DTU Library

Maximum Sum-Rank Distance Codes over Finite Chain Rings

Martinez-Penas, Umberto; Puchinger, Sven

Published in:
IEEE Transactions on Information Theory

Link to article, DOI:
10.1109/TIT.2024.3354513

Publication date:
2024

Document Version
Peer reviewed version

Citation (APA):
Martinez-Penas, U., & Puchinger, S. (in press). Maximum Sum-Rank Distance Codes over Finite Chain Rings. *IEEE Transactions on Information Theory*. https://doi.org/10.1109/TIT.2024.3354513

Maximum Sum-Rank Distance Codes over Finite Chain Rings

Umberto Martínez-Peñas, and Sven Puchinger,

***Abstract*—In this work, maximum sum-rank distance (MSRD) codes and linearized Reed-Solomon codes are extended to finite chain rings. It is proven that linearized Reed-Solomon codes are MSRD over finite chain rings, extending the known result for finite fields. For the proof, several results on the roots of skew polynomials are extended to finite chain rings. These include the existence and uniqueness of minimum-degree annihilator skew polynomials and Lagrange interpolator skew polynomials. A general cubic-complexity sum-rank Welch-Berlekamp decoder and a quadratic-complexity sum-rank syndrome decoder (under some assumptions) are then provided over finite chain rings. The latter also constitutes the first known syndrome decoder for linearized Reed–Solomon codes over finite fields. Finally, applications in Space-Time Coding with multiple fading blocks and physical-layer multishot Network Coding are discussed.**

***Index Terms*—Finite chain rings, linearized Reed-Solomon codes, maximum sum-rank distance codes, sum-rank metric, syndrome decoding, Welch-Berlekamp decoding.**

I. Introduction

The sum-rank metric, introduced in [31], is a natural generalization of both the Hamming metric and the rank metric. Codes considered with respect to the sum-rank metric over finite fields have applications in multishot Network Coding [25], [31], Space-Time Coding with multiple fading blocks [20], [38] and local repair in Distributed Storage [26]. However, codes over rings may be more suitable for physical-layer Network Coding [6], where alphabets are subsets or lattices of the complex field instead of finite fields. Similarly, finite rings derived from the complex field allows for more flexible choices of constellations to construct Space-Time codes [9], [10].

Maximum sum-rank distance (MSRD) codes are those codes whose minimum sum-rank distance attains the Singleton bound. Among known MSRD codes over finite fields, linearized Reed-Solomon codes [22] are those with smallest finite-field sizes (thus more computationally efficient) for the main parameter regimes, see [24, Table 1] and [24, Sec. 2.4]. Furthermore, they cover a wide range of parameter values and are the only known MSRD codes compatible with square matrices. Linearized Reed-Solomon codes include both generalized Reed-Solomon codes [36] and Gabidulin codes [7], whenever the sum-rank metric includes the Hamming metric and the rank metric, respectively. Reed-Solomon codes over rings were systematically studied for the first time in [35]. Gabidulin codes over Galois rings were introduced in [10], and later extended to finite principal ideal rings in [9]. Such families of Gabidulin codes over rings were proposed for Space-Time Coding in the case of a single fading block in [9], [10], and they were proposed for physical-layer singleshot Network Coding in [9].

In this work, we introduce and study MSRD codes and linearized Reed-Solomon codes over finite chain rings, together with their applications in Space-Time Coding with multiple fading blocks and physical-layer multishot Network Coding. Finite chain rings are those finite rings whose family of ideals form a chain with respect to set inclusion. They are an important subfamily of finite principal ideal rings. In fact, finite principal ideal rings are all Cartesian products of finite chain rings [28, Th. VI.2]. Moreover, finite chain rings include Galois rings (although not all finite chain rings are Galois rings [28, Th. XVII.5]), which are of the form $\mathbb{Z}_{p^r}[x]/(F)$, where p is a prime number, r is a positive integer, and $F \in \mathbb{Z}_{p^r}[x]$ is a polynomial whose reduction modulo p is irreducible. Galois rings hence include finite fields (when $r = 1$) and finite rings of the form $\mathbb{Z}_{p^r}$ (when $F = x$). Finite chain rings also include quotients of subrings of the complex field of the form $\mathbb{Z}_{2^r}[i] = \mathbb{Z}[i]/(2^r)$, where i is the imaginary unit, see [28, Th. XVII.5].

The contributions and organization of this manuscript are as follows. In Section II, we collect some preliminaries on finite chain rings. In Section III, we define the sum-rank metric over finite chain rings, together with the corresponding Singleton bound and the definition of MSRD codes. Section IV contains the theoretical building blocks for constructing linearized Reed-Solomon codes and their decoding. We extend Lam and Leroy's results [12], [13] relating roots and degrees of skew polynomials to the case of finite chain rings. As a result, we prove the existence and uniqueness of minimum-degree annihilator skew polynomials and Lagrange interpolator skew polynomials, and we describe when the corresponding extended Moore matrices are invertible. In Section V, we introduce linearized Reed-Solomon codes over finite

This work was done in part while U. Martínez-Peñas was with the Institutes of Computer Science and Mathematics, University of Neuchâtel, Switzerland. He is now with the University of Valladolid, Spain. During this period he was supported by a María Zambrano contract by the University of Valladolid, Spain, Contract no. E-47-2022-0001486 and by Grant TED2021-130358B-I00 funded by MCIN/AEI/10.13039/501100011033 and by the "European Union NextGenerationEU/PRTR" (e-mail: umberto.martinez@uva.es).

This work was done while S. Puchinger was with the Department of Applied Mathematics and Computer Science, Technical University of Denmark (DTU), Lyngby, Denmark and the Department of Electrical and Computer Engineering, Technical University of Munich, Munich, Germany. Within this period he was supported by the European Union's Horizon 2020 research and innovation program under the Marie Sklodowska-Curie grant agreement no. 713683 and by the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme, grant agreement no. 801434 (e-mail: mail@svenpuchinger.de).

Parts of this paper have been presented at The Twelfth International Workshop on Coding and Cryptography (WCC 2022), Rostock, Germany, March, 2022 [27].

chain rings and use the previous results to prove that they are MSRD. In Section VI, we provide a cubic-complexity Welch-Berlekamp decoder with respect to the sum-rank metric for linearized Reed–Solomon codes over finite chain rings that works in general. Then, in Section VII, we provide a quadratic-complexity syndrome decoder with respect to the sum-rank metric for linearized Reed–Solomon codes over finite chain rings that work under some (not very strict) assumptions. This decoder also constitutes the first known syndrome decoder for linearized Reed–Solomon codes over finite fields, to the best of our knowledge. Finally, in Section VIII, we discuss applications in Space-Time Coding with multiple fading blocks and physical-layer multishot Network Coding.

We conclude by mentioning that there are other constructions of MSRD codes in the case of finite fields, in particular using different geometric points of view [5], [24], [29], [30]. However, we leave as an open problem generalizing them to finite chain rings.

NOTATION

Let m and n be positive integers. We denote $[n] = \{1, 2, \ldots, n\}$. For a set $\mathcal{A}$, we denote by $\mathcal{A}^{m\times n}$ the set of $m \times n$ matrices with entries in $\mathcal{A}$, and we denote $\mathcal{A}^n = \mathcal{A}^{1\times n}$. All rings are considered with identity, and ring morphisms map identities to identities. Unless otherwise stated, rings are assumed to be commutative. For a ring R, we denote by R^* the set of units of R. For $a \in R$, we denote by (a) the ideal generated by a.

II. PRELIMINARIES ON FINITE CHAIN RINGS

In this preliminary section, we introduce and revisit some important properties of finite chain rings. We refer the reader to [28] for more details.

A *local ring* is a ring with only one maximal ideal, and a *chain ring* is a ring whose ideals form a chain with respect to set inclusion, thus being a local ring. Throughout this manuscript, we fix a finite chain ring R, meaning a chain ring of finite size. We will denote by $\mathfrak{m}$ the maximal ideal of R. Since R is finite and $R/\mathfrak{m}$ is a field, then it must be a finite field. We will fix the prime power $q = |R/\mathfrak{m}|$, and we denote $\mathbb{F}_q = R/\mathfrak{m}$, the *finite field* with q elements.

Let $H \in R[x]$ be a monic polynomial of degree m whose image in $\mathbb{F}_q[x]$ is irreducible. Throughout this manuscript, we will fix $S = R[x]/(H)$. The ring S is a free local Galois extension of R (hence a free R-module) of rank m with maximal ideal $\mathfrak{M} = \mathfrak{m}S$. Furthermore, the Galois group of $R \subseteq S$ is cyclic of order m, and generated by a ring automorphism $\sigma : S \longrightarrow S$ such that $R = \{a \in S \mid \sigma(a) = a\}$ and $\sigma(c) = c^q$, for some primitive element $c \in S$. Moreover, it holds that $S/\mathfrak{M} = \mathbb{F}_{q^m}$, and we have a commutative diagram

$$\begin{array}{ccc} S & \xrightarrow{\sigma} & S \\ \rho\downarrow & & \downarrow\rho \\ \mathbb{F}_{q^m} & \xrightarrow{\overline{\sigma}} & \mathbb{F}_{q^m}, \end{array} \tag{1}$$

where $\rho : S \longrightarrow S/\mathfrak{M} = \mathbb{F}_{q^m}$ is the natural projection map, and $\overline{\sigma}(a) = a^q$, for all $a \in \mathbb{F}_{q^m}$. In other words, $\rho(\sigma(a)) = \overline{\sigma}(\rho(a))$, for all $a \in S$. We will usually denote $\overline{a} = \rho(a)$, and therefore, we have that $\overline{\sigma(a)} = \overline{\sigma}(\overline{a})$, for $a \in S$.

Example 1. Let $R = \mathbb{Z}_9$, that is, the ring of integers modulo 9. It is clearly a finite chain ring with maximal ideal $\mathfrak{m} = (3)$. Its residue field is the finite field $R/\mathfrak{m} = \mathbb{F}_3 = \mathbb{Z}_3$ and $q = 3$. We may choose $H = x^2 + 1$ (i.e., $m = 2$) and construct the finite residue ring $S = R[x]/(H) = \mathbb{Z}_9[x]/(x^2 + 1)$. Denote by $\alpha \in S$ the image of x in S, which satisfies $\alpha^2 + 1 = 0$. The set S is then

$$S = \{a\alpha + b \mid a, b \in \mathbb{Z}_9\}.$$

We may define the morphism $\sigma : S \longrightarrow S$ given by $\sigma(\alpha) = \alpha^3$ and being the identity on $\mathbb{Z}_9$. It is well defined since $(\alpha^3)^2 + 1 = 0$ and it is an automorphism that generates the Galois group of S over R since $\sigma^2 = \mathrm{Id}$ is the identity map. Notice that $m = 2$ in this case and we have the residue field $S/\mathfrak{M} = \mathbb{F}_9$.

An important feature of local rings is that the group of units is formed by the elements outside of the maximal ideal. That is, $R^* = R \setminus \mathfrak{m}$ and $S^* = S \setminus \mathfrak{M}$. As stated above, S is a free R-module of rank m, and any basis of S over R has m elements. Finally, the following technical lemma will be useful for our purposes.

Lemma 1. *Let $\beta_1, \beta_2, \ldots, \beta_r \in S$ be R-linearly independent (thus $r \leq m$).*

1) *There exist $\beta_{r+1}, \ldots, \beta_m \in S$ such that $\beta_1, \beta_2, \ldots, \beta_m$ form a basis of S over R.*
2) *The projections $\overline{\beta}_1, \overline{\beta}_2, \ldots, \overline{\beta}_r \in \mathbb{F}_{q^m}$ are $\mathbb{F}_q$-linearly independent.*
3) $\beta_1, \beta_2, \ldots, \beta_r \in S^*$.

Proof. Item 1 is a particular case of [28, p. 92, ex. V.14]. Now, since $\beta_1, \beta_2, \ldots, \beta_m$ are generators of S over R, then $\overline{\beta}_1, \overline{\beta}_2, \ldots, \overline{\beta}_m$ are generators of $\mathbb{F}_{q^m}$ over $\mathbb{F}_q$, thus they are a basis since there are m of them. Thus $\overline{\beta}_1, \overline{\beta}_2, \ldots, \overline{\beta}_r$ are $\mathbb{F}_q$-linearly independent. Finally, Item 3 is [9, Lemma 2.4] but is also trivial from Item 2 since $S^* = S \setminus \mathfrak{M}$. □

III. MSRD CODES ON FINITE CHAIN RINGS

The sum-rank metric over fields was first defined in [31] under the name *extended distance*, although it was previously used in the Space-Time Coding literature [20, Sec. III]. Later, the rank metric was extended to finite principal ideal rings in [9]. In this section, we will introduce the sum-rank metric for finite chain rings.

Since R is a finite chain ring, then it is a principal ideal ring. Therefore, given a matrix $\mathbf{A} \in R^{m\times n}$, there exist two invertible matrices $\mathbf{P} \in R^{m\times m}$ and $\mathbf{Q} \in R^{n\times n}$, and a (possibly rectangular) diagonal matrix $\mathbf{D} = \mathrm{Diag}(d_1, d_2, \ldots, d_r) \in R^{m\times n}$, with $r = \min\{m, n\}$, such that $\mathbf{A} = \mathbf{PDQ}$. The elements $d_1, d_2, \ldots, d_r \in R$ are unique up to multiplication by units and permutation [40] and the diagonal matrix $\mathbf{D}$ is called the *Smith normal form* of $\mathbf{A}$. Hence we may define ranks and free ranks as in [9, Def. 3.3].

Definition 1. Given $\mathbf{A} \in R^{m\times n}$ with Smith normal form $\mathbf{D} = \mathrm{Diag}(d_1, d_2, \ldots, d_r) \in R^{m\times n}$, $r = \min\{m, n\}$, we define:

1) The rank of $\mathbf{A}$ as $\mathrm{rk}(\mathbf{A}) = |\{i \in [r] \mid d_i \neq 0\}|$.
2) The free rank of $\mathbf{A}$ as $\mathrm{frk}(\mathbf{A}) = |\{i \in [r] \mid d_i \in R^*\}|$.

