $\mathcal{C}(r,n,m,2)$.

We start by giving some examples.

Example 2.1: Consider the finite field $GF(16)$ and let α be a primitive element, i.e., $\mathcal{O}(\alpha)=15$. Then, the parity-check matrix of $\mathcal{C}(3,5,1,2)$ is given by

$$\left(\begin{array}{ccccc|ccccc|ccccc} 1 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 \\ \hline 1 & \alpha & \alpha^2 & \alpha^3 & \alpha^4 & 1 & \alpha & \alpha^2 & \alpha^3 & \alpha^4 & 1 & \alpha & \alpha^2 & \alpha^3 & \alpha^4 \\ 1 & \alpha^{14} & \alpha^{13} & \alpha^{12} & \alpha^{11} & \alpha^{10} & \alpha^9 & \alpha^8 & \alpha^7 & \alpha^6 & \alpha^5 & \alpha^4 & \alpha^3 & \alpha^2 & \alpha \end{array}\right).$$

Similarly, the parity-check matrix of $\mathcal{C}(3,5,2,2)$ is given by

$$\left(\begin{array}{ccccc|ccccc|ccccc} 1 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & \alpha & \alpha^2 & \alpha^3 & \alpha^4 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & \alpha & \alpha^2 & \alpha^3 & \alpha^4 & 0 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & \alpha & \alpha^2 & \alpha^3 & \alpha^4 \\ \hline 1 & \alpha^2 & \alpha^4 & \alpha^6 & \alpha^8 & 1 & \alpha^2 & \alpha^4 & \alpha^6 & \alpha^8 & 1 & \alpha^2 & \alpha^4 & \alpha^6 & \alpha^8 \\ 1 & \alpha^{14} & \alpha^{13} & \alpha^{12} & \alpha^{11} & \alpha^{10} & \alpha^9 & \alpha^8 & \alpha^7 & \alpha^6 & \alpha^5 & \alpha^4 & \alpha^3 & \alpha^2 & \alpha \end{array}\right).$$

Let us point out that the construction of this type of codes is valid also over the ring of polynomials modulo $M_p(x) = 1 + x + \dots + x^{p-1}$, p a prime number, as done with the Blaum-Roth (BR) codes [4]. In that case, $\mathcal{O}(\alpha) = p$, where $\alpha^{p-1} = 1 + \alpha + \dots + \alpha^{p-2}$. The construction proceeds similarly, and we denote it $\mathcal{C}(r,n,m,2;M_p(x))$. Utilizing the ring modulo $M_p(x)$ allows for XOR operations at the encoding and the decoding without look-up tables in a finite field, which is advantageous in erasure decoding [4]. It is well known that $M_p(x)$ is irreducible if and only if 2 is primitive in $GF(p)$ [9].

Next we give a lemma that is key to proving the conditions under which codes $\mathcal{C}(r,n,m,2)$ are PMDS or SD.

Lemma 2.1: Let $\alpha \in GF(2^w)$, $rn \leqslant \mathcal{O}(\alpha)$, $1 \leqslant \ell \leqslant r-1$, and, if $1 \leqslant m \leqslant n-2$, let $0 \leqslant i_0 < i_1 < i_2 < \dots < i_m \leqslant n-1$ and $0 \leqslant j_0 < j_1 < j_2 < \dots < j_m \leqslant n-1$. Consider the $(2m+2) \times (2m+2)$ matrix $M(i_0,i_1,\dots,i_m;j_0,j_1,\dots,j_m;r;n;\ell)$ given by

