UNIVERSITÉ PIERRE ET MARIE CURIE - PARIS 6

ÉCOLE DOCTORALE EDITE

THÈSE

pour obtenir le titre de

Docteur en Sciences

de l'Université Pierre et Marie Curie - Paris 6
Mention : INFORMATIQUE

Présentée et soutenue par

Pierre-Jean SPAENLEHAUER

RÉSOLUTION DE SYSTÈMES MULTI-HOMOGÈNES ET DÉTERMINANTIELS

ALGORITHMES – COMPLEXITÉ – APPLICATIONS

Thèse dirigée par Jean-Charles Faugère et Mohab Safey El Din
préparée au Laboratoire d'Informatique de Paris 6 (UPMC)

Soutenue le 9 Octobre 2012 après avis des rapporteurs :

Bernd STURMFELS	-	Professor, University of California, Berkeley
Gilles VILLARD	-	Directeur de Recherche CNRS, ENS Lyon

devant le jury composé de

Jean-Claude BAJARD	-	Professeur, Université Pierre et Marie Curie
Jean-Charles FAUGÈRE	-	Directeur de Recherche INRIA, CRI Paris-Rocquencourt
Antoine JOUX	-	Professeur associé, Université de Versailles
Mohab SAFEY EL DIN	-	Professeur, Université Pierre et Marie Curie
Bruno SALVY	-	Directeur de Recherche INRIA, ENS Lyon
Gilles VILLARD	-	Directeur de Recherche CNRS, ENS Lyon

Remerciements

Mes premiers remerciements vont à Jean-Charles Faugère et Mohab Safey El Din pour leur disponibilité, leurs encouragements, et pour le temps qu'ils ont consacré aux relectures de ce manuscrit et des articles écrits pendant ces trois ans. Ils m'ont fait découvrir le domaine du calcul formel et c'est grâce à eux que je souhaite continuer dans le monde de la recherche. Malgré leurs emplois du temps chargés, ils ont toujours pris le temps de me conseiller. Si cette thèse existe, c'est grâce à leur encadrement à la fois exigeant, mais toujours stimulant et rempli de nouveaux challenges à relever.

Je remercie vivement Gilles Villard et Bernd Sturmfels, qui m'ont fait l'honneur de rapporter ma thèse. L'intérêt qu'ils ont porté à mes travaux et leur lecture minutieuse ont grandement contribué à l'amélioration de ce manuscrit.

Je remercie chaleureusement Jean-Claude Bajard, Antoine Joux et Bruno Salvy d'avoir accepté de faire partie de mon jury de thèse. Bruno a notamment été présent depuis mes balbutiements dans le monde de la recherche lors du cours du MPRI (j'en profite pour remercier Alin Bostan, Frédéric Chyzak et Marc Giusti) qui n'est pas étranger à mon choix de me lancer dans une thèse en calcul formel il y a trois ans.

Ma thèse s'est déroulée au sein de l'équipe POLSYS du LIP6, et j'ai eu beaucoup de plaisir à y vivre durant ces trois années. Je remercie tous les membres que j'y ai cotoyés : Ludovic Perret, pour m'avoir encadré lors de mes stages de M1 et de M2 et pour m'avoir coaché lors de mon premier exposé scientifique à l'étranger ; Guénaël Renault, pour sa bonne humeur constante et les innombrables surnoms qu'il a pu m'attribuer (et pour le “MIAU”) ; Daniel Lazard, pour les discussions scientifiques toujours riches en enseignements. Durant la plus grande partie de ces trois années, j'ai passé mes journées dans le bureau 338 où règnait une ambiance toujours chaleureuse entre ses occupants : par ordre chronologique approximatif, Luk, Wei, Christopher, Chenqi, Aurélien, Louise, Mourad, Frédéric, Rina, Jules et Thibaut. Entre les débats scientifiques au tableau, les nocturnes au labo les veilles de deadlines, les pauses café, les verres après le travail et les soirées et sorties, l'ambiance était toujours à la camaraderie et au soutien lors des moments difficiles. Merci pour tous ces moments passés ensemble.

