An Explicit Construction of PMDS (Maximally Recoverable) Codes for All Parameters

Gokhan Calis and O. Ozan Koyluoglu

***Abstract*—PMDS (a.k.a. maximally recoverable) codes allow for local erasure recovery by utilizing row-wise parities and additional erasure correction through global parities. Recent works on PMDS codes focus on special case parameter settings, and a general construction for PMDS codes is stated as an open problem. This paper provides an explicit construction for PMDS codes for all parameters utilizing concatenation of Gabidulin and MDS codes, a technique orginally proposed by Rawat et al. for constructing optimal locally repairable codes. This approach allows for PMDS constructions for any parameters albeit with large field sizes. To lower the field size, a relaxation on the rate requirement is considered, and PMDS codes based on combinatorial designs are constructed.**

***Index Terms*—MDS Codes, Partial MDS Codes, Maximally Recoverable Codes, Gabidulin Codes, Combinatorial Designs.**

I. Introduction

Redundant array of independent disks (RAID) [1] architecture is used to prevent systems from data loss in case of catastrophic failures (disk failure). Popular RAID schemes include RAID 4 (one disk dedicated to parities), RAID 5 (parities are distributed to disks) and RAID 6 (similar to RAID 5 but includes second parity scheme). Maximum distance separable (MDS) codes, i.e., Reed-Solomon codes, can be utilized for erasure correcting in RAID systems, i.e., in RAID 6 to overcome the failure of two disks. However, using solid-state drives (SSDs) (instead of hard disk drives (HDD)) brought challenges, e.g., the system may experience both *disk failures* and *hard errors* which may not be realized unless the specific sector is accessed. RAID 6 architecture can tolerate such an erasure pattern, i.e., it can tolerate two sector erasures in a disk. However, the cost of recovery is expensive and partial MDS (PMDS) codes are proposed to overcome this problem by utilizing both row-wise parities and global parities to simultaneously recover from mixed failures [2].

PMDS codes tolerate mixed failures consisting of column failures (referring to a disk) in an $r \times n$ array and additional failures (sectors). Each row in the array forms an MDS code, i.e., each row forms a *local group* for erasure correction of up to m symbols. Considering $r \times n$ array over a finite field $\mathbb{F}$, PMDS codes have the following properties [3].

- m entire columns of elements are devoted to coding.
- Each row is an $[n, n-m, m+1]$ MDS code.
- In the remaining $n-m$ columns, s more elements are also devoted to coding.
- Any m elements per row plus any additional s erasures in the array can be recovered.

PMDS codes are labeled with $(m; s)$ and formal definition is given as follows [3].

Definition 1. *Let $\mathcal{C}$ be a linear $[rn, k]$ 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. $\mathcal{C}$ is an $(m; s)$ partial MDS (PMDS) code if, for any $(s_1, s_2, \ldots, s_t)$ such that each $s_j \geq 1$ and $\sum_{j=1}^{t} s_j = s$, and for any $i_1, i_2, \ldots, i_t$ such that $0 \leq i_1 < i_2 < \cdots < i_t \leq r-1$, $\mathcal{C}$ can correct up to $s_j + m$ erasures in each row i_j, $1 \leq j \leq t$.*

PMDS codes draw attention recently and code constructions are proposed in the literature, however, the parameter set ($(m; s)$ values) is limited. Explicit constructions are provided in [4], [5] for $(m; s) = (1; 1)$ and $(m; s) = (\leq 2; 2)$, in [2], [6] for $(m; s) = (\geq 1; 1)$, in [7] for $(m; s) = (\geq 1; 2)$, in [8] for $(m; s) = (1; 3)$ and $(m; s) = (1; 4)$, in [2] for $(m; s) = (1; \geq 1)$ and in [9] for $(m; s) = (\geq 1; 1)$. In all these explicit PMDS constructions, m or s is set to be 1 or 2.

Coding schemes that can be considered as relaxations to erasure recovery properties of PMDS codes include SD codes, STAIR codes, and t-level Generalized Concatenated (GC) codes. Sector Disk (SD) codes are defined similar to PMDS codes with only difference being that SD codes tolerate the erasure of any m columns plus any additional s symbols in the array [3], [5]. In STAIR codes [10], the number of disks that may simultaneously fail containing sector failures as well as the number of sector failures per disk are limited. Sector failure coverage is defined by a vector $\mathbf{e}$, and STAIR codes can tolerate any m column failures plus total of s sector failures in the remaining $n-m$ columns defined by $\mathbf{e}$. t-level Generalized Concatenated (GC) codes [6] defined by their parity-check matrix $H(n; \mathbf{u})$ as a relaxation of PMDS codes in which the erasure pattern is defined by a vector $\mathbf{u}$. For example $H(5; (2, 2, 3, 3))$ code can tolerate any 2 erasures per row from any 2 rows plus any 3 erasures per row from the remaining two rows.

Locally repairable codes (LRCs) has been studied recently [11]–[14] and these codes allow a recovery of a symbol within a corresponding *local group*, which usually has small number of nodes. Since recovery only requires small number of nodes, LRCs are shown to be useful in reducing cost of recovery. We remark that d_{min}-optimal LRCs necessarily have disjoint local groups, which make them as candidates for constructing PMDS codes. However, this approach (utilizing d_{min}-optimal LRCs) produces PMDS codes only for special parameter settings.

