An iterative algorithm for parametrization of shortest length linear shift registers over finite chain rings

M. Kuijper and R. Pinto*†

April 21, 2016

Abstract

The construction of shortest feedback shift registers for a finite sequence $S_1, \ldots, S_N$ is considered over finite chain rings, such as $\mathbb{Z}_{p^r}$. A novel algorithm is presented that yields a parametrization of all shortest feedback shift registers for the sequence of numbers $S_1, \ldots, S_N$, thus solving an open problem in the literature. The algorithm iteratively processes each number, starting with S_1, and constructs at each step a particular type of minimal basis. The construction involves a simple update rule at each step which leads to computational efficiency. It is shown that the algorithm simultaneously computes a similar parametrization for the reverse sequence $S_N, \ldots, S_1$. The complexity order of the algorithm is shown to be $O(rN^2)$.

Keywords: Iterative algorithms, minimal basis, parametrization, polynomial modules, sequences, shift registers

Mathematics Subject Classification (2010): 94A55, 16P10

1 Introduction

Minimal bases have been identified in the literature as ideal tools for various types of minimal interpolation problems. Among the most fundamental of those is the classical problem of constructing shortest feedback shift registers for a given sequence of numbers $S_1, \ldots, S_N$. This problem is motivated by coding applications, such as list decoding over finite rings as well as cryptographic applications, such as complexity analysis of ring sequences [31, 33].

To set the scene with a few simple examples: consider $N = 4$ and the sequence $1, 3, 3, 3$. Over the field $\mathbb{Z}_5$ its unique minimal characteristic polynomial is given by $d(x) = x^2 - x$, meaning that $S_{2+j} - S_{1+j} = 0$ for $j = 1$ and $j = 2$. However, if we view this sequence over the ring $\mathbb{Z}_9$ then the above minimal characteristic polynomial is no longer unique, in fact there are two other minimal characteristic polynomials, namely $x^2 + 2x$ and $x^2 + 5x$. Furthermore, a parametrization of all characteristic polynomials of degree 3 (so non-minimal) is given by $d(x) = (x + a)(x^2 - x) + b(3 - x) + 3cx + 3dx^2$, where $a, b, c, d \in \mathbb{Z}_3$. Over $\mathbb{Z}_5$ a parametrization of all characteristic polynomials of degree 3 is given by $d(x) = (x + a)(x^2 - x) + b(3 - x)$, where $a, b \in \mathbb{Z}_5$.

*M. Kuijper is with the Department of Electrical and Electronic Engineering, University of Melbourne, VIC 3010, Australia mkuijper@unimelb.edu.au; R. Pinto is with CIDMA - Center for Research and Development in Mathematics and Applications, Department of Mathematics, University of Aveiro, Aveiro, Portugal raquel@ua.pt

†Partly supported by the Australian Research Council(ARC); partly supported by Portuguese funds through the Center for Research and Development in Mathematics and Applications (CIDMA), and The Portuguese Foundation for Science and Technology (FCT - Fundação para a Ciencia e a Tecnologia), within project UID/MAT/04106/2013.

In [2, 35] such parametrizations are used for the purpose of list decoding of Reed-Solomon codes over fields. In this paper we focus on the iterative construction of such a parametrization in the general setting of finite chain rings. The recent paper [17, section 4.3] provides a conceptual framework for a noniterative solution based on minimal Gröbner bases. In the field case, there exists an efficient iterative alternative, namely the Berlekamp-Massey algorithm which has complexity order $O(N^2)$. Because of its iterative nature, this algorithm has the additional property that, at each step, it solves the problem for the data processed so far. This is a useful property for sequence related applications, for example for the computation of a sequence complexity profile. In this paper we aim for such an iterative solution for the ring case. Via a computationally simple update rule, at each step k, the algorithm constructs a minimal basis that yields a parametrization of all shortest feedback shift registers for the sequence $S_1, \ldots, S_k$ for $k = 1, \ldots, N$. Thus the construction of the minimal basis is tailored to the problem at hand.

There has been a considerable amount of literature on this subject since 1985, notably [30, 5, 6, 8, 11, 12, 20, 21, 25, 26, 27, 29]. Many different algorithms have been proposed in these works, but to our knowledge no parametrization results were presented. The bottleneck here seems to be the proof which is difficult to generalize from the field case due to the existence of zero divisors. Note that the parametrization problem for the finite chain ring case was explicitly posed as an open problem in the 1999 paper [27]. The proof of our result in this paper requires specific linear algebraic machinery—it relies on the framework developed in [15, 17] for dealing with polynomial vectors in $\mathbb{Z}_{p^r}[x]^q$. We find that the use of minimal Gröbner bases ideas enhances the insightfulness of the proof due to the fact that we can explicitly use properties such as the 'predictable leading monomial property", explained in section 2 below.

For the field case, the idea of an iterative Gröbner based algorithm is already in the 1995 paper [7]. In fact, a closer inspection shows that our algorithm in subsection 4.1 on the field case resembles the algorithm of [7]. However, our formulation differs to the extent that it allows for an interesting interpretation of some of the auxiliary polynomials as shortest feedback shift registers for the reverse sequence $S_N, \ldots, S_1$. Apart from additional insight into the role of the auxiliary polynomials, this connection with the reverse sequence also leads to results on bidirectionality which is relevant (see [32]) for cryptographic applications.

Most importantly, our new formulation for the field case enables our main result in subsection 4.2 which is an extension to sequences over a finite chain ring $\mathcal{R}$. For notational simplicity we restrict ourselves to the finite chain ring $\mathbb{Z}_{p^r}$, where p is a prime integer and r is a positive integer. However, we emphasize that our results are valid for any finite chain ring $\mathcal{R}$: replace r by the number of proper ideals of $\mathcal{R}$; replace p by π, where π is a generator of $\mathcal{R}$'s unique maximal ideal, say $\mathcal{N}$. We have that every element $a \in \mathcal{R}$ can be written uniquely as (π-*adic expansion*)

$$a = \theta_0 + \theta_1\pi + \theta_2\pi^2 + \cdots + \theta_{r-1}\pi^{r-1},$$

where $\theta_\ell \in \mathcal{R}/\mathcal{N}$ (residue field).

Our main result is Algorithm 4.11 which is an iterative algorithm that yields a parametrization of all shortest feedback shift registers for a sequence $S_1, \ldots, S_N$ in $\mathbb{Z}_{p^r}$. The algorithm constructs a particular type of minimal basis at each step. It resembles, but is not the same as the 1985 Reeds-Sloane algorithm from [30] which constructs a shortest feedback shift register for a sequence $S_1, \ldots, S_N$ in $\mathbb{Z}_{p^r}$ as a generalization of the Berlekamp-Massey algorithm. In fact, our Gröbner methodology enables a novel parametrization, thus extending Massey's parametrization result [23] to the ring case.

Further preliminary studies for this paper were conference papers [14] and [19].

2 Preliminaries

Minimal Gröbner bases are recognized as effective tools for minimal realization and interpolation problems, see e.g. [5, 7, 22]. In recent papers [16, 17] this effectiveness was ascribed to a powerful property

of minimal Gröbner bases, explicitly identified as the "Predictable Leading Monomial Property". Before recalling this property let us now recall some terminology and basic results on Gröbner bases.

Let us first present some preliminaries from [17] on polynomial vectors with coefficients in a finite chain ring $\mathcal{R}$.

Let $\mathrm{e}_1, \ldots, \mathrm{e}_q$ denote the unit (row) vectors in $\mathcal{R}^q$. The elements $x^\alpha \mathrm{e}_i$ with $i \in \{1, \ldots, q\}$ and $\alpha \in \mathbb{N}_0$ are called **monomials**. Let us consider two types of orderings on these monomials, see also the textbook [1]:

- The **Term Over Position (`top`)** ordering, defined as

$$x^\alpha \mathrm{e}_i < x^\beta \mathrm{e}_j \quad :\Leftrightarrow \quad \alpha < \beta \text{ or } (\alpha = \beta \text{ and } i < j).$$

- The **Position Over Term (`pot`)** ordering, defined as

$$x^\alpha \mathrm{e}_i < x^\beta \mathrm{e}_j \quad :\Leftrightarrow \quad i < j \text{ or } (i = j \text{ and } \alpha < \beta).$$

Clearly, whatever ordering is chosen, every nonzero element $f \in \mathcal{R}[x]^q$ can be written uniquely as

$$f = \sum_{i=1}^{L} c_i X_i,$$

where $L \in \mathbb{N}$, the c_i's are nonzero elements of $\mathcal{R}$ for $i = 1, \ldots, L$ and $X_1, \ldots, X_L$ are monomials, ordered as $X_1 > \cdots > X_L$. Using the terminology of [1] we define

- $\mathrm{lm}(f) := X_1$ as the **leading monomial** of f
- $\mathrm{lt}(f) := c_1 X_1$ as the **leading term** of f
- $\mathrm{lc}(f) := c_1$ as the **leading coefficient** of f

Writing $X_1 = x^{\alpha_1} \mathrm{e}_{i_1}$, where $\alpha_1 \in \mathbb{N}_0$ and $i_1 \in \{1, \ldots, q\}$, we define

- $\mathrm{lpos}(f) := i_1$ as the **leading position** of f
- $\deg(f) := \alpha_1$ as the **degree** of f.

Below we denote the submodule generated by polynomials f_1, ..., f_n by $\langle f_1, ..., f_n \rangle$. There are several ways to define Gröbner bases, here we adopt the definition of [1] which requires us to first define the concept of "leading term submodule".

Definition 2.1 ([1]) *Let F be a subset of $\mathcal{R}[x]^q$. Then the submodule $\mathrm{L}(F)$, defined as*

$$\mathrm{L}(F) := \langle \mathrm{lt}(f) \mid f \in F \rangle$$

*is called the **leading term submodule** of F.*

Definition 2.2 ([1]) *Let $M \subseteq \mathcal{R}[x]^q$ be a module and $G \subseteq M$. Then G is called a **Gröbner basis** of M if*

$$\mathrm{L}(G) = \mathrm{L}(M).$$

It can be shown that a Gröbner basis of a module M is a generating set for M. In order to define a concept of minimality we have the following definition.

Definition 2.3 ([1, Def. 4.1.1]) *Let $0 \neq f \in \mathcal{R}[x]^q$ and let $F = \{f_1, \ldots, f_s\}$ be a set of nonzero elements of $\mathcal{R}[x]^q$. Let $\alpha_{j_1}, \ldots, \alpha_{j_m} \in \mathbb{N}_0$ and $\beta_{j_1}, \ldots, \beta_{j_m}$ be nonzero distinct elements of $\mathcal{R}$, where $1 \leq j_i \leq s$ for $i = 1, \ldots, m$, such that*

1. $\mathrm{lm}(f) = x^{\alpha_{j_i}} \mathrm{lm}(f_{j_i})$ *for* $i = 1, \ldots, m$ *and*

2. $\mathrm{lt}(f) = \beta_{j_1} x^{\alpha_{j_1}} \mathrm{lt}(f_{j_1}) + \cdots + \beta_{j_m} x^{\alpha_{j_m}} \mathrm{lt}(f_{j_m})$.