In this manuscript, we will mainly work with linear codes in S^n. To that end, we will translate the rank metric from $R^{m\times n}$ to S^n as in [9, Sec. III-B]. For a positive integer t and an ordered basis $\boldsymbol{\alpha} = (\alpha_1, \alpha_2, \ldots, \alpha_m) \in S^m$ of S over R, we define the matrix representation map $M_{\boldsymbol{\alpha}} : S^t \longrightarrow R^{m\times t}$ by

$$M_{\boldsymbol{\alpha}}(\mathbf{c}) = \begin{pmatrix} c_{1,1} & c_{1,2} & \cdots & c_{1,t} \\ c_{2,1} & c_{2,2} & \cdots & c_{2,t} \\ \vdots & \vdots & \ddots & \vdots \\ c_{m,1} & c_{m,2} & \cdots & c_{m,t} \end{pmatrix} \in R^{m\times t}, \qquad (2)$$

where $\mathbf{c} = (c_1, c_2, \ldots, c_t) \in S^t$ and, for each $j \in [t]$, $c_{1,j}, c_{2,j}, \ldots, c_{m,j} \in R$ are the coordinates of c_j in the ordered basis $\boldsymbol{\alpha}$, that is, they are the unique scalars in R such that $c_j = \sum_{i=1}^m \alpha_i c_{i,j}$. Notice that also $\mathbf{c} = \sum_{i=1}^m \alpha_i (c_{i,1}, c_{i,2}, \ldots, c_{i,t})$. Using this matrix representation map, we define $\mathrm{rk}(\mathbf{c}) = \mathrm{rk}(M_{\boldsymbol{\alpha}}(\mathbf{c}))$ and $\mathrm{frk}(\mathbf{c}) = \mathrm{frk}(M_{\boldsymbol{\alpha}}(\mathbf{c}))$, which do not depend on the ordered basis $\boldsymbol{\alpha}$, see also [9].

We may now define the sum-rank metric for the ring extension $R \subseteq S$.

Definition 2 (**Sum-rank metric**)**.** Consider positive integers $n_1, n_2, \ldots, n_\ell$ and $n = n_1 + n_2 + \cdots + n_\ell$. We define the sum-rank weight of $\mathbf{c} \in S^n$ over R for the length partition $n = n_1 + n_2 + \cdots + n_\ell$ as

$$\mathrm{wt}_{SR}(\mathbf{c}) = \sum_{i=1}^{\ell} \mathrm{rk}\left(\mathbf{c}^{(i)}\right),$$

where $\mathbf{c} = \left(\mathbf{c}^{(1)}, \mathbf{c}^{(2)}, \ldots, \mathbf{c}^{(\ell)}\right)$ and $\mathbf{c}^{(i)} \in S^{n_i}$, for $i \in [\ell]$. We define the sum-rank metric $\mathrm{d}_{SR} : S^{2n} \longrightarrow S^n$ over R for the length partition $n = n_1 + n_2 + \cdots + n_\ell$ by

$$\mathrm{d}_{SR}(\mathbf{c}, \mathbf{d}) = \mathrm{wt}_{SR}(\mathbf{c} - \mathbf{d}),$$

for $\mathbf{c}, \mathbf{d} \in S^n$.

This definition coincides with the classical one [20], [31] when R and S are fields. Over finite chain rings, this definition coincides with the Hamming metric when $n_1 = n_2 = \ldots = n_\ell = 1$ and with the rank metric as above [9] when $\ell = 1$.

Once again, the definitions of the sum-rank weight and metric in S^n do not depend on the ordered basis $\boldsymbol{\alpha}$. Furthermore, the sum-rank metric satisfies the properties of a metric since rank weights are norms by [9, Th. 3.9]. As noted in [9, Remark 3.10], free ranks do not generally give rise to a metric nor include the Hamming metric over rings. The subring R and the length partition $n = n_1+n_2+\cdots+n_\ell$ will not be specified unless necessary.

The following result will be crucial for our purposes. It can be proven as in [26, Th. 1], but using the Smith normal form.

Lemma 2. *Given $\mathbf{c} \in S^n$, and considering the subring $R \subseteq S$ and the length partition $n = n_1 + n_2 + \cdots + n_\ell$, it holds that*

$$\mathrm{wt}_{SR}(\mathbf{c}) = \min\{ \mathrm{wt}_H(\mathbf{c}\,\mathrm{Diag}(\mathbf{A}_1, \mathbf{A}_2, \ldots, \mathbf{A}_\ell)) \mid \mathbf{A}_i \in R^{n_i\times n_i} \textit{ invertible}, i \in [\ell]\}.$$

In particular, given an arbitrary code $\mathcal{C} \subseteq S^n$ (linear or not),we have that

$$\mathrm{d}_{SR}(\mathcal{C}) = \min\{ \mathrm{d}_H(\mathcal{C}\,\mathrm{Diag}(\mathbf{A}_1, \mathbf{A}_2, \ldots, \mathbf{A}_\ell)) \mid \mathbf{A}_i \in R^{n_i\times n_i} \textit{ invertible}, i \in [\ell]\}.$$

One immediate consequence of Lemma 2 above is the following classical version of the Singleton bound, but for the sum-rank metric for the ring extension $R \subseteq S$. This bound recovers [22, Prop. 34] when R and S are fields, and it recovers [9, Prop. 3.20] when $\ell = 1$.

Proposition 1 (**Singleton bound**)**.** *Given an arbitrary code $\mathcal{C} \subseteq S^n$ (linear or not), and setting $k = \log_{|S|} |\mathcal{C}|$, it holds that*

$$\mathrm{d}_{SR}(\mathcal{C}) \leq n - k + 1.$$

We note that there exist more general Singleton bounds for the sum-rank metric over finite fields, see [5, Th. III.2]. We leave as an open problem generalizing such bounds to finite chain rings.

Thus we may define MSRD codes as follows. This definition recovers that of MSRD codes [22, Th. 4] when R and S are fields, MDS codes over finite chain rings when $n_1 = n_2 = \ldots = n_\ell = 1$, and MRD codes over finite chain rings [9, Def. 3.21] when $\ell = 1$.

Definition 3 (**MSRD codes**)**.** We say that a code $\mathcal{C} \subseteq S^n$ is a *maximum sum-rank distance* (MSRD) code over R for the length partition $n = n_1 + n_2 + \cdots + n_\ell$ if $k = \log_{|S|} |\mathcal{C}|$ is a positive integer and $\mathrm{d}_{SR}(\mathcal{C}) = n - k + 1$, where d_{SR} is considered for such a subring and length partition.

From Lemma 2, we deduce the following auxiliary lemma, which we will use in Section V to prove that linearized Reed–Solomon codes are MSRD.

Lemma 3. *Given an arbitrary code $\mathcal{C} \subseteq S^n$ (linear or not) such that $k = \log_{|S|} |\mathcal{C}|$ is a positive integer, and for the subring $R \subseteq S$ and length partition $n = n_1 + n_2 + \cdots + n_\ell$, it holds that $\mathcal{C}$ is MSRD if, and only if, the code $\mathcal{C}\,\mathrm{Diag}(\mathbf{A}_1, \mathbf{A}_2, \ldots, \mathbf{A}_\ell)$ is MDS, for all invertible matrices $\mathbf{A}_i \in R^{n_i\times n_i}$, for $i \in [\ell]$.*

IV. Skew Polynomials on Finite Chain Rings

We will extensively use *skew polynomials* [32], but defined over finite chain rings instead of fields or division rings. The ring of skew polynomials over S with morphism σ is the set $S[x;\sigma]$ formed by elements of the form $F = F_0 + F_1x + F_2x^2 + \cdots + F_dx^d$, for $F_0, F_1, F_2, \ldots, F_d \in S$ and $d \in \mathbb{N}$, as in the conventional polynomial ring. Furthermore, if $F_d \neq 0$, then we define the *degree* of F as $\deg(F) = d$, and we say that F is *monic* if $F_d = 1$. If $F = 0$, then we define $\deg(F) = -\infty$. Moreover, sums of skew polynomials and products with scalars on the left are defined as in the case of conventional polynomials. The only difference is that the product of skew polynomials is given by the rule

$$xa = \sigma(a)x,$$

for $a \in S$, together with the rule $x^ix^j = x^{i+j}$, for $i, j \in \mathbb{N}$.

In order to define linearized Reed–Solomon codes for the extension $R \subseteq S$, we will need the following definitions. We start with the following operators, considered in [14, Def. 3.1] and [15, Eq. (2.7)] for division rings. The definition can be trivially adapted to finite chain rings.

Definition 4 ([14], [15]). Fix $a \in S$ and define its ith norm as $N_i(a) = \sigma^{i-1}(a) \cdots \sigma(a)a$ for $i \in \mathbb{N}$. Now define the R-linear operator $\mathcal{D}_a^i : S \longrightarrow S$ by

$$\mathcal{D}_a^i(\beta) = \sigma^i(\beta)N_i(a), \tag{3}$$

for all $\beta \in S$, and all $i \in \mathbb{N}$. Define also $\mathcal{D}_a = \mathcal{D}_a^1$ and observe that $\mathcal{D}_a^{i+1} = \mathcal{D}_a \circ \mathcal{D}_a^i$, for $i \in \mathbb{N}$. If σ is not understood from the context, we will write $N_i^\sigma(a)$ and $\mathcal{D}_{\sigma,a}^i(\beta)$, for $i \in \mathbb{N}$, $a, \beta \in S$.

Finally, given a skew polynomial $F = \sum_{i=0}^d F_i x^i \in S[x;\sigma]$, where $d \in \mathbb{N}$, we define its operator evaluation on the pair $(a, \beta) \in S^2$ as

$$F_a(\beta) = \sum_{i=0}^d F_i \mathcal{D}_a^i(\beta) \in S.$$

Observe that F_a can be seen as an R-linear map $F_a : S \longrightarrow S$, taking $\beta \in S$ to $F_a(\beta) \in S$.

Example 2. Let the setting be as in Example 1. Choose $a = \alpha + 4$ and $\beta = \alpha$. Then

$$\mathcal{D}_a^2(\beta) = \sigma^2(\beta)\sigma(a)a = \alpha^9 \cdot (\alpha^3 + 4) \cdot (\alpha + 4) = 8\alpha.$$

We will also need the concept of *conjugacy*, introduced in [12], [13] for division rings. We adapt the definition to finite chain rings as follows.

Definition 5 (Conjugacy [12], [13]). Given $a, b \in S$, we say that they are *conjugate* in S with respect to σ if there exists $\beta \in S^*$ such that $b = a^\beta$, where

$$a^\beta = \sigma(\beta)a\beta^{-1}.$$

We now extend some results by Lam and Leroy [12], [13] to finite chain rings. These results will be crucial for defining and studying linearized Reed–Solomon codes.

The following result follows by combining [12, Th. 23] and [13, Th. 4.5], and was presented in the following form in [16, Th. 2.1]. We only consider finite fields.

Lemma 4 ([12], [13]). *If $\overline{a}_1, \overline{a}_2, \ldots, \overline{a}_\ell \in \mathbb{F}_{q^m}^*$ are pair-wise non-conjugate (with respect to $\overline{\sigma}$) and $F \in \mathbb{F}_{q^m}[x;\overline{\sigma}]$ is not zero, then*

$$\sum_{i=1}^{\ell} \dim_{\mathbb{F}_q}(\ker(F_{\overline{a}_i})) \leq \deg(F).$$

We now extend this result to the finite chain rings $R \subseteq S$ (we will give a different extension in Lemma 10). To this end, we define the *free rank* of an R-module M as the maximum size of an R-linearly independent subset of M. We will denote it by $\mathrm{frk}_R(M)$.

Theorem 1. *Let $a_1, a_2, \ldots, a_\ell \in S^*$ be such that $a_i - a_j^\beta \in S^*$, for all $\beta \in S^*$, and for $1 \leq i < j \leq \ell$. For any non-zero monic $F \in S[x;\sigma]$, we have*

$$\sum_{i=1}^{\ell} \mathrm{frk}_R(F_{a_i}^{-1}(\mathfrak{M})) \leq \deg(F).$$

Proof. If $F = F_0 + F_1 x + \cdots + F_d x^d$, where $F_0, F_1, \ldots, F_d \in S$, denote $\overline{F} = \overline{F}_0 + \overline{F}_1 x + \cdots + \overline{F}_d x^d \in \mathbb{F}_{q^m}[x;\overline{\sigma}]$. We have the following two facts:

1) We have that

$$\mathrm{frk}_R(F_a^{-1}(\mathfrak{M})) \leq \dim_{\mathbb{F}_q}(\ker(\overline{F}_{\overline{a}})).$$

We now prove this claim. From Definition 4 and the fact that $\overline{\sigma(a)} = \overline{\sigma}(\overline{a})$ (see (1)),

$$\overline{F}_{\overline{a}}(\overline{\beta}) = \overline{F_a(\beta)},$$

for all $a, \beta \in S$. This means that, if $F_a(\beta) \in \mathfrak{M}$, then $\overline{F}_{\overline{a}}(\overline{\beta}) = \overline{F_a(\beta)} = 0$. Therefore, $\overline{F_a^{-1}(\mathfrak{M})} \subseteq \ker(\overline{F}_{\overline{a}})$. By Item 2 in Lemma 1, $\mathrm{frk}_R(F_a^{-1}(\mathfrak{M})) \leq \dim_{\mathbb{F}_q}(\overline{F_a^{-1}(\mathfrak{M})})$. Thus we conclude that $\mathrm{frk}_R(F_a^{-1}(\mathfrak{M})) \leq \dim_{\mathbb{F}_q}(\overline{F_a^{-1}(\mathfrak{M})}) \leq \dim_{\mathbb{F}_q}(\ker(\overline{F}_{\overline{a}}))$.

2) For $1 \leq i < j \leq \ell$ and $\overline{\beta} \in \mathbb{F}_{q^m}^*$, we have that $\overline{a}_i \neq \overline{a}_j^{\overline{\beta}}$ since $\beta \in S^*$ and $a_i - a_j^\beta \notin \mathfrak{M}$.