The authors are with the Department of Electrical and Computer Engineering, University of Arizona. Email: {gcalis, ozan}@email.arizona.edu. This work is supported in part by NSF grant CCF 1563622.

In this study, we first propose an explicit PMDS code construction for all parameters using concatenation of Gabidulun and MDS codes, a technique originally proposed in [13] for constructing optimal locally repairable codes. Gabidulin codes are introduced in the following section, and the general PMDS construction along with examples are detailed in Section III. Then, to lower the field size requirement of this approach, we develop rate suboptimal PMDS constructions using combinatorial designs in Section IV. In particular, we will refer to the PMDS definition given above as rate-optimal PMDS, where the corresponding rate is $R^* = \frac{r(n-m)-s}{rn}$ as $k = r(n-m)-s$, and compare this optimal rate with those of codes that we construct based on combinatorial designs.

II. Maximum Rank Distance (MRD) codes

MRD codes can be described either in a vector or in a matrix notation. We provide both presentations here. Before formally defining these codes, we first define column rank, rank distance, and linearized polynomials; and, then, provide the construction of Gabidulin codes.

Definition 2 (Column rank)**.** *For a given basis of $\mathbb{F}_{q^M}$ over $\mathbb{F}_q$, the column rank of a vector $\mathbf{v} \in \mathbb{F}_{q^M}^N$ over the base field $\mathbb{F}_q$, denoted by $Rk(\mathbf{v}|\mathbb{F}_q)$, is the maximum number of linearly independent coordinates of $\mathbf{v}$ over the base field $\mathbb{F}_q$.*

We note that a basis also establishes an isomorphism between N-length vectors, in $\mathbb{F}_{q^M}^N$, to $M \times N$ matrices, in $\mathbb{F}_q^{M\times N}$. (To be more explicit, this is the mapping of $\mathbf{v} \in \mathbb{F}_{q^M}^{1\times N}$ to $\mathbf{V} \in \mathbb{F}_q^{M\times N}$ by writing out each element in $\mathbf{v} = \{v(1), \cdots, v(N)\}$, $v(i) \in \mathbb{F}_{q^M}$, using the given basis, as a vector $\mathbf{v}(i) \in \mathbb{F}_q^{1\times M}$; and, constructing the matrix $\mathbf{V} = [\mathbf{v}(1)^T, \cdots, \mathbf{v}(N)^T]$.) In addition, for the given basis, the column rank $Rk(\mathbf{v}|\mathbb{F}_q)$ is equal to $rank(\mathbf{V})$, the rank of the corresponding matrix of $\mathbf{v}$. We will utilize this connection to provide a matrix interpretation of MRD codes.

Definition 3 (Rank distance)**.** *Rank distance between two vectors is defined by $d_R(\mathbf{v}_1, \mathbf{v}_2) = Rk(\mathbf{v}_1 - \mathbf{v}_2|\mathbb{F}_q)$.*

Rank distance is a metric [15]. In matrix representation, this distance metric is equivalent to the rank of the difference between the corresponding matrices of the two vectors, i.e., $Rk(\mathbf{v}_1 - \mathbf{v}_2|\mathbb{F}_q) = rank(\mathbf{V}_1 - \mathbf{V}_2)$. Rank metric codes utilize the definition given above as the underlying metric.

Definition 4 (Matrix (array) code)**.** *A matrix code is defined as any nonempty subset of $\mathbb{F}_q^{M\times N}$. (This is also called array code [16].)*

Rank-metric code is a matrix (array) code, where the distance is the rank distance. The minimum distance of a rank-metric code $\mathcal{C} \subseteq \mathbb{F}_q^{M\times N}$ is given by

$$d_R(\mathcal{C}) = \min_{\mathbf{v}_1,\mathbf{v}_2\in\mathcal{C};\mathbf{v}_1\neq\mathbf{v}_2} d_R(\mathbf{v}_1, \mathbf{v}_2). \tag{1}$$

Remark 5 (Transposed code)**.** *The transposed code of a given rank-metric code $\mathbb{C} \subseteq \mathbb{F}_q^{M\times N}$ is given by $\mathbb{C}^T = \{\mathbf{v}^T : \mathbf{v} \in \mathcal{C}\} \subseteq \mathbb{F}_q^{N\times M}$, and is also a rank-metric code. Remarkably, $|\mathcal{C}| = |\mathcal{C}^T|$ and $d_R(\mathcal{C}) = d_R(\mathcal{C}^T)$.*

As that of the codes for Hamming metric, one can provide a Singleton bound for the rank-metric codes.

Lemma 6 ([15], [17])**.** *Consider a rank-metric code $\mathcal{C} \subseteq \mathbb{F}_q^{M\times N}$ with minimum distance $d_R(\mathcal{C}) = D$. Then,*

$$\log_q(|\mathcal{C}|) \leq \min\{M(N-D+1), M(N-D+1)\}. \tag{2}$$

Proof. Reader may refer to [15], [17] for the proof. □

The codes that achieve the bound in (2) are called *maximum rank distance* (MRD) codes. Gabidulin presented a construction of such codes for $N \leq M$ [15]. By transposition of the codes obtained by the Gabidulin construction, one can also obtain codes with $M \geq N$. Thus, MRD codes exist for all $D \leq \min\{M, N\}$. Before providing the Gabidulin construction of MRD codes, we introduce linearized polynomials.

