ISIT'2013 1569716815

An Efficient Interpolation-Based Systematic Encoder for Low-Rate Blaum-Roth Codes

Qian Guo and Haibin Kan
School of Computer Science
Fudan University, Shanghai, China
Email: {10110240007, hbkan}@fudan.edu.cn

***Abstract*—In this paper, we propose an efficient interpolation-based systematic encoder for low-rate Blaum-Roth codes. Our algorithm is based upon an equivalent definition of $[p, k]$ Blaum-Roth codes from the perspective of generator matrices. Moreover, applying the interpolation method first proposed by D.J.J. Versfeld et al. to the generator matrix, we then derive a formula to resolve the erasure-only decoding problem. Finally, we present a straightforward systematic encoder based on this formula. Compared to the encoders in [5] and [14], it is more efficient for low-rate codes.**

I. Introduction

Maximum Distance Separable (MDS) array codes have become a hot-spot in the last two decades due to their wide applications in communications over parallel channels, packet transmissions in networks, and especially storage systems, such as magnetic tapes, RAID architectures and distributed file systems, see [3], [5]- [10], [12], and many references therein.

As storage systems have grown in size and complexity, proper data redundancy is the key to providing high reliability, availability, and survivability. Therefore, low-rate Blaum-Roth codes [5] seem more applicable, due to their larger-scale redundancy—when the size of the underlying rings becomes relatively large—than EVENODD codes in [9] or other existing coding schemes used in RAID architectures.

In storage systems, the coding scheme must be systematic; otherwise it requires decoding operations in any data retrieval even without node failures. The original systematic encoding method proposed by Blaum and Roth [5] is syndrome-based and quite efficient for high-rate codes. Later, authors in [14] remove the syndrome computing step and obtain an interpolation-based encoder that is efficient for low-rate codes. Both works, however, are based on parity-check matrices; therefore, they can be improved for low-rate Blaum-Roth codes.

In this paper, we study the systematic encoding problem from the perspective of generator matrices. Thus, instead of inverting an $r \times r$ sub-matrix of the parity-check matrix, the encoder in this paper inverts a $k \times k$ matrix. Consequently, the encoding complexity is reduced when $k < r$, i.e., for the low-rate codes.

This paper is organized as follows. We first briefly describe Blaum-Roth codes in Section II, and a Vandermonde-type generator matrix is then derived in Section III. In the next two sections, we present the main results, i.e., the erasure-only decoding formula, the systematic encoder, and its complexity analysis as well. This is followed by our conclusions in Section VI.

II. Preliminary

Let $F = GF(q)$ and let p be a prime which is not the characteristic of F (i.e., $gcd(p, q) = 1$). For an integer a, let $\langle a \rangle_p$ denote the integer $b \in \{0, 1, \cdots, p-1\}$ such that $b \equiv a (mod p)$.

1) Geometric Presentation of Blaum-Roth Codes: Given an integer $n \leq p$, let $\mathcal{M}(p-1, n)$ denote the space of all $(p-1) \times n$ matrices (arrays) $\Gamma = [c_{i,j}]_{i=0,j=0}^{p-2,n-1}$ over F, and we assume that each array $\Gamma \in \mathcal{M}(p-1, n)$ has an extra all-zero row $[c_{p-1,0}, c_{p-1,1}, \ldots, c_{p-1,n-1}]$ for simplicity.

Definition 1: [5] The linear array code $\mathcal{C}(p-1, n, r)$ over F is defined as a subspace of $\mathcal{M}(p-1, n)$ consisting of all arrays $\Gamma = [c_{i,j}]_{i,j}$ which satisfy the following $p \cdot r$ linear constraints:

$$\sum_{j=0}^{n-1} c_{\langle m - jl \rangle_p, j} = 0 \tag{1}$$

where $0 \leq m \leq p-1$ and $0 \leq l \leq r-1$.

To put it another way, $\mathcal{C}(p-1, n, r)$ consists of all arrays in $\mathcal{M}(p-1, n)$ such that the entries along the p lines of slope $l, 0 \leq l \leq r-1$, sum to zero.