By 2), Lemma 4 applies and, using 1), we conclude that

$$\begin{aligned}\sum_{i=1}^{\ell} \mathrm{frk}_R(F_{a_i}^{-1}(\mathfrak{M})) &\leq \sum_{i=1}^{\ell} \dim_{\mathbb{F}_q}(\ker(\overline{F}_{\overline{a}_i})) \\ &\leq \deg(\overline{F}) \\ &= \deg(F),\end{aligned}$$

where $\deg(F) = \deg(\overline{F})$, since F is non-zero and monic. $\square$

Using Theorem 1, we may prove the existence of monic annihilator skew polynomials and Lagrange interpolating skew polynomials of the smallest possible degree. To this end, we need more auxiliary tools. First, we need the following alternative notion of evaluation, introduced in [12], [13] for division rings and based on right Euclidean division [32]. The adaptation to finite chain rings is trivial.

Definition 6 ([12], [13]). Given a skew polynomial $F \in S[x;\sigma]$ and $a \in S$, we define the remainder evaluation of F at a, denoted by $F(a)$, as the only scalar $F(a) \in S$ such that there exists $Q \in S[x;\sigma]$ with $F = Q \cdot (x - a) + F(a)$.

We will also need the product rule, given in [13, Th. 2.7] for division rings, but which holds for finite chain rings as stated below.

Lemma 5 ([13]). *Let $F, G \in S[x;\sigma]$ and $a \in S$. If $G(a) = 0$, then $(FG)(a) = 0$. If $\beta = G(a) \in S^*$, then $(FG)(a) = F(a^\beta)G(a)$.*

Another tool that we will need is the following connection between the remainder evaluation as above and the evaluation from Definition 4. It was proven in [12, Lemma 1] for division rings, but it holds for finite chain rings as stated below.

Lemma 6 ([12]). *Given $F \in S[x;\sigma]$, $a \in S$ and $\beta \in S^*$, it holds that*

$$F_a(\beta) = F(a^\beta)\beta.$$

We will show that annihilator skew polynomials and Lagrange interpolating skew polynomials exist for sequences of evaluation points as follows.

Definition 7. Consider vectors $\mathbf{a} = (a_1, a_2, \ldots, a_\ell) \in (S^*)^\ell$ and $\boldsymbol{\beta}_i = (\beta_{i,1}, \beta_{i,2}, \ldots, \beta_{i,n_i}) \in S^{n_i}$, for $i \in [\ell]$. Set $\boldsymbol{\beta} = (\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \ldots, \boldsymbol{\beta}_\ell)$. We say that $(\mathbf{a}, \boldsymbol{\beta})$ satisfies the *MSRD property* if the following conditions hold:

1) $a_i - a_j^\beta \in S^*$, for all $\beta \in S^*$ and for $1 \leq i < j \leq \ell$.
2) $\beta_{i,1}, \beta_{i,2}, \ldots, \beta_{i,n_i}$ are linearly independent over R, for $i \in [\ell]$.

Note that, by Item 3 in Lemma 1, $\beta_{i,j} \in S^*$, for $j \in [n_i]$ and $i \in [\ell]$.

Example 3. Let the setting be as in Example 1. Let $a_1 = 1$, $a_2 = \alpha + 1$. Their images $\overline{1}, \overline{\alpha} + \overline{1}$ in $\mathbb{F}_9$ satisfy

$$N_{\mathbb{F}_9/\mathbb{F}_3}(\overline{\alpha} + \overline{1}) = (\overline{\alpha}^3 + \overline{1})(\overline{\alpha} + \overline{1}) = \overline{2} \neq \overline{1} = N_{\mathbb{F}_9/\mathbb{F}_3}(\overline{1}),$$

where $N_{\mathbb{F}_9/\mathbb{F}_3}$ is the norm of the field extension $\mathbb{F}_3 \subseteq \mathbb{F}_9$. Thus by Hilbert's Theorem 90 it holds that $a_1 - a_2^\beta \in S^*$, for all $\beta \in S^*$.

Finally choose $\beta_1 = 1$ and $\beta_2 = \alpha$, which are clearly R-linearly independent, and set $\mathbf{a} = (a_1, a_2)$ and $\boldsymbol{\beta} = (\beta_1, \beta_2)$. Therefore $(\mathbf{a}, \boldsymbol{\beta}) = ((1, \alpha + 1), (1, \alpha))$ satisfies the MSRD property.

The next step is the existence of minimal annihilator skew polynomials of the "right" degree. The following proposition recovers [9, Prop. 2.5] when $\ell = 1$ and $a_1 = 1$.

Theorem 2. *Let $(\mathbf{a}, \boldsymbol{\beta})$ be as in Definition 7, and satisfying the MSRD property. Then there exist units $\gamma_{i,j} \in S^*$, with $\gamma_{1,1} = \beta_{1,1}$, and skew polynomials of the form*

$$\begin{aligned} G_{i,j} = &\left(x - a_i^{\gamma_{i,j}}\right) \cdots \left(x - a_i^{\gamma_{i,1}}\right) \cdot \\ &\left(x - a_{i-1}^{\gamma_{i-1,n_{i-1}}}\right) \cdots \left(x - a_{i-1}^{\gamma_{i-1,1}}\right) \cdots \\ &\left(x - a_1^{\gamma_{1,n_1}}\right) \cdots \left(x - a_1^{\gamma_{1,1}}\right) \in S[x;\sigma], \end{aligned}$$

of degree $\deg(G_{i,j}) = \sum_{u=1}^{i-1} n_u + j$, and such that

$$\begin{aligned} G_{i,j}(a_u^{\beta_{u,v}}) = 0, &\text{ if } 1 \leq u \leq i-1, \\ &\text{or if } u = i \text{ and } 1 \leq v \leq j, \\ G_{i,j}(a_u^{\beta_{u,v}}) \in S^*, &\text{ if } i+1 \leq u \leq \ell, \\ &\text{or if } u = i \text{ and } j+1 \leq v \leq n_i, \end{aligned}$$

and $G_{i,j}$ is unique among monic skew polynomials in $S[x;\sigma]$ satisfying such properties, for $j \in [n_i]$ and $i \in [\ell]$.

Proof. We prove the proposition by induction in the pair (i,j). For the basis step, we only need to define $G_{1,1} = x - a_1^{\beta_{1,1}}$. We have $G_{1,1,a_1}(\beta_{1,1}) = 0$ by Lemma 6. On the other hand, since $\deg(G_{1,1}) = 1$ and it is non-zero and monic, then $G_{1,1,a_u}(\beta_{u,v}) \in S^*$, if $(u,v) \neq (1,1)$, by Theorem 1 and Lemma 6.

Now, we have two cases for the inductive step. Either we go from $G_{i,j}$ to $G_{i,j+1}$, if $j < n_i$, or from G_{i,n_i} to $G_{i+1,1}$ if $i < \ell$. The process stops when $i = \ell$ and $j = n_\ell$. We will only develop the first case of induction step, since the second case is analogous.

Assume that $G_{i,j}$ satisfies the properties in the proposition and $j < n_i$. In particular, $G_{i,j}(a_i^{\beta_{i,j+1}}) \in S^*$. Thus, we may define $\gamma_{i,j+1} = G_{i,j}(a_i^{\beta_{i,j+1}})\beta_{i,j+1} \in S^*$ and

$$G_{i,j+1} = \left(x - a_i^{\gamma_{i,j+1}}\right) G_{i,j}.$$

By Lemmas 5 and 6 and the assumptions on $G_{i,j}$, we have that $G_{i,j+1}(a_u^{\beta_{u,v}}) = 0$, if $1 \leq u \leq i-1$, or if $u = i$ and $1 \leq v \leq j+1$. Since $G_{i,j+1}$ has such a set of zeros, it is non-zero, monic and of degree $\sum_{u=1}^{i-1} n_u + j + 1$, then we deduce from Theorem 1 and Lemma 6 that $G_{i,j+1}(a_u^{\beta_{u,v}}) \in S^*$, if $i+1 \leq u \leq \ell$, or if $u = i$ and $j+2 \leq v \leq n_i$.

Finally, the uniqueness of $G_{i,j}$ follows by combining Theorem 1 and Lemma 6. $\square$

We immediately deduce the following two consequences. The first of these corollaries is the existence of annihilator skew polynomials of minimum possible degree.

Corollary 1. *Let $(\mathbf{a}, \boldsymbol{\beta})$ be as in Definition 7, and satisfying the MSRD property. Then there exists a unique monic skew polynomial $F \in S[x;\sigma]$ such that $\deg(F) = n_1 + n_2 + \cdots + n_\ell$ and $F_{a_i}(\beta_{i,j}) = 0$, for $j \in [n_i]$ and $i \in [\ell]$.*

Proof. Take $F = G_{\ell,n_\ell}$ in Theorem 2. $\square$

The second corollary states the existence of a basis for Lagrange interpolation.

Corollary 2. *Let $(\mathbf{a}, \boldsymbol{\beta})$ be as in Definition 7, and satisfying the MSRD property. For each $j \in [n_i]$ and $i \in [\ell]$, there exists a unique skew polynomial $F_{i,j} \in S[x;\sigma]$ such that $\deg(F_{i,j}) = n_1 + n_2 + \cdots + n_\ell - 1$, $F_{i,j,a_i}(\beta_{i,j}) = 1$, and $F_{i,j,a_u}(\beta_{u,v}) = 0$, for all $v \in [n_i]$ and $u \in [\ell]$ with $u \neq i$ or $v \neq j$.*

Proof. Up to reordering, we may assume that $i = \ell$ and $j = n_\ell$. With notation as in Theorem 2, let $G = G_{\ell,n_\ell-1}$ if $n_\ell > 1$, or $G = G_{\ell-1,n_{\ell-1}}$ if $n_\ell = 1$. By Lemma 6, since $G(a_\ell^{\beta_{\ell,n_\ell}}) \in S^*$ and $\beta_{\ell,n_\ell} \in S^*$, then $G_{a_\ell}(\beta_{\ell,n_\ell}) \in S^*$. Hence, we are done by defining $F_{\ell,n_\ell} = G_{a_\ell}(\beta_{\ell,n_\ell})^{-1} G$. The uniqueness follows again from Theorem 1. $\square$

We may also obtain the following strengthening of Corollary 1 on monic annihilator skew polynomials. It is a generalization of [9, Prop. 3.15].

Corollary 3. *Let $a_1, a_2, \ldots, a_\ell \in S$ be such that $a_i - a_j^\beta \in S^*$, for all $\beta \in S^*$ and for $1 \leq i < j \leq \ell$. Let $\mathbf{u}_i \in S^{n_i}$ and let $t_i = \mathrm{rk}(\mathbf{u}_i)$, for $i \in [\ell]$. Set $t = t_1 + t_2 + \cdots + t_\ell$. Then there exists a monic skew polynomial $F \in S[x;\sigma]$ such that $\deg(F) = t$ and $F_{a_i}(u_{i,j}) = 0$, for $j \in [n_i]$ and for $i \in [\ell]$.*

Proof. Using the Smith normal form (see Section III), we see that there are $\boldsymbol{\alpha}_i \in S^{t_i}$ and $\mathbf{B}_i \in R^{t_i \times n_i}$ such that $\mathbf{u}_i = \boldsymbol{\alpha}_i \mathbf{B}_i$, $\mathrm{frk}(\boldsymbol{\alpha}_i) = t_i$ and $\mathrm{rk}(\mathbf{B}_i) = t_i$, for $i \in [\ell]$. In particular, $(\mathbf{a}, \boldsymbol{\alpha})$ satisfies the MSRD property (Definition 7), where $\mathbf{a} = (a_1, a_2, \ldots, a_\ell)$ and $\boldsymbol{\alpha} = (\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \ldots, \boldsymbol{\alpha}_\ell)$. By Corollary 1, there exists a monic skew polynomial $F \in S[x;\sigma]$ such that

$\deg(F) = t$ and $F_{a_i}(\alpha_{i,j}) = 0$, for $j \in [t_i]$ and for $i \in [\ell]$. Since the map F_{a_i} is R-linear and $\mathbf{u}_i = \boldsymbol{\alpha}_i \mathbf{B}_i$, we deduce that $F_{a_i}(u_{i,j}) = 0$, for $j \in [n_i]$ and for $i \in [\ell]$, and we are done. $\square$

Next we define extended Moore matrices for the ring extension $R \subseteq S$. Such matrices are a trivial adaptation of the matrices from [22, p. 604] from division rings to finite chain rings. These matrices will be used to define linearized Reed–Solomon codes and to explore further forms of Lagrange interpolation.

Definition 8. Consider vectors $\mathbf{a} = (a_1, a_2, \ldots, a_\ell) \in S^\ell$ and $\boldsymbol{\beta}_i = (\beta_{i,1}, \beta_{i,2}, \ldots, \beta_{i,n_i}) \in S^{n_i}$, for $i \in [\ell]$. Set $\boldsymbol{\beta} = (\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \ldots, \boldsymbol{\beta}_\ell)$ and $n = n_1 + n_2 + \cdots + n_\ell$. For $k \in [n]$, we define the *extended Moore matrix* $\mathbf{M}_k(\mathbf{a}, \boldsymbol{\beta}) =$

$$\left(\begin{array}{ccc|c|ccc} \beta_{1,1} & \ldots & \beta_{1,n_1} & \ldots & \beta_{\ell,1} & \ldots & \beta_{\ell,n_\ell} \\ \mathcal{D}_{a_1}(\beta_{1,1}) & \ldots & \mathcal{D}_{a_1}(\beta_{1,n_1}) & \ldots & \mathcal{D}_{a_\ell}(\beta_{\ell,1}) & \ldots & \mathcal{D}_{a_\ell}(\beta_{\ell,n_\ell}) \\ \mathcal{D}^2_{a_1}(\beta_{1,1}) & \ldots & \mathcal{D}^2_{a_1}(\beta_{1,n_1}) & \ldots & \mathcal{D}^2_{a_\ell}(\beta_{\ell,1}) & \ldots & \mathcal{D}^2_{a_\ell}(\beta_{\ell,n_\ell}) \\ \vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \vdots \\ \mathcal{D}^{k-1}_{a_1}(\beta_{1,1}) & \ldots & \mathcal{D}^{k-1}_{a_1}(\beta_{1,n_1}) & \ldots & \mathcal{D}^{k-1}_{a_\ell}(\beta_{\ell,1}) & \ldots & \mathcal{D}^{k-1}_{a_\ell}(\beta_{\ell,n_\ell}) \end{array} \right).$$

When there is confusion about σ, we will write $\mathbf{M}^\sigma_k(\mathbf{a}, \boldsymbol{\beta})$ instead of $\mathbf{M}_k(\mathbf{a}, \boldsymbol{\beta})$.

