Sub-cubic Change of Ordering for Gröner Basis: A Probabilistic Approach

Jean-Charles Faugère, Pierrick Gaudry, Louise Huot, Guénaël Renault

Sub-Cubic Change of Ordering for Gröbner Basis. A Probabilistic Approach.

Jean-Charles Faugère
INRIA, Paris Rocquencourt
Sorbonne Universités
UPMC Univ. Paris 06
CNRS, UMR 7606, LIP6
PolSys Project, Paris, France
jean-charles.faugere@inria.fr

Pierrick Gaudry
CNRS, INRIA
Université de Lorraine
Caramel Project, LORIA
UMR 7503, Nancy, France
pierrick.gaudry@loria.fr

Louise Huot Guénaël Renault
Sorbonne Universités
UPMC Univ. Paris 06
INRIA, Paris Rocquencourt
CNRS, UMR 7606, LIP6
PolSys Project, Paris, France
{louise.huot;guenael.renault}@lip6.fr

ABSTRACT

The usual algorithm to solve polynomial systems using Gröbner bases consists of two steps: first computing the DRL Gröbner basis using the F_5 algorithm then computing the LEX Gröbner basis using a change of ordering algorithm. When the Bézout bound is reached, the bottleneck of the total solving process is the change of ordering step. For 20 years, thanks to the FGLM algorithm the complexity of change of ordering is known to be cubic in the number of solutions of the system to solve.

We show that, in the generic case or up to a generic linear change of variables, the multiplicative structure of the quotient ring can be computed with no arithmetic operation. Moreover, given this multiplicative structure we propose a change of ordering algorithm for *Shape Position* ideals whose complexity is polynomial in the number of solutions with exponent ω where $2 \leq \omega < 2.3727$ is the exponent in the complexity of multiplying two dense matrices. As a consequence, we propose a new Las Vegas algorithm for solving polynomial systems with a finite number of solutions by using Gröbner basis for which the change of ordering step has a sub-cubic (*i.e.* with exponent ω) complexity and whose total complexity is dominated by the complexity of the F_5 algorithm.

In practice we obtain significant speedups for various polynomial systems by a factor up to 1500 for specific cases and we are now able to tackle some instances that were intractable.

Categories and Subject Descriptors

I.1.2 [**Computing Methodologies**]: Symbolic and Algebraic Manipulation; F.2.2 [**Theory of Computation**]: Analysis of Algorithms and Problem Complexity

General Terms

ALGORITHMS, EXPERIMENTATION, THEORY

ISSAC '14, July 23 - 25, 2014, Kobe, Japan.

http://dx.doi.org/10.1145/2608628.2608669.

Keywords

Polynomial systems, Gröbner basis, change of ordering

1. INTRODUCTION

In all this paper, we consider the fundamental problem of Polynomial System Solving (PoSSo for short). More precisely, we focus on the complexity of computing a LEX Gröbner basis of a zero-dimensional ideal. In the sequel, we denote by D the finite number of the corresponding solutions counted with multiplicities in an algebraic closure of the coefficient field.

For the particular case of approximating or computing a rational parametrization of all the solutions of a polynomial system with coefficients in a field of characteristic zero there exist algorithms with sub-cubic complexity in D. Indeed, if the number of real roots is logarithmic in D then the cost is $\widetilde{O}(12^n D^2)$ for the approximation, see [24], and if the multiplicative structure of the quotient ring is known the cost is $O\left(n2^n D^{\frac{5}{2}}\right)$ for the rational parametrization, see [5]. However, to the best of our knowledge, there is no better bound than $O(nD^3)$ for the complexity of computing a LEX Gröbner basis.

This complexity bound for solving the PoSSo problem is obtained by using the usual algorithm to compute a LEX Gröbner basis. This algorithm consists in two steps. First by computing a degree reverse lexicographical (DRL for short) Gröbner basis by using for instance the F_5 algorithm [9] whose complexity is bounded by $O\left(ne^{\omega n}d^{\omega n}\right)$ arithmetic operations [1] where d is the maximal degree of the input equations and ω is the exponent in the complexity of multiplying two dense matrices ($2 \leq \omega < 2.3727$ from [30]). Then, the LEX Gröbner basis is computed using a change of ordering algorithm [11, 13, 14] *e.g.* the FGLM algorithm whose complexity is bounded by $O\left(nD^3\right)$ arithmetic operations which is in turn bounded by $O\left(nd^{3n}\right)$ according to the Bézout bound. When $d \geq 44$ (see Figure 1) the complexity of the PoSSo problem is then bounded by $O\left(nd^{3n}\right)$ arithmetic operations.

In this paper, we propose a new probabilistic algorithm for solving the PoSSo problem. The change of ordering step (Fast FGLM on Figure 1) has a complexity in $\widetilde{O}\left(D^\omega\right)$ bounded by $\widetilde{O}\left(d^{\omega n}\right)$ where the notation $\widetilde{O}$ means that we omit the logarithmic factors in D or polynomial factors in n. As a consequence, the complexity of our algorithm (Fast PoSSo on Figure 1) and thus of the PoSSo problem is bounded by $\widetilde{O}\left(e^{\omega n}d^{\omega n}\right)$ the complexity bound of the F_5 algorithm. A deterministic version of this complexity result can be found in the extended version of this work [8] but the range of applicability of the probabilistic version is wider.

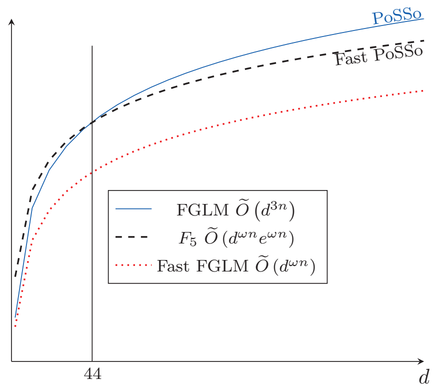

Figure 1: Dominant step in the complexity (ordinate axis) of the PoSSo problem.

In order to obtain such a complexity for solving the PoSSo problem, we first propose in Section 3 a dense and fast version of the change of ordering for *Shape Position* ideals described in [13, 14]. In order to obtain a sub-cubic variant of this algorithm, we focus on its dominant part, the computation of the Krylov iterates associated to a square matrix of size $D \times D$. We propose to use the algorithm of Keller-Gehrig [18] for computing these iterates in $\widetilde{O}(D^\omega)$ which provides us the expected complexity.

Let $\mathcal{I} \subset \mathbb{K}[x_1, \ldots, x_n]$ be the ideal of which we look for the solutions. The input of this change of ordering algorithm is the DRL Gröbner basis of $\mathcal{I}$ and the matrix representation, denoted T_n, of the multiplication by x_n in the quotient ring $\mathbb{K}[x_1, \ldots, x_n]/\mathcal{I}$. From [11], computing T_n can be done in $O\left(nD^3\right)$ arithmetic operations in $\mathbb{K}$. We thus need a way to reduce the cost, to at most sub-cubic, of the computation of T_n. Using the study of the shape of DRL Gröbner bases by Moreno-Socías [22] we actually show in Section 4 that, in the generic case, *no arithmetic* operation is required to build the matrix T_n (Theorem 8). Note that this was already heuristically known in [14]. Hence, we remove the heuristic nature of this result.

Moreover, for non-generic polynomial systems, using results of Galligo [15], Bayer and Stillman [2] and Pardue [26] about *Generic initial ideals* we prove (Corollary 14) that a generic linear change of variables bring us back to this case. As a consequence we obtain our main result about change of ordering.

THEOREM 1. *Let $\mathcal{I}$ be a generic ideal of $\mathbb{K}[x_1, \ldots, x_n]$ of dimension zero. If $\mathcal{I}$ is in* Shape Position *then given its DRL Gröbner basis its LEX Gröbner basis can be computed in $\widetilde{O}(D^\omega)$ arithmetic operations in $\mathbb{K}$. In the case where $\mathcal{I}$ is non-generic and radical we obtain the same complexity up to a change of variables chosen in a non-empty Zariski open subset of* $\mathbf{GL}(\mathbb{K}, n)$.

The radical assumption in the non-generic case is required to ensure that after a generic linear change of variables the ideal is in *Shape Position* (using the Shape Lemma [16, 19]). However, even if the ideal is not radical our algorithm (and complexity result) is still correct if the ideal is in *Shape Position* after a generic linear change of variables. The characterization of such zero-dimensional ideals has been done in [3].

Finally, by using this result in Section 5 we present our new algorithm for polynomial systems solving using Gröbner basis, its complexity and its probability of success. Moreover, as presented in the end of this section although our result seems theoretical we obtain significant improvements in practice.

Related work. Since we focus on *Shape Position* (possibly up to a linear change of variables) ideals the output of our algorithm for solving the PoSSo problem *i.e.* as a representation of the solutions, is similar (up to a normalization) to a rational univariate representation (RUR for short) introduced by Rouillier [27]. Given the multiplicative structure of the quotient ring, the complexity of the algorithm in [27] is in $O(nD^5)$ arithmetic operations in $\mathbb{K}$ which is decreased to $O(n2^n D^{\frac{5}{2}})$ in [5].

The main difference between [5, 27] and our work is that the multiplicative structure of the quotient ring is not assumed to be known. In [13, 14], for generic ideals it is heuristically stated that a sufficient part of this multiplicative structure can be known without arithmetic operation. In this work, we prove this heuristic and extend its scope of applicability.