Define

$$h := f - (\beta_{j_1} x^{\alpha_{j_1}} f_{j_1} + \cdots + \beta_{j_m} x^{\alpha_{j_m}} f_{j_m}).$$

Then we say that f ***reduces*** *to* h *modulo* F *in one step and we write*

$$f \xrightarrow{F} h.$$

If f *cannot be reduced modulo* F*, we say that* f *is* ***minimal*** *with respect to* F*.*

Lemma 2.1 ([1, Lemma 4.1.3]) *Let* f*,* h *and* F *be as in the above definition. If* $f \xrightarrow{F} h$ *then* $h = 0$ *or* $\mathrm{lm}(h) < \mathrm{lm}(f)$*.*

Definition 2.4 ([1]) *A Gröbner basis* G *is called* ***minimal*** *if all its elements* g *are minimal with respect to* $G \backslash \{g\}$*.*

Elements of a minimal Gröbner basis have the convenient property that all their leading monomials are different from each other. In the case that $\mathcal{R} = \mathbb{F}$ is a field, this implies that all leading positions are different and that there are exactly $\dim(M)$ elements in a minimal Gröbner basis for M. In fact, in the case that $\mathcal{R} = \mathbb{F}$ is a field, minimal Gröbner bases exhibit another powerful property, see the next theorem which merely formulates a well known result.

Theorem 2.2 ([1]) *Let* M *be a submodule of* $\mathbb{F}[x]^q$ *with minimal Gröbner basis* $G = \{g_1, \ldots, g_m\}$*. Then for any* $0 \neq f \in M$*, written as*

$$f = a_1 g_1 + \cdots + a_m g_m, \tag{1}$$

where $a_1, \ldots, a_m \in \mathbb{F}[x]$*, we have*

$$\mathrm{lm}(f) = \max_{1 \leq i \leq m; a_i \neq 0} (\mathrm{lm}(a_i g_i)). \tag{2}$$

Conform [17] we say that $\{g_1, \ldots, g_m\}$ has the **Predictable Leading Monomial (PLM) property**. Note that this property involves not only degree information (as in the 'predictable degree property' first introduced in [9]) but also leading position information. The above theorem holds no matter which monomial ordering is chosen; here we only consider `top` or `pot`, but one could also employ reflected versions of `top` or `pot`, as in [17] or weighted versions of `top` or `pot`, as in [2].

Theorem 2.2 leads to parametrizations of other types of minimal vectors in M, as outlined in a general formulation in the next theorem.

Theorem 2.3 *Let* M *be a submodule of* $\mathbb{F}[x]^q$ *with minimal Gröbner basis* $G = \{g_1, \ldots, g_m\}$*. Let* $\ell \in \{1, \ldots, m\}$ *and let* $\mathcal{P}$ *be a property of* g_ℓ *that is absent in* $\mathrm{span}_{\, i \neq \ell} \{g_i\}$*. Then among all elements in* M *with property* $\mathcal{P}$*,* g_ℓ *has minimal leading monomial. More specifically, a parametrization of all elements with minimal leading monomial and property* $\mathcal{P}$ *is given by:*

$$f = a_\ell g_\ell + \sum_{i \neq \ell} a_i g_i,$$

with $a_\ell \in \mathbb{F}$ *a nonzero constant and for all* $i \neq \ell$ *the polynomials* $a_i \in \mathbb{F}[x]$ *chosen such that* $\mathrm{lm}(a_i g_i) \leq \mathrm{lm}(g_\ell)$*.*

Proof Suppose $f \in M$ has property $\mathcal{P}$ and has minimal leading monomial. Obviously we can write f as a linear combination of $g_1, \ldots, g_m$. Because of the assumptions on G, it follows that this linear combination must use g_ℓ. The parametrization now follows immediately from Theorem 2.2, that is, the PLM property of G. In particular, it follows that $\mathrm{lm}(f) = \mathrm{lm}(g_\ell)$, that is, g_ℓ has minimal leading monomial among all elements in M with property $\mathcal{P}$.

The next two corollaries follow immediately from Theorem 2.3.

Corollary 2.4 *Let M be a submodule of $\mathbb{F}[x]^2$ of dimension 2 with minimal Gröbner basis $G = \{g_1, g_2\}$. Suppose that $\mathrm{lpos}(g_2) = 2$. Then g_2 is the lowest degree vector in M with 2 as leading position. A parametrization of all lowest degree vectors f that have 2 as leading position is given by*

$$f = a_2 g_2 + a_1 g_1,$$

with $a_2 \in \mathbb{F}$ a nonzero constant and the polynomial $a_1 \in \mathbb{F}[x]$ chosen such that $\mathrm{lm}(a_1 g_1) \leq \mathrm{lm}(g_2)$.

Corollary 2.5 *Let M be a submodule of $\mathbb{F}[x]^2$ of dimension 2 with minimal Gröbner basis $G = \{g_1, g_2\}$, where $g_1 = [g_{11} \quad g_{12}]$ and $g_2 = [g_{21} \quad g_{22}]$. Suppose that $g_{12}(0) = 0$ and $g_{22}(0) \neq 0$. Then g_2 is the lowest degree vector in M that satisfies $g_{22}(0) \neq 0$. More specifically, a parametrization of all lowest degree $f = [f_1 \quad f_2]$ in M that satisfy $f_2(0) \neq 0$ is given by*

$$f = a_2 g_2 + a_1 g_1,$$

with $a_2 \in \mathbb{F}$ a nonzero constant and the polynomial $a_1 \in \mathbb{F}[x]$ chosen such that $\mathrm{lm}(a_1 g_1) \leq \mathrm{lm}(g_2)$.

Proof Define $f = [f_1 \quad f_2]$ to have property $\mathcal{P}$ if $f_2(0) \neq 0$. The result then follows immediately from Theorem 2.3.

We also have the following theorem, which merely reformulates the wellknown result of [10] that the maximum degree of the full size minors of a row reduced polynomial matrix equals the sum of its row degrees, see also [2].

Theorem 2.6 *Let M be a module in $\mathbb{F}[x]^q$. Let $G = \{g_1, \ldots, g_m\}$ be a minimal Gröbner basis of M with respect to the* `top` *ordering; denote the corresponding* `top` *degrees by $\ell_i := \deg g_i$ for $i = 1, \ldots, m$. Let $\tilde{G} = \{\tilde{g}_1, \ldots, \tilde{g}_m\}$ be a minimal Gröbner basis of M with respect to the* `pot` *ordering; denote the corresponding* `pot` *degrees by $\tilde{\ell}_i := \deg \tilde{g}_i$ for $i = 1, \ldots, m$. Then*

$$\sum_{i=1}^{m} \ell_i = \sum_{i=1}^{m} \tilde{\ell}_i. \tag{3}$$

Definition 2.5 *Let M and G be as in the above theorem. Then the sum in (3) is called the* **degree** *of M, denoted by* $\deg(M)$.

3 Gröbner bases for modules in $\mathbb{Z}_{p^r}[x]^q$

In this section we turn our attention to the case where $\mathcal{R}$ is a finite chain ring, for notational simplicity assumed to be $\mathbb{Z}_{p^r}$, as explained in Section 1; here r is a positive integer and p is a prime integer. For the sake of completeness we repeat several preliminaries from [15] and [17].

3.1 Preliminaries on $\mathbb{Z}_{p^r}$

A subset of $\mathbb{Z}_{p^r}$ that plays a fundamental role in this section is the subset $\mathcal{A}_p = \{0, 1, \ldots, p-1\}$ which is isomorphic to the ring's residue field $\mathbb{Z}_{p^r}/ < p >$. Recall that any element $a \in \mathbb{Z}_{p^r}$ can be written uniquely as $a = \theta_0 + p\theta_1 + \cdots + p^{r-1}\theta_{r-1}$, where $\theta_\ell \in \mathcal{A}_p$ for $\ell = 0, \ldots, r-1$ (*p-adic expansion*).

Next, adopting terminology from [34], a constant a in $\mathbb{Z}_{p^r}$ is said to have **order** k if the additive subgroup generated by a has p^k elements. Constants of order r are called **units**. Thus the constants $1, p, p^2, \ldots, p^{r-1}$ have orders $r, r-1, r-2, \ldots, 1$, respectively. For any choice of monomial ordering (`top` or `pot`), we extend the above notion of "order" for constants to polynomial vectors as follows.

Definition 3.1 *The **order** of a nonzero polynomial vector $f \in \mathbb{Z}_{p^r}[x]^q$, is defined as the order of the constant* $\mathrm{lc}(f)$*, denoted as* $\mathrm{ord}\,(f)$.

To deal with zero divisors occurring in $\mathbb{Z}_{p^r}[x]^q$, it is useful to use the notions defined in [15] of "p-linear dependence" and "p-generator sequence" (such notions were first introduced for "constant" modules, i.e., modules in $\mathbb{Z}_{p^r}^q$ in [34]).

Definition 3.2 ([15]) *Let $\{v_1, \ldots, v_N\} \subset \mathbb{Z}_{p^r}[x]^q$. A **p-linear combination** of $v_1, \ldots, v_N$ is a vector $\sum_{j=1}^{N} a_j v_j$, where $a_j \in \mathcal{A}_p[x]$ for $j = 1, \ldots, N$. Furthermore, the set of all p-linear combinations of $v_1, \ldots, v_N$ is denoted by **p-span**$\{v_1, \ldots, v_N\}$, whereas the set of all linear combinations of $v_1, \ldots, v_N$ with coefficients in $\mathbb{Z}_{p^r}[x]$ is denoted by* $\mathrm{span}\,\{v_1, \ldots, v_N\}$.

Definition 3.3 ([15]) *An ordered sequence $(v_1, \ldots, v_N)$ of vectors in $\mathbb{Z}_{p^r}[x]^q$ is said to be a **p-generator sequence** if $p\,v_N = 0$ and $p\,v_i$ is a p-linear combination of $v_{i+1}, \ldots, v_N$ for $i = 1, \ldots, N-1$.*

Theorem 3.1 ([15]) *Let $v_1, \ldots, v_N \in \mathbb{Z}_{p^r}[x]^q$. If $(v_1, \ldots, v_N)$ is a p-generator sequence then*

$$p\text{−span}\,\{v_1, \ldots, v_N\} = \mathrm{span}\,\{v_1, \ldots, v_N\}.$$

In particular, p−span $\{v_1, \ldots, v_N\}$ is a submodule of $\mathbb{Z}_{p^r}[x]^q$.

All submodules of $\mathbb{Z}_{p^r}[x]^q$ can be written as the p-span of a p-generator sequence. In fact, if $M = \mathrm{span}\,\{g_1, \ldots, g_m\}$ then M is the p-span of the p-generator sequence

$$(g_1, pg_1, \ldots, p^{r-1}g_1, g_2, pg_2, \ldots, p^{r-1}g_2, \ldots, g_m, pg_m, \ldots, p^{r-1}g_m).$$

Definition 3.4 ([15]) *The vectors $v_1, \ldots, v_N \in \mathbb{Z}_{p^r}[x]^q$ are said to be **p-linearly independent** if the only p-linear combination of $v_1, \ldots, v_N$ that equals zero is the trivial one.*

Definition 3.5 ([15]) *Let M be a submodule of $\mathbb{Z}_{p^r}[x]^q$, written as the p-span of a p-generator sequence $(v_1, \cdots, v_N)$. Then $(v_1, \cdots, v_N)$ is called a **p-basis** of M if the vectors $v_1, \ldots, v_N$ are p-linearly independent in $\mathbb{Z}_{p^r}[x]^q$.*

For consistency with the field case, here we call the number of elements of a p-basis the **p-dimension** of M, denoted as $p\dim\,(M)$. The following definition adjusts the PLM property from the previous section to the specific structure of $\mathbb{Z}_{p^r}$.

