

# Contents

<b>1</b>	<b>Introduction</b>	<b>2</b>
1.1	Basic Operations . . . . .	2
1.2	Sparse-FGLM Algorithm . . . . .	2
<b>2</b>	<b>Previous Results on Generating Series</b>	<b>4</b>
2.1	Overview . . . . .	4
2.1.1	Basic facts on linearly recurrent sequences. . . . .	5
2.1.2	Structure of the dual. . . . .	6
2.1.3	A fundamental formula. . . . .	8
2.2	Computing the canonical generator of a linearly recurrent matrix sequence . . . . .	13
<b>3</b>	<b>Block Sparse-FGLM algorithm</b>	<b>17</b>
3.1	Proof of Correctness . . . . .	19
3.2	Example . . . . .	23
<b>4</b>	<b>Experimental Results</b>	<b>24</b>
<b>5</b>	<b>Improvements for Special cases</b>	<b>25</b>
5.1	Choice of the blocking size . . . . .	25
5.2	Decomposition of linear forms. . . . .	26

# 1 Introduction

Computing the Gröbner basis of an ideal with respect to a term ordering is an essential step in solving systems of polynomials. Certain term orderings, such as the degree reverse lexicographical ordering (*degrevlex*), the computation of the Gröbner basis faster, while other orderings, such as the lexicographical ordering (*lex*), make it easier to find the coordinates of the solutions. In particular, for a radical ideal with a finite set of points in its variety and in generic position, its lex Gröbner basis has the form

$$\{x_1 - R_1(x_n), x_2 - R_2(x_n), \dots, x_{n-1} - R_{n-1}(x_n), R_n(x_n)\}$$

where  $R_i$  is a polynomial in only  $x_n$ . The points in the variety are

$$\{(R_1(\alpha), R_2(\alpha), \dots, R_n(\alpha)) \mid \forall \alpha : R_n(\alpha) = 0\}$$

Thus, one typically first computes a Gröbner basis for the degrevlex ordering, and then converts it to a lex Gröbner basis, or a related representation, such as Rouillier's Rational Univariate Representation [14]. One such algorithm is the Sparse-FGLM, which can be seen as an application of the Wiedemann algorithm. In this document, we will introduce a new algorithm inspired by Sparse-FGLM and block Wiedemann algorithm that is easily parallelizable.

## 1.1 Basic Operations

Over a field  $\mathbb{K}$ , recall that we can compute multiplication, division with remainder, extended GCD, and square free part of polynomial of degree at most  $n$  in  $\tilde{O}(n)$  field operations ( $\tilde{O}$  omits polylogarithmic factors) [5].

## 1.2 Sparse-FGLM Algorithm

The Sparse-FGLM algorithm [4] computes the lex Gröbner basis of an ideal with runtime cubic in the dimension of the monomial basis. More precisely, given

$$\begin{array}{ll} I \subset \mathbb{K}: & \text{zero dimensional radical ideal in shape position} \\ \mathbb{B} \subset \mathbb{K}[X_1, \dots, X_n]/I: & \text{monomial basis of } \mathbb{K}[X_1, \dots, X_n]/I \\ D: & \text{dimension of } \mathbb{B} \\ \mathbf{M}_1, \dots, \mathbf{M}_n: & \text{multiplication matrices of } X_1, \dots, X_n \text{ respectively} \end{array}$$

it produces the lex Gröbner basis of  $I$  of the form  $\{P(X_1), X_2 - R_2(X_1), \dots, X_n - R_n(X_1)\}$ . The key idea is that  $P(X_1)$  is the minimal polynomial of the multiplication matrix  $\mathbf{M}_1$ , which we can find using the Wiedemann algorithm. We generate

$$S = (u^{tr} \mathbf{M}_1^i e)_{(0 \leq i < 2D)}$$

where  $u$  is a random vector and  $e$  is the coordinate vector for 1 in  $\mathbb{B}$ . Then, we find the minimum generating polynomial  $P$  by applying the Berlekamp-Massey algorithm on  $S$ . We

compute the numerator  $n$  of the generating series  $Z = \sum_{i=0}^{D-1} u^{tr} \mathbf{M}_1^i e / T^{i+1}$  by a product  $N = PZ$ . To find  $R_j(x_1)$ ,  $2 \leq j \leq n$ , we first compute the numerator  $N_j$  of the generating series  $Z_j = \sum_{i=0}^D u^{tr} \mathbf{M}_1^i \mathbf{M}_j e / T^{i+1}$  by a product  $N_j = PZ_j$ . Finally,  $R_j(x_1) = \frac{N_j}{N} \bmod P$ . More general versions of the Sparse-FGLM algorithm exist to handle non-radical ideals using the Berlekamp-Massey-Sakata algorithm [4].

## 2 Previous Results on Generating Series

### 2.1 Overview

In what follows,  $\mathbb{K}$  is a field. Let  $I$  be an ideal in  $\mathbb{K}[X_1, \dots, X_n]$  and  $Q = \mathbb{K}[X_1, \dots, X_n]/I$  be the associated residue class ring. Suppose that  $V = V(I)$  has dimension zero, and write it as  $V = \{\alpha_1, \dots, \alpha_d\}$ , with all  $\alpha_i$ 's in  $\overline{\mathbb{K}}^n$ , and  $\alpha_i = (\alpha_{i,1}, \dots, \alpha_{i,n})$  for all  $i$ . We also let  $D$  be the dimension of  $Q$ , so that  $d \leq D$ , and *we assume that  $\text{char}(\mathbb{K})$  is greater than  $D$* . In this section, we recall and generalize results from the appendix of [2], with the objective of computing a zero-dimensional parametrization of  $V$ .

The main novelty in our approach is to avoid generic coordinates as much as possible. The algorithm will decompose  $V$  into two parts: for the first part, we will be able to use  $X_1$  as a separating element; the remaining points will be dealt with using a random linear form. Throughout, we rely only the following operations: evaluations of linear forms on successive powers of a given element in  $Q$ , say  $1, t, t^2, \dots$ , and elementary operations on univariate polynomials.

The main algorithm is as follows. For the moment, we can only describe its main structure; the details of the subroutines are given in the next paragraphs.

---

#### Algorithm 1 Parametrization( $\ell, t$ )

---

**Input:**

- a linear form  $\ell$  over  $Q$
- $t = t_1 X_1 + \dots + t_n X_n$

**Output:**

- polynomials  $(P, (V_1, \dots, V_n))$
1. let  $(F, (G_1, \dots, G_n), X_1) = \text{ParametrizationX}_1(\ell)$
  2. let  $\ell' = \text{Update}(\ell, F, t)$
  3. let  $(Q, (W_1, \dots, W_n), t) = \text{ParametrizationGeneric}(\ell', t)$
  4. let  $(F^*, (G_1^*, \dots, G_n^*), t) = \text{ChangeCoordinate}(F, (G_1, \dots, G_n), t)$
  5. let  $(P, (V_1, \dots, V_n)) = \text{Union}(F^*, (G_1^*, \dots, G_n^*), Q, (W_1, \dots, W_n))$
  6. **return**  $(P, (V_1, \dots, V_n), t)$
- 

The call to  $\text{ParametrizationX}_1(\ell)$  computes a zero-dimensional parametrization of a subset  $V'$  of  $V$  for which  $X_1$  is a separating element, using values of the form  $(\ell(X_1^s))_{s \geq 0}$ . We then modify  $\ell$  (which in effect removes from  $V$  the points we just found) and apply  $\text{ParametrizationGeneric}(\ell', t)$ , to obtain a zero-dimensional parametrization of  $V'' = V - V'$  using values of the form  $(\ell'(t^s))_{s \geq 0}$ . The last two steps involve changing coordinates in  $(F, (G_1, \dots, G_n))$  (to use  $t$  as a separating variable instead), and performing the union of the two components  $V'$  and  $V''$ .

### 2.1.1 Basic facts on linearly recurrent sequences.

Consider a sequence  $(\ell_s)_{s \geq 0} \in \mathbb{K}^{\mathbb{N}}$  and the associated generating series  $Z = \sum_{s \geq 0} \ell_s T^s \in \mathbb{K}[[T]]$ . The sequence  $(\ell_s)_{s \geq 0}$  is linearly recurrent if and only if its generating series is *rational* that is, if there exist polynomials  $A, B$  in  $\mathbb{K}[T]$  such that  $Z = A/B$ ; these polynomials are unique if we assume  $\gcd(A, B) = 1$  and  $B(0) = 1$ .

We say that a degree  $m$  polynomial  $P \in \mathbb{K}[T]$  *cancels* a sequence  $(\ell_s)_{s \geq 0}$  if  $p_0 \ell_s + \dots + p_m \ell_{s+m} = 0$  for all  $s \geq 0$ , where  $p_0, \dots, p_m$  are the coefficients of  $P$ ; this is equivalent to  $\text{rev}(P)Z$  being a polynomial of degree less than  $m$ , with  $Z = \sum_{s \geq 0} \ell_s T^s$  and  $\text{rev}(P) = T^m P(1/T)$ . The *minimal polynomial* of sequence  $(\ell_s)_{s \geq 0}$  is the monic polynomial of lowest degree that cancels this sequence.

Given the closed form  $Z = A/B$  as above, we define  $\tilde{B} = T^{\max(\deg(A)+1, \deg(B))} B(1/T)$ . By construction,  $\tilde{B}$  lies in  $\mathbb{K}[T]$ , and one can check that it is the minimal polynomial of sequence  $(\ell_s)_{s \geq 0}$ . However, it is often easier to work with a closed form that has the minimal polynomial as the denominator rather than its reverse. Let  $Z' = \sum_{s \geq 0} \ell_s / T^{i+1}$  and  $P$  be a polynomial that cancels  $(\ell_s)_{s \geq 0}$ , then the sequence  $(\ell_s)_{s \geq 0}$  is linearly recurrent if and only if  $Z'P$  is a polynomial of degree less than  $\deg(P)$ . Therefore, if  $A' = Z'P$  and  $B' = P$ , then we can write the closed form  $Z' = A'/B'$ .

In both cases, if we are given the series and the denominator, then our choice for the numerator is unique. The next definition captures this idea:

**Definition 2.1.** Let  $S = (S_k)_{k \geq 0}$  be a sequence and  $P$  be a polynomial that cancels  $S$ , then  $\Omega(S, P)$  is the numerator for the closed form of  $S$  with respect to  $P$  if:

- $\Omega(S, P)$  is a polynomial of degree less than  $\deg(P)$
- $\sum_{k \geq 0} S_k / T^{k+1} = \Omega(S, P) / P$

Let  $A$  and  $A'$  be defined as above, then

$$A' = \Omega((\ell)_{s \geq 0}, P) = T^{\deg(P)-1} A(1/T)$$

This makes it possible to avoid computations using a series over  $1/T$ . More precisely, given a sequence  $(\ell_s)_{s \geq 0}$  and polynomial  $P$  that cancels it, we compute  $\Omega((\ell)_{s \geq 0}, P)$  by doing the following:

$$\begin{aligned} Z &= \sum_{s=0}^{\deg(P)} \ell_{(\deg(P)-s)} T^s \\ N &= ZP \\ \Omega((\ell)_{s \geq 0}, P) &= \sum_{s=0}^{\deg(P)-1} ([T^{s+\deg(P)}]N) T^s, \end{aligned}$$

where  $[T^{s+\deg(P)}]N$  is the coefficient of  $[T^{s+\deg(P)}]$  in  $N$ .

We give an example of this using the Fibonacci sequence, which is well known to be linearly recurrent. Since the Fibonacci sequence  $F = (1, 1, 2, 3, 5, 8, \dots)$  is defined recursively

as  $F_{s+2} = F_s + F_{s+1}$ ,  $F_0 = F_1 = 1$ , it is easy to verify that the minimal polynomial of  $F$  is  $P = T^2 - T - 1$ . Define  $Z = \sum_{s \geq 0} F_s T^s$  and  $\text{rev}(P) = 1 - T - T^2$ , then we can write  $Z$  in its closed form:

$$Z = \frac{A}{B} = \frac{1}{1 - T - T^2}$$

Now, define  $Z' = \sum_{s \geq 0} F_s / T^{s+1}$ , then we have that

$$\begin{aligned} \Omega(S, P) &= T^{\deg(P)-1} A(1/T) = T \\ \implies Z' &= \frac{T}{T^2 - T - 1} \end{aligned}$$

The sequences we consider below are of the form  $(\ell(t^s))_{s \geq 0}$ , for  $\ell$  a  $\mathbb{K}$ -linear form  $Q \rightarrow \mathbb{K}$  and  $t$  in  $Q$ . For such sequences, the following standard result will be useful.

**Lemma 2.2.** *Let  $t$  be in  $Q$  and let  $P \in \mathbb{K}[T]$  be its minimal polynomial. For a generic choice of  $\ell$  in  $\text{hom}_{\mathbb{K}}(Q, \mathbb{K})$ ,  $P$  is the minimal polynomial of the sequence  $(\ell(t^s))_{s \geq 0}$ .*

### 2.1.2 Structure of the dual.

For  $i$  in  $\{1, \dots, d\}$ , let  $Q_i$  be the local algebra at  $\alpha_i$ , that is  $Q_i = \overline{\mathbb{K}}[X_1, \dots, X_n]/I_i$ , with  $I_i$  the  $\mathfrak{m}_{\alpha_i}$ -primary component of  $I$ . By the Chinese Remainder Theorem,  $Q \otimes_{\mathbb{K}} \overline{\mathbb{K}} = \overline{\mathbb{K}}[X_1, \dots, X_n]/I$  is isomorphic to the direct product  $Q_1 \times \dots \times Q_d$ . We let  $N_i$  be the *nil-index* of  $Q_i$ , that is, the maximal integer  $N$  such that  $\mathfrak{m}_{\alpha_i}^N$  is not contained in  $I_i$ ; for instance,  $N_i = 0$  if and only if  $Q_i$  is a field, if and only if  $\alpha_i$  is a non-singular root of  $I$ . We also let  $D_i = \dim_{\overline{\mathbb{K}}}(Q_i)$ , so that we have  $D_i \geq N_i$  and  $D = D_1 + \dots + D_d$ .

Fix  $i$  in  $1, \dots, d$ . There exists a basis of the dual  $\text{hom}_{\overline{\mathbb{K}}}(Q_i, \overline{\mathbb{K}})$  consisting of linear forms  $(\lambda_{i,j})_{1 \leq j \leq D_i}$  of the form

$$\lambda_{i,j} : f \mapsto (\Lambda_{i,j}(f))(\alpha_i),$$

where  $\Lambda_{i,j}$  is the operator

$$f \mapsto \Lambda_{i,j}(f) = \sum_{\mu=(\mu_1, \dots, \mu_n) \in S_{i,j}} c_{i,j,\mu} \frac{\partial^{\mu_1 + \dots + \mu_n} f}{\partial X_1^{\mu_1} \dots \partial X_n^{\mu_n}},$$

for some finite subset  $S_{i,j}$  of  $\mathbb{N}^n$  and non-zero constants  $c_{i,j,\mu}$  in  $\overline{\mathbb{K}}$ . For instance, when  $\alpha_i$  is non-singular, we have  $D_i = 1$ , so there is only one function  $\lambda_{i,j}$ , namely  $\lambda_{i,1}$ , we write it  $\lambda_{i,1}(f) = f(\alpha_i)$ .