Contrary to the RUR where one looks for a separating variable, to compute the matrix representation of the multiplication by x_n in the quotient ring we do not need that x_n separates the variety. Hence, in some cases where the RUR is not applicable (*i.e.* after a generic linear change of coordinates the smallest variable is not separating) it is possible to compute with no arithmetic operations the corresponding multiplication matrix. Thus, we can compute its minimal polynomial and obtain the univariate polynomial of the lexicographical Gröbner basis. Note that if the RUR is not applicable then the ideal is not in *Shape Position* and our complete strategy for solving the PoSSo problem cannot be applied. However, the univariate polynomial that we have computed gives a significant information on the solutions which can be sufficient for instance in the case of finite fields.

2. NOTATIONS

From now on, $\mathbb{K}$ denotes a field and $\mathcal{I} \subset \mathbb{K}[x_1, \ldots, x_n]$ denotes an ideal in with a finite number D of solutions counted with multiplicities in $\overline{\mathbb{K}}$. A monomial of $\mathbb{K}[x_1, \ldots, x_n]$ is denoted $x^\alpha = x_1^{\alpha_1} \cdots x_n^{\alpha_n}$ with $\alpha = (\alpha_1, \ldots, \alpha_n) \in \mathbb{N}^n$. The quotient ring $\mathbb{K}[x_1, \ldots, x_n]/\mathcal{I}$ is denoted $\mathcal{R}_\mathcal{I}$. The set of invertible matrices of size $n \times n$ with coefficients in $\mathbb{K}$ is denoted $\mathbf{GL}(\mathbb{K}, n)$ and $g \cdot \mathcal{I}$ with $g \in \mathbf{GL}(\mathbb{K}, n)$ denotes the ideal $\{f(g \cdot \mathbf{x}) \mid f \in \mathcal{I}\}$ where $\mathbf{x}$ is the vector $(x_1, \ldots, x_n)$.

Once a monomial ordering $>$ on $\mathbb{K}[x_1, \ldots, x_n]$ is fixed we define

- $\mathrm{LT}_>(f)$, the leading term of f w.r.t. $>$;
- $\mathrm{in}_>(\mathcal{I}) = \{\mathrm{LT}_>(f) \mid f \in \mathcal{I}\}$, the initial ideal of $\mathcal{I}$ w.r.t. $>$;
- $\mathcal{G}_>$, the reduced Gröbner basis of $\mathcal{I}$ w.r.t. $>$;
- $\mathrm{E}_>(\mathcal{I}) = \{\mathrm{LT}_>(f) \mid f \in \mathcal{G}_>\}$, the stair of $\mathcal{I}$ w.r.t. $>$ *i.e.* a minimal set of generators of $\mathrm{in}_>(\mathcal{I})$.

The quotient ring $\mathcal{R}_\mathcal{I}$ is a $\mathbb{K}$-vector space of dimension D. Its canonical basis w.r.t. the ordering $>$ is given by

$$\mathrm{B}_>(\mathcal{I}) = \{x^\alpha \mid x^\alpha \notin \mathrm{in}_>(\mathcal{I})\} = \{\epsilon_D > \cdots > \epsilon_1 = 1\}.$$

The normal form map gives a representative of any polynomial f in $\mathcal{R}_\mathcal{I}$ w.r.t. this basis; we denote by $\mathrm{NF}_>(f)$ this unique polynomial of the form $\sum_{i=1}^D c_i \epsilon_i$ where $c_i \in \mathbb{K}$ such that $f - \mathrm{NF}_>(f) \in \mathcal{I}$.

The normal form map thus provides a representation of $\mathcal{R}_\mathcal{I}$ as a D-dimensional $\mathbb{K}$-vector space. The matrix representation of the multiplication by x_i in $\mathcal{R}_\mathcal{I}$ seen as the vector space with the basis $\mathrm{B}_>(\mathcal{I})$ is called the multiplication matrix by x_i and is denoted

T_i. The columns of this matrix thus consist of the coefficients of $\mathrm{NF}_>(x_i\epsilon_j)$ for $j = 1, \ldots, D$.

3. FAST CHANGE OF ORDERING FOR SHAPE POSITION IDEALS

In [13], Faugère & Mou propose a probabilistic algorithm which given the reduced Gröbner basis w.r.t. a monomial ordering $>_1$ of an ideal $\mathcal{I} \subset \mathbb{K}[x_1, \ldots, x_n]$ computes the LEX Gröbner basis – if it is in *Shape Position* – of $\mathcal{I}$. The idea is to take advantage of the shape of the LEX Gröbner basis (assumed to be known) to design a very efficient change of ordering algorithm.

Throughout this section, the multiplication matrix T_n is assumed to be known and $\mathcal{I}$ is assumed to be in *Shape Position*. That is to say the LEX Gröbner basis of $\mathcal{I}$ has the following shape:

$$\mathcal{G}_{>_{\mathrm{lex}}} = \{x_1 - h_1(x_n), \ldots, x_{n-1} - h_{n-1}(x_n), h_n(x_n)\}$$

where $h_1, \ldots, h_n \in \mathbb{K}[x_n]$, $\deg(h_i) < D$ for $i = 1, \ldots, n-1$ and $\deg(h_n) = D$.

Computing h_n. The polynomial h_n is then given by the minimal polynomial of the multiplication matrix T_n. In order to reduce its computation to the solving of a linear Hankel system one can use the first part of the Wiedemann probabilistic algorithm [31]. More precisely, first one computes the linearly recurrent sequence $S = [(\mathbf{r}, T_n^j\mathbf{1}) \mid j = 0, \ldots, 2D-1]$ where $\mathbf{r}$ is a random column vector of $\mathbb{K}^D$, $\mathbf{1} = (1, 0, \ldots, 0)^t$ is the vector representing the monomial 1 in $\mathcal{R}_\mathcal{I}$ and $(\cdot, \cdot)$ denotes the scalar product. Then, by using the Berlekamp-Massey algorithm [21] one computes the minimal polynomial μ of S. Finally, if $\deg(\mu) = D$ one has $\mu = h_n$.

Computing $h_1, \ldots, h_{n-1}$. Let us write $h_i = \sum_{k=0}^{D-1} c_{i,k}x_n^k$ where the $c_{i,k}$ are unknown. By noting that $x_i - h_i(x_n) \in \mathcal{I}$ one has $\mathrm{NF}_{>_1}\left(x_i - \sum_{k=0}^{D-1} c_{i,k}x_n^k\right) = 0$. By translating this equation as a linear combination in $\mathcal{R}_\mathcal{I}$ seen as a $\mathbb{K}$-vector space, then by multiplying the resulting equation by T_n^j for $j = 0, \ldots, D-1$ and taking the scalar product with $\mathbf{r}$ we deduce that

$$0 = \left(\mathbf{r}, T_n^j(T_i\mathbf{1})\right) - \sum_{k=0}^{D-1} c_{i,k}\left(\mathbf{r}, T_n^{k+j}\mathbf{1}\right). \tag{3a}$$

For each polynomial h_i for $i = 1, \ldots, n-1$ the equation (3a) allows to construct a linear Hankel system defined by the linearly recurrent sequence S of which $c_{i,k}$ for $k = 0, \ldots, D-1$ are the solutions. From [17], this linear Hankel system is non-singular since the rank of the Hankel matrix is given by the degree of the minimal polynomial of S which is exactly D in our case. Note that one can assume w.l.o.g. that $x_i \in \mathrm{B}_{>_1}(\mathcal{I})$. Hence, the vectors $\mathbf{w_i} = T_i\mathbf{1}$ are known without arithmetic operations. For more details see [13]. In [14] the authors propose a deterministic version of their algorithm for *Shape Position* ideal. Note that their algorithm computes the LEX Gröbner basis of the radical of the ideal $\mathcal{I}$ given in input if it is in *Shape Position*.