Definition 7 (Linearized polynomial)**.** *A linearized polynomial $f(g)$ over $\mathbb{F}_{q^M}$ of q-degree $K-1$ has the form*

$$f(g) = \sum_{i=0}^{K-1} c_i g^{[i]},$$

where the coefficients $c_i \in \mathbb{F}_{q^M}$, $c_{k-1} \neq 0$, and $[i] = q^i$.

We note that the linearized polynomial satisfies $f(a_1g_1 + a_2g_2) = a_1f(g_1) + a_2f(g_2)$, for a given $a_1, a_2 \in \mathbb{F}_q$ and $g_1, g_2 \in \mathbb{F}_{q^M}$. We now provide Gabidulin construction of maximum rank distance (MRD) codes.

Definition 8 (MRD (Gabidulin) codes)**.** *An $[N, K, D]$ MRD code $\mathcal{C}_{MRD}$ over the extension field $\mathbb{F}_{q^M}$ where $M \geq N$*

- *has length-K input $u_0, \cdots, u_{K-1}$ where $u_i \in \mathbb{F}_{q^M}, i = 0, \cdots, K-1$, and*
- *encodes the input to length-N codewords by*

$$x_j = f(g_j) = \sum_{i=0}^{K-1} u_i g_j^{[i]},$$

for $j = 1, \cdots, N$, where the linearized polynomial is constructed with N linearly independent, over $\mathbb{F}_q$, generator elements $\{g_1, \cdots, g_N\}$ with $g_j \in \mathbb{F}_{q^M}$; and its coefficients are selected by the length-K input vector.

We note that, the above code can be represented with a generator matrix representation, where $\mathbf{x} = \mathbf{uG}$ with $\mathbf{G} = [g_1, \cdots, g_N; \cdots; g_1^{[K-1]}, \cdots, g_N^{[K-1]}]$.

Note that, symbol erasures in the vector representation of a codeword in the Gabidulin code correspond to column erasures in the matrix representation. Here, any non zero code (vector) has a rank (norm) of at least $N-K+1$. Thus, the Gabidulin code achieves a rank distance of $D = N-K+1$, which is the maximum achiveable, and can correct $D-1$ erasures. The construction above is referred to as the Gabidulin construction of MRD codes, or, simply, Gabidulin codes [15]–[18].

III. A general construction for PMDS codes

Recently, a concatenation of MRD and MDS array codes are utilized for coding in distributed storage systems. This approach is used for constructing locally repairable codes (with and without security or bandwidth efficient local repairs) in [13], locally repairable codes with minimum bandwidth node repairs in [19], thwarting adversarial errors in [20], [21], and cooperative regenerating codes with built-in security mechanisms against node capture attacks in [22]. We utilize the same concatenation approach here to construct partial MDS (PMDS) codes.

Construction I. [An (m;s) PMDS code over an array of (r, n) symbols (r rows and n columns)] Set K=r(n-m)-s, and consider data symbols $\{u_0, \cdots, u_{K-1}\}$.

- Use [N=K+s,K,D=s+1] Gabidulin code to encode $\{u_0, \cdots, u_{K-1}\}$ to length-N codeword $(x_1, \cdots, x_N)$. That is,

  $$(x_1, \cdots, x_N) = (f(g_1), \cdots, f(g_N)),$$

  where the linearized polynomial $f(g) = u_0 g^{[0]} + \cdots + u_{K-1} g^{[K-1]}$ is constructed with N linearly independent, over $\mathbb{F}_q$, generator elements $\{g_1, \cdots, g_N\}$ each in $\mathbb{F}_{q^M}$; and its coefficients are selected by the length-K input vector. We represent this operation by writing $\mathbf{x} = \mathbf{u}\mathbf{G}_{\text{MRD}}$.
- Split resulting N=K+s=r(n-m) symbols $\{x_1, \cdots, x_N\}$ into r rows each with n-m symbols. We represent this operation by double indexing the codeword symbols, i.e., $x_{i,j}$ is the symbol at row i and column j for $i = 1, \cdots, r$, $j = 1, \cdots, n-m$. We also denote the resulting sets with the vector notation, $\mathbf{x}_{i,1:n-m} = (x_{i,1}, x_{i,2}, \cdots, x_{i,n-m})$ for row i.
- Use an [n,k=n-m,d=m+1] MDS array code for each row to construct additional parities. Representing the output symbols as $\mathbf{y}_{i,1:n}$ we have

  $$\mathbf{y}_{i,1:n} = \mathbf{x}_{i,1:n-m} G_{\text{MDS}}$$

  for each row i, where G_{MDS} is the encoding matrix of the MDS code over $\mathbb{F}_q$. For instance, if a systematic code is used, $\mathbf{x}_{i,1:n-m}$ is encoded into the vector $\mathbf{y}_{i,1:n} = (x_{i,1}, \cdots, x_{i,n-m}, p_{i,1}, \cdots, p_{i,m})$ for each row $i = 1, \cdots, r$.