Definition 3.6 ([17]) *Let $M = p\text{−span}\,\{v_1, \ldots, v_N\}$ be a submodule of $\mathbb{Z}_{p^r}[x]^q$. Then $\{v_1, \ldots, v_N\}$ has the **p-Predictable Leading Monomial (p-PLM) property** if for any $0 \neq f \in M$, written as*

$$f = a_1 v_1 + \cdots + a_N v_N, \tag{4}$$

where $a_1, \ldots, a_N \in \mathcal{A}_p[x]$, we have

$$\mathrm{lm}(f) = \max_{1 \leq i \leq N; a_i \neq 0} (\mathrm{lm}(a_i f_i)).$$

Note that, in contrast to the field case of the previous section, the above definition requires $a_i \in \mathcal{A}_p[x]$ rather than $a_i \in \mathcal{R}[x]$.

The next theorem is the analogon of Theorem 2.3; we omit its proof as it is very similar to the proof of Theorem 2.3.

Theorem 3.2 *Let $M = p\text{−span}\,\{v_1, \ldots, v_N\}$ be a submodule of $\mathbb{Z}_{p^r}[x]^q$. Assume that $\{v_1, \ldots, v_N\}$ has the p-PLM property. Let, for some $\ell \in \{1, \ldots, m\}$, $\mathcal{P}$ be a property of v_ℓ that is absent in p-linear combinations of the other v_i's. Then among all elements in M with property $\mathcal{P}$, v_ℓ has minimal leading monomial. More specifically, a parametrization of all elements f that have property $\mathcal{P}$ and minimal leading monomial is given by:*

$$f = a_\ell v_\ell + \sum_{i \neq \ell} a_i v_i,$$

with a_ℓ a nonzero constant in $\mathcal{A}_p$ and for all $i \neq \ell$ the polynomials $a_i \in \mathcal{A}_p[x]$ chosen such that $\mathrm{lm}(a_i v_i) \leq \mathrm{lm}(v_\ell)$.

The above theorem gives rise to two corollaries. The first corollary is the ring analogon of Corollary 2.4.

Corollary 3.3 *Let $M = p\text{−span}\,\{v_1, \ldots, v_{2r}\}$ be a submodule of $\mathbb{Z}_{p^r}[x]^2$. Assume that $\{v_1, \ldots, v_{2r}\}$ has the p-PLM property. Let $j^\star$ be such that $\mathrm{lpos}(v_{j^\star}) = 2$ and $\mathrm{ord}(v_{j^\star}) = r$. Then $v_{j^\star}$ is the lowest degree vector in M that has order r and leading position 2. A parametrization of all lowest degree vectors f that have order r and leading position 2 is given by*

$$f = a v_{j^\star} + \sum_{i \in \{1,\ldots,2r\} \setminus \{j^\star\}} a_i v_i,$$

where a is a nonzero constant in $\mathcal{A}_p$ and for all $i \neq j^\star$ the polynomials $a_i \in \mathcal{A}_p[x]$ chosen such that $\mathrm{lm}(a_i v_i) \leq \mathrm{lm}(v_{j^\star})$.

Proof Clearly all vectors in $\{v_1, \ldots, v_{2r}\}$ must have either different orders or different leading position, for otherwise the p-PLM property would not hold. In particular, this implies that $j^\star$ is unique. Now define f to have property $\mathcal{P}$ if ord $f = r$ and $\mathrm{lpos}(f) = 2$. It follows that this property is absent in p-linear combinations of the v_i's with $i \in \{1, \ldots, 2r\} \setminus \{j^\star\}$. The result now follows from Theorem 3.2.

The next corollary is the ring analogon of Corollary 2.5.

Corollary 3.4 *Let $M = p\text{−span}\,\{v_1, \ldots, v_{2r}\}$ be a submodule of $\mathbb{Z}_{p^r}[x]^2$. Assume that $\{v_1, \ldots, v_{2r}\}$ has the p-PLM property and write $v_i = [v_{i1} \quad v_{i2}]$ for $i = 1, \ldots, 2r$. Also assume that*

$$v_{i2}(0) = 0 \ \text{ for } i = 1, \ldots, r \text{ and } \mathrm{ord}\ v_{i2}(0) = 2r - i + 1 \ \text{ for } i = r+1, \ldots, 2r. \tag{5}$$

Then a parametrization of all lowest degree $f = [f_1 \quad f_2]$ in M with $\mathrm{ord}\ f_2(0) = r$ is given by

$$f = a_{r+1} v_{r+1} + \sum_{i \neq r+1} a_i v_i,$$

with a_{r+1} a nonzero constant in $\mathcal{A}_p$ and for all $i \neq r+1$ the polynomials $a_i \in \mathcal{A}_p[x]$ chosen such that $\mathrm{lm}(a_i v_i) \leq \mathrm{lm}(v_{r+1})$.

Proof Define $f = [f_1 \quad f_2]$ to have property $\mathcal{P}$ if ord $f_2(0) = r$, that is, $f_2(0)$ is a unit. The result now follows immediately from Theorem 3.2.

A timely question is now: do p-bases with the p-PLM property exist? Not surprisingly, the answer is affirmative, in fact the next theorem from [17] shows that such a basis can be written down immediately from a minimal Gröbner basis. Current computational packages are capable of computing minimal Gröbner bases for modules in $\mathbb{Z}_{p^r}[x]^q$. The underlying theory for this was developed in literature such as [5, 29] and references therein.

Theorem 3.5 *([17]) Let M be a submodule of $\mathbb{Z}_{p^r}[x]^q$ with minimal Gröbner basis $G = \{g_1, \ldots, g_m\}$, ordered so that $\mathrm{lm}(g_1) > \cdots > \mathrm{lm}(g_m)$. For $1 \leq j \leq m$ define*

$$\beta_j := \text{ord } (g_j) - \text{ord } (g_i),$$

where i is the smallest integer $> j$ with $\mathrm{lpos}(g_i) = \mathrm{lpos}(g_j)$. If i does not exist we define $\beta_j :=$ ord (g_j). Then $N = p\dim(M) = \beta_1 + \beta_2 + \cdots + \beta_m$ and the sequence V given as

$$V = (g_1, pg_1, \cdots, p^{\beta_1 - 1}g_1, g_2, pg_2, \cdots, p^{\beta_2 - 1}g_2, \cdots, g_m, pg_m, \cdots, p^{\beta_m - 1}g_m)$$

is a p-basis of M that has the p-PLM property.

Conform [17] we call V a **minimal Gröbner *p*-basis** of M. Note that the degrees of vectors in V are nonincreasing. In the next section we will use the above theorem in our proofs but not in our constructions, as we will construct minimal p-bases with the p-PLM property directly from the data in an iterative way.

4 Iterative algorithm

Let $\mathcal{R}$ be a finite chain ring. Consider a sequence $S_1, \ldots, S_N$ over $\mathcal{R}$. A polynomial $\lambda(x) = \lambda_0 + \lambda_1 x + \cdots + \lambda_L x^L \in \mathcal{R}[x]$, with λ_0 a unit is called a **feedback polynomial** of **length** L if

$$\lambda_0 S_{L+j} + \sum_{i=1}^{L} \lambda_i S_{L+j-i} = 0 \quad \text{for } j = 1, \ldots, N - L.$$

Note that $\lambda(x)$ is called "connection polynomial" in [23]. Also note that λ_L may be zero. Now consider the module M in $\mathcal{R}[x]^2$ defined as the rowspace of

$$\begin{bmatrix} x^{N+1} & 0 \\ -(S_N x^N + S_{N-1}x^{N-1} + \cdots + S_1 x) & 1 \end{bmatrix}. \tag{6}$$

We seek to find a lowest `top` degree vector $[\gamma(x) \quad \lambda(x)]$ in M for which $\lambda(0)$ is a unit. A useful interpretation is in terms of annihilation of sequences defined on the time-axis $\mathbb{Z}_-$: let σ be the forward shift operator, acting on sequences $\mathbf{w}$ on $\mathbb{Z}_-$ as $(\sigma\mathbf{w})(k) = w(k-1)$. Then we have

$$[\gamma(\sigma) \quad \lambda(\sigma)]\, \mathbf{b} = 0,$$

where $\mathbf{b} : \mathbb{Z}_- \mapsto \mathcal{R}^2$ is given by

$$\mathbf{b} := \left(\ldots \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ S_1 \end{bmatrix}, \begin{bmatrix} 0 \\ S_2 \end{bmatrix}, \ldots, \begin{bmatrix} 0 \\ S_N \end{bmatrix} \right). \tag{7}$$

Thus the pair of polynomials $(\gamma(x), \lambda(x))$ constitutes an annihilator for the data sequence $\mathbf{b}$. Our objective in this paper is to develop an iterative algorithm to construct feedback polynomials of shortest length. This length is called the **complexity** of the sequence. We require the algorithm to construct, at each step k, an annihilator for $\sigma^{N-k}\mathbf{b}$ of lowest `top` degree.

Remark 4.1 Note that the requirement to process $S_1, \ldots, S_k$ at step k (rather than $S_N, \ldots, S_{N-k+1}$) necessitates our formulation in terms of "feedback polynomial" λ, rather than its reverse version, denoted as d in [17]. In this paper we call d a characteristic polynomial of the sequence $S_1, \ldots, S_N$; the degree of a minimal characteristic polynomial equals the complexity of the sequence. Thus, a polynomial d written as $d(x) = d_L x^L + \cdots + d_0$ is a **characteristic polynomial** of $S_1, \ldots, S_N$ if d_L is a unit and

$$d_L S_{L+j} + \sum_{i=1}^{L} d_{L-i} S_{L+j-i} = 0 \text{ for } j = 1, \ldots, N-L.$$

Consider the reverse module M^{rev} in $\mathcal{R}[x]^2$, defined as the rowspace of

$$\begin{bmatrix} x^{N+1} & 0 \\ -(S_1 x^N + S_2 x^{N-1} + \cdots + S_N x) & 1 \end{bmatrix}. \tag{8}$$

It is easily verified that a minimal characteristic polynomial d for $S_1, \ldots, S_N$ is found in any vector $[h \quad d]$ in M^{rev} of leading position 2 that has minimal leading monomial, see [17]. Note that, by definition, whenever λ_L is a unit, a feedback polynomial $\lambda(x) = \lambda_0 + \lambda_1 x + \cdots + \lambda_L x^L$ of length L for a sequence $S_1, \ldots, S_N$ also serves as a characteristic polynomial of the reverse sequence $S_N, \ldots, S_1$; such a polynomial is called **bidirectional** as in [32], see also [28].

4.1 The field case

In this subsection we focus on the case that $\mathcal{R}$ is a field $\mathbb{F}$. The Berlekamp-Massey algorithm is a famous iterative algorithm that constructs a feedback polynomial of shortest length for a sequence $S_1, \ldots, S_N$ in $\mathbb{F}$. It processes a new data element S_k at each step k for $k = 1, \ldots, N$ and then produces a feedback polynomial of shortest length for $S_1, \ldots, S_k$. In this subsection we present an algorithm that resembles the Berlekamp-Massey algorithm but has a slightly different update rule. Our update rule yields an algorithm that iteratively constructs a minimal Gröbner basis at each step. The algorithm has the same complexity order $O(N^2)$ as the Berlekamp-Massey algorithm and shares several other properties, such as that it processes the data in a natural order and that it allows us to read off the solution at once.