Theorem 2 (Faugère & Mou [13, 14]). *Let $\mathcal{I}$ be an ideal of $\mathbb{K}[x_1, \ldots, x_n]$. Let $\mathbf{r}$ be a random column vector of $\mathbb{K}^D$ and let T be the transpose of the multiplication matrix T_n w.r.t. a monomial ordering $>$. If $\mathcal{I}$ is in* Shape Position *then given the reduced Gröbner basis of $\mathcal{I}$ w.r.t. $>$ and the vectors $T^j\mathbf{r}$ for $j = 0, \ldots, 2D-1$, there exists a probabilistic algorithm which computes the LEX Gröbner basis of $\mathcal{I}$ in $O\left(nD\log^2 D \log\log D\right)$ arithmetic operations in $\mathbb{K}$.*

If the radical of $\mathcal{I}$ is in Shape Position *then given the reduced Gröbner basis of $\mathcal{I}$ w.r.t. $>$ and the vectors $T_n^j\mathbf{1}$, $T_n^j\mathbf{w_1}, \ldots, T_n^j\mathbf{w_{n-1}}$ for $j = 0, \ldots, 2D-1$, there exists a deterministic algorithm which computes the LEX Gröbner basis of the radical of $\mathcal{I}$ in $O\left(nD^2\log D\log\log D\right)$ (by omitting logarithmic factors in q if $\mathbb{K} = \mathbb{F}_q$) arithmetic operations in $\mathbb{K}$.*

One issue remains to get a change of ordering algorithm with sub-cubic complexity in D given the matrix T_n. Indeed, in [13] the authors assume the matrix T_n sparse and compute iteratively the vectors $T^j\mathbf{r}$ or $T_n^j\mathbf{1}, T_n^j\mathbf{w_1}, \ldots, T_n^j\mathbf{w_{n-1}}$ for $j = 0, \ldots, 2D-1$. However, when T_n is dense this yields a complexity in $O\left(D^3\right)$ arithmetic operations in $\mathbb{K}$. In order to overcome this issue we use an algorithm of Keller-Gehrig [18] which computes these matrix-vector products by multiplying $O(\log D)$ matrices. More precisely, first one computes $T^2, T^4, \ldots, T^{2^{\lceil\log_2 D\rceil}}$ using binary exponentiation with $\lceil\log_2 D\rceil$ matrix products; then the vectors $T^j\mathbf{r}$ for $j = 0, \ldots, 2D-1$ are computed by induction in $\lceil\log_2 D\rceil$ steps:

$$\begin{aligned} T^2\left(T\mathbf{r} \mid \mathbf{r}\right) &= \left(T^3\mathbf{r} \mid T^2\mathbf{r}\right) \\ T^4\left(T^3\mathbf{r} \mid T^2\mathbf{r} \mid T\mathbf{r} \mid \mathbf{r}\right) &= \left(T^7\mathbf{r} \mid T^6\mathbf{r} \mid T^5\mathbf{r} \mid T^4\mathbf{r}\right) \\ T^8\left(T^7\mathbf{r} \mid T^6\mathbf{r} \mid \cdots \mid \mathbf{r}\right) &= \left(T^{15}\mathbf{r} \mid T^{14}\mathbf{r} \mid \cdots \mid T^8\mathbf{r}\right) \\ &\vdots \end{aligned} \tag{3b}$$

until the product $T^{2^{\lceil\log_2 D\rceil}}\left(T^{2^{\lceil\log_2 D\rceil}-1}\mathbf{r} \mid \cdots \mid \mathbf{r}\right)$ where the notation $(\mathbf{r}_1 | \ldots | \mathbf{r}_k)$ is the matrix with D rows and k columns obtained by joining the column vectors $\mathbf{r}_i$ vertically. As a consequence, we obtain the following result.

Proposition 3. *Let $\mathcal{I}$ be an ideal of $\mathbb{K}[x_1, \ldots, x_n]$. If $\mathcal{I}$ (resp. the radical of $\mathcal{I}$) is in* Shape Position *then given the reduced Gröbner basis of $\mathcal{I}$ w.r.t. a monomial ordering $>$ and the associated multiplication matrix T_n there exists a probabilistic (resp. deterministic) algorithm which computes the LEX Gröbner basis of $\mathcal{I}$ (resp. of the radical of $\mathcal{I}$) in $O\left(D\log D\left(D^{\omega-1} + n\log D\log\log D\right)\right)$ (resp. $O\left(nD^2\log D\left(D^{\omega-2} + \log\log D\right)\right)$ by omitting logarithmic factors in q if $\mathbb{K} = \mathbb{F}_q$) arithmetic operations in $\mathbb{K}$.*

In the next section we investigate the computation of the matrix T_n.

4. COMPUTING T_N

In this section we fix the first monomial ordering to the DRL ordering $>_{\mathrm{drl}}$. To compute the multiplication matrix T_n we need to compute the normal forms w.r.t. the DRL ordering of all the monomials $\epsilon_i x_n$ for $i = 1, \ldots, D$ with $\epsilon_i \in \mathrm{B}_{>_{\mathrm{drl}}}(\mathcal{I})$. From [11] the monomials $\epsilon_i x_n$ can be of three types.

Proposition 4 (FGLM [11]). *Let $F = \{x_j\epsilon_i \mid 1 \le i \le D$ and $1 \le j \le n\} \setminus \mathrm{B}_>(\mathcal{I})$ be the* border. *Let $t = \epsilon_i x_j$ with $i \in \{1, \ldots, D\}$ and $j \in \{1, \ldots, n\}$. One has the following three cases*

I. either $t \in \mathrm{B}_>(\mathcal{I})$ and $\mathrm{NF}_>(t) = t$;

II. or $t \in \mathrm{E}_>(\mathcal{I})$ i.e. *$t = \mathrm{LT}_>(g)$ for some $g \in \mathcal{G}_>$ hence, $\mathrm{NF}_>(t) = t - g$;*

III. or $t = x_k t'$ with $t' \in F$. Hence, denoting $\mathrm{NF}_>(t') = \sum_{l=1}^s \alpha_l\epsilon_l$ with $t' > \epsilon_s$, we have $\mathrm{NF}_>(t) = \mathrm{NF}_>(x_k\,\mathrm{NF}_>(t')) = \sum_{l=1}^s \alpha_l\,\mathrm{NF}_>(\epsilon_l x_k)$.

In this section, thanks to the study of the stairs of generic ideals by Moreno-Socías [22], we first show that for generic ideals and DRL ordering, all monomials of the form $\epsilon_i x_n$ are either in $\mathrm{B}_{>_{\mathrm{drl}}}(\mathcal{I})$ or in $\mathrm{E}_{>_{\mathrm{drl}}}(\mathcal{I})$. Hence, the multiplication matrix T_n can be computed very efficiently. Then, we show that, up to a generic linear change of variables, this result can be extended to any ideal.

In the sequel, the arithmetic operations will be the addition or the multiplication of two operands in $\mathbb{K}$ that are different from ± 1 and 0. In particular we do not consider the change of sign as an arithmetic operation.

4.1 Generic case

DEFINITION 5. *A generic sequence of polynomials F is a sequence of polynomials whose coefficients are indeterminates i.e. $F = (f_1, \ldots, f_s)$ with $f_i = \sum_\alpha c_{i,\alpha} x^\alpha$ is in $\mathbb{K}[x_1, \ldots, x_n]$ where $\mathbb{K} = k(\{c_{i,\alpha}\})$ and k is a field. A generic ideal is an ideal generated by a generic sequence of polynomials.*

In [22] it is shown that the intersection of the section of $\mathcal{R}_\mathcal{I}$ by $x_{i_1}^{d_1}, \ldots, x_{i_{n-2}}^{d_{n-2}}$ has steps of depth two and height one for any $d_1, \ldots, d_{n-2} \geq 0$ and $i_1, \ldots, i_{n-2} \leq n-1$ all pairwise distinct. We illustrate this result on Figure 2 where for fixed value of $d_1, \ldots, d_{i-1}, d_{i+1}, \ldots, d_{n-1}$ we represent the corresponding monomials of $\mathrm{E}_{>_{\mathrm{drl}}}(\mathcal{I}) \cup \mathrm{B}_{>_{\mathrm{drl}}}(\mathcal{I})$. The x_n-axis (resp. x_i-axis) corresponds to the degree in x_n (resp. x_i) of these monomials.

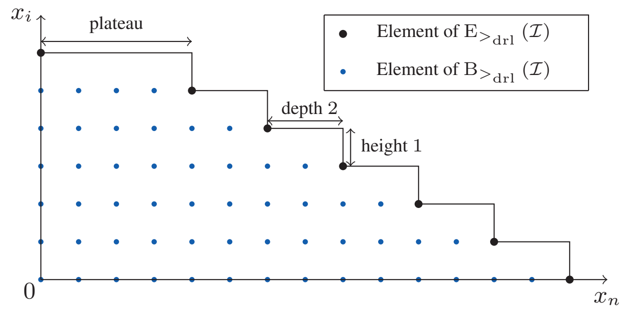

Figure 2: Intersection of sections of the quotient ring $\mathcal{R}_\mathcal{I}$ by $x_1^{d_1}, \ldots, x_{i-1}^{d_{i-1}}, x_{i+1}^{d_i}, \ldots, x_{n-1}^{d_{n-2}}$ with $\mathcal{I}$ a generic ideal.

The shape of the stair in Figure 2 is formally stated in the following theorem.

THEOREM 6 (MORENO-SOCÍAS [22]). *Let $\mathcal{I}$ be a generic ideal of $\mathbb{K}[x_1, \ldots, x_n]$ generated by $(f_1, \ldots, f_n)$. Let $\mathcal{M}$ be the set of monomials of $\mathbb{K}[x_1, \ldots, x_{n-1}]$ and $\widetilde{B}_i = \{m \in \mathcal{M} \mid m x_n^i \in \mathrm{B}_{>_{\mathrm{drl}}}(\mathcal{I})\}$. Let $\delta = \sum_{i=1}^n (\deg(f_i) - 1)$, $\delta^* = \sum_{i=1}^{n-1} (\deg(f_i) - 1)$ and $\sigma = \min\left(\delta^*, \lfloor \frac{\delta}{2} \rfloor\right)$. Let $\mu = \delta - 2\sigma$, then*

a. $\widetilde{B}_0 = \cdots = \widetilde{B}_\mu$ (plateau) and $\widetilde{B}_i = \widetilde{B}_{i+1}$ for $\mu < i < \delta$ and $i \not\equiv \delta \mod 2$ (depth two);

b. The leading term of the polynomials in $\mathcal{G}_{>_{\mathrm{drl}}}$ of degree 0 in x_n have degree at most $\sigma + 1 = \bar{\sigma}$;

c. The leading term of the polynomials in $\mathcal{G}_{>_{\mathrm{drl}}}$ of degree α in x_n with $\mu < \alpha \leq \delta + 1$ with $\alpha \not\equiv \delta \mod 2$ are all of total degree $d + \alpha$ where $d = \max(\deg(m) \mid m \in \widetilde{B}_{\alpha-1})$. Moreover, all these leading terms are exactly given by $t = m x_n^\alpha$ for all $m \in \widetilde{B}_{\alpha-1}$ of degree d (height one);

d. There is no leading term of polynomials in $\mathcal{G}_{>_{\mathrm{drl}}}$ of degree $1, \ldots, \mu$ in x_n (plateau) or of degree α in x_n with $\alpha > \delta + 1$ or $\mu \leq \alpha \leq \delta$ and $\alpha \equiv \delta \mod 2$ (depth two).