2) Algebraic Presentation of Blaum-Roth Codes: We denote the polynomial $\sum_{i=0}^{p-1} x^i$ over F by $M_p(x)$, and let $\mathcal{R}_p = \mathcal{R}_p(q)$ be the ring of polynomials of degree less than $p-1$ over F with multiplication taken modulo $M_p(x)$. Let $\mathcal{R}_p^*$ denote the multiplicative group of the polynomials in $\mathcal{R}_p$, which is relatively prime to $M_p(x)$. Similarly, we denote the polynomial ring modulo $x^p - 1$ by $R(q)$. To avoid confusion, we use the indeterminate α instead of x, when we refer to polynomials as elements of $\mathcal{R}_p$.

Thus, we obtain the following lemmas [5]:

Lemma 1: $\alpha^n \in \mathcal{R}_p^*$ and $(\alpha^m - \alpha^l) \in \mathcal{R}_p^*$, where $m \not\equiv l (mod p)$.

Lemma 2: Consider the following $r \times r$ matrix:

$$V = \begin{pmatrix} 1 & 1 & \ldots & 1 \\ \alpha_0 & \alpha_1 & \ldots & \alpha_{r-1} \\ \alpha_0^2 & \alpha_1^2 & \ldots & \alpha_{r-1}^2 \\ \vdots & \vdots & \ddots & \vdots \\ \alpha_0^{r-1} & \alpha_1^{r-1} & \ldots & \alpha_{r-1}^{r-1} \end{pmatrix}$$

over $\mathcal{R}_p$, where the α_i's are distinct and $\alpha_i = \alpha^{j_i}, 0 \le i \le r-1 (\le p-1)$. Then, the columns of V are linearly independent over $\mathcal{R}_p$.

Therefore, we can define an $[n, n-r]$ linear code over $\mathcal{R}_p$. In particular, it is MDS.

Definition 2: [5] For $r \le n \le p$, if H is the $r \times n$ matrix over $\mathcal{R}_p$ defined by

$$H = \begin{pmatrix} 1 & 1 & \dots & 1 \\ 1 & \alpha & \dots & \alpha^{n-1} \\ 1 & \alpha^2 & \dots & \alpha^{2(n-1)} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & \alpha^{(r-1)} & \dots & \alpha^{(n-1)(r-1)} \end{pmatrix}, \quad (2)$$

we then define a linear code C of length n over $\mathcal{R}_p$ with a parity-check matrix H, i.e.,

$$C = \{c \in (\mathcal{R}_p)^n | cH^T = 0\}. \quad (3)$$

In [5], the authors prove that C is identical to the code obtained by regarding the columns of the arrays in $\mathcal{C}(p-1, n, r)$ as elements in $\mathcal{R}_p$. Thus, we denote a column $[c_{0,i}, c_{1,i}, \dots, c_{p-2,i}]$ in $\mathcal{C}(p-1, n, r)$ by $c_i(\alpha) = \sum_{j=0}^{p-2} c_{j,i}\alpha^i \in \mathcal{R}_p$, when there is no confusion.

Lemma 3: [14] There is a systematic generator matrix of $[n, n-r]$ Blaum-Roth codes with the form

$$G_0 = (I|P^T)_{(n-r)\times n}.$$

In the sequel, we denote $n-r$ by k, and only consider the $[n = p, k]$ Blaum-Roth codes. This constraint is reasonable, since in other cases, namely, $n < p$, these codes can be viewed as punctured codes of $[p, k]$ Blaum-Roth codes.

3) Well-Implemented Structures of Blaum-Roth Codes: Since the three types of operations in erasure-only decoding of Blaum-Roth codes, namely, the additions, the inversions of $\alpha^i - \alpha^j (1 \le i, j \le n-1, i \ne j)$ and the operations of multiplying $a(\alpha)$ by $\alpha^m (m = 1, 2, \dots)$, are carefully designed and well-implemented, these codes are computationally efficient in storage systems. Consequently, one criterion for us to design encoding and decoding procedures for Blaum-Roth codes is to maintain their following special structures.