To summarize, we have

$$\mathbf{y} = \begin{bmatrix} \mathbf{y}_{1,1:n} \\ \mathbf{y}_{2,1:n} \\ \vdots \\ \mathbf{y}_{r,1:n} \end{bmatrix} = \begin{bmatrix} \mathbf{G}_{\text{MDS}} & & & \\ & \mathbf{G}_{\text{MDS}} & & \\ & & \ddots & \\ & & & \mathbf{G}_{\text{MDS}} \end{bmatrix} = \begin{bmatrix} \mathbf{x}_{1,1:n-m}\mathbf{G}_{\text{MDS}} \\ \mathbf{x}_{2,1:n-m}\mathbf{G}_{\text{MDS}} \\ \vdots \\ \mathbf{x}_{r,1:n-m}\mathbf{G}_{\text{MDS}} \end{bmatrix} \tag{3}$$

The resulting codeword symbols are represented in the following symbol matrix representation:

$$\begin{array}{|cccc|}\hline y_{1,1} & y_{1,2} & \cdots & y_{1,n} \\ y_{2,1} & y_{2,2} & \cdots & y_{2,n} \\ \vdots & \vdots & \vdots & \vdots \\ y_{r,1} & y_{r,2} & \cdots & y_{r,n} \\ \hline\end{array} \tag{4}$$

For the case of a systematic MDS code, we have

$$\begin{array}{|ccccccc|}\hline x_{1,1} & x_{1,2} & \cdots & x_{1,n-m} & p_{1,1} & \cdots & p_{1,m} \\ x_{2,1} & x_{2,2} & \cdots & x_{2,n-m} & p_{2,1} & \cdots & p_{2,m} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ x_{r,1} & x_{r,2} & \cdots & x_{r,n-m} & p_{r,1} & \cdots & p_{r,m} \\ \hline\end{array} \tag{5}$$

Proposition 9. *The symbol matrix resulting from Construction 1 has a total of rn symbols that are placed in r rows and n columns. Now, consider that we have m erasures per row, and an additional s erasures over the remaining symbols (referred to as $(m;s)$ erasure pattern). The remaining symbols are sufficient to decode the data symbols $u_0, \cdots, u_{K-1}$, from which the erasures in $(m;s)$ erasure pattern can be recovered by re-encoding the data.*

We first give the following examples before discussing the proof of the proposition above.

Example 10. *Consider construction of (m=1;s=1) PMDS over an array of $r = 2, n = 3$ symbols. Here, one can readily use the trivial parity (i.e., sum of codewords) at each row as an underlying MDS array code. We obtain the codeword $(x_{1,1}, x_{1,2}, p_{1,1} = x_{1,1} + x_{1,2}, x_{2,1}, x_{2,2}, p_{2,1} = x_{2,1} + x_{2,2})$. That is, we have*

$$\begin{array}{|ccc|}\hline x_1 & x_2 & x_1 + x_2 \\ x_3 & x_4 & x_3 + x_4 \\ \hline\end{array} \tag{6}$$

At this point, this symbol matrix can have m=1 erasure in each row and additional $m = 1$ erasure in the remaining symbols. For instance, one may have

$$\begin{array}{|ccc|}\hline x_1 & * & * \\ * & x_4 & x_3 + x_4 \\ \hline\end{array} \tag{7}$$

This resulting symbol array $(x_1, x_4, x_3 + x_4)$ forms a set of linearly independent evaluation points of the underlying linearized polynomial for the [N=4,K=3,D=2] Gabidulin code. By polynomial interpolation, one can then solve for the data coefficients u_0, u_1, u_2, re-encode this into codewords and construct back the full symbol matrix.

Example 11. *Consider construction of (m=2;s=3) PMDS over an array of $r = 3, n = 5$ symbols. Here, we use $[N = 9, K = 6, D = 4]$ Gabidulin code together with an $[n = 5, k = 3, d = 3]$ MDS code. We obtain, e.g., the following symbols, for the case of systematic MDS code.*

$$\begin{array}{|ccccc|}\hline x_{1,1} & x_{1,2} & x_{1,3} & p_{1,1} & p_{1,2} \\ x_{2,1} & x_{2,2} & x_{2,3} & p_{2,1} & p_{2,2} \\ x_{3,1} & x_{3,2} & x_{3,3} & p_{3,1} & p_{3,2} \\ \hline\end{array} \tag{8}$$

At this point, this symbol matrix can have m=2 erasures in each row and additional $s = 3$ erasures in the remaining symbols. For instance, one may have

$$\begin{array}{|ccccc|}\hline x_{1,1} & x_{1,2} & x_{1,3} & * & * \\ * & * & * & p_{2,1} & p_{2,2} \\ * & * & * & * & p_{3,2} \\ \hline\end{array} \tag{9}$$

This resulting symbol array $(x_{1,1}, x_{1,2}, x_{1,3}, p_{2,1}, p_{2,2}, p_{3,2})$ forms a set of linearly independent evaluation points of the underlying linearized polynomial for the [N=9,K=6,D=4]

Gabidulin code. By polynomial interpolation, one can then solve for the data coefficients $u_0, \cdots, u_5$, re-encode this into codewords and construct back the full symbol matrix.