We deduce of the previous theorem that generic ideals satisfy the following property.

PROPOSITION 7. *Let $\mathcal{I}$ be a generic ideal. Let t be a monomial in $\mathrm{E}_{>_{\mathrm{drl}}}(\mathcal{I})$ i.e. a leading term of a polynomial in the DRL Gröbner basis of $\mathcal{I}$. If x_n divides t then for all $k \in \{1, \ldots, n-1\}$, $\frac{x_k t}{x_n} \in \mathrm{in}_{>_{\mathrm{drl}}}(\mathcal{I})$.*

PROOF. This result is deduced from the shape of the stairs of $\mathcal{I}$. Let $t = x_1^{\alpha_1} \cdots x_n^{\alpha_n}$ be a leading term of a polynomial in $\mathcal{G}_{>_{\mathrm{drl}}}$ divisible by x_n *i.e.* $\alpha_n > 0$ and $m = x_1^{\alpha_1} \cdots x_{n-1}^{\alpha_{n-1}}$. We use the same notations as in Theorem 6.

From Theorem 6 item (d), since $t \in \mathrm{E}_{>_{\mathrm{drl}}}(\mathcal{I})$ and $\alpha_n > 0$ we have $\alpha_n > \mu$ and $\alpha_n \not\equiv \delta \mod 2$. Then, from Theorem 6 item (c), $\deg(m)$ is the maximal degree reached by the monomials in $\widetilde{B}_{\alpha_{n-1}}$. Thus $x_k m \notin \widetilde{B}_{\alpha_{n-1}}$ for all $k \in \{1, \ldots, n-1\}$. As a consequence, for all $k \in \{1, \ldots, n-1\}$ we have $\frac{x_k t}{x_n} \in \mathrm{in}_{>_{\mathrm{drl}}}(\mathcal{I})$. $\square$

A consequence of this result is that all the monomials of the form $\epsilon_i x_n$ where $\epsilon_i \in \mathrm{B}_{>_{\mathrm{drl}}}(\mathcal{I})$ are either of type (I) or of type (II) of Proposition 4. Hence, their normal form can be read on $\mathcal{G}_{>_{\mathrm{drl}}}$ with no arithmetic operations and the multiplication matrix T_n can be computed very efficiently. This is summarized in the following result.

THEOREM 8. *Given $\mathcal{G}_{>_{\mathrm{drl}}}$ the DRL Gröbner basis of a generic ideal $\mathcal{I}$ of dimension zero, the multiplication matrix T_n can be read from $\mathcal{G}_{>_{\mathrm{drl}}}$ with no arithmetic operation.*

PROOF. Suppose that there exists $i \in \{1, \ldots, D\}$ such that $t = x_n \epsilon_i$ is of type (III). Hence, $t = m\ \mathrm{LT}_{>_{\mathrm{drl}}}(g)$ for some $g \in \mathcal{G}_{>_{\mathrm{drl}}}$ and $\deg(m) > 1$ with $x_n \nmid m$ (otherwise $\epsilon_i \notin B$). Then, there exists $k \in \{1, \ldots, n-1\}$ such that $x_k \mid m$. By consequence, from Proposition 7, we have $\epsilon_i = \frac{m}{x_k} \cdot \frac{x_k\ \mathrm{LT}_{>_{\mathrm{drl}}}(g)}{x_n} \in \mathrm{in}_{>_{\mathrm{drl}}}(\mathcal{I})$ which yields a contradiction. Thus, all monomials $t = x_n \epsilon_i$ are either in B or in $\mathrm{E}_{>_{\mathrm{drl}}}(\mathcal{I})$ and their normal forms are known and given either by t (if $t \in B$) or by changing the sign of some polynomial $g \in \mathcal{G}_{>_{\mathrm{drl}}}$ and removing its leading term. Note that by using a linked list representation (for instance), removing the leading term of a polynomial does not require arithmetic operation. $\square$

From Theorem 8 and Proposition 3, we obtain the following result.

COROLLARY 9. *Let $\mathcal{I}$ be a generic ideal in* Shape Position*. From the DRL Gröbner basis of $\mathcal{I}$, its LEX Gröbner basis can be computed in $O(\log D (D^\omega + nD \log D \log\log D))$ arithmetic operations with a probabilistic algorithm.*

However, polynomial systems coming from applications are usually not generic. Nevertheless, this difficulty can be bypassed by applying a linear change of variables. By studying the structure of the *Generic initial ideal* (see Remark 10) of $\mathcal{I}$ – that is to say, the initial ideal of $g \cdot \mathcal{I}$ for a generic choice of g in $\mathbf{GL}(\mathbb{K}, n)$ – we will show that results of Proposition 7 and Theorem 8 can be generalized to non generic ideals, up to a generic linear change of variables. Indeed, in [15] Galligo shows that for the characteristic zero fields, the *Generic initial ideal* of any homogeneous ideal satisfies a more general property than Proposition 7. Later, Pardue [26] extends this result to the fields of positive characteristic.

REMARK 10. *Note that* Generic initial ideal *are not defined as an initial ideal whose coefficients are indeterminates. Its definition is given in Definition 11. To avoid ambiguity, in the sequel we always use the notation* $\mathbf{Gin}(\mathcal{I})$ *for* Generic initial ideal *as defined in Definition 11.*

4.2 Non-generic case

DEFINITION 11. *Let $\mathbb{K}$ be an infinite field and $\mathcal{I}$ be an homogeneous ideal of $\mathbb{K}[x_1,\ldots,x_n]$. There exists a non-empty Zariski open set $U \subset \mathbf{GL}(\mathbb{K},n)$ and a monomial ideal $\mathcal{J}$ such that $\mathrm{in}_{>_{\mathrm{drl}}}(g\cdot\mathcal{I}) = \mathcal{J}$ for all $g \in U$. The* Generic initial ideal *of $\mathcal{I}$ is denoted $\mathbf{Gin}(\mathcal{I})$ and is defined by $\mathcal{J}$.*

The proof of the existence of $\mathbf{Gin}(\mathcal{I})$ can be found in [7, p.351–358]. The next result, is a direct consequence of [2, 15, 26] and summarized in [7, p.351–358]. This result allows to extend, up to a linear change of variables, Proposition 7 to non-generic ideals.

THEOREM 12. *Let $\mathbb{K}$ be an infinite field of characteristic $p \geq 0$. Let $\mathcal{I}$ be an homogeneous ideal of $\mathbb{K}[x_1,\ldots,x_n]$ and $\mathcal{J} = \mathbf{Gin}(\mathcal{I})$. For the DRL ordering, for all generators m of $\mathcal{J}$, if x_i^t divides m and x_i^{t+1} does not divide m then for all $j < i$, the monomial $\frac{x_j}{x_i}m$ is in $\mathcal{J}$ if $t \not\equiv 0 \mod p$.*

Polynomial systems coming from applications are usually not homogeneous and Theorem 12 does not apply directly. Let $f = \sum_{i=0}^{d} f_i$ be an affine polynomial of degree d of $\mathbb{K}[x_1,\ldots,x_n]$ where f_i is an homogeneous polynomial of degree i. The homogeneous component of highest degree of f, denoted f^h, is the homogeneous polynomial f_d. Let $\mathcal{I}$ be an affine ideal *i.e.* generated by a sequence of affine polynomials. In the next proposition we highlight an homogeneous ideal having the same initial ideal than $\mathcal{I}$. This allows to extend the result of Theorem 12 to affine ideals.

