Contents

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

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

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

$$\{(R_1(\alpha), R_2(\alpha), \ldots, R_n(\alpha)) | \forall \alpha : R_n(\alpha) = 0\}$$

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 Roillier's Rational Univariate Representation [8]. One such algorithm is the Sparse-FGLM, which can be seen as an application of the Wiedemann algorithm. In this paper, we will introduce a new algorithm inspired by Sparse-FGLM and block Wiedemann algorithm that is easily parallelizable.

1.1 Basic Operations

Over an 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 polylogarithm factors) [5].

2 Previous Algorithms

2.1 Wiedemann algorithm

The Wiedemann algorithm of [12] solves a system $\mathbf{A}y = b$, where $\mathbf{A}$ is an invertible square matrix. The most important aspect of this algorithm is that the minimal polynomial $P(x) = \sum_{i=0}^{D} c_i x^i$ of $\mathbf{A}$ can be computed through a scalar sequence

$$S = (u^{tr}\mathbf{A}^i b)_{i \geq 0}$$

for a random vector u. Once we have sufficient number of terms of s, we apply the Berlekamp-Massey algorithm, which efficiently computes $P(x)$ from s. In the simplest case, where $c_0 \neq 0$, we have that

$$\begin{aligned} & P(\mathbf{A}) = 0 = c_0\mathbf{I} + c_1\mathbf{A} + \cdots + c_D\mathbf{A}^D \\ \implies & -c_0\mathbf{I} = \sum_{i=1}^{D} c_i\mathbf{A}^i \\ \implies & \mathbf{I} = \mathbf{A}(-c_0^{-1}\sum_{i=1}^{D} c_i\mathbf{A}^{i-1}) \\ \implies & b = \mathbf{A}(-c_0^{-1}\sum_{i=1}^{D} c_i\mathbf{A}^{i-1})b \end{aligned}$$

Therefore, $x = (-c_0^{-1}\sum_{i=1}^{D} c_i\mathbf{A}^{i-1})b$.

2.2 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

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

it produces the lex Gröbner basis of I of the form $\{P(X_1), X_2 - R_2(X_1), \ldots, X_n - R_n(X_1)\}$. The key idea is that $R(X_1)$ 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, $R_j(x_1) = \frac{N_j}{N} \mod P$. Other versions of this algorithm exist to handle non-radical ideals using the Berlekamp-Massey-Sakata algorithm.

3 Previously Results on Generating Series

3.1 Overview

In what follows, $\mathbb{K}$ is a field. Let I be an ideal in $\mathbb{K}[X_1, \ldots, X_n]$ and $Q = \mathbb{K}[X_1, \ldots, X_n]/I$ be the associated residue class ring. Suppose that $V = V(I)$ has dimension zero, and write it as $V = \{\boldsymbol{\alpha}_1, \ldots, \boldsymbol{\alpha}_d\}$, with all $\boldsymbol{\alpha}_i$'s in $\overline{\mathbb{K}}^n$, and $\boldsymbol{\alpha}_i = (\alpha_{i,1}, \ldots, \alpha_{i,n})$ for all i. We also let D be the dimension of Q, so that $d \leq D$, and *we assume that* $\mathrm{char}(\mathbb{K})$ *is greater than* D. In this section, we recall and generalize results from the appendix of [1], 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 $1, t, t^2, \ldots$, 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 $\mathsf{Parametrization}(\ell, t)$

Intput:

- a linear form ℓ over Q
- $t = t_1 X_1 + \cdots + t_n X_n$

Output:

- polynomials $(P, (V_1, \ldots, V_n))$

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

The call to $\mathsf{ParametrizationX}_1(\ell)$ 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 ℓ (which in effect removes from V the points we just found) and apply $\mathsf{ParametrizationGeneric}(\ell', t)$, 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 $(F, (G_1, \ldots, G_n)$ (to use t as a separating variable instead), and performing the union of the two components V' and V''.

3.1.1 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 $p_0\ell_s + \cdots + p_m\ell_{s+m} = 0$ for all $s \geq 0$, where $p_0, \ldots, p_m$ are the coefficients of P; this is equivalent to $\mathrm{rev}(P)S$ being polynomial of degree less than m, with $Z = \sum_{s\geq 0} \ell_s T^s$ and $\mathrm{rev}(P) = T^m P(1/T)$. 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 $Z' = \sum_{s\geq 0} \ell_s/T^{i+1}$, then the sequence $(\ell_s)_{s\geq 0}$ is linearly recurrent if and only if $Z'P$ is a polynomial of degree less than P. Therefore, if $A' = Z'P$ and $B' = P$, then we can write $Z' = A'/B'$

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 1. *$\Omega(Z, P)$ is the numerator for the closed form of Z with respect to P if:*

- *$\Omega(Z, P)$ is a polynomial of degree less than P*
- *$S = \Omega(Z, P)/P$*

Let A and A' be defined as above, then $\Omega(Z, \mathrm{rev}(P)) = A$ and $\Omega(Z', P) = A'$. Finally, if Z and Z' defined over the same sequence with minimal polynomial P, then

$$\Omega(Z', P) = T^{deg(P)-1}\Omega(Z, \mathrm{rev}(P))(1/T)$$

This makes it possible to avoid computations using a series over $1/T$. We give an example of this using the Fibonacci sequence, which is well known to be linearly recurrent. Since the Fibonacci sequence $F = (1, 1, 2, 3, 5, 8, \ldots)$ is defined recursively as $F^{(i+2)} = F^{(i)} + F^{(i+1)}$, $F^{(0)} = 1$, it is easy to verify that the minimal polynomial of F is $P = 1 - T - T^2$. Define $Z = \sum_{i\geq 0} F^{(i)}T^s$ and $\mathrm{rev}(P) = T^2 - T - 1$, then $\Omega(Z, \mathrm{rev}(P)) = 1$; therefore,

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

Now, define $Z' = \sum_{i\geq 0} F^{(i)}/T^{i+1}$, then we have that

$$\Omega(Z', P) = T^{deg(P)-1}\Omega(Z, \mathrm{rev}(P))(1/T) = T$$
$$\implies Z' = \frac{T}{1-T-T^2}$$

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

Lemma 1. *Let t be in Q and let $P \in \mathbb{K}[T]$ be its minimal polynomial. For a generic choice of ℓ in $\hom_{\mathbb{K}}(Q, \mathbb{K})$, P is the minimal polynomial of the sequence $(\ell(t^s))_{s\geq 0}$.*