More generally, we can always take  $\lambda_{i,1}$  of the form  $\lambda_{i,1}(f) = f(\alpha_i)$ ; for  $j > 1$ , we can then also assume that  $S_{i,j}$  does not contain  $\mu = (0, \dots, 0)$  (that is, all terms in  $\Lambda_{i,j}$  have order 1 or more). Thus, introducing new variables  $(U_{i,j})_{j=1, \dots, D_i}$ , we deduce the existence of non-zero homogeneous linear forms  $P_{i,\mu}$  in  $(U_{i,j})_{j=1, \dots, D_i}$  such that for any  $\lambda$  in  $\text{hom}_{\overline{\mathbb{K}}}(Q_i, \overline{\mathbb{K}})$ ,

there exist  $\mathbf{u}_i = (u_{i,j}) \in \overline{\mathbb{K}}^{D_i}$  such that we have

$$\begin{aligned}
\lambda : f &\mapsto \lambda(f) = \sum_{j=1}^{D_i} u_{i,j} \lambda_{i,j}(f) \\
&= \sum_{j=1}^{D_i} u_{i,j} (\Lambda_{i,j}(f))(\boldsymbol{\alpha}_i) \\
&= \sum_{j=1}^{D_i} u_{i,j} \sum_{\mu=(\mu_1, \dots, \mu_n) \in S_{i,j}} c_{i,j,\mu} \frac{\partial^{\mu_1 + \dots + \mu_n} f}{\partial X_1^{\mu_1} \dots \partial X_n^{\mu_n}}(\boldsymbol{\alpha}_i) \\
&= \sum_{\mu=(\mu_1, \dots, \mu_n) \in S_i} P_{i,\mu}(\mathbf{u}_i) \frac{\partial^{\mu_1 + \dots + \mu_n} f}{\partial X_1^{\mu_1} \dots \partial X_n^{\mu_n}}(\boldsymbol{\alpha}_i), \tag{1}
\end{aligned}$$

where  $S_i$  is the union of  $S_{i,1}, \dots, S_{i,D_i}$ , with in particular  $P_{i,(0,\dots,0)} = u_{i,1}$  and where  $P_{i,\mu}$  depends only on  $(u_{i,j})_{j=2,\dots,D_i}$  for all  $\mu$  in  $S_i$ ,  $\mu \neq (0, \dots, 0)$ . Explicitly, we can write  $P_{i,\mu} = \sum_{j \in \{1, \dots, D_i\}} c_{i,j,\mu} U_{i,j}$ .

Fix  $\lambda$  non-zero in  $\text{hom}_{\overline{\mathbb{K}}}(Q_i, \overline{\mathbb{K}})$ . We can then define its *order*  $w$  and *symbol*  $\pi$ . The former is the maximum of all  $|\mu| = \mu_1 + \dots + \mu_n$  for  $\mu = (\mu_1, \dots, \mu_n)$  in  $S_i$  such that  $P_{i,\mu}(\mathbf{u}_i)$  is non-zero; by [12, Lemma 3.3] we have  $w \leq N_i - 1$ . Then, we let

$$\pi = \sum_{\mu \in S_i, |\mu|=w} P_{i,\mu}(\mathbf{u}_i) X_1^{\mu_1} \dots X_n^{\mu_n}$$

be the *symbol* of  $\lambda$ ; by construction, this is a non-zero polynomial. In the following paragraphs, we will need the next easy lemma.

**Lemma 2.3.** *Fix  $i$  in  $\{1, \dots, d\}$ . For a generic choice of  $\lambda$  in  $\text{hom}_{\overline{\mathbb{K}}}(Q_i, \overline{\mathbb{K}})$ , and of  $t_1, \dots, t_n$  in  $\overline{\mathbb{K}}^n$ ,  $\pi_i(t_1, \dots, t_n)$  is non-zero.*

*Proof.* Let  $\Omega$  be the maximum of all  $|\mu| = \mu_1 + \dots + \mu_n$  for  $\mu = (\mu_1, \dots, \mu_n)$  in  $S_i$ , and define

$$\Pi = \sum_{\mu \in S_i, |\mu|=\Omega} P_{i,\mu} X_1^{\mu_1} \dots X_n^{\mu_n} \in \overline{\mathbb{K}}[U_{i,1}, \dots, U_{i,D_i}, X_1, \dots, X_n];$$

this is by construction a non-zero polynomial. Thus, for a generic choice of  $\mathbf{u}_i = (u_{i,1}, \dots, u_{i,D_i})$ , that define a linear form  $\lambda$  in  $\text{hom}_{\overline{\mathbb{K}}}(Q_i, \overline{\mathbb{K}})$  as in Eq. (1), and of  $t_1, \dots, t_n$  in  $\overline{\mathbb{K}}^n$ , the value  $\Pi(u_{i,1}, \dots, u_{i,D_i}, t_1, \dots, t_n)$  is non-zero. As a result, the symbol of such a linear form  $\lambda$  is  $\pi = \sum_{\mu \in S_i, |\mu|=\Omega} P_{i,\mu}(\mathbf{u}_i) X_1^{\mu_1} \dots X_n^{\mu_n}$ , and  $\pi(t_1, \dots, t_n)$  is then non-zero.  $\square$

Finally, we say a word about global objects. Fix a linear form  $\ell : Q \rightarrow \mathbb{K}$ . By the Chinese Remainder Theorem, there exist unique  $\ell_1, \dots, \ell_d$ , with  $\ell_i$  in  $\text{hom}_{\overline{\mathbb{K}}}(Q_i, \overline{\mathbb{K}})$  for all  $i$ , such that the extension  $\ell_{\overline{\mathbb{K}}} : Q \otimes_{\mathbb{K}} \overline{\mathbb{K}} \rightarrow \overline{\mathbb{K}}$  decomposes as  $\ell_{\overline{\mathbb{K}}} = \ell_1 + \dots + \ell_d$ . We call *support* of  $\ell$  the

subset  $\mathfrak{S}$  of  $\{1, \dots, d\}$  such that  $\ell_i$  is non-zero exactly for  $i$  in  $\mathfrak{S}$ . As a consequence, for all  $f$  in  $Q$ , we have

$$\begin{aligned}\ell(f) &= \ell_1(f) + \dots + \ell_d(f) \\ &= \sum_{i \in \mathfrak{S}} \ell_i(f).\end{aligned}\tag{2}$$

For  $i$  in  $\mathfrak{S}$ , we denote by  $w_i$  and  $\pi_i$  respectively the order and the symbol of  $\ell_i$ . For such a subset  $\mathfrak{S}$ , we also write  $Q_{\mathfrak{S}} = \prod_{i \in \mathfrak{S}} Q_i$  and  $V_{\mathfrak{S}} = \cup_{i \in \mathfrak{S}} \{\alpha_i\}$ .

### 2.1.3 A fundamental formula.

The following lemma gives an explicit form for a generating series of the form  $\sum_{\ell \geq 0} \ell(vt^\ell)T^\ell$ , for a linear form  $\ell : Q \rightarrow \mathbb{K}$ . A slightly less precise version of it is in [2].

**Lemma 2.4.** *Let  $\ell$  be in  $\text{hom}_{\mathbb{K}}(Q, \mathbb{K})$ , with support  $\mathfrak{S}$ , and let  $\{\pi_i \mid i \in \mathfrak{S}\}$  and  $\{w_i \mid i \in \mathfrak{S}\}$  be as above.*

*Let  $t = t_1X_1 + \dots + t_nX_n$ , for some  $t_1, \dots, t_n$  in  $\mathbb{K}$  and let  $v$  be in  $\mathbb{K}[X_1, \dots, X_n]$ . Then, we have the equality*

$$\sum_{s \geq 0} \ell(vt^s)/T^{s+1} = \sum_{i \in \mathfrak{S}} \frac{v(\alpha_i) w_i! \pi_i(t_1, \dots, t_n) + (T - t(\alpha_i))A_{v,i}}{(T - t(\alpha_i))^{w_i+1}},\tag{3}$$

for some polynomials  $\{A_{v,i} \in \overline{\mathbb{K}}[T] \mid i \in \mathfrak{S}\}$  (that depend on the choice of  $v$ ), with  $A_{v,i}$  of degree less than  $w_i$  for all  $i$  in  $\mathfrak{S}$ .

*Proof.* Take  $v$  and  $t$  as above. Consider first an operator of the form  $f \mapsto \frac{\partial^{|\mu|} f}{\partial X_1^{\mu_1} \dots \partial X_n^{\mu_n}}$ , where we write  $|\mu| = \mu_1 + \dots + \mu_n$ . Then, we have the following generating series identities, with coefficients in  $\mathbb{K}(X_1, \dots, X_n)$ :

$$\begin{aligned}\sum_{s \geq 0} \frac{\partial^{|\mu|}(vt^s)}{\partial X_1^{\mu_1} \dots \partial X_n^{\mu_n}} \frac{1}{T^{s+1}} &= \sum_{s \geq 0} \frac{\partial^{|\mu|}(vt^s/T^{s+1})}{\partial X_1^{\mu_1} \dots \partial X_n^{\mu_n}} \\ &= \frac{\partial^{|\mu|}}{\partial X_1^{\mu_1} \dots \partial X_n^{\mu_n}} \left( \sum_{s \geq 0} vt^s/T^{s+1} \right) \\ &= \frac{\partial^{|\mu|}}{\partial X_1^{\mu_1} \dots \partial X_n^{\mu_n}} \left( \frac{v}{T-t} \right) \\ &= v |\mu|! \prod_{1 \leq k \leq n} \left( \frac{\partial t}{\partial X_k} \right)^{\mu_k} \frac{1}{(T-t)^{|\mu|+1}} + \frac{P_{|\mu|}(\mathbf{X}, T)}{(T-t)^{|\mu|}} + \dots + \frac{P_1(\mathbf{X}, T)}{(T-t)} \\ &= v |\mu|! \prod_{1 \leq k \leq n} t_k^{\mu_k} \frac{1}{(T-t)^{|\mu|+1}} + \frac{P(\mathbf{X}, T)}{(T-t)^{|\mu|}},\end{aligned}$$



for some polynomials  $P_1, \dots, P_{|\mu|}, P$  in  $\mathbb{K}[\mathbf{X}, T]$  that depend on the choices of  $\mu$ ,  $v$  and  $t$ , with  $\deg(P_i, T) < i$  for all  $i$  and thus  $\deg(P, T) < |\mu|$ .

Take now a linear combination of such operators, such as  $f \mapsto \sum_{\mu \in R} c_\mu \frac{\partial^{|\mu|} f}{\partial X_1^{\mu_1} \dots \partial X_n^{\mu_n}}$ . The corresponding generating series becomes

$$\sum_{s \geq 0} \sum_{\mu \in R} c_\mu \frac{\partial^{|\mu|} (vt^s)}{\partial X_1^{\mu_1} \dots \partial X_n^{\mu_n}} \frac{1}{T^{s+1}} = v \sum_{\mu \in R} c_\mu |\mu|! \prod_{1 \leq k \leq n} t_k^{\mu_k} \frac{1}{(T-t)^{|\mu|+1}} + \sum_{\mu \in R} \frac{P_\mu(\mathbf{X}, T)}{(T-t)^{|\mu|}},$$

where each  $P_\mu$  has degree less than  $|\mu|$  in  $T$ . Let  $w$  be the maximum of all  $|\mu|$  for  $\mu$  in  $R$ . We can rewrite the above as

$$v w! \sum_{\mu \in R, |\mu|=w} c_\mu \prod_{1 \leq k \leq n} t_k^{\mu_k} \frac{1}{(T-t)^{w+1}} + \frac{A(\mathbf{X}, T)}{(T-t)^w},$$

for some polynomial  $A$  of degree less than  $w$  in  $T$ . If we let  $\pi = \sum_{\mu \in R, |\mu|=w} c_\mu X_1^{\mu_1} \dots X_n^{\mu_n}$ , this becomes

$$\sum_{s \geq 0} \sum_{\mu \in R} c_\mu \frac{\partial^{|\mu|} (vt^s)}{X_1^{\mu_1} \dots X_n^{\mu_n}} \frac{1}{T^{s+1}} = v w! \pi(t_1, \dots, t_n) \frac{1}{(T-t)^{w+1}} + \frac{A(\mathbf{X}, T)}{(T-t)^w}.$$

Applying this formula to the sum in Eq. (2), we obtain the claim in the lemma.  $\square$

The most useful consequence of the previous lemma is the following interpolation formula, where we fix a subset  $\mathfrak{S}$  of  $\{1, \dots, d\}$ . The mapping  $t : V_{\mathfrak{S}} \rightarrow \overline{\mathbb{K}}$  defined by  $\alpha_i \mapsto t(\alpha_i)$  plays a special role in the formula in the lemma; this leads us to the following definitions.

- We consider  $\ell$  and  $t$  as in Lemma 2.4, such that  $\ell$  has support  $\mathfrak{S}$ .
- $\mathfrak{T}$  is the subset of  $\mathfrak{S}$  consisting of all indices  $i$  such that
  - $\pi_i(t_1, \dots, t_n)$  is non-zero;
  - $t(\alpha_{i'}) \neq t(\alpha_i)$  for  $i' \neq i$  in  $\mathfrak{S}$ .
- $\{r_1, \dots, r_c\}$  are the pairwise distinct values taken by  $t$  on  $V_{\mathfrak{S}}$ , for some  $c \leq |\mathfrak{S}|$ .
- $\mathfrak{t}$  is the set of all indices  $j$  in  $\{1, \dots, c\}$  such that
  - the fiber  $t^{-1}(r_j) \subset V_{\mathfrak{S}}$  contains a single point, written  $\alpha_{\sigma_j}$ ;
  - the point  $\alpha_{\sigma_j}$  is in  $\mathfrak{T}$  (equivalently,  $\pi_{\sigma_j}(t_1, \dots, t_n)$  is non-zero).

Remark that  $j \mapsto \sigma_j$  induces a one-to-one correspondence between  $\mathfrak{t}$  and  $\mathfrak{T}$ .

**Lemma 2.5.** *Let  $\ell$ ,  $t$  and all other notation be as above. Let further  $M$  be the minimal polynomial of  $t$  in  $Q_{\mathfrak{S}}$ . Suppose that  $M$  is also the minimal polynomial of the sequence  $(\ell(t^s))_{s \geq 0}$ , then there exist non-zero constants  $\{c_j \mid j \in \mathfrak{t}\}$  such that for  $v$  in  $\mathbb{K}[X_1, \dots, X_n]$ ,*

$$\Omega((\ell(vt^s))_{s \geq 0}, M) = c_j v(\alpha_{\sigma_j}) \quad \text{for all } j \text{ in } \mathfrak{t}.$$

*Proof.* For  $j = 1, \dots, c$ , we write  $T_j$  for the set of all indices  $i$  in  $\mathfrak{S}$  such that  $t(\alpha_i) = r_j$ ; the sets  $T_1, \dots, T_c$  form a partition of  $\mathfrak{S}$ . When  $T_j$  has cardinality 1, we thus have  $T_j = \{\sigma_j\}$ .

