Constructions of Partial MDS Codes Over Small Fields

Ryan Gabrys, *Member, IEEE*, Eitan Yaakobi, *Senior Member, IEEE*, Mario Blaum, *Life Fellow, IEEE*, and Paul H. Siegel, *Life Fellow, IEEE*

***Abstract*— Partial MDS (PMDS) codes are a class of erasure-correcting array codes that combine local correction of the rows with global correction of the array. An $m \times n$ array code is called an $(r; s)$ PMDS code if each row belongs to an $[n, n-r, r+1]$ MDS code and the code can correct erasure patterns consisting of r erasures in each row together with s more erasures anywhere in the array. While a recent construction by Calis and Koyluoglu generates $(r; s)$ PMDS codes for all r and s, its field size is exponentially large. In this paper, a family of PMDS codes with field size $\mathcal{O}\left(\max\{m, n^{r+s}\}^s\right)$ is presented for the case where $r = \mathcal{O}(1), s = \mathcal{O}(1)$.**

***Index Terms*— Partial MDS codes, sector-disk codes, locally recoverable codes.**

I. Introduction

ERASURE-CORRECTING array codes are a class of codes mainly used to protect data in a Redundant Array of Independent Disks (RAID) architecture against catastrophic disk failures. Every column in the array corresponds to sectors from the same disk, such that a disk failure is modeled by a column erasure. Column failures are the widely studied model in the literature and different solutions such as RAID 5 and RAID 6 are designed specifically for this failure model. However, introducing solid state drives (SSDs) to the enterprise industry has brought new challenges in the design of a RAID architecture. Namely, existing solutions for RAID architectures are no longer adequate since SSDs may experience both disk failures and hard errors. To overcome this problem, Partial MDS (PMDS) codes were proposed in [1].

PMDS codes are designed to tolerate this mixed type of failures. We say that a code consisting of $m \times n$ arrays is an $(r; s)$ PMDS code if each row in the codeword arrays belongs to an $[n, n-r, r+1]$ MDS code, and furthermore, it is possible to correct s more arbitrary erasures in the array. The construction of PMDS codes for all r and s is a challenging task. In [1] the problem was solved when either $s = 1$ or $r = 1$. In [9], it was shown that it is possible to construct codes with field size $\mathcal{O}((mn)^{s/2})$. However, this construction only holds for the case where $m = 2$. In [2] codes were constructed for $s = 2$ and arbitrary r and recently in [3] a code construction was given for all r and s using Gabidulin codes. However, this construction requires a field of size $O(n^{mn})$. We note that in addition to this ongoing work, a generalization of PMDS codes was recently introduced in [8].

Our goal is constructing PMDS codes over fields of relatively small size. In fact, [6] gives a construction of PMDS codes over smaller fields for the case $r = 1$. In this paper, we extend the work of [6] to the case where $r > 1$. Our $(r; s)$ PMDS codes have field size $\mathcal{O}\left(\max\{m, n^{r+s}\}^s\right)$ when $r = \mathcal{O}(1), s = \mathcal{O}(1)$.

The rest of the paper is organized as follows. In Section II, we formally define PMDS codes. In Section III, we present a code, which we denote by $\mathcal{C}(\Gamma, \beta)$, using its parity check matrix, and we state necessary and sufficient conditions for $\mathcal{C}(\Gamma, \beta)$ to be a PMDS code. In Section IV we present a construction of matrices satisfying these conditions while using small fields. In Section V, we consider PMDS codes for certain parameters that improve upon the more general construction provided in Section IV. Lastly, Section VI concludes the paper.

II. Definitions and Preliminaries

For a prime power q, $\mathbb{F}_q$ is the finite field of size q. We assume in this paper that q is a power of two. A linear code of length n and dimension k over $\mathbb{F}_q$ will be denoted by $[n, k]_q$ or $[n, k, d]_q$, where d denotes its minimum distance. For an integer $n \geqslant 1$, the set $\{1, 2, \ldots, n\}$ will be denoted by $[n]$. We begin with the definition of PMDS codes and sector-disk codes from [1] and [2].

Definition 1: Let $\mathcal{C}$ be a linear $[mn, m(n-r)-s]$ code over a field such that when codewords are taken row-wise as $m \times n$ arrays, each row belongs to an $[n, n-r, r+1]$ MDS code. Given $(\sigma_1, \sigma_2, \ldots, \sigma_t)$ such that for $j \in [t]$, $\sigma_j \geqslant 1$, we say

Manuscript received October 11, 2017; revised August 27, 2018; accepted November 27, 2018. Date of publication December 28, 2018; date of current version May 20, 2019. This work was supported in part by the Center for Memory and Recording Research (CMRR) at UCSD, in part by NSF under Grant CCF-1405119 and Grant CCF-1619053, in part by the United States–Israel Binational Science Foundation (BSF) under Grant 2015816, and in part by the Israel Science Foundation under Grant 1624/14. R. Gabrys was supported by the NISE Program at SSC Pacific. E. Yaakobi was supported by CMRR at UCSD. This paper was presented at the IEEE International Symposium on Information Theory [5].

R. Gabrys is with the Spawar Systems Center, San Diego, San Diego, CA 92115 USA (e-mail: ryan.gabrys@navy.mil).

E. Yaakobi is with the Department of Computer Science, Technion—Israel Institute of Technology, Haifa 32000, Israel (e-mail: yaakobi@cs.technion.ac.il).

M. Blaum is with the IBM Research Division, Almaden Research Center, San Jose, CA 95120 USA (e-mail: Mario.Blaum@ibm.com).

P. H. Siegel is with the Department of Electrical and Computer Engineering, University of California at San Diego, La Jolla, CA 92093 USA, and also with the Center for Memory and Recording Research, University of California at San Diego, La Jolla, CA 92093 USA (e-mail: psiegel@ucsd.edu).

Communicated by P. Gopalan, Associate Editor for Coding Theory.

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TIT.2018.2890201

that $\mathcal{C}$ is an $(r;\sigma_1,\sigma_2,\ldots,\sigma_t)$-erasure correcting code if, for any $1 \leqslant i_1 < i_2 < \cdots < i_t \leqslant m$, $\mathcal{C}$ can correct up to $\sigma_j + r$ erasures in row i_j of an array in $\mathcal{C}$. We say that $\mathcal{C}$ is an $(r;s)$ partial-MDS (PMDS) code if, for every $(\sigma_1,\sigma_2,\ldots,\sigma_t)$ where $\sum_{j=1}^{t}\sigma_j = s$, $\mathcal{C}$ is an $(r;\sigma_1,\sigma_2,\ldots,\sigma_t)$-erasure correcting code.

We will also be interested in sector-disk codes, which are defined next. Sector-disk codes are a special type of PMDS code, and they are designed to handle the case where many erasures occur in the same set of r columns.

Definition 2: Let $\mathcal{C}$ be a linear $[mn, m(n-r)-s]$ code over a field such that when codewords are taken row-wise as $m \times n$ arrays, each row belongs to an $[n, n-r, r+1]$ MDS code. Then $\mathcal{C}$ is an $(r;s)$ sector-disk (SD) code if, for any $\ell_1,\ell_2,\ldots,\ell_r$ such that $0 \leqslant \ell_1 < \ell_2 < \cdots < \ell_r \leqslant n-1$, for any $(\sigma_1,\sigma_2,\ldots,\sigma_t)$ such that $\sigma_j \geqslant 1$ and $\sum_{j=1}^{t}\sigma_j = s$, and for any $i_1,i_2,\ldots,i_t$ such that $0 \leqslant i_1 < i_2 < \cdots < i_t \leqslant m$, $\mathcal{C}$ can correct up to $\sigma_j + r$ erasures in row i_j, $1 \leqslant j \leqslant t$, of an array in $\mathcal{C}$ provided that locations $\ell_1,\ldots,\ell_r$ in each of the rows i_j have been erased.

We note that PMDS codes are closely related to locally recoverable codes (LRCs) [11] in that both LRC and PMDS codes allow correcting errors locally by accessing a subset of the symbols across a codeword. For example, since by definition each row in an $(r;s)$ PMDS code is in an $[n, n-r, r+1]$ MDS code, if a row experiences at most r erasures, then it is possible to recover the erased symbols by only reading the symbols from that row. Hence, PMDS codes can be viewed as a special case of LRC codes. LRC codes that optimize the minimum distance were obtained in [11]. However, the requirement of Definition 1 is more stringent than the optimization of the minimum distance, and the codes from [11] are not PMDS in general.

We refer to a set $E_{(r;s)} \subseteq [m]\times[n]$ as an $(r;s)$-*erasure set* if it corresponds to an erasure pattern that can be corrected by an $(r;s)$ PMDS code. The next example illustrates a (2;2) PMDS code and a corresponding erasure set.

Example 1: Assume $m=3, n=5, r=2$, and $s=2$. Then we can interpret our codewords as arrays having the following form:

$$\boldsymbol{x} = \begin{pmatrix} c_1 & c_2 & c_3 & p_1^{(1)} & p_2^{(1)} \\ c_4 & c_5 & c_6 & p_3^{(1)} & p_4^{(1)} \\ c_7 & p_1^{(2)} & p_2^{(2)} & p_5^{(1)} & p_6^{(1)} \end{pmatrix}, \tag{1}$$

where there are 7 information symbols and 8 parity symbols. Given this setup, a PMDS code with these parameters is able to correct at most 2 erasures in any row using only the symbols from that row. Furthermore, it can correct 2 more erasures occurring anywhere in the codeword matrix.

Suppose $\boldsymbol{x}$ was stored, where $\boldsymbol{x}$ belongs to a $(2;2)$ PMDS code. Suppose, further that $\boldsymbol{x}$ experiences erasures resulting in the vector $\boldsymbol{y}$ shown below (the symbols in bold correspond to the erasures):

$$\boldsymbol{y} = \begin{pmatrix} \boldsymbol{c}_1 & c_2 & c_3 & p_1^{(1)} & \boldsymbol{p}_2^{(1)} \\ \boldsymbol{c}_4 & \boldsymbol{c}_5 & \boldsymbol{c}_6 & \boldsymbol{p}_3^{(1)} & p_4^{(1)} \\ c_7 & \boldsymbol{p}_1^{(2)} & p_2^{(2)} & \boldsymbol{p}_5^{(1)} & p_6^{(1)} \end{pmatrix}.$$

The erasures in $\boldsymbol{x}$ can be described by the following $(2;2)$-erasure set:

$$E_{(2;2)} = \{(1,1),(1,5),(2,1),(2,2),(2,3),(2,4),(3,2),(3,4)\}.$$

Let $\mathbb{F}_q$ be a field of size q and M be a positive integer. For a given basis of $\mathbb{F}_{q^M}$ over $\mathbb{F}_q$, every element $\boldsymbol{v} \in \mathbb{F}_{q^M}$ can also be represented as a length-M vector over $\mathbb{F}_q$. A set of elements $\boldsymbol{v}_1,\ldots,\boldsymbol{v}_N \in \mathbb{F}_{q^M}$ is said to be *linearly independent* over $\mathbb{F}_q$ if the length-M vectors representing these N elements are linearly independent over $\mathbb{F}_q$.