Example 4. Let the setting be as in Example 1. Let $\mathbf{a} = (a_1, a_2) = (1, \alpha + 1)$ and $\boldsymbol{\beta} = (\beta_1, \beta_2) = (1, \alpha)$. In Example 3, we saw that $(\mathbf{a}, \boldsymbol{\beta}) = ((1, \alpha+1), (1, \alpha))$ satisfies the MSRD property. Notice that $\ell = m = n_1 = n_2 = 2$. If we set $k = 3$, then the corresponding extended Moore matrix is

$$\mathbf{M}_3(\mathbf{a}, \boldsymbol{\beta}) = \left(\begin{array}{cc|cc} 1 & \alpha & 1 & \alpha \\ 1 & -\alpha & \alpha + 1 & 8\alpha + 1 \\ 1 & \alpha & 2 & 2\alpha \end{array} \right).$$

The following result gives a sufficient condition for extended Moore matrices over finite chain rings to be invertible, and it may be of interest on its own.

Theorem 3. *Let $(\mathbf{a}, \boldsymbol{\beta})$ be as in Definition 7, and satisfying the MSRD property. Let $n = n_1 + n_2 + \cdots + n_\ell$. Then the square extended Moore matrix $\mathbf{M}_n(\mathbf{a}, \boldsymbol{\beta})$ is invertible.*

Proof. Let $F_{i,j} \in S[x; \sigma]$ be as in Corollary 2, for $j \in [n_i]$ and $i \in [\ell]$. Then, for the appropriate ordering, the coefficients of such skew polynomials (they are of degree $n - 1$) form the rows of the inverse of $\mathbf{M}_n(\mathbf{a}, \boldsymbol{\beta})$. $\square$

From Theorem 3, we may obtain the following Lagrange interpolation theorem, which we will use later for decoding and may be of interest on its own.

Theorem 4. *Let $(\mathbf{a}, \boldsymbol{\beta})$ be as in Definition 7, and satisfying the MSRD property. Let $c_{i,j} \in S$, for $j \in [n_i]$ and $i \in [\ell]$. Then there exists a unique skew polynomial $F \in S[x; \sigma]$ such that $\deg(F) \leq n_1 + n_2 + \cdots + n_\ell - 1$, and $F_{a_i}(\beta_{i,j}) = c_{i,j}$, for $j \in [n_i]$ and $i \in [\ell]$.*

V. Linearized Reed–Solomon Codes

In this section, we extend the definition of linearized Reed–Solomon codes [22] to finite chain rings, thus providing a first explicit construction of MSRD codes over finite chain rings (that are not fields).

Definition 9. Let $(\mathbf{a}, \boldsymbol{\beta})$ be as in Definition 7, and satisfying the MSRD property. For $k \in [n]$, we define the k-dimensional linearized Reed–Solomon code as the linear code $\mathcal{C}_k(\mathbf{a}, \boldsymbol{\beta}) \subseteq S^n$ with generator matrix $\mathbf{M}_k(\mathbf{a}, \boldsymbol{\beta})$ as in Definition 8. When there is confusion about σ, we will write $\mathcal{C}^\sigma_k(\mathbf{a}, \boldsymbol{\beta})$ instead of $\mathcal{C}_k(\mathbf{a}, \boldsymbol{\beta})$.

This definition coincides with [22, Def. 31] when R and S are fields. It coincides with Gabidulin codes over finite chain rings [9, Def. 3.22] when $\ell = 1$ and generalized Reed–Solomon codes over finite chain rings [35, Def. 22] when $m = n_1 = n_2 = \ldots = n_\ell = 1$.

The main result of this section is the following.

Theorem 5. *Let $(\mathbf{a}, \boldsymbol{\beta})$ be as in Definition 7, and satisfying the MSRD property. For $k \in [n]$, the linearized Reed–Solomon code $\mathcal{C}_k(\mathbf{a}, \boldsymbol{\beta}) \subseteq S^n$ is a free S-module of rank k and an MSRD code over R for the length partition $n = n_1 + n_2 + \cdots + n_\ell$.*

Proof. Let $\mathbf{A}_i \in R^{n_i \times n_i}$ be invertible, for $i \in [\ell]$. By the R-linearity of σ, we have that

$$\begin{aligned} &\mathcal{C}_k(\mathbf{a}, \boldsymbol{\beta}) \operatorname{Diag}(\mathbf{A}_1, \mathbf{A}_2, \ldots, \mathbf{A}_\ell) \\ &= \mathcal{C}_k(\mathbf{a}, \boldsymbol{\beta} \operatorname{Diag}(\mathbf{A}_1, \mathbf{A}_2, \ldots, \mathbf{A}_\ell)), \end{aligned}$$

which is also a linearized Reed–Solomon code, since $(\mathbf{a}, \boldsymbol{\beta} \operatorname{Diag}(\mathbf{A}_1, \mathbf{A}_2, \ldots, \mathbf{A}_\ell))$ also satisfies the MSRD property since $\mathbf{A}_1, \mathbf{A}_2, \ldots, \mathbf{A}_\ell$ are invertible. Therefore, from Lemma 3, we see that we only need to prove that $\mathcal{C}_k(\mathbf{a}, \boldsymbol{\beta})$ is MDS and a free S-module of rank k. Both properties follow from the fact that any $k \times k$ square submatrix of $\mathbf{M}_k(\mathbf{a}, \boldsymbol{\beta})$ is invertible by Theorem 3. $\square$

This result coincides with [22, Th. 4] when R and S are fields, with [9, Th. 3.24] over finite chain rings when $\ell = 1$, and with [35, Prop. 23 & Cor. 24] over finite chain rings when $m = n_1 = n_2 = \ldots = n_\ell = 1$.

Next, we show how to explicitly construct sequences $(\mathbf{a}, \boldsymbol{\beta})$ satisfying the MSRD property. In this way, we have explicitly constructed linearized Reed–Solomon codes for the finite chain ring extension $R \subseteq S$.

The R-linearly independent elements $\beta_{i,1}, \beta_{i,2}, \ldots, \beta_{i,n_i} \in S^*$ can be chosen as subsets of any basis of S over R, for $i \in [\ell]$. The more delicate part is choosing the elements $a_1, a_2, \ldots, a_\ell \in S$. We now show two ways to do this. The proof of the following proposition is straightforward.

Proposition 2. *Let $\ell \in [q-1]$ and let $\gamma \in \mathbb{F}^*_{q^m}$ be a primitive element, that is, $\mathbb{F}^*_{q^m} = \{\gamma^0, \gamma^1, \ldots, \gamma^{q^m-2}\}$. Such an element always exists [18, Th. 2.8]. Take elements $a_1, a_2, \ldots, a_\ell \in S^*$ such that $\overline{a}_i = \gamma^{i-1}$, for $i \in [\ell]$. Then $a_1, a_2, \ldots, a_\ell \in S^*$ are such that $a_i - a_j^\beta \in S^*$, for all $\beta \in S^*$ and all $1 \leq i < j \leq \ell$.*

Another possibility is to choose elements from R^* when $q - 1$ and m are coprime.

Proposition 3. *Assume that $q-1$ and m are coprime and let $\ell \in [q-1]$. Given $a_1, a_2, \ldots, a_\ell \in R^*$, it holds that $a_i - a_j \in R^*$ for all $1 \leq i < j \leq \ell$ if, and only if, $\overline{a}_1, \overline{a}_2, \ldots, \overline{a}_\ell \in \mathbb{F}^*_q$ are all distinct. Moreover, if that is the case, then $a_i - a_j^\beta \in S^*$, for all $\beta \in S^*$ and all $1 \leq i < j \leq \ell$.*

Proof. The first part is trivial, since $R^* = R \setminus \mathfrak{m}$ and $\mathfrak{m} = \ker(\rho)$. Now, since $q-1$ and m are coprime, it follows from [23, Lemma 26] that $\overline{a}_1, \overline{a}_2, \ldots, \overline{a}_\ell \in \mathbb{F}_q^*$ are pair-wise non-conjugate. Therefore, $a_i - a_j^\beta \in S^*$, for all $\beta \in S^*$ and all $1 \leq i < j \leq \ell$. □

Observe that in the previous two propositions, the maximum length of the vector $(a_1, a_2, \ldots, a_\ell)$ is $\ell = q-1$. In the next proposition, we show that this is indeed the maximum possible.

Proposition 4. *Let $(a_1, a_2, \ldots, a_\ell) \in (S^*)^\ell$ be such that $a_i - a_j^\beta \in S^*$, for all $\beta \in S^*$ and all $1 \leq i < j \leq \ell$. Then $\ell \leq q-1$.*

Proof. By the hypothesis on $a_1, a_2, \ldots, a_\ell$, we deduce that $\overline{a}_1, \overline{a}_2, \ldots, \overline{a}_\ell \in \mathbb{F}_{q^m}^*$ are pair-wise non-conjugate. Now, as shown in [19] (see also [22, Prop. 45]), there are at most $q-1$ non-zero conjugacy classes in $\mathbb{F}_{q^m}$ with respect to $\overline{\sigma}$, that is, $\ell \leq q-1$. □

In particular, we have shown the existence of linear MSRD codes of any rank for the extension $R \subseteq S$ as detailed in the following corollary.

Corollary 4. *Let $\ell \in [q-1]$, $n_i \in [m]$ for $i \in [\ell]$, and let $k \in [n]$, where $n = n_1 + n_2 + \cdots + n_\ell$. Then there exists a linear code $\mathcal{C} \subseteq S^n$ that is a free S-module of rank k and is MSRD over R for the length partition $n = n_1 + n_2 + \cdots + n_\ell$.*

For linearized Reed–Solomon codes, notice that these are the same parameter restrictions as in the finite-field case [22, Sec. 4.2].

We observer that, in the case of finite fields and square matrices ($m = n_1 = n_2 = \ldots = n_\ell$), we have the upper bound $\ell \leq q + \left\lfloor \frac{d-3}{n} \right\rfloor$ [5, Th. VI.12]. This bound might hold also for finite chain rings, but we leave it as an open problem. Furthermore, in the Hamming-metric case ($m = n_1 = \ldots = n_\ell = 1$) it is conjectured that $\ell \leq q+1$ in general. Hence being able to attain the number of blocks $\ell = q-1$ is close to the known upper bounds on ℓ for the case of finite fields. In the non-square case ($m > n_i$), one may construct MSRD codes with an unrestricted number of blocks, see [24, Subsec. 4.5].

Finally, we show that duals of linearized Reed–Solomon codes are again linearized Reed–Solomon codes. For a linear code $\mathcal{C} \subseteq S^n$, we define its dual as $\mathcal{C}^\perp = \{\mathbf{d} \in S^n \mid \mathbf{c} \cdot \mathbf{d} = 0\}$, where $\cdot$ denotes the usual Euclidean inner product in S^n. The following lemma follows from [8, Th. 3.1].

Lemma 7 ([8]). *Given a linear code $\mathcal{C} \subseteq S^n$, we have that $\mathcal{C}^\perp$ is a free module if and only if, so is $\mathcal{C}$. In such a case, if $\mathcal{C}$ is of rank k, then $\mathcal{C}^\perp$ is of rank $n-k$. Furthermore, $\mathcal{C}^{\perp\perp} = \mathcal{C}$.*

Using this lemma, we may prove that the dual of a linearized Reed–Solomon code is again a linearized Reed–Solomon code in the same way as in [25, Th. 4].

Theorem 6. *Let $(\mathbf{a}, \boldsymbol{\beta})$ be as in Definition 7, and satisfying the MSRD property. There exists a vector $\boldsymbol{\delta} = \left(\boldsymbol{\delta}^{(1)}, \boldsymbol{\delta}^{(2)}, \ldots, \boldsymbol{\delta}^{(\ell)}\right) \in S^n$, where $\boldsymbol{\delta}^{(i)} = \left(\delta_1^{(i)}, \delta_2^{(i)}, \ldots, \delta_{n_i}^{(i)}\right) \in S^{n_i}$ and $\delta_1^{(i)}, \delta_2^{(i)}, \ldots, \delta_{n_i}^{(i)}$ are R-linearly independent, for $i \in [\ell]$, and such that*

$$\mathcal{C}_k^\sigma(\mathbf{a}, \boldsymbol{\beta})^\perp = \mathcal{C}_{n-k}^{\sigma^{-1}}\left(\sigma^{-1}(\mathbf{a}), \boldsymbol{\delta}\right), \tag{4}$$

for $k \in [n-1]$, where $\sigma^{-1}(\mathbf{a}) = (\sigma^{-1}(a_1), \sigma^{-1}(a_2), \ldots, \sigma^{-1}(a_\ell))$. Notice that $(\sigma^{-1}(\mathbf{a}), \boldsymbol{\delta})$ also satisfies the MSRD property.

Furthermore, if $q-1$ and m are coprime, and $a_1, a_2, \ldots, a_\ell \in R^*$ are such that $a_i - a_j \in R^*$ for all $1 \leq i < j \leq \ell$ (see Proposition 3), then

$$\mathcal{C}_k^\sigma(\mathbf{a}, \boldsymbol{\beta})^\perp = \mathcal{C}_{n-k}^{\sigma^{-1}}(\mathbf{a}, \boldsymbol{\delta}). \tag{5}$$

We will use the form of the dual of a linearized Reed–Solomon code shown in (5) to describe a quadratic-time decoding algorithm in Section VII.

VI. A Welch-Berlekamp Decoder

In this section, we present a Welch-Berlekamp sum-rank error-correcting algorithm for the linearized Reed–Solomon codes from Definition 9. The decoder is based on the original one by Welch and Berlekamp [2]. Welch-Berlekamp decoders for the sum-rank metric in the case of fields were given in [1], [3], [25], listed in decreasing order of computational complexity. Our decoder has cubic complexity over the ring S and is analogous to the works listed above. In Section VII, we will present a decoder with quadratic complexity, but which only works if $q-1$ and m are coprime. The decoder in this section works for all cases.