Take an arbitrary  $v$  in  $\mathbb{K}[X_1, \dots, X_n]$  and let us collect terms in Eq. (3) as

$$\begin{aligned} \sum_{s \geq 0} \ell(vt^s) \frac{1}{T^{s+1}} &= \sum_{j \in \{1, \dots, c\}} \sum_{i \in T_j} \frac{v(\alpha_i) w_i! \pi_i(t_1, \dots, t_n) + (T - r_j) A_{v,i}}{(T - r_j)^{w_i+1}} \\ &= \sum_{j \in \mathfrak{s}} \frac{v(\alpha_{\sigma_j}) w_{\sigma_j}! \pi_{\sigma_j}(t_1, \dots, t_n) + (T - r_j) A_{v, \sigma_j}}{(T - r_j)^{w_{\sigma_j}+1}} \\ &\quad + \sum_{j \in \{1, \dots, c\} - \mathfrak{s}} \frac{\sum_{i \in T_j} \left( [v(\alpha_i) w_i! \pi_i(t_1, \dots, t_n) + (T - r_j) A_{v,i}] (T - r_j)^{y_j - (w_i+1)} \right)}{(T - r_j)^{y_j}}, \end{aligned}$$

where  $y_j$  is the maximum of all  $w_i$  for  $i$  in  $T_j$ . Remark that for  $v = 1$ , our condition that  $\pi_i$  is non-zero for  $i$  in  $\mathfrak{T}$  implies that in the second line, together with our assumption on the characteristic of  $\mathbb{K}$ , imply that all terms in the first sum are non-zero and in reduced form.

After simplifying terms in the second sum, we can rewrite the expression above as

$$\sum_{s \geq 0} \ell(vt^s) \frac{1}{T^{s+1}} = \sum_{j \in \mathfrak{t}} \frac{v(\alpha_{\sigma_j}) w_{\sigma_j}! \pi_{\sigma_j}(t_1, \dots, t_n) + (T - r_j) A_{v, \sigma_j}}{(T - r_j)^{w_{\sigma_j}+1}} + \sum_{j \in \{1, \dots, c\} - \mathfrak{s}} \frac{D_{v,j}}{(T - r_j)^{z_{v,j}}},$$

for some positive integers  $\{z_{v,j} \mid j \in \{1, \dots, c\} - \mathfrak{s}\}$  and polynomials  $\{D_{v,j} \mid j \in \{1, \dots, c\} - \mathfrak{s}\}$  such that for all  $j$  in  $\{1, \dots, c\} - \mathfrak{s}$ , we have  $\deg(D_{v,j}) < z_{v,j}$  and  $\gcd(D_{v,j}, (T - r_j)) = 1$ ; the integers  $z_{v,j}$  are uniquely determined by these conditions, except if  $r_j = 0$ , in which case we set  $z_{v,j} = \deg(D_{v,j}) + 1$ . Some of the polynomials  $D_{v,j}$  may vanish, so we let  $\mathfrak{u}_v \subset \{1, \dots, c\} - \mathfrak{s}$  be the set of all  $j$  for which this is not the case. We then arrive at our final form for this sum, namely

$$\sum_{s \geq 0} \ell(vt^s) \frac{1}{T^{s+1}} = \sum_{j \in \mathfrak{t}} \frac{v(\alpha_{\sigma_j}) w_{\sigma_j}! \pi_{\sigma_j}(t_1, \dots, t_n) + (T - r_j) A_{v, \sigma_j}}{(T - r_j)^{w_{\sigma_j}+1}} + \sum_{j \in \mathfrak{u}_v} \frac{D_{v,j}}{(T - r_j)^{z_{v,j}}}, \quad (4)$$

where all terms in the second sum are non-zero and in reduced form (and similarly for the first sum, for  $v = 1$ ). This implies that the minimal polynomial of the sequence  $(\ell(vt^s))_{s \geq 0}$  is

$$M_v = \prod_{j \in \mathfrak{t}} (T - r_j)^{\zeta_j} \prod_{j \in \mathfrak{u}_v} (T - r_j)^{z_{v,j}},$$

for some integers  $\{\zeta_j \leq w_{\sigma_j} + 1 \mid j \in \mathfrak{t}\}$ ; for  $v = 1$ , we actually have  $\zeta_j = w_{\sigma_j} + 1$  for all such  $j$ .

Now, for  $v = 1$ , we assume that the minimal polynomial of the sequence  $(\ell(t^s))_{s \geq 0}$  is the minimal polynomial  $M$  of  $t$  in  $Q_{\mathfrak{S}}$ . Writing  $\mathfrak{u} = \mathfrak{u}_1$  and  $z_k = z_{1,k}$  for all  $k$  in  $\mathfrak{u}$ , we can thus write it as

$$M = \prod_{j \in \mathfrak{t}} (T - r_j)^{w_{\sigma_j}+1} \prod_{j \in \mathfrak{u}} (T - r_j)^{z_j}.$$

Since it is the minimal polynomial of  $t$  in  $Q_{\mathfrak{S}}$ , it also cancels the sequence  $(\ell(vt^s))_{s \geq 0}$  for any  $v$ , so that for all  $v$ ,  $\mathfrak{u}_v$  is contained in  $\mathfrak{u}$  and  $M_v$  divides  $M$ . Remark also that the integer

$\delta = \deg(M)$  is given by

$$\delta = \sum_{j \in \mathfrak{t}} (w_{\sigma_j} + 1) + \sum_{j \in \mathfrak{u}} z_j.$$

For  $v$  arbitrary in  $\mathbb{K}[X_1, \dots, X_n]$ , since  $M$  cancels the sequence  $(\ell(vt^s))_{s \geq 0}$ ,  $\Omega((\ell(vt^s))_{s \geq 0}, M)$  in the statement of the lemma is indeed a polynomial of degree less than  $\delta$  (this proves our first claim). We can then rewrite the sum in Eq. (4) as  $\sum_{s \geq 0} \ell(vt^s)/T^{s+1} = \Omega((\ell(vt^s))_{s \geq 0}, M)/M$ , with

$$\begin{aligned} \Omega((\ell(vt^s))_{s \geq 0}, M) &= \sum_{j \in \mathfrak{t}} \left( [v(\alpha_{\sigma_j}) w_{\sigma_j}! \pi_{\sigma_j}(t_1, \dots, t_n) + (T - r_j) \tilde{A}_{v, \sigma_j}] \prod_{\iota \in \mathfrak{t} - \{j\}} (T - r_\iota)^{w_{\sigma_\iota} + 1} \right) \prod_{j \in \mathfrak{u}} (T - r_j)^{z_j} \\ &\quad + \left( \prod_{j \in \mathfrak{t}} (T - r_j)^{w_{\sigma_j} + 1} \right) \sum_{j \in \mathfrak{u}_v} \left( (T - r_j)^{z_j - z_{v,j}} D_{v,j} \prod_{\iota \in \mathfrak{u} - \{j\}} (T - r_\iota)^{z_\iota} \right), \end{aligned}$$

In particular, for  $k$  in  $\mathfrak{t}$ , the value  $\Omega((\ell(vt^s))_{s \geq 0}, M)(r_k)$  is

$$\begin{aligned} \Omega((\ell(vt^s))_{s \geq 0}, M)(r_k) &= v(\alpha_{\sigma_k}) w_{\sigma_k}! \pi_{\sigma_k}(t_1, \dots, t_n) \prod_{\iota \in \mathfrak{t} - \{k\}} (r_\iota - r_k)^{w_{\sigma_\iota} + 1} \prod_{j \in \mathfrak{u}} (r_j - r_k)^{z_k} \\ &= v(\alpha_{\sigma_k}) c_k, \end{aligned}$$

with

$$c_k = w_{\sigma_k}! \pi_{\sigma_k}(t_1, \dots, t_n) \prod_{\iota \in \mathfrak{t} - \{k\}} (r_\iota - r_k)^{w_{\sigma_\iota} + 1} \prod_{j \in \mathfrak{u}} (r_j - r_k)^{z_k}$$

for  $k$  in  $\mathfrak{t}$ . This is a non-zero constant, independent of  $v$ , which finishes the proof of the lemma.  $\square$

As an application, the following algorithm shows how to compute a zero-dimensional parametrization of  $V_{\mathfrak{S}}$ .

---

**Algorithm 2** ParametrizationGeneric( $\ell, t$ )

---

**Input:**

- a linear form  $\ell$  over  $Q_{\mathfrak{S}}$
- $t = t_1 X_1 + \dots + t_n X_n$

**Output:** polynomials  $(P, V_1, \dots, V_n)$

1. let  $M$  be the minimal polynomial of the sequence  $(\ell(t^s))_{s \geq 0}$  and let  $\delta$  be its degree
  2. let  $P$  be the squarefree part of  $M$
  3. let  $C_1 = \Omega((\ell(t^s))_{s \geq 0}, M)$
  4. **for**  $i = 1, \dots, n$  **do**
    - (a) let  $C_{X_i} = \Omega((\ell(X_i t^s))_{s \geq 0}, M)$
  5. **return**  $((P, C_{X_1}/C_1 \bmod P, \dots, C_{X_n}/C_1 \bmod P), t)$
-

**Lemma 2.6.** *Suppose that  $\ell$  is a generic element of  $\text{hom}_{\overline{\mathbb{K}}}(Q_{\mathfrak{S}}, \overline{\mathbb{K}})$  and that  $t$  is a generic linear form. Then the output  $((P, V_1, \dots, V_n), t)$  of  $\text{Parametrization}(\ell, t)$  is a zero-dimensional parametrization of  $V_{\mathfrak{S}}$ .*

*Proof.* A generic choice of  $t$  separates the points of  $V_{\mathfrak{S}}$ , and we saw in Lemma 2.3 that for a generic choice of  $\ell$ ,  $\pi_i(t_1, \dots, t_n)$  vanishes for no  $i$  in  $\mathfrak{S}$ . As a result,  $\mathfrak{T} = \mathfrak{S}$ . Besides, we recall that for a generic  $\ell$  in  $\text{hom}_{\overline{\mathbb{K}}}(Q_{\mathfrak{S}}, \overline{\mathbb{K}})$ , the minimal polynomials of  $(\ell(t^s))_{s \geq 0}$  and of  $t$  are the same (Lemma 2.2).

Thus, the polynomial  $M$  we compute at step 1 is indeed the minimal polynomial of  $t$ , and we can apply the previous lemma, and for any root  $r_j$  of  $P$ , and  $i = 1, \dots, n$ , we have

$$\frac{C_{X_i}(r_j)}{C_1(r_j)} = \frac{\Omega((\ell(X_i t^s))_{s \geq 0}, M)(r_j)}{\Omega((\ell(t^s))_{s \geq 0}, M)(r_j)} = \frac{c_j \alpha_{\sigma_j, i}}{c_j} = \alpha_{\sigma_j, i},$$

so that  $C_{X_i}/C_1 \bmod P$  is the  $i$ th polynomial in the zero-dimensional parametrization of  $V$  corresponding to  $t$ .  $\square$

We demonstrate how this algorithm works through a small example. Let  $I = \langle (X_1 - 1)(X_2 - 2), (X_1 - 3)(X_2 - 4) \rangle \subset GF(101)[X_1, X_2]$ , then clearly  $V(I) = \{(1, 4), (3, 2)\}$  and  $X_1$  separates the points of  $V$ . We choose a random linear form

$$\ell : f \in I \mapsto \mathbb{N}, \ell(f) = 17f(1, 4) + 33f(3, 2)$$

Then we have

$$\begin{aligned} \ell(X_1^i) &= 17 \cdot 1^i + 33 \cdot 3^i \\ \ell(X_2 X_1^i) &= 17 \cdot 4 \cdot 1^i + 33 \cdot 2 \cdot 3^i \end{aligned}$$

We define an infinite series for each sequence

$$\begin{aligned} Z_1 &= \sum_{i=0}^{\infty} \ell(X_1^i) / T^{i+1} = \frac{17}{T-1} + \frac{33}{T-3} = \frac{17(T-3) + 33(T-1)}{(T-1)(T-3)} \\ Z_2 &= \sum_{i=0}^{\infty} \ell(X_2 X_1^i) / T^{i+1} = \frac{17 \cdot 4}{T-1} + \frac{33 \cdot 2}{T-3} = \frac{17 \cdot 4(T-3) + 33 \cdot 2(T-1)}{(T-1)(T-3)} \end{aligned}$$

$Z_1$  and  $Z_2$  have a common denominator  $P = (t-1)(t-3)$ , whose roots are the coordinates of  $X_1$  in  $V(I)$ . Now, let

$$\begin{aligned} R_2 &= \frac{\Omega((\ell(X_2 X_1^i))_{i \geq 0}, P)}{\Omega((\ell(X_1^i))_{i \geq 0}, P)} \bmod P \\ &= \frac{17 \cdot 4(T-3) + 33 \cdot 2(T-1)}{17(T-3) + 33(T-1)} \bmod P \\ &= \frac{4(T-3) + 2(T-1)}{(T-3) + (T-1)} \bmod P \end{aligned}$$

Now,  $R_2(1) = 4$  and  $R_2(3) = 2$  as needed.

## 2.2 Computing the canonical generator of a linearly recurrent matrix sequence

We first present the notion of linear recurrence for sequences of matrices over a field  $\mathbb{K}$ , which extends the well-known notion for sequences in  $\mathbb{K}^{\mathbb{N}}$ .

**Definition 2.7** ([10, Sec. 3]). Let  $\mathcal{S} = (S_k)_{k \in \mathbb{N}} \subset \mathbb{K}^{m \times n'}$  be a matrix sequence. Then,

- a polynomial  $p = \sum_{0 \leq k \leq d} p_k T^k \in \mathbb{K}[T]$  is said to be a scalar relation for  $\mathcal{S}$  if  $\sum_{0 \leq k \leq d} p_k S_{\delta+k} = 0$  holds for all  $\delta \geq 0$ ;
- a polynomial vector  $\mathbf{p} = \sum_{0 \leq k \leq d} p_k T^k \in \mathbb{K}[T]^{1 \times m}$  is said to be a (left, vector) relation for  $\mathcal{S}$  if  $\sum_{0 \leq k \leq d} p_k S_{\delta+k} = 0$  holds for all  $\delta \geq 0$ ;
- $\mathcal{S}$  is said to be linearly recurrent if there exists a nontrivial scalar relation for  $\mathcal{S}$ .

For designing efficient algorithms it will be useful to rely on operations on polynomials or truncated series, hence the following characterization of vector relations.