Proof of Proposition 9. We first provide a lemma, which is a summary of the observations given in [13] for the scenario considered here. (In particular, we have scalar symbols here as compared to the vector case considered in [13].)

Lemma 12. *Consider the code given in Construction 1, where the Gabidulin codeword $\mathbf{x} = [x_1, \cdots, x_N] = [f(g_1), \cdots, f(g_N)]$ in $\mathbb{F}_{q^M}^N$ is partitioned into symbol vectors $\mathbf{x}_{i,1:n-m} = (x_{i,1}, \cdots, x_{i,n-m})$ for row $i = 1, \cdots, r$, and each is encoded into symbols $\mathbf{y}_{i,1:n}$ through G_{MDS}. Consider a set $\mathcal{S}$ which is the union of l_i symbols from row i (symbols in $\mathbf{y}_{i,1:n}$). Then, the symbols in $\mathcal{S}$ correspond to the evaluations of the underlying linearized polynomial $f(\cdot)$ at $\sum_{i=1}^{r} \min\{l_i, k\}$ linearly independent (over $\mathbb{F}_q$) points from $\mathbb{F}_{q^m}$.*

The proof of this lemma is provided in Appendix A. We utilize this Lemma in the following.

Corollary 13. *Consider the code given in Construction I and an erasure pattern which leaves l_i number of remaining symbols in row i in the symbol matrix. In such a scenario, if $\sum_{i=1}^{r} \min\{l_i, k\} \geq K$, then, the erasure pattern can be recovered from the remaining symbols. In particular, the remaining symbols result in $\sum_{i=1}^{r} \min\{l_i, k\}$ linearly independent evaluation points for the underlying polynomial (see Lemma 12). And, when this number is greater than or equal to K, the data symbols $u_0, \cdots, u_{K-1}$ can be decoded via polynomial interpolation, from which the pre-erasure situation of the array can be recovered by re-encoding the symbols.*

For a given (m;s) erasure scenario over an array of (r, n) symbols (r rows and n columns), we have m erasures in each row and additional s_i erasures per row, resulting in a total of $rm + \sum_{i=1}^{r} s_i = rm + s$ erasures. In Construction 1, after erasing m symbols from each row, we are left with $n - m$ symbols in r rows. Now, having s_i number of additional erasures in each row will result in having $l_i = n - m - s_i$ number of symbols at row i. Therefore, as the underlying MDS code has a dimension of $k = n - m$, the number of linearly independent evaluations at hand is $\sum_{i=1}^{r} \min\{l_i, k\} = \sum_{i=1}^{r} l_i = r(n-m) - \sum_{i=1}^{r} s_i = r(n-m) - s = K$. Therefore, any (m;s) erasure pattern can be recovered with Construction 1. □

We note that the construction above allows for construction of PMDS for any m and s, but with a field size of $q^{r(n-m)}$, whereas the existing literature on PMDS codes only works for limited range of m or s (with lower field sizes). In the following, we relax the optimal rate requirement in PMDS codes and provide constructions with lower field sizes.

IV. Rate suboptimal PMDS codes through combinatorial designs

We first note that one can readily utilize product codes [23] to obtain PMDS type erasure correction properties. We first report the rate from such an approach and then detail our constructions utilizing combinatorial designs.

Construction II. Assume we have a data matrix of size $(r - s) \times (n - m)$.

- First, create m parities per row by using $[n, n-m]$ MDS code, which expands the matrix to $(r - s) \times n$.
- Apply column-wise $[r, r-s]$ MDS codes to the expanded matrix in the first step to obtain $r \times n$ matrix.

Observe that this code construction tolerates any m erasures per row plus any s while being sup-optimal in the rate since $R^{(II)} = \frac{(r-s)(n-m)}{rn}$.

We will now utilize two combinatorial block designs, projective planes of order-p and resolvable balanced incomplete block design (RBIBD) in order to construct rate suboptimal PMDS codes.

Definition 14. *A (v, κ, λ)-balanced incomplete block design (BIBD) has v points distributed into blocks of size κ such that any pair of points are contained in λ blocks. For a (v, κ, λ)-BIBD, every point occurs in $t = \frac{\lambda(v-1)}{\kappa-1}$ blocks and the design has exactly $r = \frac{\lambda(v^2-v)}{\kappa^2-\kappa}$ blocks. A $(v = p^2 + p + 1, \kappa = p + 1, \lambda = 1)$-BIBD with $p \geq 2$ is called a projective plane of order p.*

Definition 15. *A resolvable balanced incomplete block design (RBIBD) with parameters $(v, r, t, \kappa, \lambda)$ is an arrangement of v points into r blocks of size κ, where $\kappa < v$, such that each point appears in t parallel classes, and every pair of points appears in λ blocks. The five parameters must satisfy $vt = r\kappa$ and $\lambda(v - 1) = t(\kappa - 1)$*

We first provide an example. Assume a data $\mathcal{D} = \mathbf{u}$ of size 9 contains 3 sub-data ($\mathcal{D}_i = [u_{i,1:3}]$) each of size 3, i.e.,

$$\mathbf{u} = \{u_{1,1}, u_{1,2}, u_{1,3}, u_{2,1}, u_{2,2}, u_{2,3}, u_{3,1}, u_{3,2}, u_{3,3}\}.$$