3.1.2 Structure of the dual.

For i in $\{1, \ldots, d\}$, let Q_i be the local algebra at $\boldsymbol{\alpha}_i$, that is $Q_i = \overline{\mathbb{K}}[X_1, \ldots, X_n]/I_i$, with I_i the $\mathfrak{m}_{\boldsymbol{\alpha}_i}$-primary component of I. By the Chinese Remainder Theorem, $Q \otimes_{\mathbb{K}} \overline{\mathbb{K}} = \overline{\mathbb{K}}[X_1, \ldots, X_n]/I$ is isomorphic to the direct product $Q_1 \times \cdots \times Q_d$. 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 $\boldsymbol{\alpha}_i$ is a non-singular root of I. We also let $D_i = \dim_{\mathbb{K}}(Q_i)$, so that we have $D_i \geq N_i$ and $D = D_1 + \cdots + D_d$.

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

$$\lambda_{i,j} : f \mapsto (\Lambda_{i,j}(f))(\boldsymbol{\alpha}_i),$$

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

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

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 $\boldsymbol{\alpha}_i$ is non-singular, we have $D_i = 1$, so there is only one function $\lambda_{i,j}$, namely $\lambda_{i,1}$, we write it $\lambda_{i,1}(f) = f(\boldsymbol{\alpha}_i)$.

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

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

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

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 μ in S_i, $\mu \neq (0,\dots,0)$. Explicily, we can write $P_{i,\mu} = \sum_{j\in\{1,\dots,D_i\} \text{ such that } \mu\in S_{i,j}} c_{i,j,\mu}U_{i,j}$.

Fix λ non-zero in $\hom_{\overline{\mathbb{K}}}(Q_i,\overline{\mathbb{K}})$. We can then define its *order* w and *symbol* π. The former is the maximum of all $|\mu| = \mu_1+\cdots+\mu_n$ for $\mu = (\mu_1,\dots,\mu_n)$ in S_i such that $P_{i,\mu}(\boldsymbol{u}_i)$ is non-zero; by [7, Lemma 3.3] we have $w \leq N_i - 1$. Then, we let

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

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

Lemma 2. *Fix i in $\{1,\dots,d\}$. For a generic choice of λ in $\hom_{\overline{\mathbb{K}}}(Q_i,\overline{\mathbb{K}})$, and of $t_1,\dots,t_n$ in $\overline{\mathbb{K}}^n$, $\pi_i(t_1,\dots,t_n)$ is non-zero.*

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

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

this is by construction a non-zero polynomial. Thus, for a generic choice of $\boldsymbol{u}_i = (u_{i,1},\dots,u_{i,D_i})$, that define a linear form λ in $\hom_{\overline{\mathbb{K}}}(Q_i,\overline{\mathbb{K}})$ as in (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 λ is $\pi = \sum_{\mu\in S_i,\ |\mu|=\Omega} P_{i,\mu}(\boldsymbol{u}_i)X_1^{\mu_1}\cdots X_n^{\mu_n}$, and $\pi(t_1,\dots,t_n)$ is then non-zero. □

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

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

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

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

3.1.3 A fundamental formula.

The following lemma gives an explicit form for a generating series of the form $\sum_{\ell\geq 0}\ell(vt^i)T^\ell$, for a linear form $\ell : Q \to \mathbb{K}$. A slightly less precise version of it is in [1].

Lemma 3. *Let ℓ be in $\hom_{\mathbb{K}}(Q,\mathbb{K})$, 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 $t = t_1X_1 + \cdots + t_nX_n$, 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

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

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 $f\mapsto \frac{\partial^{|\mu|}f}{\partial X_1^{\mu_1}\cdots\partial X_n^{\mu_n}}$, where we write $|\mu| = \mu_1+\cdots+\mu_n$. Then, we have the following generating series identities, with coefficients in $\mathbb{K}(X_1,\dots,X_n)$:

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

for some polynomials $P_1, \ldots, P_{|\mu|}, P$ in $\mathbb{K}[\boldsymbol{X}, T]$ that depend on the choices of μ, 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 $f \mapsto \sum_{\mu \in R} c_\mu \frac{\partial^{|\mu|} f}{\partial X_1^{\mu_1} \cdots \partial X_n^{\mu_n}}$. The corresponding generating series becomes

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

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

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

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

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

Applying this formula to the sum in (2), we obtain the claim in the lemma. □

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

- We consider ℓ and t as in Lemma 3, such that ℓ has support $\mathfrak{S}$.
- $\mathfrak{T}$ is the subset of $\mathfrak{S}$ consisting of all indices i such that
  - $\pi_i(t_1, \ldots, t_n)$ is non-zero;
  - $t(\boldsymbol{\alpha}_{i'}) \neq t(\boldsymbol{\alpha}_i)$ for $i' \neq i$ in $\mathfrak{S}$.
- $\{r_1, \ldots, r_c\}$ 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 $\{1, \ldots, c\}$ such that
  - the fiber $t^{-1}(r_j) \subset V_{\mathfrak{S}}$ contains a single point, written $\boldsymbol{\alpha}_{\sigma_j}$;
  - the point $\boldsymbol{\alpha}_{\sigma_j}$ is in $\mathfrak{T}$ (equivalently, $\pi_{\sigma_j}(t_1, \ldots, t_n)$ is non-zero).

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

Lemma 4. *Let ℓ, t and all other notation be as above. Let further M be the minimal polynomial of t in $Q_{\mathfrak{S}}$, let δ be its degree and let $B = T^\delta M(1/T)$. Suppose that M is also the minimal polynomial of the sequence $(\ell(t^s))_{s \geq 0}$. Then, the following holds:*

- *for v in $\mathbb{K}[X_1,\dots,X_n]$, the power series $C_v = \Big(\sum_{s\geq 0}\ell(vt^s)T^s\Big)B$ is actually a polynomial of degree less than δ;*
- *there exist non-zero constants $\{c_j \mid j\in\mathfrak{t}\}$ such that for v in $\mathbb{K}[X_1,\dots,X_n]$, the polynomial $\tilde{C}_v = T^{\delta-1}C_v(1/T)$ satisfies*

$$\tilde{C}_v(r_j) = c_j v(\boldsymbol{\alpha}_{\sigma_j}) \quad \textit{for all } j \textit{ in } \mathfrak{t}.$$