A closer inspection shows that our algorithm resembles the algorithm of [7] albeit that our formulation uses 2×2 polynomial matrices, as in Berlekamp's original work [3], see also its formulation in the textbook [4] and earlier work [18]. This formulation facilitates explicit use of the PLM property yielding a parametrization of all solutions as well as a result on the reverse sequence, see Theorem 4.7 below. Furthermore, it facilitates an extension to sequences over the finite ring $\mathbb{Z}_{p^r}$, presented in Subsection 4.2 below. This extension proves nontrivial as it involves a careful use of the minimal Gröbner p-bases of the previous section.

Last but not least, the main result of this subsection (Theorem 4.7) shows that the auxiliary polynomials generated by the algorithm characterize the recurrence structure of the reverse sequence $S_k, \ldots, S_1$. Note that the Berlekamp-Massey algorithm generates similar auxiliary polynomials. Despite the vast literature on the Berlekamp-Massey algorithm, to our knowledge these auxiliary polynomials have not been interpreted in terms of the reverse sequence before.

In this subsection we focus on modules in $\mathbb{F}[x]^q$, where $\mathbb{F}$ is a field. We iteratively construct minimal Gröbner bases in a computationally efficient way. In order to be able to do this we first need to answer the following question: given a set of vectors in M, how do we recognize this set as a minimal Gröbner basis? The next lemma considers the special case in which M is a full rank module; the lemma holds for either `top` or `pot` monomial ordering and uses Definition 2.5 of a module's "degree", see also [2, Thm 7].

Lemma 4.2 *Let $M \in \mathbb{F}[x]^q$ be a module of dimension q and degree δ and let $G = \{g_1(x), \ldots, g_q(x)\} \subset M$. Then G is a minimal Gröbner basis of M if and only if the following two conditions hold:*

i) $\sum_{i=1}^{q} \deg g_i(x) = \delta$;

ii) all leading positions of the vectors $g_1(x), \ldots, g_q(x)$ *are different.*

Proof The "only if" part (i) follows immediately from Definition 2.5, whereas part (ii) follows from Definition 2.4. To prove the "if" part, let us assume that (i) and (ii) hold and let $\tilde{G} = \{\tilde{g}_1, \ldots, \tilde{g}_q\}$ be a minimal Gröbner basis of M. Then all leading positions of $\tilde{g}_1(x), \ldots, \tilde{g}_q(x)$ are different and we can order their leading monomials as $\mathrm{lm}(\tilde{g}_1) > \mathrm{lm}(\tilde{g}_2) > \cdots > \mathrm{lm}(\tilde{g}_q)$. Let $i \in \{1, \ldots, q\}$. Also, without restrictions, we may assume that g_i and $\tilde{g}_i$ have the same leading position. The predictable leading monomial property of $\tilde{G}$ (see Theorem 2.2) now implies that g_i is a linear combination of $\tilde{g}_1, \ldots, \tilde{g}_q$ that uses $\tilde{g}_i$ and it follows that $\mathrm{lm}(g_i) \geq \mathrm{lm}(\tilde{g}_i)$. Since g_i and $\tilde{g}_i$ have the same leading position, this implies that $\deg g_i \geq \deg \tilde{g}_i$. Consequently, by condition i) of the theorem

$$\delta = \sum_{i=1}^{q} \deg g_i \geq \sum_{i=1}^{q} \deg \tilde{g}_i = \delta,$$

where the last equality holds because of $\deg M = \delta$. It follows that $\deg g_i = \deg \tilde{g}_i$ for $i = 1, \ldots, q$. We also conclude that $\mathrm{lt}(g_i) = a_i \, \mathrm{lt}(\tilde{g}_i)$ for some nonzero constant a_i in $\mathbb{F}$ for $i = 1, \ldots, q$. As a result $L(G) = L(\tilde{G}) = L(M)$, so that, by Definition 2.2, G is a Gröbner basis for M. Furthermore, clearly G cannot be reduced, so that G is a minimal Gröbner basis for M.

In the next algorithm the unit vectors e_1 and e_2 are defined as $e_1 := [1 \quad 0]$ and $e_2 := [0 \quad 1]$; the algorithm is iterative and produces, at each step k, a matrix which is denoted by R_k; the two rows of R_k are denoted by g_1^k and g_2^k, respectively. Recall that σ denotes the forward shift operator.

Algorithm 4.3 Input data: $S_1, \ldots, S_N$.
Initialization: *Define*

$$R_0(x) := \begin{bmatrix} x & 0 \\ 0 & 1 \end{bmatrix}.$$

Proceed iteratively as follows for $k = 1, \ldots, N$**.**

- *Define the error sequence*

$$\mathbf{e}^k := R_{k-1}(\sigma)\mathbf{b}_k, \text{ written as } \mathbf{e}^k = (\ldots, \Delta^k), \tag{9}$$

  where $\mathbf{b}_k$ *is given as* $\mathbf{b}_k := \sigma^{N-k}\mathbf{b}$*, with* $\mathbf{b}$ *given by (7).*

- *Denote* $\Delta^k = \begin{bmatrix} \Delta_1^k & \Delta_2^k \end{bmatrix}^T$*, where* Δ^k *is given by (9).*

- *Define* $\mathcal{P}^k := \{i \in \{1, 2\} : \Delta_i^k \neq 0\}$.

- *Define* $i^\star$ *as the largest index* i *in* $\mathcal{P}^k$ *for which* $\mathrm{lm}(g_{i^\star}^{k-1})$ *is minimal (with respect to* `top` *monomial ordering).*

- *Define the update matrix* $E_k(x) := \frac{x}{\Delta_{i^\star}^k} e_1^T e_{i^\star} + e_2^T(-\Delta_2^k e_1 + \Delta_1^k e_2)$*, in other words*

$$E_k(x) = \begin{bmatrix} \frac{x}{\Delta_1^k} & 0 \\ -\Delta_2^k & \Delta_1^k \end{bmatrix} \text{ if } i^\star = 1$$

$$E_k(x) = \begin{bmatrix} 0 & \frac{x}{\Delta_2^k} \\ -\Delta_2^k & \Delta_1^k \end{bmatrix} \text{ if } i^\star = 2$$

- *Define* $R_k(x) := E_k(x)R_{k-1}(x)$.

Output: $R(x) := R_N(x)$.

Lemma 4.4 *Let $S_1, \ldots, S_N$ be a sequence over a field $\mathbb{F}$ and let $k \in \{1, \ldots, N\}$. Let $\mathbf{e}^k$ and Δ_j^k be defined as in Algorithm 4.3. Then*

i) $\sigma\mathbf{e}^k$ *equals the zero sequence*

ii) $\Delta_1^k = 1$.

Proof Clearly (i) and (ii) hold for $k = 1$. Let us now proceed by induction and assume that the lemma holds for some $k \in \{1, \ldots, N-1\}$. To prove (i), we observe that

$$\sigma\mathbf{e}^{k+1} = \sigma R_k(\sigma)\mathbf{b}_{k+1} = R_k(\sigma)\mathbf{b}_k = E_k(\sigma)R_{k-1}(\sigma)\mathbf{b}_k = E_k(\sigma)\mathbf{e}_k.$$

Using the induction hypothesis (i), it follows that

$$\sigma\mathbf{e}^{k+1} = E_k(\sigma)(\ldots, 0, 0, \Delta^k),$$

which equals the zero sequence by definition of the matrix $E_k(x)$. In other words, (i) holds. In order to prove (ii), we first observe that, by definition, $g_1^k(x) = \frac{x}{\Delta_1^k} g_1^{k-1}(x)$ if $i^* = 1$, or $g_1^k(x) = \frac{x}{\Delta_2^k} g_2^{k-1}(x)$ if $i^* = 2$. Thus, if $i^* = 1$ then

$$g_1^k(\sigma)\mathbf{b}_{k+1} = \frac{1}{\Delta_1^k} g_1^{k-1}(\sigma)\sigma\mathbf{b}_{k+1} = \frac{1}{\Delta_1^k} g_1^{k-1}(\sigma)\mathbf{b}_k = \frac{1}{\Delta_1^k}(\ldots, 0, 0, \Delta_1^k) = (\ldots, 0, 0, 1),$$

in other words $\Delta_1^{k+1} = 1$. Similarly, if $i^* = 2$ then also $g_1^k(\sigma)\mathbf{b}_{k+1} = (\ldots, 0, 0, 1)$, so that also in this case $\Delta_1^{k+1} = 1$.

Lemma 4.5 *Let $S_1, \ldots, S_N$ be a sequence over a field $\mathbb{F}$ and let $k \in \{0, \ldots, N\}$. Let R_k be the matrix obtained by applying Algorithm 4.3 to $S_1, \ldots, S_k$. Denote the two rows of R_k by $g_1^k := [g_{11}^k \;\; g_{12}^k]$ and $g_2^k := [g_{21}^k \;\; g_{22}^k]$. Then*

i) $R_k(\sigma)\mathbf{b}_k = 0$

ii) $\deg g_1^k + \deg g_2^k = k + 1$ *with respect to the* `top` *monomial ordering*

iii) $\mathrm{lpos}(g_1^k) \neq \mathrm{lpos}(g_2^k)$ *with respect to the* `top` *monomial ordering:*

iv) $g_1^k(0) = [0 \quad 0]$ *and* $g_{22}^k(0) = 1$

Proof Clearly all statements hold for $k = 0$. Let us now proceed by induction and assume that the lemma holds for some $k \in \{0, \ldots, N-1\}$. To prove (i), we observe that

$$R_{k+1}(\sigma)\mathbf{b}_{k+1} = E_{k+1}(\sigma)R_k(\sigma)\mathbf{b}_{k+1} = E_{k+1}(\sigma)\mathbf{e}^{k+1}.$$

Because of Lemma 4.4 (i) we thus have

$$R_{k+1}(\sigma)\mathbf{b}_{k+1} = E_{k+1}(\sigma)(\ldots, 0, 0, \Delta^{k+1})$$

which equals the zero sequence by definition of E_{k+1}. This proves (i). Further, by definition and Lemma 4.4 (ii), in the update operation $R_{k+1} = E_{k+1}R_k$ the degree of exactly one row of R_k is increased by 1, so that (ii) holds by induction. Similarly, it follows straightforwardly from the definition of E_{k+1} that (iii) and (iv) hold by induction.

Property 4.6 *Algorithm 4.3 has complexity order $O(N^2)$.*

Proof At each step k, the computation of the discrepancy Δ_2^k has complexity order $O(N)$. Also, the degrees of the entries of the update matrices $E_1, E_2, \ldots, E_{N-1}$ as well as of the matrix R_0 are at most 1. This implies that the rows of R_{k-1} have degree at most k for $1 \leq k \leq N$, so they certainly have degree $\leq N$. Using this as well as the sparseness of the matrix E_k, we establish that the update step $R_k(x) := E_k(x)R_{k-1}(x)$ has complexity order $O(N)$ as well. Since there are N steps this results in an overall complexity order $O(N^2)$.

Theorem 4.7 *Let $S_1, \ldots, S_N$ be a sequence over a field $\mathbb{F}$ and let R be the matrix obtained by applying Algorithm 4.3 to $S_1, \ldots, S_N$. Denote the two rows of R by $g_1 = [g_{11} \;\; g_{12}]$ and $g_2 = [g_{21} \;\; g_{22}]$; denote $\tilde{L} := \deg g_1$ and $L := \deg g_2$ with respect to the* `top` *monomial ordering. Then the complexity of the sequence equals L and g_{22} is a feedback polynomial of shortest length L. More specifically, a parametrization of all shortest length feedback polynomials is given by*

$$ag_{22} + bg_{12}, \tag{10}$$

where a is a nonzero constant in $\mathbb{F}$ and $b \in \mathbb{F}[x]$ such that $\deg b \leq L - \tilde{L}$.

