Contents

1 Introduction **2**

1.1 Basic Operations 2

2 Previous Algorithms **3**

2.1 Wiedemann algorithm 3

2.2 Sparse-FGLM Algorithm 3

3 Blocking **5**

3.1 Linearly recurrent sequences 5

3.2 Matrix sequences and generators 6

3.3 Computing the Minimal Polynomial 6

4 Block Sparse-FGLM algorithm **9**

4.1 Computing a scalar numerator 10

4.2 Structure of the dual 12

4.3 Using a generic separating element 13

4.4 Computing with X_1 16

5 Experimental Results **17**

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 [7]

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 $O\tilde{}(n)$ field operations ($O\tilde{}$ omits polylogarithm factors) [4].

2 Previous Algorithms

2.1 Wiedemann algorithm

The Wiedemann algorithm of [11] 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 [3] 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 $\mathbb{B}$
$\mathbf{T}_1, \ldots, \mathbf{T}_n$:	multiplication matrices of $x_1, \ldots, x_n$ respectively

it produces the lex Gröbner basis of I of the form $\{R_1(x_1), x_2 - R_2(x_1), \ldots, x_n - R_n(x_1)\}$. The key idea is that $R(x_1)$ is the minimal polynomial of the multiplication matrix $\mathbf{T}_1$, which we can find using the Wiedemann algorithm. We generate

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

where u is a random vector and e is the coordinate vector for 1 in $\mathbb{B}$. Then, we find the minimum generating polynomial $P(t)$ by applying the Berlekamp-Massey algorithm on S. If we rewrite $P(t)$ in x_1, we get $R_1(x_1)$ We compute the numerator n of the generating series $Z = \sum_{i=0}^{D} s^{(i)}/t^{i+1}$ by a product $N = PZ$. To find $R_j(x_1)$, $2 \leq j \leq n$, we generate $s_j = (u^{tr}\mathbf{T}_1^i\mathbf{T}_j e)_{(0\leq i<D)}$ along with the numerator N_j of the generating series $Z_j = \sum_{i=0}^{D} s_j^{(i)}/t^{i+1}$ by a product $N_j = PZ_j$. Finally, $R_j(x_1) = \frac{N_j}{N} \mod P$. Also, other versions of this algorithm exist to handle non-radical ideals using the Berlekamp-Massey-Sakata algorithm.

3 Blocking

In Sparse-FGLM algorithm, generating the matrix sequence $L = (u^{tr}\mathbf{T}_1^i)_{(0\leq i<2D)}$ 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 [5][9]. As with the Wiedemann algorithm, we are mainly interested in computing the minimal polynomial of a matrix from a sequence. In this section, we will review basic terminology and definitions as well as present a proof that choosing the blocking matrices generically will produce the correct output, which had previously been proven by Kaltofen and Villard.

3.1 Linearly recurrent sequences

We first review basic facts on linearly recurrent sequences. Consider a sequence $(\ell_i)_{i\geq 0} \in \mathbb{K}^{\mathbb{N}}$ and the associated generating series $S = \sum_{i\geq 0} \ell_i t^i \in \mathbb{K}[[t]]$. The sequence (ℓ_i) is linearly recurrent if and only if its generating series is *rational*- that is, if there exist polynomials N, D in $\mathbb{K}[t]$ such that $S = N/D$; these polynomials are unique if we assume $\gcd(N, D) = 1$ and $D(0) = 1$. When this is the case, given a degree bound δ on such N and D, we can recover them by means of a rational reconstruction algorithm [4]. In the same vein, we say that a degree m polynomial $P \in \mathbb{K}[t]$ *cancels* a sequence ℓ if $p_0\ell_i + \cdots + p_m\ell_{i+d} = 0$ for all $i \geq 0$, where $p_0, \ldots, p_m$ are the coefficients of P; this is equivalent to $\mathrm{rev}(P)S$ being a polynomial, with $S = \sum_{i\geq 0} \ell_i t^i$ and $\mathrm{rev}(P) = t^m P(1/t)$.