Je remercie l'ensemble de l'équipe POLSYS, notamment ceux que j'ai cotoyés pendant des périodes plus courtes : Guillaume, Sajjad, Martin, Alexandre, Adrien, Elias et Jérémy. Merci également aux membres de l'équipe PEQUAN, et en particulier Pierre, Stef et Fabienne avec qui j'ai partagé un bureau lors de mon arrivée au LIP6.

Durant ces trois années de thèse, j'ai aussi eu le plaisir d'enseigner, et je remercie Valérie Ménissier-Morain, Jean-Claude Bajard, Bernd Amann et Béatrice Bérard de m'avoir donné cette opportunité.

Durant cette période qu'est la thèse, les doutes et remises en question sont inévitables. Je remercie ma mère, ma soeur, mon père ainsi que mes amis qui m'ont toujours soutenu, chacun à leur manière. Je pense notamment à Thomas, Hélène, Mathieu, Ariadna, Mélanie, Rodolfo, Laurent et Kévin. Merci à Ben et Fred, mes compagnons de musique depuis maintenant plus de onze ans. Une mention spéciale à Thomas, pour les soirées guitare et les (nombreuses) versions acoustiques de “Sober”, qui ont été le leitmotiv de cette dernière année de thèse.

Enfin, je ne peux conclure qu'en remerciant Sophie pour sa présence et son soutien indéfectible durant cette période intense de fin de thèse.

Résumé

De nombreux systèmes polynomiaux multivariés apparaissant en Sciences de l'Ingénieur possèdent une structure algébrique spécifique. En particulier, les structures multi-homogènes, déterminantielles et les systèmes booléens apparaissent dans une variété d'applications. Une méthode classique pour résoudre des systèmes polynomiaux passe par le calcul d'une base de Gröbner de l'idéal associé au système. Cette thèse présente de nouveaux outils pour la résolution de tels systèmes structurés.
D'une part, ces outils permettent d'obtenir sous des hypothèses de généricité des bornes de complexité du calcul de base de Gröbner de plusieurs familles de systèmes polynomiaux structurés (systèmes bilinéaires, systèmes déterminantiels, systèmes définissant des points critiques, systèmes booléens). Ceci permet d'identifier des familles de systèmes pour lequels la complexité arithmétique de résolution est polynomiale en le nombre de solutions.
D'autre part, cette thèse propose de nouveaux algorithmes qui exploitent ces structures algébriques pour améliorer l'efficacité du calcul de base de Gröbner et de la résolution (systèmes multi-homogènes, systèmes booléens). Ces résultats sont illustrés par des applications concrètes en cryptologie (cryptanalyse des systèmes MinRank et ASC), en optimisation et en géométrie réelle effective (calcul de points critiques).

Abstract

Multivariate polynomial systems arising in Engineering Science often carry algebraic structures related to the problems they stem from. In particular, multi-homogeneous, determinantal structures and boolean systems can be met in a wide range of applications. A classical method to solve polynomial systems is to compute a Gröbner basis of the ideal associated to the system. This thesis provides new tools for solving such structured systems in the context of Gröbner basis algorithms.
On the one hand, these tools bring forth new bounds on the complexity of the computation of Gröbner bases of several families of structured systems (bilinear systems, determinantal systems, critical point systems, boolean systems). In particular, it allows the identification of families of systems for which the complexity of the computation is polynomial in the number of solutions.
On the other hand, this thesis provides new algorithms which take profit of these algebraic structures for improving the efficiency of the Gröbner basis computation and of the whole solving process (multi-homogeneous systems, boolean systems). These results are illustrated by applications in cryptology (cryptanalysis of MinRank), in optimization and in effective real geometry (critical point systems).

Contents

Introduction

Problem statement

Investigating algebraic systems from a computational viewpoint is of first importance since such systems arise in many areas of Engineering Sciences and Computer Science. For instance, the security of several cryptographic primitives is strongly related to the difficulty of solving algebraic systems. Such systems also appear naturally in optimization problems, when the constraints are given by polynomial equalities or inequalities. Among other applications, Effective Geometry, Computer Aided Geometric Design, Game Theory, Control Theory are areas where such systems arise frequently.

