

Repairing Reed–Solomon Codes Evaluated on Subspaces

Amit Berman*, Sarit Buzaglo*, Avner Dor*, Yaron Shany*, and Itzhak Tamo†*

*Samsung Semiconductor Israel R&D Center, 2 Shoham St., Ramat Gan, 5251003, Israel

†Department of Electrical Engineering-Systems, Tel Aviv University, Tel Aviv 6997801, Israel

{amit.berman, sarit.b, avner.dor, yaron.shany}@samsung.com, zactamo@gmail.com

Abstract—We consider the repair problem for Reed–Solomon (RS) codes, evaluated on an \mathbb{F}_q -linear subspace $U \subseteq \mathbb{F}_{q^m}$ of dimension d , where q is a prime power, m is a positive integer, and \mathbb{F}_q is the Galois field of size q . For the case of $q \geq 3$, we show the existence of a linear repair scheme for the RS code of length $n = q^d$ and codimension q^s , $s < d$, evaluated on U , in which each of the $n - 1$ surviving nodes transmits only r symbols of \mathbb{F}_q , provided that $ms \geq d(m - r)$. For the case of $q = 2$, we prove a similar result, with some restrictions on the evaluation linear subspace U . Our proof is based on a probabilistic argument, however the result is not merely an existence result; the success probability is fairly large (at least $1/3$) and there is a simple criterion for checking the validity of the randomly chosen linear repair scheme. Our result extend the construction of Dau–Milenkovich to the range $r < m - s$, for a wide range of parameters.

I. INTRODUCTION

Erasure codes are widely used for increasing the reliability of distributed storage systems. In such systems, data is encoded and stored on several *nodes*, where each storage node corresponds to one coordinate of the erasure code. To minimize storage overhead due to coding, erasure codes used in practice are typically Maximum Distance Separable (MDS) codes. While an erasure code can typically recover from several node failures (i.e., from more than a single erasure), a single-node failure is the most common type of failure [10]. Hence, there is an interest in finding MDS codes that can efficiently repair a single node failure.

To repair a single node failure, the system has to download part of the content of some of the surviving nodes, called *helper nodes*. The total amount of data downloaded from the helper nodes is called the *repair bandwidth*. A code designed for minimizing the repair bandwidth is called a *regenerating code*. Regenerating codes have been studied extensively since the introduction of the subject in [3]. A convenient way to measure the repair bandwidth is through the concept of *sub-packetization*, where data is divided to smaller units of a fixed size and each helper node transmits some function of these units. A common approach, which is adopted in this paper, is to utilize a sub-packetization is to consider codes over the extension field \mathbb{F}_{q^m} (over \mathbb{F}_q), where each data node is composed of m symbols of \mathbb{F}_q , hence, the units in the sub-packetization are \mathbb{F}_q -symbols. A code that is defined over \mathbb{F}_{q^m} is called an *array code* of *sub-packetization* m , if $m \geq 2$, and is called a *scalar code*, otherwise.

For an MDS code of length n and dimension k over \mathbb{F}_{q^m} , the *cut-set bound* [3] states that the repair bandwidth is at least $hm/(h+1-k)$ \mathbb{F}_q -symbols, where h is the maximum number of helper nodes that participates in a single node repair. Thus, the repair bandwidth is minimized when h takes its maximum possible value of $n-1$. In this paper we consider only the case $h = n-1$, for which the cut-set bound reads $(n-1)m/(n-k)$. An MDS array code achieving the cut-set bound is called a *minimum storage regenerating (MSR)* code. By now, there are several constructions of MSR array codes (see, e.g., [11] and [13]).

Guruswami and Wootters (GW) [5] introduced a useful characterization of *linear* repair scheme for linear MDS codes in terms of appropriate codewords of the dual codes. In the same paper, Guruswami and Wootters also introduced a linear repair scheme for Reed–Solomon (RS) codes over \mathbb{F}_{q^m} , of full-length (i.e., their evaluation-set is the entire field) and of codimension q^{m-1} . This linear repair scheme is optimal, that is, it achieves the minimum possible repair bandwidth of any linear repair scheme with the same code parameters and sub-packetization. The result of Guruswami and Wootters was later extended by Dau and Milenkovich (DM) [2], who presented linear repair schemes for RS codes with higher dimensions, which is optimal only for RS codes.

While the schemes of [2] and [5] are optimal for full-length RS codes, where the number of data units m in the sub-packetization is logarithmic in the length, they are quite far from the cut-set bound. Until recently, it was an open question whether scalar MDS codes, particularly RS codes, can achieve the cut-set bound. This question was answered in [12], where an explicit evaluation set was presented for which the corresponding RS codes achieve the cut-set bound. For practical implementation, however, this construction is infeasible, since it requires m to be exponential in $n \log n$, where n is the code length [12]. For this reason, there is both a practical and a theoretical interest to further explore the tradeoff between the number of data units (m) and the repair bandwidth of RS codes and to find additional repair schemes. This direction has been recently pursued in [4], [7], and [8].

In this paper we consider linear repair schemes for RS codes evaluated on an \mathbb{F}_q -linear subspace $U \subseteq \mathbb{F}_{q^m}$, in which each surviving node transmits r \mathbb{F}_q -symbols for the repair of the failed node. When q is greater than two, we show the existence of such a linear repair scheme for every choice of U , provided that $ms \geq d(m - r)$, where q^s is the codimension of the RS code and d is the dimension of U . For the case of $q = 2$,

we prove that such a linear repair scheme exists for every choice of U , whenever $ms \geq d(m-r) + 1$, and for many \mathbb{F}_q -linear subspaces, when $ms = d(m-r)$. Our result translates to a practical probabilistic algorithm that outputs with high probability a linear repair scheme for the code, since success probability is fairly large (at least $1/3$) and there is a simple algorithm for checking the validity of the construction. Our result generalizes the result of Dau and Milenkovich and the ‘‘scheme in one coset’’ presented in [7] and [8].

A useful property of our scheme is a duality property between the pair of parameters d and r , and the pair of parameters $m-r$ and $m-d$. Namely, assume C is an RS code that is evaluated on an \mathbb{F}_q -linear subspace of dimension d , and that our construction generates a linear repair scheme for C in which each surviving node transmits r \mathbb{F}_q -symbols. Then there is an explicit way to derive a linear repair scheme for an RS code that is evaluated on an \mathbb{F}_q -linear subspace of dimension $m-r$, in which each surviving node transmits $m-d$ \mathbb{F}_q -symbols.

We also present an explicit construction for the special cases where d or $m-r$ divides m and for a specific choice of U . When d divides m , we set U to be the subfield \mathbb{F}_{q^d} , and present an explicit construction that is almost identical to the scheme of Li *et al.* [8]. Notice that, our existence result supports a much wider parameters range as d may not divide m and the evaluation set may be *any* subspace of dimension d . The case that $m-r$ divides m follows immediately from the duality of our scheme.

The rest of this paper is organized as follows. In Section II we present some of the basic concepts that are used throughout the paper. In particular, we recall the concept of a linear repair scheme and review the important result from [5] that provides a convenient criterion for the existence of a linear repair scheme. In Section III, we present some general results on linear repair schemes for RS codes that are evaluated on linear subspaces. The main result of the paper is given in Section IV. In Section V we show explicit constructions, where d or $m-r$ divides m . Some specific examples are given in Section VI and we conclude the paper in Section VII.

II. PRELIMINARIES

The set of all polynomials in the variable X with coefficients taken from a field \mathbb{F} is denoted by $\mathbb{F}[X]$. The degree of a polynomial $f \in \mathbb{F}[X]$ is denoted by $\deg(f)$. For a subset S of an \mathbb{F}_q -linear space, the \mathbb{F}_q -linear subspace that is spanned by S is denoted by $\text{Span}_q(S)$ and the rank of S (the dimension of $\text{Span}_q(S)$) is denoted by $\text{rank}_q(S)$. For a vector $\mathbf{s} = (s_1, \dots, s_\ell) \in \mathbb{F}_q^\ell$ we will write $\text{rank}_q(\mathbf{s})$, for $\text{rank}_q(s_1, \dots, s_\ell)$. As usual, for a matrix $A \in \mathbb{F}_q^{n \times n}$, the rank of A over \mathbb{F}_q is denoted by $\text{rank}_q(A)$.

Let $V \subseteq \mathbb{F}_q^m$ be an \mathbb{F}_q -subspace of dimension r with a basis $B = \{b_1, \dots, b_r\}$ and let $S = \{b_1, \dots, b_m\}$ be a basis for \mathbb{F}_q^m that contains B . For $x \in \mathbb{F}_q^m$, the *projection* of x to V , x_V , is the unique element $v \in V$ such that $x = v + w$, for some (unique) $w \in \text{Span}_q(S \setminus B)$. For an element $u \in \mathbb{F}_q^m$, consider the \mathbb{F}_q -linear map $F_u : \mathbb{F}_q^m \rightarrow \mathbb{F}_q^m$ defined by $F_u(x) \stackrel{\text{def}}{=} u \cdot x$, and let $[u]_S \in \mathbb{F}_q^{m \times m}$ be the matrix representation of F_u

by right multiplication, according to the basis S . That is, if $\mathbf{x} = (x_1, x_2, \dots, x_m) \in \mathbb{F}_q^m$ is the vector representation of $x \in \mathbb{F}_q^m$ according to the basis S , then $[u]_S \cdot \mathbf{x}^T$ is the vector representation of $F_u(x)$ according to the basis S . We denote by $[u]_{B,S} \in \mathbb{F}_q^{r \times m}$ the matrix consisting of the r rows of $[u]_S$ corresponding to the elements of B . Note that, right multiplication by $[u]_{B,S}$ represents the linear map that maps x to the projection of $F_u(x)$ to V . Similarly, we denote by $[u]_{S,B} \in \mathbb{F}_q^{m \times r}$ the matrix consisting of the r columns of $[u]_S$ corresponding to the elements of B . Right multiplication by $[u]_{S,B}$ represents the linear map that projects x to V and multiplies the result by u .