**Lemma 2.8.** Consider a matrix sequence  $\mathcal{S} = (S_k)_{k \in \mathbb{N}} \subset \mathbb{K}^{m \times n'}$  and its generating series  $\mathbf{Z} = \sum_{k \geq 0} S_k / T^{k+1} \in \mathbb{K}[[T^{-1}]]^{m \times n'}$ . Then,  $\mathbf{p} \in \mathbb{K}[T]^{1 \times m}$  is a vector relation for  $\mathcal{S}$  if and only if the entries of  $\mathbf{q} = \mathbf{p}\mathbf{Z}$  are in  $\mathbb{K}[T]$ ; furthermore, in this case,  $\deg(\mathbf{q}) < \deg(\mathbf{p})$ .

*Proof.* Let  $\mathbf{p} = \sum_{0 \leq k \leq d} p_k T^k$ . For  $\delta \geq 0$ , the coefficient of  $\mathbf{q}$  of degree  $-\delta - 1 < 0$  is  $\sum_{0 \leq k \leq d} p_k S_{\delta+k}$ . Hence the equivalence, by definition of a relation. The degree comparison is clear since  $\mathbf{Z}$  has only terms of (strictly) negative degree.  $\square$

Concerning the algebraic structure of the set of vector relations, we have the following basic result, which can be found for example in [19, 10, 17].

**Lemma 2.9.** The sequence  $\mathcal{S}$  is linearly recurrent if and only if the set of left vector relations for  $\mathcal{S}$  is a  $\mathbb{K}[T]$ -submodule of  $\mathbb{K}[T]^{1 \times m}$  of rank  $m$ .

*Proof.* The set of vector relations for  $\mathcal{S}$  is a  $\mathbb{K}[T]$ -submodule of  $\mathbb{K}[T]^{1 \times m}$ , and hence is free of rank at most  $m$  [3, Chap. 12].

If  $\mathcal{S}$  is linearly recurrent, let  $p \in \mathbb{K}[T]$  be a nontrivial scalar relation for  $\mathcal{S}$ . Then each vector  $[0 \cdots 0 \ p \ 0 \cdots 0]$  with  $p$  at index  $1 \leq i \leq m$  is a vector relation for  $\mathcal{S}$ , hence  $\mathcal{S}$  has rank  $m$ . Conversely, if  $\mathcal{S}$  has rank  $m$ , then it has a basis with  $m$  vectors, which form a matrix in  $\mathbb{K}[T]^{m \times m}$ ; the determinant of this matrix is a nontrivial scalar relation for  $\mathcal{S}$ .  $\square$

Note however that a matrix sequence may admit nontrivial vector relations and have no scalar relation (and therefore not be linearly recurrent with the present definition); in this case the module of vector relations has rank less than  $m$ .

**Definition 2.10.** Let  $\mathcal{S} \subset \mathbb{K}^{m \times n'}$  be linearly recurrent. A (left) matrix generator for  $\mathcal{S}$  is a matrix in  $\mathbb{K}[T]^{m \times m}$  whose rows form a basis of the module of left vector relations for  $\mathcal{S}$ . This basis is said to be

- minimal if the matrix is row reduced [21, 9];

- canonical if the matrix is in Popov form [13, 9].

Note that the canonical generator is also a minimal generator; furthermore, all matrix generators  $\mathbf{P} \in \mathbb{K}[T]^{m \times m}$  for  $\mathcal{S}$  have the same determinantal degree  $\deg(\det(\mathbf{P}))$ , which we denote by  $\Delta(\mathcal{S})$ . We now show that minimal matrix generators are denominators in some irreducible fraction description of the generating series of the sequence. This is a direct consequence of Lemmas 2.8 and 2.9 and of basic properties of polynomial matrices.

**Corollary 2.11.** *A matrix sequence  $\mathcal{S} = (S_k)_{k \in \mathbb{N}} \subset \mathbb{K}^{m \times n'}$  is linearly recurrent if and only if its generating series  $\mathbf{Z} = \sum_{k \geq 0} S_k/T^{k+1} \in \mathbb{K}[[T^{-1}]]^{m \times n'}$  can be written as a matrix fraction  $\mathbf{Z} = \mathbf{P}^{-1}\mathbf{Q}$  where  $\mathbf{P} \in \mathbb{K}[T]^{m \times m}$  is nonsingular and  $\mathbf{Q} \in \mathbb{K}[T]^{m \times n'}$ . In this case, we have  $\text{rdeg}(\mathbf{Q}) < \text{rdeg}(\mathbf{P})$  and  $\deg(\det(\mathbf{P})) \geq \Delta(\mathcal{S})$ , and  $\mathbf{P}$  is a matrix generator of  $\mathcal{S}$  if and only if  $\deg(\det(\mathbf{P})) = \Delta(\mathcal{S})$  or, equivalently, the fraction  $\mathbf{P}^{-1}\mathbf{Q}$  is irreducible (that is,  $\mathbf{U}\mathbf{P} + \mathbf{V}\mathbf{Q} = \mathbf{I}$  for some polynomial matrices  $\mathbf{U}$  and  $\mathbf{V}$ ).*

We remark that we may also consider vector relations operating on the right: in particular, Lemma 2.8 shows that if the sequence is linearly recurrent then these right relations form a submodule of  $\mathbb{K}[T]^{n' \times 1}$  of rank  $n'$ . Thus, a linearly recurrent sequence also admits a right canonical generator.

Now, we focus on our algorithmic problem: given a linearly recurrent sequence, find a minimal matrix generator. We assume the availability of bounds  $(d_\ell, d_r)$  on the degrees of the left and right canonical generators, which allow us to control the number of terms of the sequence we will access during the algorithm. Since taking the Popov form of a reduced matrix does not change the degree, any left minimal matrix generator  $\mathbf{P}$  has the same degree  $\deg(\mathbf{P})$  as the left canonical generator: thus,  $d_\ell$  is also a bound on the degree of any left minimal generator. The same remark holds for  $d_r$  and right minimal generators. (These bounds  $d_\ell, d_r$  are the same as  $\gamma_1, \gamma_2$  in [17, Def. 4.6 and 4.7]; see also  $\delta_l, \delta_r$  in [20, Sec. 4.2].)

**Lemma 2.12.** *Let  $\mathcal{S} = (S_k)_k \subset \mathbb{K}^{m \times n'}$  be linearly recurrent and let  $d_r \in \mathbb{N}$  be such that the right canonical matrix generator of  $\mathcal{S}$  has degree at most  $d_r$ . Then,  $\mathbf{p} = \sum_{0 \leq k \leq d} p_k T^k \in \mathbb{K}[T]^{1 \times m}$  is a left relation for  $\mathcal{S}$  if and only if  $\sum_{0 \leq k \leq d} p_k S_{\delta+k} = 0$  holds for  $\delta \in \{0, \dots, d_r - 1\}$ .*

*Proof.* Since the right canonical generator  $\mathbf{P} \in \mathbb{K}[T]^{n' \times n'}$  is in column Popov form, we have  $\mathbf{P} = \mathbf{L}\text{Diag}(T^{t_1}, \dots, T^{t_n}) - \mathbf{Q}$  where  $\text{cdeg}(\mathbf{Q}) < \text{cdeg}(\mathbf{P}) = (t_1, \dots, t_n)$  componentwise and  $\mathbf{L} \in \mathbb{K}^{n' \times n'}$  is unit upper triangular. We define the matrix  $\mathbf{U} = \text{Diag}(T^{d_r - t_1}, \dots, T^{d_r - t_n})\mathbf{L}^{-1}$ , which is in  $\mathbb{K}[T]^{n' \times n'}$  since  $d_r \geq \deg(\mathbf{P}) = \max_j t_j$ . Then, the columns of the right multiple  $\mathbf{P}\mathbf{U} = T^{d_r}\mathbf{I}_n - \mathbf{Q}\mathbf{U}$  are right relations for  $\mathcal{S}$ , and we have  $\deg(\mathbf{Q}\mathbf{U}) < d_r$ . As a consequence, writing  $\mathbf{Q}\mathbf{U} = \sum_{0 \leq k < d_r} Q_k T^k$ , we have  $S_{d_r+\delta} = \sum_{0 \leq k < d_r} S_{k+\delta} Q_k$  for all  $\delta \geq 0$ .

Assuming that  $\sum_{0 \leq k \leq d} p_k S_{\delta+k} = 0$  holds for all  $\delta \in \{0, \dots, d_r - 1\}$ , we prove by induction that this holds for all  $\delta \in \mathbb{N}$ . Let  $\delta \geq d_r - 1$  and assume that this identity holds for all

integers up to  $\delta$ . Then, the identity concluding the previous paragraph implies that

$$\begin{aligned} \sum_{0 \leq k \leq d} p_k S_{\delta+1+k} &= \sum_{0 \leq k \leq d} p_k \left( \sum_{0 \leq j < d_r} S_{\delta+1+k-d_r+j} Q_j \right) \\ &= \sum_{0 \leq j < d_r} \underbrace{\left( \sum_{0 \leq k \leq d} p_k S_{\delta+1-d_r+j+k} \right)}_{=0 \text{ since } \delta+1-d_r+j \leq \delta} Q_j = 0, \end{aligned}$$

and the proof is complete.  $\square$

(A similar result is in [17, Thm. 4.5].)

The fast computation of matrix generators is usually handled via algorithms for computing minimal approximant bases [19, 17, 7]. The next result gives the main idea behind this approach. This result is similar to [17, Thm. 4.6] (see also [17, Thm. 4.7, 4.8, 4.9, 4.10]), but in some sense the reversal is on the input sequence rather than on the output matrix generator.

We recall from [18, 1] that, given a matrix  $\mathbf{F} \in \mathbb{K}[T]^{m \times n'}$  and an integer  $d \in \mathbb{N}$ , the set of *approximants* for  $\mathbf{F}$  at order  $d$  is defined as

$$\mathcal{A}(\mathbf{F}, d) = \{\mathbf{p} \in \mathbb{K}[T]^{1 \times m} \mid \mathbf{p}\mathbf{F} = 0 \bmod T^d\}.$$

Then, the next theorem shows that relations for  $\mathcal{S}$  can be retrieved as subvectors of approximants at order about  $d_\ell + d_r$  for a matrix involving the first  $d_\ell + d_r$  entries of  $\mathcal{S}$ .

**Theorem 2.13.** *Let  $\mathcal{S} = (S_k)_k \subset \mathbb{K}^{m \times n'}$  be a linearly recurrent sequence and let  $(d_\ell, d_r) \in \mathbb{N}^2$  be such that the left (resp. right) canonical matrix generator of  $\mathcal{S}$  has degree  $\leq d_\ell$  (resp.  $\leq d_r$ ).*

*For  $d > 0$ , define*

$$\mathbf{F} = \begin{bmatrix} \sum_{0 \leq k < d} S_k T^{d-k-1} \\ -\mathbf{I}_{n'} \end{bmatrix} \in \mathbb{K}[T]^{(m+n') \times n'}. \quad (5)$$

*For any relation  $\mathbf{p} \in \mathbb{K}[T]^{1 \times m}$  for  $\mathcal{S}$ , there exists  $\mathbf{r} \in \mathbb{K}[T]^{1 \times n'}$  such that  $\deg(\mathbf{r}) < \deg(\mathbf{p})$  and  $[\mathbf{p} \ \mathbf{r}] \in \mathcal{A}(\mathbf{F}, d)$ . Assuming  $d \geq d_r + 1$ , for any vectors  $\mathbf{p} \in \mathbb{K}[T]^{1 \times m}$  and  $\mathbf{r} \in \mathbb{K}[T]^{1 \times n'}$ , if  $[\mathbf{p} \ \mathbf{r}] \in \mathcal{A}(\mathbf{F}, d)$  and  $\deg([\mathbf{p} \ \mathbf{r}]) \leq d - d_r - 1$ , then  $\mathbf{p}$  is a relation for  $\mathcal{S}$ .*

*As a corollary, if  $\mathbf{B} \in \mathbb{K}[T]^{(m+n') \times (m+n')}$  is a basis of  $\mathcal{A}(\mathbf{F}, d_\ell + d_r + 1)$ , then*

- *if  $\mathbf{B}$  is in Popov form then its  $m \times m$  leading principal submatrix is the canonical matrix generator for  $\mathcal{S}$ ;*
- *if  $\mathbf{B}$  is row reduced then it has exactly  $m$  rows of degree  $\leq d_\ell$ , and the corresponding submatrix  $[\mathbf{P} \ \mathbf{R}]$  of  $\mathbf{B}$  is such that  $\mathbf{P} \in \mathbb{K}[T]^{1 \times m}$  is a minimal matrix generator for  $\mathcal{S}$ .*

*Proof.* From Lemma 2.8, if  $\mathbf{p}$  is a relation for  $\mathcal{S}$  then  $\mathbf{q} = \mathbf{p}\mathbf{Z}$  has polynomial entries, where  $\mathbf{Z} = \sum_{k \geq 0} S_k T^{-k-1}$ . Then, the vector  $\mathbf{r} = -\mathbf{p}(\sum_{k \geq d} S_k T^{d-k-1})$  has polynomial entries, has degree less than  $\deg(\mathbf{p})$ , and is such that  $[\mathbf{p} \ \mathbf{r}]\mathbf{F} = \mathbf{q}T^d$ , hence  $[\mathbf{p} \ \mathbf{r}] \in \mathcal{A}(\mathbf{F}, d)$ .

Conversely, if  $[\mathbf{p} \ \mathbf{r}] \in \mathcal{A}(\mathbf{F}, d)$  we have  $\mathbf{p}(\sum_{0 \leq k < d} S_k T^{d-k-1}) = \mathbf{r} \bmod T^d$ . Since  $d \geq d_r + 1$  and  $\deg([\mathbf{p} \ \mathbf{r}]) \leq d - d_r - 1$ , this implies that the coefficients of degree  $d - d_r$  to  $d - 1$  of  $\mathbf{p}(\sum_{0 \leq k < d} S_k T^{d-k-1})$  are zero. Then, Lemma 2.12 shows that  $\mathbf{p}$  is a relation for  $\mathcal{S}$ .

Finally, the two items are straightforward consequences.  $\square$

Then, using fast approximant basis algorithms, we obtain the next result.

**Corollary 2.14.** *Let  $\mathcal{S} \subset \mathbb{K}^{m \times n'}$  be a linearly recurrent sequence and let  $d = d_\ell + d_r + 1$ , where  $(d_\ell, d_r) \in \mathbb{N}^2$  are such that the left (resp. right) canonical matrix generator of  $\mathcal{S}$  has degree  $\leq d_\ell$  (resp.  $\leq d_r$ ). Then,*

- *using the algorithm of [6]: if  $n' \in \Omega(m)$ , a left minimal matrix generator for  $\mathcal{S}$  can be computed in  $O(n'^\omega \mathbf{M}(d) \log(d))$  operations in  $\mathbb{K}$ ;*
- *using the algorithm of [22]: if  $n' \in O(m)$ , a left minimal matrix generator for  $\mathcal{S}$  can be computed in  $O(m^\omega \mathbf{M}(n'd/m) \log(n'd))$  operations in  $\mathbb{K}$ ;*
- *using the algorithm of [8]: the left canonical matrix generator for  $\mathcal{S}$  can be computed in  $O((m + n')^{\omega-1} \mathbf{M}(n'd) \log(n'd)^3)$  operations in  $\mathbb{K}$ .*

In our algorithm, we will always use  $n' = m$ ; therefore we can find the left canonical matrix generator in  $O(m^\omega \mathbf{M}(d) \log(d))$  operations in  $\mathbb{K}$ .