III. General Construction of PMDS Codes

A PMDS code can be seen either as an $m \times n$ array code or as a code of length mn. We will interchangeably use these two options to represent the code. From the context, it will be clear which option we are using. We assume that q is a prime power such that $q \geqslant n$ and $\mathbb{F}_{q^M}$ is an extension field of $\mathbb{F}_q$. Let $\beta \in \mathbb{F}_q$ be a primitive element, and let $\Gamma = (\alpha_{1,1},\alpha_{1,2},\ldots,\alpha_{m,n}) \in \left(\mathbb{F}_{q^M}\right)^{mn}$ be a sequence of mn distinct elements. Let $\mathcal{C}(\Gamma,\beta) \subseteq \left(\mathbb{F}_{q^M}\right)^{mn}$ be defined so that $\boldsymbol{x} = (x_{1,1},\ldots,x_{m,n}) \in \mathcal{C}(\Gamma,\beta)$ if and only if

$$0^{k-1}\cdot x_{i,1} + \sum_{j=2}^{n}\beta^{(k-1)(j-1)}x_{i,j} = 0, \quad \text{for } i \in [m],\ k \in [r] \tag{2}$$

$$\sum_{i=1}^{m}\sum_{j=1}^{n}\alpha_{i,j}^{q^{\ell-1}}x_{i,j} = 0, \quad \text{for } \ell \in [s], \tag{3}$$

where $0^0 = 1$. For shorthand, let $H^{(\beta)}$ be the matrix

$$H^{(\beta)} = \begin{pmatrix} 1 & 1 & 1 & \cdots & 1 \\ 0 & \beta & \beta^2 & \cdots & \beta^{n-1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & \beta^{r-1} & \beta^{2(r-1)} & \cdots & \beta^{(r-1)(n-1)} \end{pmatrix} \in (\mathbb{F}_q)^{r\times n},$$

and note that $0^{k-1}\cdot x_{i,1} + \sum_{j=2}^{n}\beta^{(k-1)(j-1)}x_{i,j} = 0$ for all $k \in [r]$ is equivalent to $H^{(\beta)}\cdot(x_{i,1},\ldots,x_{i,n})^T = \boldsymbol{0}$. Our main aim in this section is to consider necessary and sufficient conditions for $\mathcal{C}(\Gamma,\beta)$ to be a PMDS code.

In the rest of this paper, we will refer to an $(r;0)$-erasure set with exactly r erasures in each row as an $(r;0)$-erasure vector $S \in ([n]^r)^m$ and denote it by $S = (\boldsymbol{s}_1,\ldots,\boldsymbol{s}_m)$, where for $i \in [m]$, $\boldsymbol{s}_i = (s_{i,1},\ldots,s_{i,r}) \in [n]^r$, and $s_{i,1} < s_{i,2} < \cdots < s_{i,r}$. For example, suppose we store the vector $\boldsymbol{x}$ from (1) and we receive

$$\boldsymbol{y} = \begin{pmatrix} \boldsymbol{c}_1 & \boldsymbol{c}_2 & c_3 & p_1^{(1)} & p_2^{(1)} \\ c_4 & c_5 & \boldsymbol{c}_6 & \boldsymbol{p}_3^{(1)} & p_4^{(1)} \\ \boldsymbol{c}_7 & p_1^{(2)} & p_2^{(2)} & p_5^{(1)} & \boldsymbol{p}_6^{(1)} \end{pmatrix},$$

so that a $(2;0)$-erasure vector occurred, marked by the symbols in bold. Then we represent this erasure pattern using the vector $S = (\boldsymbol{s}_1,\boldsymbol{s}_2,\boldsymbol{s}_3) = \big((1,2),(3,4),(1,5)\big)$.

For a matrix A with n columns and a vector $\boldsymbol{s} = (s_1,\ldots,s_r) \in [n]^r$, where $s_1 < s_2 < \cdots < s_r$, let $A_{\boldsymbol{s}}$ be the submatrix of A given by the columns indexed by the set $\{s_1,s_2,\ldots,s_r\}$. Similarly, for a vector $\boldsymbol{u}$ of length n,

$\boldsymbol{u}_{\boldsymbol{s}}$ is the sub-vector of $\boldsymbol{u}$ given by the indices of the set $\{s_1, s_2, \ldots, s_r\}$. By a slight abuse of notation, for a vector $\boldsymbol{s} = (s_1, \ldots, s_r) \in [n]^r$, we denote by $\overline{\boldsymbol{s}} \in [n]^{n-r}$ the vector

$$\overline{\boldsymbol{s}} = (\overline{s}_1, \overline{s}_2, \ldots, \overline{s}_{n-r}), \tag{4}$$

which contains, in increasing order, all the values in $[n]$ that do not appear in $\boldsymbol{s}$. For example, if $\boldsymbol{s} = (1, 3, 7, 8) \in [9]^4$, then $\overline{\boldsymbol{s}} = (2, 4, 5, 6, 9)$.

Next, we state necessary and sufficient conditions for the construction of the code $\mathcal{C}(\Gamma, \beta)$ to generate PMDS codes. Here we use the notation $\boldsymbol{x} = (\boldsymbol{x}_1, \ldots, \boldsymbol{x}_m)$ to represent our codewords, where for $i \in [m]$, $\boldsymbol{x}_i$ is a length-n row vector. We also assume that β and the sequence Γ are given so they determine the code $\mathcal{C}(\Gamma, \beta)$. We start with the following claim.

Claim 1: If $\boldsymbol{x} = (\boldsymbol{x}_1, \ldots, \boldsymbol{x}_m) \in \mathcal{C}(\Gamma, \beta)$, then for any $(r; 0)$-erasure vector $S = (\boldsymbol{s}_1, \ldots, \boldsymbol{s}_m) \in ([n]^r)^m$, the following equality holds for all $i \in [m]$:

$$(\boldsymbol{x}_i)^T_{\boldsymbol{s}_i} = ((H^{(\beta)}_{\boldsymbol{s}_i})^{-1} \cdot H^{(\beta)}_{\overline{\boldsymbol{s}}_i}) \cdot (\boldsymbol{x}_i)^T_{\overline{\boldsymbol{s}}_i}.$$

Proof: Since $H^{(\beta)}$ is a parity check matrix of an $[n, n-r]$ MDS code, any r columns of $H^{(\beta)}$ are linearly independent, and thus the matrix $H^{(\beta)}_{\boldsymbol{s}_i} \in \mathbb{F}_q^{r \times r}$ has an inverse matrix $(H^{(\beta)}_{\boldsymbol{s}_i})^{-1}$. According to (2), we have

$$\mathbf{0}^T = H^{(\beta)} \cdot \boldsymbol{x}_i^T = H^{(\beta)}_{\boldsymbol{s}_i} \cdot (\boldsymbol{x}_i)^T_{\boldsymbol{s}_i} + H^{(\beta)}_{\overline{\boldsymbol{s}}_i} \cdot (\boldsymbol{x}_i)^T_{\overline{\boldsymbol{s}}_i}.$$

After multiplying both sides on the left by the matrix $(H^{(\beta)}_{\boldsymbol{s}_i})^{-1}$ we get

$$\begin{aligned}\mathbf{0}^T &= (H^{(\beta)}_{\boldsymbol{s}_i})^{-1} \cdot H^{(\beta)}_{\boldsymbol{s}_i} \cdot (\boldsymbol{x}_i)^T_{\boldsymbol{s}_i} + (H^{(\beta)}_{\boldsymbol{s}_i})^{-1} \cdot H^{(\beta)}_{\overline{\boldsymbol{s}}_i} \cdot (\boldsymbol{x}_i)^T_{\overline{\boldsymbol{s}}_i} \\ &= (\boldsymbol{x}_i)^T_{\boldsymbol{s}_i} + ((H^{(\beta)}_{\boldsymbol{s}_i})^{-1} \cdot H^{(\beta)}_{\overline{\boldsymbol{s}}_i}) \cdot (\boldsymbol{x}_i)^T_{\overline{\boldsymbol{s}}_i}\end{aligned}$$

which implies the statement in the claim. ∎

Next, we recall the notion of t-wise independence [6].

*Definition 3: Let $\mathbb{F}$ be a field. A multiset $\mathcal{S} \subseteq \mathbb{F}$ is **t-wise independent** over a subfield $\mathbb{F}' \subseteq \mathbb{F}$ if for every $T \subseteq \mathcal{S}$ such that $|T| \leqslant t$, the elements of T are linearly independent over $\mathbb{F}'$.*

For any $(r; 0)$-erasure vector $S = (\boldsymbol{s}_1, \ldots, \boldsymbol{s}_m) \in ([n]^r)^m$, let $\mathcal{C}^S$ be the code obtained by puncturing $\mathcal{C}$ in the positions corresponding to the erasure locations referenced by the vector S. The next claim is straightforward.

Claim 2: A code $\mathcal{C}$ is an $(r; s)$ PMDS code if and only if $\mathcal{C}^S$ is an $[m(n-r), m(n-r)-s]$ MDS code for any $(r; 0)$-erasure vector S.

The following lemma is well known. For completeness, we provide a proof in Appendix.

Lemma 1 (cf. [7]): Let $\mathbb{F}_{q^M}$ be an extension field of $\mathbb{F}_q$, and $G \in \left(\mathbb{F}_{q^M}\right)^{s \times s}$ be the following matrix

$$G = \begin{pmatrix} \alpha_1 & \alpha_2 & \cdots & \alpha_s \\ \alpha_1^q & \alpha_2^q & \cdots & \alpha_s^q \\ \vdots & \vdots & \ddots & \vdots \\ \alpha_1^{q^{s-1}} & \alpha_2^{q^{s-1}} & \cdots & \alpha_s^{q^{s-1}} \end{pmatrix},$$

where $\alpha_1, \alpha_2, \ldots, \alpha_s \in \mathbb{F}_{q^M}$. Then G has rank s if and only if the elements $\alpha_1, \alpha_2, \ldots, \alpha_s$ are linearly independent over $\mathbb{F}_q$.