We encode each of these sub-data with $[10, 3]$ MDS code and represent the resulting elements with $\mathcal{P}_1 = p_{1,1:10}$ for $\mathcal{D}_1$, $\mathcal{P}_2 = p_{2,1:10}$ for $\mathcal{D}_2$, and $\mathcal{P}_3 = p_{3,1:10}$ for $\mathcal{D}_3$. We have

$$\left[\ u_{i,1:3}\ \right] \begin{bmatrix} 1 & 1 & \ldots & 1 \\ \alpha_{i,1} & \alpha_{i,2} & \ldots & \alpha_{i,10} \\ \alpha_{i,1}^2 & \alpha_{i,2}^2 & \ldots & \alpha_{i,10}^2 \end{bmatrix} = \left[\ p_{i,1:10}\ \right]. \tag{10}$$

These elements are grouped in a specific way placed into array as represented in Fig. 1 where each codeword symbol contains two elements each coming from two of the different sets $\mathcal{P}_1, \mathcal{P}_2, \mathcal{P}_3$. Thus, each row now can be taken as $[5, 3]$ MDS code. Here, we can think of the generator matrix $\mathbf{G}$ of overall code $\mathcal{C}$ as consisting of 15 thick columns each of size 2 thin columns (corresponding to 2 different sub-data). Note that, the code can tolerate erasure of any $m = 2$ symbols per row plus any $s = 3$ symbols hence allowing recovery from PMDS erasure pattern since the remaining 12 elements (6 symbols) have at least 3 elements (3 thin columns) per sub-data from which each of the sub-data can be recovered and so is the original array.

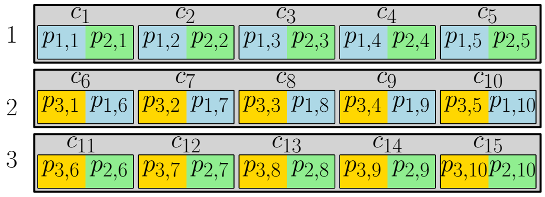
1
c_1 c_2 c_3 c_4 c_5
$p_{1,1}$ $p_{2,1}$ $p_{1,2}$ $p_{2,2}$ $p_{1,3}$ $p_{2,3}$ $p_{1,4}$ $p_{2,4}$ $p_{1,5}$ $p_{2,5}$
2
c_6 c_7 c_8 c_9 c_{10}
$p_{3,1}$ $p_{1,6}$ $p_{3,2}$ $p_{1,7}$ $p_{3,3}$ $p_{1,8}$ $p_{3,4}$ $p_{1,9}$ $p_{3,5}$ $p_{1,10}$
3
c_{11} c_{12} c_{13} c_{14} c_{15}
$p_{3,6}$ $p_{2,6}$ $p_{3,7}$ $p_{2,7}$ $p_{3,8}$ $p_{2,8}$ $p_{3,9}$ $p_{2,9}$ $p_{3,10}$ $p_{2,10}$

Fig. 1. MDS codewords corresponding to each sub-data are placed as symbols of the code according to the underlying projective plane. This code tolerates any $(m = 2; s = 3)$ PMDS erasure pattern.

The general construction using a projective plane of order p is as follows.

Construction III. Assume we have a data $\mathcal{D}$ of size $r(n-m)$, and consider a projective plane of order p with PMDS parameters satisfying $(n-m)p = s$ and $r = p^2+p+1$.

- Partition $\mathcal{D}$ into $r = p^2+p+1$ sub-data, where $p = \frac{s}{n-m}$.
- Encode each sub-data using $[n(p+1), n-m]$ MDS code and distribute the resulting $n(p+1)$ elements for each sub-data evenly to $p+1$ different rows (according to the underlying projective plane).

As a result of this construction, symbols in each row stores elements from $p+1$ distinct sub-data, hence a row can be considered as an $[n, n-m]$ MDS code since puncturing np coordinates from $[n(p+1), n-m]$ MDS code results in $[n, n-m]$ MDS code. We now show that erasure of any m symbols per row plus any s symbols can be tolerated.

Proof. Consider the generator matrix $\mathbf{G}$ which has r sub-block-matrix (corresponding to the rows), each having n thick columns (corresponding to the symbols in each row). Furthermore, each of these thick columns also have $p+1$ thin columns. Erasure of any m nodes per row is same as puncturing any m thick columns from each of the r sub-block-matrix. In addition, any s erasures corresponds to puncturing any additional s thick columns.