The implementation of additions requires $(p-1)$-bit operations, by just adding bit by bit, and the inversions of $\alpha^i - \alpha^j (1 \le i, j \le n-1, i \ne j)$ can be implemented by the following lemma:

Lemma 4 ([5]): Let $a(\alpha) = \sum_{i=0}^{p-2} a_i\alpha^i$ and $\widetilde{a}_i = a_i - \frac{1}{p}\sum_{j=0}^{p-1} a_j, 0 \le i \le p-1$, where $a_{p-1} = 0$. Then, for any $m \not\equiv l (mod p)$, the coefficients of the unique solution $b(\alpha) = \sum_{i=0}^{p-2} b_i\alpha^i$ for

$$a(\alpha) = (\alpha^m - \alpha^l)b(\alpha)$$

in $\mathcal{R}_p$ are given by the recursion

$$b_{<-k(m-l)-1>_p} = b_{<-(k-1)(m-l)-1>_p} + \widetilde{a}_{<-(k-1)(m-l)+l-1>_p},$$

$$1 \le k \le p-1,$$

with $b_{p-1} = 0$.

The remaining is the implementation of multiplying $a(\alpha) \in \mathcal{R}_p$ by $\alpha^m (m = 1, 2, \dots)$. As presented in [5], we break the multiplications modulo $M_p(x)$ into two pieces: the first step involves a multiplication modulo $x^p - 1$, which is simply a cyclic-shift of p-vector over F; the second step, called the rectifying operation, rectifies the result modulo $M_p(x)$. Therefore, if $b(\alpha) = \sum_{i=0}^{p-2} b_i\alpha^i \in \mathcal{R}_p$ and $a(\alpha) = \alpha^m b(\alpha) (mod M_p(x))$, we then have,

$$a_i = b_{<i-m>_p} - b_{<-m-1>_p}, 0 \le i \le p-2. \quad (4)$$

III. A Generator Matrix of Blaum-Roth Codes

A Blaum-Roth code C does not always form a vector space since the reducibility of the polynomial $M_p(x)$ is determined by p[1]. However, it forms a module over the ring $\mathcal{R}_p$. Furthermore, the module is a finitely generated free module, and we investigate its generators in this section.

Theorem 1: Let C be an $[n = p, k]$ Blaum-Roth code. Then, the matrix G is a generator matrix of C, where

$$G = \begin{pmatrix} 1 & \alpha & \dots & \alpha^{p-1} \\ 1 & \alpha^2 & \dots & \alpha^{2(p-1)} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & \alpha^k & \dots & \alpha^{k(p-1)} \end{pmatrix}.$$

Proof: We denote $G = (h_1, h_2, \dots, h_k)^T$, where,

$$h_i = (1, \alpha^i, \dots, \alpha^{i(p-1)})^T.$$

For every $j = 0, 1, \dots, r-1$, $1 \le (i+j) < p$ and

$$\sum_{t=0}^{p-1} \alpha^{(i+j)t} = \frac{\alpha^{(i+j)p} - 1}{\alpha^{i+j} - 1} = 0.$$

The last equation follows from two facts:

1) $\alpha^{i+j} \ne 1$, since $1 \le (i+j) < p$.
2) $\alpha^p = 1$.

Therefore, $Hh_i = 0$ and h_i^T is a codeword of the given $[p, k]$ Blaum-Roth code. Then, h_i^T can be represented by the linear combination of the rows of the generator matrix G_0 in Lemma 3, i.e.,

$$G = L_{k\times k}G_0.$$

Since the first k columns of G_0 form an identity matrix, it can be derived that

$$L = \begin{pmatrix} 1 & \alpha & \dots & \alpha^{k-1} \\ 1 & \alpha^2 & \dots & \alpha^{2(k-1)} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & \alpha^k & \dots & \alpha^{k(k-1)} \end{pmatrix}.$$

According to Lemma 1 and Lemma 2, L is invertible over $\mathcal{R}_p$. Thus, G is a generator matrix of C due to the fact that G_0 is a generator matrix of C. ■

[1]The polynomial $M_p(x)$ is irreducible if and only if q is a primitive element in $GF(p)$.