### 3 Block Sparse-FGLM algorithm

In this section, we will show how to extend the Sparse-FGLM to using blocking methods. Steel's method also uses the Block Wiedemann algorithm to compute the minimal polynomial of  $\mathbf{M}_i$ , for which the roots provide the appropriate values for  $X_i$ , but uses the “evaluation” method for the rest (another Gröbner Basis computation with that variable evaluated at each root of the minimal polynomial) [15]. Our algorithm computes the rest of the lex Gröbner basis directly.

Given:

$I \subset \mathbb{K}$ :	zero dimensional ideal in shape position
$\mathbb{B} \subset \mathbb{K}[X_1, \dots, X_n]/I$ :	monomial basis of $\mathbb{K}[X_1, \dots, X_n]/I$
$D$ :	dimension of $\mathbb{B}$
$\mathbf{M}_1, \dots, \mathbf{M}_n$ :	multiplication matrices of $X_1, \dots, X_n$ respectively
$t$ :	random linear combination of $X_i$ 's
$\mathbf{M}$ :	multiplication matrix of $t$

we compute a lex Gröbner basis that have the same points in its variety as the radical of  $I$  (note that we do not assume that  $I$  is radical). This is because we introduce another variable  $t$  which, generically, separates the points in the variety. We also assume that the base field  $\mathbb{K}$  has characteristic larger than  $D$ . More precisely, we want to find polynomials  $(R, R_1, \dots, R_n)$  such that for all  $\alpha$  that is a factor of  $R$ ,  $\{(R_1(\alpha), \dots, R_n(\alpha))\} = V(I)$

The main idea is to compute the minimal polynomial of  $\mathbf{M}$  by using the block Wiedemann algorithm. This is done by applying the Matrix Berlekamp-Massey algorithm of Section 2.2 on the sequence  $S = (\mathbf{U}^{tr} \mathbf{M}^i \mathbf{V})_{0 \leq i < 2\lceil \frac{D}{m} \rceil}$ , for some  $m \in \mathbb{N}$  and  $U, V \in \mathbb{K}^{D \times m}$ , and taking the largest invariant factor. We find the rest of the polynomials in the output by extracting the scalar closed forms from the matrix closed forms. Lastly, we denote a matrix  $\mathbf{N}$  with all its terms of negative exponents removed as  $(\mathbf{N})_{T \geq 0}$

---

**Algorithm 3** Block Sparse-FGLM( $\mathbf{M}, \mathbf{M}_1, \dots, \mathbf{M}_n, m$ )

---

**Input:**

- $\mathbf{M}, \mathbf{M}_1, \dots, \mathbf{M}_n$  defined as above
- dimension of the blocks  $m \in \mathbb{N}$

**Output:** polynomials  $(R, R_1, \dots, R_n)$

1. choose  $\mathbf{U}, \mathbf{V} \in \mathbb{K}^{D \times m}$
  2.  $S = (\mathbf{U}^{tr} \mathbf{M}^i \mathbf{V})_{0 \leq i < 2d}$ , with  $d = \frac{D}{m}$
  3.  $\mathbf{F}^{U,M,V} = \text{MatrixBerlekampMassey}(S)$
  4.  $\mathbf{N}^* = (\mathbf{F}^{U,M,V} \sum_{i=0}^{d-1} (\mathbf{U}^{tr} \mathbf{M}^i e) / T^{i+1})_{T \geq 0}$
  5.  $P = \text{largest invariant factor of } \mathbf{F}^{U,M,V}$
  6.  $R = \text{SquareFreePart}(P)$
  7.  $a = [0 \ \dots \ 0 P] (\mathbf{F}^{U,M,V})^{-1}$
  8.  $N = a \mathbf{N}^*$
  9. for  $j = 1 \dots n$ :
    - 9.1.  $\mathbf{N}_j^* = (\mathbf{F}^{U,M,V} \sum_{i=0}^{d-1} (\mathbf{U}^{tr} \mathbf{M}^i \mathbf{M}_j e) / T^{i+1})_{T \geq 0}$
    - 9.2.  $N_j = a \mathbf{N}_j^*$
    - 9.3.  $R_j = N_j / N \text{ mod } R$
- 

In lines 2,4, and 9.1 of the above algorithm, we use  $(\mathbf{U}^{tr} \mathbf{M}^i)_{i \geq 0}$ , which we compute once in line 2 and store. Computing this sequence is the bottleneck of the algorithm, but this can be parallelized. Let  $\mathbf{U} \in \mathbb{K}^{D \times m}$ , then we can write  $\mathbf{U}$  as  $[u_1 u_2 \dots u_m]$  where  $u_k \in \mathbb{K}^{D \times 1}$ ,  $1 \leq k \leq m$ , is a vector. Now,

$$\mathbf{U}^{tr} \mathbf{M}^i = [u_1^{tr} \dots u_m^{tr}] \mathbf{M}^i = [u_1^{tr} \mathbf{M}^i \dots u_m^{tr} \mathbf{M}^i],$$

where each product  $u_k^{tr} \mathbf{M}^i$  is independent of others. Therefore, to generate the sequence  $(\mathbf{U}^{tr} \mathbf{M}^i)_{i \geq 0}$ , we generate, in parallel, the sequences  $(u_k^{tr} \mathbf{M}^i)_{i \geq 0}$ ,  $1 \leq k \leq m$ , and concatenate them together as above. Let  $c$  be the sparsity of  $\mathbf{M}$ , then generating  $(u_k^{tr} \mathbf{M}^i)_{0 \leq i < 2d}$  costs  $O(dcD^2) = O(cD^3/m)$  field operations. If we are able to compute  $m$  vector-matrix products in parallel at once, the total cost of generating  $(\mathbf{U}^{tr} \mathbf{M}^i)_{0 \leq i < 2d}$  is  $O(cD^3/m)$ . We can then compute  $(\mathbf{U}^{tr} \mathbf{M}^i \mathbf{V})_{0 \leq i < 2d}$  by a product

$$\begin{bmatrix} \mathbf{U}^{tr} \\ \mathbf{U}^{tr} \mathbf{M} \\ \mathbf{U}^{tr} \mathbf{M}^2 \\ \vdots \\ \mathbf{U}^{tr} \mathbf{M}^{2d} \end{bmatrix} \mathbf{V} = \begin{bmatrix} \mathbf{U}^{tr} \mathbf{V} \\ \mathbf{U}^{tr} \mathbf{M} \mathbf{V} \\ \mathbf{U}^{tr} \mathbf{M}^2 \mathbf{V} \\ \vdots \\ \mathbf{U}^{tr} \mathbf{M}^{2d} \mathbf{V} \end{bmatrix}$$

in cost  $O(m^{\omega-2} D^2)$ .

Recall, from Section 2.2, that we can compute the minimum generating polynomial matrix in  $O(m^{\omega-1} D \log(D/m))$ . We can compute  $P$  and  $a$  in lines 5 and 7 respectively by using the high-order lifting algorithm of [16, algorithm 5]. By [16, Cor. 16], the cost of this is

$O(m^\omega d \log(m)) = O(m^{\omega-1} D \log(m))$ . Now, for any  $v \in \mathbb{K}^{D \times 1}$ , computing  $(\mathbf{U}\mathbf{M}^i v)_{0 \leq i < d}$  costs  $O(D^2)$  by using the above method. Recall, from Section 2.1.1, that we can compute  $\mathbf{N}_v^* = \mathbf{F}^{U,M,V} \sum_{i=0}^{d-1} (\mathbf{U}^{tr} \mathbf{M}^i v) / T^{i+1}$  by first doing a product  $\mathbf{F}^{U,M,V} \sum_{i=0}^{d-1} (\mathbf{U}^{tr} \mathbf{M}^i v) T^i$ . Since both  $\mathbf{F}^{U,M,V}$  and  $\sum_{i=0}^{d-1} (\mathbf{U}^{tr} \mathbf{M}^i v) T^i$  have degree at most  $d$ , this product can be done in cost  $O(m(md)) = O(mD)$ . Now,  $\mathbf{N}_v^*$  has degree less than  $d$  and  $a$  has degree at most  $D$ , so computing  $a\mathbf{N}_v^*$  costs  $O(mD)$ . Therefore, computing  $N, N_1, \dots, N_j$  costs  $O(nmD)$ . Since  $\mathbf{F}^{U,M,V}$  has degree  $d$  and  $a$  has degree  $D$  at most, we can compute  $\cdot$ . Since  $P, R, N, N_j$ ,  $1 \leq j \leq n$  have degree at most  $D$ , lines 6 and 9.3 costs  $O(D)$ . Finally, we get the cost  $O(cD^3/m + m^{\omega-2}D^2 + m^{\omega-1}D \log(D) + nmD)$ .

### 3.1 Proof of Correctness

First, we present a proof that one can compute the minimal polynomial of  $\mathbf{M}$  from the output of Matrix Berlekamp-Massey if we choose the blocking matrices  $\mathbf{U}$  and  $\mathbf{V}$  generically. This result had been previously been proven by Kaltofen and Villard [11][20].

Let  $\mathbf{M}$  be in  $\mathbb{K}^{D \times D}$  and  $\mathbf{U}, \mathbf{V} \in \mathbb{K}^{D \times m}$  be two blocking matrices. Now, define two matrix sequences  $S_V = (\mathbf{M}^i \mathbf{V})_{(i \geq 0)}$  and  $S_{U,V} = (\mathbf{U}^{tr} \mathbf{M}^i \mathbf{V})_{(i \geq 0)}$  denote their minimum generating matrix polynomial as  $\mathbf{F}^{M,V}$  and  $\mathbf{F}^{U,M,V}$  respectively. Let  $s_1, \dots, s_r$  be the invariant factors of  $T\mathbf{I} - \mathbf{M}$ , ordered in such a way that  $s_r | s_{r-1} | \dots | s_1$ , and let  $d_i = \deg(s_i)$  for all  $i$ ; for  $i > r$ , we let  $s_i = 1$ , with  $d_i = 0$ . We define  $\nu = d_1 + \dots + d_m \leq D$  and  $\delta = \lceil \nu/m \rceil \leq \lceil D/m \rceil$ . We also denote by  $\sigma_1, \dots, \sigma_k$  the invariant factors of  $\mathbf{F}^{U,M,V}$ , for some  $k \leq m$ . As above, for  $i > k$ , we let  $\sigma_i = 1$ .

**Theorem 3.1.** *For a generic choice of  $\mathbf{U}$  and  $\mathbf{V}$ , we have:*

- $\mathbf{F}^{U,M,V}$  has degree  $\delta$ ;
- $s_i = \sigma_i$  for  $1 \leq i \leq m$ .

*Proof.* We denote by  $\langle \mathbf{V} \rangle$  the vector space generated by the columns of  $\mathbf{V}, \mathbf{M}\mathbf{V}, \mathbf{M}^2\mathbf{V}, \dots$ . We also write  $D_V = \dim(\langle \mathbf{V} \rangle)$ .

First, we prove that for any  $\mathbf{V}$  in  $\mathbb{K}^{D \times m}$ , for a generic  $\mathbf{U}$  in  $\mathbb{K}^{D \times m}$ ,  $\mathbf{F}^{U,M,V} = \mathbf{F}^{M,V}$ . Indeed, by [20, Lemma 4.2], there exists matrices  $\mathbf{P}_V$  in  $\mathbb{K}^{D \times D_V}$  and  $\mathbf{M}_V \in \mathbb{K}^{D_V \times D_V}$ , with  $\mathbf{P}_V$  of full rank  $D_V$ , and where  $\mathbf{M}_V$  is a matrix of the restriction of  $\mathbf{M}$  to  $\langle \mathbf{V} \rangle$ , such that  $\mathbf{F}^{U,M,V} = \mathbf{F}^{M,V}$  if and only if the dimension of the span of  $[\mathbf{Z} \mathbf{B}_V \mathbf{Z} \mathbf{B}_V^2 \mathbf{Z} \dots]$  is equal to  $D_V$ , with  $\mathbf{B}_V = \mathbf{M}_V^\perp$  and  $\mathbf{Z} = \mathbf{P}_V^\perp \mathbf{U} \in \mathbb{K}^{D_V \times m}$ .

We prove that this is the case for a generic  $\mathbf{U}$ . By construction, one can find a basis of  $\langle \mathbf{V} \rangle$  in which the matrix of  $\mathbf{M}_V$  is block-companion, with  $m' \leq m$  blocks (take the  $\mathbf{M}_V$ -span of the first column of  $\mathbf{V}$ , then of the second column, working modulo the previous vector space, etc.) Thus,  $\mathbf{B}_V$  is similar to a block-companion matrix with  $m'$  blocks as well; since  $\mathbf{Z}$  has  $m$  columns,  $S$  has full dimension  $N_V$  for a generic  $\mathbf{Z}$  (and for a generic  $\mathbf{U}$ , since  $P_V$  has rank  $N_V$ ). Thus, for generic choices of  $\mathbf{U}$  and  $\mathbf{V}$ ,  $\mathbf{F}^{U,M,V} = \mathbf{F}^{M,V}$ .

Let us next introduce a matrix  $\mathcal{V}$  of indeterminates of size  $D \times m$ , and let  $\mathbf{F}^{M,\mathcal{V}}$  be the minimal generating polynomial of the “generic” sequence  $(\mathbf{M}^i \mathcal{V})_{i \geq 0}$ . The notation  $\langle \mathcal{V} \rangle$  and

$D_{\mathcal{V}}$  are defined as above. In particular, by [20, Proposition 6.1], the minimal generating polynomial  $\mathbf{F}^{M,\mathcal{V}}$  has degree  $\delta$  and determinantal degree  $\nu$ .

Now, for a generic  $\mathbf{V}$  in  $\mathbb{K}^{D \times m}$ ,  $D_V = D_{\mathcal{V}}$ . Indeed,  $\langle \mathcal{V} \rangle$  is the span of  $K_{\mathcal{V}} = [\mathcal{V} \mathbf{M} \mathcal{V} \cdots \mathbf{M}^{N-1} \mathcal{V}]$ , whereas  $\langle \mathbf{V} \rangle$  is the span of  $[\mathbf{V} \mathbf{M} \mathbf{V} \cdots \mathbf{M}^{N-1} \mathbf{V}]$ . Take a maximal non-zero minor  $\mu$  of  $K_{\mathcal{V}}$ ; as soon as  $\mu(Y) \neq 0$ , we have equality of the dimensions. On the other hand, by [20, Lemma 4.3], for any  $\mathbf{V}$  (including  $\mathcal{V}$ ), the degree of  $\mathbf{F}^{M,V}$  is equal to the first index  $d$  such that  $\dim(\text{span}([\mathbf{V} \mathbf{M} \mathbf{V} \cdots \mathbf{M}^{d-1} \mathbf{V}])) = D_V$ . As a result, for generic  $\mathbf{V}$ ,  $\mathbf{F}^{M,V}$  and  $\mathbf{F}^{M,\mathcal{V}}$  have the same degree, that is,  $\delta$ . The first item is proved.