Puncturing any m thick columns from each of the r sub-block-matrix has the same effect on each sub-data. However, the additional s erasures may have different effect on different sub-data depending on the erasure pattern. Since any two blocks in the projective plane has only one common point, any $s \geq 2$ thick columns contains at least one common sub-data. Considering the worst case of having all s punctured thick columns containing at least one common sub-data, the remaining thick columns contain at least $n(p+1)-m(p+1)-s$ thin columns for each of the sub-data. Since we have $p = \frac{s}{n-m}$ in the code construction, we have at least $n(p+1)-m(p+1)-p(n-m) = n-m$ thin columns for each of the sub-data. Therefore, using these $n-m$ thin columns, we can decode each of the sub-data using the underlying MDS code from which the original array can be reconstructed. □

Although this construction requires lower field size, it is not rate optimal. The original data is of size $r(n-m)$ and storage cost is $rn(p+1)$ yielding rate as $R^{(III)} = \frac{n-m}{n(p+1)}$ and we have $\frac{R^{(III)}}{R^*} = \frac{p^2+p+1}{(p+1)(p^2+1)}$. One observation is that with projective plane construction, the system may tolerate even more than any s additional erasures (since construction is designed to tolerate the worst case of s). For example, using projective plane of order $p = 1$ for $(m = 2, r = 3, n = 5)$ we can tolerate %100 of $s \leq 3$, %64.29 of $s = 4$ and none of $s \geq 5$ erasures.

Construction IV. Assume we have a data $\mathcal{D}$ and consider an $(v, \kappa, \lambda = 1)$-RBIBD satisfying $s = \frac{(n-m)(v-\kappa)}{\kappa-1}$ and $r = \frac{v(v-1)}{\kappa(\kappa-1)}$.

- Partition $\mathcal{D}$ into v sub-data.
- Encode each sub-data using $[\frac{n(v-1)}{\kappa-1}, n-m]$ MDS code and distribute the resulting $\frac{n(v-1)}{\kappa-1}$ elements for each sub-data according to the underlying RBIBD.

A row in $r \times n$ array stores symbols from the same set of κ sub-data and since each sub-data is repeated $\frac{v-1}{\kappa-1}$ times, each row stores n symbols for each of the κ sub-data. In other words, each row can be represented by a block of RBIBD. PMDS codes need to tolerate any m erasures per row plus any s erasures. Any m erasures per row results in erasure of $\frac{m(v-1)}{\kappa-1}$ for each sub-data. Furthermore, assume the worst case that is the additional s erasures also occur involving a common sub-data, then at least $\frac{n(v-1)}{\kappa-1} - \frac{m(v-1)}{\kappa-1} - s$ symbols remain for each sub-data. Since $s = \frac{(n-m)(v-\kappa)}{\kappa-1}$, we have at least $n-m$ symbols for each sub-data, which is enough to decode each sub-data using the underlying MDS code and from which the original data can be decoded. This code construction yields rate suboptimal PMDS as $R^{(IV)} = \frac{(n-m)(\kappa-1)}{n(v-1)}$ and we have $\frac{R^{(IV)}}{R^*} = \frac{(\kappa-1)v}{v^2-v-v\kappa+\kappa^2}$.

ACKNOWLEDGEMENT

The authors would like to thank Mario Blaum of IBM and Eitan Yaakobi of Technion for discussions on PMDS codes and for comments on the previous version of this manuscript.

APPENDIX

The proof of this observation follows from the linearized property of the polynomial utilized in the Gabidulin code. (The proof is given in Lemma 9 and 23 of [13] for the general case of having vector symbols. See also [19]. We provide a summary here for the scalar case.) Consider row i which encodes symbols $\mathbf{x}_{i,1:n-m}$ into $\mathbf{y}_{i:1:n} = \mathbf{x}_{i,n-m} G_{\text{MDS}}$. Here, representing the corresponding evaluation points with $g_{i,j}$ for row i, we have, as $k = n-m$, $\mathbf{x}_{i,1:k} = (f(g_{i,1}), f(g_{i,2}), \cdots, f(g_{i,k}))$. Now, denoting G_{MDS} as $k \times n$ matrix with entries $[G_{h,j}]$ for $h = 1, \cdots, k$ and $j = 1, \cdots, n$, we have $y_{i,j} = \sum\limits_{h=1}^{k} f(g_{i,h}) G_{h,j}$. At this point, due to the linearized property of $f(\cdot)$, we have $y_{i,j} = f(\sum\limits_{h=1}^{k} G_{h,j} g_{i,h})$. Denote this new evaluation points as $\tilde{g}_{i,j} = \sum\limits_{h=1}^{k} G_{h,j} g_{i,h}$. These evaluation points given by $\tilde{g}_{i,1:n}$ span the space spanned by the set $g_{i,1:k}$. Now, consider a set $\mathcal{S}_i \subseteq \{1, \cdots, n\}$ of size l_i. Due to the full rankness of the matrix G_{MDS}, the set of points $\{\tilde{g}_{i,j}, j \in \mathcal{S}_i\}$ span a $\min\{l_i, k\}$ dimensional space in the space spanned by $g_{i,1:k}$. Furthermore, as the points in different rows, say $g_{i,1:k}$ and $g_{\tilde{i},1:k}$ for $i \neq \tilde{i}$, are independent, we have linear independence of $\tilde{g}_{i,1:n}$ and $\tilde{g}_{\tilde{i},1:n}$ for any $i \neq \tilde{i}$. Therefore, the symbols in $\mathcal{S} = \cup_{i=1}^{r} \mathcal{S}_i$ correspond to the evaluations of the underlying linearized polynomial $f(\cdot)$ at $\sum_{i=1}^{r} \min\{l_i, k\}$ linearly independent (over $\mathbb{F}_q$) points from $\mathbb{F}_{q^m}$.

References

[1] G. A. Gibson, *Redundant disk arrays: Reliable, parallel secondary storage*. MIT press Cambridge, MA, 1992, vol. 368.