Lastly, suppose we are given a generating series in q, $\sum_{i\geq 0} c^i q^i$ and we want to rewrite it in terms of $t = \frac{1}{q}$. We have that

$$\begin{aligned}
\sum_{i\geq 0} c^i q^i &= \frac{1}{1-cq} \\
\implies \sum_{i\geq 0} c^i/t^i &= \frac{1}{1-c(1/t)} \\
&= \frac{t}{t-c} \\
\implies \sum_{i\geq 0} c^i/t^{i+1} &= \frac{1}{t-c}
\end{aligned}$$

3.2 Matrix sequences and generators

In this section, we will extend the idea of linearly recurrent sequences to matrix sequences. 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 $\sum_{i\geq 0} S^{(i)}t^i = \mathbf{D}^{-1}\mathbf{N}$. Similarly, we can also define a generating polynomial which *cancels* the sequence S.

Definition 1. *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,*

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

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 2. *The generating matrix polynomial in Popov form for a matrix sequence is called the minimum generating matrix polynomial (see [5, definition 2.3])*

Definition 3. *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

3.3 Computing the Minimal Polynomial

Let $\mathbf{A}$ be in $\mathbb{K}^{D\times D}$ and $\mathbf{U}, \mathbf{V} \in \mathbb{K}^{D\times M}$ be two blocking matrices for some $M \leq D$. Now, define two matrix sequences $S_V = (\mathbf{A}^i\mathbf{V})_{(i\geq 0)}$ and $S_{U,V} = (\mathbf{U}^{tr}\mathbf{A}^i\mathbf{V})_{(i\geq 0)}$ denote their minimum generating matrix polynomial as $\mathbf{F}^{A,V}$ and $\mathbf{F}^{U,A,V}$ respectively. As with the Wiedemann algorithm, we may not be able to find the minimal polynomial of $\mathbf{A}$ from $S_{U,V}$ for some unlucky choices of $\mathbf{U}$ and $\mathbf{V}$. In this section, we will sketch a proof that a generic choice of $\mathbf{U}$ and $\mathbf{V}$ will produce the correct result.

Let $s_1, \dots, s_r$ be the invariant factors of $tI - \mathbf{A}$, 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 $\nu = d_1 + \cdots + d_M \leq D$ and $\delta = \lceil \nu/M \rceil \leq \lceil D/M \rceil$. We also denote by $\sigma_1, \cdots, \sigma_t$ the invariant factors of $\mathbf{F}^{X,A,Y}$, for some $t \leq M$. As above, for $i > t$, we let $\sigma_i = 1$.

Theorem 4. *For a generic choice of X and Y, we have:*

- *$F_X^{A,Y}$ has degree δ;*
- *$s_i = \sigma_i$ for $1 \leq i \leq M$.*

Proof. We denote by $\langle Y \rangle$ the vector space generated by the columns of $Y, AY, A^2Y, \dots$. We also write $N_Y = \dim(\langle Y \rangle)$.

First, we prove that for any Y in $\mathbb{K}^{N \times M}$, for a generic X in $\mathbb{K}^{N \times M}$, $F_X^{A,Y} = F^{A,Y}$. Indeed, by [10, Lemma 4.2], there exists matrices P_Y in $\mathbb{K}^{N \times N_Y}$ and $A_Y \in \mathbb{K}^{N_Y \times N_Y}$, with P_Y of full rank N_Y, and where A_Y is a matrix of the restriction of A to $\langle Y \rangle$, such that $F_X^{A,Y} = F^{A,Y}$ if and only if the dimension of the span of $[Z \; B_Y Z \; B_Y^2 Z \; \cdots]$ is equal to N_Y, with $B_Y = A_Y^\perp$ and $Z = P_Y^\perp X \in \mathbb{K}^{N_Y \times M}$.

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

Let us next introduce a matrix $\mathscr{Y}$ of indeterminates of size $N \times M$, and let $F^{A,\mathscr{Y}}$ be the minimal generating polynomial of the "generic" sequence $(A^i \mathscr{Y})_{i \geq 0}$. The notation $\langle \mathscr{Y} \rangle$ and $N_{\mathscr{Y}}$ are defined as above. In particular, by [10, Proposition 6.1], the minimal generating polynomial $F^{A,\mathscr{Y}}$ has degree δ and determinantal degree ν.