Proof. For $j=1,\dots,c$, we write T_j for the set of all indices i in $\mathfrak{S}$ such that $t(\boldsymbol{\alpha}_i)=r_j$; 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 (3) as

$$\begin{aligned}\sum_{s\geq 0}\ell(vt^s)T^s &= \sum_{j\in\{1,\dots,c\}}\sum_{i\in T_j}\frac{v(\boldsymbol{\alpha}_i)w_i!\,\pi_i(t_1,\dots,t_n)T^{w_i}+(1-r_jT)A_{v,i}}{(1-r_jT)^{w_i+1}}\\ &= \sum_{j\in\mathfrak{s}}\frac{v(\boldsymbol{\alpha}_{\sigma_j})w_{\sigma_j}!\,\pi_{\sigma_j}(t_1,\dots,t_n)T^{w_{\sigma_j}}+(1-r_jT)A_{v,\sigma_j}}{(1-r_jT)^{w_{\sigma_j}+1}}\\ &\quad+\sum_{j\in\{1,\dots,c\}-\mathfrak{s}}\frac{\sum_{i\in T_j}\Big(\big[v(\boldsymbol{\alpha}_i)w_i!\,\pi_i(t_1,\dots,t_n)T^{w_i}+(1-r_jT)A_{v,i}\big](1-r_jT)^{y_j-(w_i+1)}\Big)}{(1-r_jT)^{y_j}},\end{aligned}$$

where y_j is the maximum of all w_i for i in T_j. Remark that for $v=1$, our condition that π_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 simplyfing terms in the second sum, we can rewrite the expression above as

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

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 $\gcd(D_{v,j},(1-r_jT))=1$; 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

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

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 $\{\zeta_j \le w_{\sigma_j}+1 \mid j \in \mathfrak{t}\}$; 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 $(\ell(t^s))_{s\ge 0}$ 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 $(\ell(vt^s))_{s\ge 0}$ 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

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

and $B = T^\delta M(1/T)$ satisfies

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

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

$$\begin{aligned} C_v = &\sum_{j\in\mathfrak{t}} \Big(\big[v(\boldsymbol{\alpha}_{\sigma_j}) w_{\sigma_j}!\, \pi_{\sigma_j}(t_1,\dots,t_n) T^{w_{\sigma_j}} + (1-r_jT) A_{v,\sigma_j} \big] \prod_{\iota\in\mathfrak{t}-\{j\}} (1-r_\iota T)^{w_{\sigma_\iota}+1} \Big) \Big(\prod_{j\in\mathfrak{u}} (1-r_jT)^{z_j} \Big) \\ &+ \Big(\prod_{j\in\mathfrak{t}} (1-r_jT)^{w_{\sigma_j}+1} \Big) \sum_{j\in\mathfrak{u}_v} \Big(D_{v,j} (1-r_jT)^{z_j - z_{v,j}} \prod_{\iota\in\mathfrak{u}-\{j\}} (1-r_\iota T)^{z_\iota} \Big). \end{aligned}$$

Since $\delta-1$ is an upper bound on the degree of C_v, we can then define $\tilde{C}_v = T^{\delta-1} C_v(1/T)$, that is,

$$\begin{aligned} \tilde{C}_v = &\sum_{j\in\mathfrak{t}} \Big(\big[v(\boldsymbol{\alpha}_{\sigma_j}) w_{\sigma_j}!\, \pi_{\sigma_j}(t_1,\dots,t_n) + (T-r_j) \tilde{A}_{v,\sigma_j} \big] \prod_{\iota\in\mathfrak{t}-\{j\}} (T-r_\iota)^{w_{\sigma_\iota}+1} \Big) \prod_{j\in\mathfrak{u}} (T-r_j)^{z_j} \\ &+ \Big(\prod_{j\in\mathfrak{t}} (T-r_j)^{w_{\sigma_j}+1} \Big) \sum_{j\in\mathfrak{u}_v} \Big(\tilde{D}_{v,j} (1-r_jT)^{z_j - z_{v,j}} \prod_{\iota\in\mathfrak{u}-\{j\}} (T-r_\iota)^{z_\iota} \Big), \end{aligned}$$

with $\tilde{A}_{v,\sigma_j} = T^{w_{\sigma_j}-1} A_{v,\sigma_j}(1/T) \in \overline{\mathbb{K}}[T]$ for j in $\mathfrak{t}$ and and $\tilde{D}_{v,j} = T^{z_{v,j}-1} D_{v,j}(1/T)$ for j in $\mathfrak{u}_v$. In particular, for k in $\mathfrak{t}$, the value $\tilde{C}_v(r_k)$ is

$$\begin{aligned} \tilde{C}_v(r_k) &= v(\boldsymbol{\alpha}_{\sigma_k}) w_{\sigma_k}!\, \pi_{\sigma_k}(t_1,\dots,t_n) \prod_{\iota\in\mathfrak{t}-\{k\}} (r_\iota - r_k)^{w_{\sigma_\iota}+1} \prod_{j\in\mathfrak{u}} (r_j - r_k)^{z_k} \\ &= v(\boldsymbol{\alpha}_{\sigma_k}) c_k, \end{aligned}$$

with

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

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

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

Algorithm 2 ParametrizationGeneric(ℓ, t)

Input:

- a linear form ℓ over $Q_{\mathfrak{S}}$
- $t = t_1X_1 + \cdots + t_nX_n$

Output: polynomials $(P, V_1, \ldots, V_n)$

1. let M be the minimal polynomial of the sequence $(\ell(t^s))_{s\geq 0}$ and let δ be its degree
2. let P be the squarefree part of M
3. let $B = T^\delta M(1/T)$
4. let $C_1 = B\big(\sum_{s<\delta}\ell(t^s)T^s\big) \mod T^\delta$
5. let $\tilde{C}_1 = T^{\delta-1}C_1(1/T)$
6. **for** $i = 1,\ldots,n$ **do**
   (a) let $C_{X_i} = B\big(\sum_{s<\delta}\ell(X_it^s)T^s\big) \mod T^\delta$
   (b) let $\tilde{C}_{X_i} = T^{\delta-1}C_{X_i}(1/T)$
7. **return** $(P, \tilde{C}_{X_1}/\tilde{C}_1 \bmod P, \ldots, \tilde{C}_{X_n}/\tilde{C}_1 \bmod P)$

Lemma 5. *Suppose that ℓ is a generic element of $\hom_{\overline{\mathbb{K}}}(Q_{\mathfrak{K}}, \overline{\mathbb{K}})$ and that t is a generic linear form. Then the output $((P, V_1, \ldots, V_n), t)$ of* Parametrization(ℓ, 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 that for a generic choice of ℓ, $\pi_i(t_1,\ldots,t_n)$ vanishes for no i in $\mathfrak{S}$. As a result, $\mathfrak{T} = \mathfrak{S}$. Besides, we recall that for a generic ℓ in $\hom_{\overline{\mathbb{K}}}(Q_{\mathfrak{S}}, \overline{\mathbb{K}})$, the minimal polynomials of $(\ell(t^s))_{s\geq 0}$ and of t are the same (Lemma 1).