$$\left(\begin{array}{cccc|cccc} 1 & 1 & \dots & 1 & 0 & 0 & \dots & 0 \\ \alpha^{i_0} & \alpha^{i_1} & \dots & \alpha^{i_m} & 0 & 0 & \dots & 0 \\ \alpha^{2i_0} & \alpha^{2i_1} & \dots & \alpha^{2i_m} & 0 & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ \alpha^{(m-1)i_0} & \alpha^{(m-1)i_1} & \dots & \alpha^{(m-1)i_m} & 0 & 0 & \dots & 0 \\ \hline 0 & 0 & \dots & 0 & 1 & 1 & \dots & 1 \\ 0 & 0 & \dots & 0 & \alpha^{j_0} & \alpha^{j_1} & \dots & \alpha^{j_m} \\ 0 & 0 & \dots & 0 & \alpha^{2j_0} & \alpha^{2j_1} & \dots & \alpha^{2j_m} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & \alpha^{(m-1)j_0} & \alpha^{(m-1)j_1} & \dots & \alpha^{(m-1)j_m} \\ \hline \alpha^{mi_0} & \alpha^{mi_1} & \dots & \alpha^{mi_m} & \alpha^{mj_0} & \alpha^{mj_1} & \dots & \alpha^{mj_m} \\ \alpha^{-i_0} & \alpha^{-i_1} & \dots & \alpha^{-i_m} & \alpha^{-n\ell-j_0} & \alpha^{-n\ell-j_1} & \dots & \alpha^{-n\ell-j_m} \end{array}\right)$$

Let

$$\Delta(i_0,i_1,\dots,i_m;j_0,j_1,\dots,j_m;r;n;\ell) = \det M(i_0,i_1,\dots,i_m;j_0,j_1,\dots,j_m;r;n;\ell).$$

Then,

$$\Delta(i_0,i_1,\dots,i_m;j_0,j_1,\dots,j_m;r;n;s;\ell) = \left(\prod_{0 \leqslant u < v \leqslant m} \left(\alpha^{i_u} \oplus \alpha^{i_v}\right)\left(\alpha^{j_u} \oplus \alpha^{j_v}\right)\right)\left(\alpha^{-\sum_{u=0}^{m} i_u} \oplus \alpha^{-n\ell-\sum_{u=0}^{m} j_u}\right).^1 \quad (4)$$

Proof: For simplicity, let us denote

$$\Delta = \Delta(i_0,i_1,\dots,i_m;j_0,j_1,\dots,j_m;r;n;\ell).$$

[1]Here and throughout the paper the notation $\oplus$ denotes the addition or summation operations.

Consider the $m \times (m+1)$ matrices

$$M = \begin{pmatrix} 1 & 1 & \dots & 1 \\ \alpha^{i_0} & \alpha^{i_1} & \dots & \alpha^{i_m} \\ \alpha^{2i_0} & \alpha^{2i_1} & \dots & \alpha^{2i_m} \\ \vdots & \vdots & \ddots & \vdots \\ \alpha^{(m-1)i_0} & \alpha^{(m-1)i_1} & \dots & \alpha^{(m-1)i_m} \end{pmatrix}$$

and

$$M' = \begin{pmatrix} 1 & 1 & \dots & 1 \\ \alpha^{j_0} & \alpha^{j_1} & \dots & \alpha^{j_m} \\ \alpha^{2j_0} & \alpha^{2j_1} & \dots & \alpha^{2j_m} \\ \vdots & \vdots & \ddots & \vdots \\ \alpha^{(m-1)j_0} & \alpha^{(m-1)j_1} & \dots & \alpha^{(m-1)j_m} \end{pmatrix}.$$

For each u, $0 \leqslant u \leqslant m$, let M_u and M'_u denote the $m \times m$ Vandermonde matrices obtained from deleting column u from M and M' respectively. Also, for $0 \leqslant u, v \leqslant 2m+1$, $u \neq v$, let $X^{(u,v)}$ be the $(2m) \times (2m)$ matrix obtained from removing columns u and v and the last two rows from $M(i_0, i_1, \dots, i_m; j_0, j_1, \dots, j_m; r; n; \ell)$.