Let $S = (\boldsymbol{s}_1, \ldots, \boldsymbol{s}_m) \in ([n]^r)^m$ be an $(r; 0)$-erasure vector. For $i \in [m]$, let $A^{(\beta)}_{\boldsymbol{s}_i} \in (\mathbb{F}_q)^{r \times (n-r)}$ be the matrix

$$A^{(\beta)}_{\boldsymbol{s}_i} = (H^{(\beta)}_{\boldsymbol{s}_i})^{-1} \cdot H^{(\beta)}_{\overline{\boldsymbol{s}}_i}.$$

Let $\boldsymbol{\alpha}_i$ be the vector $\boldsymbol{\alpha}_i = (\alpha_{i,1}, \alpha_{i,2}, \ldots, \alpha_{i,n})$, and denote $\boldsymbol{\alpha}_{\boldsymbol{s}_i} = (\alpha_{i,s_{i,1}}, \alpha_{i,s_{i,2}}, \ldots, \alpha_{i,s_{i,r}})$, where $\boldsymbol{s}_i = (s_{i,1}, s_{i,2}, \ldots, s_{i,r}) \in [n]^r$, and $\boldsymbol{\alpha}_{\overline{\boldsymbol{s}}_i} = (\alpha_{i,\overline{s}_{i,1}}, \alpha_{i,\overline{s}_{i,2}}, \ldots, \alpha_{i,\overline{s}_{i,n-r}})$, where $\overline{\boldsymbol{s}}_i = (\overline{s}_{i,1}, \overline{s}_{i,2}, \ldots, \overline{s}_{i,n-r})$ is the vector defined in (4). Lastly, for $i \in [m]$ we denote by $\boldsymbol{\gamma}_{\boldsymbol{s}_i} = (\gamma_{\boldsymbol{s}_i,1}, \gamma_{\boldsymbol{s}_i,2}, \ldots, \gamma_{\boldsymbol{s}_i,n-r}) \in \left(\mathbb{F}_{q^M}\right)^{n-r}$ the vector

$$\boldsymbol{\gamma}_{\boldsymbol{s}_i} = \boldsymbol{\alpha}_{\boldsymbol{s}_i} \cdot A^{(\beta)}_{\boldsymbol{s}_i} + \boldsymbol{\alpha}_{\overline{\boldsymbol{s}}_i}. \tag{5}$$

We are now ready to prove a necessary and sufficient condition for $\mathcal{C}(\Gamma, \beta)$ to be a PMDS code.

Lemma 2: The code $\mathcal{C}(\Gamma, \beta)$ is an $(r; s)$ PMDS code if and only if for any $(r; 0)$-erasure vector $S = (\boldsymbol{s}_1, \ldots, \boldsymbol{s}_m) \in ([n]^r)^m$, the set $T(S) = \{\gamma_{\boldsymbol{s}_i,j}\}_{i \in [m], j \in [n-r]}$ is s-wise independent over $\mathbb{F}_q$.

Proof: According to Claim 2, the code $\mathcal{C}(\Gamma, \beta)$ is an $(r; s)$ PMDS code if and only if for any $(r; 0)$-erasure vector $S = (\boldsymbol{s}_1, \ldots, \boldsymbol{s}_m) \in ([n]^r)^m$, the code $\mathcal{C}(\Gamma, \beta)^S$ is an $[m(n-r), m(n-r)-s]$ MDS code. Let $S = (\boldsymbol{s}_1, \ldots, \boldsymbol{s}_m) \in ([n]^r)^m$ be an $(r; 0)$-erasure vector. From (3), we see that for $\ell \in [s]$, the global parity constraints imply

$$0 = \sum_{i=1}^{m} \sum_{j=1}^{n} \alpha_{i,j}^{q^{\ell-1}} x_{i,j} = \sum_{i=1}^{m} \boldsymbol{\alpha}_i^{q^{\ell-1}} \boldsymbol{x}_i^T,$$

where $\boldsymbol{\alpha}_i^{q^{\ell-1}} = (\alpha_{i,1}^{q^{\ell-1}}, \alpha_{i,2}^{q^{\ell-1}}, \ldots, \alpha_{i,n}^{q^{\ell-1}})$ and $\boldsymbol{x}_i = (x_{i,1}, \ldots, x_{i,n})$ for $i \in [m]$. We also denote $\boldsymbol{\alpha}_{\boldsymbol{s}_i}^{q^{\ell-1}} = \left(\alpha_{i,s_{i,1}}^{q^{\ell-1}}, \alpha_{i,s_{i,2}}^{q^{\ell-1}}, \ldots, \alpha_{i,s_{i,r}}^{q^{\ell-1}}\right)$ and define $\boldsymbol{\alpha}_{\overline{\boldsymbol{s}}_i}^{q^{\ell-1}}$ similarly. From Claim 1, we can write $(\boldsymbol{x}_i)^T_{\boldsymbol{s}_i} = A^{(\beta)}_{\boldsymbol{s}_i} \cdot (\boldsymbol{x}_i)^T_{\overline{\boldsymbol{s}}_i}$ and therefore we have that

$$\begin{aligned} 0 &= \sum_{i=1}^{m} \boldsymbol{\alpha}_i^{q^{\ell-1}} \boldsymbol{x}_i^T \\ &= \sum_{i=1}^{m} \left(\boldsymbol{\alpha}_{\boldsymbol{s}_i}^{q^{\ell-1}} \cdot (\boldsymbol{x}_i)^T_{\boldsymbol{s}_i} + \boldsymbol{\alpha}_{\overline{\boldsymbol{s}}_i}^{q^{\ell-1}} \cdot (\boldsymbol{x}_i)^T_{\overline{\boldsymbol{s}}_i} \right) \\ &= \sum_{i=1}^{m} \left(\boldsymbol{\alpha}_{\boldsymbol{s}_i}^{q^{\ell-1}} \cdot A^{(\beta)}_{\boldsymbol{s}_i} \cdot (\boldsymbol{x}_i)^T_{\overline{\boldsymbol{s}}_i} + \boldsymbol{\alpha}_{\overline{\boldsymbol{s}}_i}^{q^{\ell-1}} \cdot (\boldsymbol{x}_i)^T_{\overline{\boldsymbol{s}}_i} \right) \\ &= \sum_{i=1}^{m} \left(\boldsymbol{\alpha}_{\boldsymbol{s}_i}^{q^{\ell-1}} \cdot A^{(\beta)}_{\boldsymbol{s}_i} + \boldsymbol{\alpha}_{\overline{\boldsymbol{s}}_i}^{q^{\ell-1}} \right) \cdot (\boldsymbol{x}_i)^T_{\overline{\boldsymbol{s}}_i}, \\ &\overset{(a)}{=} \sum_{i=1}^{m} \left(\boldsymbol{\alpha}_{\boldsymbol{s}_i} \cdot A^{(\beta)}_{\boldsymbol{s}_i} + \boldsymbol{\alpha}_{\overline{\boldsymbol{s}}_i} \right)^{q^{\ell-1}} \cdot (\boldsymbol{x}_i)^T_{\overline{\boldsymbol{s}}_i}, \\ &= \sum_{i=1}^{m} \left(\boldsymbol{\gamma}_{\boldsymbol{s}_i} \right)^{q^{\ell-1}} \cdot (\boldsymbol{x}_i)^T_{\overline{\boldsymbol{s}}_i}, \end{aligned}$$

where equality (a) holds since the matrix $A^{(\beta)}_{\boldsymbol{s}_i}$ is over $\mathbb{F}_q$. We conclude that a parity check matrix for the code $\mathcal{C}(\Gamma, \beta)^S$ will be given by (6), as shown at the top of the next page. Therefore, from Lemma 1, the code $\mathcal{C}(\Gamma, \beta)^S$ is an MDS code if and only if the set of $m(n-r)$ elements given by the entries of the vectors $\boldsymbol{\gamma}_{\boldsymbol{s}_i}$ for $i \in [m]$ is s-wise independent over $\mathbb{F}_q$. ∎

$$\begin{pmatrix} \gamma^1_{\boldsymbol{s}_1,1} & \gamma^1_{\boldsymbol{s}_1,2} & \gamma^1_{\boldsymbol{s}_1,3} & \cdots & \gamma^1_{\boldsymbol{s}_1,n-r} & \gamma^1_{\boldsymbol{s}_2,1} & \gamma^1_{\boldsymbol{s}_2,2} & \cdots & \gamma^1_{\boldsymbol{s}_2,n-r} & \cdots & \cdots & \cdots & \gamma^1_{\boldsymbol{s}_m,1} & \cdots & \gamma^1_{\boldsymbol{s}_m,n-r} \\ \gamma^q_{\boldsymbol{s}_1,1} & \gamma^q_{\boldsymbol{s}_1,2} & \gamma^q_{\boldsymbol{s}_1,3} & \cdots & \gamma^q_{\boldsymbol{s}_1,n-r} & \gamma^q_{\boldsymbol{s}_2,1} & \gamma^q_{\boldsymbol{s}_2,2} & \cdots & \gamma^q_{\boldsymbol{s}_2,n-r} & \cdots & \cdots & \cdots & \gamma^q_{\boldsymbol{s}_m,1} & \cdots & \gamma^q_{\boldsymbol{s}_m,n-r} \\ \gamma^{q^2}_{\boldsymbol{s}_1,1} & \gamma^{q^2}_{\boldsymbol{s}_1,2} & \gamma^{q^2}_{\boldsymbol{s}_1,3} & \cdots & \gamma^{q^2}_{\boldsymbol{s}_1,n-r} & \gamma^{q^2}_{\boldsymbol{s}_2,1} & \gamma^{q^2}_{\boldsymbol{s}_2,2} & \cdots & \gamma^{q^2}_{\boldsymbol{s}_2,n-r} & \cdots & \cdots & \cdots & \gamma^{q^2}_{\boldsymbol{s}_m,1} & \cdots & \gamma^{q^2}_{\boldsymbol{s}_m,n-r} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots & \cdots & \cdots & \cdots & \vdots & \ddots & \vdots \\ \gamma^{q^{s-1}}_{\boldsymbol{s}_1,1} & \gamma^{q^{s-1}}_{\boldsymbol{s}_1,2} & \gamma^{q^{s-1}}_{\boldsymbol{s}_1,3} & \cdots & \gamma^{q^{s-1}}_{\boldsymbol{s}_1,n-r} & \gamma^{q^{s-1}}_{\boldsymbol{s}_2,1} & \gamma^{q^{s-1}}_{\boldsymbol{s}_2,2} & \cdots & \gamma^{q^{s-1}}_{\boldsymbol{s}_2,n-r} & \cdots & \cdots & \cdots & \gamma^{q^{s-1}}_{\boldsymbol{s}_m,1} & \cdots & \gamma^{q^{s-1}}_{\boldsymbol{s}_m,n-r} \end{pmatrix}. \quad (6)$$

Notice that for the case $r = 1$, we have the following corollary which is similar to the result from [6].