As usual, an $[n, k]_q$ code C is a linear code of length n and dimension k , over the field \mathbb{F}_q . The *dual code* of an $[n, k]_q$ code C , $C^* \subseteq \mathbb{F}_q^n$,¹ is an $[n, n-k]_q$ code defined by

$$C^* \stackrel{\text{def}}{=} \left\{ (x_1, \dots, x_n) \in \mathbb{F}_q^n : \forall \mathbf{c} \in C, \sum_{i=1}^n c_i x_i = 0 \right\}.$$

A. The Trace Map and the Trace Dual Basis

The *trace map*, $\text{Tr}_{q,m} : \mathbb{F}_{q^m} \rightarrow \mathbb{F}_q$, is defined by

$$\text{Tr}_{q,m}(x) \stackrel{\text{def}}{=} x + x^q + x^{q^2} + \dots + x^{q^{m-1}}.$$

For ease of notation, we denote the trace map by Tr , when q and m are clear from the context. For a basis of \mathbb{F}_{q^m} over \mathbb{F}_q , $S = \{b_1, \dots, b_m\}$, the *trace dual basis* of S , $S' = \{b'_1, \dots, b'_m\}$, is a basis of \mathbb{F}_{q^m} over \mathbb{F}_q for which $\text{Tr}(b'_i b_j) = 1$ if $i = j$ and $\text{Tr}(b'_i b_j) = 0$ otherwise. Note that, for every basis there exists a unique trace dual basis. For $x \in \mathbb{F}_{q^m}$ with $x = \sum_{i=1}^m x_i b_i$, $x_i \in \mathbb{F}_q$, we have that $x_i = \text{Tr}(x b'_i)$, $1 \leq i \leq m$.

Let $B = \{b_1, b_2, \dots, b_r\} \subseteq S$ and let $V = \text{Span}_q(B)$. The *trace-orthogonal* subspace of V , V^\perp , is defined by

$$V^\perp \stackrel{\text{def}}{=} \{x \in \mathbb{F}_{q^m} : \forall v \in V, \text{Tr}(vx) = 0\}.$$

Notice that $\{b'_{r+1}, \dots, b'_m\} \subseteq S$ is a basis for V^\perp .

Lemma 1. *Let $S = \{b_1, b_2, \dots, b_m\}$ be a basis of \mathbb{F}_{q^m} over \mathbb{F}_q , let $B \subseteq S$, and let $V = \text{Span}_q(S \setminus B)$. For $u \in \mathbb{F}_{q^m}$, $w \in V^\perp$, and $x \in \mathbb{F}_{q^m}$ we have that $x = u \cdot w$ if and only if*

$$\mathbf{w} \cdot [u]_{B,S} = \mathbf{x},$$

where \mathbf{w} is the vector representation of w according to the basis B' and \mathbf{x} is the vector representation of x according to the basis S' .

Proof. Let $[u]_S = (u_{ij})$. Then for all $1 \leq j \leq m$, the j th column of $[u]_S$ is the vector representation of $b_j \cdot u$ according to the basis S . Hence, for all $1 \leq i \leq m$, $u_{i,j} = \text{Tr}(b'_i \cdot b_j \cdot u)$. Thus, $[u]_S^T = [u]_{S'}$.

Let $\hat{\mathbf{w}}$ be the vector representation of w according to the basis S' . We have that

$$\hat{\mathbf{w}}[u]_S = \mathbf{x} \Leftrightarrow [u]_S \hat{\mathbf{w}}^T = \mathbf{x}^T \Leftrightarrow x = u \cdot w.$$

¹We use the superscript $*$ instead of the conventional notation $^\perp$ to denote the dual code. The later is used throughout this paper to denote a different type of duality that is defined through the trace map and has a more prominent role in this paper.

Since $w \in V^\perp$ it follows that \hat{w} as zero entries in indices corresponding to elements of $S' \setminus B'$ and hence also in indices corresponding to elements of $S \setminus B$. Thus,

$$\hat{w}[u]_S = \mathbf{x} \Leftrightarrow \mathbf{w}[u]_{B,S} = \mathbf{x},$$

which concludes the proof. \square

Denote by $\text{hom}_q(\mathbb{F}_{q^m}, \mathbb{F}_q)$ the set of all \mathbb{F}_q -linear functionals from \mathbb{F}_{q^m} to \mathbb{F}_q . The set $\text{hom}_q(\mathbb{F}_{q^m}, \mathbb{F}_q)$ is an \mathbb{F}_q -linear space. It is well known that $\text{hom}_q(\mathbb{F}_{q^m}, \mathbb{F}_q)$ is isomorphic to \mathbb{F}_{q^m} . More precisely, any linear functional $\mathbb{F}_{q^m} \rightarrow \mathbb{F}_q$ is of the form $y \mapsto \text{Tr}(xy)$ for a unique $x \in \mathbb{F}_{q^m}$.

Finally, denote by $\sigma : \mathbb{F}_{q^m} \rightarrow \mathbb{F}_{q^m}$ the Frobenius map, defined by $\sigma(x) = x^q$. Notice that σ is an \mathbb{F}_q -linear map.

B. Reed–Solomon Codes

Let $A = \{a_1, \dots, a_n\} \subseteq \mathbb{F}_q$ be a subset of n elements, $1 \leq n \leq q$, and let $k \leq n$ be a positive integer. The Reed–Solomon (RS) code, $\text{RS}(A, k)_q$ is defined as

$$\text{RS}(A, k)_q \stackrel{\text{def}}{=} \left\{ (f(a_1), \dots, f(a_n)) : \begin{array}{l} f \in \mathbb{F}_q[X] \\ \deg(f) \leq k-1 \end{array} \right\}.$$

The set A is called the *evaluation set* of $\text{RS}(A, k)_q$ and we say that the code $\text{RS}(A, k)_q$ is *evaluated on* A . The code $\text{RS}(A, k)_q$ is a linear code of length n , dimension k , and minimum distance $n - k + 1$. Thus, $\text{RS}(A, k)_q$ is an MDS code and can correct up to $n - k$ erasures.

For $\mathbf{v} = (v_1, \dots, v_n)$, $v_i \in \mathbb{F}_q \setminus \{0\}$, $1 \leq i \leq n$, the *Generalized Reed–Solomon* (GRS) code, $\text{GRS}(A, k, \mathbf{v})_q$, is defined as

$$\text{GRS}(A, k, \mathbf{v})_q \stackrel{\text{def}}{=} \{(v_1 c_1, \dots, v_n c_n) : \mathbf{c} \in \text{RS}(A, k)_q\}.$$

We refer to the vector \mathbf{v} as the *GRS scaling vector* of $\text{GRS}(A, k, \mathbf{v})_q$.

It is well known that the dual of a GRS code is yet another GRS code (see, e.g., [6, Thm. 5.1.6, p. 66]),

$$\text{GRS}(A, k, \mathbf{v})^* = \text{GRS}(A, n - k, \mathbf{v}'),$$

where $\mathbf{v}' = (v'_1, \dots, v'_n)$ is given by

$$v'_i = \frac{v_i}{\prod_{j \neq i} (a_i - a_j)}, \quad 1 \leq i \leq n. \quad (1)$$

C. Linear Repair Schemes

In what follows, we review the definition of a linear repair scheme, and the important result of Guruswami–Wootters [5] that provides a criterion to validate a linear repair scheme. The result of Dau–Milenkovich (DM) [2] on linear repair schemes for RS codes is also given. In Section III we focus only on linear repair scheme of RS codes evaluated on \mathbb{F}_q -subspaces of \mathbb{F}_{q^m} .

For a linear code $C \subset \mathbb{F}_{q^m}^n$ and for $1 \leq i \leq n$, an \mathbb{F}_q -linear *repair scheme* for the i th node (coordinate) of codewords in C , in which a surviving node transmits at most r \mathbb{F}_q -symbols, consists of the following.

- 1) A set of \mathbb{F}_q -linear functionals,

$$L = \left\{ g_{j,t} \in \text{hom}_q(\mathbb{F}_{q^m}, \mathbb{F}_q) : \begin{array}{l} 1 \leq j \leq n, j \neq i, \\ 1 \leq t \leq r \end{array} \right\},$$

of size $|L| = (n - 1)r$.

- 2) An \mathbb{F}_q -linear map $f: \mathbb{F}_q^{|L|} \rightarrow \mathbb{F}_{q^m}$, such that for all $(c_1, c_2, \dots, c_n) \in C$, we have

$$c_i = f \left(\left((g_{j,t}(c_j))_{\substack{1 \leq j \leq n, j \neq i \\ 1 \leq t \leq r}} \right) \right). \quad (2)$$

Remark 2. It can be easily verified that if there exists some function f for which (2) holds, then there is also an \mathbb{F}_q -linear map for which (2) holds. Hence, there is no loss of generality in restricting f to be a linear map.

The *repair bandwidth*, b_i , of the above repair scheme for node i , is defined as $b_i \stackrel{\text{def}}{=} \log_2(q) \cdot (n - 1)r$, which is the total number of bits transmitted from the helper nodes in order to repair the erased node i .