Polynomial System Solving (PoSSo for short) and *elimination theory* have a long history and were already studied by Lagrange in the 18th century. Following works initiated by Kronecker and Hilbert, a new milestone was reached in the beginning of the 20th century by Macaulay with the definition of the multivariate resultant. The next algorithmic breakthrough was obtained by Buchberger in his Ph.D. thesis [Buc65] where he defined the notion of Gröbner bases and gave the first algorithm to compute them.

With the advent of computers and computer algebra during the last decades, Gröbner basis algorithms have been thoroughly investigated. In particular, the F_4 algorithm [Fau99] uses linear algebra to obtain huge speed-ups compared to Buchberger algorithm. In the F_5 algorithm [Fau02], a new criterion is used to avoid useless computations. These algorithms are nowadays among the most standard techniques to solve *symbolically* polynomial systems coming from applications.

From a theoretical viewpoint, the PoSSo problem in finite fields is NP-hard: it is intrinsically of exponential complexity in the number of variables. Indeed, the Bézout bound states that the number of solutions of generic systems with as many equations as variables over an algebraically closed field is exponential in the number of variables. However, systems coming from practical applications are not generic: they carry structures arising from the problem they stem from. In particular, their number of solutions is less than that of a dense generic system.

In this thesis, we will mainly focus on determinantal, multi-homogeneous and quadratic boolean systems, which arise for instance in Cryptology, in Optimization and in Real geometry. Experimentally, these systems are easier to solve than generic dense systems of the same degrees. Consequently, it is natural to ask the following questions:

1. What is the asymptotic complexity of Gröbner basis algorithms when the input is such a system?

2. Can we solve generic determinantal, multi-homogeneous and critical point systems with a complexity which is asymptotically *polynomial* in the number of solutions?

3. Explain the experimental behavior observed in the case of structured systems. Can the practical timings and memory requirements for solving structured systems be estimated *a priori*?

4. Can we design variants of Gröbner basis algorithms dedicated to such systems in order to obtain practical speed-ups?

Motivations

Applications in Cryptology, Geometry and Optimization

Systems of polynomial equations arise in several applicative fields. For instance, in Cryptology, several schemes can be modeled by polynomial systems such that their solutions correspond to secret information. Therefore, the security of such cryptosystems is directly related to the difficulty of solving the corresponding algebraic systems. This process of retrieving secret information by solving algebraic systems is called *Algebraic Cryptanalysis*. These polynomial systems are usually structured since the cryptosystems they stem from have to verify a set of properties. We give below examples of such properties.

- **Trapdoor.** In asymmetric Cryptology, the plaintext should be easily recoverable from the ciphertext once the secret key is known. This yields structure that can be exploited algebraically. Recent examples of such algebraic cryptanalysis are HFE (and variants) [FJ03] and IP [FP06]. In both cases, structured systems have to be solved.

- **Key reduction.** The McEliece cryptosystem is a typical example of an asymmetric scheme whose main drawback is the size of the keys. Therefore, tremendous efforts have been made to reduce the sizes of the keys. This is generally achieved by adding structure to the cryptosystem (see e.g. [BCGO09, MB09]). However, in [FOPT10], the authors show that the key can be retrieved by solving a "quasi-bilinear" system and that the corresponding structure in the algebraic system leads to significant reduction of the security.

- **Zero-knowledge authentication.** In zero-knowledge authentication schemes, someone wants to prove their identity (i.e. to prove that they know a secret which is not shared with anybody else) without revealing any information. This is usually achieved by designing a protocol where the prover has to be able to answer a family of problems with the secret. It means that there are invariance properties which can be translated to the corresponding algebraic system. A typical example is the MinRank authentication scheme [Cou01]. In Section 8.2, we show how this structure can be algorithmically exploited by solving a determinantal system.

Another field of application is geometry over the real field $\mathbb{R}$ and optimization. Indeed, local optima of a polynomial function P under polynomial constraints $f_1 = \ldots = f_p = 0$ are reached at points where a Jacobian matrix is rank defective. Therefore these points can be computed by considering the algebraic system $(f_1, \ldots, f_p)$ and the maximal minors of the latter Jacobian matrix. Computing critical points of such applications is also an important routine of the so-called *critical point method*, which can be used for answering several problems in real geometry: quantifier elimination [HS11], deciding whether a semi-algebraic set is empty or not, computing at least one point by connected component in a semi-algebraic set [SS03], answering connectivity queries [Can93, SS10],...