If $0 \leqslant u, v \leqslant m$, $u \neq v$,

$$X^{(u,v)} = \left(\begin{array}{c|c} P & \underline{0} \\ \hline \underline{0} & M' \end{array}\right),$$

where P denotes an $m \times (m-1)$ matrix and $\underline{0}$ are zero matrices. Notice that $X^{(u,v)}$ has rank smaller than $2m$, since the first m rows have rank smaller than m. Thus,

$$\det\left(X^{(u,v)}\right) = 0 \quad \text{for } 0 \leqslant u, v \leqslant m,\ u \neq v.$$

If $0 \leqslant u \leqslant m$ and $m+1 \leqslant v \leqslant 2m+1$,

$$X^{(u,v)} = \left(\begin{array}{c|c} M_u & \underline{0} \\ \hline \underline{0} & M'_{v-m-1} \end{array}\right).$$

By properties of determinants,

$$\det\left(X^{(u,v)}\right) = \left(\det(M_u)\right)\left(\det(M'_{v-m-1})\right)$$

for $0 \leqslant u \leqslant m$, $m+1 \leqslant v \leqslant 2m+1$. Similarly,

$$\det\left(X^{(u,v)}\right) = \left(\det(M'_{u-m-1})\right)\left(\det(M_v)\right)$$

for $m+1 \leqslant u \leqslant 2m+1$, $0 \leqslant v \leqslant m$, and

$$\det\left(X^{(u,v)}\right) = 0,$$

for $m+1 \leqslant u, v \leqslant 2m+1$, $u \neq v$.

Expanding the determinant Δ from the bottom row, and then from the next to bottom row, we obtain

$$\begin{aligned}
\Delta &= \left(\bigoplus_{u=0}^{m} \alpha^{-i_u} \bigoplus_{\substack{v=0 \\ v\neq u}}^{m} \alpha^{mi_v} \det\left(X^{(u,v)}\right)\right) \\
&\quad \oplus \left(\bigoplus_{u=0}^{m} \alpha^{-i_u} \bigoplus_{v=m+1}^{2m+1} \alpha^{mj_{v-m-1}} \det\left(X^{(u,v)}\right)\right) \\
&\quad \oplus \left(\bigoplus_{u=m+1}^{2m+1} \alpha^{-n\ell-j_{u-m-1}} \bigoplus_{\substack{v=0 \\ v\neq u}}^{m} \alpha^{mi_v} \det\left(X^{(u,v)}\right)\right) \\
&\quad \oplus \left(\bigoplus_{u=m+1}^{2m+1} \alpha^{-n\ell-j_{u-m-1}} \bigoplus_{v=m+1}^{2m+1} \alpha^{mj_{v-m-1}} \det\left(X^{(u,v)}\right)\right) \\
&= \left(\bigoplus_{u=0}^{m} \alpha^{-i_u} \bigoplus_{v=0}^{m} \alpha^{mj_v} \det(M_u)\det(M'_v)\right) \\
&\quad \oplus \left(\bigoplus_{u=0}^{m} \alpha^{-n\ell-j_u} \bigoplus_{v=0}^{m} \alpha^{mi_v} \det(M_v)\det(M'_u)\right) \\
&= \left(\bigoplus_{u=0}^{m} \alpha^{-i_u} \det(M_u)\right)\left(\bigoplus_{u=0}^{m} \alpha^{mj_u} \det(M'_u)\right) \\
&\quad \oplus \left(\bigoplus_{u=0}^{m} \alpha^{-n\ell-j_u} \det(M'_u)\right)\left(\bigoplus_{u=0}^{m} \alpha^{mi_u} \det(M_u)\right).
\end{aligned} \tag{5}$$

Let

$$W_0 = \begin{pmatrix} 1 & 1 & \dots & 1 \\ \alpha^{i_0} & \alpha^{i_1} & \dots & \alpha^{i_m} \\ \alpha^{2i_0} & \alpha^{2i_1} & \dots & \alpha^{2i_m} \\ \vdots & \vdots & \ddots & \vdots \\ \alpha^{(m-1)i_0} & \alpha^{(m-1)i_1} & \dots & \alpha^{(m-1)i_m} \\ \alpha^{mi_0} & \alpha^{mi_1} & \dots & \alpha^{mi_m} \end{pmatrix}$$