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 $i = 1,\ldots,n$, we have

$$\frac{\tilde{C}_{X_i}(r_j)}{\tilde{C}_1(r_j)} = \frac{c_j\alpha_{\sigma_j,i}}{c_j} = \alpha_{\sigma_j,i},$$

so that $\tilde{C}_{X_i}/\tilde{C}_1 \bmod P$ is the ith polynomial in the zero-dimensional parametrization of V corresponding to t. □

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

$$\ell(X_1^i) = 17 \cdot 1^i + 33 \cdot 3^i$$
$$\ell(X_2X_1^i) = 17 \cdot 4 \cdot 1^i + 33 \cdot 2 \cdot 3^i$$

We define an infinite series for both sequences

$$Z_1 = \sum_{i=0}^{\infty} \ell(X_1^i)/T^{i+1} = \frac{17}{T-1} + \frac{33}{T-3} = \frac{17(T-3) + 33(T-1)}{(T-1)(T-3)}$$
$$Z_2 = \sum_{i=0}^{\infty} \ell(X_2X_1^i)/T^{i+1} = \frac{17 \cdot 4}{T-1} + \frac{33 \cdot 2}{t-3} = \frac{17 \cdot 4(T-3) + 33 \cdot 2(T-1)}{(T-1)(T-3)}$$

S_1 and S_2 have a common denominator $P = (t - 1)(t - 3)$, whose roots are the coordinates of x_1 in $V(I)$. Now, let

$$\begin{aligned} R_2 &= \frac{\Omega(S_2, P)}{\Omega(S_1, P)} \mod P \\ &= \frac{17 \cdot 4(T-3) + 33 \cdot 2(T-1)}{17(T-3) + 33(T-1)} \mod P \\ &= \frac{4(T-3) + 2(T-1)}{(T-3) + (T-1)} \mod P \end{aligned}$$

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

3.1.4 Decomposition of linear forms.

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

For $i = 1, \ldots, d$, let us write ν_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 $\{1, \ldots, c\}$, we define m_j as the maximum of all ν_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 $M = \prod_{j \in \{1,\ldots,c\}} (T - r_j)^{m_j}$.

Recall that a linear form $\ell : Q \to \overline{\mathbb{K}}$ can be written uniquely as $\ell = \sum_{i\in\{1,\dots,d\}} \ell_i$, with $\ell_i : Q_i \to \overline{\mathbb{K}}$; collecting terms, ℓ 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 ℓ, 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 6. *Let ℓ be in $\hom_{\mathbb{K}}(Q, \mathbb{K})$, let M be the minimal polynomial of X_1 in Q, let δ be its degree and let $B = T^\delta M(1/T)$. Then, the following holds:*

- *the power series $A = \left(\sum_{s\geq 0} \ell(X_1^s)T^s\right)B$ is actually a polynomial of degree less than δ;*
- *the polynomial $\tilde{A} = T^{\delta-1}A(1/T)$ satisfies*

$$\tilde{A}(r_j) = \lambda_j(1)M'(r_j) \quad \text{for all } j \text{ such that } m_j = 1.$$

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

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

Using Lemma 3 with $t = X_1$ and $v = 1$, any sum $\sum_{s\geq 0} \lambda_j(X_1^s)T^s$ in the second summand can be rewritten as

$$\frac{C_j}{(1 - r_jT)^{v_j}},$$

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 ℓ_i takes the form

$$\ell_i : f \mapsto (\Lambda_i(f))(\boldsymbol{\alpha}_i),$$

where Λ_i is a differential operator that does not involve $\partial/\partial X_1$. Since all terms of positive order in Λ_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

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

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

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

for some integers $\{x_j \mid j \in b\}$ such that $\deg(D_j) < x_j$ holds, and with D_j and $1 - r_jT$ 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 $\delta = \sum_{j\in\mathfrak{e}} 1 + \sum_{j\in\mathfrak{f}} m_j$ and $B = \prod_{j\in\mathfrak{e}}(1 - r_jT)\prod_{j\in\mathfrak{f}}(1 - r_jT)^{m_j}$. 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, $A = \big(\sum_{s\geq 0}\ell(X_1^s)T^s\big)B$ is indeed a polynomial of degree less than δ, given by

$$\begin{aligned} A = &\sum_{j\in\mathfrak{e}}\Big(\lambda_j(1)\prod_{\iota\in\mathfrak{e}-\{j\}}(1 - r_\iota T)\Big)\Big(\prod_{j\in\mathfrak{f}}(1 - r_jT)^{m_j}\Big) \\ &+ \sum_{j\in\mathfrak{f}}\Big(D_j(1 - r_jT)^{m_j - x_j}\prod_{\iota\in\mathfrak{f}-\{j\}}(1 - r_jT)^{m_\iota}\Big)\Big(\prod_{j\in\mathfrak{e}}(1 - r_jT)\Big). \end{aligned}$$

The reciprocal polynomial $\tilde{A} = T^{\delta-1}A(1/T)$ is then

$$\begin{aligned} \tilde{A} = &\sum_{j\in\mathfrak{e}}\Big(\lambda_j(1)\prod_{\iota\in\mathfrak{e}-\{j\}}(T - r_\iota)\Big)\Big(\prod_{j\in\mathfrak{f}}(T - r_j)^{m_j}\Big) \\ &+ \sum_{j\in\mathfrak{f}}\Big(\tilde{D}_j(T - r_j)^{m_j - x_j}\prod_{\iota\in\mathfrak{f}-\{j\}}(T - r_j)^{m_\iota}\Big)\Big(\prod_{j\in\mathfrak{e}}(T - r_j)\Big), \end{aligned}$$

with $\tilde{D}_j = T^{x_j-1}D_j(1/T)$ for all j in $\mathfrak{f}$. This implies that

$$\tilde{A}(r_k) = \lambda_k(1)\prod_{\iota\in\mathfrak{e}-\{k\}}(r_k - r_\iota)\prod_{j\in\mathfrak{f}}(r_k - r_j)^{m_j} = \lambda_k(1)M'(r_k)$$