Now, for a generic Y in $\mathbb{K}^{N \times M}$, $N_Y = N_{\mathscr{Y}}$. Indeed, $\langle \mathscr{Y} \rangle$ is the span of $[\mathscr{Y} \; A\mathscr{Y} \; \cdots \; A^{N-1}\mathscr{Y}]$, whereas $\langle Y \rangle$ is the span of $[Y \; AY \; \cdots \; A^{N-1}Y]$. 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 [10, Lemma 4.3], for any Y (including $\mathscr{Y}$), the degree of $F^{A,Y}$ is equal to the first index d such that $\dim(\mathrm{span}([Y \; AY \; \cdots \; A^{d-1}Y])) = N_Y$. As a result, for generic Y, $F^{A,Y}$ and $F^{A,\mathscr{Y}}$ have the same degree, that is, δ. The first item is proved.

We conclude by proving that for generic X, Y, the invariant factors $\sigma_1, \dots, \sigma_M$ of $F_X^{A,Y}$ are $s_1, \dots, s_M$. By [5, Theorem 2.12], for any X and Y in $\mathbb{K}^{N \times M}$, for $i = 1, \dots, M$, the i^{th} invariant factor σ_i of $F_X^{A,Y}$ divides s_i, so that $\deg(\det(F_X^{A,Y})) \leq \nu$, with equality if and only if $\sigma_i = s_i$ for all $i \leq M$.

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

$$\mathrm{Hk}_{e,d}(Y) = \begin{bmatrix} I \\ A \\ A^2 \\ \vdots \\ A^{e-1} \end{bmatrix} \begin{bmatrix} Y & AY & A^2Y & \cdots & A^{d-1}Y \end{bmatrix}$$

By [5, Eq. (2.6)], $\mathrm{rank}(\mathrm{Hk}_{e,d}(Y)) = \deg(\det(F^{A,Y}))$ for $d \geq \deg(F^{A,Y})$ and $e \geq N$. We take $e = N$, so that $\mathrm{rank}(\mathrm{Hk}_{N,d}(Y)) = \deg(\det(F^{A,Y}))$ for $d \geq \deg(F^{A,Y})$. On the other hand, the sequence $\mathrm{rank}(\mathrm{Hk}_{N,d}(Y))$ is constant for $d \geq N$; as a result, $\mathrm{rank}(\mathrm{Hk}_{N,N}(Y)) = \deg(\det(F^{A,Y}))$. For the same reason, we also have $\mathrm{rank}(\mathrm{Hk}_{N,N}(\mathscr{Y})) = \deg(\det(F^{A,\mathscr{Y}}))$, so that for a generic Y, $F^{A,Y}$ and $F^{A,\mathscr{Y}}$ have the same determinantal degree, that is, ν. As a result, for generic X and Y, we also have $\deg(\det(F_X^{A,Y})) = \nu$, and the conclusion follows. □

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

Finally, we show that if $\mathbf{Z} = \sum_{i=0}^{\infty}(\mathbf{X}^{tr}\mathbf{A}^i\mathbf{Y})/t^{i+1}$ and $\mathbf{Z} = (\mathbf{F}^{X,A,Y})^{-1}\mathbf{N}$, then for every row of $\mathbf{N}$, its row degree is strictly less than the corresponding row degree of $\mathbf{F}^{X,A,Y}$. Since the highest power of the entries in $\mathbf{Z}$ is t^{-1}, the entries of the product $\mathbf{F}^{X,A,Y}\mathbf{Z}$ must have degree of at most $\delta - 1$. Furthermore, since the sequence $(\mathbf{A}^i)_{i\geq 0}$ is linearly recurrent, both $\mathbf{N}$ and $\mathbf{F}^{X,A,Y}$ are polynomial matrices; therefore, any terms of $\mathbf{Z}$ with degree less than $t^{-\delta}$ will vanish in the product. This shows that once we have $\mathbf{F}^{X,A,Y}$, we can compute $\mathbf{N}$ with $\lceil D/M \rceil \geq \delta$ terms of $(\mathbf{X}\mathbf{A}^i\mathbf{Y})_{i\geq 0}$.

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 used the Block Wiedemann algorithm to compute the minimal polynomial of T_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) [8]. 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 $\mathbb{B}$
$T_1, \ldots, T_n$:	multiplication matrices of $x_1, \ldots, x_n$ respectively
x:	random linear combination of x_i's
T:	multiplication matrix of x

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 x which, generically, separates the points in the variety. We assume that the base field $\mathbb{K}$ has characteristic larger than D.