Throughout this section, we fix $(\mathbf{a}, \boldsymbol{\beta})$ as in Definition 7, and satisfying the MSRD property. Let

$$b_{i,j} = a_i^{\beta_{i,j}},$$

for $j \in [n_i]$ and for $i \in [\ell]$. Next fix a dimension $k \in [n-1]$, and consider the linearized Reed–Solomon code $\mathcal{C}_k(\mathbf{a}, \boldsymbol{\beta}) \subseteq S^n$ (Definition 9). The number of sum-rank errors that it can correct is

$$t = \left\lfloor \frac{\mathrm{d}_{SR}\left(\mathcal{C}_k(\mathbf{a}, \boldsymbol{\beta})\right) - 1}{2} \right\rfloor = \left\lfloor \frac{n-k}{2} \right\rfloor. \tag{6}$$

Let $\mathbf{c} \in \mathcal{C}_k(\mathbf{a}, \boldsymbol{\beta})$ be any *codeword*, let $\mathbf{e} \in S^n$ be an *error vector* such that $\mathrm{wt}_{SR}(\mathbf{e}) \leq t$, and define the *received word* as

$$\mathbf{r} = \mathbf{c} + \mathbf{e} \in S^n. \tag{7}$$

Since $\mathrm{wt}_{SR}(\mathbf{e}) \leq t$ and $2t+1 \leq \mathrm{d}_{SR}\left(\mathcal{C}_k(\mathbf{a}, \boldsymbol{\beta})\right)$, there is a unique solution $\mathbf{c} \in \mathcal{C}_k(\mathbf{a}, \boldsymbol{\beta})$ to the decoding problem.

We start by defining the auxiliary vectors

$$\begin{aligned} \mathbf{c}' &= \mathbf{c} \cdot \mathrm{Diag}(\boldsymbol{\beta})^{-1}, \\ \mathbf{e}' &= \mathbf{e} \cdot \mathrm{Diag}(\boldsymbol{\beta})^{-1}, \text{ and} \\ \mathbf{r}' &= \mathbf{r} \cdot \mathrm{Diag}(\boldsymbol{\beta})^{-1}. \end{aligned} \tag{8}$$

By Lagrange interpolation (Theorem 4) and Lemma 6, there exist unique skew polynomials $F, G, R \in S[x;\sigma]$, all of degree less than n, such that

$$F(\mathbf{b}) = \mathbf{c}', \quad G(\mathbf{b}) = \mathbf{e}', \quad \text{and} \quad R(\mathbf{b}) = \mathbf{r}', \tag{9}$$

which denote component-wise remainder evaluation (Definition 6). Following the original idea of the Welch–Berlekamp decoding algorithm, we want to find a non-zero monic skew polynomial $L \in S[x;\sigma]$ with $\deg(L) \leq t$ and such that

$$(LR)(\mathbf{b}) = (LF)(\mathbf{b}). \tag{10}$$

However, since we do not know F, we look instead for non-zero $L, Q \in S[x;\sigma]$ such that L is monic, $\deg(L) \leq t$, $\deg(Q) \leq t+k-1$ and

$$(LR)(\mathbf{b}) = Q(\mathbf{b}). \tag{11}$$

In the following two lemmas, we show that (10) and (11) can be solved, and once L and Q are obtained, F may be obtained in quadratic time (by Euclidean division).

Lemma 8. *There exists a non-zero monic skew polynomial $L \in S[x;\sigma]$ with $\deg(L) \leq t$ satisfying (10). In particular, there exist non-zero $L, Q \in S[x;\sigma]$ such that L is monic, $\deg(L) \leq t$, $\deg(Q) \leq t+k-1$ and (11) holds.*

Proof. By Corollary 3, there exists a non-zero monic skew polynomial $L \in S[x;\sigma]$ such that $\deg(L) \leq t$ and $L_{a_i}(e_{i,j}) = 0$, for $j \in [n_i]$ and for $i \in [\ell]$. From the definitions and Lemma 6, it follows that

$$(LG)(b_{i,j}) = L_{b_{i,j}}(G(b_{i,j})) = L_{b_{i,j}}(e'_{i,j}) = L_{a_i}(e_{i,j}) = 0,$$

for $j \in [n_i]$ and for $i \in [\ell]$. Since $R(\mathbf{b}) = F(\mathbf{b}) + G(\mathbf{b})$, we conclude that

$$(L(R-F))(\mathbf{b}) = (LG)(\mathbf{b}) = 0$$

by Lemma 5. In other words, L satisfies (10) and we are done. $\square$

Lemma 9. *If $L, Q \in S[x;\sigma]$ are such that L is monic, $\deg(L) \leq t$, $\deg(Q) \leq t+k-1$ and (11) holds, then*

$$Q = LF.$$

Proof. First, by (11) and the product rule (Lemma 5),

$$\text{if} \quad (F-R)(b_{i,j}) = 0, \quad \text{then} \quad (LF-Q)(b_{i,j}) = 0,$$

for $j \in [n_i]$ and for $i \in [\ell]$. From this fact, and using Lemmas 2 and 6, the reader may deduce that

$$\begin{aligned} &\mathrm{wt}_{SR}\left((LF-Q)(\mathbf{b}) \cdot \mathrm{Diag}(\boldsymbol{\beta})\right) \\ &\leq \mathrm{wt}_{SR}\left((F-R)(\mathbf{b}) \cdot \mathrm{Diag}(\boldsymbol{\beta})\right) \leq t. \end{aligned}$$

Therefore, we may apply Lemma 8 to LF and Q, instead of F and R. Thus there exists a non-zero monic $L_0 \in S[x;\sigma]$ such that $\deg(L_0) \leq t$ and

$$(L_0(LF-Q))(\mathbf{b}) = \mathbf{0}.$$

Now observe that

$$\deg\left(L_0(LF-Q)\right) \leq 2t+k-1 < n.$$

By Lemma 6 and Theorem 4, we conclude that

$$L_0(LF-Q) = 0.$$

Since L_0 is non-zero and monic, we conclude that $LF = Q$ and we are done. $\square$

Finally, once we find non-zero skew polynomials $L, Q \in S[x;\sigma]$ such that L is monic, $\deg(L) \leq t$, $\deg(Q) \leq t+k-1$ and (11) holds, then we may find F by left Euclidean division, since $Q = LF$ by Lemma 9 above. Observe that left Euclidean division is possible in $S[x;\sigma]$ since σ is invertible. Finding L and Q using R and $\mathbf{b}$ (which are known) amounts to solving a system of linear equations derived from (11) using the Smith normal form, as in the Gabidulin case, see [9, Sec. III-D]. Using this method, the decoding algorithm has an overall complexity of $\mathcal{O}(n^3)$ operations over the ring S.

VII. A Quadratic Syndrome Decoder

In this section, we extend the syndrome decoder from [33] to linearized Reed–Solomon codes when $q-1$ and m are coprime. This decoder also constitutes the first known syndrome decoder for linearized Reed–Solomon codes over finite fields, to the best of our knowledge. Note that the algorithm [33, Alg. 2], and the skew polynomial version ([33, Alg. 1]) of the Byrne-Fitzpatrick algorithm [4] it is based upon, are given in those works for Galois rings, a particular case of finite chain rings (and not all finite chain rings are Galois rings, see [28, Th. XVII.5]). However, we notice that such algorithms work for finite chain rings in general. For such a generalization, we need the following observation. For the finite chain ring S, there exists an element $\pi \in S$ such that the maximal ideal of S is $\mathfrak{M} = (\pi)$, and all ideals of S are of the form $\mathfrak{M}^i = (\pi^i)$, for $i \in [r]$, where r is the smallest positive integer such that $\pi^r = 0$, and thus $\mathfrak{M}^r = 0$, see [9, Sec. II-B]. With this representation of the ideal chain of S, one can extend mutatis mutandis [33, Alg. 1] and the proof of its correctness and complexity to general finite chain rings. For the convenience of the reader, we include [33, Alg. 1] for a finite chain ring S in Algorithm 1. Here, we also denote $\mathrm{lt}(F) = x^{\deg(F)}$, for $F \in S[x;\sigma]$, and $\prec$ is any total order in the set $\{(x^n, 0) \mid n \in \mathbb{N}\} \cup \{(0, x^n) \mid n \in \mathbb{N}\}$ compatible with multiplication by x^k, for all $k \in \mathbb{N}$. For left Gröbner bases, see [33, Sec. III]. Finally, mod denotes modulo on the right, that is, we say that $F \equiv G \bmod H$ if, and only if, H divides $F - G$ on the right.

Throughout this section, we fix a pair $(\mathbf{a}, \boldsymbol{\beta})$ as in Definition 7. We will assume that $a_1, a_2, \ldots, a_\ell \in R^*$ satisfy $a_i - a_j \in R^*$ for all $1 \leq i < j \leq \ell$ (i.e., $\overline{a}_1, \overline{a}_2, \ldots, \overline{a}_\ell \in \mathbb{F}_q^*$ are all distinct) and that $\beta_{i,1}, \beta_{i,2}, \ldots, \beta_{i,n_i} \in S$ are R-linearly independent, for $i \in [\ell]$. Hence $(\mathbf{a}, \boldsymbol{\beta})$ satisfies the MSRD property by Proposition 3 since we are assuming that $q-1$ and m are coprime. In particular, fixing a dimension $k \in [n-1]$ and a linearized Reed–Solomon code $\mathcal{C}_k^\sigma(\mathbf{a}, \boldsymbol{\beta}) \subseteq S^n$, we have that $\mathcal{C}_k^\sigma(\mathbf{a}, \boldsymbol{\beta})^\perp = \mathcal{C}_{n-k}^{\sigma^{-1}}(\mathbf{a}, \boldsymbol{\delta})$ by Theorem 6, where $(\mathbf{a}, \boldsymbol{\delta})$ also satisfies the MSRD property.

We consider the same error-correcting scenario as in Section VI. That is, $t = \lfloor (n-k)/2 \rfloor$ as in (6) and $\mathbf{r} = \mathbf{c} + \mathbf{e} \in S^n$ as in (7), for a fixed codeword $\mathbf{c} \in \mathcal{C}_k^\sigma(\mathbf{a}, \boldsymbol{\beta})$ and an error vector $\mathbf{e} \in S^n$, where we may assume that $t = \mathrm{wt}_{SR}(\mathbf{e})$.

We start by extending [33, Def. 2]. Notice that $a_i^{-1} - (a_j^{-1})^\beta \in S^*$, for all $\beta \in S^*$, with respect to σ^{-1}, since the same property holds for a_i and a_j with respect to σ by assumption, for all $1 \leq i < j \leq \ell$.

Algorithm 1: SkewByrneFitzpatrick [33, Alg. 1]

Input : $U \in S[x;\sigma]$ and $m \in \mathbb{Z}_{>0}$.
Output: Left Gröbner basis of the left $S[x;\sigma]$-module

$$\mathcal{M} := \left\{(F,G) \in S[x;\sigma]^2 \mid FU \equiv G \bmod x^m\right\}.$$

let $\mathcal{B}_0 := \left\{(\pi^i, 0) \mid i \in \{0,1,\ldots,r-1\}\right\} \cup \left\{(0,\pi^i) \mid i \in \{0,1,\ldots,r-1\}\right\}$
for $k \in \{0,1,\ldots,m-1\}$ **do**
  for *each* $(F_i,G_i) \in \mathcal{B}_k$ **do**
    compute the discrepancy $\zeta_i := (F_iU - G_i)_k$ (where $(\cdot)_k$ denotes the kth coefficient)
  for *each* $(F_i,G_i) \in \mathcal{B}_k$ **do**
    if $\zeta_i = 0$ **then**
      put $(F_i,G_i) \in \mathcal{B}_{k+1}$
      continue
    if *there is* $(F_j,G_j) \in \mathcal{B}_k$ *with* $\mathrm{lt}(F_j,G_j) \prec \mathrm{lt}(F_i,G_i)$ *and* ζ_j *divides* ζ_i **then**
      put $(F_i,G_i) - Q(F_j,G_j)$ in $\mathcal{B}_{k+1}$, where $Q \in S$ with $\zeta_i = Q\zeta_j$
    else
      put (xF_i, xG_i) in $\mathcal{B}_{k+1}$
return $\mathcal{B}_m$

Definition 10. We say that $F = \sum_{i=0}^d F_i x^i \in S[x;\sigma^{-1}]$, $d \in \mathbb{N}$, is primitive if it is not a zero divisor, i.e., $F_i \in S^*$ for some $i \in [d]$, i.e., $\overline{F} \in \mathbb{F}_{q^m}[x;\overline{\sigma}^{-1}] \setminus \{0\}$. We say that $\Lambda \in S[x;\sigma^{-1}]$ is an annihilator of $\mathbf{e} \in S^n$ if it is primitive, $\Lambda_{a_i^{-1}}(e_{i,j}) = 0$ for $j \in [n_i]$ and $i \in [\ell]$ and it has minimum possible degree among primitive skew polynomials in $S[x;\sigma^{-1}]$ satisfying such a property.

Notice that, here, $\Lambda_{a_i^{-1}}(e_{i,j})$ is the operator evaluation (Definition 4) with respect to σ^{-1}. We will not specify this in the notation $\Lambda_{a_i^{-1}}(e_{i,j})$ since we wrote that $\Lambda \in S[x;\sigma^{-1}]$, hence emphasizing the use of σ^{-1} for Λ instead of σ.

We need some preliminary auxiliary properties on the zeros of skew polynomials over finite chain rings. This result extends Lemma 4 in a different direction than Theorem 1.