For a code C , the *automorphism group* of C , $\text{Aut}(C)$, is the set of all permutations τ of $\{1, \dots, n\}$, such that $\tau \cdot C = C$, where for $\mathbf{c} = (c_1, \dots, c_n) \in C$, $\tau \cdot \mathbf{c} \stackrel{\text{def}}{=} (c_{\tau(1)}, \dots, c_{\tau(n)})$.² The group $\text{Aut}(C)$ is called *transitive* if for all $1 \leq i, j \leq n$, there exists $\tau \in \text{Aut}(C)$ with $\tau(i) = j$. If C has a transitive automorphism group, then a linear repair scheme of C for *some* node can be “permuted” in order to become a linear repair scheme for *any* node.

In this paper we are interested in linear repair schemes for RS codes evaluated on \mathbb{F}_q -subspaces. Henceforth, $U \subseteq \mathbb{F}_{q^m}$ is an \mathbb{F}_q -subspace of dimension $d \leq m$. For a positive integer $s < d$, we denote by $C(U, s)$ the RS code evaluated on U with codimension q^s , i.e.,

$$C(U, s) \stackrel{\text{def}}{=} \text{RS}(U, q^d - q^s)_{q^m}.$$

Clearly, $C(U, s)$ is invariant under any permutation that is a translation by an element of U , and hence we have the following well-known lemma.

Lemma 3. *The code $C(U, s)$ has a transitive automorphism group.*

From Lemma 3 it follows that if $C(U, s)$ has a linear repair scheme for some node i , then it has a linear repair scheme for all nodes.

The following theorem by Guruswami–Wootters [5] plays an important role in the proof of the Dau–Milenkovich scheme and is also useful for the proof of the main result of this paper.

Theorem 4. *A linear code $C \subseteq \mathbb{F}_{q^m}^n$ has an \mathbb{F}_q -linear repair scheme for the i th node in which every surviving node transmits at most r \mathbb{F}_q -symbols, if and only if there exist m dual codewords $\mathbf{u}_\ell = (u_{\ell,1}, \dots, u_{\ell,n}) \in C^*$, $1 \leq \ell \leq m$, with the following properties.*

- 1) $\text{rank}_q(u_{1,j}, \dots, u_{m,j}) \leq r$, for all $j \neq i$.
- 2) $\text{rank}_q(u_{1,i}, \dots, u_{m,i}) = m$.

Remark 5. *As observed in [5], a repair scheme for one GRS scaling vector is automatically also a repair scheme for all GRS scaling vectors. In detail, a repair scheme for $\text{GRS}(A, k, \mathbf{v})$, can be converted to a repair scheme for $\text{GRS}(A, k, \mathbf{v}')$ in the following obvious way. When working*

²The automorphism group is indeed a group with composition as its group operation.

with the latter code, each surviving node j multiplies its content by v_j/v'_j , before using the existing repair scheme, and then the repaired value of the i th node is multiplied by v'_i/v_i . In particular, when repairing RS codes, we may assume without loss of generality that the dual code in the criterion of Theorem 4 is also an RS code. When this sort of argument will be used ahead, we will say that some relevant vectors are in the dual code up to GRS scaling.

As mentioned in the introduction, the main result of this paper can be viewed as a generalization of the Dau–Milenkovich (DM) [2] scheme. The DM scheme is given in the following theorem.

Theorem 6. *For a set $A \subseteq \mathbb{F}_{q^m}$ of size n , where $q^s < n \leq q^m$ the code $RS(A, n - q^s)_{q^m}$ has a linear repair scheme in which each surviving node has to transmit $m - s$ \mathbb{F}_q -symbols for the repair of the erased node.*

For an \mathbb{F}_q -linear subspace U of dimension d and for $1 \leq s < d$, the result of Dau and Milenkovich given in Theorem 6 states that the code $C(U, s)$ has a linear repair scheme in which each helper node transmits at most $r = m - s$ \mathbb{F}_q -symbols for the repair of the erased node. The main contribution of this paper is to show that a lower value of r can be used for the same s ; in fact, r can be as low as $m(d - s)/d$ (with some restrictions on the choice of U for the case $q = 2$ and $ms = d(m - r)$).

III. LINEAR REPAIR SCHEMES FOR RS CODES EVALUATED ON \mathbb{F}_q -LINEAR SUBSPACES

In this section we introduce some results that will be useful in Section IV, where we present and prove the main result of this paper.

The result of Guruswami–Wooters, presented in Theorem 4, provides a criterion to determine if a linear code has a linear repair scheme. For the code $C(U, s)$, the following proposition provides an equivalent criterion for the existence of a linear repair scheme that will be useful for the proof of our main theorem.

Proposition 7. *Let $V \subset \mathbb{F}_{q^m}$ be an \mathbb{F}_q -linear subspace of dimension r , let B_1 be a basis of V , and let B_2 be a set of $m - r$ vectors, such that $S = B_1 \cup B_2$ is a basis of \mathbb{F}_{q^m} over \mathbb{F}_q . For a basis $\{u_1, u_2, \dots, u_d\}$ of U , consider the matrix $M \in \mathbb{F}_q^{d(m-r) \times m(s+1)}$ defined by*

$$M \stackrel{\text{def}}{=} \left(\begin{array}{c|ccc} [u_1]_{B_2, S} & [u_1^q]_{B_2, S} & \cdots & [u_1^{q^s}]_{B_2, S} \\ [u_2]_{B_2, S} & [u_2^q]_{B_2, S} & \cdots & [u_2^{q^s}]_{B_2, S} \\ \vdots & \vdots & \ddots & \vdots \\ [u_d]_{B_2, S} & [u_d^q]_{B_2, S} & \cdots & [u_d^{q^s}]_{B_2, S} \end{array} \right). \quad (3)$$

Write $M = (M_1 | M_2)$, where M_1 consists of the first m columns of M , and M_2 consists of the remaining ms columns. If the column space of M_1 is contained in the column space of M_2 , then $C(U, s)$ has an \mathbb{F}_q -linear repair scheme in which each surviving node has to transmit at most r \mathbb{F}_q -symbols.

Proof. Assume that the column space of M_1 is contained in the column space of M_2 . We will prove the existence of a

linear repair scheme for the node corresponding to evaluation on $0 \in U$ and by Lemma 3, conclude the existence of a linear repair scheme for all nodes.

Let $W = \text{Span}_q(B_2)$. Any linear combination of the columns of M_1 over \mathbb{F}_q can be interpreted as a vector of the form

$$((a_0 u_1)_W, (a_0 u_2)_W, \dots, (a_0 u_d)_W)^T,$$

for some $a_0 \in \mathbb{F}_{q^m}$ (recall that x_W is the projection of x to W). Similarly, any linear combination of the columns of M_2 over \mathbb{F}_q can be interpreted as a vector of the form

$$\left(\left(\sum_{\ell=1}^s a_\ell u_1^{q^\ell} \right)_W, \left(\sum_{\ell=1}^s a_\ell u_2^{q^\ell} \right)_W, \dots, \left(\sum_{\ell=1}^s a_\ell u_d^{q^\ell} \right)_W \right)^T,$$

for some $a_1, a_2, \dots, a_s \in \mathbb{F}_{q^m}$.

Since the column space of M_1 is contained in the column space of M_2 , it follows that for every $a_0 \in \mathbb{F}_{q^m}$, there exist $a_1, \dots, a_s \in \mathbb{F}_{q^m}$ such that $(a_0 u)_W = -(a_1 u^q + \dots + a_s u^{q^s})_W$, for all $u \in U$. Equivalently, the polynomial

$$f(X) = a_0 X + a_1 X^q + \dots + a_s X^{q^s}$$

satisfies that $f(u)_W = 0$, for all $u \in U$, and hence $f(U) \subseteq V$.

In particular, if we write $S = \{b_1, \dots, b_m\}$, then for every $1 \leq j \leq m$, there exist $a_{j,1}, a_{j,2}, \dots, a_{j,s}$, such that

$$f_j(X) \stackrel{\text{def}}{=} b_j X + a_{j,1} X^q + \dots + a_{j,s} X^{q^s}$$

maps U to V .

For $1 \leq j \leq m$, set

$$g_j(X) \stackrel{\text{def}}{=} f_j(X)/X.$$

Then for all $1 \leq j \leq m$, $\deg(g_j) \leq q^s - 1$ and hence, the evaluation of g_j on U is a codeword of $C(U, s)^*$, $\mathbf{x}_j = (x_{j,u})_{u \in U}$. Now, for all $u \in U \setminus \{0\}$, we have

$$\{x_{j,u}\}_{u \in U}^m = \{f_j(u)/u\}_{u \in U}^m \subseteq \frac{1}{u} \cdot V,$$

so that $\text{rank}_q(\{x_{j,u}\}_{u \in U}^m) \leq \dim(V) = r$. Moreover, since $x_{j,0} = b_j$, for all $1 \leq j \leq m$, we have that

$$\text{rank}_q(\{x_{j,0}\}_{j=1}^m) = \text{rank}_q(S) = m.$$

The proof follows From Theorem 4. \square

For M_1, M_2 defined in Proposition 7, a sufficient condition that the column space of M_1 is contained in the column space of M_2 is that the column space of M_2 is equal to $\mathbb{F}_q^{d(m-r)}$, or equivalently, M_2 is of full rank and $ms \geq d(m - r)$.

Definition 8. *A pair (U, V) of \mathbb{F}_q -linear subspaces of dimensions d and r , respectively, is called a good pair, if the corresponding matrix M_2 is of full rank and $ms \geq d(m - r)$.*

Notice that, although the matrix M is defined through a basis B_1 for V and some completion of B_1 to a basis S for \mathbb{F}_{q^m} over \mathbb{F}_q , the goodness of the pair (U, V) does not depend on the choice of these bases.