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

To find the matrix numerator, we do a product $N^* = G\sum_{i=0}^{d} \frac{S^{(i)}}{t^{i+1}}$. We can find a scalar numerator N by taking the first entry of $[0\ldots 0P]G^{-1}N^*$. To find $R_j(x_1)$, $1 \le j \le n$, we generate $S_j^{(i)} = (U^{tr}T^iT_jV)_{0\le id}$, compute $N_j^* = G\sum_{i=0}^{d} \frac{S^{(i)}}{t^{i+1}}$, and set N^j as the first entry of $[0\ldots 0P]G^{-1}N_j^*$. Finally, $R_j(x_1) = N_j/N \mod P$

Block Sparse-FGLM:

```
1. choose $U, V \in \mathbb{K}^{m\times D}$
2. $S = (UT^iV^t)_{0\le i<2d}$, with $d = \frac{D}{m}$
3. $G = $ MatrixBerlekampMassey$(S)$ and $N^* = G\sum_{i\ge 0} \frac{S^{(i)}}{t^{i+1}}$
4. $P = $ largest invariant factor of $G$ and $R = $ SquareFreePart$(P)$
5. $a = [0\ \cdots 0P]G^{-1}$
6. $N = $ first entry of $aN^*$
7. for $j = 1\ldots n$:
```

7.1. $S_j = (UT_1^iT_jV^t)_{0\le i<d}$ and $N_j^* = G\sum_{i\ge 0}\frac{S_j^{(i)}}{t^{i+1}}$

7.2. $N_j =$ first entry of aN_j^*

7.3. $R_j = N_j/N$ mod R

We will demonstrate how step 7.3 computes R_j 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 linear form

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

(later, we will prove that for a radical ideal I, every linear form looks like $\ell = c_1f(\alpha_1)+\cdots+c_nf(\alpha_n)$ for all $\alpha\in V(I)$). 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 can define a generating series for both sequences,

$$S_1 = \sum_{i=0}^{\infty}\ell(x_1^i)/t^{i+1} = \frac{17}{t-1}+\frac{33}{t-3} = \frac{17(t-3)+33(t-1)}{(t-1)(t-3)}$$
$$S_2 = \sum_{i=0}^{\infty}\ell(x_2x_1^i)/t^{i+1} = \frac{17\cdot 4}{t-1}+\frac{33\cdot 2}{t-3} = \frac{17\cdot 4(t-3)+33\cdot 2(t-1)}{(t-1)(t-3)}$$

S_1 and S_2 have a common denominator $P = (t-1)(t-3)$, whose roots are the coordinates of x_1 in $V(I)$. If we apply step 7.3, we get

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

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

4.1 Computing a scalar numerator

We first prove that one can recover a scalar numerator through a matrix numerator.

Lemma 5. *Let $Z = \sum_{i=0}^{\infty} S^{(i)}/t^{i+1}$, then each entry of Z is in the form $N_{i,j}/P$, where P is the minimum scalar generator*

Proof. Rewrite U and V as $U = [u_1 u_2 \cdots u_M], V = [v_1 v_2 \cdots v_M]$, then

$$S^{(i)} = \begin{bmatrix} u_1^{tr} T^i v_1 & u_1^{tr} T^i v_2 & \cdots & u_1^{tr} T^i v_M \\ u_2^{tr} T^i v_1 & \cdots & \cdots & u_2^{tr} T^i v_M \\ \vdots & \ddots & \ddots & \vdots \\ u_M^{tr} T^i v_1 & \cdots & \cdots & u_M^{tr} T^i v_M \end{bmatrix}$$

Thus,

$$Z = \begin{bmatrix} \sum u_1^{tr} T^i v_1 / t^{i+1} & \cdots & \cdots & \sum u_1^{tr} T^i v_M / t^{i+1} \\ \vdots & \ddots & \ddots & \vdots \\ \sum u_M^{tr} T^i v_1 / t^{i+1} & \cdots & \cdots & \sum u_M^{tr} T^i v_M / t^{i+1} \end{bmatrix}$$