Corollary 3: The code $\mathcal{C}(\Gamma, \beta)$ is a $(1; s)$ PMDS code if and only if for any $(1; 0)$-erasure vector $S = (s_1, \ldots, s_m) \in [n]^m$, the set

$$T(S) = \left\{\alpha_{i,j} + \alpha_{i,s_i}\right\}_{i\in[m],j\in[n]\setminus\{s_i\}}$$

is s-wise independent over $\mathbb{F}_2$.

Proof: For the case $r = 1$, notice that the matrix $H^{(\beta)}$ is simply a single row of n ones, that is $H^{(\beta)} = (1, 1, \ldots, 1)$. Let $S = (s_1, \ldots, s_m) \in [n]^m$ be a $(1; 0)$-erasure vector. (Note that each s_i, $i \in [m]$ represents a single position.) Then, for all $i \in [m]$ the matrix $A^{(\beta)}_{\boldsymbol{s}_i}$ becomes a single row of $n - 1$ ones, and the vector $\boldsymbol{\gamma}_{s_i} = (\gamma_{i,1}, \gamma_{i,2}, \ldots, \gamma_{i,n-1})$ becomes

$$\begin{aligned} \boldsymbol{\gamma}_{s_i} &= \boldsymbol{\alpha}_{s_i} \cdot A^{(\beta)}_{\boldsymbol{s}_i} + \boldsymbol{\alpha}_{\overline{\boldsymbol{s}}_i} \\ &= \alpha_{i,s_i} \cdot (1, 1, \ldots, 1) + (\alpha_{i,1}, \alpha_{i,2}, \ldots, \alpha_{i,s_i-1}, \\ &\quad \alpha_{i,s_i+1}, \ldots, \alpha_{i,n}) \\ &= (\alpha_{i,1} + \alpha_{i,s_i}, \ldots, \alpha_{i,s_i-1} + \alpha_{i,s_i}, \alpha_{i,s_i+1} \\ &\quad + \alpha_{i,s_i}, \ldots, \alpha_{i,n} + \alpha_{i,s_i}). \end{aligned}$$

Therefore, the set $T(S) = \{\gamma_{s_i,j}\}_{i\in[m],j\in[n-1]}$ is given by

$$T(S) = \left\{\alpha_{i,j} + \alpha_{i,s_i}\right\}_{i\in[m],j\in[n]\setminus\{s_i\}}.$$

■

The next corollary will be used in Section V-B2 to construct $(1; s)$-SD codes.

Corollary 4: The code $\mathcal{C}(\Gamma, \beta)$ is a $(1; s)$-SD code if and only if for any $k \in [n]$, the set

$$T = \left\{\alpha_{i,j} + \alpha_{i,k}\right\}_{i\in[m],j\in[n]\setminus\{k\}}$$

is s-wise independent over $\mathbb{F}_2$.

According to (2) and (3), the code $\mathcal{C}(\Gamma, \beta)$ is determined by the choice of $\beta \in \mathbb{F}_q$, the primitive element in $\mathbb{F}_q$, and the sequence $\Gamma = (\alpha_{1,1}, \alpha_{1,2}, \ldots, \alpha_{m,n}) \in \left(\mathbb{F}_{q^M}\right)^{mn}$ over the extension field $\mathbb{F}_{q^M}$. Hence, the necessary and sufficient condition given in Lemma 2 for the code $\mathcal{C}(\Gamma, \beta)$ to be an $(r; s)$ PMDS code involves both β and Γ. However, in our construction, β can be chosen to be any primitive element in $\mathbb{F}_q$. The next corollary exploits this fact to simplify the task of finding a suitable set Γ defined over a small field. It replaces the necessary and sufficient condition of Lemma 2 with a sufficient condition involving the s-wise independence of sets of elements whose definition is independent of the choice of β. In the next section we discuss how to choose sequences Γ that satisfy this condition over small fields.

Corollary 5: The code $\mathcal{C}(\Gamma, \beta)$ is an $(r; s)$ PMDS code if for any $(r; 0)$-erasure vector $S = (\boldsymbol{s}_1, \ldots, \boldsymbol{s}_m) \in ([n]^r)^m$ and any m full-rank matrices $V_1, \ldots, V_m \in \mathbb{F}_q^{r\times(n-r)}$, the set

$$\widehat{T}(S) = \{\widehat{\gamma}_{\boldsymbol{s}_i,j}\}_{i\in[m],j\in[n-r]},$$

is s-wise independent over $\mathbb{F}_q$, where for $i \in [m]$

$$\widehat{\boldsymbol{\gamma}}_{\boldsymbol{s}_i} = (\widehat{\gamma}_{\boldsymbol{s}_i,1}, \widehat{\gamma}_{\boldsymbol{s}_i,2}, \ldots, \widehat{\gamma}_{\boldsymbol{s}_i,n-r}) = \boldsymbol{\alpha}_{\boldsymbol{s}_i} \cdot V_i + \boldsymbol{\alpha}_{\overline{\boldsymbol{s}}_i} \in \mathbb{F}_q^{n-r}. \quad (7)$$

Proof: The proof follows immediately from Lemma 2 by setting $V_i = A^{(\beta)}_{\boldsymbol{s}_i}$. ■

IV. PMDS Codes Over Small Fields

Note again that the condition in Corollary 5 is a condition only on the sequence Γ. Thus we say that a sequence Γ is an $(r; s)$-*PMDS sequence* of size $m \times n$ if it satisfies the condition from Corollary 5. As mentioned above, our goal is to find $(r; s)$-PMDS sequences of size $m \times n$ over a field with the smallest possible size. We denote by $\phi(n, d, q)$ the minimum redundancy of an $[n, k, d]_q$ code. In the following, we assume that $s = O(1)$, $r = O(1)$. We will make use of the following useful lemma, based upon BCH codes and Reed-Solomon codes, in subsequent derivations.

Lemma 6 (cf. [10, Problem 8.9; Ch. 8, pp. 203]): Suppose $n = q^\mu - 1$, and $d \leqslant q^{\lceil \frac{\mu}{2} \rceil} + 1$. Then there exists a BCH code where $\phi(n, d, q)$ satisfies

$$\phi(n, d, q) \leqslant \min\left\{1 + \left\lceil \left(1 - \frac{1}{q}\right)(d - 2)\right\rceil \cdot \mu, \left\lceil \left(1 - \frac{1}{q}\right)(d - 1)\right\rceil \cdot \mu\right\}.$$

For the case where $n \leqslant q$ there exists an (extended) Reed-Solomon code where

$$\phi(n, d, q) \leqslant d - 1.$$

A. First Construction of PMDS Sequences Over Small Fields

Our first result on PMDS codes with small field size, which is an extension of the work in [6] for the case where $r > 1$, is stated in the next lemma. The purpose of this construction is to illustrate one straightforward way to construct the sequence Γ from $\mathcal{C}(\Gamma, \beta)$. The construction given in the next section improves upon the construction in this section in many cases, and it represents our main result.

Lemma 7: If the set of elements in the sequence $\Gamma = (\alpha_{1,1}, \ldots, \alpha_{m,n})$ is $(r + 1)$s-wise independent over $\mathbb{F}_q$,

then it is an $(r;s)$-PMDS sequence. Hence, there exists an $(r;s)$-PMDS sequence over a field $\mathbb{F}_{q^M}$, where

$$M = \phi(mn, (r+1)s+1, q), \tag{8}$$

and $q \geqslant n$ is a prime power. In particular, there exists an $(r;s)$ PMDS code with field size at most $n(2mn)^{(r+1)s-1}$ when $q = n$ and $2mn$ is a power of q.

Proof: The result follows by noting that, according to (7), for any $(r;0)$-erasure vector $S = (\boldsymbol{s}_1, \ldots, \boldsymbol{s}_m) \in ([n]^r)^m$ and any m full-rank matrices $V_1, \ldots, V_m \in \mathbb{F}_q^{r\times(n-r)}$, each element in the set $\widehat{T}(S) = \{\widehat{\gamma}_{\boldsymbol{s}_i,j}\}_{i\in[m],j\in[n-r]}$, from Corollary 5 can be written as a linear combination over $\mathbb{F}_q$ of exactly $r+1$ elements from Γ. This follows since according to Corollary 5, we can write

$$\widehat{\boldsymbol{\gamma}}_{\boldsymbol{s}_i} = (\widehat{\gamma}_{\boldsymbol{s}_i,1}, \widehat{\gamma}_{\boldsymbol{s}_i,2}, \ldots, \widehat{\gamma}_{\boldsymbol{s}_i,n-r}) = \boldsymbol{\alpha}_{\boldsymbol{s}_i} \cdot V_i + \boldsymbol{\alpha}_{\overline{\boldsymbol{s}}_i} \in \mathbb{F}_q^{n-r}.$$

for $i \in [m]$. Since $\boldsymbol{s}_i \in [n]^r$, it follows that $\widehat{\gamma}_{\boldsymbol{s}_i,j}$ is the result of linearly combining at most $r+1$ elements from Γ. Therefore, any s elements in $\widehat{T}(S)$ are linear combinations over $\mathbb{F}_q$ of at most $(r+1)s$ elements from Γ. Thus, if the set of elements in the sequence Γ is $(r+1)s$-wise independent over $\mathbb{F}_q$, it follows that the set $\widehat{T}(S)$ is s-wise independent over $\mathbb{F}_q$ as well, and hence Γ is an $(r;s)$-PMDS sequence.

A construction of such a sequence Γ whose elements are $(r+1)s$-wise independent over $\mathbb{F}_q$ is as follows. Let H be a parity check matrix of an $[mn, k, (r+1)s+1]_q$ code C with redundancy $\rho = mn - k$, where q is the smallest prime power such that $q \geqslant n$. We set Γ to consist of the mn columns of H, which are mn elements over the field $\mathbb{F}_{q^\rho}$. Since the minimum distance of the code C is $(r+1)s+1$, every $(r+1)s$ columns of H are linearly independent over $\mathbb{F}_q$ and so are every $(r+1)s$ elements from the sequence Γ. By choosing a code C where $\rho = \phi(mn, (r+1)s+1, q)$ we get the result stated in (8). From the bound in Lemma 6, we get that there exists a code of length $2mn-1$, minimum distance $(r+1)s+1$, and redundancy $1 + \log_q(2mn)((r+1)s-1)$. By shortening the code, we arrive at a code of length mn with minimum distance $(r+1)s+1$, which, since $q = n$, implies there is a PMDS code with field size at most $n(2mn)^{(r+1)s-1}$ for the case where $2mn$ is a power of q. ∎

According to Lemma 7, we already deduce that there exists an $(r;s)$ PMDS code over a field which is polynomial in mn, in contrast to the exponential field size reported in [3]. We now present a further improvement on the field size.