PROPOSITION 13. *Let $\mathcal{I} = \langle f_1,\ldots,f_s\rangle$ be an affine ideal. If $(f_1^h,\ldots,f_s^h)$ is a regular sequence of polynomials, then there exists a non-empty Zariski open subset $U_a \subset \mathbf{GL}(\mathbb{K},n)$ such that for all $g \in U_a$, $\mathrm{E}_{>_{\mathrm{drl}}}(g\cdot\mathcal{I}) = \mathrm{E}_{>_{\mathrm{drl}}}\left(\mathbf{Gin}\left(\mathcal{I}^h\right)\right)$.*

PROOF. Let f be a polynomial. We denote by f^a the polynomial $f - f^h$. Let $t \in \mathrm{in}_{>_{\mathrm{drl}}}(\mathcal{I})$, there exists $f \in \mathcal{I}$ such that $\mathrm{LT}_{>_{\mathrm{drl}}}(f) = t$. Since, $f \in \mathcal{I}$ and $(f_1^h,\ldots,f_s^h)$ is assumed to be a regular sequence then there exist $h_1,\ldots,h_s \in \mathbb{K}[x_1,\ldots,x_n]$ s.t. $f = \sum_{i=1}^s h_i f_i = \sum_{i=1}^s h_i f_i^h + \sum_{i=1}^s h_i f_i^a$ with $\deg(h_i f_i) \leq \deg(f)$ for all $i \in \{1,\ldots,s\}$ and there exists $j \in \{1,\ldots,s\}$ such that $\deg(h_j f_j) = \deg(f)$. By consequence, $0 \neq \sum_{i=1}^s h_i f_i^h \in \mathcal{I}^h$ where $\mathcal{I}^h$ is the ideal generated by $\{f_1^h,\ldots,f_s^h\}$ and $\mathrm{LT}_{>_{\mathrm{drl}}}(f) = \mathrm{LT}_{>_{\mathrm{drl}}}\left(\sum_{i=1}^s h_i f_i^h\right)$. Thus, $\mathrm{in}_{>_{\mathrm{drl}}}(\mathcal{I})$ is included in $\mathrm{in}_{>_{\mathrm{drl}}}\left(\mathcal{I}^h\right)$. It is straightforward that $\mathrm{in}_{>_{\mathrm{drl}}}\left(\mathcal{I}^h\right) \subset \mathrm{in}_{>_{\mathrm{drl}}}(\mathcal{I})$ hence $\mathrm{in}_{>_{\mathrm{drl}}}\left(\mathcal{I}^h\right) = \mathrm{in}_{>_{\mathrm{drl}}}(\mathcal{I})$.

For all $g \in \mathbf{GL}(\mathbb{K},n)$, since g is invertible the sequence $(g\cdot f_1,\ldots,g\cdot f_s)$ is also regular. Indeed, if there exists $i \in \{1,\ldots,s\}$ such that $g\cdot f_i$ is a divisor of zero in the quotient ring $\mathbb{K}[x_1,\ldots,x_n]/\langle g\cdot f_1,\ldots,g\cdot f_{i-1}\rangle$ then f_i is a divisor of zero in $\mathbb{K}[x_1,\ldots,x_n]/\langle f_1,\ldots,f_{i-1}\rangle$. Hence,

$$\mathrm{in}_{>_{\mathrm{drl}}}(g\cdot\mathcal{I}) = \mathrm{in}_{>_{\mathrm{drl}}}\left((g\cdot\mathcal{I})^h\right).$$

Moreover, g is a linear change of variables thus it preserves the degree. Hence, for all $f \in \mathcal{I}$, we have $(g\cdot f)^h = g\cdot f^h$. Finally, let U_a be the non-empty Zariski open subset of $\mathbf{GL}(\mathbb{K},n)$ such that for all $g \in U_a$, we have the equality $\mathrm{in}_{>_{\mathrm{drl}}}\left(g\cdot\mathcal{I}^h\right) = \mathbf{Gin}\left(\mathcal{I}^h\right)$. Thus, for all $g \in U_a$, we have $\mathrm{in}_{>_{\mathrm{drl}}}(g\cdot\mathcal{I}) = \mathrm{in}_{>_{\mathrm{drl}}}\left((g\cdot\mathcal{I})^h\right) = \mathrm{in}_{>_{\mathrm{drl}}}\left(g\cdot\mathcal{I}^h\right) = \mathbf{Gin}\left(\mathcal{I}^h\right)$. $\square$

Hence, from the previous proposition, for a random linear change of variables $g \in \mathbf{GL}(\mathbb{K},n)$ we have $\mathrm{in}_{>_{\mathrm{drl}}}(g\cdot\mathcal{I}) = \mathbf{Gin}\left(\mathcal{I}^h\right)$. Thus from Theorem 12, for all generators m of the monomial ideal $\mathrm{in}_{>_{\mathrm{drl}}}(g\cdot\mathcal{I})$ (*i.e.* m is a leading term of a polynomial in the DRL Gröbner basis of $g\cdot\mathcal{I}$) if x_n^t divides m and x_n^{t+1} does not divide m then for all $j < n$ we have $\frac{x_j}{x_n}m \in \mathrm{in}_{>_{\mathrm{drl}}}(g\cdot\mathcal{I})$ if $t \not\equiv 0 \mod p$. Therefore, in the same way as for generic ideals, the multiplication matrix T_n of $g\cdot\mathcal{I}$ can be read from its DRL Gröbner basis.

Moreover, the Shape Lemma [16, 19] states that radical ideals have, up to a generic linear change of variables, a LEX Gröbner basis in *Shape Position*. Hence, one can compute very efficiently the multiplication matrix T_n and then use the algorithm presented in Section 3 to compute the LEX Gröbner basis of $g\cdot\mathcal{I}$. This is summarized in the following corollary.

COROLLARY 14. *Let $\mathbb{K}$ be an infinite field of characteristic $p \geq 0$. Let $\mathcal{I} = \langle f_1,\ldots,f_n\rangle$ be a zero-dimensional ideal of $\mathbb{K}[x_1,\ldots,x_n]$ s.t. $(f_1^h,\ldots,f_n^h)$ is a regular sequence. There exists a non-empty Zariski open subset U of $\mathbf{GL}(\mathbb{K},n)$ such that for all $g \in U$, the arithmetic complexity of computing the multiplication matrix by x_n of $g\cdot\mathcal{I}$ given its DRL Gröbner basis can be done without arithmetic operation. If $p > 0$ this is true only if $\deg_{x_n}(m) \not\equiv 0 \mod p$ for all $m \in \mathrm{E}_{>_{\mathrm{drl}}}(g\cdot\mathcal{I})$. Consequently, under the same hypotheses and if $\mathcal{I}$ is a radical ideal, the complexity of computing the LEX Gröbner basis of $g\cdot\mathcal{I}$ given its DRL Gröbner basis can be bounded by $O(\log D(D^\omega + nD\log D\log\log D))$ (or $O(nD^\omega\log D)$ with a deterministic algorithm) arithmetic operations in $\mathbb{K}$.*

Following this result, we propose another algorithm for polynomial systems solving.

5. FAST ALGORITHM FOR SOLVING THE POSSO PROBLEM

Let $\mathcal{S} \subset \mathbb{K}[x_1,\ldots,x_n]$ be a polynomial system generating a radical ideal denoted $\mathcal{I}$. For any $g \in \mathbf{GL}(\mathbb{K},n)$, from the solutions of $g\cdot\mathcal{I}$ one can easily recover the solutions of $\mathcal{I}$. Let U be the non-empty Zariski open subset of $\mathbf{GL}(\mathbb{K},n)$ such that for all $g \in U$, $\mathrm{in}_{>_{\mathrm{drl}}}(g\cdot\mathcal{I}) = \mathbf{Gin}\left(\mathcal{I}^h\right)$. If g is chosen in U then the multiplication matrix T_n can be computed very efficiently. Indeed, from Section 4 all the monomials of the form $\epsilon_i x_n$ for $i = 1,\ldots,D$ are in $\mathrm{B}_{>_{\mathrm{drl}}}(g\cdot\mathcal{I})$ or in $\mathrm{E}_{>_{\mathrm{drl}}}(g\cdot\mathcal{I})$ and their normal form are easily known. Moreover, from the Shape Lemma [16, 19], there exists U' a non-empty Zariski open subset of $\mathbf{GL}(\mathbb{K},n)$ such that for all $g \in U'$ the ideal $g\cdot\mathcal{I}$ admits a LEX Gröbner basis in *Shape Position*. If g is also chosen in U' then we can use the algorithm presented in Section 3 to compute the LEX Gröbner basis of $g\cdot\mathcal{I}$. Hence, we propose in Algorithm 1 a Las Vegas algorithm to solve the PoSSo problem. This is a randomized algorithm whose output (which can be *fail*) is always correct. The end of this section is devoted to evaluate its complexity and its probability of success *i.e.* when the algorithm does not return *fail*.

REMARK 15. *Let $\mathcal{S} = \{f_1,\ldots,f_n\}$ be a polynomial system of $\mathbb{K}[x_1,\ldots,x_n]$. Note that Algorithm 1 can be used if $(f_1^h,\ldots,f_n^h)$ is a regular sequence or not. However, if it is not regular then the complexity and the probability of failure are not well understood.*

REMARK 16. *The test in Line 5 in Algorithm 1 is performed by the beginning of the deterministic algorithm of Proposition 3. Indeed, this algorithm computes for sure the univariate polynomial in $\mathbb{K}[x_n]$ in the LEX Gröbner basis of $\langle g\cdot\mathcal{S}\rangle$. If this polynomial is of degree D then the ideal is in* Shape Position.

REMARK 17. *At Line 6 of Algorithm 1 we use the deterministic version of the change of ordering for* Shape Position *ideals (Section 3) so that the probability of failure of Algorithm 1 does not*

Algorithm 1: Fast PoSSo.

Input : A polynomial system $\mathcal{S} \subset \mathbb{K}[x_1, \ldots, x_n]$ generating a radical ideal.

Output: g in $\mathbf{GL}(\mathbb{K}, n)$ and the LEX Gröbner basis of $\langle g \cdot \mathcal{S} \rangle$ or *fail*.