Since each entry of Z is a scalar sequence with the same minimum generator, we can rewrite them in their closed form $N_{i,j}/P$. □

Lemma 6. *Let $\mathscr{D} = AGB$ be the Smith normal form of G and $s_1, \cdots s_M$ be invariant factors of G such that $s_M | s_{M-1} | \cdots | s_1$. Then, there exists a vector $\tilde{u}$ such that $\tilde{u}G = [0, \cdots, 0, s_1]$*

Proof. Let $[b_1, \cdots, b_M]$ be the last row of 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

$$\begin{aligned} (wA)A^{-1}\mathscr{D} &= [\frac{s_1 b_1}{s_M}, \frac{s_1 b_2}{s_{M-1}}, \cdots, \frac{s_1 b_{M-1}}{s_2}, b_M] \begin{bmatrix} s_M & & \\ & \ddots & \\ & & s_1 \end{bmatrix} \\ &= [s_1 b_1, s_1 b_2, \cdots, s_1 b_M] \\ &= [0, \cdots, 0, s_1] B \end{aligned}$$

Therefore, if we choose $\tilde{u} = wA$, we get $\tilde{u}G = (wA)A^{-1}\mathscr{D}B^{-1} = [0, \cdots, 0, s_1]$ as needed. □

Theorem 7. *The first entry of the last row of $\tilde{u}N^*$ is the numerator of the generating function for $(u_M^{tr} T^i v_1)_{i \geq 0}$*

Proof. By lemma 5 ,

$$N^* = G \sum_{i=0}^{\infty} S^{(i)} / t^{i+1} = G \begin{bmatrix} N_{1,1}/P & \cdots & N_{1,M}/P \\ \vdots & \ddots & \vdots \\ N_{M,1}/P & \cdots & N_{M,M}/P \end{bmatrix}$$

By theorem 4, the i^{th} invariant factor of $tI - A$ is equal to the i^{th} invariant factor of G for generic choice of U, V. Thus, $s_1 = P$ and by lemma 6

$$\begin{aligned}
\tilde{u}N^* &= \tilde{u}G \begin{bmatrix} N_{1,1}/P & \cdots & N_{1,M}/P \\ \vdots & \ddots & \vdots \\ N_{M,1}/P & \cdots & N_{M,M}/P \end{bmatrix} \\
&= [0, \cdots, 0, P] \begin{bmatrix} N_{1,1}/P & \cdots & N_{1,M}/P \\ \vdots & \ddots & \vdots \\ N_{M,1}/P & \cdots & N_{M,M}/P \end{bmatrix} \\
&= [N_{M,1}, \cdots, N_{M,M}]
\end{aligned}$$

Since P is known to be the minimum generator of the sequences, we have that $N_{M,1}$ is the numerator of generating function for $(u_M^{tr} T^i v_1)_{i \geq 0}$. □

By following the same steps, we also have that the first entry of $\tilde{u}N_j^*$ is the product of the generating function for $(u_M^{tr} T^i T_j v_1)_{i \geq 0}$ with P.

4.2 Structure of the dual

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 $V = \{\boldsymbol{\alpha}_1, \ldots, \boldsymbol{\alpha}_d\}$, 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].

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 is 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 α_i is a non-singular root of I. We also let $D_i = \dim_{\overline{\mathbb{K}}}(Q_i) \geq N_i$, so that $D = D_1 + \cdots + D_d$. Fix i in $1, \ldots, d$. There exists a basis of the dual $\hom_{\overline{\mathbb{K}}}(Q_i, \overline{\mathbb{K}})$ consisting of linear forms $(\lambda_{i,j})_{1 \leq j \leq D_i}$ of the form

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

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 constants $c_{i,j,\mu}$ in $\overline{\mathbb{K}}$. For instance, when $\boldsymbol{\alpha}_i$ is non-singular, there is (up to a scalar multiple) only one function $\lambda_{i,j}$, say $\lambda_{i,1}$, and it takes the form $\lambda_{i,1}(f) = f(\boldsymbol{\alpha}_i)$.