There are also several other fields where structured algebraic systems appear: game theory [HKL$^+$11], control theory [Hen08], computer aided geometric design [ELLS09], coding theory [OJ02],...

Polynomial System Solving

Representation of the solutions

Before going further, it is important to state what we mean by "Solving Systems of Polynomial Equations". Let $\mathbb{K}$ be a perfect field (i.e. all its finite extensions are separable; finite fields and fields of characteristic 0 are perfect) and $f_1(x_1, \ldots, x_n) = \cdots = f_m(x_1, \ldots, x_n) = 0$ be an algebraic system in $\mathbb{K}[x_1, \ldots, x_n]$. In this thesis, we mainly consider systems which have finitely-many solutions in the algebraic closure $\overline{\mathbb{K}}$ of $\mathbb{K}$ (i.e. 0-*dimensional systems*). A good representation of the solutions is another system from which properties of the solutions can be read off easily.

Solving 0-dimensional systems in algebraically closed fields. A good representation of the solutions in $\overline{\mathbb{K}}$ is given by a *rational parametrization*: it is given by a univariate polynomial $h \in \mathbb{K}[u]$ and by n rational functions $g_1, \ldots, g_n \in \mathbb{K}(u)$ such that the solutions of the polynomial system are parametrized by the solutions of h:

$$\begin{gathered} f_1(x_1, \ldots, x_n) = \cdots = f_m(x_1, \ldots, x_n) = 0 \\ \Updownarrow \\ \exists u \in \overline{\mathbb{K}}, h(u) = 0, x_1 = g_1(u), \ldots, x_n = g_n(u). \end{gathered}$$

Such representation does not always exist. However it exists after almost all linear change of coordinates on the x_i variables. Under genericity assumptions such a parametrization is given by a lexicographical *Gröbner basis* of the ideal $\langle f_1, \ldots, f_m \rangle$.

Solving in finite fields. For several applications (especially in Cryptology), we want to find solutions of polynomial systems in $\mathbb{K}^n$, where $\mathbb{K}$ is a finite field. In that case, we want the list of solutions as vectors in $\mathbb{K}^n$. For some applications, we only need one solution of the system. These vectors in $\mathbb{K}^n$ can be easily computed as soon as a *lexicographical Gröbner basis* of the ideal $\langle f_1, \ldots, f_m \rangle$ is known.

A Gröbner basis is a set of generators of the ideal verifying useful properties. Consequently the specification of what we mean by "Solving Polynomial System" in this thesis is

---

Algorithm 1 Specification: Solving 0-Dimensional Polynomial Systems

---

Input: $f_1, \ldots, f_m \in \mathbb{K}[x_1, \ldots, x_n]$ such that these polynomials vanish on finitely-many points in the algebraic closure $\overline{\mathbb{K}}^n$.
Output: G a *lexicographical Gröbner basis* of the ideal $\langle f_1, \ldots, f_m \rangle$.

We focus in this thesis on the *arithmetic complexity* of the algorithms involved, i.e. the number of operations in $\mathbb{K}$. In the case of finite fields, this provides good estimates of the running time of Gröbner basis engines.

Gröbner basis algorithms

Gröbner bases were introduced by Buchberger in his Ph.D. thesis [Buc65] to solve the so-called *Ideal Membership Problem*, i.e. given a finite family of polynomials $f_1, \ldots, f_m, h \in \mathbb{K}[x_1, \ldots, x_n]$, deciding whether h belongs to the ideal $\langle f_1, \ldots, f_m \rangle$. The main idea to solve this problem is to use *pseudo-division* algorithms: given a monomial ordering $\prec$ and denoting by $\mathsf{LM}_\prec(\cdot)$ the leading monomial of a polynomial, if g_1 and g_2 are two polynomials and $\mathsf{LM}_\prec(g_2)$ divides $\mathsf{LM}_\prec(g_1)$, we can define the *top-reduction* of g_1 by g_2:

$$g_1 \xrightarrow{g_2} g_1 - \frac{\mathsf{LM}_\prec(g_1)}{\mathsf{LM}_\prec(g_2)} g_2.$$