B. Second Construction of PMDS Sequences Over Small Fields

In this section, we give a construction of $(r;s)$-PMDS sequences with a smaller field size than that achieved in Lemma 7 for many cases. Recall from the previous subsection that we designed our sequence Γ (used in $\mathcal{C}(\Gamma, \beta)$) so that each element in Γ is $(r+1)s$-wise independent over $\mathbb{F}_q$. As we will discuss in more detail, it turns out that this condition in many cases is too strong since each element in the set $\widehat{T}(S)$ from Corollary 5 can be written as a linear combination of at most $r+1$ elements that are contained *in the same row* as each other (when we interpret our codewords as $m \times n$ arrays). The construction discussed in this section exploits this fact and results in codebooks with small field sizes.

Our construction uses tensor product codes, which were introduced in [12]. Our approach will be to use the columns of the parity check matrix of a tensor product code as the elements of the sequence Γ. As we will soon discuss, tensor product codes have the special property that they can correct structured errors, which implies that certain (structured) subsets of columns in a parity check matrix of the code are linearly independent. In the following, we detail how to use this property to design the sequence Γ.

Let us first review the definition of tensor product codes, focusing on the case of erasure correction. A code $\mathcal{C} \subseteq (\mathbb{F}_q)^{m\times n}$ is called an $[m, n; t_1, t_2]$ *erasure-correcting code* if it can correct any erasure pattern of the form $\mathbf{E} = (E_1, \ldots, E_m)$, where $E_i \subseteq [n]$, for all $i \in [m]$, and

1) $|\{i : E_i \neq \emptyset\}| \leqslant t_1$,
2) $|E_i| \leqslant t_2$, for all $i \in [m]$.

In other words, such a code is required to correct erasures in at most t_1 rows, while in each row there are at most t_2 erasures. Such an erasure pattern will be called a $(t_1; t_2)$-*erasure pattern*.

The following theorem draws a connection between $[m, n; t_1, t_2]$ erasure-correcting codes and $(r;s)$-PMDS sequences and gives the main result of this section. In Corollary 10, we give an upper bound on the field size of an $(r;s)$-PMDS code which is a consequence of the theorem below.

Theorem 8: Let $\mathcal{C}_{TP}$ be an $[m, n; s, r+s]$ erasure-correcting code over $\mathbb{F}_q$ with redundancy ρ and parity check matrix $H_{TP} = (\alpha_{1,1}, \ldots, \alpha_{m,n}) \in (\mathbb{F}_{q^\rho})^{mn}$. Then, the sequence $\Gamma_{TP} = (\alpha_{1,1}, \ldots, \alpha_{m,n})$ is an $(r;s)$-PMDS sequence of size $m \times n$.

Proof: Assume $S = (\boldsymbol{s}_1, \ldots, \boldsymbol{s}_m) \in ([n]^r)^m$ is an $(r;0)$-erasure vector and $V_1, \ldots, V_m \in \mathbb{F}_q^{r\times(n-r)}$ are m full-rank matrices which determine the set $\widehat{T}(S) = \{\widehat{\gamma}_{\boldsymbol{s}_i,j}\}_{i\in[m],j\in[n-r]}$ as specified in (7). We will show that $\widehat{T}(S)$ is s-wise independent over $\mathbb{F}_q$. Any s elements from the set $\widehat{T}(S)$ can be expressed as a linear combination of elements from the sequence Γ_{TP}. For $i \in [m]$, we denote by $E_i \subseteq [n]$ the locations of elements which belong to this linear combination from the i-th row of Γ_{TP}, that is, from the elements $\alpha_{i,1}, \ldots, \alpha_{i,n}$. Consider the vector $\mathbf{E} = (E_1, \ldots, E_m)$ and note that it is an $(s; r+s)$-erasure pattern. Since the code $\mathcal{C}_{TP}$ is an $[m, n; s, r+s]$ erasure-correcting code, every collection of columns from the parity matrix H_{TP} which correspond to an $(s; r+s)$-erasure pattern is linearly independent. Hence, the set of elements from Γ_{TP} at locations in $\cup_{i=1}^m E_i$ is linearly independent over $\mathbb{F}_q$ and, therefore, so is every set of s elements from $\widehat{T}(S)$. ∎

According to Theorem 8, the task of finding $(r;s)$-PMDS sequences can now be translated to the construction of $[m, n; s, r+s]$ erasure-correcting codes over $\mathbb{F}_q$ with small redundancy ρ. We review the construction of such codes as presented in [12]. Specifically, the parity check matrix for an $[m, n; s, r+s]$ erasure-correcting code $\mathcal{C}_{TP}$ over $\mathbb{F}_q$ (where $q = n$) is given as follows:

1) Assume that $H' \in \mathbb{F}_q^{(r+s)\times n}$ is a parity check matrix for an $[n, n-r-s]_q$ MDS code.
2) Assume that $H'' \in \mathbb{F}_{q^{r+s}}^{R\times m}$ is a parity check matrix for an $[m, m-R, s+1]_{q^{r+s}}$ code.
3) Representing every column in H' as an element in $\mathbb{F}_{q^{r+s}}$, we let $\mathcal{C}_{TP}$ be a code over $\mathbb{F}_q$ with the parity check matrix $H'' \otimes H'$ formed by taking the tensor product of H'' and H'.

The existence of an $[n, n-r-s]_q$ MDS code in 1) where $q = n$ follows from the second statement regarding the existence of extended Reed-Solomon codes in Lemma 6.

Theorem 9 (cf. [12]): The code $\mathcal{C}_{TP}$ is an $[m, n; s, r+s]$ erasure-correcting code over $\mathbb{F}_q$ with redundancy $\rho = (r+s)R$.

We are now ready to provide an upper bound on the field size for $(r; s)$ PMDS codes by providing an upper bound on R using the bound on the redundancy of BCH codes from Lemma 6.

Corollary 10: There exists an $(r; s)$ PMDS code with field size $q^{(r+s)\phi(m,s+1,q^{r+s})}$, which is at most $\mathcal{O}\Big(\max\{m, n^{r+s}\}^s\Big)$ when $q = n$ and $m+1$ is a power of q^{r+s}.

Proof: According to Theorem 8 and Theorem 9, the field size of the constructed $(r; s)$-PMDS sequence and thus the $(r; s)$ PMDS code is $q^{R(r+s)}$, where R was defined above and $q \geqslant n$. Choosing $R = \phi(m, s+1, q^{r+s})$, we get an $(r; s)$ PMDS code with field size $q^{(r+s)\phi(m,s+1,q^{r+s})}$. If $m \leqslant q^{r+s}$ then $R = s$ and the field size becomes $q^{s(r+s)} = n^{s(r+s)}$. Otherwise if $m > q^{r+s}$ and $m+1$ is a power of q^{r+s}, then according to the bound in Lemma 6, we have that $R \leqslant 1 + \left\lceil \left(1 - \frac{1}{q^{r+s}}\right)(s-1) \right\rceil \cdot \log_{q^{r+s}}(m+1)$ and thus the field size becomes

$$q^{(r+s)\left(1+\left\lceil\left(1-\frac{1}{q^{r+s}}\right)(s-1)\right\rceil\cdot\log_{q^{r+s}}(m+1)\right)} = q^{r+s}(m+1)^{s-1} = \mathcal{O}(m^s).$$

In conclusion, it follows that there exists a family of codes which has field size at most $\mathcal{O}\Big(\max\{m, n^{r+s}\}^s\Big)$. ∎

Corollary 10 states our best result for the field size of PMDS codes for arbitrary fixed r and s. Next, we turn to an example that illustrates our code construction. We note that the purpose of the following example is not to identify a code with small field size but to show an illustration and highlight the ideas.

Example 2: Consider a code with the following parameters:

1) $m = 3$, $n = 8$
2) $s = 2, r = 2$.

Our construction will be described through the parity check matrix for a $(2; 2)$ PMDS code $\mathcal{C}_{2;2}$, which we denote as $\mathcal{H}_{2;2}$. Under this setup, our codewords can be interpreted as 3×8 arrays with 8 parity symbols and 16 information symbols. The parity check matrix has the following form:

$$\mathcal{H}_{2;2} = \begin{pmatrix} H & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & H & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & H \\ v_1 & \dots & v_{24} \\ v_1^8 & \dots & v_{24}^8 \end{pmatrix}. \tag{9}$$

Suppose ζ is a primitive element in $\mathbb{F}_8$. Then $H = \begin{pmatrix} 1 & 1 & 1 & \dots & 1 \\ 0 & \zeta & \zeta^2 & \dots & \zeta^7 \end{pmatrix} \in \mathbb{F}_8^{2\times 8}$, which is a parity check matrix for an $[8, 8-2]_8$ extended Reed-Solomon code. From (9), it is straightforward to see that $\mathcal{C}_{2;2}$ can correct any $(2; 0)$-erasure set. In addition, we see from (9) that the first $3 \cdot 2$ rows of $\mathcal{H}_{2;2}$ correspond to constraints that allow us to locally recover from errors. The remaining two rows in $\mathcal{H}_{2;2}$ correspond to the global constraints that allow us to correct an additional 2 erasures.