More generally, we will always take $\lambda_{i,1}$ of the form $\lambda_{i,1}(f) = f(\boldsymbol{\alpha}_i)$, and for $j > 1$, and $\mu = (0, \ldots, 0)$, we set $c_{i,j,\mu} = 0$ (so all terms in $\Lambda_{i,j}$ have order 1 or more). Thus, introducing parameters $\ell_{i,1}$ and $\boldsymbol{\ell}_i = (\ell_{i,m})_{m=2,\dots,D_i}$, we deduce the existence of homogeneous linear forms $P_{i,\mu}$ such that any λ in $\hom_{\overline{\mathbb{K}}}(Q_i, \overline{\mathbb{K}})$ can be written as

$$\lambda : f \mapsto (\Lambda(f))(\boldsymbol{\alpha}_i), \quad \text{with} \quad \Lambda(f) = \sum_{\mu=(\mu_1,\dots,\mu_n)\in S_i} c_\mu \frac{\partial^{\mu_1+\cdots+\mu_n} f}{\partial X_1^{\mu_1} \cdots \partial X_n^{\mu_n}},$$

where S_i is the union of $S_{i,1}, \dots, S_{i,D_i}$, and where the coefficients c_μ can be descibed in parametric manner as $c_{(0,\dots,0)} = \ell_{i,1}$ and $c_\mu = P_{i,\mu}(\boldsymbol{\ell}_i)$ for all μ in S_i, $\mu \neq (0, \dots, 0)$.

If λ is non-zero, we can then define its *order* ω 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 c_μ is non-zero; by [6, Lemma 3.3] we have $\omega \leq N_i - 1$. Then, we let

$$\pi = \sum_{\mu \in S_i,\ |\mu|=\omega} c_\mu X_1^{\mu_1} \cdots X_n^{\mu_n}$$

be the *symbol* of λ; by construction, this is a non-zero polynomial.

Finally, we say a word about global objects. Fix a linear form $\lambda : Q \to \mathbb{K}$. By the Chinese Remainder Theorem, there exist unique $\lambda^{(1)}, \dots, \lambda^{(d)}$, with $\lambda^{(i)}$ in $\hom_{\overline{\mathbb{K}}}(Q_i, \overline{\mathbb{K}})$ for all i, such that the extension $\lambda_{\overline{\mathbb{K}}} : Q \otimes_{\mathbb{K}} \overline{\mathbb{K}} \to \overline{\mathbb{K}}$ decomposes as $\lambda_{\overline{\mathbb{K}}} = \lambda^{(1)} + \cdots + \lambda^{(d)}$. We call *support* of λ the subset $\mathfrak{S}$ of $\{1, \dots, d\}$ such that $\lambda^{(i)}$ is non-zero exactly for i in $\mathfrak{S}$. As a consequence, for all f in Q, we have

$$\begin{aligned} \lambda(f) &= \lambda^{(1)}(f) + \cdots + \lambda^{(d)}(f) \\ &= \sum_{i \in \mathfrak{S}} \lambda^{(i)}(f). \end{aligned} \tag{1}$$

4.3 Using a generic separating element

We now show how to compute a parametrization of $V(I)$ through the introduction of a generic combination of variables $x = \beta_1 X_1 + \cdots + \beta_n X_n$.

Lemma 8. *Let ℓ be in $\hom_{\mathbb{K}}(Q, \mathbb{K})$, suppose that ℓ is supported on some subset $\mathfrak{S}$ of $\{1, \dots, d\}$ and let $\{\pi_i \mid i \in \mathfrak{S}\}$ and $\{w_i \mid i \in \mathfrak{S}\}$ be as above.*

Let $x = \beta_1 X_1 + \cdots + \beta_n X_n$, for some $\beta_1, \dots, \beta_n$ in $\mathbb{K}$ and let v be in $\mathbb{K}[X_1, \dots, X_n]$. Then, we have the equality

$$\sum_{\ell \geq 0} \ell(vx^i) t^\ell = \sum_{i \in \mathfrak{S}} \frac{v(\boldsymbol{\alpha}_i)\, w_i!\, \pi_i(\beta_1, \dots, \beta_n) t^{w_i} + (1 - x(\boldsymbol{\alpha}_i)t) A_i}{(1 - x(\boldsymbol{\alpha}_i)t)^{w_i+1}}, \tag{2}$$

for some polynomials $A_1, \dots, A_d$ in $\overline{\mathbb{K}}[t]$, with A_i of degree less than w_i for all i in $\mathfrak{S}$.