Lemma 10. *If $F \in S[x;\sigma]$ is primitive and $c_1, c_2, \ldots, c_\ell \in S^*$ are such that $c_i - c_j^\beta \in S^*$ for all $\beta \in S^*$ and all $1 \le i < j \le \ell$, then*

$$\sum_{i=1}^{\ell} \mathrm{rk}_R(\ker(F_{c_i})) \le \deg(F).$$

Proof. Let $r_i = \ker(F_{c_i})$, for $i \in [\ell]$. Using the Smith normal form, we see that there are R-linearly independent elements $b_{i,1}, \ldots, b_{i,r_i} \in S$ and non-zero $\lambda_{i,1}, \ldots, \lambda_{i,r_i} \in R$ such that $\ker(F_{c_i}) = \langle \lambda_{i,1}b_{i,1}, \ldots, \lambda_{i,r_i}b_{i,r_i}\rangle_R$, for $i \in [\ell]$. Since R is a chain ring, we may assume that there exists $k \in [\ell]$ such that $\lambda_{i,j} | \lambda_{k,r_k}$ for all $j \in [r_i]$ and all $i \in [\ell]$. Therefore, since $F_{c_i}(\lambda_{i,j}b_{i,j}) = \lambda_{i,j}F_{c_i}(b_{i,j}) = 0$, we see that $(\lambda_{k,r_k}F)_{c_i}(b_{i,j}) = 0$, for all $j \in [r_i]$ and all $i \in [\ell]$.

Assume that $\deg(F) < r_1 + r_2 + \cdots + r_\ell$. Then we deduce that $\lambda_{k,r_k}F = 0$ by Theorem 1. However, since $\lambda_{k,r_k} \neq 0$, then F is not primitive, a contradiction. Therefore, $\deg(F) \ge r_1 + r_2 + \cdots + r_\ell$ and we are done. $\square$

We next extend [33, Lemma 4].

Lemma 11. *Any annihilator of $\mathbf{e} \in S^n$ has degree $t = \mathrm{wt}_{SR}(\mathbf{e})$. In addition, if $\mathrm{rk}(\mathbf{e}_i) = \mathrm{frk}(\mathbf{e}_i)$, for $i \in [\ell]$, then there is a unique monic annihilator of $\mathbf{e}$.*

Proof. Let $\Lambda \in S[x;\sigma^{-1}]$ be an annihilator of $\mathbf{e}$. First, $\deg(\Lambda) \le t$ by Corollary 3. Second, if $\deg(\Lambda) < t$, then Λ would not be primitive by Lemma 10. Hence $\deg(\Lambda) = t$.

Now assume that $\mathrm{rk}(\mathbf{e}_i) = \mathrm{frk}(\mathbf{e}_i)$, for $i \in [\ell]$. First, there exists a monic annihilator $\Lambda \in S[x;\sigma^{-1}]$ of $\mathbf{e}$ by Corollary 3. Let $\Lambda' \in S[x;\sigma^{-1}]$ be another annihilator of $\mathbf{e}$. Note that $t = \deg(\Lambda) = \deg(\Lambda')$. Since Λ is monic, we may perform right Euclidean division, i.e., there are $Q, R \in S[x;\sigma^{-1}]$ with $\deg(R) < t$ and $\Lambda' = Q\Lambda + R$. By Lemmas 5 and 6, we have that $R_{a_i^{-1}}(e_{i,j}) = 0$, for $j \in [n_i]$ and $i \in [\ell]$. Since $\deg(R) < \sum_{i=1}^{\ell} \mathrm{frk}_R(\mathbf{e}_i)$, we deduce that $R = 0$ by Theorem 4. In other words, $\Lambda' = Q\Lambda$, where $Q \in S^*$, and thus Λ is the unique monic annihilator of $\mathbf{e}$. $\square$

We will define syndromes as usual.

Definition 11. Let $h = n - k$ and define the syndrome vector $\mathbf{s} = \mathbf{e}\mathbf{M}_h^{\sigma^{-1}}(\mathbf{a}, \boldsymbol{\delta})^\intercal \in S^h$. We define the syndrome skew polynomial $s = \sum_{i=0}^{h-1} s_i x^i \in S[x;\sigma^{-1}]$, where $\mathbf{s} = (s_0, s_1, \ldots, s_{h-1})$.

In order to prove the key equation between annihilators of $\mathbf{e}$ and the syndrome skew polynomial s, we need the following two lemmas. The first one follows directly from the Smith normal form.

Lemma 12. *For $i \in [\ell]$, there exist $\boldsymbol{\alpha}_i \in S^{t_i}$ and $\mathbf{B}^{(i)} \in R^{t_i \times n_i}$ such that $\mathbf{e}_i = \boldsymbol{\alpha}_i\mathbf{B}^{(i)}$, $t_i = \mathrm{frk}(\boldsymbol{\alpha}_i) = \mathrm{rk}(\mathbf{B}^{(i)})$ and $\langle \alpha_{i,1}, \ldots, \alpha_{i,t_i}\rangle_R = \langle e_{i,1}, \ldots, e_{i,n_i}\rangle_R$. In particular, we have that $\Lambda_{a_i^{-1}}(\alpha_{i,j}) = 0$, for all $j \in [t_i]$ and $i \in [\ell]$, for any annihilator $\Lambda \in S[x;\sigma]$ of $\mathbf{e}$.*

The second lemma can be found in [12, Lemma 1].

Lemma 13 ([12]). *For all $a \in S$ and all integers $0 \le j \le i$, it holds that*

$$N_i^\sigma(a) = \sigma^{i-j}\left(N_j^\sigma(a)\right) N_{i-j}^\sigma(a).$$

We may now provide the key equation.

Theorem 7 (Key Equation). *Let $\Lambda \in S[x;\sigma^{-1}]$ be an annihilator of $\mathbf{e}$. There exists $\Omega \in S[x;\sigma^{-1}]$ with $\deg(\Omega) < t$ and*

$$\Omega \equiv \Lambda s \mod x^h. \tag{12}$$

Proof. If we set $d_{i,u} = \sum_{j=1}^{n_i} \mathbf{B}_{u,j}^{(i)}\delta_{i,j} \in S$, for $u \in [t_i]$ and

$i \in [\ell]$, then we have that

$$\begin{aligned} s_v &= \sum_{i=1}^{\ell}\sum_{j=1}^{n_i} e_{i,j}\mathcal{D}^v_{\sigma^{-1},a_i}(\delta_{i,j}) \\ &= \sum_{i=1}^{\ell}\sum_{j=1}^{n_i}\sum_{u=1}^{t_i} \alpha_{i,u}\mathbf{B}^{(i)}_{u,j}\mathcal{D}^v_{\sigma^{-1},a_i}(\delta_{i,j}) \\ &= \sum_{i=1}^{\ell}\sum_{u=1}^{t_i}\sum_{j=1}^{n_i} \alpha_{i,u}\mathcal{D}^v_{\sigma^{-1},a_i}\left(\mathbf{B}^{(i)}_{u,j}\delta_{i,j}\right) \\ &= \sum_{i=1}^{\ell}\sum_{u=1}^{t_i} \alpha_{i,u}\mathcal{D}^v_{\sigma^{-1},a_i}(d_{i,u}), \end{aligned} \tag{13}$$

for $v = 0, 1, \ldots, h-1$.

Note that $\deg(\Lambda) = t$ and $\deg(s) = h-1$, and therefore, $\Lambda s = \sum_{v=0}^{t+h-1}(\Lambda s)_v$. For $v = t, t+1, \ldots, h-1$, we have that

$$\begin{aligned} (\Lambda s)_v &= \sum_{l=0}^{v}\Lambda_{v-l}\sigma^{-v+l}(s_l) \\ &\overset{(a)}{=} \sum_{i=1}^{\ell}\sum_{u=1}^{t_i}\sum_{l=0}^{v}\Lambda_{v-l}\sigma^{-v+l}\left(\alpha_{i,u}\mathcal{D}^l_{\sigma^{-1},a_i}(d_{i,u})\right) \\ &= \sum_{i=1}^{\ell}\sum_{u=1}^{t_i}\sum_{l=0}^{v}\Lambda_{v-l}\sigma^{-v+l}(\alpha_{i,u})\,\sigma^{-v+l}\left(N_l^{\sigma^{-1}}(a_i)\right)\sigma^{-v+l}\left(\sigma^{-l}(d_{i,u})\right) \\ &\overset{(b)}{=} \sum_{i=1}^{\ell}\sum_{u=1}^{t_i}\sum_{l=0}^{v}\Lambda_{v-l}\sigma^{-v+l}(\alpha_{i,u})\,N_{v-l}^{\sigma^{-1}}\left(a_i^{-1}\right)N_v^{\sigma^{-1}}(a_i)\sigma^{-v}(d_{i,u}) \\ &= \sum_{i=1}^{\ell}\sum_{u=1}^{t_i}\sum_{l=0}^{v}\Lambda_{v-l}\mathcal{D}^{v-l}_{\sigma^{-1},a_i^{-1}}(\alpha_{i,u})\,\mathcal{D}^v_{\sigma^{-1},a_i}(d_{i,u}) \\ &= \sum_{i=1}^{\ell}\sum_{u=1}^{t_i}\mathcal{D}^v_{\sigma^{-1},a_i}(d_{i,u})\left(\sum_{l=0}^{v}\Lambda_l\mathcal{D}^l_{\sigma^{-1},a_i^{-1}}(\alpha_{i,u})\right) \\ &\overset{(c)}{=} \sum_{i=1}^{\ell}\sum_{u=1}^{t_i}\mathcal{D}^v_{\sigma^{-1},a_i}(d_{i,u})\left(\sum_{l=0}^{t}\Lambda_l\mathcal{D}^l_{\sigma^{-1},a_i^{-1}}(\alpha_{i,u})\right) \\ &\overset{(d)}{=} 0, \end{aligned}$$

where we have used the formula (13) for s_l in (a) since $l \leq v \leq h-1$, Lemma 13 in (b), the fact that $\Lambda_l = 0$ if $t < l \leq v$ in (c), and Lemma 12 in (d). □

The next ingredient is the following extension of [1, Th. 7]. From now on, we will also define

$$\widetilde{\beta}_{i,j} = \sigma^{k-1}(\beta_{i,j})a_i^{k-1},$$

for $j \in [n_i]$ and $i \in [\ell]$. Note that $\widetilde{\beta}_{i,1}, \ldots, \widetilde{\beta}_{i,n_i} \in S^*$ are also R-linearly independent, since $a_i \in R^*$, σ is an automorphism and $\beta_{i,1}, \ldots, \beta_{i,n_i} \in S^*$ are R-linearly independent.

Theorem 8. *Recall that $\mathbf{r} = \mathbf{c} + \mathbf{e}$, where $\mathbf{c} \in \mathcal{C}_k^\sigma(\mathbf{a}, \boldsymbol{\beta})$ and $\mathbf{e} \in S^n$ is such that $t = \mathrm{wt}_{SR}(\mathbf{e})$. Assume that we have non-zero $U, V \in S[x;\sigma^{-1}]$ such that*

1) *U is primitive,*
2) *$Us - V \equiv 0 \bmod x^h$,*
3) *$\deg(U) \leq t$,*
4) *$\deg(V) < \deg(U)$.*

Then U is an annihilator of $\mathbf{e}$ and in particular, $\deg(U) = t$. Moreover,

$$UR \equiv U\widetilde{F} \mod G,$$

where $R, \widetilde{F}, G \in S[x;\sigma^{-1}]$ are the unique skew polynomials with

$$\begin{array}{llll} R_{a_i^{-1}}(\widetilde{\beta}_{i,j}) & = r_{i,j} & \textit{and} & \deg(R) < n, \\ \widetilde{F}_{a_i^{-1}}(\widetilde{\beta}_{i,j}) & = c_{i,j} & \textit{and} & \deg(\widetilde{F}) < k, \\ G_{a_i^{-1}}(\widetilde{\beta}_{i,j}) & = 0 & \textit{and} & \deg(G) = n, \end{array}$$

and G is the unique monic annihilator of $\widetilde{\boldsymbol{\beta}} \in S^n$.

Proof. Since $\deg(V) < t$, $2t-1 < h$ and $Us - V \equiv 0 \bmod x^h$, then $(Us)_i = 0$, for $i = t, t+1, \ldots, 2t-1$. This may be rewritten as

$$\begin{pmatrix} \sigma^0(s_t) & \sigma^{-1}(s_{t-1}) & \ldots & \sigma^{-t}(s_0) \\ \sigma^0(s_{t+1}) & \sigma^{-1}(s_t) & \ldots & \sigma^{-t}(s_1) \\ \vdots & \vdots & \ddots & \vdots \\ \sigma^0(s_{2t-1}) & \sigma^{-1}(s_{2t-2}) & \ldots & \sigma^{-t}(s_{t-1}) \end{pmatrix}\begin{pmatrix} u_0 \\ u_1 \\ \vdots \\ u_t \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix},$$

where $U = \sum_{i=0}^{t} u_i x^i$. If we denote by $\mathbf{S} \in S^{t\times(t+1)}$ the matrix above, then by (13) we have a decomposition

$$\mathbf{S} = \mathbf{D}\mathbf{A},$$

where

$$\mathbf{D} = (\mathbf{D}_1|\ldots|\mathbf{D}_\ell) \in S^{t\times t}, \quad \mathbf{A} = \begin{pmatrix} \mathbf{A}_1 \\ \hline \vdots \\ \hline \mathbf{A}_\ell \end{pmatrix} \in S^{t\times(t+1)},$$

and

$$\mathbf{D}_i = \begin{pmatrix} \mathcal{D}^t_{\sigma^{-1},a_i}(d_{i,1}) & \mathcal{D}^t_{\sigma^{-1},a_i}(d_{i,2}) & \ldots & \mathcal{D}^t_{\sigma^{-1},a_i}(d_{i,t_i}) \\ \mathcal{D}^{t+1}_{\sigma^{-1},a_i}(d_{i,1}) & \mathcal{D}^{t+1}_{\sigma^{-1},a_i}(d_{i,2}) & \ldots & \mathcal{D}^{t+1}_{\sigma^{-1},a_i}(d_{i,t_i}) \\ \vdots & \vdots & \ddots & \vdots \\ \mathcal{D}^{2t-1}_{\sigma^{-1},a_i}(d_{i,1}) & \mathcal{D}^{2t-1}_{\sigma^{-1},a_i}(d_{i,2}) & \ldots & \mathcal{D}^{2t-1}_{\sigma^{-1},a_i}(d_{i,t_i}) \end{pmatrix} \in S^{t\times t_i},$$