Consequently, the leading monomial of the reduced polynomial is smaller than that of g_1. This can be seen as a term rewriting rule $\mathsf{LM}(g_2) \to \mathsf{LM}_\prec(g_2) - g_2$. The set of such rewriting rules for the polynomials $f_1, \ldots, f_m$ is *Noetherian* but *not confluent*.

A Gröbner basis of the ideal is a family of polynomials generating the same ideal such that this set of rewriting rules is confluent. Therefore, a polynomial belongs to the ideal if and only if it reduces to zero.

The main principle of Buchberger's algorithm is to find critical pairs, to reduce them and to add the newly found rules to the rewriting system. This operation is repeated until the system becomes confluent.

In the last decades, the algorithms F_4 [Fau99] and F_5 [Fau02] improved Buchberger's algorithm. In the F_4 algorithm, row echelon form computations are used to reduce simultaneously several critical pairs. In the F_5 algorithm, a criterion detects useless critical pairs and thus avoids their reduction. These two improvements led to huge practical speed-ups for computing Gröbner bases.

Another important algorithm is the so-called FGLM algorithm [FGLM93, FM11]. This algorithm is used for 0-dimensional systems (i.e. systems which have finitely-many solutions); it takes as input a Gröbner basis for some monomial ordering $\prec_1$ and another monomial ordering $\prec_2$ and it outputs a Gröbner basis for $\prec_2$.

Solving strategy for 0-dimensional systems. The FGLM algorithm is central for solving 0-dimensional systems since it is usually more efficient to compute first a Gröbner basis for the so-called *graded reverse lexicographical ordering* (grevlex) with the F_5 algorithm and then to convert it into a Gröbner basis for the *lexicographical ordering* (lex) by using the FGLM algorithm. Indeed, the degrees of the polynomials occurring in the grevlex Gröbner basis are significantly smaller than the degrees in the lex basis. Hence the F_5 algorithm computes grevlex Gröbner bases more efficiently than lex bases. Moreover, the complexity of the FGLM algorithm is well understood and is polynomial in the number of solutions of the system. This solving strategy (i.e. using successively the F_5 algorithm and the FGLM algorithm) is used in most of the chapters of this thesis.

Related algorithms

There exist a wide range of methods and algorithms for solving algebraic systems. In this section, a few of the most standard methods for solving polynomial systems are briefly described. It is not easy to compare these methods since they all have their own specificities and their complexity bounds do not involve the same parameters of the systems.

Resultants. Historically, the first algorithms for eliminating variables were obtained by computing resultants. If $f, g \in \mathbb{K}[t]$ are two univariate polynomials, their resultant (i.e. the determinant of the Sylvester matrix) is a polynomial function of their coefficients which is equal to zero if and only if the two polynomials share a common root. This notion was generalized by Macaulay to the multivariate case: if $f_1, \ldots, f_n \in R[x_1, \ldots, x_n]$ are *homogeneous* polynomials (where R is a unique factorization domain), their multivariate resultant is a polynomial function of their coefficients that is zero if they share a common *non-zero* root. This can be used for elimination as follows: if $\mathbf{F} = (f_1, \ldots, f_n) \in \mathbb{K}[x_1, \ldots, x_n]^n$ is a non-homogeneous family of polynomials, we can treat each f_i as a polynomial in the ring $\mathbb{K}[x_1][x_2, \ldots, x_n]$. By adding a homogenizing variable h, we obtain a homogeneous system of n equations in n unknowns in $\mathbb{K}[x_1][x_2, \ldots, x_n, h]$ (coefficients are in $\mathbb{K}[x_1]$). Their multivariate resultant is a univariate polynomial in $\mathbb{K}[x_1]$ and its roots correspond to the first coordinates of the solutions of the system $f_1 = \cdots = f_n = 0$ in generic situations.