$$W_1 = \begin{pmatrix} 1 & 1 & \dots & 1 \\ \alpha^{i_0} & \alpha^{i_1} & \dots & \alpha^{i_m} \\ \alpha^{2i_0} & \alpha^{2i_1} & \dots & \alpha^{2i_m} \\ \vdots & \vdots & \ddots & \vdots \\ \alpha^{(m-1)i_0} & \alpha^{(m-1)i_1} & \dots & \alpha^{(m-1)i_m} \\ \alpha^{-i_0} & \alpha^{-i_1} & \dots & \alpha^{-i_m} \end{pmatrix}$$

$$W'_0 = \begin{pmatrix} 1 & 1 & \dots & 1 \\ \alpha^{j_0} & \alpha^{j_1} & \dots & \alpha^{j_m} \\ \alpha^{2j_0} & \alpha^{2j_1} & \dots & \alpha^{2j_m} \\ \vdots & \vdots & \ddots & \vdots \\ \alpha^{(m-1)j_0} & \alpha^{(m-1)j_1} & \dots & \alpha^{(m-1)j_m} \\ \alpha^{mj_0} & \alpha^{mj_1} & \dots & \alpha^{mj_m} \end{pmatrix}$$

$$W'_1 = \begin{pmatrix} 1 & 1 & \dots & 1 \\ \alpha^{j_0} & \alpha^{j_1} & \dots & \alpha^{j_m} \\ \alpha^{2j_0} & \alpha^{2j_1} & \dots & \alpha^{2j_m} \\ \vdots & \vdots & \ddots & \vdots \\ \alpha^{(m-1)j_0} & \alpha^{(m-1)j_1} & \dots & \alpha^{(m-1)j_m} \\ \alpha^{-m\ell-j_0} & \alpha^{-m\ell-j_1} & \dots & \alpha^{-m\ell-j_m} \end{pmatrix}$$

Notice that, by properties of determinants and of Vandermonde determinants,

$$\det(W_0) = \bigoplus_{u=0}^{m} \alpha^{mi_u} \det(M_u) = \prod_{0\leqslant u<v\leqslant m} (\alpha^{i_u} \oplus \alpha^{i_v})$$

$$\det(W_1) = \bigoplus_{u=0}^{m} \alpha^{-i_u} \det(M_u) = \alpha^{-\sum_{u=0}^{m} i_u} \prod_{0\leqslant u<v\leqslant m} (\alpha^{i_u} \oplus \alpha^{i_v})$$

$$\det(W_0') = \bigoplus_{u=0}^{m} \alpha^{mj_u} \det(M_u') = \prod_{0\leqslant u<v\leqslant m} (\alpha^{j_u} \oplus \alpha^{j_v})$$

$$\det(W_1') = \bigoplus_{u=0}^{m} \alpha^{-m\ell-j_u} \det(M_u') = \alpha^{-m\ell-\sum_{u=0}^{m} j_u} \prod_{0\leqslant u<v\leqslant m} (\alpha^{j_u} \oplus \alpha^{j_v}).$$

So, (5) becomes

$$\begin{aligned}\Delta &= \begin{pmatrix} \det(W_0) & \det(W_0') \\ \det(W_1) & \det(W_1') \end{pmatrix} \\ &= \left(\prod_{0\leqslant u<v\leqslant m} \left(\alpha^{i_u} \oplus \alpha^{i_v}\right) \left(\alpha^{j_u} \oplus \alpha^{j_v}\right) \right) \\ &\quad \cdot \det \begin{pmatrix} 1 & 1 \\ \alpha^{-\sum_{u=0}^{m} i_u} & \alpha^{-m\ell-\sum_{u=0}^{m} j_u} \end{pmatrix}\end{aligned}$$

and (4) follows. ∎

Lemma 2.1 is valid also over the ring of polynomials modulo $M_p(x)$, p prime, where $rn < p$. Let us illustrate it with an example for $m=1$ and $m=2$.