$$\mathbf{A}_i = \begin{pmatrix} \mathcal{D}^0_{\sigma^{-1},a_i^{-1}}(\alpha_{i,1}) & \mathcal{D}^1_{\sigma^{-1},a_i^{-1}}(\alpha_{i,1}) & \ldots & \mathcal{D}^t_{\sigma^{-1},a_i^{-1}}(\alpha_{i,1}) \\ \mathcal{D}^0_{\sigma^{-1},a_i^{-1}}(\alpha_{i,2}) & \mathcal{D}^1_{\sigma^{-1},a_i^{-1}}(\alpha_{i,2}) & \ldots & \mathcal{D}^t_{\sigma^{-1},a_i^{-1}}(\alpha_{i,2}) \\ \vdots & \vdots & \ddots & \vdots \\ \mathcal{D}^0_{\sigma^{-1},a_i^{-1}}(\alpha_{i,t_i}) & \mathcal{D}^1_{\sigma^{-1},a_i^{-1}}(\alpha_{i,t_i}) & \ldots & \mathcal{D}^t_{\sigma^{-1},a_i^{-1}}(\alpha_{i,t_i}) \end{pmatrix} \in S^{t_i\times(t+1)},$$

for $i \in [\ell]$. Since $d_{i,1}, \ldots d_{i,t_i} \in S$ are R-linearly independent, we deduce from Theorem 3 that $\mathbf{D} \in S^{t\times t}$ is invertible. Thus we have that

$$\mathbf{S}\mathbf{u} = \mathbf{D}\mathbf{A}\mathbf{u} = \mathbf{0} \iff \mathbf{A}\mathbf{u} = \mathbf{0},$$

which means that $U_{a_i^{-1}}(\alpha_{i,j}) = 0$, for all $j \in [t_i]$ and $i \in [\ell]$. Since $\langle \alpha_{i,1}, \ldots, \alpha_{i,t_i}\rangle_R = \langle e_{i,1}, \ldots, e_{i,n_i}\rangle_R$, then $U_{a_i^{-1}}(e_{i,j}) = 0$, for $j \in [n_i]$ and $i \in [\ell]$. Since $\deg(U) \leq t$ and it is primitive, then U is an annihilator of $\mathbf{e}$. Finally, we have

$$0 = U_{a_i^{-1}}(e_{i,j}) = U_{a_i^{-1}}(r_{i,j} - c_{i,j}) =$$

$$U_{a_i^{-1}}(R_{a_i^{-1}}(\widetilde{\beta}_{i,j}) - \widetilde{F}_{a_i^{-1}}(\widetilde{\beta}_{i,j})) = (U(R-\widetilde{F}))_{a_i^{-1}}(\widetilde{\beta}_{i,j}),$$

for $j \in [n_i]$ and $i \in [\ell]$, using Lemmas 5 and 6 in the last equality. Since G is a monic annihilator of $\widetilde{\boldsymbol{\beta}}$, then we deduce that G divides $U(R-\widetilde{F})$ on the right, and we are done. □

Finally, we show that we may recover the skew polynomial associated to $\mathbf{c}$ for the pair $(\mathbf{a}, \boldsymbol{\beta})$ and σ from that for the pair $(\mathbf{a}^{-1}, \widetilde{\boldsymbol{\beta}})$ and σ^{-1}.

Proposition 5. *Let* $F = \sum_{u=0}^{k-1} F_u x^u \in S[x;\sigma]$ *and* $\widetilde{F} = \sum_{u=0}^{k-1} \widetilde{F}_u x^u \in S[x;\sigma^{-1}]$ *be related by* $\widetilde{F}_{k-u-1} = F_u$*, for* $u = 0, 1, \ldots, k-1$*. Then*

$$F_{a_i}(\beta_{i,j}) = \widetilde{F}_{a_i^{-1}}(\widetilde{\beta}_{i,j}),$$

where $\widetilde{\beta}_{i,j} = \sigma^{k-1}(\beta_{i,j}) a_i^{k-1}$*, for* $j \in [n_i]$ *and* $i \in [\ell]$*.*

Proof. Since $a_i \in R^*$ and $\widetilde{F}_{a_i^{-1}}$ is R-linear, then for $j \in [n_i]$ and $i \in [\ell]$, we have

$$\begin{aligned}
\widetilde{F}_{a_i^{-1}}(\widetilde{\beta}_{i,j}) &= \widetilde{F}_{a_i^{-1}}(a^{k-1}\sigma^{k-1}(\beta_{i,j})) \\
&= a^{k-1}\widetilde{F}_{a_i^{-1}}(\sigma^{k-1}(\beta_{i,j})) \\
&= a^{k-1}\sum_{u=0}^{k-1} \widetilde{F}_u \sigma^{-u}(\sigma^{k-1}(\beta_{i,j})) a_i^{-u} \\
&= \sum_{u=0}^{k-1} F_{k-u-1}\sigma^{k-u-1}(\beta_{i,j}) a_i^{k-u-1} \\
&= F_{a_i}(\beta_{i,j}).
\end{aligned}$$

□

The algorithm in [1, alg. 2] can be extended to our case as shown in Algorithm 2. In the following theorem, we prove its correctness and give its complexity.

Algorithm 2: SyndromeDecoder

```
Input  : r ∈ S^n
Output: If there is a
        c = (F_{a_1}(β_{1,1}), ..., F_{a_ℓ}(β_{ℓ,n_ℓ})) ∈ C_k^σ(a, β)
        with F ∈ S[x; σ], deg(F) < k and
        d_SR(r, c) ≤ (n−k)/2, then F.
        Otherwise "decoding failure".
s := H r^T
s := Σ_{i=0}^{n−k−1} s_i x^i ∈ S[x; σ^{−1}]
B := SkewByrneFitzpatrick(s, n − k)
(Λ, Ω) := element of B of minimal degree among all
    (U, V) ∈ B with deg(U) > deg(V) and U primitive.
R := unique R ∈ S[x; σ^{−1}] such that
    R_{a_i^{−1}}(β̃_{i,j}) = r_{i,j}, for all i, j, with deg(R) < n.
G := unique G ∈ S[x; σ^{−1}] such that G_{a_i^{−1}}(β̃_{i,j}) = 0,
    for all i, j, with deg(G) < n.
Ψ := ΛR mod G
(F̃, T) := quotient and remainder of left division of Ψ
    by Λ.
F := skew polynomial obtained by F̃_{k−u−1} = F_u, for
    u = 0, 1, ..., k − 1.
if T = 0 and
    d_SR(r, (F_{a_1}(β_{1,1}), ..., F_{a_ℓ}(β_{ℓ,n_ℓ}))) ≤ (n−k)/2 and
    deg(F) < k then
    return F
else
    return "decoding failure"
```

Theorem 9. *Algorithm 2 is correct and has a complexity of* $\mathcal{O}(rn^2)$ *operations in* S*, where* r *is the smallest positive integer such that* $\pi^r = 0$ *or* $\mathfrak{M}^r = 0$*.*

Proof. Using Algorithm 1, we obtain a pair $U, V \in S[x;\sigma^{-1}]$ with U primitive, $\deg(U) > \deg(V)$, $Us \equiv V \bmod x^h$ and $\deg(U)$ minimal among pairs with these properties. By Theorem 7, there is a pair (Λ, Ω) satisfying such properties and with $\deg(\Lambda) = t$. Thus we have $\deg(U) \leq t$.

Since $t \leq h/2$, then U is an annihilator of $\mathbf{e}$ by Theorem 8. Moreover, we have $UR \equiv U\widetilde{F} \bmod G$, with notation as in Theorem 8.

Since $\deg(U\widetilde{F}) = \deg(U) + \deg(\widetilde{F}) < t + k - 1 < n = \deg(G)$, then we may obtain $U\widetilde{F}$ by right division of UR by G. Note that R and G may be computed from the received word $\mathbf{r}$ and the pair $(\mathbf{a}, \boldsymbol{\beta})$, and that G is monic, hence right division by G is well defined. Next, since U is primitive, then we may divide $U\widetilde{F} \neq 0$ by U on the left and we obtain $\widetilde{F}$. Finally, we compute F from $\widetilde{F}$ as in Proposition 5, where F is the skew polynomial whose coefficients contain the message encoded by the sent codeword $\mathbf{c}$, and we are done.

Finally, the complexity of the skew polynomial Byrne-Fitzpatrick Algorithm 1 has a complexity of $\mathcal{O}(rn^2)$ operations in S by [1, Th. 3] (the extension from Galois rings to general finite chain rings is straightforward as mentioned at the beginning of the section). The other operations that appear in Algorithm 2 can be implemented with a complexity of $\mathcal{O}(n^2)$ operations in S by [1, Lemma 8]. □

VIII. Applications

In this section, we briefly discuss applications of MSRD codes over finite chain rings, and in particular, the linearized Reed–Solomon codes from Definition 9. We will only focus on applications in Space-Time Coding and Multishot Network Coding, and we will only briefly discuss how to adapt ideas from the literature to the case of MSRD codes over finite chain rings.

A. Space-Time Coding with Multiple Fading Blocks

Space-time codes [41] are used in wireless communication, in scenarios of multiple input/multiple output antenna transmission. Such codes utilize space diversity (via multiple antennas) and time diversity (via interleaving up to some delay constraint) in order to reduce the fading of the channel.

In the case of one fading block, codewords are seen as matrices in $\mathcal{A}^{n_t \times T}$, where $\mathcal{A} \subseteq \mathbb{C}$ is the signal constellation (a subset of the complex field), n_t is the number of transmit antennas and T is the time delay. In particular, the code is a subset $\mathcal{C} \subseteq \mathcal{A}^{n_t \times T}$. In this scenario, the code achieves transmit diversity gain d (or simply code diversity) if the rank of the difference of any two matrices in the code is at least d, see [20], [41]. Large code diversity is desirable, but it competes with the symbol rate of the code, defined as

$$\frac{1}{T}\log_{|\mathcal{A}|}|\mathcal{C}|.$$

The symbol rate is an important parameter when the constellation $\mathcal{A}$ is constrained or we wish it to be as small as possible. See the discussion in [37]. The diversity-rate tradeoff is expressed in a Singleton-type bound, and codes attaining equality in such a bound may be obtained by mapping a maximum rank distance (MRD) code over a finite field, such as a Gabidulin code, into the constellation $\mathcal{A} \subseteq \mathbb{C}$. This may be done via Gaussian integers [21] or Eisenstein integers [34].

The case of multiple fading blocks, say L, was first investigated in [20]. In this case, the codewords are matrices in $\mathcal{A}^{n_t \times LT}$, which can be thought of as L matrices of size $n_t \times T$, that is tuples in $(\mathcal{A}^{n_t \times T})^L$. In this case, a code diversity d is attained if the minimum sum-rank distance of the code is at least d. For this reason, space-time codes with optimal rate-diversity tradeoff in the multiblock case may be obtained by mapping MSRD codes over finite fields to the constellation $\mathcal{A} \subseteq \mathbb{C}$. This was observed in [20], and linearized Reed–Solomon codes were first used for this purpose in [38]. As shown there, the use of linearized Reed–Solomon codes allows one to attain optimal rate-diversity while minimizing the time delay T, and while the constellation size $|\mathcal{A}|$ grows linearly in L, in contrast with previous space-time codes, whose constellation sizes grow exponentially in L. See also [37].

In [9, Sec. VI-A], it was shown how to translate any MRD code over a finite principal ideal ring into a space-time code over a complex constellation with optimal rate-diversity tradeoff. This result can be immediately generalized to the multiblock case by using MSRD codes over finite principal ideal rings, as follows. We omit the proof.

Theorem 10. *Let Q be a principal ideal ring such that there exists a surjective ring morphism $\varphi : Q \longrightarrow R$ (recall that R is a finite chain ring). Let $\varphi^* : R \longrightarrow Q$ be such that $\varphi \circ \varphi^* = \mathrm{Id}$. Extend both maps component-wise to tuples of matrices in $(Q^{n_t \times T})^L$ and $(R^{n_t \times T})^L$. If $\mathcal{C} \subseteq (R^{n_t \times T})^L$ is an MSRD code, then so is $\varphi^*(\mathcal{C}) \subseteq (Q^{n_t \times T})^L$, of the same dimension and minimum sum-rank distance. In particular, if $Q \subseteq \mathbb{C}$, then $\varphi^*(\mathcal{C})$ is a space-time code with optimal rate-diversity tradeoff for L fading blocks.*

Examples may be constructed easily. For instance, we may choose $Q = \mathbb{Z}[i]$ and $R = \mathbb{Z}_{2^r}[i] = \mathbb{Z}[i]/(2^r)$, where r is a positive integer (see also the introduction). In this case, we have that $q = 2^r$, and we may construct linearized Reed–Solomon codes corresponding to $L = 2^r - 1$ fading blocks.

B. Physical-Layer Multishot Network Coding

Linear network coding [17] permits maximum information flow from a source to several sinks (*multicast*) in one use of the network (a *single shot*). In such a communication scenario, MRD codes can correct a given number of link errors and packet erasures with the maximum possible information rate, without knowledge and independently of the transfer matrix or topology of the network (*universal error correction*), see [11], [39]. In the case of a number ℓ of uses of the network (*multishot Network Coding*), the minimum sum-rank distance of the code determines how many link errors and packet erasures the code can correct in total throughout the ℓ shots of the network, without knowledge and independently of the transfer matrices and network topology [31]. MSRD codes, in particular linearized Reed–Solomon codes, over finite fields were used in this scenario in [25].

In [6], a similar model was developed for physical-layer Network Coding, where the network code lies in some constellation (a subset of the complex numbers), which may be identified with finite principal ideal rings of the form $\mathbb{Z}[i]/(q)$, for a positive integer q. Such rings are finite chain rings if $q = 2^r$ and r is a positive integer, see the introduction.