1 Choose any g in $\mathbf{GL}(\mathbb{K}, n)$;

2 Compute $\mathcal{G}_{>_{\mathrm{drl}}}$ the DRL Gröbner basis of $g \cdot \mathcal{S}$;

3 **if** *T_n can be read from $\mathcal{G}_{>_{\mathrm{drl}}}$* **then**

4   Extract T_n from $\mathcal{G}_{>_{\mathrm{drl}}}$;

5   **if** *$\langle \mathcal{G}_{>_{\mathrm{drl}}} \rangle$ is in* Shape Position **then**

6   From T_n and $\mathcal{G}_{>_{\mathrm{drl}}}$ compute $\mathcal{G}_{>_{\mathrm{lex}}}$ using the deterministic algorithm of Proposition 3;

7   **return** g and $\mathcal{G}_{>_{\mathrm{lex}}}$;

8 **return** *fail*;

depend on the probability of failure of Wiedemann algorithm. Nevertheless, in practice when $\mathbb{K}$ is sufficiently large we can use the probabilistic version of the change of ordering for Shape Position *ideals.*

Algorithm 1 succeeds if the three following conditions are satisfied

1. $g \in \mathbf{GL}(\mathbb{K}, n)$ is chosen in a non-empty Zariski open set U' such that for all $g \in U'$, $g \cdot \mathcal{I}$ has a LEX Gröbner basis in *Shape Position*;

2. $g \in \mathbf{GL}(\mathbb{K}, n)$ is chosen in a non-empty Zariski open set U such that for all $g \in U$, $\mathrm{in}_{>_{\mathrm{drl}}}(g \cdot \mathcal{I}) = \mathbf{Gin}(\mathcal{I}^h)$;

3. $p = 0$ or $p > 0$ and for all $m \in \mathrm{E}_{>_{\mathrm{drl}}}(g \cdot \mathcal{I})$, $\deg_{x_n}(m) \not\equiv 0 \mod p$.

The existence of the non-empty Zariski open subset U' is proven in [16, 19]. Conditions (1) and (2) are satisfied if $g \in U \cap U'$. Since, U and U' are open and dense, $U \cap U'$ is also a non-empty Zariski open set.

5.1 Probability of success of Algorithm 1

Usually the coefficient field of the polynomials is the field of rational numbers or a finite field. Assume that $\mathbb{K} = \mathbb{F}_q$ or $\mathbb{K} = \mathbb{Q}$ and we randomly choose in a finite subset of $\mathbb{Q}$ of size q. The Schwartz-Zippel lemma [28, 32] allows to bound the probability that the conditions (1) and (2) are not satisfied by $\frac{d}{q}$ where d is the degree of the polynomial defining $U \cap U'$. Thus, in order to bound this failure probability we need to estimate the degree of the polynomials defining U and U'.

Construction of U'. Let $\mathcal{I} = \langle f_1, \ldots, f_n \rangle$ be a radical ideal of $\mathbb{K}[x_1, \ldots, x_n]$. Since $\mathcal{I}$ is radical, all its solutions are distinct. Therefore, let $a_i = (a_{i,1}, \ldots, a_{i,n}) \in \overline{\mathbb{K}}^n$ be an element of the algebraic set of solutions of $\mathcal{I}$ (recall that the cardinality of this set is D). Let g be a given matrix in $\mathbf{GL}(\mathbb{K}, n)$. We denote by $v_i = (v_{i,1}, \ldots, v_{i,n})$ the point obtained after transformation of a_i by g, *i.e* $v_i = g \cdot a_i^t$. To ensure that $g \cdot \mathcal{I}$ admits a LEX Gröbner basis in *Shape Position*, g should be such that $v_{i,n} \neq v_{j,n}$ for all couples of integers (i, j) verifying $1 \leq j < i \leq D$. Hence, let $\mathbf{g} = (\mathbf{g}_{i,j})$ be a $(n \times n)$ matrix of unknowns, the polynomial $P_{U'}$ defining the non-empty Zariski open subset U' is then given as the determinant of the Vandermonde matrix associated to $\mathbf{v}_{i,n}$ for $i = 1, \ldots, D$ where $\mathbf{v}_i = (\mathbf{v}_{i,1}, \ldots, \mathbf{v}_{i,n}) = \mathbf{g} \cdot a_i^t$. Therefore, we know exactly the degree of $P_{U'}$ which is $\frac{D(D-1)}{2}$.

Construction of U. The non-empty Zariski open subset U is constructed (see [7, p.351–358]) as the intersection of non-empty Zariski open subsets $U_1, \ldots, U_\delta$ of $\mathbf{GL}(\mathbb{K}, n)$ where δ is the maximum degree of the generators of $\mathbf{Gin}(\mathcal{I}^h)$. Let d be a fixed degree. Let $\mathbb{K}[x_1, \ldots, x_n]_d = R_d$ be the set of homogeneous polynomials of degree d of $\mathbb{K}[x_1, \ldots, x_n]$. Let $\{f_1, \ldots, f_{t_d}\}$ be a vector basis of $\mathcal{I}_d^h = \mathcal{I}^h \cap R_d$. Let $\mathbf{g} = (\mathbf{g}_{i,j})$ be a $(n \times n)$ matrix of unknowns and let M be a matrix representation of the map $\mathcal{I}_d^h \to \mathbf{g} \cdot \mathcal{I}_d^h$ defined as follow:

$$M = (M_{i,j}) = \begin{array}{c} \begin{matrix} m_1 & \cdots & m_N \end{matrix} \\ \begin{bmatrix} \star & \cdots & \star \\ \vdots & \ddots & \vdots \\ \star & \cdots & \star \end{bmatrix} \end{array} \begin{array}{l} \\ \mathbf{g} \cdot f_1 \\ \vdots \\ \mathbf{g} \cdot f_{t_d} \end{array}$$

where $M_{i,j}$ is the coefficient of m_j in $\mathbf{g} \cdot f_i$ and $\{m_1, \ldots, m_N\}$ is the set of monomials in R_d. In [2, 7], the polynomial P_{U_d} defining U_d is constructed as a particular minor of size t_d of M. Since each coefficient in M is a polynomial in $\mathbb{K}[\mathbf{g}_{1,1}, \ldots, \mathbf{g}_{n,n}]$ of degree d, the degree of P_{U_d} is $d \cdot t_d$. Finally, since U_d is open and dense for all $d = 1, \ldots, \delta$ we deduce that $U = \cap_{d=1}^{\delta} U_d$ is a non-empty Zariski open set whose defining polynomial, P_U, is of degree $\sum_{d=1}^{\delta} d \cdot t_d \leq \delta \sum_{d=1}^{\delta} t_d$. Moreover, $D = \dim_{\mathbb{K}}(\mathbb{K}[x_1, \ldots, x_n]/\mathcal{I}^h) = \sum_{d=0}^{\delta} \dim_{\mathbb{K}}(R_d/\mathcal{I}_d^h)$.

Thus, $\sum_{d=0}^{\delta} \dim_{\mathbb{K}}(\mathcal{I}_d^h) = \sum_{d=0}^{\delta} \dim_{\mathbb{K}}(R_d) - D = \binom{n+\delta}{n} - D$. By consequence, $\deg(P_U) \leq \delta\left(\binom{n+\delta}{n} - D\right)$.

For ideals generated by $f_1, \ldots, f_n$ s.t. $(f_1^h, \ldots, f_n^h)$ is a regular sequence, according to the Macaulay bound, δ can be bounded by $\sum_{i=1}^{n}(\deg(f_i) - 1) + 1$. Note that the Macaulay bound gives also a bound on $\deg_{x_n}(m)$ for all $m \in \mathrm{E}_{>_{\mathrm{drl}}}(g \cdot \mathcal{I})$. To conclude, the probability that conditions (1) and (2) are satisfied is greater than

$$1 - \frac{1}{q}\left(\frac{D(D-1)}{2} + (\Delta - n + 1)\left(\binom{\Delta+1}{n} - D\right)\right),$$

where $\Delta = \sum_{i=1}^{n} \deg(f_i)$ and if $p = 0$ or $p > \sum_{i=1}^{n}(\deg(f_i) - 1) + 1$ then condition (3) is satisfied.