Such resultant techniques have been extended to a general theory including specific systems (see e.g. [EM09, DE03] for multi-homogeneous resultants and [Bus04] for determinantal resultants).

Geometric resolution. The Geometric resolution was proposed in [GLS01]. It relies on geometric techniques such as lifting points into curves by using Newton iteration and then intersecting them with hypersurfaces. This algorithm is probabilistic since it relies on random choices of linear changes of coordinates but the probability that it fails is negligible. It has been implemented in the MAGMA package `Kronecker`[1]. From a theoretical viewpoint, one of the main feature of this algo-

[1] available at `http://lecerf.perso.math.cnrs.fr/software/kronecker/distribution.html`

rithm is that its complexity is polynomial in the maximum of the degrees of the intermediate ideals $\langle f_1\rangle, \langle f_1, f_2\rangle, \ldots, \langle f_1, \ldots, f_m\rangle$.

Homotopy continuation. In the last decades, tremendous efforts have been put into semi-numerical algorithms for solving numerically systems of polynomial equations. One of the most successful framework, from the viewpoint of numerical stability as well as efficiency, is the homotopy continuation method. In order to solve a system $f_1 = \cdots = f_m = 0$ which has $\mathsf{DEG}(\langle \mathbf{F}\rangle)$ isolated solutions, the general idea is to start with another system with $\mathsf{DEG}(\langle \mathbf{F}\rangle)$ known solutions, and then to deform step by step the system, and recompute the approximate solutions of the deformed system. This is done usually with Newton iteration techniques. At the end of the path, we obtain approximate solutions of the system $f_1 = \cdots = f_m = 0$. Variants of homotopy methods dedicated to multi-homogeneous and determinantal systems have been also proposed [MS87, HSS98]. Also, efficient implementations of these tools are available in the packages `Bertini`[2] and `PHCpack`[3] [Ver11], and there exist tools for certifying the correctness of the approximations [BL09, HS10].

Structured systems

In this thesis, we focus essentially on four kinds of structured systems:

1. (**Multi-homogeneous systems.**) These systems are homogeneous with respect to several blocks of variables. Roughly speaking, they generalize multi-linear systems by allowing higher degrees. They arise in practical applications as soon as there are blocks of variables representing quantities of different nature.

2. (**Determinantal systems.**) These systems are related to the so-called *Generalized MinRank Problem*: given a matrix M whose entries are multivariate polynomials, find the points where the rank of the evaluation of M is at most a given value $r \in \mathbb{N}$. These points are zeros of all minors of size $r+1$ of M.

3. (**Critical point systems.**) The critical points of a polynomial map restricted to an algebraic variety V are defined by the points of the variety such that a Jacobian matrix is rank defective. Consequently, they are the intersection of V and of the solutions of a generalized MinRank problem. Computing these points is a central subroutine of several algorithms in Optimization and in Effective Real Geometry.

4. (**Quadratic boolean systems.**) Searching for boolean solutions of quadratic polynomial systems is a crucial NP-hard problem and the security of several modern multivariate cryptosystems directly relies on its difficulty. Properly speaking, these systems are not really structured. The structure comes from the fact that we are searching for solutions in the field GF_2 (and not in its algebraic closure): the Fröbenius relations $x_i^2 = x_i$ add a specific combinatorial structure to the ideal generated by the polynomials. Moreover, the tools used for investigating these systems (Hilbert series, degree of regularity,...) are similar to those used for structured systems.

In the next section, we present the main results obtained. We focus on four aspects of these structured systems:

1. (**Complexity.**) New asymptotic complexity bounds for Gröbner basis algorithms when the input is such a system.

2. (**Algorithms.**) New Gröbner basis algorithms dedicated to these systems.

[2] available at `http://www.nd.edu/~sommese/bertini/`
[3] available at `http://homepages.math.uic.edu/~jan/download.html`

3. (**Structural results.**) Theoretical results on the combinatorial structure of ideals generated by these systems under genericity assumptions.

4. (**Applications in Cryptology.**) We present results obtained by applying the complexity and theoretical results to systems arising from applications.