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 $t = t_1X_1 + \cdots + t_nX_n$, 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 $X_1, \ldots, X_n$), 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 $\{1, \ldots, d\}$, $\boldsymbol{\alpha}_i$ is in V' if and only if:

- for i' in $\{1, \ldots, d\}$, 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 $\mathfrak{A} \subset \{1, \ldots, d\}$ the set of corresponding indices i, and we let $\mathfrak{B} = \{1, \ldots, d\} - \mathfrak{a}$, 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 σ_j is in $\mathfrak{A}$. In other words, j is in $\mathfrak{a}$ if and only if T_j has cardinality 1 and Q_{σ_j} is reduced. The algorithm in this paragraph will compute a zero-dimenstional parametrization of $V_{\mathfrak{A}}$; we use the following lemma to perform this decomposition of V.

Lemma 7. *Let j be in $\{1,\dots,c\}$ such that $m_j = 1$, let λ be a linear form over $\prod_{i\in T_j} Q_i$ and let $t = t_2X_2 + \cdots + t_nX_n$. 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 λ 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 λ can be uniquely written as a sum $\lambda = \sum_{i\in T_j} \ell_i$, where each ℓ_i is in $\hom_{\overline{\mathbb{K}}}(Q_i, \overline{\mathbb{K}})$. The fact that all ν_i are equal to 1 then implies that each ℓ_i takes the form

$$\ell_i : f \mapsto (\Lambda_i(f))(\boldsymbol{\alpha}_i),$$

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

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

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

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

which gives

$$\begin{aligned}
a &= \sum_{i\in T_j} u_{i,1}\\
b &= \sum_{i\in T_j} u_{i,1}t(\boldsymbol{\alpha}_i) + \sum_{i\in T_j, 2\le r\le n} P_{i,r}t_r\\
c &= \sum_{i\in T_j} u_{i,1}t(\boldsymbol{\alpha}_i)^2 + 2\sum_{i\in T_j, 2\le r\le n} t(\boldsymbol{\alpha}_i)P_{i,r}t_r + 2\sum_{i\in T_j, 2\le r\le s\le n} P_{i,r,s}t_rt_s.
\end{aligned}$$

Suppose first that j is in $\mathfrak{a}$. Then, $T_j = \{\sigma_j\}$, so we have only one term Λ_{σ_j} to consider, and Q_{σ_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

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

so that we have $ac = b^2$, for *any* choice of λ and t. Now, we suppose that j is not in $\mathfrak{a}$, and we prove that for a generic choice of λ and t, $ac - b^2$ is non-zero. The quantity $ac - b^2$ is a polynomial in the coefficients $(\boldsymbol{u}_i)_{i \in T_j}$, 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 σ in T_j, Q_σ is not reduced. In this case, there exists as least one index ρ 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 $\boldsymbol{u}_{\sigma'}$ to 0, for σ' 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 $\mathfrak{a}$, we can assume that T_j has cardinality at least 2, with Q_σ reduced for all σ in T_j (so that $P_{\sigma,r}$ and $P_{\sigma,r,s}$ vanish for all such σ and all r, s). Suppose that σ and σ' are two indices in T_j; we set all indices $u_{\sigma'',1}$ to zero, for σ'' in $T_j - \{\sigma, \sigma'\}$. We are left with

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

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

The previous lemmas allow us to write Algorithm $\mathsf{ParametrizationX}_1$. 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 $\prod_{j \in \mathfrak{a}}(T - r_j)$, and we conclude using the same kind of calculations as in ParametrizationGeneric.

Algorithm 3 ParametrizationX$_1(\ell, t)$

Input:

- a linear form ℓ over Q
- a linear form $t = t_2X_2 + \cdots + t_nX_n$

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

1. let M be the minimal polynomial of the sequence $(\ell(X_1^s))_{s\geq 0}$ and let δ be its degree
2. let $P = \prod_{r \text{ root of } M \text{ of multiplicity } 1}(T - r)$
3. let $B = T^\delta M(1/T)$
4. let t be a random linear form in $X_2, \ldots, X_n$
5. **for** $i = 0, 1, 2$ **do**
   (a) let $A_i = \left(\sum_{s<\delta} \ell(t^i X_1^s)T^s\right)B \mod T^\delta$
   (b) let $\tilde{A}_i = T^{\delta-1}A_i(1/T)$
6. let $P = \gcd(P, \tilde{A}_0\tilde{A}_2 - \tilde{A}_1^2)$
7. **for** $i = 2, \ldots, n$ **do**
   (a) let $A_{X_i} = \left(\sum_{s<\delta} \ell(X_2X_1^s)T^s\right)B \mod T^\delta$
   (b) let $\tilde{A}_{X_i} = T^{\delta-1}A_{X_i}(1/T)$
8. **return** $((P, T, \tilde{A}_{X_2}/\tilde{A}_1 \bmod P, \ldots, \tilde{A}_{X_n}/\tilde{A}_1 \bmod P), X_1)$

Lemma 8. *Suppose that ℓ is a generic element of $\hom_{\mathbb{K}}(Q, \overline{\mathbb{K}})$ and that t is a generic linear form. Then the output $((P, V_1, \ldots, V_n), X_1)$ of* ParametrizationX$_1(\ell, t)$ *is a zero-dimensional parametrization of $V_{\mathfrak{A}}$.*

Proof. Lemma 1 shows that for a generic choice of ℓ, M is the minimal polynomial of X_1, so that we indeed have $P = \prod_{j\in\{1,\ldots,c\}, m_j=1}(T - r_j)$. Let then r_j be one of these roots; by Lemma 6, for $i = 0, 1, 2$ we have $\tilde{A}_i(r_j) = M'(r_j)(t^i \cdot \lambda_j)(1)$, where $\lambda_j = \sum_{i\in T_j} \ell_i$, and the ℓ_i's are the components of ℓ.