Lemma 9. *The goodness of the pair (U, V) does not depend on the choice of the basis $\{u_1, u_2, \dots, u_d\}$ for U .*

Proof. Let $\{w_1, w_2, \dots, w_d\}$ be another basis for U and let $A = (a_{i,j}) \in \mathbb{F}_q^{d \times d}$ be the non-singular matrix such that

$$w_i = \sum_{j=1}^d a_{i,j} u_j,$$

for all $1 \leq i \leq d$. Consider the matrix $\tilde{A} = A \otimes I_{m-r}$, where I_{m-r} is the $(m-r) \times (m-r)$ identity matrix and the operation \otimes is the tensor product of matrices. Then \tilde{A} is a $d(m-r) \times d(m-r)$ non-singular matrix and hence the matrix $\tilde{A} \cdot M_2$ has the same rank as M_2 . The proof of the Lemma follows from the fact that

$$\tilde{A} \cdot M_2 = \begin{pmatrix} [w_1^q]_{B_2, S} & \cdots & [w_1^{q^s}]_{B_2, S} \\ [w_2^q]_{B_2, S} & \cdots & [w_2^{q^s}]_{B_2, S} \\ \vdots & \ddots & \vdots \\ [w_d^q]_{B_2, S} & \cdots & [w_d^{q^s}]_{B_2, S} \end{pmatrix}$$

For $x_1, x_2, \dots, x_\ell \in \mathbb{F}_{q^m}$, define

$$T(x_1, \dots, x_\ell; s) \stackrel{\text{def}}{=} \begin{pmatrix} x_1^q & x_2^q & \cdots & x_\ell^q \\ x_1^{q^2} & x_2^{q^2} & \cdots & x_\ell^{q^2} \\ \vdots & \vdots & \ddots & \vdots \\ x_1^{q^s} & x_2^{q^s} & \cdots & x_\ell^{q^s} \end{pmatrix} \in \mathbb{F}_{q^m}^{s \times \ell}.$$

Proposition 10. *For every two positive integers ℓ, s and for $x_1, \dots, x_\ell \in \mathbb{F}_{q^m}$, if $\rho = \text{rank}_q(x_1, \dots, x_\ell)$ then the rank of $T(x_1, \dots, x_\ell; s)$ (over \mathbb{F}_{q^m}) is $\min\{s, \rho\}$.*

Proof. Let $\{y_1, \dots, y_\rho\} \subseteq \{x_1, \dots, x_\ell\}$ be a basis for $\text{Span}_q(x_1, \dots, x_\ell)$. Since the Frobenius map, $\sigma_q(x) \equiv x^q$, is an \mathbb{F}_q -linear map, it follows that all columns of $T = T(x_1, \dots, x_\ell; s)$ are linear combinations of those of $T_1 = T(y_1, \dots, y_\rho; s)$. Hence $\text{rank}(T) = \text{rank}(T_1)$ and it is sufficient to prove that $\text{rank}(T_1) = \min\{s, \rho\}$. For this, it is sufficient to consider the case where $s \leq \rho$, because for $s \geq \rho + 1$, $T(y_1, \dots, y_\rho; \rho)$ appears in the first rows of T_1 and if $T(y_1, \dots, y_\rho; \rho)$ is of full rank then $\text{rank}(T_1) = \rho$.

For a vector $\mathbf{a} = (a_1, \dots, a_s) \in \mathbb{F}_{q^m}^s$ such that $\mathbf{a}T_1 = \mathbf{0}$, consider the polynomial

$$g(X) = a_1 X^q + a_2 X^{q^2} + \cdots + a_s X^{q^s}.$$

Then y_1, y_2, \dots, y_ρ are all roots of $g(X)$ and since $g(X)$ is an \mathbb{F}_q -linear map, it follows that all elements of $\text{Span}_q(y_1, \dots, y_\rho)$ are roots of $g(X)$. Let $f(X) \in \mathbb{F}_{q^m}[X]$ be the polynomial

$$f(X) = a_1^{q^{m-1}} X + a_2^{q^{m-1}} X^q + \cdots + a_s^{q^{m-1}} X^{q^{s-1}}.$$

Since the Frobenius map is an \mathbb{F}_q -linear map, and since $a^{q^m} = a$, for all $a \in \mathbb{F}_{q^m}$, it follows that $f(X)^q = g(X)$. Hence, all the roots of $g(X)$ are roots of $f(X)$ as well and $f(X)$ has at least q^ρ roots. However, the degree of $f(X)$ is $q^{s-1} < q^\rho$. It follows that $f(X)$ must be the zero polynomial and $a_\ell^{q^{m-1}} = a_\ell = 0$, for all $1 \leq \ell \leq s$.

We showed that if $s \leq \rho$ then the rows of T_1 are linearly independent over \mathbb{F}_{q^m} , which concludes the proof. \square

The following two propositions provide useful characterizations of a good pair of subspaces (U, V) .

Proposition 11. *The following conditions are equivalent.*

- 1) *The pair (U, V) is good.*
- 2) *For every basis $\{u_1, \dots, u_d\}$ of U and for all $v'_1, \dots, v'_d \in V^\perp$ for which*

$$T(u_1, u_2, \dots, u_d; s) \cdot (v'_1, v'_2, \dots, v'_d)^T = \mathbf{0} \quad (4)$$

we have that $v'_1 = v'_2 = \cdots = v'_d = 0$.

- 3) *For every basis $B'_2 = \{b'_1, \dots, b'_{m-r}\}$ of V^\perp and for all $w_1, \dots, w_{m-r} \in U$ for which*

$$T(w_1, w_2, \dots, w_{m-r}; s) \cdot (b_1, b'_2, \dots, b'_{m-r})^T = \mathbf{0} \quad (5)$$

we have that $w_1 = w_2 = \cdots = w_{m-r} = 0$.

Proof. Let $\{u_1, u_2, \dots, u_d\}$ be any basis for U . We first prove that conditions (1) and (2) are equivalent. By definition, the pair (U, V) is good if and only if M_2 is of full rank and $ms \geq d(m-r)$. The latter holds if and only if $\mathbf{x} = \mathbf{0}$ is the only vector in $\mathbb{F}_q^{d(m-r)}$ for which $\mathbf{x}M_2 = \mathbf{0}$.

A vector $\mathbf{x} \in \mathbb{F}_q^{d(m-r)}$ can be represented by d chunks of length $m-r$, such that the i th chunk is the vector representation of some element $v'_i \in V^\perp$, $1 \leq i \leq d$, according to the basis B'_2 . By Lemma 1 we have that $\mathbf{x}M_2 = \mathbf{0}$ is equivalent to

$$\sum_{i=1}^d v'_i u_i^{q^\ell} = 0,$$

for all $1 \leq \ell \leq s$, which is equivalent to equation (4).

Hence, (U, V) is good if and only if for every basis $\{u_1, u_2, \dots, u_d\}$ of U and for all $v'_1, v'_2, \dots, v'_d \in V^\perp$, equation (4) implies

$$v'_1 = v'_2 = \cdots = v'_d = 0.$$

Next, we show that conditions (2) and (3) are equivalent. Let $w_1, w_2, \dots, w_{m-r} \in U$ and let $A = (a_{i,j}) \in \mathbb{F}_q^{d \times (m-r)}$ be the matrix for which $w_j = \sum_{i=1}^d u_i a_{i,j}$, for all $1 \leq j \leq m-r$. Then, for every basis $\{b'_1, b'_2, \dots, b'_{m-r}\}$ of V^\perp ,

$$\begin{aligned} T(w_1, w_2, \dots, w_{m-r}; s) \cdot (b'_1, b'_2, \dots, b'_{m-r})^T &= \\ T(u_1, u_2, \dots, u_d; s) \cdot A \cdot (b'_1, b'_2, \dots, b'_{m-r})^T &= \\ T(u_1, u_2, \dots, u_d; s) \cdot (v'_1, v'_2, \dots, v'_d)^T, & \end{aligned}$$

for $v'_1, v'_2, \dots, v'_d \in V$, such that $v'_i = \sum_{j=1}^{m-r} a_{i,j} b'_j$, $1 \leq i \leq d$.

Hence,

$$T(w_1, w_2, \dots, w_{m-r}; s) \cdot (b'_1, b'_2, \dots, b'_{m-r})^T = \mathbf{0},$$

if and only if

$$T(u_1, u_2, \dots, u_d; s) \cdot (v'_1, v'_2, \dots, v'_d)^T = \mathbf{0}.$$

Therefore, if condition (2) holds and w_1, w_2, \dots, w_{m-r} satisfy equation (5), then $v'_1 = v'_2 = \cdots = v'_d = 0$ and hence A is the zero matrix. Thus, $w_1 = w_2 = \cdots = w_{m-r} = 0$ and condition (3) holds as well. Similarly, condition (3) implies condition (2). \square

For a positive integer i , define $U^{\wedge q^i} \stackrel{\text{def}}{=} \{u^{q^i} : u \in U\}$. Note that, since the Frobenius map is \mathbb{F}_q -linear, it follows that $U^{\wedge q^i}$ is an \mathbb{F}_q -linear subspace as well. The next proposition, that is useful for deriving explicit code constructions, suggests a duality between a linear repair scheme of $C(U, s)$, in which each surviving node as to transmits r \mathbb{F}_q -symbols, and a linear repair scheme of $C(V^\perp, s)$, in which each surviving node as to transmits $m - d$ \mathbb{F}_q -symbols.

Proposition 12. *The pair (U, V) is good if and only if $(V^\perp, (U^{\wedge q^{s+1}})^\perp)$ is good.*

Proof. We will show that if (U, V) is good then $(V^\perp, (U^{\wedge q^{s+1}})^\perp)$ is also good. Similar arguments can be used to prove the other direction.