Genericity. Many results in this thesis are true *under genericity assumptions*. This means that there holds for almost all systems of a given shape (multi-homogeneous, determinantal, critical point systems,...). This is usually achieved by considering the coefficients of these systems as formal parameters (which are thus algebraically independent). Then properties of this *generic system* are proved, and then it is sufficient to show that for almost every specialization of these parameters, the specialized system verifies the same properties (by "almost all", we mean "outside a Zariski proper closed subset of the space of coefficients").

Main results

Complexity results

We give in this thesis new complexity bounds for solving these systems. In the following, ω is a feasible exponent for the matrix multiplication ($\omega = 2.373$ with Williams' algorithm [Vas11]).

One of the main tools used for the analysis of the combinatorial structure of ideals is the so-called *Hilbert series*. It provides information on the combinatorial structure of graded algebras, and is related to the ranks of the matrices that appear during the execution of the F_5 algorithm.

If R is a graded $\mathbb{K}$-algebra, its Hilbert series is the power series

$$\mathsf{HS}_R(t) = \sum_{d\in\mathbb{N}} \dim_{\mathbb{K}}(R_d) t^d \in \mathbb{N}[[t]]$$

where R_d is the $\mathbb{K}$-vector space of homogeneous elements of degree d.

In this introduction, we focus on the Hilbert series of the quotient algebra $\mathbb{K}[x_1, \ldots, x_n]/I$, where I is a 0-dimensional ideal. In particular, when a system $f_1 = \cdots = f_m = 0$ has finitely-many solutions and when the ideal generated by the homogeneous parts of highest degrees $\langle f_1^h, \ldots, f_m^h\rangle$ has dimension 0, then the Hilbert series

$$\mathsf{HS}_{\mathbb{K}[x_1,\ldots,x_n]/\langle f_1^h,\ldots,f_m^h\rangle}(t)$$

is a polynomial and we can read from it the so-called *degree of regularity* of the system:

$$\mathrm{d}_{\mathrm{reg}}(f_1, \ldots, f_m) = 1 + \deg(\mathsf{HS}_{\mathbb{K}[x_1,\ldots,x_n]/\langle f_1^h,\ldots,f_m^h\rangle}).$$

The degree of regularity of algebraic systems is an important indicator of the complexity of Gröbner basis computations, since a Gröbner basis with respect to the *reverse graded lexicographical* ordering of $\langle f_1, \ldots, f_m\rangle$ can be computed within $O(m\binom{n+\mathrm{d}_{\mathrm{reg}}(f_1,\ldots,f_m)}{n}^\omega)$ arithmetic operations in $\mathbb{K}$. The degree of regularity actually bounds the highest degree reached during the computation of the Gröbner basis with the F_4 Algorithm. This value is a strong indicator of the complexity of the computation since the sizes of the largest matrices that have to be reduced during the F_4 algorithm are exponential in the degree of regularity.

Another central indicator of the complexity is the *degree* of the ideal. When a system has finitely-many solutions, this value corresponds to the number of solutions counted with multiplicities. For homogeneous 0-dimensional systems, it can be read off from the Hilbert series:

$$\mathsf{DEG}(\langle f_1^h,\ldots,f_m^h\rangle)=\mathsf{HS}_{\mathbb{K}[x_1,\ldots,x_n]/\langle f_1^h,\ldots,f_m^h\rangle}(1).$$

In this case, a *lexicographical* Gröbner basis of the ideal $\langle f_1,\ldots,f_m\rangle$ which gives an explicit algebraic description of the solutions can be computed within

$$O\left(m\binom{n+\mathrm{d_{reg}}(f_1,\ldots,f_m)}{n}^{\omega}+n\,\mathsf{DEG}(\langle f_1^h,\ldots,f_m^h\rangle)^3\right)$$

arithmetic operations by using the algorithms F_5 and FGLM. Consequently, our goal is to give explicit formulas for the Hilbert series of structured systems under genericity assumptions, which then provide complexity bounds. We report below the new formulas that we have obtained for ideals generated by polynomial families having the previously mentioned structure.