As a result, the value of $\tilde{A}_0\tilde{A}_2 - \tilde{A}_1^2$ at r_j is (up to the non-zero factor $M'(r_j)^2$) equal to the quantity $ac - b^2$ defined in Lemma 7, so for a generic choice of ℓ and t, it vanishes if and only if j is in $\mathfrak{a}$. Thus, after Step 6, 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 6. Indeed, for j in $\mathfrak{a}$, T_j is simply equal to $\{\sigma_j\}$, so that we have, for $i = 2, \ldots, n$,

$$\tilde{A}_1(r_j) = M'(r_j)\lambda_j(1) \quad \text{and} \quad \tilde{A}_{X_i}(r_j) = M'(r_j)(X_i \cdot \lambda_j)(1) = M'(r_j)\lambda_j(X_i).$$

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

$$\frac{\tilde{A}_{X_j}(r_j)}{\tilde{A}_1(r_j)} = \frac{M'(r_j)\lambda_{j,1}\alpha_{j,i}}{M'(r_j)\lambda_{j,1}} = \alpha_{j,i}.$$

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. □

3.2 Generating series with matrix coefficients

In Sparse-FGLM algorithm, generating the matrix sequence $L = (u^{tr}\mathbf{T}_1^i)_{(0 \le i < 2D)}$ is the bottleneck of this algorithm. The most efficient way to compute $L^{(i)}$ is to compute $L^{(i-1)}\mathbf{T}_1$; however, this requires the terms of L to be computed sequentially. Therefore, it is natural to consider using blocking methods; that is, using sequences of small matrices instead of scalar sequences. Computing each term of such sequences will take longer, but this is easily parallelizable. Therefore, if blocking reduces the number of terms needed, we can expect an overall speed up.

The idea of extending the Wiedemann algorithm to using blocking methods is due to Coppersmith [2]. The formal analysis of Coppersmith's algorithm was done by Kaltofen, Villard, and others [6][10]. As with the Wiedemann algorithm, we are mainly interested in computing the minimal polynomial of a matrix from a sequence.

Let S be a matrix sequence, then it is linearly recurrent if and only if there exits polynomial matrices $\mathbf{D}$ and $\mathbf{N}$ such that $\sum_{i \ge 0} S^{(i)} t^i = \mathbf{D}^{-1}\mathbf{N}$. Similarly, we can also define a generating polynomial which *cancels* the sequence S.

Definition 2. *A generating polynomial matrix of S is a polynomial with matrix coefficients $\mathbf{F} = \sum_{i=0}^{\nu} \mathbf{W}_i t^i$ that satisfies,*

$$\mathbf{W}_0 S^{(\alpha)} + \mathbf{W}_1 S^{(\alpha+1)} + \cdots + \mathbf{W}_\nu S^{(\alpha+\nu)} = 0, \forall \alpha \ge 0$$

We can find the coefficients $\mathbf{W}_i$ by solving the system

$$\begin{aligned}
0 &= \mathbf{W}_0 S^{(0)} + \mathbf{W}_1 S^{(1)} + \cdots + \mathbf{W}_\nu S^{(\nu)} \\
0 &= \mathbf{W}_0 S^{(1)} + \mathbf{W}_1 S^{(2)} + \cdots + \mathbf{W}_\nu S^{(\nu+1)} \\
&\vdots \\
0 &= \mathbf{W}_0 S^{(d)} + \mathbf{W}_1 S^{(d+1)} + \cdots + \mathbf{W}_\nu S^{(\nu+d)}
\end{aligned}$$

which is equivalent to finding the kernel of the block Hankel matrix

$$\begin{bmatrix} S^{(0)} & \cdots & S^{(\nu)} \\ \vdots & \ddots & \vdots \\ S^{(d)} & \cdots & S^{(\nu+d)} \end{bmatrix}$$

Many generating polynomial matrices are possible, but some may add additional factors when used in calculations. Thus, as with the scalar case, it is necessary to define minimality of generators.

Definition 3. *The generating matrix polynomial in Popov form for a matrix sequence is called the minimum generating matrix polynomial (see [6, definition 2.3])*

Definition 4. *A polynomial matrix* $\mathbf{F}$ *is in Popov form if*

-

There are several algorithms to compute the minimum generating matrix polynomial, but we will use a generalization of the Berlekamp-Massey algorithm [2]. TODO: add specification of pm-basis

Lastly, by theorem 2.7 of [3], the largest invariant factor of the minimum generating matrix polynomial is the minimum scalar generator of S.

4 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) [9]. Our algorithm computes the rest of the lex Gröbner basis directly.

Given:

$I \subset \mathbb{K}$:	zero dimensional ideal in shape position
$\mathbb{B} \subset \mathbb{K}[X_1, \ldots, X_n]/I$:	monomial basis of $\mathbb{K}[X_1, \ldots, X_n]/I$
D:	dimension of $\mathbb{B}$
$\mathbf{M}_1, \ldots, \mathbf{M}_n$:	multiplication matrices of $X_1, \ldots, X_n$ 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 $(R, R_1, \ldots, R_n)$ such that for all α that is a factor of R, $\{(R_1(\alpha), \ldots, R_n(\alpha))\} = V(I)$

As with the scalar case, we need to compute the minimal polynomial of $\mathbf{M}$. We choose an integer m and two matrices of random entries $\mathbf{U}, \mathbf{V} \in \mathbb{K}^{D \times m}$. Then, we generate a matrix sequence $S = (\mathbf{U}^{tr}\mathbf{M}^i\mathbf{V})_{0 \leq i < 2d}$ with $d = \lceil \frac{D}{m} \rceil$ and find the minimum generating polynomial matrix $\mathbf{F}^{U,M,V}$ of S by applying the matrix Berlekamp-Massey algorithm. We set P to be the largest invariant factor of $\mathbf{F}^{U,M,V}$ and R to be the square free part of P.

To find the matrix numerator, we do a product $\mathbf{N}^* = \mathbf{F}^{U,M,V} \sum_{i=0}^{d} (\mathbf{U}^{tr}\mathbf{M}^i e)/T^{i+1}$. We can find a scalar numerator by computing $N = [0 \ldots 0P](\mathbf{F}^{U,M,V})^{-1}\mathbf{N}^*$. To find $R_j(X_1)$, $1 \leq j \leq n$, we compute $\mathbf{N}_j^* = \mathbf{F}^{U,M,V} \sum_{i=0}^{d} (\mathbf{U}^{tr}\mathbf{M}^i\mathbf{M}_j e)/T^{i+1}$ and set $N^j = [0 \ldots 0P](\mathbf{F}^{U,M,V})^{-1}\mathbf{N}_j^*$. Finally, $R_j(x_1) = N_j/N \mod P$.