We conclude by proving that for generic  $\mathbf{U}, \mathbf{V}$ , the invariant factors  $\sigma_1, \dots, \sigma_m$  of  $\mathbf{F}^{U,M,V}$  are  $s_1, \dots, s_m$ . By [11, Theorem 2.12], for any  $\mathbf{U}$  and  $\mathbf{V}$  in  $\mathbb{K}^{D \times m}$ , for  $i = 1, \dots, m$ , the  $i^{\text{th}}$  invariant factor  $\sigma_i$  of  $\mathbf{F}^{U,M,V}$  divides  $s_i$ , so that  $\deg(\det(\mathbf{F}^{U,M,V})) \leq \nu$ , with equality if and only if  $\sigma_i = s_i$  for all  $i \leq m$ .

For  $\mathbf{V}$  as above and any integers  $e, d$ , we let  $\text{Hk}_{e,d}(\mathbf{V})$  be the block Hankel matrix

$$\text{Hk}_{e,d}(\mathbf{V}) = \begin{bmatrix} \mathbf{I} \\ \mathbf{M} \\ \mathbf{M}^2 \\ \vdots \\ \mathbf{M}^{e-1} \end{bmatrix} [\mathbf{V} \quad \mathbf{M}\mathbf{V} \quad \mathbf{M}^2\mathbf{V} \quad \cdots \quad \mathbf{M}^{d-1}\mathbf{V}]$$

By [11, Eq. (2.6)],  $\text{rank}(\text{Hk}_{e,d}(\mathbf{V})) = \deg(\det(\mathbf{F}^{M,V}))$  for  $d \geq \deg(\mathbf{F}^{M,V})$  and  $e \geq D$ . We take  $e = D$ , so that  $\text{rank}(\text{Hk}_{D,d}(\mathbf{V})) = \deg(\det(\mathbf{F}^{M,V}))$  for  $d \geq \deg(\mathbf{F}^{M,V})$ . On the other hand, the sequence  $\text{rank}(\text{Hk}_{N,d}(\mathbf{V}))$  is constant for  $d \geq D$ ; as a result,  $\text{rank}(\text{Hk}_{D,D}(\mathbf{V})) = \deg(\det(\mathbf{F}^{T,V}))$ . For the same reason, we also have  $\text{rank}(\text{Hk}_{D,D}(\mathcal{V})) = \deg(\det(\mathbf{F}^{M,\mathcal{V}}))$ , so that for a generic  $\mathbf{U}$ ,  $\mathbf{F}^{M,V}$  and  $\mathbf{F}^{M,\mathcal{V}}$  have the same determinantal degree, that is,  $\nu$ . As a result, for generic  $\mathbf{U}$  and  $\mathbf{V}$ , we also have  $\deg(\det(\mathbf{F}^{U,M,V})) = \nu$ , and the conclusion follows.  $\square$

It is well known that the largest invariant factor of  $T\mathbf{I} - \mathbf{M}$  is the minimal polynomial of  $\mathbf{M}$ . Therefore, the above theorem shows that if we choose the entries of  $\mathbf{U}$  and  $\mathbf{V}$  randomly, we will have that the largest invariant factor of  $\mathbf{F}^{U,M,V}$  is the minimal polynomial of  $\mathbf{M}$  with high probability.

Next, we prove that one can recover a scalar numerator through a matrix numerator.

**Lemma 3.2.** *Let  $\mathbf{a}$  be defined as line 7 of Algorithm 3, then  $\mathbf{a}$  has polynomial entries.*

*Proof.* Let  $\mathcal{D} = \mathbf{A}\mathbf{F}^{U,M,V}\mathbf{B}$  be the Smith normal form of  $\mathbf{F}^{U,M,V}$  and  $s_1, \dots, s_m$  be invariant factors of  $\mathbf{F}^{U,M,V}$  such that  $s_m | s_{m-1} | \cdots | s_1$ . Let  $[b_1, \dots, b_m]$  be the last row of  $\mathbf{B}$  and  $w =$

$[\frac{s_1 b_1}{s_m}, \frac{s_1 b_2}{s_{m-1}}, \dots, \frac{s_1 b_{m-1}}{s_2}, b_m]$  (since  $s_i | s_1$ ), then

$$\begin{aligned} (w\mathbf{A})\mathbf{A}^{-1}\mathcal{D} &= [\frac{s_1 b_1}{s_m}, \frac{s_1 b_2}{s_{m-1}}, \dots, \frac{s_1 b_{m-1}}{s_2}, b_m] \begin{bmatrix} s_m & & \\ & \ddots & \\ & & s_1 \end{bmatrix} \\ &= [s_1 b_1, s_1 b_2, \dots, s_1 b_m] \\ &= [0, \dots, 0, s_1]\mathbf{B} \end{aligned}$$

Therefore, if  $a = w\mathbf{A}$ , we get  $a\mathbf{F}^{U,M,V} = (w\mathbf{A})\mathbf{A}^{-1}\mathcal{D}\mathbf{B}^{-1} = [0, \dots, 0, s_1]$  as needed. Since both  $w$  and  $\mathbf{A}$  have polynomial entries,  $a$  must also have polynomial entries.  $\square$

**Theorem 3.3.** *Let  $U = [u_1, u_2, \dots, u_m]$  be in  $\mathbb{K}^{D \times m}$ ,  $\mathbf{F}^{U,M,V}$  be the minimum generator of  $(\mathbf{U}^t \mathbf{M}^i \mathbf{V})_{i \geq 0}$ , and  $P$  be the minimal polynomial of  $\mathbf{M}$ . For any  $v \in \mathbb{K}^D$ , if  $\mathbf{N}^* = (\mathbf{F}^{U,M,V} \sum_{i=0}^{d-1} (\mathbf{U}^t \mathbf{M}^i v) / T^{i+1})_{T \geq 0}$  and  $N = a\mathbf{N}^*$ , then  $N = \Omega((u_m^t \mathbf{M}^i v)_{i \geq 0}, P)$ .*

*Proof.* Let  $Z = \sum_{i=0}^{\infty} (\mathbf{U}^t \mathbf{M}^i v) / T^{i+1}$  and  $\mathbf{N}^{*'} = \mathbf{F}^{U,M,V} Z$ . Since the highest power of the entries in  $Z$  is  $T^{-1}$ , the entries of the product  $\mathbf{N}^{*'}$  must have degree less than  $d = \deg(\mathbf{F}^{U,M,V})$ . Furthermore, since  $\mathbf{F}^{U,M,V}$  cancels the sequence  $(\mathbf{U}^t \mathbf{M}^i \mathbf{V})_{i \geq 0}$ ,  $\mathbf{N}^{*'}$  is a polynomial matrix and does not have any terms of negative exponents. This means that all terms in  $Z$  with degree less than  $T^{-d}$  must vanish in the product. Therefore,

$$\mathbf{N}^* = \mathbf{N}^{*'}$$

Now, rewrite  $\mathbf{U}$  as  $\mathbf{U} = [u_1 u_2 \dots u_m]$ , then

$$\mathbf{N}^* = \mathbf{F}^{U,M,V} \sum_{i=0}^{\infty} \mathbf{U}^{tr} \mathbf{M}^i v / T^{i+1} = \begin{bmatrix} \sum_{i \geq 0} u_1^{tr} \mathbf{M}^i v / T^{i+1} \\ \vdots \\ \sum_{i \geq 0} u_m^{tr} \mathbf{M}^i v / T^{i+1} \end{bmatrix}$$

Recall  $\Omega$  from Definition 2.1. By rewriting each  $\sum_{i \geq 0} u_j^{tr} \mathbf{M}^i v$  in its closed form, we get

$$\mathbf{N}^* = \mathbf{F}^{U,M,V} \begin{bmatrix} \Omega((u_1^{tr} \mathbf{M}^i v)_{i \geq 0}, P) / P \\ \vdots \\ \Omega((u_m^{tr} \mathbf{M}^i v)_{i \geq 0}, P) / P \end{bmatrix}$$

By Theorem 3.1, the  $i^{th}$  invariant factor of  $T\mathbf{I} - \mathbf{M}$  is equal to the  $i^{th}$  invariant factor of  $\mathbf{F}^{U,M,V}$  for generic choice of  $\mathbf{U}, \mathbf{V}$ . Thus,  $s_1 = P$  and by Lemma 3.2

$$\begin{aligned} a\mathbf{N}^* &= a\mathbf{F}^{U,M,V} \begin{bmatrix} \Omega((u_1^{tr} \mathbf{M}^i v)_{i \geq 0}, P) / P \\ \vdots \\ \Omega((u_m^{tr} \mathbf{M}^i v)_{i \geq 0}, P) / P \end{bmatrix} \\ &= [0, \dots, 0, P] \begin{bmatrix} \Omega((u_1^{tr} \mathbf{M}^i v)_{i \geq 0}, P) / P \\ \vdots \\ \Omega((u_m^{tr} \mathbf{M}^i v)_{i \geq 0}, P) / P \end{bmatrix} \\ &= \Omega((u_m^{tr} \mathbf{M}^i v)_{i \geq 0}, P) \end{aligned}$$

Therefore,  $N = \Omega((u_m^{tr} \mathbf{M}^i v)_{i \geq 0}, P)$  as needed.  $\square$

Finally, we conclude by seeing what happens when we pick specific values for  $v$  and applying Theorem 3.3. In line 4 of Algorithm 3, we have  $v = e$ , where  $e$  is the coordinate vector for  $1 \in \mathbb{B}$ . By applying Theorem 3.3, we have that

$$\mathbf{N}^* = \mathbf{F}^{U,M,V} \sum_{i \geq 0} \mathbf{U}^t \mathbf{M} e / T^{i+1},$$

and

$$N = a\mathbf{N}^* = \Omega((u_m^{tr} \mathbf{M}^i e)_{i \geq 0}, P).$$

By construction,  $\mathbf{M}^i e$  gives the coordinate vector of  $t^i$  in  $\mathbb{B}$ , so

$$\sum_{i \geq 0} u_m^{tr} \mathbf{M}^i e / T^{i+1} = \sum_{i \geq 0} \ell_{u_m}(t^i) / T^{i+1}$$

where  $\ell_{u_m}$  is the linear form associated with  $u_m$ . Therefore,

$$N = \Omega((u_m^{tr} \mathbf{M}^i e)_{i \geq 0}, P) = \Omega((\ell_{u_m}(t^i))_{i \geq 0}, P)$$

Similarly, in line 7.1 of Algorithm 3, we compute  $\mathbf{N}_j^* = \mathbf{F}^{U,M,V} \sum_{i \geq 0} \mathbf{U}^{tr} \mathbf{M} \mathbf{M}_j e / T^{i+1}$ . By Theorem 3.3,

$$N_j = a\mathbf{N}_j^* = \Omega((u_m^{tr} \mathbf{M}^i \mathbf{M}_j e)_{i \geq 0}, P)$$

Again, by construction,  $\mathbf{M}_j e$  is the coordinate vector for  $X_j$  in  $\mathbb{B}$  and  $\mathbf{M}^i \mathbf{M}_j e$  is the coordinate vector for  $X_j t^i$  in  $\mathbb{B}$ . Therefore,

$$N_j = \Omega((u_m^{tr} \mathbf{M}^i \mathbf{T}_j e)_{i \geq 0}, P) = \Omega((\ell_{u_m}(X_j t^i))_{i \geq 0}, P)$$

By the correctness of Algorithm 2, we have the proof of correctness for the Block Sparse-FGLM algorithm.

### 3.2 Example

First, we give an example in the radical case. Let

$$I = \langle -16X_1^2 - 15X_1X_2 - 14X_2^2 - 48X_1 + 26, 35X_1X_2 + 47X_1 - 46X_2 - 47 \rangle \subset GF(101)[X_1, X_2]$$

We choose  $t = 2X_1 + 53X_2$ , with multiplication matrices:

$$\mathbf{M}_1 = \begin{bmatrix} 85 & 0 & 37 & 0 \\ 69 & 85 & 15 & 0 \\ 100 & 91 & 19 & 1 \\ 1 & 10 & 68 & 0 \end{bmatrix} \quad \mathbf{M}_2 = \begin{bmatrix} 36 & 1 & 0 & 0 \\ 42 & 0 & 85 & 1 \\ 51 & 0 & 91 & 0 \\ 95 & 0 & 10 & 0 \end{bmatrix} \quad \mathbf{M} = \begin{bmatrix} 58 & 53 & 74 & 0 \\ 41 & 69 & 91 & 53 \\ 75 & 81 & 13 & 2 \\ 88 & 20 & 60 & 0 \end{bmatrix}$$

We choose  $m = 2$  and choose  $U, V \in GF(101)^{D \times m}$  of random entries:

$$\mathbf{U}^{tr} = \begin{bmatrix} 84 & 38 \\ 29 & 58 \\ 80 & 43 \\ 7 & 82 \end{bmatrix} \quad \mathbf{V} = \begin{bmatrix} 6 & 97 \\ 83 & 58 \\ 101 & 95 \\ 59 & 89 \end{bmatrix}$$

We compute the matrix sequence  $S = (\mathbf{U}^{tr} \mathbf{M}^i \mathbf{V})_{i \geq 0}$  and its minimum generating matrix polynomial  $\mathbf{G}$

$$S = \left( \begin{bmatrix} 92 & 75 \\ 83 & 51 \end{bmatrix}, \begin{bmatrix} 57 & 82 \\ 23 & 16 \end{bmatrix}, \begin{bmatrix} 54 & 93 \\ 70 & 66 \end{bmatrix}, \begin{bmatrix} 50 & 77 \\ 26 & 76 \end{bmatrix} \right)$$

$$\mathbf{F}^{U,M,V} = \begin{bmatrix} T^2 + 76T + 8 & 87T + 31 \\ 100T + 46 & T^2 + 87T + 44 \end{bmatrix}$$

The biggest invariant factor of  $\mathbf{F}^{U,M,V}$  is

$$P = T^4 + 62T^3 + 85T^2 + 69T + 37 = R$$

since  $P$  is square free. Now, we compute  $\mathbf{N}^*$  and  $a$ :

$$\mathbf{N}^* = [84T + 46, 38T + 65]$$

$$a = [T + 55, T^2 + 76T + 8]$$

Finally, we find the scalar numerator  $N = a\mathbf{N}^*$ :

$$N = 100T^3 + 26T^2 + 33T + 18$$

To get  $R_1(T)$ , we compute  $\mathbf{N}_1^*$  and  $N_1 = a\mathbf{N}_1^*$ :

$$\mathbf{N}_1^* = [79T + 8, 100T + 23]$$

$$N_1 = 100T^3 + 26T^2 + 33T + 18$$

Lastly,

$$R_1(T) = N_1/N \mod R$$

$$= 61T^3 + 75T^2 + 85T + 23$$

We compute  $R_2(t) = 32T^3 + 41T^2 + 94T + 22$  in the same way. As a sanity check, we see that  $V(I)$  has one point in  $GF(101)^2$ :  $(54, 79)$  and  $P$  has one factor in  $GF(101)$ : 53. Now,  $R_1(53) = 54$  and  $R_2(53) = 79$  as expected.