Let $\{b'_1, \dots, b'_{m-r}\}$ be a basis for V^\perp . Assume that (U, V) is good and that $(w_1^{q^{s+1}}, w_2^{q^{s+1}}, \dots, w_{m-r}^{q^{s+1}}) \in U^{\wedge q^{s+1}}$ satisfies

$$T(b'_1, b'_2, \dots, b'_{m-r}; s) \cdot (w_1^{q^{s+1}}, w_2^{q^{s+1}}, \dots, w_{m-r}^{q^{s+1}})^T = \mathbf{0}.$$

Then, for all $1 \leq \ell \leq s$, we have

$$\sum_{j=1}^{m-r} w_j^{q^{s+1}} b'_j{}^\ell = 0.$$

Since $t = s + 1 - \ell$ satisfies that $1 \leq t \leq s$, we can rewrite the equations as

$$\sum_{j=1}^{m-r} w_j^{q^{s+1}} b'_j{}^{q^t} = 0.$$

Raising the t th equation to the power of $q^{m+\ell-(s+1)}$ and using the fact that $x^{q^m} = x$, for all $x \in \mathbb{F}_{q^m}$, we get that for all $1 \leq \ell \leq s$,

$$\sum_{j=1}^{m-r} w_j^{q^\ell} b'_j = 0,$$

or equivalently

$$T(w_1, w_2, \dots, w_{m-r}; s) \cdot (b'_1, b'_2, \dots, b'_{m-r})^T = \mathbf{0}.$$

Since (U, V) is good, it follows from condition (3) of Proposition 11 that $w_1 = w_2 = \dots = w_{m-r} = 0$, and hence $w_j^{q^{s+1}} = 0$, for all $1 \leq j \leq m - r$. Thus, by condition (2) of Proposition 11 we have that $(V^\perp, (U^{\wedge q^{s+1}})^\perp)$ is good. \square

Let Ω be the set of all vectors in $\mathbb{F}_{q^m}^{m-r}$ whose entries are \mathbb{F}_q -linearly independent, i.e., for $\mathbf{x} \in \mathbb{F}_{q^m}^{m-r}$, $\mathbf{x} \in \Omega$ if and only if $\text{rank}_q(\mathbf{x}) = m - r$.

Lemma 13.

$$|\Omega| > \frac{q^{m(m-r)}(q-1-q^{-r})}{q-1}.$$

Proof. The size of Ω is given by

$$\begin{aligned} |\Omega| &= \prod_{j=0}^{m-r-1} (q^m - q^j) \\ &= q^{m(m-r)} \prod_{j=0}^{m-r-1} (1 - q^{-m+j}). \end{aligned}$$

A straightforward induction on n shows that for all n positive real numbers x_1, \dots, x_n , we have that $\prod_{j=1}^n (1 - x_j) \geq 1 - \sum_{j=1}^n x_j$. Hence,

$$\begin{aligned} \frac{|\Omega|}{q^{m(m-r)}} &= \prod_{j=0}^{m-r-1} (1 - q^{-(m-j)}) \\ &\geq 1 - \sum_{j=0}^{m-r-1} q^{-(m-j)} \\ &= 1 - \sum_{j=r+1}^m q^{-j} \\ &> 1 - \sum_{j=r+1}^{\infty} q^{-j} \\ &= 1 - \frac{q^{-(r+1)}}{1 - q^{-1}} = 1 - \frac{q^{-r}}{q-1}, \end{aligned}$$

as required. \square

Let

$$\text{Bad}(U) \stackrel{\text{def}}{=} \left\{ \mathbf{x} \in \Omega : \begin{array}{l} \exists (u_1, \dots, u_{m-r}) \in U^{m-r} \setminus \{\mathbf{0}\}, \\ \text{s.t. } T(u_1, \dots, u_{m-r}; s) \cdot \mathbf{x}^T = \mathbf{0}^T \end{array} \right\}.$$

For a pair (U, V) of \mathbb{F}_q -linear subspaces of dimensions d and r , respectively, let $\mathbf{v}' \in \Omega$ be such that $V^\perp = \text{Span}_q(\mathbf{v}')$. It follows from Proposition 11 that (U, V) is good if and only if $\mathbf{v}' \in \Omega \setminus \text{Bad}(U)$. In the next section, we will show the existence of a good pair (U, V) under certain conditions. For this purpose, it will be useful to upper bound the size of $\text{Bad}(U)$.

Lemma 14. *For $\mathbf{u} = (u_1, \dots, u_{m-r}) \in U^{m-r} \setminus \{\mathbf{0}\}$, let $\rho = \text{rank}_q(\mathbf{u})$ and define the set*

$$\text{Bad}(\mathbf{u}) \stackrel{\text{def}}{=} \{ \mathbf{x} \in \Omega : T(u_1, \dots, u_{m-r}; s) \cdot \mathbf{x}^T = \mathbf{0}^T \}.$$

Then the following holds.

- 1) If $\rho \leq s$ then $\text{Bad}(\mathbf{u}) = \emptyset$.
- 2) If $\rho \geq s + 1$ then

$$|\text{Bad}(\mathbf{u})| < q^{m(m-r-s)}. \quad (6)$$

Proof. Let $\{w_1, \dots, w_\rho\}$ be a basis for $\text{Span}_q(\mathbf{u})$. Since the Frobenius map is \mathbb{F}_q -linear, it follows that there exists a (unique) matrix $N \in \mathbb{F}_q^{\rho \times (m-r)}$ such that

$$T(u_1, \dots, u_{m-r}; s) = T(w_1, \dots, w_\rho; s) \cdot N.$$

To prove (1), assume that $\rho \leq s$. It follows from Proposition 10 that the rank of $T = T(u_1, \dots, u_\rho; s)$ is ρ , and therefore the columns of T are \mathbb{F}_{q^m} -linearly independent. Hence, $T \cdot N \cdot \mathbf{x}^T = \mathbf{0}^T$ if and only if $N \cdot \mathbf{x}^T = \mathbf{0}^T$. However, for all $\mathbf{x} \in \Omega$, we have that $N \cdot \mathbf{x}^T \neq \mathbf{0}$. This is true since all entries of N belong to \mathbb{F}_q , N is not the zero matrix, and the entries of \mathbf{x} are \mathbb{F}_q -linearly independent. Hence, $T(u_1, \dots, u_{m-r}; s) \cdot \mathbf{x}^T \neq \mathbf{0}^T$, for all $\mathbf{x} \in \Omega$, which implies that $\text{Bad}(\mathbf{u}) = \emptyset$.

For the proof of (2), assume that $\rho \geq s + 1$ (note that since $\rho \leq m - r$, this implies in particular that $s \leq m - r - 1$). It follows from Proposition 10, that

$\text{rank}_{\mathbb{F}_q}(T(u_1, \dots, u_{m-r}; s)) = s$. Hence, the \mathbb{F}_q -dimension of the right-kernel of $T(u_1, \dots, u_{m-r}; s)$ is $m - r - s$. Thus,

$$|\text{Bad}(\mathbf{u})| < q^{m(m-r-s)}. \quad (7)$$

□

Lemma 15. *Let a be a common factor of m and d and assume that $U \subseteq \mathbb{F}_{q^m}$ is an \mathbb{F}_{q^a} -linear subspace of dimension d/a . Then,*

$$|\text{Bad}(U)| < \frac{q^{d(m-r)-ms}}{q^a - 1} q^{m(m-r)} \quad (8)$$

Proof. Define an equivalence relation \sim on $\mathbb{F}_{q^m}^{m-r} \setminus \{\mathbf{0}\}$ by setting $\mathbf{x} \sim \mathbf{y}$ if and only if there exists $\beta \in \mathbb{F}_{q^a} \setminus \{0\}$ such that $\mathbf{x} = \beta \cdot \mathbf{y}$. Note that, since U is a vector space over \mathbb{F}_{q^a} , the equivalence class of any $\mathbf{u} \in U^{m-r} \setminus \{\mathbf{0}\}$ is contained in $U^{m-r} \setminus \{\mathbf{0}\}$.

Let $\text{Reps} \subset U^{m-r} \setminus \{\mathbf{0}\}$ be a set consisting of a single representative from each equivalence class of \sim in $U^{m-r} \setminus \{\mathbf{0}\}$, and note that $|\text{Reps}| = (q^{d(m-r)} - 1)/(q^a - 1)$. Note also that, as $T(\beta \cdot \mathbf{u}; s) = \text{diag}(\{\beta^i\}_{i=1}^s) \cdot T(\mathbf{u}; s)$ ($\text{diag}(\beta_1, \dots, \beta_n)$ is the diagonal $n \times n$ matrix D with $D_{i,i} = \beta_i$, $1 \leq i \leq n$), for all $\beta \in \mathbb{F}_{q^a}$, $\mathbf{u} \sim \mathbf{u}'$ implies $\text{Bad}(\mathbf{u}) = \text{Bad}(\mathbf{u}')$. It follows from the above comments and from Lemma 14 that

$$\begin{aligned} |\text{Bad}(U)| &= \left| \bigcup_{\mathbf{u} \in \text{Reps}} \text{Bad}(\mathbf{u}) \right| \leq \sum_{\mathbf{u} \in \text{Reps}} |\text{Bad}(\mathbf{u})| \\ &< \frac{q^{d(m-r)} - 1}{q^a - 1} q^{m(m-r-s)} < \frac{q^{d(m-r)-ms}}{q^a - 1} q^{m(m-r)} \end{aligned}$$

□

An intriguing question is whether or not, for all U , all pairs (U, V) are good, or equivalently, does $|\text{Bad}(U)| = 0$. The answer to this question is given in the next proposition. The proof can be found in the appendix.