In Lemma 2, we showed that if certain linear combinations of the elements from the set $\{v_1, v_2, \dots, v_{24}\}$ are linearly independent over $\mathbb{F}_8$, then $\mathcal{C}_{2;2}$ is indeed a $(2; 2)$ PMDS code. To guarantee the condition in Lemma 2 is satisfied, we make use of the following matrices H' and H''. Let

$$H' = \begin{pmatrix} 1 & 1 & 1 & 1 & \cdots & 1 \\ 0 & \zeta & \zeta^2 & \zeta^3 & \cdots & \zeta^7 \\ 0 & \zeta^2 & \zeta^4 & \zeta^6 & \cdots & \zeta^{14} \\ 0 & \zeta^3 & \zeta^6 & \zeta^9 & \cdots & \zeta^{21} \end{pmatrix} \in \mathbb{F}_8^{4\times 8},$$

and let $\eta \in \mathbb{F}_{8^4}$

$$H'' = \begin{pmatrix} 1 & 1 & 1 \\ \eta & \eta^2 & \eta^3 \end{pmatrix} \in \mathbb{F}_{8^4}^{2\times 3}.$$

Since every column in H' has dimension 4 over $\mathbb{F}_8$, we can uniquely represent each column in H' with an element from $\mathbb{F}_{4096}$. In particular, suppose we can write H' as

$$H' = \begin{pmatrix} 1 & 1 & 1 & \cdots & 1 \\ 0 & \zeta & \zeta^2 & \cdots & \zeta^7 \\ 0 & \zeta^2 & \zeta^4 & \cdots & \zeta^{14} \\ 0 & \zeta^3 & \zeta^6 & \cdots & \zeta^{21} \end{pmatrix} = \begin{pmatrix} \eta^{i_1} & \eta^{i_2} & \cdots & \eta^{i_8} \end{pmatrix} \in \mathbb{F}_{8^4}^{1\times 8}.$$

Then, taking the tensor product of H'' and H' (where H' is assumed to be over $\mathbb{F}_{8^4}$), denoted $H'' \otimes H'$, we get:

$$H'' \otimes H' = \begin{pmatrix} 1 \cdot H' & 1 \cdot H' & 1 \cdot H' \\ \eta \cdot H' & \eta^2 \cdot H' & \eta^3 \cdot H' \end{pmatrix} \in \mathbb{F}_{8^4}^{2\times 24} = \begin{pmatrix} \eta^{i_1} & \cdots & \eta^{i_8} & \eta^{i_1} & \cdots & \eta^{i_8} & \eta^{i_1} & \cdots & \eta^{i_8} \\ \eta^{i_1+1} & \cdots & \eta^{i_8+1} & \eta^{i_1+2} & \cdots & \eta^{i_8+2} & \eta^{i_1+3} & \cdots & \eta^{i_8+3} \end{pmatrix}$$

Since $\eta \in \mathbb{F}_{8^4}$, we can represent each column of the matrix $H'' \otimes H'$ as an element from $\mathbb{F}_{8^8}$. Suppose we can write

$$H'' \otimes H' = (\omega_1, \omega_2, \dots, \omega_{24}) \in \mathbb{F}_{8^8}^{24}.$$

Then, by setting $\{v_1, v_2, \dots, v_{24}\} = \{\omega_1, \omega_2, \dots, \omega_{24}\}$ results in a $(2; 2)$-partial MDS code according to Theorem 8. Notice that we have constructed a code $\mathcal{C}_{2;2}$ of length 24 over $\mathbb{F}_{8^8}$ where clearly $8^8 = n^{(r+s)s}$ since $r = s = 2$ and $n = 8$.

In the next section, we consider the field sizes of PMDS codes for some special cases.

V. Codes for Special Cases

This section is organized as follows. In Section V-A, we show that our general result for the field size using the second construction from Section IV-B can be improved for the case where $r = 1$ and $2 \leqslant s \leqslant 4$. In Section V-B,

we present new constructions of PMDS codes that give small field sizes for small m, and we give a construction for $(1;s)$-SD codes with field size $\mathcal{O}\left(n^{\frac{s+1}{2}}m^{s-2}\right)$ when s is odd.

A. Special Cases of the Second Construction

In the following, we show that improvements are possible on the bound from Corollary 10 for the case of $r=1$ and $s \in \{2,3,4\}$.

1) $r = 1$: Recall from Corollary 3 that for $r=1$ the set $T(S)$ needs to be s-wise independent only over $\mathbb{F}_2$. Consequently, repeating the same ideas as in Theorem 8, we have the following construction for the case $r=1$.

Corollary 11: Let $\mathcal{C}_{TP}$ be an $[m,n;s,s+1]$ erasure-correcting code over $\mathbb{F}_2$ with redundancy ρ and parity check matrix $H_{TP} = (\alpha_{1,1},\ldots,\alpha_{m,n}) \in \mathbb{F}_{2^\rho}^{mn}$. Then, the sequence $\Gamma_{TP} = (\alpha_{1,1},\ldots,\alpha_{m,n}) \in \mathbb{F}_{2^\rho}^{mn}$ is a $(1;s)$-PMDS sequence of size $m \times n$.

The next corollary bounds the field size for the case $r=1$. Similar to before, we first give the construction of an $[m,n;s,s+1]$ erasure-correcting code using techniques from [12]:

1) Suppose $H' \in \mathbb{F}_2^{\ell \times n}$ is a parity check matrix for an $[n, n-\ell, s+2]_2$ code.
2) Suppose $H'' \in \mathbb{F}_{2^\ell}^{R \times m}$ is a parity check matrix for an $[m, m-R, s+1]_{2^\ell}$ code.
3) Representing every column in H' as an element in $\mathbb{F}_{2^\ell}$, we let $\mathcal{C}_{TP}$ be a code over $\mathbb{F}_2$ with the parity check matrix $H'' \otimes H'$ formed by taking the tensor product of H'' and H'. Note that $\mathcal{C}_{TP}$ has redundancy ℓR.

Note that the primary difference between the construction presented here and the one from [6] is that we use binary BCH codes to form the matrices H' and, as will be discussed after Corollary 12, this results in a reduction in the required field size of the resulting PMDS code in certain cases.

Using Corollary 11, we get the following analog of Corollary 10 for the case $r=1$.

Corollary 12: There exists a $(1;s)$ PMDS code with field size at most $(n+1)^{\lceil \frac{s+1}{2} \rceil}(m+1)^{s-1}$ provided $n+1$ is a power of 2 and $m+1$ is a power of 2^ℓ.

Proof: According to Lemma 6, we see that $\ell \leqslant \lceil \frac{s+1}{2} \rceil \log_2(n+1)$ provided $n+1$ is a power of two. As before, we can choose $R = \phi(m, s+1, 2^\ell)$, and thus assuming $m+1$ is a power of 2^ℓ, the code redundancy is given by

$$\begin{aligned}
&\ell \cdot \phi(m, s+1, 2^\ell) \\
&\quad = \ell \cdot \left(1 + \left\lceil \left(1 - \frac{1}{2^\ell}\right)(s-1) \right\rceil \cdot \log_{2^\ell}(m+1)\right) \\
&\quad \leqslant \ell \cdot (1 + (s-1) \cdot \log_{2^\ell}(m+1)).
\end{aligned}$$

Note that $2^\ell \leqslant (n+1)^{\lceil \frac{s+1}{2} \rceil}$, and so the field size is at most

$$\begin{aligned}
2^{\ell \cdot \left(1 + (s-1)\cdot \log_{2^\ell}(m)\right)} &\leqslant 2^\ell \cdot 2^{\ell \cdot (s-1) \cdot \log_{2^\ell}(m+1)} \\
&\leqslant (n+1)^{\lceil \frac{s+1}{2} \rceil}(m+1)^{s-1}. \quad \blacksquare
\end{aligned}$$

Thus, it can be seen that when $s \geqslant 5$, n is large enough, and m and n satisfy the conditions stated in Corollary 12, the result improves upon the bound in [6], where the authors showed an achievable field size of $O((mn)^{(s-1)\cdot(1-\frac{1}{2^m})})$. In addition, note that for the case where $s=2$, the previous corollary results in a $(1;2)$ PMDS code with linear field size.

2) PMDS Codes for $2 \leqslant s \leqslant 4$: For $s=2$, the proof of Corollary 10 shows that there exists an $(r;2)$ PMDS code with field size $q^{(r+2)\phi(m,3,q^{r+2})}$, where q is smallest prime power where $q \geqslant n$. If $m \leqslant q^{r+2}$, then $\phi(m,3,q^{r+2}) = 2$ and the required field size is at least $n^{2(r+2)}$; otherwise $\phi(m,3,q^{r+2}) = \lceil \log_{q^{r+2}}(m(q^{r+2}-1)+1) \rceil$, and the size of the field is at least mn^{r+2}. In either case, the construction from [2] requires a smaller field size, given by $r((m+1)(n-m-1)+1)$. However, that construction works only for $s=2$.

For $s=3$, $s=4$, we have the following results, obtained by combining the tensor product construction with the non-binary codes from [4] which were also used in [6].

Corollary 13: There exists an $(r;3)$ PMDS code with field size $n^{r+3}m^{1.5}$ if $m > n^{r+3}$, and $n^{3(r+3)}$ otherwise.

Proof: Following the ideas from the proof of Corollary 10, we deduce that the field size of the constructed $(r;3)$ PMDS code is given by

$$q^{(r+3)\phi(m,4,q^{r+3})}.$$

Here we use the result from [4], which states that $\phi(m,4,q^{r+3}) = 1.5\log_{q^{r+3}}(m) + 1$ for $m > q^{r+3}$, and $\phi(m,4,q^{r+3}) = 3$, otherwise. Assuming that n is a prime we can set $q=n$, and, in the former case, we see that

$$\begin{aligned}
q^{(r+3)\phi(m,4,q^{r+3})} &= q^{(r+3)(1.5\log_{q^{r+3}}(m)+1)} = q^{r+3}m^{1.5} \\
&= n^{r+3}m^{1.5}. \quad \blacksquare
\end{aligned}$$

A similar claim can be proved for the case $s=4$.

Corollary 14: There exists an $(r;4)$ PMDS code with field size $n^{3(r+4)}m^{\frac{7}{3}}$ if $m > n^{r+4}$, and $n^{4(r+4)}$ otherwise.

Proof: From Corollary 10, the field size of the $(r;4)$ PMDS code is

$$q^{(r+4)\phi(m,5,q^{r+4})}.$$

From [4], we can set $\phi(m,5,q^{r+4}) = \frac{7}{3}\log_{q^{r+4}}(m) + 3$ for $m > q^{r+4}$, and $\phi(m,5,q^{r+4}) = 4$ otherwise. Assuming that n is a prime, we can set $q=n$, which gives

$$\begin{aligned}
q^{(r+4)\phi(m,5,q^{r+4})} &= q^{(r+4)(\frac{7}{3}\log_{q^{r+4}}(m)+3)} = q^{3(r+4)}m^{\frac{7}{3}} \\
&= n^{3(r+4)}m^{\frac{7}{3}}.
\end{aligned}$$

The second statement follows simply by setting $\phi(m,5,q^{r+4}) = 4$. $\blacksquare$

B. New Constructions of PMDS and SD Codes

1) PMDS Codes for Small m: In this section, we construct PMDS codes for small m. The constructions from Section IV-B can be used to construct codes with field sizes at most $\mathcal{O}\left(\max\{m, n^{r+s}\}^s\right)$. Here, we show it is possible to construct PMDS codes with field sizes at most $\mathcal{O}\left(n^{s(r+\log_2 s)}\right)$ when $m \leqslant n-1$, and $s \geqslant 10$.