Furthermore

- *in case that $\mathrm{lpos}(g_2) = 2$ then the feedback polynomial g_{22} is bidirectional and (10) also parametrizes all bidirectional minimal characteristic polynomials of the reverse sequence $S_N, \ldots, S_1$.*
- *in case that $\mathrm{lpos}(g_2) = 1$ then the complexity of the reverse sequence $S_N, \ldots, S_1$ equals $\tilde{L}$ and g_{12} is a minimal characteristic polynomial of $S_N, \ldots, S_1$. More specifically, a parametrization of all minimal characteristic polynomials of $S_N, \ldots, S_1$ is then given by*

  $$ag_{12} + bg_{22}, \tag{11}$$

  where a is a nonzero constant in $\mathbb{F}$ and $b \in \mathbb{F}[x]$ such that $\deg b \leq \tilde{L} - L$. In particular, any choice of $b \in \mathbb{F}[x]$ such that $\deg b \leq \tilde{L} - L$ and $b(0) \neq 0$ gives a bidirectional minimal characteristic polynomial of $S_N, \ldots, S_1$.

Proof From Lemma 4.2, Lemma 4.5(ii) and Lemma 4.5(iii), it follows that for all $k \in \{0, \ldots, N\}$ the set $\{g_1^k, g_2^k\}$ is a minimal Gröbner basis for the row space of

$$\begin{bmatrix} x^{k+1} & 0 \\ -(S_k x^k + S_{k-1}x^{k-1} + \cdots + S_1 x) & 1 \end{bmatrix}.$$

Thus it has the PLM property of Theorem 2.2 and the theorem now follows immediately from Corollary 2.5. The statements on the reverse sequence follow immediately from Remark 4.1 and Corollary 2.4 (note that $\mathrm{lpos}(g_2) = 1$ implies that $\mathrm{lpos}(g_1) = 2$).

Example 4.8 *Consider the sequence $S_1, S_2, S_3, S_4, S_5 = 4, 0, 4, 4, 2$ over the field $\mathbb{Z}_5$. Application of Algorithm 4.3 yields:*

$$\begin{aligned}
\Delta^1 &= \begin{bmatrix} 1 \\ 4 \end{bmatrix}, \quad \mathcal{P}^1 = \{1, 2\}, \quad i^* = 2, \quad R_1(x) = \begin{bmatrix} 0 & 4x \\ 1 & 1 \end{bmatrix} R_0(x) = \begin{bmatrix} 0 & 4x \\ x & 1 \end{bmatrix}; \\
\Delta^2 &= \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \quad \mathcal{P}^2 = \{1\}, \quad i^* = 1, \quad R_2(x) = \begin{bmatrix} x & 0 \\ 0 & 1 \end{bmatrix} R_1(x) = \begin{bmatrix} 0 & 4x^2 \\ x & 1 \end{bmatrix}; \\
\Delta^3 &= \begin{bmatrix} 1 \\ 4 \end{bmatrix}, \quad \mathcal{P}^3 = \{1, 2\}, \quad i^* = 2, \quad R_3(x) = \begin{bmatrix} 0 & 4x \\ 1 & 1 \end{bmatrix} R_2(x) = \begin{bmatrix} 4x^2 & 4x \\ x & 4x^2 + 1 \end{bmatrix}; \\
\Delta^4 &= \begin{bmatrix} 1 \\ 4 \end{bmatrix}, \quad \mathcal{P}^4 = \{1, 2\}, \quad i^* = 1, \quad R_4(x) = \begin{bmatrix} x & 0 \\ 1 & 1 \end{bmatrix} R_3(x) = \begin{bmatrix} 4x^3 & 4x^2 \\ 4x^2 + x & 4x^2 + 4x + 1 \end{bmatrix}; \\
\Delta^5 &= \begin{bmatrix} 1 \\ 4 \end{bmatrix}, \quad \mathcal{P}^5 = \{1, 2\}, \quad i^* = 2, \quad R_5(x) = \begin{bmatrix} 0 & 4x \\ 1 & 1 \end{bmatrix} R_4(x) = \begin{bmatrix} x^3 + 4x^2 & x^3 + x^2 + 4x \\ 4x^3 + 4x^2 + x & 3x^2 + 4x + 1 \end{bmatrix}.
\end{aligned}$$

By the above theorem, the complexity of the sequence equals $L = 3$ and $3x^2 + 4x + 1$ is a shortest length feedback polynomial. The complexity of the reverse sequence $2, 4, 4, 0, 4$ equals $\tilde{L} = 3$ and $x^3 + x^2 + 4x$ serves as a minimal characteristic polynomial of $2, 4, 4, 0, 4$. From the parametrization (11) we see that there is only one monic bidirectional minimal characteristic polynomial with value 1 at $x = 0$, namely $(x^3 + x^2 + 4x) + (3x^2 + 4x + 1) = x^3 + 4x^2 + 3x + 1$.

Remark 4.9 The earlier paper [18] formulates the Berlekamp-Massey algorithm in a similar format as Algorithm 4.3. From this it is clear that Algorithm 4.3 differs from the Berlekamp-Massey algorithm only in the definition of $i^\star$. More precisely, in the Berlekamp-Massey algorithm $i^\star$ equals the largest integer i in $\mathcal{P}^k$ such that g_i^{k-1} has minimal degree. Application of the Berlekamp-Massey algorithm in the above example gives the same first three steps leading to $R_3(x) = \begin{bmatrix} 4x^2 & 4x \\ x & 4x^2+1 \end{bmatrix}$. However, the next two steps give a different result:

$$\Delta^4 = \begin{bmatrix} 1 \\ 4 \end{bmatrix}, \quad \mathcal{P}^4 = \{1, 2\}, \quad i^* = 2, \quad R_4(x) = \begin{bmatrix} 0 & 4x \\ 1 & 1 \end{bmatrix} R_3(x) = \begin{bmatrix} 4x^2 & x^3 + 4x \\ 4x^2 + x & 4x^2 + 4x + 1 \end{bmatrix};$$
$$\Delta^5 = \begin{bmatrix} 1 \\ 4 \end{bmatrix}, \quad \mathcal{P}^5 = \{1, 2\}, \quad i^* = 2, \quad R_5(x) = \begin{bmatrix} 0 & 4x \\ 1 & 1 \end{bmatrix} R_4(x) = \begin{bmatrix} x^3 + 4x^2 & x^3 + x^2 + 4x \\ 3x^2 + x & x^3 + 4x^2 + 3x + 1 \end{bmatrix}.$$

In particular we see that here the rows of R_4 do not constitute a minimal Gröbner basis, since both rows have leading position 2. Similarly, the rows of R_5 do not constitute a minimal Gröbner basis. Thus this example illustrates a main difference between the Berlekamp-Massey algorithm and our Algorithm 4.3: by keeping track of not just degrees but also leading positions, our algorithm produces a minimal Gröbner basis, whereas the Berlekamp-Massey algorithm does not necessarily produce a minimal Gröbner basis since it only keeps track of degrees. The advantage of the Gröbner formulation is twofold: firstly, it allows for results on the reverse sequence, as detailed in Theorem 4.7; secondly, it allows for a transparent extension to the ring case, as detailed in the next subsection.

Remark 4.10 Throughout this subsection no special assumptions on the field $\mathbb{F}$ are made, thus the results are valid for both finite and infinite fields.

4.2 The ring case

In this subsection we focus on the ring case. Once the sequence $S_1, S_2, \ldots, S_N$ takes its values in a ring with zero divisors rather than a field, Berlekamp-Massey type algorithms such as Algorithm 4.3 are no longer applicable — its update matrices are no longer well defined because of the existence of zero divisors in the ring. Thus an alternative algorithm is needed. In this subsection we present such an algorithm. We focus on a finite sequence $S_1, \ldots, S_N$ from $\mathbb{Z}_{p^r}$ and seek to construct a feedback polynomial of shortest length (including parametrization) by iteratively processing the data in the natural order $S_1, \ldots, S_N$. Again, our key object of interest is the module M defined as the row space of (6). Our algorithm constructs a $2r \times 2$ polynomial matrix R whose rows are a p-basis for M that has the p-PLM property. We remark that the operations implemented at each step are in $\mathbb{Z}_{p^r}$. The algorithm has complexity order $O(rN^2)$. The r appears because of the number of rows of the matrix R that is constructed at each step.

In the next algorithm the $2r$ rows of the $2r \times 2$ matrix R_k are denoted by $v_1^k, \ldots, v_{2r}^k$, whereas the $2r$ rows of the $2r \times 2r$ identity matrix are denoted by $e_1, e_2, \ldots, e_{2r}$.

Algorithm 4.11 Input data: $S_1, \ldots, S_N$.

Initialization: *Define*

$$R_0(x) := \begin{bmatrix} x & 0 \\ px & 0 \\ \vdots & \vdots \\ p^{r-1}x & 0 \\ 0 & 1 \\ 0 & p \\ \vdots & \vdots \\ 0 & p^{r-1} \end{bmatrix}.$$

Proceed iteratively as follows for $k = 1, \ldots, N$.

- *Define the error sequence*
$$\mathbf{e}^k := (\ldots, \Delta^k) := R_{k-1}(\sigma)\mathbf{b}_k,$$
where $\mathbf{b}_k$ is given as $\mathbf{b}_k := \sigma^{N-k}\mathbf{b}$, with $\mathbf{b}$ given by (7).

- *Denote $\Delta^k = \begin{bmatrix} \Delta_1^k & \cdots & \Delta_{2r}^k \end{bmatrix}^T$ and define $\mathcal{P}_0^k := \{i \in \{1, \ldots, 2r\} : \Delta_i^k = 0\}$.*

- *For $i = 1, 2, \ldots, 2r$ and $i \notin \mathcal{P}_0^k$, write $\Delta_i^k = \theta_i^k p^{\ell_i^k - 1}$, where θ_i^k is a unit and $\ell_i^k \in \{1, \ldots, r\}$.*

- *For $j = 1, \ldots, r$, define $\mathcal{P}_j^k := \{i \in \{1, \ldots, 2r\} : \ell_i^k = j\}$.*

- *For $j = 1, \ldots, r$, define $i_j^\star$ as the largest index i in $\mathcal{P}_j^k$ for which $\mathrm{lm}(v_i^{k-1})$ is minimal.*

- *Define the update matrix $E_k(x)$ as*
$$E_k(x) := \sum_{j=1}^{r} \frac{x}{\theta_{i_j^*}^k} e_j^T e_{i_j^*} + \sum_{i \in \mathcal{P}_0^k} e_i^T e_i + \sum_{j=1}^{r} A_j + \sum_{j=1}^{r} B_j,$$
where
$$A_j = \sum_{i \in \mathcal{P}_j^k \setminus \{j, i_j^\star\}} e_i^T(-\theta_i^k e_{i_j^\star} + \theta_{i_j^\star}^k e_i)$$
and
$$B_j = e_{i_j^\star}^T(-\theta_{i_j^\star}^k e_j + \theta_j^k e_{i_j^\star}).$$

- *Define $R_k(x) := E_k(x)R_{k-1}(x)$.*

Output: $R(x) := R_N(x)$.

Lemma 4.12 *Let $S_1, \ldots, S_N$ be a sequence over $\mathbb{Z}_{p^r}$. Let $k \in \{1, \ldots, N\}$ and let $\mathbf{e}^k$ and Δ_j^k be defined as in Algorithm 4.11. Then*

i) $\sigma\mathbf{e}^k$ equals the zero sequence

ii) $\Delta_j^k = p^{j-1}$ for $j = 1, \ldots, r$.