IV. Interpolation-Based Erasure-Only Decoding

Suppose there occur r erasures[2] (and no errors). Let $c = [c_0(\alpha), c_1(\alpha), \ldots, c_{p-1}(\alpha)] \in C$ be the transmitted codeword and the erased locations be $m_1, m_2, \ldots, m_r$. We denote the remaining locations by $n_1, n_2, \ldots, n_k$, where $k = p - r$.

A. Formula and Proof

Theorem 2: Given a Blaum-Roth code over $\mathcal{R}_p$ with a parity-check matrix H defined in Definition 2, the erased symbols can be written as:

$$c_{m_i}(\alpha) = \sum_{u=1}^{k} c_{n_u}(\alpha)\alpha^{m_i - n_u}\frac{f_u(\alpha^{m_i})}{f'(\alpha^{n_j})} \quad i = 1, 2, \ldots, r, \quad (5)$$

where[3]

$$f(y) = \prod_{j=1}^{k}(y - \alpha^{n_j}) \quad (6)$$

and

$$f_u(y) = \frac{f(y)}{(y - \alpha^{n_u})} = \prod_{j=1 j\neq u}^{k}(y - \alpha^{n_j}). \quad (7)$$

Proof: Let $[s_0(\alpha), s_1(\alpha), \ldots, s_{k-1}(\alpha)]$ be the information symbols. For the given Blaum-Roth code, we have that

$$(c_0(\alpha), c_1(\alpha), \ldots, c_{p-1}(\alpha)) = (s_0(\alpha), s_1(\alpha), \ldots, s_{k-1}(\alpha))G,$$

where G is a generator matrix as defined in Theorem 1.

Moreover, the polynomial $f_u(y)$ has another form, i.e.,

$$f_u(y) = \sum_{t=0}^{k-1} f_{u,t}y^t, \quad (8)$$

where $f_{i,t} \in \mathcal{R}_p$.

Consider a matrix D, where

$$D = \begin{pmatrix} f_{1,0} & f_{1,1} & \cdots & f_{1,k-1} \\ f_{2,0} & f_{2,1} & \cdots & f_{2,k-1} \\ \vdots & \vdots & \ddots & \vdots \\ f_{k,0} & f_{k,1} & \cdots & f_{k,k-1} \end{pmatrix}.$$

Left-multiplying D on G', a sub-matrix of the generator matrix G, which consists of the columns with indices $n_1, n_2, \ldots, n_k$, we obtain that

$$DG' = \begin{pmatrix} \alpha^{n_1}f_1(\alpha^{n_1}) & & & \\ & \alpha^{n_2}f_2(\alpha^{n_2}) & & \\ & & \ddots & \\ & & & \alpha^{n_k}f_k(\alpha^{n_k}) \end{pmatrix}. \quad (9)$$

Lemma 1 and Lemma 2 demonstrate that the matrix on the RHS of the above equation and G' are both invertible over $\mathcal{R}_p$. Then, D is invertible over $\mathcal{R}_p$ and we denote its inverse by D^{-1}.

Similarly, we have $(DG)_{u,j} = \alpha^j f_u(\alpha^j)$, for $1 \le u \le k$, $0 \le j \le p-1$, where $(DG)_{u,j}$ stands for the element in the u-th row and j-th column of the matrix DG.

Denote $(G')^{-1}G$ by B. Thus,

$$B = (G')^{-1}G = (DG')^{-1}(DG),$$

and the elements in the matrix B can be written as:

$$B_{u,j} = \frac{\alpha^j f_u(\alpha^j)}{\alpha^{n_u} f_u(\alpha^{n_u})},$$

for all $1 \le u \le k$, $0 \le j \le p-1$.

Since the formal derivative of $f(y)$ is $f'(y) = (y - \alpha^{n_u})f'_u(y) + f_u(y)$, we get $f'(\alpha^{n_u}) = f_u(\alpha^{n_u})$, for $1 \le u \le k$.

Thus, B can be written as:

$$B_{u,j} = \frac{\alpha^j f_u(\alpha^j)}{\alpha^{n_u} f'(\alpha^{n_u})}, \quad (10)$$

for all $1 \le u \le k$, $0 \le j \le p-1$.