We present a construction of a code $\mathcal{C}_{m<n} \subseteq \mathbb{F}_q^{mn}$ that is an $[m,n;s,r+s]$ erasure-correcting code, which according to Theorem 8, implies the existence of an $(r;s)$-PMDS code. We begin by introducing some useful notation. For $\ell \in [s]$, let

$H_{(\ell)} \in \mathbb{F}_q^{(r+\ell)\times n}$ be a parity check matrix for an $[n, n-r-\ell, r+\ell+1]$ MDS code (where $q \geqslant n$). Let

$$\mathcal{H}_{m<n} = \begin{pmatrix} H_{(s)} & \cdots & H_{(s)} \\ \beta \cdot H_{(\lfloor \frac{s}{2} \rfloor)} & \cdots & \beta^m \cdot H_{(\lfloor \frac{s}{2} \rfloor)} \\ \beta^2 \cdot H_{(\lfloor \frac{s}{3} \rfloor)} & \cdots & \beta^{m\cdot 2} \cdot H_{(\lfloor \frac{s}{3} \rfloor)} \\ \vdots & \ddots & \vdots \\ \beta^{r+s-1} \cdot H_{(1)} & \cdots & \beta^{m\cdot(r+s-1)} \cdot H_{(1)} \end{pmatrix}$$

be a parity check matrix for $\mathcal{C}_{m<n}$ where β has order at least $m+1$. The next corollary follows from Corollary 3 using the same ideas as in the proof of Theorem 8.

Corollary 15: Let $\mathcal{C}_P$ be a code over $\mathbb{F}_q$ with redundancy ρ and parity check matrix $H_P = (\alpha_{1,1}, \ldots, \alpha_{m,n}) \in (\mathbb{F}_{q^\rho})^{mn}$ that is able to correct any $(r; s)$-erasure set. Then, the sequence $\Gamma_P = (\alpha_{1,1}, \ldots, \alpha_{m,n})$ is an $(r; s)$-PMDS sequence of size $m \times n$.

We have the following theorem.

Theorem 16: The code $\mathcal{C}_{m<n} \subseteq \mathbb{F}_q^{mn}$ can correct any $(r; s)$ erasure set $\mathbf{E} = (E_1, E_2, \ldots, E_m)$.

Proof: Let $\epsilon = |\{i : E_i \neq \emptyset\}|$. Notice that since $\mathbf{E}$ is an $(r; s)$ erasure set, then there exists a row which has at most $r + \lfloor \frac{s}{\epsilon} \rfloor$ erasures. Furthermore, notice that for every $k \geqslant j$ where $j, k \in [s]$, $H_{(\lfloor \frac{s}{k} \rfloor)}$ (from $\mathcal{H}_{m<n}$) is equal to the first $r + \lfloor \frac{s}{k} \rfloor$ rows of $H_{(\lfloor \frac{s}{j} \rfloor)}$. Let $\boldsymbol{e} \in \mathbb{F}_q^{mn}$ be a vector that has non-zero entries in positions in $\mathbf{E}$. Then, we need to show $\mathcal{H}_{m<n} \cdot \boldsymbol{e} \neq \mathbf{0}$, which implies that the columns of $\mathcal{H}_{m<n}$ corresponding to the non-zero entries in $\boldsymbol{e}$ are linearly independent. Suppose $\boldsymbol{e} = (\boldsymbol{e}_1, \boldsymbol{e}_2, \ldots, \boldsymbol{e}_m)$. Given ϵ, we know that there exists an index k where

$$H_{(\lfloor \frac{s}{\epsilon} \rfloor)} \cdot \boldsymbol{e}_k \neq \mathbf{0},$$

since there exists a row that has at most $r + \lfloor \frac{s}{\epsilon} \rfloor$ erasures. Thus, $1 \leqslant |\{j : H_{(\lfloor \frac{s}{\epsilon} \rfloor)} \cdot \boldsymbol{e}_j \neq \mathbf{0}\}| \leqslant \epsilon$. Therefore, we have that

$$\begin{pmatrix} H_{(1)} & \cdots & H_{(1)} \\ \beta \cdot H_{(\lfloor \frac{s}{\epsilon} \rfloor)} & \cdots & \beta^m \cdot H_{(\lfloor \frac{s}{\epsilon} \rfloor)} \\ \vdots & \ddots & \vdots \\ \beta^\epsilon \cdot H_{(\lfloor \frac{s}{\epsilon} \rfloor)} & \cdots & \beta^{m\epsilon} \cdot H_{(\lfloor \frac{s}{\epsilon} \rfloor)} \end{pmatrix} \cdot \boldsymbol{e} \neq \mathbf{0}$$

which implies $\mathcal{H}_{m<n} \cdot \boldsymbol{e} \neq \mathbf{0}$ as desired. ∎

Since the dimension of $H_{(\ell)}$ is $r+\ell$ over $\mathbb{F}_q$, it follows that the dimension of the matrix $\mathcal{H}_{m<n}$ over $\mathbb{F}_q$ is at most

$$\begin{aligned} \sum_{j=1}^{s} r + \left\lfloor \frac{s}{j} \right\rfloor &\leqslant rs + s\sum_{j=1}^{s} \frac{1}{j} \\ &\leqslant rs + s(\ln s + 1) \\ &\leqslant s(r + \log_2\ s), \end{aligned}$$

for $s \geqslant 10$.

2) (1;s)-SD Codes With Small Field Sizes: Using our framework, we show that there exists a $(1; s)$-SD code with field size $\mathcal{O}\left(n^{\frac{s+1}{2}} m^{s-2}\right)$ when s is odd. To satisfy the condition in Corollary 4, we construct a matrix $\mathcal{H}_{1;s} = (h_{1,1}, \ldots, h_{m,n})$ such that the set

$$T(\mathcal{H}_{1;s}) = \{h_{i,j} + h_{i,k}\}_{i\in[m], j\in[n]\setminus\{k\}}, \tag{10}$$

is s-wise independent.

The construction of the matrix $\mathcal{H}_{1;s}$ is a simple modification of our construction for $(1; s)$-PMDS codes. We show here that it is possible to reduce the minimum distance of the code whose parity check matrix is H'' (described below). The resulting construction shows the existence of $(1; s)$-SD codes whose field size is $\mathcal{O}\left(n^{\frac{s+1}{2}} m^{s-2}\right)$ when s is odd. For smaller s, the non-binary codes from [4] can be used to provide improvements in the same manner as Section V-A2. Suppose $s \geqslant 3$ is an odd number. We begin by stating the construction for $(1; s)$-SD codes.

1) Suppose $H' \in \mathbb{F}_2^{\ell\times n}$ is a parity check matrix for an $[n, n-\ell, s+2]_2$ code.
2) Suppose $H'' \in \mathbb{F}_{2^\ell}^{R\times m}$ is a parity check matrix for an $[m, m-R, s]_{2^\ell}$ code whose top row is all-ones.
3) Representing every column in H' as an element in $\mathbb{F}_{2^\ell}$, we let $\mathcal{H}_{1;s}$ be the sequence $H'' \otimes H'$ formed by taking the tensor product of H'' and H', with dimension ℓR over $\mathbb{F}_2$.

We note that there are many possible choices for H'' that satisfy the second condition, such as a normalized Reed-Solomon code.

Lemma 17: For odd s, $T(\mathcal{H}_{1;s})$ is s-wise independent.

Proof: Let $\boldsymbol{e} \in \mathbb{F}_2^{mn} = (\boldsymbol{e}_1, \boldsymbol{e}_2, \ldots, \boldsymbol{e}_m) = (e_{1,1}, \ldots, e_{m,n})$ be such that $e_{i,j} = 1$ if and only if $h_{i,j}$ appears as an addend in one of the terms from the set $T(\mathcal{H}_{1;s})$. We prove the result by showing $\mathcal{H}_{1;s} \cdot \boldsymbol{e}^T \neq 0$ unless $\boldsymbol{e} = \mathbf{0}$.

Suppose first that $|\{i : \boldsymbol{e}_i \neq \mathbf{0}| \leqslant s-1$. In this case we cannot have $\mathcal{H}_{1;s} \cdot \boldsymbol{e}^T = \mathbf{0}$ since $\mathcal{H}_{1;s}$ is the parity check matrix for an $[m, n; s-1, s+1]$ code over $\mathbb{F}_2$ and $\boldsymbol{e}$ is correctable by an $[m, n; s-1, s+1]$ erasure-correcting code. Suppose then that $|\{i : \boldsymbol{e}_i \neq \mathbf{0}| = s$. In this case, we have that for any $\boldsymbol{e}_i \neq \mathbf{0}$, $|\{j : e_{i,j} \neq 0\}| = 2$ and $e_{i,k} = 1$. Let $H' = (h'_1, h'_2, \ldots, h'_n)$ and let $i_1, \ldots, i_s$ denote the indices of the non-zero rows in $\boldsymbol{e}$. Since the top row of H'' is all-ones and s is odd, then

$$\begin{aligned} \begin{pmatrix} H' & H' & H' & \cdots & H' \end{pmatrix} \cdot \boldsymbol{e}^T &= \sum_{\ell=1}^{s} h'_{i_\ell,k} + h'_{i_\ell,j_\ell} \\ &= h'_{i_1,k} + \sum_{\ell=1}^{s} h'_{i_\ell,j_\ell}. \end{aligned}$$

Since H' is a parity check matrix for a code with minimum distance $s+2$, $h'_{i_1,k} + \sum_{\ell=1}^{s} h'_{i_\ell,j_\ell} \neq \mathbf{0}$, and the result follows. ∎

Using the same ideas as in the proof of Corollary 12, we have the following result.

Corollary 18: There exists an $(1; s)$-SD code with field size at most $n^{\frac{s+1}{2}} m^{s-2}$ when s is odd provided $n+1$ is a power of 2 and $m+1$ is a power of 2^ℓ.