## 4 Experimental Results

name	n	D	Sparsity	m = 1	m = 4	m = 8	deg(P) = D
rand3-1	3	512	0.08	2.853	0.798	0.592	yes
rand3-2	3	1728	0.05	70.626	21.941	13.381	yes
rand3-3	3	3375	0.04	418.708	113.336	77.259	yes
randsq3-1	3	1728	0.05	70.633	19.770	13.249	no
randsq3-2	3	2744	0.04	241.162	63.778	44.829	no
randsq3-2	3	4096	0.04	699.944	187.988	128.88	no
bannwarth	5	284	0.27	1.027	0.364	0.220	yes
cyclic7	7	924	0.02	7.038	1.954	1.448	yes
eco12	12	1024	0.13	62.294	16.427	11.539	yes
katsura8	9	256	0.28	1.207	0.338	0.249	yes
katsura9	10	512	0.26	9.210	2.453	1.865	yes
katsura10	11	1024	0.25	71.243	18.867	13.218	yes
sot1	5	8694	0.001	762.063	208.25	143.281	yes
vor2	6	574	0.24	10.238	2.749	1.972	yes



## 5 Improvements for Special cases

### 5.1 Choice of the blocking size

In general, experiments show that choosing  $1 \leq m \leq D$  that matches the number of cores is optimal in Section 4. However, when  $m = 1$ , blocking has no effect and there is no parallelization to be had. In particular, the minimum generating polynomial matrix  $\mathbf{F}^{U,M,V}$  is equal to the minimum polynomial  $P$  of  $\mathbf{M}$ . Therefore,  $a = P(\mathbf{F}^{U,M,V})^{-1} = 1$  and we can omit steps 7,8, and 9.2 in Algorithm 3.

If the multiplication matrices are dense, then it is better to compute the product  $\mathbf{UM}^i$  using a single matrix multiplication, rather than in parallel. When  $m = D$ ,  $\mathbf{U}, \mathbf{V} \in \mathbb{K}^{D \times D}$  are square matrices and we can write  $\mathbf{UMV} = \mathbf{U}'\mathbf{M}$ ; therefore, we can generate sequences of the form  $(\mathbf{UM}^i)_{i \geq 0}$  rather than  $(\mathbf{UM}^i\mathbf{V})_{i \geq 0}$ . Now,  $d = D/m = 1$  and we generate the sequence  $S = (\mathbf{U}, \mathbf{UM})$ . The minimum generating polynomial matrix of  $S$  is

$$\mathbf{F}^{U,M} = T - \mathbf{UMU}^{-1},$$

and

$$((T - \mathbf{UMU}^{-1})(\mathbf{U}/T + \mathbf{UM}/T^2))_{T \geq 0} = \mathbf{U}.$$

For  $\mathbf{M}_j$ ,  $1 \leq j \leq n$ , we get the sequence  $(\mathbf{UM}_j, \mathbf{UMM}_j)$  and

$$((T - \mathbf{UMU}^{-1})(\mathbf{UM}_j/T + \mathbf{UMM}_j/T))_{T \geq 0} = \mathbf{UM}_j.$$

Thus, we get the following algorithm:

---

**Algorithm 4** Dense Block Sparse-FGLM( $\mathbf{M}, \mathbf{M}_1, \dots, \mathbf{M}_n$ )

---

**Input:**

- $\mathbf{M}, \mathbf{M}_1, \dots, \mathbf{M}_n$  defined as above

**Output:** polynomials  $(R, R_1, \dots, R_n)$

1. choose  $\mathbf{U} \in \mathbb{K}^{D \times D}$
  2.  $\mathbf{F}^{U,M} = T - \mathbf{UMU}^{-1}$
  3.  $\mathbf{N}^* = \mathbf{U}$
  4.  $P$  = largest invariant factor of  $\mathbf{F}^{U,M}$
  5.  $R$  = SquareFreePart( $P$ )
  6.  $a = [0 \ \dots \ 0P](\mathbf{F}^{U,M,V})^{-1}$
  7.  $N = a\mathbf{N}^*e$
  8. for  $j = 1 \dots n$ :
    - 8.1.  $\mathbf{N}_j^* = \mathbf{UM}_j$
    - 8.2.  $N_j = a\mathbf{N}_j^*e$
    - 8.3.  $R_j = N_j/N \bmod R$
  9. return  $((R, R_1, \dots, R_n), t)$
-

Let  $\omega$  be an exponent such that matrices of size  $n$  can be multiplied in  $O(n^\omega)$ . In this case, computing  $\mathbf{F}^{U,M}$  costs  $O(D^\omega)$  and computing  $N, N_1, \dots, N_n$  costs  $O(nD^2)$  since entries of  $a$  has roughly degree  $D$ . We can solve for  $a$  by using the high-order lifting algorithm of [16, algorithm 5]. By [16, Cor. 16], the cost of this is  $O(D^\omega \log(D))$ . Putting everything together, we get the final cost of  $O(D^\omega \log(D) + nD^2)$ .

## 5.2 Decomposition of linear forms.

In this paragraph, we work over the whole  $V$  (so  $\mathfrak{S} = \{1, \dots, d\}$ ). Specializing our previous discussion to the case  $t = X_1$ , we let  $r_1, \dots, r_c$  be the pairwise distinct values taken by  $X_1$  on  $V$ , for some  $c \leq d$ . For  $j = 1, \dots, c$ , we write  $T_j$  for the set of all indices  $i$  in  $\{1, \dots, d\}$  such that  $\alpha_{i,1} = r_j$ ; the sets  $T_1, \dots, T_c$  form a partition of  $\{1, \dots, d\}$ . When  $T_j$  has cardinality 1, we denote it as  $T_j = \{\sigma_j\}$ , for some index  $\sigma_j$  in  $\{1, \dots, d\}$ , so that  $\alpha_{\sigma_j,1} = r_j$ .

For  $i = 1, \dots, d$ , let us write  $\nu_i$  for the degree of the minimal polynomial of  $X_1$  in  $Q_i$ ; thus, this polynomial is  $(T - \alpha_{i,1})^{\nu_i}$ . For  $j$  in  $\{1, \dots, c\}$ , we define  $m_j$  as the maximum of all  $\nu_i$ , for  $i$  in  $T_j$ . As a result, the minimal polynomial of  $X_1$  in  $\prod_{j \in T_j} Q_j$  is  $(T - r_j)^{m_j}$ , and the minimal polynomial of  $X_1$  in  $Q$  is  $M = \prod_{j \in \{1, \dots, c\}} (T - r_j)^{m_j}$ .

Recall that a linear form  $\ell : Q \rightarrow \overline{\mathbb{K}}$  can be written uniquely as  $\ell = \sum_{i \in \{1, \dots, d\}} \ell_i$ , with  $\ell_i : Q_i \rightarrow \overline{\mathbb{K}}$ ; collecting terms,  $\ell$  may also be written as  $\ell = \sum_{j \in \{1, \dots, c\}} \lambda_j$ , with  $\lambda_j = \sum_{i \in T_j} \ell_i$ . Given such an  $\ell$ , we first explain how to compute values of the form  $\lambda_j(1)$ . We will do this for some values of  $j$  only, namely those  $j$  for which  $m_j = 1$ .

**Lemma 5.1.** *Let  $\ell$  be in  $\text{hom}_{\mathbb{K}}(Q, \mathbb{K})$ , let  $M$  be the minimal polynomial of  $X_1$  in  $Q$  and let  $\delta$  be its degree. Then, the polynomial  $\Omega((\ell(X_1^s))_{s \geq 0}, M)$  satisfies*

$$\Omega((\ell(X_1^s))_{s \geq 0}, M)(r_j) = \lambda_j(1)M'(r_j) \quad \text{for all } j \text{ such that } m_j = 1.$$

*Proof.* Let  $\mathfrak{e}$  be the set of all indices  $j$  in  $\{1, \dots, c\}$  such that  $m_j = 1$ , and let  $\mathfrak{f} = \{1, \dots, c\} - \mathfrak{e}$ ; this definition allows us to split the sum as

$$\begin{aligned} \sum_{s \geq 0} \ell(X_1^s)/T^{s+1} &= \sum_{j \in \{1, \dots, c\}} \sum_{i \in T_j} \sum_{s \geq 0} \ell_i(X_1^s)/T^{s+1} \\ &= \sum_{j \in \mathfrak{e}} \sum_{i \in T_j} \sum_{s \geq 0} \ell_i(X_1^s)/T^{s+1} + \sum_{j \in \mathfrak{f}} \sum_{i \in T_j} \sum_{s \geq 0} \ell_i(X_1^s)/T^{s+1}. \end{aligned}$$

Using Lemma 2.4 with  $t = X_1$  and  $v = 1$ , any sum  $\sum_{s \geq 0} \lambda_j(X_1^s)/T^{s+1}$  in the second summand can be rewritten as

$$\frac{C_j}{(T - r_j)^{v_j}},$$

for some integer  $v_j$ , and for some polynomial  $C_j$  of degree less than  $v_j$ . Next, take  $j$  in  $\mathfrak{e}$ . Since  $m_j = 1$ ,  $\nu_i = 1$  for all  $i$  in  $T_j$ , so that each such  $\ell_i$  takes the form

$$\ell_i : f \mapsto (\Lambda_i(f))(\alpha_i),$$

where  $\Lambda_i$  is a differential operator that does not involve  $\partial/\partial X_1$ . Since all terms of positive order in  $\Lambda_i$  involve one of  $\partial/\partial X_2, \dots, \partial/\partial X_n$ , they cancel  $X_1^s$  for  $s \geq 0$ . Thus,  $\ell_i(X_1^s)$  can be rewritten as  $\ell_{i,1}\alpha_{i,1}^s$ , for some constant  $\ell_{i,1}$ , and the generating series of these terms is

$$\frac{\ell_{i,1}}{T - \alpha_{i,1}} = \frac{\ell_{i,1}}{T - r_j}.$$

Remarking that we can write  $\ell_{i,1} = \ell_i(1)$ , altogether, the sum in question can be written

$$\begin{aligned} \sum_{s \geq 0} \ell(X_1^s)/T^{s+1} &= \sum_{j \in \mathfrak{e}} \frac{\sum_{i \in T_j} \ell_i(1)}{T - r_j} + \sum_{j \in \mathfrak{f}} \frac{D_j}{(T - r_j)^{x_j}} \\ &= \sum_{j \in \mathfrak{e}} \frac{\lambda_j(1)}{T - r_j} + \sum_{j \in \mathfrak{f}} \frac{D_j}{(T - r_j)^{x_j}} \end{aligned}$$

for some integers  $\{x_j \mid j \in b\}$  such that  $\deg(D_j) < x_j$  holds, and with  $D_j$  and  $T - r_j$  coprime; if  $r_j = 0$ , we take  $x_j = \deg(D_j) + 1$ . In particular, the minimal polynomial of  $(\ell(X_1^s))_{s \geq 0}$  is  $N = \prod_{j \in \mathfrak{e}} (T - r_j) \prod_{j \in \mathfrak{f}} (T - r_j)^{x_j}$ .

On the other hand, the minimal polynomial  $M$  of  $X_1$  can be rewritten as  $M = \prod_{j \in \mathfrak{e}} (T - r_j) \prod_{j \in \mathfrak{f}} (T - r_j)^{m_j}$ , so that  $\delta = \sum_{j \in \mathfrak{e}} 1 + \sum_{j \in \mathfrak{f}} m_j$ . The minimal polynomial of the sequence  $\ell(X_1^s)$  divides  $M$ , so that  $x_j \leq m_j$  holds for all  $j$  in  $\mathfrak{f}$ . As a result,  $\Omega((\ell(X_1^s))_{s \geq 0}, M)$  exists and is given by

$$\begin{aligned} \Omega((\ell(X_1^s))_{s \geq 0}, M) &= \sum_{j \in \mathfrak{e}} \left( \lambda_j(1) \prod_{\iota \in \mathfrak{e} - \{j\}} (T - r_\iota) \right) \left( \prod_{j \in \mathfrak{f}} (T - r_j)^{m_j} \right) \\ &\quad + \sum_{j \in \mathfrak{f}} \left( (T - r_j)^{m_j - x_j} D_j \prod_{\iota \in \mathfrak{f} - \{j\}} (T - r_\iota)^{m_\iota} \right) \left( \prod_{j \in \mathfrak{e}} (T - r_j) \right). \end{aligned}$$

This implies that

$$\Omega((\ell(X_1^s))_{s \geq 0}, M)(r_k) = \lambda_k(1) \prod_{\iota \in \mathfrak{e} - \{k\}} (r_k - r_\iota) \prod_{j \in \mathfrak{f}} (r_k - r_j)^{m_j} = \lambda_k(1) M'(r_k)$$

holds for all  $k$  in  $\mathfrak{e}$ . □

We then show how to use this result to avoid (as much as possible) using a generic linear form  $t = t_1 X_1 + \dots + t_n X_n$ , and how to use (say)  $X_1$  instead to compute a zero-dimensional parametrization of a subset of  $V$ ; this is motivated by the fact that the multiplication matrix by  $X_1$  is expected to be sparser than that of  $t$  (since the matrix of  $t$  is a combination of those of  $X_1, \dots, X_n$ ), sometimes by a substantial amount. Of course, there is no guarantee that  $X_1$  is a separating element for  $V$ . As a result, we will compute a decomposition of  $V$  into two components  $V'$  and  $V''$ ;  $X_1$  will be a separating element for  $V'$ , whereas we will use a generic linear form to describe  $V''$ .

More precisely, we characterize the set  $V'$  mentioned above as follows: for  $i$  in  $\{1, \dots, d\}$ ,  $\alpha_i$  is in  $V'$  if and only if:

- for  $i'$  in  $\{1, \dots, d\}$ , with  $i' \neq i$ ,  $\alpha_{i',1} \neq \alpha_{i,1}$ ;
- $Q_i$  is a reduced algebra (equivalently,  $I_i$  is radical).

We denote by  $\mathfrak{A} \subset \{1, \dots, d\}$  the set of corresponding indices  $i$ , and we let  $\mathfrak{B} = \{1, \dots, d\} - \mathfrak{A}$ , so that we have  $V' = V_{\mathfrak{A}}$  and  $V'' = V_{\mathfrak{B}}$ . Remark that  $X_1$  is a separating element for  $V'$ .

Correspondingly, we define  $\mathfrak{a}$  as the set of all indices  $j$  in  $\{1, \dots, c\}$  such that  $\sigma_j$  is in  $\mathfrak{A}$ . In other words,  $j$  is in  $\mathfrak{a}$  if and only if  $T_j$  has cardinality 1 and  $Q_{\sigma_j}$  is reduced. The algorithm in this paragraph will compute a zero-dimensional parametrization of  $V_{\mathfrak{A}}$ ; we use the following lemma to perform this decomposition of  $V$ .