By the definition of the matrix G', we have that

$$(c_{n_1}(\alpha), c_{n_2}(\alpha), \ldots, c_{n_k}(\alpha)) = (s_0(\alpha), s_1(\alpha), \ldots, s_{k-1}(\alpha))G',$$

which implies that

$$\begin{aligned}(c_0(\alpha), \ldots, c_{p-1}(\alpha)) &= (s_0(\alpha), s_1(\alpha), \ldots, s_{k-1}(\alpha))G \\ &= (s_0(\alpha), s_1(\alpha), \ldots, s_{k-1}(\alpha))G'(G')^{-1}G \\ &= (c_{n_1}(\alpha), c_{n_2}(\alpha), \ldots, c_{n_k}(\alpha))B\end{aligned}$$

Hence, by the matrix multiplication, we know:

$$c_{m_i}(\alpha) = \sum_{u=1}^{k} c_{n_u}(\alpha)\frac{\alpha^{m_i} f_u(\alpha^{m_i})}{\alpha^{n_u} f'(\alpha^{n_u})},$$

for $i = 1, 2, \ldots, r$. ∎

Remark 1: The form of the matrix B in Formula (10) is first derived in [1] by interpolation, when considering erasure-only decoding for Reed-Solomon codes. It is the reason that we name our encoder an interpolation-based algorithm. Although this formula is similar to the famed Forney's formula [13], the computing steps are totally different.[4] We adopt this form just for brevity.

V. Systematic Encoder

In this section, we regard the encoding problem as a special case of the erasure-only decoding problem, in which the first $p-r$ columns serve as information columns and the remaining r columns are uniquely determined by applying the erasure-only decoding formula in Section IV. The encoding procedure is straightforward, but efficient for low-rate codes.

[2] If the number of erasures are less than r, i.e., there are t error-free columns, where $t \ge k$, then the first k error-free columns of the array are chosen to be substituted in this decoding formula.

[3] The polynomials with indeterminate y represent the polynomials with coefficients in $\mathcal{R}_p$.

[4] To preserve the special structure of Blaum-Roth codes, neither the values of the elementary symmetric functions of $f(y)$ nor their formal derivatives are calculated. In addition, the divide and conquer techniques, e.g., FFT methods, cannot be employed to accelerate the algorithm.

Theorem 3: For a Blaum-Roth code over $\mathcal{R}_p$ with a parity-check matrix H defined in Definition 2, if the information symbols are $(c_0(\alpha), c_1(\alpha), \ldots, c_{k-1}(\alpha))$, then the array $(s_0(\alpha), s_1(\alpha), \ldots, s_{p-1}(\alpha))$ is a codeword of the given Blaum-Roth codes, where

$$s_j(\alpha) = \begin{cases} c_j(\alpha) & \text{for } j = 0, 1, \ldots, k-1 \\ \sum_{t=0}^{k-1} c_t(\alpha)\alpha^{j-t}\frac{f_t(\alpha^j)}{f'(\alpha^t)} & \text{otherwise.} \end{cases} \tag{11}$$

The polynomials in the above formula are defined as,

$$f(y) = \prod_{t=0}^{k-1}(y - \alpha^t) \tag{12}$$

and

$$f_t(y) = \frac{f(y)}{(y-\alpha^t)} = \prod_{i=0, i\neq t}^{k-1}(y - \alpha^i). \tag{13}$$

Proof: The proof is straightforward, by just letting the subscripts $(n_1, n_2, \ldots, n_k, m_1, m_2, \ldots, m_r)$ in Theorem 2 be $(0, 1, \ldots, k-1, k, k+1, \ldots, p-1)$. ■

Corollary 1: For a Blaum-Roth code over $\mathcal{R}_p$ with a parity-check matrix H defined in Definition 2, if we accept the notations of polynomials $f(y)$ and $f_i(y)$ in the above theorem, then the matrix

$$B = (I|P)_{k\times p}$$

is a systematic generator matrix, where

$$P_{i,j} = \frac{\alpha^{j+k} f_i(\alpha^{j+k})}{\alpha^i f'(\alpha^i)},$$

for $0 \le i \le k-1$, $0 \le j \le r-1$.