Proof: According to Lemma 6, we see that $\ell \leqslant (\frac{s+1}{2})\log_2(n+1)$ provided $n+1$ is a power of two. As before, we can choose $R = \phi(m, s, 2^\ell)$, and thus if $m+1$ is a power of 2^ℓ, the code redundancy is given by

$$\begin{aligned} \ell \cdot \phi(m, s, 2^\ell) &= \ell \cdot \left(1 + \left\lceil \left(1 - \frac{1}{2^\ell}\right)(s-2) \right\rceil \cdot \log_{2^\ell}(m+1)\right) \\ &\leqslant \ell \cdot (1 + (s-2) \cdot \log_{2^\ell}(m+1)). \end{aligned}$$

Note that $2^{\ell} \leqslant (n+1)^{\frac{s+1}{2}}$, and so the field size is at most

$$\begin{aligned} 2^{\ell \cdot \left(1+(s-2)\cdot \log_{2^\ell}(m)\right)} &\leqslant (n+1)^{\frac{s+1}{2}} \cdot 2^{\ell \cdot (s-2)\cdot \log_{2^\ell}(m+1)} \\ &\leqslant (n+1)^{\frac{s+1}{2}} (m+1)^{s-2}, \end{aligned}$$

which implies the desired result. ∎

VI. Conclusion and Future Work

In this work, we presented a family of PMDS codes which achieve field size $\mathcal{O}\left(\max\{m, n^{r+s}\}^s\right)$. Although this work shows that it is possible to explicitly construct PMDS codes with polynomial field sizes for $r > 1$, many important problems remain. In particular, it remains an open problem whether there exist families of PMDS codes with smaller field sizes.

Appendix
Proof of Lemma 1

Proof of Lemma 1: Recall that $G \in \mathbb{F}_{q^m}^{s \times s}$. We prove the statement in the lemma by induction on s. For the case where $s = 1$, the lemma clearly holds.

For the inductive step, assume the lemma holds for all $s \leqslant t-1$. Now suppose that $s = t$. Consider the following variable matrix:

$$G(x) = \begin{pmatrix} x & \alpha_2 & \cdots & \alpha_t \\ x^q & \alpha_2^q & \cdots & \alpha_t^q \\ \vdots & \vdots & \ddots & \vdots \\ x^{q^{t-1}} & \alpha_2^{q^{t-1}} & \cdots & \alpha_t^{q^{t-1}} \end{pmatrix}.$$

Clearly, the set of roots of the polynomial $\det(G(x))$ contains the elements $\alpha_2, \alpha_3, \ldots, \alpha_t$, and since $\det(G(x))$ is a linearized polynomial, any linear combination of $\alpha_2, \alpha_3, \ldots, \alpha_t$ over the field $\mathbb{F}_q$ is also a root of $\det(G(x))$. Note that these are all the roots of $\det(G(x))$ since it is a polynomial of degree q^{t-1}. Thus, letting $\boldsymbol{\alpha} = (\alpha_2, \alpha_3, \ldots, \alpha_t)$, we can write

$$\det(G(x)) = C \cdot \prod_{\boldsymbol{u} \in \mathbb{F}_q^{t-1}} (x - \boldsymbol{u} \cdot \boldsymbol{\alpha}^T) \tag{11}$$

where $C \in \mathbb{F}_{q^m}$ is a constant. Setting $x = \alpha_1$, we get

$$\det(G) = C \cdot \prod_{\boldsymbol{u} \in \mathbb{F}_q^{t-1}} (\alpha_1 - \boldsymbol{u} \cdot \boldsymbol{\alpha}^T). \tag{12}$$

Now, if G has full rank, then $\det(G) \neq 0$. From (12), we conclude that $C \neq 0$ and $\prod_{\boldsymbol{u} \in \mathbb{F}_q^{t-1}} (\alpha_1 - \boldsymbol{u} \cdot \boldsymbol{\alpha}^T) \neq 0$. The constant C is equal to the coefficient of the term $x^{q^{t-1}}$ in the expression for the determinant in (11). Thus,

$$C = \det \begin{pmatrix} \alpha_2 & \alpha_3 & \cdots & \alpha_t \\ \alpha_2^q & \alpha_3^q & \cdots & \alpha_t^q \\ \vdots & \vdots & \ddots & \vdots \\ \alpha_2^{q^{t-2}} & \alpha_3^{q^{t-2}} & \cdots & \alpha_t^{q^{t-2}} \end{pmatrix}. \tag{13}$$

The condition $C \neq 0$ implies that the argument of the determinant in (13) is a full rank matrix. Therefore, by the induction hypothesis, the elements of $\boldsymbol{\alpha}$ are linearly independent over $\mathbb{F}_q$. The condition $\prod_{\boldsymbol{u} \in \mathbb{F}_q^{t-1}} (\alpha_1 - \boldsymbol{u} \cdot \boldsymbol{\alpha}^T) \neq 0$ implies that α_1 is not an $\mathbb{F}_q$-linear combination of elements in $\boldsymbol{\alpha}$. It follows that the set $\{\alpha_1, \alpha_2, \alpha_3, \ldots, \alpha_s\}$ is linearly independent over $\mathbb{F}_q$.

Conversely, if the set $\{\alpha_1, \alpha_2, \alpha_3, \ldots, \alpha_s\}$ is linearly independent over $\mathbb{F}_q$, each of the factors of the form $(\alpha_1 - \boldsymbol{u} \cdot \boldsymbol{\alpha}^T)$ in (12) is nonzero, and, by the induction hypothesis, C is also nonzero. Therefore, $\det(G) \neq 0$, implying that G is full rank. This completes the induction. ∎

References

[1] M. Blaum, J. L. 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, Jul. 2013.

[2] M. Blaum, J. S. Plank, M. Schwartz, and E. Yaakobi, "Construction of partial MDS and sector-disk codes with two global parity symbols," *IEEE Trans. Inf. Theory*, vol. 62, no. 5, pp. 2673–2681, May 2016.

[3] G. Calis and O. O. Koyluoglu, "A general construction for PMDS codes," *IEEE Commun. Lett.*, vol. 21, no. 3, pp. 452–455, Mar. 2017.

[4] I. Dumer, "Nonbinary double-error-correcting codes designed by means of algebraic varieties," *IEEE Trans. Inf. Theory*, vol. 41, no. 6, pp. 1657–1666, Nov. 1995.

[5] R. Gabrys, E. Yaakobi, M. Blaum, and P. H. Siegel, "Constructions of partial MDS codes over small fields," in *Proc. IEEE Int. Symp. Inf. Theory*, Aachen, Germany, Jun. 2017, pp. 1–5.

[6] 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, Sep. 2014.

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

[8] P. Gopalan, G. Hu, S. Kopparty, S. Saraf, C. Wang, and S. Yekhanin, "Maximally recoverable codes for grid-like topologies," in *Proc. 28th Annu. ACM-SIAM Symp. Discrete Algorithms*, Barcelona, Spain, Jan. 2017, pp. 2092–2108.

[9] G. Hu and S. Yekhanin, "New constructions of SD and MR codes over small finite fields," in *Proc. IEEE Int. Symp. Inf. Theory*, Barcelona, Spain, Jul. 2016, pp. 1591–1595.

[10] R. Roth, *Introduction to Coding Theory*, Cambridge, U.K.: Cambridge Univ. Press, 2006.

[11] 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.

[12] J. Wolf, "On codes derivable from the tensor product of check matrices," *IEEE Trans. Inf. Theory*, vol. IT-11, no. 2, pp. 281–284, Apr. 1965.

Ryan Gabrys (M'11) is a scientist at SPAWAR Systems Center Pacific. He received the B.S. degree in mathematics and computer science from the University of Illinois at Champaign-Urbana in 2005. In 2014, he received a Ph.D. in electrical engineering at the University of California at Los Angeles. His research interests broadly lie in the areas of theoretical computer science and electrical engineering, including bioinformatics, combinatorics, coding theory, and signal processing.

Eitan Yaakobi (S'07–M'12–SM'17) is an Assistant Professor at the Computer Science Department at the Technion — Israel Institute of Technology. He received the B.A. degrees in computer science and mathematics, and the M.Sc. degree in computer science from the Technion — Israel Institute of Technology, Haifa, Israel, in 2005 and 2007, respectively, and the Ph.D. degree in electrical engineering from the University of California, San Diego, in 2011. Between 2011-2013, he was a postdoctoral researcher in the department of Electrical Engineering at the California Institute of Technology. His research interests include information and coding theory with applications to non-volatile memories, associative memories, DNA storage, and data storage and retrieval. He received the Marconi Society Young Scholar in 2009 and the Intel Ph.D. Fellowship in 2010-2011.

Mario Blaum (S'84–M'85–SM'92–F'00–LF'17) was born in Buenos Aires, Argentina. He received the degree of Licenciado from the University of Buenos Aires in 1977, the M. Sc. degree from the Israel Institute of Technology (Technion) in 1981 and the Ph. D. degree from the California Institute of Technology (Caltech) in 1984, all these degrees in Mathematics. In 1985 he was a Research Fellow at the Department of Electrical Engineering at Caltech and that year he joined the IBM Research Division at the Almaden Research Center. In 2003, his division was transferred to Hitachi Global Storage Technologies, where he was a Research Staff Member until 2009, year in which he rejoined the IBM Almaden Research Center. Since 2001, he is an Academic Advisor at the Universidad Complutense of Madrid, Spain.

Paul H. Siegel (M'82–SM'90–F'97–LF'19) received the S.B. and Ph.D. degrees in mathematics from the Massachusetts Institute of Technology, Cambridge, MA, USA, in 1975 and 1979, respectively. He held a Chaim Weizmann Postdoctoral Fellowship with the Courant Institute, New York University, New York, NY, USA. He was with the IBM Research Division, San Jose, CA, USA, from 1980 to 1995. He joined the faculty at the University of California, San Diego, CA, USA, in 1995, where he is currently a Professor of electrical and computer engineering with the Jacobs School of Engineering. He is affiliated with the Center for Memory and Recording Research where he holds an Endowed Chair and served as Director from 2000 to 2011. His research interests include information theory and communications, particularly coding and modulation techniques, with applications to digital data storage and transmission. He is a Member of the National Academy of Engineering. He was a Member of the Board of Governors of the IEEE Information Theory Society from 1991 to 1996 and from 2009 to 2014. He was the 2015 Padovani Lecturer of the IEEE Information Theory Society. He was a recipient of the 2007 Best Paper Award in Signal Processing and Coding for Data Storage from the Data Storage Technical Committee of the IEEE Communications Society. He was the co-recipient of the 1992 IEEE Information Theory Society Paper Award and the 1993 IEEE Communications Society Leonard G. Abraham Prize Paper Award. He served as a Co-Guest Editor of the 1991 Special Issue on Coding for Storage Devices of the IEEE TRANSACTIONS ON INFORMATION THEORY. He served as an Associate Editor of Coding Techniques of the IEEE TRANSACTIONS ON INFORMATION THEORY from 1992 to 1995, and as the Editor-in-Chief from 2001 to 2004. He was also a Co-Guest Editor of the 2001 two-part issue on The Turbo Principle: From Theory to Practice and the 2016 issue on Recent Advances in Capacity Approaching Codes of the IEEE JOURNAL ON SELECTED AREAS IN COMMUNICATIONS.