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

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

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

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

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

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

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

$$\sum_{\ell \geq 0} \sum_{\mu \in R} c_\mu \frac{\partial^{|\mu|}(vx^\ell)}{X_1^{\mu_1} \cdots X_n^{\mu_n}} t^\ell = v\,w!\,\pi(\beta_1, \dots, \beta_n) \frac{t^w}{(1-xt)^{w+1}} + \frac{A(\boldsymbol{X}, t)}{(1-xt)^w}.$$

Applying this formula to the sum in (1), we obtain the claim in the lemma. □

Let us suppose that

- for all i in $\mathfrak{S}$, $\pi_i(\beta_1, \dots, \beta_n)$ is nonzero;
- x is a separating element for $\{\boldsymbol{\alpha}_i \mid i \in \mathfrak{S}\}$.

Take $v = 1$ in the previous lemma, and let us rewrite the sum in (2) as A/B, with

$$A = \sum_{i \in \mathfrak{S}} \Big(\big[w_i!\, \pi_i(\beta_1, \dots, \beta_n) t^{w_i} + (1 - x(\boldsymbol{\alpha}_i)t) A_i \big] \prod_{j \in \mathfrak{S} - \{i\}} (1 - x(\boldsymbol{\alpha}_j)t)^{w_j+1} \Big)$$

$$B = \prod_{i \in \mathfrak{S}} (1 - x(\boldsymbol{\alpha}_i)t)^{w_i+1}.$$

We claim that A and B are coprime. Indeed, any root of B is of the form $1/x(\boldsymbol{\alpha}_i)$ for i in $\mathfrak{S}$; we claim that all values $A(1/x(\boldsymbol{\alpha}_i))$ are nonzero. Indeed, for $i' \neq i$, the term $\prod_{j \in \mathfrak{S} - \{i'\}} (1 - x(\boldsymbol{\alpha}_j)t)^{w_j+1}$ vanishes at $t = 1/x(\boldsymbol{\alpha}_i)$, whereas our two assumptions respectively imply that $w_i!\, \pi_i(\beta_1, \dots, \beta_n) t^{w_i} + (1 - x(\boldsymbol{\alpha}_i)t) A_i$ and $\prod_{j \in \mathfrak{S} - \{i\}} (1 - x(\boldsymbol{\alpha}_j)t)^{w_j+1}$ are nonzero at $t = 1/x(\boldsymbol{\alpha}_i)$, so that $A(1/x(\boldsymbol{\alpha}_i))$ is nonzero.

As a result, given $2D$ terms of the sequence $\ell(x^i)$, we can reconstruct A and B as above. We claim that we can actually recover the polynomial $\tilde{B} = \prod_{i \in \mathfrak{S}} (t - x(\boldsymbol{\alpha}_i))^{w_i+1}$. Remark that $\tilde{B}$ may not agree with the reversed polynomial $\mathrm{rev}(B) = t^{\deg(B)} B(1/t) = \prod_{i \in \mathfrak{S}, x(\boldsymbol{\alpha}_i) \neq 0} (t - x(\boldsymbol{\alpha}_i))^{w_i+1}$; knowing B (and thus $\mathrm{rev}(B)$), we need to determine whether there exists i_0 in $\mathfrak{S}$ such that $x(\boldsymbol{\alpha}_{i_0}) = 0$, and if so, the corresponding exponent w_{i_0}.

- Suppose first that all values $x(\boldsymbol{\alpha}_i)$ are nonzero. Then, B has degree $\sum_{i \in \mathfrak{S}} (w_i + 1)$, so that we have the inequality $\deg(B) > \deg(A)$.
- Suppose now that $x(\boldsymbol{\alpha}_{i_0}) = 0$, for some i_0 in $\mathfrak{S}$. Then, B has degree $\sum_{i \in \mathfrak{S} - \{i_0\}} (w_i + 1)$; let us verify that we have $\deg(A) \geq \deg(B)$ in this case. For i in $\mathfrak{S} - \{i_0\}$, the term $\big[w_i!\, \pi_i(\beta_1, \dots, \beta_n) t^{w_i} + (1 - x(\boldsymbol{\alpha}_i)t) A_i \big] \prod_{j \in \mathfrak{S} - \{i\}} (1 - x(\boldsymbol{\alpha}_j)t)^{w_j+1}$ has degree less than $\deg(B)$, whereas that same term for $i = i_0$ has degree $\sum_{i \in \mathfrak{S}} (w_i + 1) - 1 \geq \deg(B)$.