Proposition 16. *Assume that $1 \leq s < d < m$ and $ms \geq d(m-r)$. Then for every \mathbb{F}_q -linear subspace $U \subseteq \mathbb{F}_{q^m}$ of dimension d , the following holds.*

- 1) If $r \geq m - s$ then $|\text{Bad}(U)| = 0$.
- 2) If $r < m - s$ then $|\text{Bad}(U)| > 0$.

IV. EXISTENCE OF LINEAR REPAIR SCHEMES

In this section we present and prove the main result of the paper, namely, the existence of a linear repair scheme for $C(U, s)$, in which surviving nodes transmit at most r symbols from \mathbb{F}_q .

Theorem 17. *The code $C(U, s)$ has an \mathbb{F}_q -linear repair scheme in which each surviving node transmits r \mathbb{F}_q -symbols, provided that one of the following conditions holds.*

- 1) $q \geq 3$ and $ms \geq d(m-r)$.
- 2) $q = 2$, $r \geq 2$, and $ms \geq d(m-r) + 1$.
- 3) $q = 2$, $ms = d(m-r)$ and U is a \mathbb{F}_{q^a} -linear subspace of \mathbb{F}_{q^m} of dimension d/a , for $a = \text{gcd}(m, d)$.

Notice that, the third condition of Theorem 17 includes a more strict restriction on U , i.e., U is required to be an \mathbb{F}_{q^a} -subspace of \mathbb{F}_{q^m} of dimension d/a . This requirement on U

is stronger, since any such subspace of \mathbb{F}_{q^m} is also an \mathbb{F}_q -subspace of dimension d . In addition, if d and m are co-prime, i.e., $a = 1$, the equality $s = d(m-r)/m$ implies that $r = m$ and $s = 0$, and hence $C(U, s)$ is an RS code of length $n = q^d$ and dimensions $k = n - 1$. This special case trivially holds, since such a code can correct any node failure when all surviving nodes transmit their entire content.

The proof of Theorem 17 involves a probabilistic argument in which an \mathbb{F}_q -subspace $V \subset \mathbb{F}_{q^m}$ of dimension r is chosen uniformly at random. If the pair (U, V) is good then by Proposition 7 a linear repair scheme for the code $C(U, s)$ is guaranteed. Moreover, the goodness of the pair (U, V) can be verified, using Gaussian elimination, in polynomial time. We will show that the probability that (U, V) is good is fairly large (at least $1/3$) and thus obtain a practical probabilistic algorithm to construct the promised repair scheme for each subspace U guaranteed by Theorem 17.

In what follows, we assume that $\mathbf{v}' = (v'_1, \dots, v'_{m-r})$ is a vector drawn uniformly at random from the set Ω , i.e., $\mathbf{v}' \in \mathbb{F}_{q^m}^{m-r}$ and $\text{rank}_q(\mathbf{v}') = m - r$. The proof of Theorem 17 will follow immediately from the next theorem and corollary.

Theorem 18. *Let $\mathbf{v}' \in \Omega$ and let $V \subset \mathbb{F}_{q^m}$ be the \mathbb{F}_q -linear subspace of dimension r such that $V^\perp = \text{Span}_q(\mathbf{v}')$. For a positive integer a , if a is a common factor of m and d , and $U \subseteq \mathbb{F}_{q^m}$ is an \mathbb{F}_{q^a} -subspace of dimension d/a , the probability that (U, V) is good is at least*

$$1 - \frac{q^{d(m-r)-ms}}{q^a - 1} \cdot \frac{q - 1}{q - 1 - q^{-r}}. \quad (9)$$

Proof. We will use a counting argument based on Proposition 11. By Proposition 11, (U, V) is good if and only if $\mathbf{v}' \in \Omega \setminus \text{Bad}(U)$. Hence,

$$\text{Prob}((U, V) \text{ is good}; U) = 1 - \frac{|\text{Bad}(U)|}{|\Omega|}. \quad (10)$$

Combining (10) with Lemmas 13 and 15 we have that

$$\text{Prob}((U, V) \text{ is good}) > 1 - \frac{q^{d(m-r)-ms}}{q^a - 1} \cdot \frac{q - 1}{q - 1 - q^{-r}}. \quad \square$$

Corollary 19. *If $U \subseteq \mathbb{F}_{q^m}$ is an \mathbb{F}_q -subspace of dimension d and p is the probability that (U, V) is good, then the following statements hold.*

- 1) If $q \geq 3$ and $ms \geq d(m-r)$ then $p \geq 2/5$.
- 2) If $q = 2$, $r \geq 2$, and $ms \geq d(m-r) + 1$ then $p \geq 1/3$.
- 3) Let $a = \text{gcd}(m, d)$. If $q = 2$, $a \geq 2$, $ms = d(m-r)$, and U is also an \mathbb{F}_{q^a} -subspace of \mathbb{F}_{q^m} of dimension d/a , then $p \geq 1/3$.

Proof. Let h be the right hand side of (9). Then h is minimized when $ms - d(m-r)$, r , q , and a are minimized. If the conditions of (1) hold, then the minimum of h is obtained for $q = 3$, $r = 1$, $ms = d(m-r)$, and $a = 1$ and is equal to $2/5$. If the conditions of (2) hold, then the minimum of h is obtained for $r = 2$, $ms = d(m-r) + 1$, and $a = 1$ and is equal to $1/3$. Lastly, for the conditions of (3), the minimum of $1/3$ is obtained for $a = 2$ and $r = 1$. □

V. EXPLICIT CONSTRUCTIONS

In this section we present explicit constructions of linear repair schemes for $C(U, s)$, for a specific choice of the \mathbb{F}_q -linear subspace U , where $(m - r)$ divides m or d divides m .

First, we present a construction for the code $C(U, s)$, for some \mathbb{F}_q -linear subspace $U \subseteq \mathbb{F}_{q^m}$, where $m - r$ divides m . Recall that by Proposition 7, it suffices to show an explicit choice of an \mathbb{F}_q -linear subspace V of \mathbb{F}_{q^m} of dimension r such that the pair (U, V) is good.

Proposition 20. *Assume that $(m - r)$ divides m , $d < m$, and $ms \geq d(m - r)$. Let $\alpha \in \mathbb{F}_{q^m}$ be a primitive element and let $U \stackrel{\text{def}}{=} \text{Span}_q(1, \alpha, \dots, \alpha^{d-1})$. Then the pair $(U, V = \mathbb{F}_{q^{m-r}}^\perp)$ is good.*

Proof. First, notice that since $m - r$ divides m , it follows that $\mathbb{F}_{q^{m-r}}$ is a subfield of \mathbb{F}_{q^m} and therefore, V is a well defined \mathbb{F}_q -subspace of \mathbb{F}_{q^m} of dimension r . It suffices to prove the case where s takes its minimum possible value $\lceil d(m-r)/m \rceil$, since if we show that M_2 from Proposition 7 is of full rank for the minimal s then it also holds for larger values of s . In particular, we may assume that $s \leq m - r$.

By condition (2) of Proposition 11 it is sufficient to prove that for all $f \in \mathbb{F}_{q^{m-r}}[X]$ with $\deg(f) < d$, if $f(\alpha^q) = f(\alpha^{q^2}) = \dots = f(\alpha^{q^s}) = 0$ then f must be the zero polynomial.

If $x \in \mathbb{F}_{q^m}$ is a root of $f(X)$, then so are the conjugates $x^{q^{j(m-r)}}$, for all $1 \leq j < m/(m-r)$. Hence, it is sufficient to prove that if all elements of

$$R \stackrel{\text{def}}{=} \left\{ \alpha^{q^{j(m-r)+i}} : 1 \leq i \leq s, 0 \leq j < m/(m-r) \right\} \quad (11)$$

are roots of f , then f is the zero polynomial.

Since $s \leq m - r$, all exponents $j(m-r) + i$ of q appearing in (11) are positive, distinct, and smaller than $m+1$. It follows that $|R| = sm/(m-r) \geq d > \deg(f)$, and thus f must be the zero polynomial, as required. \square

By Proposition 20 and by the duality of the goodness property given in Proposition 12, we conclude that if d divides $m - r$ then the pair $(\mathbb{F}_{q^d}, \tilde{U})$ is good, for some \mathbb{F}_q -linear subspace $\tilde{U} \subseteq \mathbb{F}_{q^m}$ of dimension $m - r$ that can be derived from $\text{Span}_q(1, \alpha, \dots, \alpha^{m-r-1})$. Thus, we have an explicit construction of a linear repair scheme for $C(\mathbb{F}_{q^d}, s)$, in which each surviving node has to transmit at most r \mathbb{F}_q -symbols for the repair of the erased node. This construction is also a straightforward generalization of the DM scheme and is similar to the ‘‘scheme in one coset’’ proposed by Li *et al.* [8]. The result is summarized in the next proposition, to which we present an alternative proof that is based on a simple but useful argument.