Proof Clearly (i) and (ii) hold for $k = 1$. Let us now proceed by induction and assume that the lemma holds for some $k \in \{1, \ldots, N-1\}$. To prove (i), we observe that

$$\sigma \mathbf{e}^{k+1} = \sigma R_k(\sigma)\mathbf{b}_{k+1} = R_k(\sigma)\mathbf{b}_k = E_k(\sigma)R_{k-1}(\sigma)\mathbf{b}_k = E_k(\sigma)\mathbf{e}_k.$$

Using the induction hypothesis (i), it follows that

$$\sigma \mathbf{e}^{k+1} = E_k(\sigma)(\ldots, 0, 0, \Delta^k),$$

which equals the zero sequence by definition of the matrix $E_k(x)$. In other words, (i) holds. In order to prove (ii), let $j \in \{1, \ldots, r\}$. By definition, the jth row of the update matrix $E_k(x)$ has all zeros except for the i_j^*-entry which equals $x/\theta_{i_j^*}^k$. Thus $v_j^k(x)$, which is the jth row of the matrix $R_k(x)$, equals

$$v_j^k(x) = \frac{x}{\theta_{i_j^*}^k} v_{i_j^*}^{k-1}(x).$$

Using induction hypothesis (i), it then follows that the sequence $(\ldots, 0, 0, \Delta_j^{k+1})$ equals

$$v_j^k(\sigma)\mathbf{b}_{k+1} = \frac{1}{\theta_{i_j^*}^k} v_{i_j^*}^{k-1}(\sigma)\sigma\mathbf{b}_{k+1} = \frac{1}{\theta_{i_j^*}^k} v_{i_j^*}^{k-1}(\sigma)\mathbf{b}_k = \frac{1}{\theta_{i_j^*}^k}(\ldots, 0, 0, \Delta_{i_j^*}^k) = (\ldots, 0, 0, p^{j-1}).$$

Thus (ii) holds.

Lemma 4.13 *Let $S_1, \ldots, S_N$ be a sequence over $\mathbb{Z}_{p^r}$ and let $k \in \{0, \ldots, N\}$. Let R_k be the matrix obtained by applying Algorithm 4.11 to $S_1, \ldots, S_k$. Denote the rows of R_k by $v_1^k, \ldots, v_{2r}^k$.*

i) $R_k(\sigma)\mathbf{b}_k = 0$

ii) $\deg v_1^k + \ldots + \deg v_{2r}^k = r(k+1)$ *with respect to* `top` *monomial ordering*

iii) if $i, j \in \{1, \ldots, 2r\}$, with $i \neq j$, then $\mathrm{lpos}(v_i^k) = \mathrm{lpos}(v_j^k) \Rightarrow \mathrm{ord}(v_i^k) \neq \mathrm{ord}(v_j^k)$ *with respect to* `top` *monomial ordering*

Proof Clearly the lemma holds for $k = 0$. Let us now proceed by induction and assume that the lemma holds for some $k \in \{0, \ldots, N-1\}$. To prove (i), we observe that

$$R_{k+1}(\sigma)\mathbf{b}_{k+1} = E_{k+1}(\sigma)R_k(\sigma)\mathbf{b}_{k+1} = E_{k+1}(\sigma)\mathbf{e}^{k+1}.$$

Because of Lemma 4.12 (i) we thus have

$$R_{k+1}(\sigma)\mathbf{b}_{k+1} = E_{k+1}(\sigma)(\ldots, 0, 0, \Delta^{k+1})$$

which equals the zero sequence by definition of E_{k+1}. This proves (i). Further, by definition and Lemma 4.12 (ii), in the update operation $R_{k+1} = E_{k+1}R_k$ the degrees of exactly r rows of R_k are increased by 1, so that (ii) holds by induction. Similarly, it follows straightforwardly from the definition of E_{k+1} that (iii) holds by induction.

Property 4.14 *Algorithm 4.11 has complexity order $O(rN^2)$.*

Proof At each step k, the total computation of the discrepancies $\Delta^k_{r+1}, \ldots, \Delta^k_{2r}$ has complexity order $O(rN)$. Also, the degrees of the entries of the update matrices $E_1, E_2, \ldots, E_{N-1}$ as well as of the matrix R_0 are at most 1. This implies that the rows of R_{k-1} have degree at most k for $1 \leq k \leq N$, so they certainly have degree $\leq N$. Using this, as well as the sparseness of the matrix E_k, we establish that the update step $R_k(x) := E_k(x)R_{k-1}(x)$ has complexity order $O(rN)$ as well. Since there are N steps this results in an overall complexity order $O(rN^2)$.

The following lemma is a key ingredient in the proof of our later main result. The proof is highly nontrivial and relies on the preceding lemma's.

Lemma 4.15 *Let $S_1, \ldots, S_N$ be a sequence over $\mathbb{Z}_{p^r}$, let M be the module defined as the row space of (6). Let R be the matrix obtained by applying Algorithm 4.11 to $S_1, \ldots, S_N$. Denote the rows of R by $v_1, \ldots, v_{2r}$. Then $\{v_1, \ldots, v_{2r}\}$ is a p-basis of M that has the p-PLM property.*

Proof Let $\tilde{V} = (\tilde{v}_1, \ldots, \tilde{v}_{2r})$ be a minimal Gröbner p-basis of M, as defined in Theorem 3.5. Note that, by definition,

$$\mathrm{lm}(\tilde{v}_{i+1}) \leq \mathrm{lm}(\tilde{v}_i) \text{ for } i = 1, \ldots, 2r-1. \tag{12}$$

and for $i < j$ we have

$$\mathrm{lm}(\tilde{v}_i) = \mathrm{lm}(\tilde{v}_j) \Rightarrow \mathrm{ord}(\tilde{v}_i) > \mathrm{ord}(\tilde{v}_j). \tag{13}$$

As a result, defining $G_1 := \{\tilde{v} \in \tilde{V} \mid \mathrm{lpos}(\tilde{v}) = 1\}$ and $G_2 := \{\tilde{v} \in \tilde{V} \mid \mathrm{lpos}(\tilde{v}) = 2\}$, there exists a bijection $\phi : G_1 \to G_2$ such that $\mathrm{ord}(\phi(\tilde{v})) = r + 1 - \mathrm{ord}(\tilde{v})$ for all $\tilde{v} \in G_1$. Clearly deg det col $(\tilde{v}, \phi(\tilde{v})) =$ deg $\tilde{v}$ + deg $\phi(\tilde{v})$ for all $\tilde{v} \in G_1$. On the other hand,

$$\text{col } (\tilde{v}, \phi(\tilde{v})) = U(x) \begin{bmatrix} x^{N+1} & 0 \\ -(S_N x^N + S_{N-1}x^{N-1} + \cdots + S_1 x) & 1 \end{bmatrix},$$

for some polynomial matrix $U(x)$, so that deg $\tilde{v}$ + deg $\phi(\tilde{v}) \geq N + 1$ for all $\tilde{v} \in G_1$. As a result,

$$\sum_{i=1}^{2r} \deg \tilde{v}_i \geq r(N+1). \tag{14}$$

Let us now examine $\{v_1, \ldots, v_{2r}\}$, where $v_1, \ldots, v_{2r}$ are the rows of R. It follows from Lemma 4.13 (iii) that, for $j = 1, 2$, there are r vectors in $\{v_1, \ldots, v_{2r}\}$ of leading position j that each have a different order. This implies that there exists a permutation g on $\{1, 2, \ldots, 2r\}$, such that $\mathrm{lpos}(v_{g(i)}) = \mathrm{lpos}(\tilde{v}_i)$ and $\mathrm{ord}(v_{g(i)}) = \mathrm{ord}(\tilde{v}_i)$ for $i = 1, \ldots, 2r$. Also, $v_{g(i)}$ can be expressed as a p-linear combination of the $\tilde{v}_j$'s. By Theorem 3.5, the sequence $(\tilde{v}_1, \ldots, \tilde{v}_{2r})$ has the p-PLM property, so that this linear combination must involve $\tilde{v}_i$ and it follows that $\mathrm{lm}(v_{g(i)}) \geq \mathrm{lm}(\tilde{v}_i)$. Since we are using the `top` monomial ordering, this implies that deg $(v_{g(i)}) \geq$ deg $(\tilde{v}_i)$. It now follows from (14) and Lemma 4.13 (ii) that equality must hold, that is, deg $(v_{g(i)}) =$ deg $(\tilde{v}_i)$ for $i = 1, \ldots, 2r$. In summary we thus have for $i = 1, \ldots, 2r$

$$\mathrm{lm}(v_{g(i)}) = \mathrm{lm}(\tilde{v}_i) \quad \text{and} \quad \mathrm{ord}(v_{g(i)}) = \mathrm{ord}(\tilde{v}_i). \tag{15}$$

We next prove by induction that $(v_{g(1)}, \ldots, v_{g(2r)})$ is a p-generator sequence whose p-span equals M. First ($i = 2r$) we observe that we must have $v_{g(2r)} = a_{2r}\tilde{v}_{2r}$ for some constant unit a_{2r} in $\mathbb{Z}_{p^r}$. Since $(\tilde{v}_1, \ldots, \tilde{v}_{2r})$ is a p-generator sequence, it follows that

$$pv_{g(2r)} = a_{2r}p\tilde{v}_{2r} = 0 \tag{16}$$

and $\tilde{v}_{2r} = a_{2r}^{-1}v_{g(2r)} \in p\text{−span } \{v_{g(2r)}\}$. Proceeding by induction, we assume that for some $i = k + 1 \in \{2, \ldots, 2r\}$ the sequence $(v_{g(i)}, \cdots, v_{g(2r)})$ is a p-generator sequence with

$$p\text{−span } (v_{g(i)}, \cdots, v_{g(2r)}) = p\text{−span } (\tilde{v}_i, \cdots, \tilde{v}_{2r}).$$

Since $(\tilde{v}_1, \ldots, \tilde{v}_{2r})$ is a p-basis of M, we can write

$$v_{g(k)} = \sum_{j=1}^{2r} a_j \tilde{v}_j$$

for some polynomial $a_j \in \mathcal{A}_p[x]$. The p-PLM property of $(\tilde{v}_1, \ldots, \tilde{v}_{2r})$ together with (12), (13) and (15) implies that $a_j = 0$ for $j < k$ and that a_k is a nonzero constant. Thus,

$$v_{g(k)} = a_k \tilde{v}_k + v \text{ with } v \in p\text{−span } (\tilde{v}_{k+1}, \cdots, \tilde{v}_{2r}) \text{ and } a_k \text{ a constant unit in } \mathbb{Z}_{p^r}.$$