Algorithm 4 Block Sparse-FGLM($\mathbf{M}, \mathbf{M}_1, \ldots, \mathbf{M}_n, m$)

Input:

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

Output: polynomials $(R, R_1, \ldots, R_n)$

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

4.1 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 [6][11].

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 $s_1, \ldots, s_r$ be the invariant factors of $T\mathbf{I} - \mathbf{M}$, ordered in such a way that $s_r | s_{r-1} | \ldots | s_1$, and let $d_i = \deg(s_i)$ for all i; for $i > r$, we let $s_i = 1$, with $d_i = 0$. We define $\nu = d_1 + \cdots + d_m \leq D$ and $\delta = \lceil \nu/m \rceil \leq \lceil D/m \rceil$. We also denote by $\sigma_1, \cdots, \sigma_k$ 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 5. *For a generic choice of* $\mathbf{U}$ *and* $\mathbf{V}$*, we have:*

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

Proof. We denote by $\langle \mathbf{V} \rangle$ the vector space generated by the columns of $\mathbf{V}, \mathbf{MV}, \mathbf{M}^2\mathbf{V}, \ldots$. 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 [11, 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 $[\mathbf{Z}\ \mathbf{B}_V\mathbf{Z}\ \mathbf{B}_V^2\mathbf{Z}\ \cdots]$ 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 $\mathscr{Y}$ of indeterminates of size $D \times m$, and let $\mathbf{F}^{M,\mathscr{Y}}$ be the minimal generating polynomial of the "generic" sequence $(\mathbf{M}^i\mathscr{Y})_{i \geq 0}$. The notation $\langle \mathscr{Y} \rangle$ and $D_{\mathscr{Y}}$ are defined as above. In particular, by [11, Proposition 6.1], the minimal generating polynomial $\mathbf{F}^{M,\mathscr{Y}}$ has degree δ and determinantal degree ν.

Now, for a generic $\mathbf{V}$ in $\mathbb{K}^{D \times m}$, $D_V = D_{\mathscr{Y}}$. Indeed, $\langle \mathscr{Y} \rangle$ is the span of $K_{\mathscr{Y}} = [\mathscr{Y}\ \mathbf{M}\mathscr{Y}\ \cdots\ \mathbf{M}^{N-1}\mathscr{Y}]$, whereas $\langle \mathbf{V} \rangle$ is the span of $[\mathbf{V}\ \mathbf{MV}\ \cdots\ \mathbf{M}^{N-1}\mathbf{V}]$. Take a maximal non-zero minor μ of $K_{\mathscr{Y}}$; as soon as $\mu(Y) \neq 0$, we have equality of the dimensions. On the other hand, by [11, Lemma 4.3], for any $\mathbf{V}$ (including $\mathscr{Y}$), the degree of $\mathbf{F}^{M,V}$ is equal to the first index d such that $\dim(\mathrm{span}([\mathbf{V}\ \mathbf{MV}\ \cdots\ \mathbf{M}^{d-1}\mathbf{V}])) = D_V$. As a result, for generic $\mathbf{V}$, $\mathbf{F}^{M,V}$ and $\mathbf{F}^{M,\mathscr{Y}}$ have the same degree, that is, δ. The first item is proved.

We conclude by proving that for generic $\mathbf{U}, \mathbf{V}$, the invariant factors $\sigma_1, \ldots, \sigma_m$ of $\mathbf{F}^{U,M,V}$ are $s_1, \ldots, s_m$. By [6, Theorem 2.12], for any $\mathbf{U}$ and $\mathbf{V}$ in $\mathbb{K}^{D \times m}$, for $i = 1, \ldots, m$, the i^{th} invariant factor σ_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 $\mathrm{Hk}_{e,d}(\mathbf{V})$ be the block Hankel matrix

$$\mathrm{Hk}_{e,d}(\mathbf{V}) = \begin{bmatrix} \mathbf{I} \\ \mathbf{M} \\ \mathbf{M}^2 \\ \vdots \\ \mathbf{M}^{e-1} \end{bmatrix} \begin{bmatrix} \mathbf{V} & \mathbf{MV} & \mathbf{M}^2\mathbf{V} & \cdots & \mathbf{M}^{d-1}\mathbf{V} \end{bmatrix}$$

By [6, Eq. (2.6)], $\mathrm{rank}(\mathrm{Hk}_{e,d}(\mathbf{V})) = \deg(\det(\mathbf{F}^{M,V}))$ for $d \geq \deg(\mathbf{F}^{M,V})$ and $e \geq D$. We take $e = D$, so that $\mathrm{rank}(\mathrm{Hk}_{D,d}(\mathbf{V})) = \deg(\det(\mathbf{F}^{M,V}))$ for $d \geq \deg(\mathbf{F}^{M,V})$. On the other hand, the sequence $\mathrm{rank}(\mathrm{Hk}_{N,d}(\mathbf{V}))$ is constant for $d \geq D$; as a result, $\mathrm{rank}(\mathrm{Hk}_{D,D}(\mathbf{V})) = \deg(\det(\mathbf{F}^{T,V}))$. For the same reason, we also have $\mathrm{rank}(\mathrm{Hk}_{D,D}(\mathscr{Y})) = \deg(\det(F^{M,\mathscr{Y}}))$, so that for a generic $\mathbf{U}$, $\mathbf{F}^{M,V}$ and $\mathbf{F}^{M,\mathscr{Y}}$ have the same determinantal degree, that is, ν. 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. □

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 9. *Let $\boldsymbol{a}$ be defined as line 5 of algorithm 4, then $\boldsymbol{a}$ has polynomial entries.*

Proof. Let $\mathscr{D} = \mathbf{A}\mathbf{F}^{U,M,V}\mathbf{B}$ be the Smith normal form of $\mathbf{F}^{U,M,V}$ and $s_1, \cdots s_m$ be invariant factors of $\mathbf{F}^{U,M,V}$ such that $s_m | s_{m-1} | \cdots | s_1$. Let $[b_1, \cdots, b_m]$ be the last row of $\mathbf{B}$ and $w = [\frac{s_1 b_1}{s_m}, \frac{s_1 b_2}{s_{m-1}}, \cdots, \frac{s_1 b_{m-1}}{s_2}, b_m]$ (since $s_i | s_1$), then