Example 2.2: Let $m=1$, then

$$M(i_0,i_1;j_0,j_1;r;n;\ell) = \left(\begin{array}{cc|cc} 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ \hline \alpha^{i_0} & \alpha^{i_1} & \alpha^{j_0} & \alpha^{j_1} \\ \alpha^{-i_0} & \alpha^{-i_1} & \alpha^{-n\ell-j_0} & \alpha^{-n\ell-j_1} \end{array}\right)$$

and

$$\Delta(i_0,i_1;j_0,j_1;r;n;s;\ell) = \left(\alpha^{i_0} \oplus \alpha^{i_1}\right) \left(\alpha^{j_0} \oplus \alpha^{j_1}\right) \left(\alpha^{-i_0-i_1} \oplus \alpha^{-n\ell-j_0-j_1}\right).$$

If $m=2$, Lemma 2.1 gives

$$M(i_0,i_1,i_2;j_0,j_1,j_2;r;n;\ell) = \left(\begin{array}{ccc|ccc} 1 & 1 & 1 & 0 & 0 & 0 \\ \alpha^{i_0} & \alpha^{i_1} & \alpha^{i_2} & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & \alpha^{j_0} & \alpha^{j_1} & \alpha^{j_2} \\ \hline \alpha^{2i_0} & \alpha^{2i_1} & \alpha^{2i_2} & \alpha^{j_0} & \alpha^{j_1} & \alpha^{j_2} \\ \alpha^{-i_0} & \alpha^{-i_1} & \alpha^{-i_2} & \alpha^{-n\ell-j_0} & \alpha^{-n\ell-j_1} & \alpha^{-n\ell-j_2} \end{array}\right)$$

and

$$\begin{aligned}&\Delta(i_0,i_1,i_2;j_0,j_1,j_2;r;n;s;\ell) \\ &= \left(\alpha^{i_0} \oplus \alpha^{i_1}\right) \left(\alpha^{i_0} \oplus \alpha^{i_2}\right) \left(\alpha^{i_1} \oplus \alpha^{i_2}\right) \left(\alpha^{j_0} \oplus \alpha^{j_1}\right) \\ &\left(\alpha^{j_0} \oplus \alpha^{j_2}\right) \left(\alpha^{j_1} \oplus \alpha^{j_2}\right) \left(\alpha^{-i_0-i_1-i_2} \oplus \alpha^{-n\ell-j_0-j_1-j_2}\right).\end{aligned}$$

In the next section we study codes $\mathcal{C}(r,n,m,2;f_\alpha(x))$ and $\mathcal{C}(r,n,m,2;M_p(x))$ as SD and PMDS codes.

III. Construction of SD and PMDS codes

Let us start with our main result for SD codes.

Theorem 3.1: The codes $\mathcal{C}(r,n,m,2;f_\alpha(x))$ and $\mathcal{C}(r,n,m,2;M_p(x))$ are SD.

Proof: Assume that m columns have been erased and in addition we have two random erasures. Assume first that these two random erasures occurred in the same row ℓ of the stripe. The rows that are different from ℓ are corrected since each one of them has m erasures, which are handled by the horizontal code, that is, each horizontal code is given by the parity-check matrix H_0, which is the parity-check matrix of a RS code that can correct up to m erasures [9]. Thus, we have to solve a linear system with $m+2$ unknowns. Without loss of generality, assume that the erasures in row ℓ occurred in locations $i_0, i_1, \ldots, i_m, i_{m+1}$, where $0 \leqslant i_0 < i_1 < \cdots < i_m < i_{m+1} \leqslant n$. According to the parity-check matrix of the code as given by (1), (2), and (3), there will be a unique solution if and only if the $(m+2)\times(m+2)$ matrix

$$\begin{pmatrix} 1 & 1 & \ldots & 1 & 1 \\ \alpha^{i_0} & \alpha^{i_1} & \ldots & \alpha^{i_m} & \alpha^{i_{m+1}} \\ \alpha^{2i_0} & \alpha^{2i_1} & \ldots & \alpha^{2i_m} & \alpha^{2i_{m+1}} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \alpha^{mi_0} & \alpha^{mi_1} & \ldots & \alpha^{mi_m} & \alpha^{mi_{m+1}} \\ \alpha^{-n\ell-i_0} & \alpha^{-n\ell-i_1} & \ldots & \alpha^{-n\ell-i_m} & \alpha^{-n\ell-i_{m+1}} \end{pmatrix}$$

is invertible. By taking $\alpha^{-n\ell}$ in the last row as a common factor, and by multiplying each column j, $0 \leqslant j \leqslant m+1$, by α^{i_j}, this matrix is transformed into a Vandermonde matrix, which is always invertible in a field and also in the ring of polynomials modulo $M_p(x)$ [4].