Remark 2: We can derive a systematic encoder from Corollary 1, by calculating the systematic generator matrix directly. However, this method totally destroys the designed structures of Blaum-Roth codes. Thus, it does not perform well when p is relatively large, and its encoding complexity is $O(krp^2)$, the same as the encoder in [4].

Taking the designed structures of Blaum-Roth codes into consideration, we write Formula (11) in the following form:

$$s_j(\alpha) = \sum_{t=0}^{k-1} \frac{c_t(\alpha)}{\alpha^t f_t(\alpha^t)} \cdot \frac{\alpha^j f(\alpha^j)}{\alpha^j - \alpha^t},$$

and present this encoder, namely, Algorithm 1, formally. The encoding procedure consists of only three types of operations: additions, inversions of $\alpha^i - \alpha^j (1 \le i, j \le n-1, i \ne j)$ and the multiplication of $a(\alpha)$ by $\alpha^m (m = 1, 2, \ldots)$, which can be implemented efficiently as presented in Section II-3. Furthermore, their costs are $p-1$, $2(p-1)$, and $p-1$[5] field operations over F, respectively.

Remark 3:The final step of the encoder involves a bunch of multiplications. For simplicity, we first carry out all calculations modulo $x^p - 1$; after all the multiplications are performed, we obtain the elements in $\mathcal{R}_p$ applying Formula (4). This little trick reduces the encoding complexity.

[5]This is the cost of the rectifying operation. Besides that, a multiplication operation requires a cyclic-shift of p-vector over F as well.

Algorithm 1 Systematic Encoder for Blaum-Roth Codes

Input: the information columns $(c_0(\alpha), c_1(\alpha), \ldots, c_{k-1}(\alpha))$
Output: the encoded columns $(s_k(\alpha), s_{k+1}(\alpha), \ldots, s_{p-1}(\alpha))$

/***Step** 1 (Calculate the values of $\frac{c_t(\alpha)}{\alpha^t f'(\alpha^t)}$).*/
for $t \leftarrow 0$ to $k-1$ **do**
  $a_t(\alpha) \leftarrow \frac{c_t(\alpha)}{\alpha^t}$
  for $s \leftarrow 0$ to $k-1$ **do**
    if $(s = t)$ continue
    else $a_t(\alpha) \leftarrow \frac{a_t(\alpha)}{\alpha^j - \alpha^t}$
  end for
end for
/***Step** 2 (Calculate the values of $\sum_{t=0}^{k-1} \frac{c_t(\alpha)}{\alpha^t f'(\alpha^t)(\alpha^j - \alpha^t)}, j = k, k+1, \ldots, p-1$).*/
for $j \leftarrow k$ to $p-1$ **do**
  $b_j(\alpha) \leftarrow 0$
  for $t \leftarrow 0$ to $k-1$ **do**
    $b_j(\alpha) \leftarrow b_j(\alpha) + \frac{a_t(\alpha)}{\alpha^j - \alpha^t}$
  end for
end for
/***The Final Step**/
for $j \leftarrow k$ to $p-1$ **do**
  $s_j(\alpha) \leftarrow \alpha^{j(k+1)} b_j(\alpha)$
  for $t \leftarrow 0$ to $k-1$ **do**
    $s_j(\alpha) \leftarrow s_j(\alpha) - s_j(\alpha)\alpha^{p+t-j}$
  end for
end for

A. Instance

The next example illustrates the encoding procedure for the low-rate code $\mathcal{C}(4, 5, 3)$ over F_2. In this case, $k = 2$ and $r = 3$. We assume that the information symbols are $c_0(\alpha) = 1 + \alpha + \alpha^3$ and $c_1(\alpha) = 1 + \alpha^2$. This problem, therefore, can be viewed as a decoding task with the received array

1	1	?	?	?
1	0	?	?	?
0	1	?	?	?
1	0	?	?	?

.