**Lemma 5.2.** *Let  $j$  be in  $\{1, \dots, c\}$  such that  $m_j = 1$ , let  $\lambda$  be a linear form over  $\prod_{i \in T_j} Q_i$  and let  $t = t_2 X_2 + \dots + t_n X_n$ . Define constants  $a, b, c$  in  $\overline{\mathbb{K}}$  by*

$$a = \lambda(1), \quad b = \lambda(t), \quad c = \lambda(t^2).$$

*Then,  $j$  is in  $\mathfrak{a}$  if and only if, for a generic choice of  $\lambda$  and  $t$ ,  $ac = b^2$ .*

*Proof.* The assumption that  $m_j = 1$  means that for all  $i$  in  $T_j$ ,  $\nu_i = 1$ . The linear form  $\lambda$  can be uniquely written as a sum  $\lambda = \sum_{i \in T_j} \ell_i$ , where each  $\ell_i$  is in  $\text{hom}_{\overline{\mathbb{K}}}(Q_i, \overline{\mathbb{K}})$ . The fact that all  $\nu_i$  are equal to 1 then implies that each  $\ell_i$  takes the form

$$\ell_i : f \mapsto (\Lambda_i(f))(\alpha_i),$$

where  $\Lambda_i$  is a differential operator that does not involve  $\partial/\partial X_1$ . Thus, as in Eq. (1), we can write a general  $\Lambda_i$  of this form as

$$\Lambda_i : f \mapsto u_{i,1}f + \sum_{2 \leq r \leq n} P_{i,r}(u_{i,2}, \dots, u_{i,D_i}) \frac{\partial}{\partial X_j} f + \sum_{2 \leq r \leq s \leq n} P_{i,r,s}(u_{i,2}, \dots, u_{i,D_i}) \frac{\partial^2}{\partial X_j \partial X_k} f + \tilde{\Lambda}_i(f),$$

where all terms in  $\tilde{\Lambda}_i$  have order at least 3,  $\mathbf{u}_i = (u_{i,1}, \dots, u_{i,D_i})$  are parameters and  $(P_{i,r})_{2 \leq r \leq n}$  and  $(P_{i,r,s})_{2 \leq r \leq s \leq n}$  are linear forms in  $u_{i,2}, \dots, u_{i,D_i}$ . We obtain

$$\begin{aligned} \Lambda_i(1) &= u_{i,1} \\ \Lambda_i(t) &= u_{i,1}t + \sum_{2 \leq r \leq n} P_{i,r}t_r \\ \Lambda_i(t^2) &= u_{i,1}t^2 + 2t \sum_{2 \leq r \leq n} P_{i,r}t_r + 2 \sum_{2 \leq r \leq s \leq n} P_{i,r,s}t_r t_s, \end{aligned}$$

which gives

$$\begin{aligned} a &= \sum_{i \in T_j} u_{i,1} \\ b &= \sum_{i \in T_j} u_{i,1}t(\alpha_i) + \sum_{i \in T_j, 2 \leq r \leq n} P_{i,r}t_r \\ c &= \sum_{i \in T_j} u_{i,1}t(\alpha_i)^2 + 2 \sum_{i \in T_j, 2 \leq r \leq n} t(\alpha_i)P_{i,r}t_r + 2 \sum_{i \in T_j, 2 \leq r \leq s \leq n} P_{i,r,s}t_r t_s. \end{aligned}$$

Suppose first that  $j$  is in  $\mathbf{a}$ . Then,  $T_j = \{\sigma_j\}$ , so we have only one term  $\Lambda_{\sigma_j}$  to consider, and  $Q_{\sigma_j}$  is reduced, so that all coefficients  $P_{\sigma_j,r}$  and  $P_{\sigma_j,r,s}$  vanish. Thus, we are left in this case with

$$a = u_{\sigma_j,1}, \quad b = u_{\sigma_j,1}t(\alpha_{\sigma_j}), \quad c = u_{\sigma_j,1}t(\alpha_{\sigma_j})^2,$$

so that we have  $ac = b^2$ , for *any* choice of  $\lambda$  and  $t$ . Now, we suppose that  $j$  is not in  $\mathbf{a}$ , and we prove that for a generic choice of  $\lambda$  and  $t$ ,  $ac - b^2$  is non-zero. The quantity  $ac - b^2$  is a polynomial in the coefficients  $(\mathbf{u}_i)_{i \in T_j}$ , and  $(t_i)_{i \in \{2, \dots, n\}}$ , and we have to show that it is not identically zero. We discuss two cases; in both of them, we prove that a suitable specialization of  $ac - b^2$  is non-zero.

Suppose first that for at least one index  $\sigma$  in  $T_j$ ,  $Q_\sigma$  is not reduced. In this case, there exists at least one index  $\rho$  in  $\{2, \dots, n\}$  such that  $P_{\sigma,\rho}(u_{\sigma,2}, \dots, u_{\sigma,D_\sigma})$  is not identically zero (**todo**: explain better). Let us set all  $\mathbf{u}_{\sigma'}$  to 0, for  $\sigma'$  in  $T_j - \{\sigma\}$ , as well as  $u_{\sigma,1}$ , and all  $t_r$  for  $r \neq \rho$ . Then, under this specialization,  $ac - b^2$  becomes  $-(P_{\sigma,\rho}(u_{\sigma,2}, \dots, u_{\sigma,D_\sigma})t_\rho)^2$ , which is non-zero, so that  $ac - b^2$  itself must be non-zero.

Else, since  $j$  is not in  $\mathbf{a}$ , we can assume that  $T_j$  has cardinality at least 2, with  $Q_\sigma$  reduced for all  $\sigma$  in  $T_j$  (so that  $P_{\sigma,r}$  and  $P_{\sigma,r,s}$  vanish for all such  $\sigma$  and all  $r, s$ ). Suppose that  $\sigma$  and  $\sigma'$  are two indices in  $T_j$ ; we set all indices  $u_{\sigma'',1}$  to zero, for  $\sigma''$  in  $T_j - \{\sigma, \sigma'\}$ . We are left with

$$a = u_{\sigma,1} + u_{\sigma',1}, \quad b = u_{\sigma,1}t(\alpha_\sigma) + u_{\sigma',1}t(\alpha_{\sigma'}), \quad c = u_{\sigma,1}t(\alpha_\sigma)^2 + u_{\sigma',1}t(\alpha_{\sigma'})^2.$$

Then,  $ac - b^2$  is equal to  $2u_{\sigma,1}u_{\sigma',1}(t(\alpha_\sigma) - t(\alpha_{\sigma'}))^2$ , which is non-zero, since  $\alpha_\sigma \neq \alpha_{\sigma'}$ .  $\square$

The previous lemmas allow us to write Algorithm ParametrizationX<sub>1</sub>. After computing  $M$ , we determine its factor  $P = \prod_{j \in \{1, \dots, c\}, m_j=1} (T - r_j)$ . We split this polynomial further using the previous results in order to find  $\prod_{j \in \mathbf{a}} (T - r_j)$ , and we conclude using the same kind of calculations as in ParametrizationGeneric.

---

**Algorithm 5** ParametrizationX<sub>1</sub>( $\ell, t$ )

---

**Input:**

- a linear form  $\ell$  over  $Q$
- a linear form  $t = t_2X_2 + \dots + t_nX_n$

**Output:** polynomials  $((P, V_1, \dots, V_n), X_1)$

1. let  $M$  be the minimal polynomial of the sequence  $(\ell(X_1^s))_{s \geq 0}$  and let  $\delta$  be its degree
  2. let  $P = \prod_{r \text{ root of } M \text{ of multiplicity } 1} (T - r)$
  3. let  $t$  be a random linear form in  $X_2, \dots, X_n$
  4. **for**  $i = 0, 1, 2$  **do**
    - (a) let  $A_i = \Omega((\ell(t^i X_1^s))_{s \geq 0}, M)$
  5. let  $P = \gcd(P, A_0A_2 - A_1^2)$
  6. **for**  $i = 2, \dots, n$  **do**
    - (a) let  $A_{X_i} = \Omega((\ell(X_i X_1^s))_{s \geq 0}, M)$
  7. **return**  $((P, T, A_{X_2}/A_1 \bmod P, \dots, A_{X_n}/A_1 \bmod P), X_1)$
-

**Lemma 5.3.** *Suppose that  $\ell$  is a generic element of  $\text{hom}_{\overline{\mathbb{K}}}(Q, \overline{\mathbb{K}})$  and that  $t$  is a generic linear form. Then the output  $((P, V_1, \dots, V_n), X_1)$  of  $\text{ParametrizationX}_1(\ell, t)$  is a zero-dimensional parametrization of  $V_{\mathfrak{A}}$ .*

*Proof.* Lemma 2.2 shows that for a generic choice of  $\ell$ ,  $M$  is the minimal polynomial of  $X_1$ , so that we indeed have  $P = \prod_{j \in \{1, \dots, c\}, m_j=1} (T - r_j)$ . Let then  $r_j$  be one of these roots; by Lemma 5.1, for  $i = 0, 1, 2$  we have  $A_i(r_j) = M'(r_j)(t^i \cdot \lambda_j)(1)$ , where  $\lambda_j = \sum_{i \in T_j} \ell_i$ , and the  $\ell_i$ 's are the components of  $\ell$ .

As a result, the value of  $A_0 A_2 - A_1^2$  at  $r_j$  is (up to the non-zero factor  $M'(r_j)^2$ ) equal to the quantity  $ac - b^2$  defined in Lemma 5.2, so for a generic choice of  $\ell$  and  $t$ , it vanishes if and only if  $j$  is in  $\mathfrak{a}$ . Thus, after Step 5,  $P$  is equal to  $\prod_{j \in \mathfrak{a}} (T - r_j)$ .

The last step is to compute the zero-dimensional parametrization of  $V_{\mathfrak{A}}$ . This is done using again Lemma 5.1. Indeed, for  $j$  in  $\mathfrak{a}$ ,  $T_j$  is simply equal to  $\{\sigma_j\}$ , so that we have, for  $i = 2, \dots, n$ ,

$$A_1(r_j) = M'(r_j)\lambda_j(1) \quad \text{and} \quad A_{X_i}(r_j) = M'(r_j)(X_i \cdot \lambda_j)(1) = M'(r_j)\lambda_j(X_i).$$

Now, since  $j$  is in  $\mathfrak{a}$ ,  $Q_{\sigma_j}$  is reduced, so that there exists a constant  $\lambda_{j,1}$  such that for all  $f$  in  $\overline{\mathbb{K}}[X_1, \dots, X_n]$ ,  $\lambda_j(f)$  takes the form  $\lambda_{j,1}f(\alpha_{\sigma_j})$ . This shows that, as claimed,

$$\frac{A_{X_j}(r_j)}{A_1(r_j)} = \frac{\Omega((\ell(X_j X_1^s))_{s \geq 0}, M)(r_j)}{\Omega((\ell(t X_1^s))_{s \geq 0}, M)(r_j)} = \frac{M'(r_j)\lambda_{j,1}\alpha_{j,i}}{M'(r_j)\lambda_{j,1}} = \alpha_{j,i}.$$

For  $i = 1$ , since we use  $X_1$  as a separating variable for  $V_{\mathfrak{A}}$ , we simply add the polynomial  $T$  to our list.  $\square$

## References

- [1] B. Beckermann and G. Labahn. A uniform approach for the fast computation of matrix-type Padé approximants. *SIAM J. Matrix Anal. Appl.*, 15(3):804–823, 1994.
- [2] A. Bostan, B. Salvy, and É. Schost. Fast algorithms for zero-dimensional polynomial systems using duality. *Appl. Algebra Engrg. Comm. Comput.*, 14:239–272, 2003.
- [3] D. S. Dummit and R. M. Foote. *Abstract Algebra*. John Wiley & Sons, 2004.
- [4] J.-C. Faugère and C. Mou. Sparse FGLM algorithms. *J. Symbolic Comput.*, 80(3):538–569, 2017.
- [5] J. Gathen and J. Gerhard. *Modern Computer Algebra*. Cambridge University Press, Cambridge, third edition, 2013.
- [6] P. Giorgi, C.-P. Jeannerod, and G. Villard. On the complexity of polynomial matrix computations. In *ISSAC’03*, pages 135–142. ACM, 2003.
- [7] P. Giorgi and R. Lebreton. Online order basis algorithm and its impact on the block Wiedemann algorithm. In *ISSAC’14*, pages 202–209. ACM, 2014.
- [8] C.-P. Jeannerod, V. Neiger, É. Schost, and G. Villard. Fast computation of minimal interpolation bases in Popov form for arbitrary shifts. In *ISSAC’16*, pages 295–302. ACM, 2016.
- [9] T. Kailath. *Linear Systems*. Prentice-Hall, 1980.
- [10] E. Kaltofen and G. Villard. On the complexity of computing determinants. In *ISSAC’01*, pages 13–27. ACM, 2001.
- [11] E. Kaltofen and G. Villard. On the complexity of computing determinants. *Comput. Complexity*, 13(3-4):91–130, 2004.
- [12] B. Mourrain. Isolated points, duality and residues. *Journal of Pure and Applied Algebra*, 117/118:469–493, 1997. Algorithms for algebra (Eindhoven, 1996).
- [13] V. M. Popov. Invariant description of linear, time-invariant controllable systems. *SIAM Journal on Control*, 10(2):252–264, 1972.
- [14] F. Rouillier. Solving zero-dimensional systems through the rational univariate representation. *Appl. Algebra Engrg. Comm. Comput.*, 9(5):433–461, 1999.
- [15] A. Steel. Direct solution of the (11,9,8)-minrank problem by the block Wiedemann algorithm in Magma with a Tesla GPU. In *PASCO’15*, pages 2–6. ACM, 2015.
- [16] A. Storjohann. High-order lifting and integrality certification. *J. Symbolic Comput.*, 36:613–648, 2003.

- [17] W. J. Turner. *Black box linear algebra with the LINBOX library*. PhD thesis, North Carolina State University, 2002.
- [18] M. Van Barel and A. Bultheel. A general module theoretic framework for vector M-Padé and matrix rational interpolation. *Numer. Algorithms*, 3:451–462, 1992.
- [19] G. Villard. Further analysis of Coppersmith’s block Wiedemann algorithm for the solution of sparse linear systems (extended abstract). In *ISSAC’97*, pages 32–39. ACM, 1997.
- [20] G. Villard. A study of Coppersmith’s block Wiedemann algorithm using matrix polynomials. Technical report, LMC-IMAG, Report 975 IM, 1997.
- [21] W. A. Wolovich. *Linear Multivariable Systems*, volume 11 of *Applied Mathematical Sciences*. Springer-Verlag New-York, 1974.
- [22] W. Zhou and G. Labahn. Efficient algorithms for order basis computation. *J. Symbolic Comput.*, 47(7):793–819, 2012.