Consider now the case in which the two random failures occur in different rows. Specifically, assume that columns $i_0, i_1, \ldots, i_{m-1}$ were erased, where $0 \leqslant i_0 < i_1 < \ldots < i_{m-1} \leqslant n-1$, and in addition, entries (ℓ,t) and (ℓ',t') were erased, where $t,t' \notin \{i_0,i_1,\ldots,i_{m-1}\}$ and $0 \leqslant \ell < \ell' \leqslant r-1$. Again, using the parity-check matrix of the code as given by (1), (2), and (3), there will be a unique solution if and only if the $(2m+2)\times(2m+2)$ matrix

$$\left(\begin{array}{ccccc|ccccc} 1 & \ldots & 1 & 1 & & 0 & \ldots & 0 & 0 \\ \alpha^{i_0} & \ldots & \alpha^{i_{m-1}} & \alpha^{t} & & 0 & \ldots & 0 & 0 \\ \alpha^{2i_0} & \ldots & \alpha^{2i_{m-1}} & \alpha^{2t} & & 0 & \ldots & 0 & 0 \\ \vdots & \ddots & \vdots & \vdots & & \vdots & \ddots & \vdots & \vdots \\ \alpha^{(m-1)i_0} & \ldots & \alpha^{(m-1)i_{m-1}} & \alpha^{(m-1)t} & & 0 & \ldots & 0 & 0 \\ \hline 0 & \ldots & 0 & 0 & & 1 & \ldots & 1 & 1 \\ 0 & \ldots & 0 & 0 & & \alpha^{i_0} & \ldots & \alpha^{i_{m-1}} & \alpha^{t'} \\ 0 & \ldots & 0 & 0 & & \alpha^{2i_0} & \ldots & \alpha^{2i_{m-1}} & \alpha^{2t'} \\ \vdots & \ddots & \vdots & \vdots & & \vdots & \ddots & \vdots & \vdots \\ 0 & \ldots & 0 & 0 & & \alpha^{(m-1)i_0} & \ldots & \alpha^{(m-1)i_{m-1}} & \alpha^{(m-1)t'} \\ \hline \alpha^{mi_0} & \ldots & \alpha^{mi_{m-1}} & \alpha^{mt} & & \alpha^{mi_0} & \ldots & \alpha^{mi_{m-1}} & \alpha^{mt'} \\ \alpha^{-n\ell-i_0} & \ldots & \alpha^{-n\ell-i_{m-1}} & \alpha^{-n\ell-t} & & \alpha^{-n\ell'-i_0} & \ldots & \alpha^{-n\ell'-i_{m-1}} & \alpha^{-n\ell'-t'} \end{array}\right)$$

is invertible. Taking $\alpha^{-n\ell}$ as a common factor in the last row, we obtain the matrix

$$M(i_0,i_1,i_2,\ldots,i_{m-1},t;i_0,i_1,i_2,\ldots,i_{m-1},t';r;n;\ell'-\ell)$$

as defined in Lemma 2.1, whose determinant, by (4), is given by

$$\begin{aligned}\Delta(i_0,i_1,i_2,\ldots,i_{m-1},t;i_0,i_1,i_2,\ldots,i_{m-1},t';r;n;\ell'-\ell)\\ =\left(\prod_{0\leqslant u<v\leqslant m-1}\left(\alpha^{i_u}\oplus\alpha^{i_v}\right)^2\right)\\ \cdot\left(\prod_{0\leqslant u\leqslant m-1}\left(\alpha^{i_u}\oplus\alpha^{t}\right)\left(\alpha^{i_u}\oplus\alpha^{t'}\right)\right)\\ \cdot\alpha^{-\sum_{u=0}^{m-1}i_u}\left(\alpha^{-t}\oplus\alpha^{-n(\ell'-\ell)-t'}\right).\end{aligned}$$