We finish encoding by the three-step procedure of Algorithm 1.

Step 1): Firstly, we solve the following equations:

$$(\alpha + 1)a_0(\alpha) = c_0(\alpha),$$

$$\alpha(\alpha + 1)a_1(\alpha) = c_1(\alpha),$$

for $a_0(\alpha)$, $a_1(\alpha)$. Applying Lemma 4 iteratively, we compute $a_0(\alpha)$ as follows:

$$a_0(\alpha) \leftarrow \frac{c_0(\alpha)}{\alpha + 1} = \frac{1 + \alpha + \alpha^3}{\alpha + 1} = \alpha^3 + \alpha^2.$$

Similarly,

$$a_1(\alpha) \leftarrow \alpha^3 + \alpha^2 + \alpha.$$

Step 2): From now on, we need to compute the values of $\frac{a(\alpha)}{\alpha^m - \alpha^l}$, where $m \not\equiv l \ (mod\ 5)$. Using the recursions in Lemma

TABLE I
COMPARISONS OF THE ENCODING COMPLEXITY BETWEEN ALGORITHM 1 AND THE SYSTEMATIC ENCODER IN [14]

	Algorithm 1	Encoder in [14]
Step 1	$2k(k-1)(p-1)$	rkp
Step 2	$3(p-1)kr$	$3(p-1)kr$
Step 3	rkp	$2r(r-1)(p-1)$
Rectifying operations	$(p-1)$	$(p-1)$

4, we receive that

$$b_2(\alpha) = \frac{a_0(\alpha)}{1+\alpha^2} + \frac{a_1(\alpha)}{\alpha+\alpha^2} = 1.$$

Computing $b_3(\alpha)$ and $b_4(\alpha)$ in a similar manner, we end up with $b_3(\alpha) = \alpha^3$ and $b_4(\alpha) = \alpha$.

Step 3): In the final step, our computation is implemented in the polynomial ring $GF(2)[x]$ modulo $x^5 - 1$. We already have that $b_2(x) = 1$, $b_3(x) = x^3$ and $b_4(x) = x$. Noting that the columns $2, 3,$ and 4 are unknown, we obtain

$$\begin{aligned} s_2(x) &= b_2(x)x^2(1+x^2)(x+x^2) \\ &= x^2(1+x^2)(x+x^2) \\ &= 1+x+x^3+x^4. \end{aligned}$$

Taking the results modulo $M_5(x)$, we have that $s_2(\alpha) = \alpha^2$. The values of $s_3(\alpha)$ and $s_4(\alpha)$ are calculated in a similar manner, ending up with $s_3(\alpha) = \alpha+\alpha^2+\alpha^3$ and $s_4(\alpha) = \alpha^2$.

Remark 4:This example provides us some intuitions why Algorithm 1 is more efficient for low-rate Blaum-Roth codes, than the encoder in [14]. The reason is that, while the complexity of the last two steps of Algorithm 1 is the same as that of the first two steps of the encoder in [14], the remaining step of Algorithm 1 requires less inversions of $\alpha^i - \alpha^j$ than the last step of the encoder in [14].

B. Complexity Analysis

We analyze the complexity[6] of the encoding algorithm step by step: Besides $k-1$ rectifying operations, the first step requires $k(k-1)$ inversions of $\alpha^i - \alpha^j$, i.e, $(2k+1)(k-1)(p-1)$ additions over F; The second one requires $3(p-1)kr$ additions over F; The last one requires rkp additions over F and r rectifying operations. Thus, the overall complexity of the encoder sums up to $2k(k-1)(p-1) + (4p-3)kr$ additions over F and $p-1$ rectifying operations, i.e., $(p-1)^2$ additions over F.

It is presented in [14] that the systematic encoder in [14] is more efficient than the original systematic encoder in [5] for low-rate Blaum-Roth codes, and we then show that Algorithm 1 can do even better. As illustrated in Table I, the encoding complexity of the last two steps of Algorithm 1 is the same as that of the first two steps of the encoder in [14]. Moreover, they require equal overall rectifying operations. The only different part is that the first step in Algorithm 1 costs $2k(k-1)(p-1)$ additions over F, while the required number of additions in the last step of the systematic encoder in [14] is $2r(r-1)(p-1)$. For low-rate codes, i.e., $k < r$, it is obvious that Algorithm 1 is more efficient. Thus, among the three known systematic encoders, Algorithm 1 is the best one for low-rate Blaum-Roth codes.