5.2 Complexity of Algorithm 1

The matrix T_n can be read from $\mathcal{G}_{>_{\mathrm{drl}}}$ (test in Line 3 of Algorithm 1) if all the monomials of the form $\epsilon_i x_n$ are either in $\mathrm{B}_{>_{\mathrm{drl}}}(\langle \mathcal{G}_{>_{\mathrm{drl}}} \rangle)$ or in $\mathrm{E}_{>_{\mathrm{drl}}}(\langle \mathcal{G}_{>_{\mathrm{drl}}} \rangle)$. Let $F_n = \{\epsilon_i x_n \mid i = 1, \ldots, D\}$, the test in Line 3 is equivalent to test if $F_n \subset \mathrm{B}_{>_{\mathrm{drl}}}(\langle \mathcal{G}_{>_{\mathrm{drl}}} \rangle) \cup \mathrm{E}_{>_{\mathrm{drl}}}(\langle \mathcal{G}_{>_{\mathrm{drl}}} \rangle)$. Since F_n contains exactly D monomials and $\mathrm{B}_{>_{\mathrm{drl}}}(\langle \mathcal{G}_{>_{\mathrm{drl}}} \rangle) \cup \mathrm{E}_{>_{\mathrm{drl}}}(\langle \mathcal{G}_{>_{\mathrm{drl}}} \rangle)$ contains at most $(n+1)D$ monomials; testing if F_n is included in $\mathrm{B}_{>_{\mathrm{drl}}}(\langle \mathcal{G}_{>_{\mathrm{drl}}} \rangle) \cup \mathrm{E}_{>_{\mathrm{drl}}}(\langle \mathcal{G}_{>_{\mathrm{drl}}} \rangle)$ can be done in at most $O(nD^2)$ elementary operations which can be decreased to $O(D)$ elementary operations if we use a hash table. Hence, the cost of the test in Line 4 of Algorithm 1 is negligible in comparison to the complexity of the algorithm in Proposition 3. Hence, the complexity of Algorithm 1 is given by the complexity of F_5 algorithm to compute the DRL Gröbner basis of $g \cdot \mathcal{I}$ and Proposition 3. From [20], the complexities of computing the DRL Gröbner basis of $g \cdot \mathcal{I}$ or $\mathcal{I}$ are the same. Since it is straightforward to see that the number of solutions of these two ideals are also the same we obtain the main result of this paper about the complexity of the PoSSo problem.

THEOREM 18. *Let $\mathbb{K}$ be a field of characteristic zero or a finite field $\mathbb{F}_q$ of sufficiently large characteric p. Let $\mathcal{S} = \{f_1, \ldots, f_n\} \subset \mathbb{K}[x_1, \ldots, x_n]$ be a polynomial system generating a zero-dimen-*

sional radical ideal $\mathcal{I} = \langle \mathcal{S} \rangle$ of degree D. If $(f_1^h, \ldots, f_n^h)$ is a regular sequence such that the degree of each polynomial is uniformly bounded by a parameter d then there exists a Las Vegas algorithm which solves the PoSSo problem in $O(ne^{\omega n} d^{\omega n} + n \log D(D^\omega + D \log D \log \log D))$ arithmetic operations.

PROOF. When $\left(f_1^h, \ldots, f_n^h\right)$ is a regular sequence of polynomials the complexity of computing the DRL Gröbner basis of $\langle \mathcal{S} \rangle$ or $\langle g \cdot \mathcal{S} \rangle$ is bounded by [1, 20] $O\left(n\binom{nd+1}{n}^\omega\right) = O\left(ne^{\omega n} d^{\omega n}\right)$ arithmetic operations in $\mathbb{K}$. From this DRL Gröbner basis, according to Corollary 14, the multiplication matrix T_n can be computed without arithmetic operations in $\mathbb{K}$. Finally, from T_n and the DRL Gröbner basis, thanks to Proposition 3 and Corollary 14 the LEX Gröbner basis can be computed by a probabilistic (respectively deterministic) algorithm in $O(D \log D(D^{\omega-1} + n \log D \log \log D))$ (respectively $O(nD^\omega \log D)$) arithmetic operations in $\mathbb{K}$. $\square$

As previously mentioned, according to the Bézout bound the number of solutions D is bounded by the product of the degrees of the input equations. Since this bound is generically reached we get the following corollary.

COROLLARY 19. *Let $\mathbb{K}$ be a field of characteristic zero or a finite field $\mathbb{F}_q$ of sufficiently large characteristic. Let $\mathcal{S} = \{f_1, \ldots, f_n\} \subset \mathbb{K}[x_1, \ldots, x_n]$ be a generic polynomial system generating a radical ideal. If the degree of each polynomial in $\mathcal{S}$ is equal to a parameter d then there exists a Las Vegas algorithm which solves the PoSSo problem in $\widetilde{O}(e^{\omega n} D^\omega)$ arithmetic operations in $\mathbb{K}$.*

5.3 Benchmarks

In this section we discuss the impact of Algorithm 1 on the practical resolution of the PoSSo problem. Note that algorithms of Proposition 3 to compute the LEX Gröbner basis given the multiplication matrix T_n is of theoretical interest. Indeed, although in theory ω is bounded by 2.3727 in practice in our knowledge the best implementation of the matrix product uses Strassen algorithm [29]. For instance this algorithm is implanted in MAGMA [4] or in LINBOX [6]. Thus, in practice $\omega = \log_2(7) \sim 2.8073$.

As a consequence, in practice the sparse version of Faugère and Mou [13, 14] is much more efficient than the fast version using dense matrix multiplication. Hence, in the following experiments we use the *sparse* version of change of ordering. In Table 1, we give the time to compute the LEX Gröbner basis using the usual algorithm (F_5 followed by a change of odering algorithm) and Algorithm 1. This time is divided into three steps, the first is the time to compute the DRL Gröbner basis using F_5 algorithm, the second is the time to compute the multiplication matrix T_n and the last part is the time to compute the LEX Gröbner basis given T_n using the algorithm in [13]. Since, this algorithm takes advantage of the sparsity of the matrix T_n we also give its density. We also give the number of normal forms to compute (*i.e.* the number of terms of the form $\epsilon_i x_n$ that are not in $\mathrm{B}_{>_{\mathrm{drl}}}(\mathcal{I})$ or in $\mathrm{E}_{>_{\mathrm{drl}}}(\mathcal{I})$ (resp. in $\mathrm{B}_{>_{\mathrm{drl}}}(g \cdot \mathcal{I})$ or in $\mathrm{E}_{>_{\mathrm{drl}}}(g \cdot \mathcal{I})$).

The experiments are performed on various polynomial systems such as random systems (n dense polynomials of degree d with random coefficients), systems coming from economical problems [23] named "Eco" and systems coming from the resolution of the elliptic curve discrete logarithm problem on Edwards curves [10] named "Edwards (weights)" or on the "Well Know Group" 3 of the IPSEC Oakley key determination [12, 25] named "Oakley".

We also present experiments on a "pathological" case for our algorithm in the sense that the system in input is already a DRL Gröbner basis. Thus, while the usual algorithm does not have to compute the DRL Gröbner basis, our algorithm needs to compute the DRL Gröbner basis of $g \cdot \mathcal{I}$. The system in input is of the form $\mathcal{S} = \{f_1, \ldots, f_n\} \subset \mathbb{F}_{65521}[x_1, \ldots, x_n]$ with $\mathrm{LT}_{>_{\mathrm{drl}}}(f_i) = x_i^2$. Hence, the monomials in the basis $\mathrm{B}_{>_{\mathrm{drl}}}(\mathcal{I})$ are all the monomials of degree at most one in each variable. The degree of the ideal D is then 2^n. The monomials $\epsilon_i x_n$ that are not in $\mathrm{B}_{>_{\mathrm{drl}}}(\mathcal{I})$ or in $\mathrm{E}_{>_{\mathrm{drl}}}(\langle \mathcal{S} \rangle)$ are of the form $x_n^2 m$ where m is a monomial in $x_1, \ldots, x_{n-1}$ of total degree greater than zero and linear in each variable. By consequence, using the usual algorithm we have to compute $2^{n-1} - 1$ normal forms to compute only T_n. In the whole of Table 1, we omit the coefficient field when it is $\mathbb{F}_{65521}$.

REMARK 20. *In practice, before applying a linear change of variables we check if the system is in the generic case (corresponding to the column labelled "Generic case" in Table 1). We apply a change of variables only if it is required to compute very efficiently the matrix T_n and to obtain a* Shape Position *ideal. In this generic case, our algorithm is as efficient as the usual algorithm but provide a better complexity bound. It is the case for instance when solving the elliptic curve discrete logarithm problem as in [10].*

One can note in Table 1 that in the usual algorithm, when the system is not in the "generic case", the bottleneck of the resolution of the PoSSo problem is the change of ordering due to the construction of the multiplication matrix T_n. Since our algorithm allows to compute very efficiently the matrix T_n (for instance for the pathological example with $n = 11$, less than one second in comparison to 7 520 seconds for the usual algorithm), the most time consuming step becomes the computation of the DRL Gröbner basis.

Moreover, still when the system is not in the "generic case" the total running time of our algorithm is far less than that of the usual algorithm. For instance, for the system "Edwards" with $n = 5$ the PoSSo problem can now be solved in less than six hours whereas we could not solve this instance of the PoSSo problem using the usual algorithm.

REMARK 21. *In Table 1, for the "Oakley" example we do not give the time for the change of ordering using the usual algorithm because it is not implemented in FGb for $\mathbb{K} = \mathbb{F}_{2^{31}}$. However, this example shows that our method still works in characteristic two. Indeed, with the usual algorithm we need to compute 480 normal forms to compute T_n while with our algorithm the number of normal forms is decreased to 0.*

We do not have explanations for all the benefits in practice of our method. Especially why the computation of the LEX Gröbner basis is speeded up for the "Eco" examples while the density of the matrix is increased. This is probably due to a particular structure. In general our method seems more efficient in practice. Actually, for the moment we do not find any counterexample.

6. ACKNOWLEDGMENTS

The authors would like to thanks André Galligo and Daniel Lazard for fruitful discussions about generic initial ideals and Vanessa Vitse for providing us with the pathological example. This work was partly supported by the HPAC grant of the French National Research Agency (HPAC ANR-11-BS02-013).