[2] M. Blaum, J. Hafner, and S. Hetzler, "Partial-MDS codes and their application to RAID type of architectures," *IEEE Trans. Inf. Theory*, vol. 59, no. 7, pp. 4510–4519, July 2013.

[3] M. Blaum, J. S. Plank, M. Schwartz, and E. Yaakobi, "Construction of partial MDS (PMDS) and Sector-Disk (SD) codes with two global parity symbols," *CoRR*, vol. abs/1401.4715, Aug. 2014.

[4] M. Blaum and J. S. Plank, "Construction of two SD codes," *CoRR*, vol. abs/1305.1221, May 2013. [Online]. Available: http://arxiv.org/abs/1305.1221

[5] J. S. Plank and M. Blaum, "Sector-Disk (SD) erasure codes for mixed failure modes in RAID systems," *Trans. Storage*, vol. 10, no. 1, pp. 4:1–4:17, Jan. 2014.

[6] M. Blaum and S. Hetzler, "Generalized concatenated types of codes for erasure correction," *CoRR*, vol. abs/1406.6270, Jun. 2014.

[7] M. Blaum, J. S. Plank, M. Schwartz, and E. Yaakobi, "Partial mds (pmds) and sector-disk (sd) codes that tolerate the erasure of two random sectors," in *Proc. 2014 IEEE International Symposium on Information Theory (ISIT 2014)*, Honolulu, HI, Jul. 2014.

[8] P. Gopalan, C. Huang, B. Jenkins, and S. Yekhanin, "Explicit maximally recoverable codes with locality," *IEEE Trans. Inf. Theory*, vol. 60, no. 9, pp. 5245–5256, Sept 2014.

[9] J. Chen, K. W. Shum, Q. Yu, and C. W. Sung, "Sector-disk codes and partial mds codes with up to three global parities," in *Proc. 2015 IEEE International Symposium on Information Theory (ISIT 2015)*, Hong Kong, Jun. 2015.

[10] M. Li and P. P. Lee, "STAIR codes: a general family of erasure codes for tolerating device and sector failures in practical storage systems," in *Proc. 12th USENIX Conference on File and Storage Technologies*, Santa Clara, CA, Feb. 2014.

[11] F. Oggier and A. Datta, "Self-repairing homomorphic codes for distributed storage systems," in *Proc. 2011 IEEE INFOCOM*, Shanghai, China, Apr. 2011.

[12] P. Gopalan, C. Huang, H. Simitci, and S. Yekhanin, "On the locality of codeword symbols," *IEEE Trans. Inf. Theory*, vol. 58, no. 11, pp. 6925–6934, Nov. 2012.

[13] A. S. Rawat, O. O. Koyluoglu, N. Silberstein, and S. Vishwanath, "Optimal locally repairable and secure codes for distributed storage systems," *IEEE Trans. Inf. Theory*, vol. 60, no. 1, pp. 212–236, Jan. 2014.

[14] I. Tamo and A. Barg, "A family of optimal locally recoverable codes," *IEEE Trans. Inf. Theory*, vol. 60, no. 8, pp. 4661–4676, Aug 2014.

[15] E. M. Gabidulin, "Theory of codes with maximum rank distance," *Problemy Peredachi Informatsii*, vol. 21, no. 1, pp. 3–16, 1985.

[16] R. M. Roth, "Maximum-rank array codes and their application to crisscross error correction," *IEEE Trans. Inf. Theory*, vol. 37, no. 2, pp. 328–336, Mar 1991.

[17] P. Delsarte, "Bilinear forms over a finite field, with applications to coding theory," *Journal of Combinatorial Theory, Series A*, vol. 25, no. 3, pp. 226–241, 1978.

[18] F. J. MacWilliams and N. J. A. Sloane, *The theory for error-correcting codes*. North-Holland, 1977.

[19] G. M. Kamath, N. Silberstein, N. Prakash, A. S. Rawat, V. Lalitha, O. O. Koyluoglu, P. V. Kumar, and S. Vishwanath, "Explicit MBR all-symbol locality codes," in *Proc. 2013 IEEE International Symposium on Information Theory (ISIT 2013)*, Istanbul, Turkey, Jul. 2013.

[20] N. Silberstein, A. S. Rawat, and S. Vishwanath, "Error resilience in distributed storage via rank-metric codes," in *Proc. 50th Annual Allerton Conference on communication, control and computing*, Monticello, IL, Oct. 2012.

[21] ——, "Error-correcting regenerating and locally repairable codes via rank-metric codes," *IEEE Trans. Inf. Theory*, vol. 61, no. 11, pp. 5765–5778, Nov 2015.

[22] O. Koyluoglu, A. Rawat, and S. Vishwanath, "Secure cooperative regenerating dodes for distributed storage systems," *IEEE Trans. Inf. Theory*, vol. 60, no. 9, pp. 5228–5244, Sep. 2014.

[23] P. Elias, "Error-free coding," *Information Theory, Transactions of the IRE Professional Group on*, vol. 4, no. 4, pp. 29–37, Sep. 1954.

[24] D. R. Stinson, *Combinatorial designs: construction and analysis*. Springer, 2004.