For simplicity, redefine $\ell\leftarrow\ell'-\ell$, hence, $1\leqslant\ell\leqslant r-1$. Each binomial $\left(\alpha^{i_u}\oplus\alpha^{i_v}\right)$, $\left(\alpha^{i_u}\oplus\alpha^{t}\right)$ and $\left(\alpha^{i_u}\oplus\alpha^{t'}\right)$ above is invertible, so it remains to be proven that $\left(\alpha^{-t}\oplus\alpha^{-n\ell-t'}\right)$ is invertible. If it is not, $n\ell+t'-t\equiv 0\pmod{\mathcal{O}(\alpha)}$. But

$$\begin{aligned}0<n\ell+t'-t&\leqslant n(r-1)+t'-t\\ &=nr-\left(n-(t'-t)\right)\leqslant nr-1<\mathcal{O}(\alpha),\end{aligned}$$

so, $n\ell+t'-t\not\equiv 0\pmod{\mathcal{O}(\alpha)}$. ∎

Next, let us prove a similar result for PMDS codes. In fact, codes $\mathcal{C}(r,n,m,2;f_\alpha(x))$ and $\mathcal{C}(r,n,m,2;M_p(x))$ are not PMDS, but we will obtain PMDS codes with a modification that requires a larger field or ring. Let

$$N \;=\; (m+1)(n-m-1)+1 \tag{6}$$

$\alpha\in GF(2^w)$ and $rN\leqslant\mathcal{O}(\alpha)$. As in the case of SD codes, we construct a PMDS code consisting of $r\times n$ arrays over $GF(2^w)$, such that m of the columns correspond to parity and in addition, two extra symbols also correspond to parity. When read row-wise, the codewords belong in an $[rn,r(n-m)-2]$ code over $GF(2^w)$. Specifically, let $\mathcal{C}'(r,n,m,2;f_\alpha(x))$ be the $[rn,r(n-m)-2]$ code whose $(mr+2)\times rn$ parity-check matrix is given by H', which is identical to H in (1), except for the bottom two rows are defined as:

$$\left(\;H'_1\;\middle|\;H'_2\;\middle|\;\ldots\;\middle|\;H'_r\;\right)$$

where, for $1\leqslant j\leqslant r$,

$$H'_j=\begin{pmatrix}1 & \alpha^m & \alpha^{2m} & \ldots & \alpha^{m(n-1)}\\ \alpha^{-(j-1)N} & \alpha^{-(j-1)N-1} & \alpha^{-(j-1)N-2} & \ldots & \alpha^{-(j-1)N-(n-1)}\end{pmatrix}.$$

As before, the construction is also valid over the ring of polynomials $M_p(x)$, p prime, in which case we denote the codes $\mathcal{C}'(r,n,m,2;M_p(x))$. Let us give an example.

Example 3.1: Let $n=5$, $m=1$ and $r=3$. According to (6), $N=(2)(3)+1=7$. Thus, we need $\mathcal{O}(\alpha)>rN=21$. For instance we may consider the field $GF(32)$ and α primitive in $GF(32)$, i.e., $\mathcal{O}(\alpha)=31>21$ (we can also handle $r=4$ in this example). Thus, the parity-check matrix of $\mathcal{C}'(3,5,1,2;f_\alpha(x))$ is given by

$$\left(\begin{array}{ccccc|ccccc|ccccc}
1&1&1&1&1&0&0&0&0&0&0&0&0&0&0\\
0&0&0&0&0&1&1&1&1&1&0&0&0&0&0\\
\hline
0&0&0&0&0&0&0&0&0&0&1&1&1&1&1\\
\hline
1&\alpha&\alpha^2&\alpha^3&\alpha^4&1&\alpha&\alpha^2&\alpha^3&\alpha^4&1&\alpha&\alpha^2&\alpha^3&\alpha^4\\
\hline
1&\alpha^{30}&\alpha^{29}&\alpha^{28}&\alpha^{27}&\alpha^{24}&\alpha^{23}&\alpha^{22}&\alpha^{21}&\alpha^{20}&\alpha^{17}&\alpha^{16}&\alpha^{15}&\alpha^{14}&\alpha^{13}
\end{array}\right).$$

Theorem 3.2: The codes $\mathcal{C}'(r,n,m,2;f_\alpha(x))$ and $\mathcal{C}'(r,n,m,2;M_p(x))$ are PMDS.