Then $pv_{g(k)} = a_k p\tilde{v}_k + pv \in p$−span $(\tilde{v}_{k+1}, \cdots, \tilde{v}_{2r})$, so that $pv_{g(k)} \in p$−span $(v_{g(k+1)}, \cdots, v_{g(2r)})$ by the induction hypothesis. As a result, $\tilde{v}_k = a_k^{-1} v_{g(k)} - a_k^{-1} v \in p$−span $\{v_{g(k)}, \ldots, v_{g(2r)}\}$. In conclusion, for $i = k$ we have $(v_{g(i)}, \ldots, v_{g(2r)})$ is a p-generator sequence and p−span $\{v_{g(i)}, \ldots, v_{g(2r)}\} = p$−span $\{\tilde{v}_i, \ldots, \tilde{v}_{2r}\}$. By induction it now follows that $(v_{g(1)}, \ldots, v_{g(2r)})$ is a p-generator sequence with p−span $\{v_{g(1)}, \ldots, v_{g(2r)}\} = p$−span $\{\tilde{v}_1, \ldots, \tilde{v}_{2r}\} = M$. Finally, we prove that $\{v_1, \ldots, v_{2r}\}$ has the p-PLM property. For this, let

$$f = a_1 v_1 + \cdots + a_{2r} v_{2r} \tag{17}$$

with $a_1, \ldots, a_{2r} \in \mathcal{A}_p[x]$. Evidently $\mathrm{lm}(f) \leq \max_{1 \leq i \leq 2r; a_i \neq 0}(\mathrm{lm}(a_i v_i))$. As a result, in order to prove the p-PLM property we need only prove that this upperbound is reached. By grouping together all vectors $a_i v_i$ in (17) that have the same leading position we write

$$f = f_1 + f_2,$$

where $f_j = 0$ if position j is not used in (17). It now follows from the p-adic decomposition and Lemma 4.13 (iii) that $\mathrm{lpos}(f_j) = j$ for $j = 1, 2$ whenever $f_j \neq 0$. More specifically, we then have $\mathrm{lm}(f_j) = \mathrm{lm}(a_{\ell_j} v_{\ell_j})$ for some $\ell_j \in \{1, \ldots, 2r\}$. In case either $f_1 = 0$ or $f_2 = 0$ the p-PLM property then follows immediately. In case both f_1 and f_2 are nonzero, we recall that their leading positions differ so that, without restrictions, we may assume that $\mathrm{lm}(f_1) < \mathrm{lm}(f_2)$. Then $\mathrm{lm}(f) = \mathrm{lm}(f_2) = \mathrm{lm}(a_{\ell_2} v_{\ell_2})$, which proves the p-PLM property. The property implies, in particular, that $\{v_1, \ldots, v_{2r}\}$ is a p-basis of M.

Lemma 4.16 *Let $S_1, \ldots, S_N$ be a sequence over $\mathbb{Z}_{p^r}$ and let $k \in \{0, \ldots, N\}$. Let R_k be the matrix obtained by applying Algorithm 4.11 to $S_1, \ldots, S_k$, with rows $v_1^k, \ldots, v_{2r}^k$; denote $v_j^k := \begin{bmatrix} v_{j1}^k & v_{j2}^k \end{bmatrix}$ for $j = 1, \ldots, 2r$. Then*

i) $v_j^k(0) = [0 \quad 0]$ *for* $j = 1, \ldots, r$

ii) $\mathrm{ord}(v_{j2}^k(0)) = 2r - j + 1$ *for* $j = r+1, \ldots, 2r$ *with respect to the* `top` *monomial ordering*

iii) $\mathrm{lm}(v_j^k) \geq \mathrm{lm}(v_{j+1}^k)$ *for* $j = r+1, \ldots, 2r-1$ *with respect to the* `top` *monomial ordering.*

Proof All conditions are obviously satisfied for $k = 0$. Let us now proceed by induction and assume that the lemma holds for some $k \in \{0, 1, \ldots, N-1\}$.
To prove (i), first note that, by Lemma 4.12 (ii), we have $j \in \mathcal{P}_j^k$ for $j \in \{1, \ldots, r\}$. As a result, for any $j \in \{1, \ldots, r\}$ we have $v_j^{k+1}(x) = x v_j^k(x)$, if $i_j^* = j$, and $v_j^{k+1}(x) = \frac{x}{\theta_{i_j^*}^{k+1}} v_{i_j^*}^k(x)$, otherwise. Thus $v_j^{k+1}(0) = [0 \;\; 0]$. To prove (ii), let $j \in \{r+1, \ldots, 2r\}$. We distinguish two cases:
Case 1: $j \in \mathcal{P}_0^k$. Then $v_j^{k+1}(x) = v_j^k(x)$ and (ii) follows immediately by induction hypothesis (ii).
Case 2: $j \in \mathcal{P}_\ell^k$ for some $\ell \in \{1, \ldots, r\}$, i.e., $\Delta_j^k = \theta_j^k p^{\ell-1}$ for some unit θ_j^k. We distinguish four subcases:

Case 2A: $i^\star_\ell = \ell$. Then $\theta^k_{i^\star_\ell} = 1$ so that $v^{k+1}_j(x) = -\theta^k_j v^k_\ell(x) + v^k_j(x)$. Since $v^k_\ell(0) = [0 \;\; 0]$ by induction hypothesis (i), it follows that $v^{k+1}_j(0) = v^k_j(0)$ so that (ii) holds by induction hypothesis (ii).
Case 2B: $i^\star_\ell = j$. Then again $v^{k+1}_j(x) = -\theta^k_j v^k_\ell(x) + v^k_j(x)$ and the reasoning proceeds as in case 2A.
Case 2C: $i^\star_\ell > j$. Then $v^{k+1}_j(x) = -\theta^k_j v^k_{i^\star_\ell}(x) + \theta^k_{i^\star_\ell} v^k_j(x)$. By induction hypothesis (ii), $\mathrm{ord}(v^k_{i^\star_\ell 2}(0)) < \mathrm{ord}(v^k_{j2}(0))$, so that $\mathrm{ord}(v^{k+1}_{j2}(0)) = \mathrm{ord}(v^k_{j2}(0)) = 2r - j + 1$.
Case 2D: $i^\star_\ell < j$ and $i^\star_\ell \neq \ell$. By definition of $i^\star_\ell$ and induction hypothesis (iii) this case cannot happen.
To prove (iii), let $j \in \{r+1, \ldots, 2r-1\}$. Because of Lemma 4.15, we can write pv^{k+1}_j as a p-linear combination of $v^{k+1}_1, \ldots, v^{k+1}_{2r}$. Because of (i) and (ii) above, this p-linear combination must use v^{k+1}_{j+1} and it follows that $\mathrm{lm}(pv^k_j) \geq \mathrm{lm}(v^k_{j+1})$ which implies that $\mathrm{lm}(v^k_j) \geq \mathrm{lm}(v^k_{j+1})$, i.e. (iii) holds.

We now present our main result.

Theorem 4.17 *Let $S_1, \ldots, S_N$ be a sequence over $\mathbb{Z}_{p^r}$ and let R be the matrix obtained by applying Algorithm 4.11 to $S_1, \ldots, S_N$. Denote the rows of R by $v_1, \ldots, v_{2r}$; denote $L := \deg v_{r+1}$ with respect to the* `top` *monomial ordering. Then the complexity of the sequence equals L and $v_{(r+1)2}$ is a feedback polynomial of shortest length L. More specifically, a parametrization of all shortest length feedback polynomials is given by*

$$a v_{(r+1)2} + \sum_{j \in \{1,\ldots,2r\} \setminus \{r+1\}} a_j v_{j2},$$

where a is a nonzero constant in $\mathcal{A}_p$ and for all $j \neq r+1$ the polynomial $a_j \in \mathcal{A}_p[x]$ chosen such that $\deg(a_j) \leq L - \deg v_j$. Furthermore, let $j^\star$ be such that $\mathrm{lpos}(v_{j^\star}) = 2$ and $\mathrm{ord}(v_{j^\star}) = r$. Let $\tilde{L} := \deg v_{j^\star}$. Then the complexity of the reverse sequence $S_N, \ldots, S_1$ equals $\tilde{L}$ and $v_{j^\star 2}$ is a minimal characteristic polynomial of $S_N, \ldots, S_1$. More specifically, a parametrization of all minimal characteristic polynomials of $S_N, \ldots, S_1$ is given by

$$a v_{j^\star 2} + \sum_{j \in \{1,\ldots,2r\} \setminus \{j^\star\}} a_j v_{j2}, \tag{18}$$

with a a nonzero constant in $\mathcal{A}_p$ and for all $j \neq j^\star$ the polynomials $a_j \in \mathcal{A}_p[x]$ chosen such that $\deg(a_j) \leq \tilde{L} - \deg v_j$. In particular,

- *in case $j^\star = r+1$ then $v_{j^\star 2}$ is bidirectional and (18) also parametrizes all bidirectional minimal characteristic polynomials of the reverse sequence $S_N, \ldots, S_1$*
- *in case $j^\star \neq r+1$ then any choice of $a_{r+1} \in \mathcal{A}_p[x]$ such that $\deg(a_{r+1}) \leq \tilde{L} - \deg v_{r+1}$ and $a_{r+1}(0) \neq 0$ gives a bidirectional minimal characteristic polynomial of $S_N, \ldots, S_1$.*

Proof The first parametrization follows immediately from Lemma 4.15, Lemma 4.16 and Corollary 3.4. Let us now consider the reverse sequence in order to prove the second parametrization (18). From Remark 4.1 we know that a minimal characteristic polynomial of $S_N, \ldots, S_1$ is given by a vector of M with leading position 2 and order r, of minimal degree. By Lemma 4.15 the set $\{v_1, \ldots, v_{2r}\}$ is a p-basis of M with the p-PLM property. Corollary 3.3 now implies (18).

The following example illustrates the workings of Algorithm 4.11 and shows how Theorem 4.17 is used.

Example 4.18 *Consider the sequence $S_1, S_2, S_3, S_4, S_5 = 6, 3, 1, 5, 6$ over the ring $\mathbb{Z}_9$ (thus data as in the example in [30]). Application of Algorithm 4.11 yields:*

$$\Delta^1 = \begin{bmatrix} 1 \\ 3 \\ 6 \\ 0 \end{bmatrix}, \;\; \mathcal{P}^1_0 = \{4\}, \;\; \mathcal{P}^1_1 = \{1\}, \;\; \mathcal{P}^1_2 = \{2, 3\}, \;\; i^*_1 = 1, \;\; i^*_2 = 3,$$

$$R_1(x) = \begin{bmatrix} x & 0 & 0 & 0 \\ 0 & 0 & 5x & 0 \\ 0 & 7 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} R_0(x) = \begin{bmatrix} x^2 & 0 \\ 0 & 5x \\ -6x & 1 \\ 0 & 3 \end{bmatrix};$$

$$\Delta^2 = \begin{bmatrix} 1 \\ 3 \\ 3 \\ 0 \end{bmatrix}, \ \mathcal{P}_0^2 = \{4\}, \ \mathcal{P}_1^2 = \{1\}, \ \mathcal{P}_2^2 = \{2,3\}, \ i_1^* = 1, \ i_2^* = 3,$$

$$R_2(x) = \begin{bmatrix} x & 0 & 0 & 0 \\ 0 & 0 & x & 0 \\ 0 & -1 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} R_1(x) = \begin{bmatrix} x^3 & 0 \\ 3x^2 & x \\ -6x & 4x+1 \\ 0 & 3 \end{bmatrix};$$

$$\Delta^3 = \begin{bmatrix} 1 \\ 3 \\ 4 \\ 3 \end{bmatrix}, \ \mathcal{P}_0^3 = \emptyset, \ \mathcal{P}_1^3 = \{1,3\}, \ \mathcal{P}_2^3 = \{2,4\}, \ i_1^* = 3, \ i_2^* = 4,$$

$$R_3(x) = \begin{bmatrix} 0 & 0 & x/4 & 0 \\ 0 & 0 & 0 & x \\ -4 & 0 & 1 & 0 \\ 0 & -1 & 0 & 1 \end{bmatrix} R_2(x) = \begin{bmatrix} 3x^2 & x^2+7x \\ 0 & 3x \\ -4x^3-6x & 4x+1 \\ 6x^2 & 8x+3 \end{bmatrix};$$