Proposition 21. *Assume that d divides m , $d < m$, and that $ms = d(m - r)$. Then the code $C(\mathbb{F}_{q^d}, s)$ has a linear repair scheme in which each surviving node has to transmit at most r \mathbb{F}_q -symbols for the repair of the erased node.*

Proof. Let $\{b_1, \dots, b_{m/d}\}$ be any basis for \mathbb{F}_{q^m} over \mathbb{F}_{q^d} , and let $\{b'_1, \dots, b'_{m/d}\}$ be its dual basis. For a polynomial $f(X) \in \mathbb{F}_{q^m}[X]$ of degree at most $k - 1$, where $k = q^d - q^s$, there

Repair Scheme	q	m	n	k	r	b
Prop. 21	2	8	14	10	4	52
Naive	2	8	14	10	8	80
DM	2	8	14	10	-	54
Thm. 17	2	15	64	48	5	315
Naive	2	15	64	48	15	720
DM	2	15	64	48	11	693

TABLE I

exist polynomials $f_j(X) \in \mathbb{F}_{q^d}[X]$ of degree at most $k - 1$ such that

$$f(X) = b_1 f_1(X) + \dots + b_{m/d} f_{m/d}(X).$$

Hence, the codeword $\mathbf{c} = (f(\alpha))_{\alpha \in \mathbb{F}_{q^d}}$ can be represented by the m/d codewords of $\text{RS}(\mathbb{F}_{q^d}, q^d - q^s)_{\mathbb{F}_{q^d}}$, $\mathbf{c}_j = (f_j(\alpha))_{\alpha \in \mathbb{F}_{q^d}}$. In addition, for $\beta \in \mathbb{F}_{q^d}$ and for all $1 \leq j \leq m/d$,

$$f_j(\beta) = \text{Tr}_{q^d, m/d}(f(\beta) \cdot b'_j). \quad (12)$$

This implies that a linear repair scheme of $\text{RS}(\mathbb{F}_{q^d}, q^d - q^s)_{\mathbb{F}_{q^d}}$, in which each surviving node transmits at most r' \mathbb{F}_q -symbols, results in a linear repair scheme for $C(\mathbb{F}_{q^d}, s)$ in which each surviving node transmits at most $r = r' m/d$ \mathbb{F}_q -symbols.

By the DM scheme, for every $1 \leq s < d$, $\text{RS}(\mathbb{F}_{q^d}, q^d - q^s)_{\mathbb{F}_{q^d}}$ has a linear repair scheme in which each surviving node has to transmit at most $r' = d - s$ symbols, which concludes the proof. \square

VI. EXAMPLES

In Table I we consider two specific examples of linear codes with linear repair schemes that are obtained from our constructions and compare their bandwidth to known linear repair schemes of these codes.

We first consider the well known $[14, 10]_{\mathbb{F}_{2^8}}$ GRS code deployed at the Facebook Hadoop Analytic cluster (see, e.g., [5, Sec. V.C] and references therein). Using Proposition 21, we construct $C(U, s)$ code over \mathbb{F}_{2^8} with $U = \mathbb{F}_{2^4}$, $s = 2$ and $r = 4$. The code $C(U, s)$ is a $[16, 12]_{2^8}$ code. We then shorten this code to obtain a $[14, 10]_{\mathbb{F}_{2^8}}$ code with a linear repair scheme in which $r = 4$ and the bandwidth is $b = 52$. This construction was also given in [8]. A naive decoding of an RS code over with \mathbb{F}_{2^8} with dimension 10 has bandwidth 80, while the linear repair scheme from [2] achieves a bandwidth of 54, where not all surviving node transmitting the same number of bits.

The second code we consider is $C(U, s = 4)$, where U is an \mathbb{F}_{2^3} -subspace of $\mathbb{F}_{2^{15}}$ of dimension two. Hence, U is an \mathbb{F}_2 -subspace of dimension 6 and from Theorem 17, $C(U, s)$ has a linear repair scheme in which $r = 5$. This code is a $[64, 48]_{2^{15}}$ RS code. The bandwidth of a naive approach and the main scheme from [2] are presented in Table I.

Lastly, we consider the case $q = 2$, $r \geq 2$, $s = 1$, and $ms = d(m - r)$, where $\gcd(m, d) > 1$. A linear repair scheme for these parameters is guaranteed by Theorem 17. The constructed RS codes have two parity symbols. Since $m = d(m - r)$, it follows that $r = m(d - 1)/d$ and the bandwidth is $(n - 1)m(d - 1)/d$, where $n = q^d$. A construction of linear repair schemes for RS codes of codimension 2 over \mathbb{F}_{2^m} is also given in [5, Thm. 10], with repair bandwidth $3(n - 1)m/4$, where $n \leq 2^{m/2+1}$ is the length of the code. This

shows that in general, the bandwidth of the scheme of Theorem 18 is not minimal. Note that, the scheme of [5, Thm. 10] is imbalanced, in the sense that about half of the surviving nodes transmit half of their content, while the remaining surviving nodes transmit their entire content. Moreover, the evaluation set in this scheme is not a linear subspace.

VII. CONCLUSION

In this work we studied the repair problem for RS codes, evaluated on an \mathbb{F}_q -linear subspace $U \subseteq \mathbb{F}_{q^m}$ of dimension d . For this class of RS codes, we showed the existence of linear repair schemes, in which each surviving node transmits at most r \mathbb{F}_q -symbols for the repair of the erased node, for a wide range of parameters. This result relies on the existence of an \mathbb{F}_q -linear subspace $V \subseteq \mathbb{F}_{q^m}$ of dimension r for which the pair (U, V) is good. It also yields a practical probabilistic construction of a linear repair scheme. We also showed that if $r < m - s$, where q^s is the codimension of the RS code, and if V is chosen uniformly at random, then the probability that (U, V) is good is strictly less than one. Thus, in this case, the probabilistic construction is not trivial in the sense that not every pair (U, V) is good. Our results expand the Dau-Milenkovich scheme and one of the schemes of Li et al., for a wide range of parameters, where $r < m - s$.

Another contribution of this paper is that the presented scheme as a duality property in the following sense; A good pair (U, V) of \mathbb{F}_q -linear subspaces of dimensions d and r can be used to construct a good pair of \mathbb{F}_q -linear subspaces of dimensions $m-r$ and $m-d$, (V^\perp, \tilde{U}) , where $\tilde{U} = (U^{\wedge q^{s+1}})^\perp$. This duality property is useful for explicit constructions.

For a wide range of parameters, our scheme provides RS codes of codimension q^s , where the minimal value of s is $d(m-r)/m$. For future research, it will be interesting to understand if the this scheme is optimal for RS codes evaluated on linear subspaces.

APPENDIX

The purpose of this Appendix is to prove Proposition 16. That is, to show that if $1 \leq s < d < m$ and $ms \geq d(m-r)$, then for every \mathbb{F}_q -linear subspace $U \subseteq \mathbb{F}_{q^m}$ of dimension d , there exists an \mathbb{F}_q -linear subspace $V \subseteq \mathbb{F}_{q^m}$ of dimension r for which the pair (U, V) is not good if and only if $r < m - s$. In fact, we prove a somewhat stronger result, namely that for $r < m - s$, there exists a pair (U, V) that is not good even in the weaker sense, as the corresponding matrices M_1, M_2 defined in Proposition 7 satisfy that the column space of M_1 does not contained in the column space of M_2 . Similarly, if $r \geq m - s$, every pair (U, V) is good in the weaker sense.

We first prove the first part of Proposition 16, that is, the case $r \geq m - s$. The proof follows the lines of the proof of the DM scheme stated in Theorem 6.

Proof of Proposition 16 Part (1). First notice that, it is sufficient to prove the claim for $r = m - s$, since every \mathbb{F}_q -linear subspace V of dimension $r > m - s$ contains an \mathbb{F}_q -linear subspace, W , of dimension $m - s$, and if (U, W) is good then (U, V) is good (this holds even in the weaker sense).

As shown in the proof of Proposition 7, if the column space of M_1 is contained in the column space of M_2 (over \mathbb{F}_q), then for every $a_0 \in \mathbb{F}_{q^m}$, there exist $a_1, a_2, \dots, a_s \in \mathbb{F}_{q^m}$, such that the \mathbb{F}_q -linearized polynomial $f(X) = a_0X + a_1X^q + \dots + a_sX^{q^s}$ maps U to V . The other direction also holds.

Let V be an \mathbb{F}_q -subspace of dimension $r = m - s$. The image polynomial of V , $f_V^{\text{Im}}(X) = b_0X + b_1X^q + \dots + b_sX^{q^s} \in \mathbb{F}_{q^m}[X]$ is an \mathbb{F}_q -linearized polynomial of degree q^s that maps \mathbb{F}_{q^m} onto V . In particular $f_V^{\text{Im}}(U) \subseteq V$. Notice that, there exists a unique image polynomial of V , for all V (see [1] and the references therein). The kernel of $f_V^{\text{Im}}(X)$ is an \mathbb{F}_q -linear subspace of \mathbb{F}_{q^m} of dimension s . This implies that all the roots of $f_V^{\text{Im}}(X)$ are distinct, i.e., $f_V^{\text{Im}}(X)$ is separable, and hence $b_0 \neq 0$.

Now, for $a_0 \in \mathbb{F}_{q^m}$, let $f_{a_0}(X) = f_V^{\text{Im}}(a_0 \cdot b_0^{-1}X)$. Then $f_{a_0}(X) = a_0X + a_1X^q + \dots + a_sX^{q^s}$, for some $a_1, \dots, a_s \in \mathbb{F}_{q^m}$, and $f_{a_0}(X)$ maps \mathbb{F}_{q^m} to V . In particular, $f_{a_0}(U) \subseteq V$. \square

Next, we prove the second part of Proposition 16, namely, the case $r < m - s$. But first, we need the following lemma.

Lemma 22. *For a pair (U, V) of \mathbb{F}_q -linear subspaces of \mathbb{F}_{q^m} of dimensions d and r , respectively, and for the corresponding matrices M_1 and M_2 as defined in Proposition 7, the following are equivalent.*

- 1) *The column space of M_1 is contained in the column space of M_2 .*
- 2) *For every basis $\{u_1, \dots, u_d\}$ of U and for all $v'_1, \dots, v'_d \in V^\perp$ such that,*

$$T(u_1, u_2, \dots, u_d; s) \cdot (v'_1, v'_2, \dots, v'_d)^T = \mathbf{0} \quad (13)$$

we have that $\sum_{i=1}^d u_i v'_i = 0$.