Theorem 3.2 is proven similarly to Theorem 3.1. For reasons of space, we omit the proof here.

IV. Conclusion

We have described a construction for SD codes and PMDS codes where the number of additional sectors, s equals two. The minimal field size required by the construction for SD codes is only the total number of sectors in the array, and in the case of PMDS codes, at most of quadratic order in the total number of sectors.

Further results, not described in detail due to the space limit, are a construction for $(m;s)$ SD codes restricted to not having two of the s erased sectors in the same disk, and a construction for $(m;s)$-like PMDS codes with extra redundancy of order $O(s\log s)$.

As related work, let us mention a recent paper [6] that gives constructions of $(1;s)$ PMDS codes trying to minimize the size of the field. In fact, $(1;s)$ PMDS, called Maximally Recoverable codes in [6], satisfy also the requirements of Locally Repairable codes [10], [13]. Additionally, the recently-defined STAIR codes relax the failure-coverage of SD codes in order to allow for general constructions [8].

Acknowledgment

This work is supported by the National Science Foundation, under grant CSR-1016636, and by an IBM Faculty Award. The work of Eitan Yaakobi was done while he was with the Electrical Engineering Department, California Institute of Technology, Pasadena, CA 91125, U.S.A.

References

[1] M. Blaum, "Construction of PMDS and SD Codes extending RAID 5," arXiv:1305.0032 [cs.IT], April 2013.
[2] M. Blaum, J. L. Hafner and S. Hetzler, "Partial-MDS Codes and their Application to RAID Type of Architectures," IEEE Trans. on Information Theory, vol. IT-59, pp. 4510-4519, July 2013.
[3] M. Blaum and J. S. Plank, "Construction of two SD Codes," arXiv:1305.1221 [cs.IT], May 2013.
[4] M. Blaum and R. M. Roth, "New Array Codes for Multiple Phased Burst Correction," IEEE Trans. on Information Theory, vol. IT-39, pp. 66-77, January 1993.
[5] G. A. Gibson, "Redundant Disk Arrays," MIT Press, 1992.
[6] P. Gopalan, C. Huang, B. Jenkins and S. Yekhanin, "Explicit Maximally Recoverable Codes with Locality," arXiv:1307.4150v2 [cs.IT], July 2013.
[7] C. Huang, H. Simitci, Y. Xu, A. Ogus, B. Calder, P. Gopalan, J. Li and S. Yekhanin, "Erasure Coding in Windows Azure Storage," 2012 USENIX Annual Technical Conference, Boston, June 2012.
[8] M. Li and P. C. Lee, "STAIR codes: A general family of erasure codes for tolerating device and sector failures in practical storage systems," Fail," FAST 14, Santa Clara, CA, February 2014.
[9] F. J. MacWilliams and N. J. A. Sloane, "The Theory of Error-Correcting Codes," North Holland, Amsterdam, 1977.
[10] D.S. Papailiopoulos and A.G. Dimakis, "Locally repairable codes," Proc. IEEE International Symposium on Information Theory, pp. 2771–2775, July 2012.
[11] J. S. Plank and M. Blaum, "Sector-Disk (SD) Erasure Codes for Mixed Failure Modes in RAID Systems," ACM Transactions on Storage, Vol. 10, No. 1, Article 4, January 2014.
[12] J. S. Plank, M. Blaum and J. L. Hafner, "SD Codes: Erasure Codes Designed for How Storage Systems Really Fail," FAST 13, San Jose, CA, February 2013.
[13] I. Tamo and A. Barg, "A family of optimal locally recoverable codes," submitted to IEEE Trans. on Information Theory.