Using Theorem 4.17, we conclude from the matrix R_3 that the sequence $6,3,1$ has highest complexity possible, namely $L = 3$, whereas its reverse sequence $1,3,6$ has complexity $\tilde{L} = 2$ and minimal characteristic polynomial $x^2 + 7x$. A parametrization of all monic minimal characteristic polynomials of $1,3,6$ is given by $x^2 + 7x + b(8x+3)$, where $b \in \mathbb{Z}_9$. This parametrization shows that there exists no bidirectional characteristic polynomial for this sequence. Proceeding with the next element $S_4 = 5$ of the sequence we obtain

$$\Delta^4 = \begin{bmatrix} 1 \\ 3 \\ 0 \\ 5 \end{bmatrix}, \ \mathcal{P}_0^4 = \{3\}, \ \mathcal{P}_1^4 = \{1,4\}, \ \mathcal{P}_2^4 = \{2\}, \ i_1^* = 4, \ i_2^* = 2,$$

$$R_4(x) = \begin{bmatrix} 0 & 0 & 0 & x/5 \\ 0 & x & 0 & 0 \\ 0 & 0 & 1 & 0 \\ -5 & 0 & 0 & 1 \end{bmatrix} R_3(x) = \begin{bmatrix} 3x^3 & 7x^2+6x \\ 0 & 3x^2 \\ -4x^3-6x & 4x+1 \\ 0 & 4x^2+3 \end{bmatrix};$$

$$\Delta^5 = \begin{bmatrix} 1 \\ 3 \\ 8 \\ 4 \end{bmatrix}, \ \mathcal{P}_0^5 = \emptyset, \ \mathcal{P}_1^5 = \{1,3,4\}, \ \mathcal{P}_2^5 = \{2\}, \ i_1^* = 4, \ i_2^* = 2,$$

$$R_5(x) = \begin{bmatrix} 0 & 0 & 0 & x/4 \\ 0 & x & 0 & 0 \\ 0 & 0 & 4 & -8 \\ -4 & 0 & 0 & 1 \end{bmatrix} R_4(x) = \begin{bmatrix} 0 & x^3+3x \\ 0 & 3x^3 \\ 2x^3-6x & 4x^2+7x+7 \\ -3x^3 & 3x^2+3x+3 \end{bmatrix}.$$

By the above theorem, the complexity of the sequence equals $L = 3$ and $4x^2 + 7x + 7$ is a shortest length feedback polynomial, normalized to $7x^2 + x + 1$. It is not unique: a parametrization of all normalized shortest length feedback polynomials of length 3 is given by

$$7x^2 + x + 1 + a(x^3 + 3x), \tag{19}$$

where a is a constant in $\mathbb{Z}_9$. The complexity of the reverse sequence $6,5,1,3,6$ equals $\tilde{L}=3$ and x^3+3x serves as a minimal characteristic polynomial of $6,5,1,3,6$. It is not unique, a parametrization of all monic minimal characteristic polynomials of $6,5,1,3,6$ is given by

$$x^3+3x+b(4x^2+7x+7),$$

where b is a constant in $\mathbb{Z}_9$. For comparison, in our notation, the algorithm of [30] produces the matrix

$$\begin{bmatrix} \star & \star \\ \star & \star \\ x^3-6x & x^3+7x^2+4x+1 \\ -3x^3 & 5x^3+3x^2+3 \end{bmatrix}$$

rather than $R_5(x)$. Thus it produces the shortest feedback polynomial x^3+7x^2+4x+1. We verify that this polynomial is indeed in our parametrization (19), namely for the parameter choice $a=1$. Note that it follows from the above parametrization (18) that x^3+7x^2+4x+1 is the unique monic bidirectional minimal characteristic polynomial of $6,5,1,3,6$ that has constant term 1.

5 Conclusions

In his 1969 paper [23] Massey showed that the Berlekamp-Massey algorithm yields a parametrization of all shortest feedback shift registers for a given finite sequence $S_1,\ldots,S_N$ of elements of a field. The main contribution of our paper is an iterative algorithm of complexity $O(rN^2)$ that yields such a parametrization when $S_1,\ldots,S_N$ are elements of a finite chain ring such as $\mathbb{Z}_{p^r}$. Although relying on nontrivial theories of p-Gröbner bases and p-linear dependence, the algorithm is highly practical as well as efficient, as we illustrated in an example. It is thus shown in this paper that it is possible to have as much "grip" on this fundamental problem in the ring case as in the field case, despite the existence of zero divisors.

Existing methods for the ring case, such as in [13, 20, 24, 25, 27, 30] yield a solution but no parametrization. Our algorithm can also be used to yield a parametrization of all feedback shift registers of a certain pre-specified length, analogous to [2] which has proved useful for list decoding. For the field case (any field, not just $\mathbb{Z}_p$), our algorithm resembles a normalized version of the Gröbner-based iterative algorithm of [7].

We have shown that our algorithm also simultaneously yields all shortest feedback shift registers for the reverse sequence $S_N,\ldots,S_1$. This is an additional fundamental result that, to our knowledge, has not appeared in the literature before.

6 Acknowledgment

We would like to thank Anna-Lena Trautmann as well as the anonymous reviewers for helpful comments.

References

[1] W. W. Adams and P. Loustaunau. *An introduction to Gröbner Bases*, volume 3 of *Graduate Stud. Math.* American Mathematical Society, 1994.

[2] M. Ali and M. Kuijper. A parametric approach to list decoding of Reed-Solomon codes using interpolation. *IEEE Trans. Inf. Th.*, IT-57:6718–6728, 2011.

[3] E.R. Berlekamp. *Algebraic Coding Theory.* McGraw-Hill, New York, 1968.

[4] R.E. Blahut. *Theory and practice of error control codes.* Addison-Wesley, 1983.

[5] E. Byrne and P. Fitzpatrick. Gröbner bases over Galois rings with an application to decoding alternant codes. *J. Symbolic Comput.*, 31:565–584, 2001.

[6] E. Byrne and P. Fitzpatrick. Hamming metric decoding of alternant codes over Galois rings. *IEEE Trans. Inf. Th.*, 48:683–694, 2002.

[7] P. Fitzpatrick. On the key equation. *IEEE Trans. Inf. Th*, 41:1290–1302, 1995.

[8] P. Fitzpatrick and S. Jennings. Comparison of two algorithms for decoding BCH codes. In *Proceedings 1997 IEEE International Symposium on Information Theory (ISIT'97, Ulm, Germany)*, page 325, 1997.

[9] G.D. Forney. Convolutional codes I: Algebraic structure. *IEEE Trans. Inf. Th*, 16:720–738, 1970. (note: correction in vol. IT-17, p.360, 1971).

[10] G.D. Forney, Jr. Minimal bases of rational vector spaces, with applications to multivariable linear systems. *SIAM J. Control*, 13:493–520, 1975.

[11] E.V. Gorbatov. Standard basis of a polynomial ideal over commutative Artinian chain ring. *Discrete Math. Appl.*, 14:75–101, 2004.

[12] E.V. Gorbatov. Standard basis concordant with the norm and computations in ideals and polylinear recurring sequences. *J. Math. Sci.*, 139:6672–6707, 2006.

[13] J.C. Interlando, R. Palazzo, and M. Elia. On the decoding of Reed-solomon and BCH codes over integer residue rings. *IEEE Trans. Inf. Th.*, IT-43:1013–1021, 1997.

[14] M. Kuijper and R. Pinto. Parametrization of linear recurrence relations by row reduction for sequences over a finite ring. In *Proc. 18th International Symposium on Mathematical Theory of Networks and Systems (MTNS)*, pages 1–12, Virginia Tech, USA, July 2008.

[15] M. Kuijper, R. Pinto, and J. W. Polderman. The predictable degree property and row reducedness for systems over a finite ring. *Linear Algebra and its Applications*, 425:776–796, 2007.

[16] M. Kuijper and K. Schindelar. The predictable leading monomial property for polynomial vectors over a ring. In *Proceedings 2010 IEEE International Symposium in Information Theory (ISIT)*, pages 1133–1137, Austin, Texas, USA, 2010.

[17] M. Kuijper and K. Schindelar. Minimal Gröbner bases and the predictable leading monomial property. *Linear Alg. Appl.*, 434:104–116, 2011.

[18] M. Kuijper and J.C. Willems. On constructing a shortest linear recurrence relation. *IEEE Trans. Aut. Control*, 42:1554–1558, 1997.

[19] M. Kuijper, X. Wu, and U. Parampalli. Behavioral models over rings—minimal representations and applications to coding and sequences. In *Proceedings of the 16th IFAC World Congress*, pages 1–6, Prague, Czech Republic, July 4-8, 2005.

[20] V.L. Kurakin. The Berlekamp-Massey algorithm over finite rings, modules, and bimodules. *Discrete Math. Appl.*, 8:441–474, 1998.

[21] V.L. Kurakin, A.S. Kuzmin, A.V. Mikhalev, and A.A. Nechaev. Linear recurring sequences over rings and modules. *J. Math. Sci.*, 76:2793–2915, 1995.

[22] K. Lee and M.E. O'Sullivan. List decoding of Reed-Solomon codes from a Gröbner basis perspective. *J. Symbolic Comput.*, 43:645–658, 2008.

[23] J.L. Massey. Shift-register synthesis and BCH decoding. *IEEE Trans. Inf. Th.*, IT-15:122–127, 1969.

[24] A.V. Michalev and A.A. Nechaev. Linear recurring sequences over modules. *Acta Applicandae Mathematicae*, 42:161–202, 1996.

[25] A.A. Nechaev. Linear recurring sequences over commutative rings. *Discrete Math. Appl.*, 2:659–683, 1992.

[26] A.A. Nechaev and D.A. Mikhailov. Canonical generating system of a monic polynomial ideal over a commutative artinian chain ring. *Discrete Math. Appl.*, 11:545–586, 2001.

[27] G. Norton. On minimal realization over a finite chain ring. *Designs, Codes and Cryptography*, 16:161–178, 1999.

[28] G. Norton. Minimal polynomial algorithms for finite sequences. *IEEE Trans. Inf. Th.*, 56:4643–4645, 2010.

[29] G. Norton and A. Salagean. Cyclic codes and minimal strong Gröbner bases over a principal ideal ring. *Finite Fields and their Applications*, 9:237–249, 2003.

[30] J.A. Reeds and N.J.A. Sloane. Shift-register synthesis (modulo m). *SIAM J. Computing*, 14:505–513, 1985.

[31] R.A. Rueppel. *Analysis and Design of Stream Cyphers*. Springer-Verlag, New York NY, USA, 1986.

[32] A. Salagean. An algorithm for computing minimal bidirectional linear recurrence relations. *IEEE Trans. Inf. Th.*, 55:4695–4700, 2009. ; correction, vol. IT-56, p.4180, 2010.

[33] I.E. Shparlinski. *Cryptographic applications of analytic number theory: complexity lower bounds and pseudorandomness*, volume 22. Birkhäuser, 2013.

[34] V.V. Vazirani, H. Saran, and B.S. Rajan. An efficient algorithm for constructing minimal trellises for codes over finite abelian groups. *IEEE Trans. Inf. Th.*, 42:1839–1854, 1996.

[35] Y. Wu. New list decoding algorithms for Reed-Solomon and BCH codes. *IEEE Trans. Inf. Th.*, 54:3611–3630, 2008.