7. REFERENCES

[1] M. Bardet, J.-C. Faugère, B. Salvy, and B. Yang. Asymptotic behaviour of the degree of regularity of semi-regular polynomial systems. In P. Gianni, editor, *The Effective Methods in Algebraic Geometry Conference, Mega 2005*, pages 1 –14, May 2005.

[2] D. Bayer and M. Stillman. A theorem on refining division orders by the reverse lexicographic order. *Duke Mathematical Journal*, 55(2):321–328, 1987.

System	n	D	Algorithm	Generic case	$\mathcal{G}_{>_{\mathrm{drl}}}$	T_n	#NF	Density	$\mathcal{G}_{>_{\mathrm{lex}}}$	Total
Random $d=2$	15	32 768	usual	$\times$	**1 580s**	41.5s	0	18.52%	1 330s	2 950s
			This work	Yes	**1 580s**	41.5s	0	18.52%	1 330s	2 950s
Random $d=6$	6	46 656	usual	$\times$	632s	20.3s	0	8.36%	**1 700s**	2 350s
			This work	Yes	632s	20.3s	0	8.36%	**1 700s**	2 350s
Random $d=30$	3	27 000	usual	$\times$	48.7s	0.9s	0	2.20%	**95.6s**	145s
			This work	Yes	48.7s	0.9s	0	2.20%	**95.6s**	145s
Oakley $\mathbb{F}_{2^{31}}$	5	4 096	usual	$\times$	2.97s	Rem. 21	480			
			This work	No	**2.85s**	0.08s	0	6.79%	2.44s	5.37s
Eco	13	2 048	usual	$\times$	28.2s	**36.5s**	1 153	12.09%	0.43s	65.1s
			This work	No	**12.0s**	0.18s	0	27.52%	0.23s	12.4s
	14	4 096	usual	$\times$	176s	**1 100s**	2 353	11.50%	1.47s	1 280s
			This work	No	**57.0s**	0.74s	0	26.41%	1.23s	59.0s
	15	8 192	usual	$\times$	1 030s	$>$ **2 days**	4 853			$>$ 2 days
			This work	No	**348s**	3.47s	0	24.95%	30.6s	382s
Edwards	5	65 536	usual	$\times$	12 300s	$>$ **2 days**				$>$ 2 days
			This work	No	**12 300s**	40.8s	0	9.31%	7 820s	20 200s
Edwards weights	5	65 536	usual	$\times$	566s	15.1s	0	3.30%	**2 150s**	2 730s
			This work	Yes	566s	15.1s	0	3.30%	**2 150s**	2 730s
Pathological	9	512	usual	$\times$	0s	**12.8s**	255	32.81%	0.01s	12.8s
			This work	No	$<$ 0.01s	$<$ 0.01s	0	23.68%	$<$ 0.01s	$<$ 0.01s
	11	2 048	usual	$\times$	0s	**7 520s**	1 023	31.93%	23.0s	7 540s
			This work	No	**5.02s**	0.15s	0	21.53%	0.13s	5.28s
	16	65 536	usual	$\times$	0s	$>$ **2 days**	32 767			$>$ 2 days
			This work	No	**38 100s**	195s	0	18.33%	14 300s	52 600s

Table 1: Comparison of the usual algorithm for solving the PoSSo problem and Algorithm 1, the proposed algorithm. Computation with FGb on a 3.47 GHz Intel Xeon X5677 CPU.

[3] E. Becker, T. Mora, M. G. Marinari, and C. Traverso. The shape of the shape lemma. In *Proceedings of the international symposium on Symbolic and algebraic computation*, ISSAC '94, pages 129–133, New York, NY, USA, 1994. ACM.

[4] W. Bosma, J. Cannon, and C. Playoust. The Magma algebra system. I. The user language. *Journal of Symbolic Computation*, 24(3–4):235–265, 1997.

[5] A. Bostan, B. Salvy, and E. Schost. Fast algorithms for zero-dimensional polynomial systems using duality. *Applicable Algebra in Engineering, Communication and Computing*, 14(4):239–272, 2003.

[6] J.-G. Dumas, T. Gautier, M. Giesbrecht, P. Giorgi, B. Hovinen, E. Kaltofen, B. D. Saunders, W. J. Turner, and G. Villard. LinBox: A generic library for exact linear algebra. In A. M. Cohen, X.-S. Gao, and N. Takayama, editors, *ICMS'2002, Proceedings of the 2002 International Congress of Mathematical Software, Beijing, China*, pages 40–50. World Scientific Pub., August 2002.

[7] D. Eisenbud. *Commutative Algebra with a View Toward Algebraic Geometry*. Springer, 1995.

[8] J.-C. Faugère, P. Gaudry, L. Huot, and G. Renault. Polynomial systems solving by fast linear algebra, 2013. http://arxiv.org/abs/1304.6039.

[9] J.-C. Faugère. A new efficient algorithm for computing Gröbner bases without reduction to zero (F5). In *Proceedings of the 2002 international symposium on Symbolic and algebraic computation*, ISSAC '02, pages 75–83, New York, NY, USA, 2002. ACM.

[10] J.-C. Faugère, P. Gaudry, L. Huot, and G. Renault. Using symmetries in the index calculus for elliptic curves discrete logarithm. *Journal of Cryptology*, pages 1–41, 2013. doi 10.1007/s00145-013-9158-5.

[11] J.-C. Faugère, P. Gianni, D. Lazard, and T. Mora. Efficient computation of zero-dimensional Gröbner bases by change of ordering. *Journal of Symbolic Computation*, 16(4):329–344, 1993.

[12] J.-C. Faugère, L. Huot, A. Joux, G. Renault, and V. Vitse. Symmetrized summation polynomials: Using small order torsion points to speed up elliptic curve index calculus, 2014. To appear in the proceedings of the 33rd Annual International Conference on the Theory and Applications of Cryptographic Techniques (EUROCRYPT 2014).

[13] J.-C. Faugère and C. Mou. Fast algorithm for change of ordering of zero-dimensional Gröbner bases with sparse multiplication matrices. In *ISSAC '11: Proceedings of the 2011 international symposium on Symbolic and algebraic computation*, ISSAC '11, pages 1–8, New York, NY, USA, 2011. ACM.

[14] J.-C. Faugère and C. Mou. Sparse FGLM algorithms. http://hal.inria.fr/hal-00807540, 2013.

[15] A. Galligo. *A Propos du Théorème de Préparation de Weierstrass*. PhD thesis, Institut de Mathématique et Sciences Physiques de l'Université de Nice, 1973.

[16] P. Gianni and T. Mora. Algebraic solution of systems of polynomial equations using Gröbner bases. In *Applied Algebra, Algebraic Algorithms and Error Correcting Codes, Proceedings of AAECC-5, volume 356 of LNCS*, pages 247–257. Springer, 1989.

[17] E. Jonckheere and C. Ma. A simple Hankel interpretation of the Berlekamp-Massey algorithm. *Linear Algebra and its Applications*, 125:65–76, 1989.

[18] W. Keller-Gehrig. Fast algorithms for the characteristic polynomial. *Theor. Comput. Sci.*, 36:309–317, June 1985.

[19] Y. N. Lakshman. On the complexity of computing a Gröbner basis for the radical of a zero dimensional ideal. In *Proceedings of the twenty-second annual ACM symposium on Theory of computing*, STOC '90, pages 555–563, New York, NY, USA, 1990. ACM.

[20] D. Lazard. Gröbner bases, gaussian elimination and resolution of systems of algebraic equations. In J. van Hulzen, editor, *Computer Algebra*, volume 162 of *Lecture Notes in Computer Science*, pages 146–156. Springer Berlin / Heidelberg, 1983.

[21] J. Massey. Shift-register synthesis and bch decoding. *Information Theory, IEEE Transactions on*, 15(1):122–127, 1969.

[22] G. Moreno-Socías. Degrevlex Gröbner bases of generic complete intersections. *Journal of Pure and Applied Algebra*, 180(3):263–283, 2003.

[23] A. Morgan. *Solving Polynominal Systems Using Continuation for Engineering and Scientific Problems*. Society for Industrial and Applied Mathematics, Philadelphia, PA, USA, 2009.

[24] B. Mourrain and V. Y. Pan. Asymptotic acceleration of solving multivariate polynomial systems of equations. In *Proceedings of the thirtieth annual ACM symposium on Theory of computing*, pages 488–496. ACM, 1998.

[25] H. Orman. The oakley key determination protocol, 1998.

[26] K. Pardue. *Nonstandard Borel-Fixed Ideals*. PhD thesis, Brandeis University, 1994.

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

[28] J. T. Schwartz. Fast probabilistic algorithms for verification of polynomial identities. *J. ACM*, 27(4):701–717, Oct. 1980.

[29] V. Strassen. Gaussian elimination is not optimal. *Numerische Mathematik*, 13(4):354–356, 1969.

[30] V. Vassilevska Williams. Multiplying matrices faster than Coppersmith-Winograd. In *Proceedings of the 44th symposium on Theory of Computing*, pages 887–898. ACM, 2012.

[31] D. H. Wiedemann. Solving sparse linear equations over finite fields. *IEEE Trans. Inf. Theor.*, 32(1):54–62, 1986.

[32] R. Zippel. Probabilistic algorithms for sparse polynomials. In E. Ng, editor, *Symbolic and Algebraic Computation*, volume 72 of *Lecture Notes in Computer Science*, pages 216–226. Springer Berlin Heidelberg, 1979.