  Thus, A itself has degree greater than (or equal to) $\deg(B)$, and the difference $\deg(A) - \deg(B)$ is precisely w_{i_0}.

In particular, in either case, we can determine the integer $\delta = \sum_{i \in \mathfrak{S}} (w_i + 1) - 1$, which is an upper bound on the degree of A. This allows us to define $\tilde{A} = t^\delta A(1/t)$, that is,

$$\tilde{A} = \sum_{i \in \mathfrak{S}} \Big(\big[w_i!\, \pi_i(\beta_1, \dots, \beta_n) + (t - x(\boldsymbol{\alpha}_i)) \tilde{A}_i \big] \prod_{j \in \mathfrak{S} - \{i\}} (t - x(\boldsymbol{\alpha}_j))^{w_j+1} \Big),$$

with $\tilde{A}_i = t^{w_i - 1} A_i(1/t) \in \mathbb{K}[t]$. In particular, the value $\tilde{A}(x(\boldsymbol{\alpha}_i))$ is

$$\tilde{A}(x(\boldsymbol{\alpha}_i)) = w_i!\, \pi_i(\beta_1, \dots, \beta_n) \prod_{j \in \mathfrak{S} - \{i\}} (x(\boldsymbol{\alpha}_i) - x(\boldsymbol{\alpha}_j))^{w_j+1},$$

which is nonzero. In other words, $\tilde{A}$ is a unit modulo $C = \prod_{i \in \mathfrak{S}} (t - x(\boldsymbol{\alpha}_i))$.

Let us finally consider the formula in Lemma 8 for an arbitrary v, still under our two assumptions above. The sum in (2) can now be rewritten as A_v/B, with

$$A_v = \sum_{i \in \mathfrak{S}} \Big(\big[v(\boldsymbol{\alpha}_i) w_i!\, \pi_i(\beta_1, \dots, \beta_n) t^{w_i} + (1 - x(\boldsymbol{\alpha}_i)t) A_i \big] \prod_{j \in \mathfrak{S} - \{i\}} (1 - x(\boldsymbol{\alpha}_j)t)^{w_j+1} \Big).$$

We do not claim that A_v and B are necessarily coprime, but since B is known, we can recover A_v using only D terms in the sequence $\ell(vx^i)$. As above, we can deduce $\tilde{A}_v = t^\delta A_v(1/t)$, that is,

$$\tilde{A}_v = \sum_{i \in \mathfrak{S}} \Big(\big[v(\boldsymbol{\alpha}_i) w_i! \, \pi_i(\beta_1, \ldots, \beta_n) + (t - x(\boldsymbol{\alpha}_i))\tilde{A}_i \big] \prod_{j \in \mathfrak{S} - \{i\}} (t - x(\boldsymbol{\alpha}_j))^{w_j+1} \Big).$$

Now, the value of $\tilde{A}_v$ at $x(\boldsymbol{\alpha}_i)$ is

$$\begin{aligned} \tilde{A}_v(x(\boldsymbol{\alpha}_i)) &= v(\boldsymbol{\alpha}_i) w_i! \, \pi_i(\beta_1, \ldots, \beta_n) \prod_{j \in \mathfrak{S} - \{i\}} (x(\boldsymbol{\alpha}_i) - x(\boldsymbol{\alpha}_j))^{w_j+1} \\ &= v(\boldsymbol{\alpha}_i) \tilde{A}(x(\boldsymbol{\alpha}_i)). \end{aligned}$$

Thus, the polynomial

$$P_v = \frac{\tilde{A}_v}{\tilde{A}} \bmod C$$

is well-defined, and $P_v(x(\boldsymbol{\alpha}_i)) = v(\boldsymbol{\alpha}_i)$ holds for all i in $\mathfrak{S}$.

4.4 Computing with X_1

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

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

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

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

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

[8] 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.

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

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

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