- 3) *For every basis $B'_2 = \{b'_1, \dots, b'_{m-r}\}$ of V^\perp and for all $w_1, \dots, w_{m-r} \in U$, such that*

$$T(w_1, w_2, \dots, w_{m-r}; s) \cdot (b'_1, b'_2, \dots, b'_{m-r})^T = \mathbf{0} \quad (14)$$

we have that $\sum_{j=1}^{m-r} w_j b'_j = 0$.

Proof. We first prove that conditions (1) and (2) are equivalent. Let u_1, u_2, \dots, u_d be a basis for U . Note that, the column space of M_1 is contained in the column space of M_2 if and only if the left kernel of M_2 is contained in the left kernel of M_1 . Equivalently, for all $\mathbf{x} \in \mathbb{F}_q^{d(m-r)}$ for which $\mathbf{x}M_2 = \mathbf{0}$, we have that $\mathbf{x}M_1 = \mathbf{0}$.

The proof proceeds along the lines of the proof of Proposition 11, by representing the equations $\mathbf{x}M_2 = \mathbf{0}$ and $\mathbf{x}M_1 = \mathbf{0}$ as equation (13) and $\sum_{i=1}^d v'_i u_i = 0$, respectively.

To prove that conditions (2) and (3) are equivalent, we follow the lines of the corresponding part of the proof of Proposition 11. \square

Proof of Proposition 16 Part (2). By Lemma 22, given an \mathbb{F}_q -linear subspace $U \subseteq \mathbb{F}_{q^m}$ of dimension d , we need to show the existence of an \mathbb{F}_q -subspace $V \subseteq \mathbb{F}_{q^m}$ of dimension r , such that for some $w_1, w_2, \dots, w_{m-r} \in U$ and for some basis $B'_2 = \{b'_1, b'_2, \dots, b'_{m-r}\}$ of V^\perp we have that

$$T(w_1, \dots, w_{m-r}; s) \cdot (b'_1, b'_2, \dots, b'_{m-r})^T = \mathbf{0}$$

and

$$\sum_{j=1}^{m-r} w_j b'_j \neq 0.$$

Let $\tau = \min\{d, m - r\}$ and let $w_1, \dots, w_\tau \in U$ be any \mathbb{F}_q -linearly independent elements. Consider the matrix $T_1 = T(w_1, \dots, w_\tau; \tau - 1)$. By Proposition 10, we have that the rank of T_1 is $\tau - 1$. Hence there exist $b'_1, b'_2, \dots, b'_\tau \in \mathbb{F}_{q^m}$, not all zeros, such that $T_1 \cdot (b'_1, b'_2, \dots, b'_\tau)^T = \mathbf{0}$.

Next, we will show that b'_1, \dots, b'_τ are \mathbb{F}_q -linearly independent. Assume to the contrary that $b'_\tau = \sum_{i=1}^{\tau-1} a_i b'_i$, for some $a_1, \dots, a_{\tau-1} \in \mathbb{F}_q$. Then,

$$\begin{aligned} \mathbf{0} &= T_1 \cdot (b'_1, b'_2, \dots, b'_\tau)^T \\ &= T_1 \cdot (b'_1, b'_2, \dots, b'_{\tau-1}, \sum_{i=1}^{\tau-1} a_i b'_i)^T \\ &= T(w_1, w_2, \dots, w_{\tau-1}; \tau - 1) \cdot (b'_1, \dots, b'_{\tau-1})^T \\ &\quad + T(w_\tau; \tau - 1)(a_1 b'_1, a_2 b'_2, \dots, a_{\tau-1} b'_{\tau-1})^T \\ &= T(w'_1, w'_2, \dots, w'_{\tau-1}; \tau - 1) \cdot (b'_1, \dots, b'_{\tau-1})^T, \end{aligned}$$

where $w'_i = w_i + a_i w_\tau$, $1 \leq i \leq \tau - 1$. Since w_1, \dots, w_τ are \mathbb{F}_q -linearly independent, it follows that $w'_1, \dots, w'_{\tau-1}$ are also \mathbb{F}_q -linearly independent. By Proposition 10 we have that $T(w'_1, w'_2, \dots, w'_{\tau-1}; \tau - 1)$ is non-singular, hence $b'_1, \dots, b'_{\tau-1}$ must all be zeros and we derived a contradiction.

Define $w_{\tau+1} = \dots = w_{m-r} = 0$ and choose $b'_{\tau+1}, \dots, b'_{m-r}$ such that $B'_2 = \{b'_1, b'_2, \dots, b'_{m-r}\}$ is a basis for some \mathbb{F}_q -linear subspace V^\perp of dimension $m - r$.

Then, since $r < m - s$, it follows that $s < m - r$, and hence, recalling that $s < d$, we have $s < \tau$. Thus,

$$\begin{aligned} T(w_1, w_2, \dots, w_{m-r}; s) \cdot (b'_1, b'_2, \dots, b'_{m-r})^T &= \\ T(w_1, w_2, \dots, w_\tau; s) \cdot (b'_1, b'_2, \dots, b'_\tau)^T &= \mathbf{0}. \end{aligned}$$

Finally, we need to show that $\sum_{i=1}^{\tau} w_i b'_i \neq 0$. Let $\sigma^{-1} : \mathbb{F}_{q^m} \rightarrow \mathbb{F}_{q^m}$ be the inverse of the Frobenius map, $\sigma : \mathbb{F}_{q^m} \rightarrow \mathbb{F}_{q^m}$ ($\sigma(x) = x^q$, for all $x \in \mathbb{F}_{q^m}$). Let $z_i = \sigma^{-1}(w_i)$, for $1 \leq i \leq \tau$. We have that z_1, z_2, \dots, z_τ are \mathbb{F}_q -linearly independent and hence, by Proposition 10, $T(z_1, z_2, \dots, z_\tau; \tau)$ is non-singular over \mathbb{F}_{q^m} . Thus, $T(z_1, z_2, \dots, z_\tau; \tau) \cdot (b'_1, b'_2, \dots, b'_\tau)^T \neq \mathbf{0}$. Writing the equations, we have that for some $0 \leq \ell \leq \tau - 1$,

$$\sum_{i=1}^{\tau} w_i^{q^\ell} b'_i \neq 0.$$

However, b'_1, \dots, b'_τ satisfy that for all $1 \leq \ell \leq \tau - 1$,

$$\sum_{i=1}^{\tau} w_i^{q^\ell} b'_i = 0,$$

and thus $\sum_{i=1}^{\tau} w_i b'_i \neq 0$, which concludes the proof. \square

REFERENCES

- [1] E. Ben-Sasson and S. Kopparty, "Affine dispers from subspace polynomials," STOC 09, pp. 65–74.
- [2] H. Dau and O. Milenkovich, "Optimal repair schemes for some families of full-length Reed–Solomon codes," arXiv:1701.04120
- [3] A. G. Dimakis, P. B. Godfrey, Y. Wu, M. J. Wainwright, and K. Ramchadran, "Network coding for distributed storage systems," *IEEE Trans. Inform. Theory*, vol. 56, no. 9, pp. 4539–4551, Sept. 2010.

- [4] V. Guruswami and H. Jiang, "Near optimal repair of Reed–Solomon codes with low sub-packetization," in *Proc. ISIT 2019*, pp. 1077–1081.
- [5] V. Guruswami and M. Wootters, "Repairing Reed–Solomon codes," *IEEE Trans. Inform. Theory*, vol. 63, no. 9, pp. 5684–5698, Sept. 2017.
- [6] J. I. Hall, *Notes on Coding Theory*, available online at <http://users.math.msu.edu/users/jhall/classes/codenotes/Topstuff.pdf>
- [7] W. Li, Z. Wang, and H. Jafarkhani, "A tradeoff between the sub-packetization size and the repair bandwidth for Reed–Solomon codes," in *Proc. 55-th Annual Allerton Conf.*, Oct. 3–6 2017, pp. 942–949.
- [8] W. Li, Z. Wang, and H. Jafarkhani, "On the sub-packetization size and the repair bandwidth of Reed–Solomon codes," *IEEE Trans. Inform. Theory*, vol. 65, no. 9, pp. 5484–5502, Sept. 2019.
- [9] R. Lidl and H. Niederreiter, *Finite Fields*. CUP, 2nd Ed., 2000.
- [10] K. V. Rashmi, N. B. Shah, D. Gu, H. Kuang, D. Borthaker, and K. Ramchadran, "A Hitchhiker's guide to fast and efficient data reconstruction in erasure-coded data centers," in *ACM SIGCOMM Computer Communication Review*, vol. 44, no. 4, ACM, 2014, pp. 331–342.
- [11] I. Tamo, Z. Wang, and J. Bruck, "Zigzag codes: MDS array codes with optimal rebuilding," *IEEE Trans. Inform. Theory*, vol. 59, no. 3, pp. 1597–1616, Mar. 2013.
- [12] I. Tamo, M. Ye, and A. Barg, "The repair problem for Reed–Solomon codes: Optimal repair of single and multiple erasures," *IEEE Trans. Inform. Theory*, vol. 65, no. 5, pp. 2673–2695, May 2019.
- [13] M. Ye and A. Barg, "Explicit constructions of high-rate MDS array codes with optimal repair bandwidth," *IEEE Trans. Inform. Theory*, vol. 63, no. 4, pp. 2001–2014, Apr. 2017.