[6]The encoding complexity is represented by the required number of additions over F.

VI. CONCLUDING REMARKS

In this paper, we propose an interpolation-based systematic encoder for low-rate Blaum-Roth codes. We investigate this encoding problem from the perspective of generator matrices, and show that this encoder is more efficient for low-rate Blaum-Roth codes, than those in [5] and [14].

ACKNOWLEDGMENT

The authors would like to thank the anonymous reviewers for their invaluable help. This work was done while the first author was a visiting student at the Institute of Network Coding, The Chinese University of Hong Kong, and he would like to thank Prof. Raymond W. Yeung for his hospitality.

This work was supported by the National Natural Science Foundations of China (Grant No. 61170208), Shanghai Key Program of Basic Research (Grant No. 12JC1401400) and Shanghai Shuguang Project (Grant No. 10SG01).

REFERENCES

[1] D.J.J. Versfeld, James N. Ridley, H.C. Ferreira and A.S.J. Helberg, "On Systematic Generator Matrices for Reed-Solomon Codes", *IEEE Trans. Information Theory,* vol. 56, no. 6, June 2010.
[2] Ron M.Roth, *"Introdution to Coding Theory"*, Cambridge, 2006.
[3] M. Blaum, P. Farrell, and H. van Tilborg, "Array codes," in *Handbook of Coding Theory,* V. Pless and W. Huffman, Eds. Amsterdam, The Netherlands: Elsevier Science B.V, 1998.
[4] J. Blomer, M. Kalfane, R. Karp, M. Karpinski, M. Luby, and D. Zuckerman, An XOR-Based Erasure-Resilient Coding Scheme, Technical Report TR-95-048, ICSI, Berkeley, Calif., Aug. 1995.
[5] M. Blaum and R.M. Roth, "New Array Codes for Multiple Phased Burst Correction," *IEEE Trans. Information Theory,* vol. 39, no. 1, pp. 66-77, Jan. 1993.
[6] O. Keren and Simon Litsyn, "A Class of Array Codes Correcting Multiple Column Erasures," *IEEE Trans. Information Theory,* vol. 43, no. 6, pp. 1843-1851, Nov. 1997.
[7] O. Keren and Simon Litsyn. "Codes Correcting Phased Burst Erasures," *IEEE Transactions on Information Theory,* Vol.44, no. 1, pp: 416-420, July, 1998
[8] M. Blaum and R. Roth, "On lowest density MDS codes," *IEEE Trans. Information Theory,* vol. 45, no. 1, pp. 46-59, Jan. 1999.
[9] M. Blaum, J. Brady, J. Bruck, and J. Menon, "EVENODD: An efficient scheme for tolerating double disk failures in RAID architectures," *IEEE Trans. Comput.*, vol. 45, pp. 192-202, 1995.
[10] M. Blaum, J. Bruck, and A. Vardy, "MDS array codes with independent parity symbols," *IEEE Trans. Information Theory,* vol. 42, pp. 529-542, 1996.
[11] Ron M. Roth, and Gadiel Seroussi, "On generator matrices of MDS codes," *IEEE Trans. Information Theory,* vol. IT-31, pp. 826–830, 1985.
[12] Xu, L. "Maximizing Burst Erasure-Correction Capability of MDS Codes." *Communications, IEEE Transactions on,* Vol.54, no. 11, pp: 1901-1904, 2006
[13] G. D. Forney, "On decoding BCH codes," *IEEE Trans. Information Theory,* vol. IT-11, pp. 549-557, 1965.
[14] Qian Guo, and Haibin Kan, "On systematic encoding for Blaum-Roth codes," *Proc. IEEE Int. Symp. on Information Theory(ISIT)*, St.Pertersburg, pp. 2353 - 2357, July 2011.