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

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

□

Theorem 6. *Let $U = [u_1, u_2, \ldots, u_m]$ 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 $\mathbf{N}^* = \mathbf{F}^{U,M,V}\sum_{i=0}^{d-1}(\mathbf{U}^t\mathbf{M}^i v)/T^{i+1}$ with all terms of negative exponents removed and $N = a\mathbf{N}^*$, then $N = \Omega(\sum_{i\geq 0}(u_m^t\mathbf{M}^i v)/T^{i+1}, P)$.*

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 $(\mathbf{U}\mathbf{M}^i\mathbf{V})_{i\geq 0}$, $\mathbf{N}^{*'}$ is a polynomial matrix and does not have terms of negative exponents. This means any terms in Z with degree less than d must vanish in the product. Therefore,

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

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

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

Recall Ω from definition 1. By rewriting each $\sum_{i\geq 0} u_j\mathbf{M}^i v$ in its closed form, we get

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

By theorem 5, 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 9

$$\begin{aligned} a\mathbf{N}^* &= a\mathbf{F}^{U,M,V} \begin{bmatrix} \Omega(\sum_{i\geq 0} u_1^{tr}\mathbf{M}^i v/T^{i+1}, P)/P \\ \vdots \\ \Omega(\sum_{i\geq 0} u_m^{tr}\mathbf{M}^i v/T^{i+1}, P)/P \end{bmatrix} \\ &= [0,\cdots,0,P] \begin{bmatrix} \Omega(\sum_{i\geq 0} u_1^{tr}\mathbf{M}^i v/T^{i+1}, P)/P \\ \vdots \\ \Omega(\sum_{i\geq 0} u_m^{tr}\mathbf{M}^i v/T^{i+1}, P)/P \end{bmatrix} \\ &= \Omega(\sum_{i\geq 0} u_m^{tr}\mathbf{M}^i v/T^{i+1}, P) \end{aligned}$$

Therefore, $N = \Omega(\sum_{i\geq 0} u_m^{tr}\mathbf{M}^i v/T^{i+1}, P)$ as needed. □

Finally, we conclude by seeing what happens when we pick specific values for v and apply theorem 6. In line 3 of algorithm 4, we compute $\mathbf{N}^* = \mathbf{F}^{U,M,V}\sum_{i=0}^{d-1}\mathbf{U}^t\mathbf{M}e/T^{i+1}$, where e is the coordinate vector of 1 in $\mathbb{B}$. By applying theorem 6,

$$N = a\mathbf{N}^* = \Omega(\sum_{i\geq 0} u_m^{tr}\mathbf{M}^i e/T^{i+1}, P)$$

By construction, $\mathbf{M}^i e$ gives the coordinate vector of t^i in $\mathbb{B}$, so $\sum_{i=0}^{d-1} u_m^{tr}\mathbf{M}^i e/T^{i+1} = \sum_{i\geq 0}\ell_{u_m}(t^i)/T^{i+1}$, where ℓ_{u_m} is the linear form associated with u_m. Therefore,

$$N = \Omega(\sum_{i\geq 0} u_m^{tr}\mathbf{M}^i e/T^{i+1}, P) = \Omega(\sum_{i\geq 0}\ell(t^i)/T^{i+1}, P)$$

.

Similarly, in line 7.1 of algorithm 4, 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 6,

$$N_j = a\mathbf{N}_j^* = \Omega(\sum_{i\geq 0} u_m^{tr}\mathbf{M}^i\mathbf{M}_j e/T^{i+1}, P)$$

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,

$$N_j = \Omega(\sum_{i\geq 0} u_m^{tr}\mathbf{M}^i\mathbf{T}_j e/T^{i+1}, P) = \Omega(\sum_{i\geq 0}\ell_{u_m}(X_j t^i)/T^{i+1}, P)$$

By the correctness of algorithm 2, we have the proof of correctness for algorithm 4.

4.2 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:

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

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

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

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

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

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

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,

$$\begin{aligned} R_1(t) &= N_1/N \mod R \\ &= 61t^3 + 75t^2 + 85t + 23 \end{aligned}$$

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 $GF(101) \times GF(101)$: $(54, 79)$ and P has one factor in $GF(101)$: 53. Now, $R_1(53) = 54$ and $R_2(53) = 79$ as expected.

5 Experimental Results

References

[1] A. Bostan, B. Salvy, and É. Schost. Fast algorithms for zero-dimensional polynomial systems using duality. *Applicable Algebra in Engineering, Communication and Computing*, 14:239–272, 2003.

[2] D. Coppersmith. Solving linear equations over GF(2): block Lanczos algorithm. *Linear Algebra and its Applications*, 192:33–60, 1993.

[3] G. Yuhasz E. Kaltofen. On the matrix berlekamp-massey algorithm. *ACM Trans. Algor.*, 2013.

[4] J.-C. Faugère and C. Mou. Sparse FGLM algorithms. *Journal of Symbolic Computation*, 80(3):538–569, 2017.

[5] J. Gathen and J. Gerhard. *Modern Computer Algebra*. Cambridge University Press, Cambridge, third edition, 2013.

[6] E. Kaltofen and G. Villard. On the complexity of computing determinants. *Comput. Complexity*, 13(3-4):91–130, 2004.

[7] B. Mourrain. Isolated points, duality and residues. *Journal of Pure and Applied Algebra*, 117/118:469–493, 1997. Algorithms for algebra (Eindhoven, 1996).

[8] F. Rouillier. Solving zero-dimensional systems through the rational univariate representation. *Applicable Algebra in Engineering, Communication and Computing*, 9(5):433–461, 1999.

[9] Allan Steel. Direct solution of the (11,9,8)-minrank problem by the block wiedemann algorithm in magma with a tesla gpu. *PASCO'15*, 2015.

[10] G. Villard. Further analysis of coppersmith's block wiedemann algorithm for the solution of sparse linear systems. *ISSAC'97*, pages 32–39, 1997.

[11] G. Villard. A study of Coppersmith's block Wiedemann algorithm using matrix polynomials. Technical report, LMC-IMAG, Report 975 IM, 1997.

[12] D. Wiedemann. Solving sparse linear equations over finite fields. *IEEE Trans. Information Theory*, IT-32:54–62, 1986.