Just as in the case of finite fields, we may consider ℓ shots in physical-layer Network Coding, with network-code alphabet $R = \mathbb{Z}[i]/(q)$, which is a finite chain ring if $q = 2^r$ as above. In this case, codes are subsets $\mathcal{C} \subseteq R^{m \times n}$, where m is the packet length and n is the number of outgoing links from the source. Using the matrix representation map (2), we may consider $\mathcal{C} \subseteq S^n$, and we may then choose $\mathcal{C}$ to be S-linear, such as a linearized Reed–Solomon code (Definition 9), which may attain the value $\ell = q-1 = 2^r-1$. If $\mathbf{c} \in \mathcal{C}$ is transmitted, then the output of an ℓ-shot linearly coded network over R with at most t link errors and ρ packet erasures is

$$\mathbf{y} = \mathbf{c}\mathbf{A} + \mathbf{e} \in S^N,$$

where $\mathbf{e} \in S^N$ is such that $\mathrm{wt}_{SR}(\mathbf{e}) \leq t$, and $\mathbf{A} = \mathrm{Diag}(\mathbf{A}_1, \mathbf{A}_2, \ldots, \mathbf{A}_\ell)$ is such that $\mathrm{frk}(\mathbf{A}) \geq n - \rho$, where $\mathbf{A}_i \in R^{n_i \times N_i}$ is the transfer matrix of the ith shot, for $i \in [\ell]$, where $N = N_1 + N_2 + \cdots N_\ell$ and $n = n_1 + n_2 + \cdots + n_\ell$. For simplicity, we will only consider the *coherent* scenario, meaning that we assume that the transfer matrix $\mathbf{A}$ is known to the receiver.

We may now prove that the minimum sum-rank distance of the code $\mathcal{C}$ gives a sufficient condition for error and erasure correction in the scenario described above. We omit necessary conditions for brevity.

Theorem 11. *In the scenario described above, fix $\rho = n - \mathrm{frk}(\mathbf{A})$. If*

$$2t + \rho + 1 \leq \mathrm{d}_{SR}(\mathcal{C}),$$

then there exists a decoder $D_{\mathbf{A}} : S^n \longrightarrow \mathcal{C}$ sending $D_{\mathbf{A}}(\mathbf{c}\mathbf{A} + \mathbf{e}) = \mathbf{c}$, for all $\mathbf{c} \in \mathcal{C}$ and all $\mathbf{e} \in S^n$ with $\mathrm{wt}_{SR}(\mathbf{e}) \leq t$. In particular, if $\mathcal{C}$ is an MSRD code over R for the length partition $n = n_1 + n_2 + \cdots + n_\ell$, then it may correct t link errors, ρ packet erasures and achieves an information rate of

$$\frac{n - 2t - \rho}{n}.$$

Proof. Let $\mathbf{C} \in R^{m \times n}$, $\mathbf{B} \in R^{n \times N}$ and $r = \mathrm{frk}(\mathbf{B})$. We may assume that the first r columns of $\mathbf{B}$, which form a matrix $\mathbf{B}' \in R^{n \times r}$, are R-linearly independent. By [28, p. 92, ex. V.14], there exists a matrix $\mathbf{B}'' \in R^{n \times (n-r)}$ such that $\mathbf{B}_2 = (\mathbf{B}', \mathbf{B}'') \in R^{n \times n}$ is invertible. Therefore,

$$\mathrm{rk}(\mathbf{C}\mathbf{B}) \geq \mathrm{rk}(\mathbf{C}\mathbf{B}') \geq \mathrm{rk}(\mathbf{C}\mathbf{B}_2) - (n-r) = \mathrm{rk}(\mathbf{C}) - (n-r).$$

In the second inequality we have used that $\mathrm{rk}(\mathbf{C}\mathbf{B}_2) \leq \mathrm{rk}(\mathbf{C}\mathbf{B}') + \mathrm{rk}(\mathbf{C}\mathbf{B}'') \leq \mathrm{rk}(\mathbf{C}\mathbf{B}') + (n-r)$. Hence, we deduce that

$$\mathrm{d}_{SR}(\mathcal{C}\mathbf{A}) \geq \mathrm{d}_{SR}(\mathcal{C}) - \rho \geq 2t + 1.$$

Therefore, there exists a decoder $D_1 : S^N \longrightarrow \mathcal{C}\mathbf{A}$ sending $D_1(\mathbf{c}\mathbf{A} + \mathbf{e}) = \mathbf{c}\mathbf{A}$, for all $\mathbf{c} \in \mathcal{C}$ and all $\mathbf{e} \in S^n$ with $\mathrm{wt}_{SR}(\mathbf{e}) \leq t$.

Now, let again $\mathbf{C} \in R^{m \times n}$, $\mathbf{B} \in R^{n \times N}$ and $r = \mathrm{frk}(\mathbf{B})$, with notation as above. Assume that $\mathbf{C}\mathbf{B} = 0$. Then $\mathbf{C}\mathbf{B}' = 0$, which implies that $\mathbf{CB}_2 = \mathbf{C}(\mathbf{B}', \mathbf{B}'') = (0, \mathbf{CB}'')$. Therefore,

$$\mathrm{rk}(\mathbf{C}) = \mathrm{rk}(\mathbf{CB}_2) = \mathrm{rk}(\mathbf{CB}'') \leq n - r.$$

Using this fact and $\rho \leq \mathrm{d}_{SR}(\mathcal{C})$, it is easy to see that the map $\mathcal{C} \longrightarrow \mathcal{C}\mathbf{A}$ consisting in multiplying by $\mathbf{A}$ is injective. Hence we deduce that there exists a decoder $D_2 : \mathcal{C}\mathbf{A} \longrightarrow \mathcal{C}$ sending $D_2(\mathbf{cA}) = \mathbf{c}$, for all $\mathbf{c} \in \mathcal{C}$. Thus, we conclude by defining $D_{\mathbf{A}} = D_2 \circ D_1$. ☐

To conclude, we briefly describe how to decode when using a linearized Reed–Solomon code as in Definition 9. Similarly to the proof above, we let $\mathbf{A}' \in R^{n \times r}$ be of full free rank r, formed by some r columns of $\mathbf{A}$, where $r = \mathrm{frk}(\mathbf{A}) = n - \rho$. In the same way as in [25, Sec. V-F], if $\mathcal{C}$ is a linearized Reed–Solomon code, then so is $\mathcal{C}\mathbf{A}'$, since $\mathbf{A}'$ has full free rank. Therefore, we may apply the decoders from Sections VI and VII to $\mathcal{C}\mathbf{A}'$ and recover the sent codeword $\mathbf{c} \in \mathcal{C}$.

References

[1] H. Bartz, T. Jerkovits, S. Puchinger, and J. Rosenkilde, "Fast decoding of codes in the rank, subspace, and sum-rank metric," *IEEE Trans. Info. Theory*, vol. 67, no. 8, pp. 5026–5050, 2021.

[2] E. R. Berlekamp and L. Welch, "Error correction of algebraic block codes," 1986, U.S. Patent No. 4,633,470.

[3] D. Boucher, "An algorithm for decoding skew Reed–Solomon codes with respect to the skew metric," *Des., Codes, Crypto.*, vol. 88, pp. 1991–2005, 2020.

[4] E. Byrne and P. Fitzpatrick, "Hamming metric decoding of alternant codes over Galois rings," *IEEE Trans. Info. Theory*, vol. 48, no. 3, pp. 683–694, 2002.

[5] E. Byrne, H. Gluesing-Luerssen, and A. Ravagnani, "Fundamental properties of sum-rank-metric codes," *IEEE Trans. Info. Theory*, vol. 67, no. 10, pp. 6456–6475, 2021.

[6] C. Feng, D. Silva, and F. R. Kschischang, "An algebraic approach to physical-layer network coding," *IEEE Trans. Info. Theory*, vol. 59, no. 11, pp. 7576–7596, 2013.

[7] E. M. Gabidulin, "Theory of codes with maximum rank distance," *Problems Inform. Transmission*, vol. 21, no. 1, pp. 1–12, 1985.

[8] T. Honold and I. Landjev, "Linear codes over finite chain rings," *The Electronic Journal of Combinatorics*, vol. 7, pp. R11–R11, 2000.

[9] H. T. Kamche and C. Mouaha, "Rank-metric codes over finite principal ideal rings and applications," *IEEE Trans. Info. Theory*, vol. 65, no. 12, pp. 7718–7735, 2019.

[10] T. Kiran and B. Rajan, "Optimal STBCs from codes over Galois rings," in *IEEE International Conference on Personal Wireless Communications*, 2005, pp. 120–124.

[11] R. Kötter and F. R. Kschischang, "Coding for errors and erasures in random network coding," *IEEE Trans. Info. Theory*, vol. 54, no. 8, pp. 3579–3591, 2008.

[12] T. Y. Lam, "A general theory of Vandermonde matrices," *Expositiones Mathematicae*, vol. 4, pp. 193–215, 1986.

[13] T. Y. Lam and A. Leroy, "Vandermonde and Wronskian matrices over division rings," *J. Algebra*, vol. 119, no. 2, pp. 308–336, 1988.

[14] ——, "Hilbert 90 theorems over divison rings," *Transactions of the American Mathematical Society*, vol. 345, no. 2, pp. 595–622, 1994.

[15] A. Leroy, "Pseudolinear transformations and evaluation in Ore extensions," *Bulletin of the Belgian Mathematical Society*, vol. 2, no. 3, pp. 321–347, 1995.

[16] ——, "Noncommutative polynomial maps," *Journal of Algebra and its Applications*, vol. 11, no. 04, p. 1250076, 2012.

[17] S.-Y. R. Li, R. Yeung, and N. Cai, "Linear network coding," *IEEE Trans. Info. Theory*, vol. 49, no. 2, pp. 371–381, 2003.

[18] R. Lidl and H. Niederreiter, *Finite Fields*, ser. Encyclopedia of Mathematics and its Applications. Amsterdam: Addison-Wesley, 1983, vol. 20.

[19] S. Liu, F. Manganiello, and F. R. Kschischang, "Matroidal structure of skew polynomial rings with application to network coding," *Finite Fields App.*, vol. 46, no. C, pp. 326–346, 2017.

[20] H.-F. Lu and P. V. Kumar, "A unified construction of space-time codes with optimal rate-diversity tradeoff," *IEEE Trans. Info. Theory*, vol. 51, no. 5, pp. 1709–1730, May 2005.

[21] P. Lusina, E. Gabidulin, and M. Bossert, "Maximum rank distance codes as space-time codes," *IEEE Trans. Info. Theory*, vol. 49, no. 10, pp. 2757–2760, 2003.

[22] U. Martínez-Peñas, "Skew and linearized Reed-Solomon codes and maximum sum rank distance codes over any division ring," *J. Algebra*, vol. 504, pp. 587–612, 2018.

[23] ——, "Sum-rank BCH codes and cyclic-skew-cyclic codes," *IEEE Trans. Info. Theory*, vol. 67, no. 8, pp. 5149–5167, 2021.

[24] ——, "A general family of MSRD codes and PMDS codes with smaller field sizes from extended Moore matrices," *SIAM J. Disc. Math.*, vol. 36, no. 3, pp. 1868–1886, 2022.

[25] U. Martínez-Peñas and F. R. Kschischang, "Reliable and secure multishot network coding using linearized Reed–Solomon codes," *IEEE Trans. Info. Theory*, vol. 65, no. 8, pp. 4785–4803, 2019.

[26] ——, "Universal and dynamic locally repairable codes with maximal recoverability via sum-rank codes," *IEEE Trans. Info. Theory*, vol. 65, no. 12, pp. 7790–7805, 2019.

[27] U. Martínez-Peñas and S. Puchinger, "Maximum sum-rank distance codes over finite chain rings," in *Proc. Workshop on Coding and Cryptography (WCC)*, 2022, pp. 1–10.

[28] B. R. McDonald, *Finite rings with identity*. Marcel Dekker Inc., 1974, vol. 28.

[29] A. Neri, "Twisted linearized Reed-Solomon codes: A skew polynomial framework," *Journal of Algebra*, vol. 609, pp. 792–839, 2022.

[30] A. Neri, P. Santonastaso, and F. Zullo, "The geometry of one-weight codes in the sum-rank metric," *J. Combinatorial Theory, S. A*, vol. 194, p. 105703, 2023.

[31] R. W. Nóbrega and B. F. Uchôa-Filho, "Multishot codes for network coding using rank-metric codes," in *Proc. 2010 Third IEEE Int. Workshop on Wireless Network Coding*, 2010, pp. 1–6.

[32] O. Ore, "Theory of non-commutative polynomials," *Annals of Mathematics (2)*, vol. 34, no. 3, pp. 480–508, 1933.

[33] S. Puchinger, J. Renner, A. Wachter-Zeh, and J. Zumbrägel, "Efficient decoding of Gabidulin codes over Galois rings," in *Proc. IEEE Int. Symp. Info. Theory*. IEEE, 2021, pp. 25–30.

[34] S. Puchinger, S. Stern, M. Bossert, and R. Fischer, "Space-time codes based on rank-metric codes and their decoding," in *2016 International Symposium on Wireless Communication Systems (ISWCS)*, 2016, pp. 125–130.

[35] G. Quintin, M. Barbier, and C. Chabot, "On generalized Reed–Solomon codes over commutative and noncommutative rings," *IEEE Trans. Info. Theory*, vol. 59, no. 9, pp. 5882–5897, 2013.

[36] I. S. Reed and G. Solomon, "Polynomial codes over certain finite fields," *J. Soc. Ind. and Appl. Math.*, vol. 8, no. 2, pp. 300–304, 1960.

[37] M. Shehadeh, "Space–time codes from sum-rank codes," Master's thesis, Edward S. Rogers Sr. Dept. Elect. & Comput. Eng., Univ. Toronto, Toronto, ON, Canada, 2020.

[38] M. Shehadeh and F. R. Kschischang, "Space–time codes from sum-rank codes," *IEEE Trans. Info. Theory*, vol. 68, no. 3, pp. 1614–1637, 2021.

[39] D. Silva, F. R. Kschischang, and R. Kötter, "A rank-metric approach to error control in random network coding," *IEEE Trans. Info. Theory*, vol. 54, no. 9, pp. 3951–3967.

[40] H. J. S. Smith, "On systems of linear indeterminate equations and congruences," *Philosophical Transactions of the Royal Society of London*, vol. 151, pp. 293–326, 1861.

[41] V. Tarokh, N. Seshadri, and A. Calderbank, "Space-time codes for high data rate wireless communication: performance criterion and code construction," *IEEE Trans. Info. Theory*, vol. 44, no. 2, pp. 744–765, 1998.