1. (**Bilinear systems.**) The first kind of multi-homogeneous systems that are encountered in practical applications are bilinear systems, and more particularly *affine bilinear systems* where each polynomial $f_i\in\mathbb{K}[x_1,\ldots,x_{n_x},y_1,\ldots,y_{n_y}]$ has the following shape:

   $$f_i=\sum_{\substack{1\le j\le n_x\\1\le k\le n_y}}a_{j,k}^{(i)}x_jy_k+\sum_{1\le j\le n_x}b_j^{(i)}x_j+\sum_{1\le k\le n_y}c_k^{(i)}y_k+d^{(i)},$$
   $$a_{j,k}^{(i)},b_j^{(i)},c_k^{(i)},d^{(i)}\in\mathbb{K}.$$

   Under genericity assumptions on the input system, we prove a new complexity bound on the complexity of computing Gröbner bases of affine bilinear systems with as many equations as unknowns:

   > **Result.** *Under genericity assumptions, the arithmetic complexity of computing a graded reverse lexicographical Gröbner basis of an affine bilinear system $f_1,\ldots,f_{n_x+n_y}\in\mathbb{K}[x_1,\ldots,x_{n_x},y_1,\ldots,y_{n_y}]$ with the F_4 Algorithm is bounded by*
   >
   > $$O\left(\min(n_x,n_y)(n_x+n_y)\binom{n_x+n_y+\min(n_x+2,n_y+2)}{\min(n_x+2,n_y+2)}^{\omega}\right).$$

   The main feature of this complexity bound is that the exponential part depends mainly on $\min(n_x,n_y)$. Consequently, this bound is polynomial in the number of variables when the size of one block is fixed. For instance, if $n_x=2$, the complexity is bounded by $O(n_y^{1+4\omega})$. This should be compared with the best previous bound available (which does not take into account the bilinear structure): the Macaulay bound for generic dense quadratic systems yields a complexity bound $O\left((n_x+n_y)\binom{2(n_x+n_y)+1}{n_x+n_y}^{\omega}\right)$. When $n_x=2$, this latter bound becomes $\widetilde{O}(4^{n_y\omega})$ which is exponential in n_y.

   The bound is proved by showing that during the execution of the Algorithms F_4 and F_5, the degrees of all polynomials occurring are bounded above by $\min(n_x,n_y)+2$ (see Section 6.5.5). This explains why in practice, bilinear systems with unbalanced sizes of blocks of variables are easier to solve than balanced ones. We also propose a dedicated variant of the F_5 Algorithm to compute Gröbner bases of multi-homogeneous ideals. Although there is no efficient low-level implementation of it so far, we expect important practical speed-ups (see Section 6.5.1).

2. (**Affine multi-homogeneous systems of bi-degree** $(D,1)$**.**) The complexity result for bilinear systems is generalized for affine systems of bi-degree $(D,1)$: we give an algorithm to compute

a rational parametrization of such systems. Its arithmetic complexity is bounded from above by

$$O\left(\binom{n_x+n_y}{n_x-1}\binom{D(n_x+n_y)+1}{n_x}^{\omega}+n_x\left(D^{n_x}\binom{n_x+n_y}{n_x}\right)^3\right).$$

Notice that this complexity is polynomial in the number of variables ($n = n_x + n_y$) when the size n_x of the first block is fixed. This bound comes from the fact that the biggest polynomials arising during the computations are polynomials of degree $(D-1)n_x+Dn_y+1$ in n_x variables. For instance, for $D = 3$, $n_x = 5$, $n_y = 2$, the highest degree reached is 17. To the best of our knowledge, the previous best bound is obtained by considering the system as generic dense of degree $D + 1$: in that case the biggest polynomials occurring during the computations are polynomials of degree $D(n_x + n_y) + 1$ in $n_x + n_y$ variables. For $D = 3$, $n_x = 5$, $n_y = 2$, this degree is 22.

3. (**Determinantal systems.**) Actually, the results on bilinear systems and systems of bidegree $(D, 1)$ have been achieved by investigating determinantal ideals. Indeed, solutions of such systems correspond to points where an associated Jacobian matrix is rank defective: its maximal minors simultaneously vanish.

   Let $r \in \mathbb{N}$ be an integer and M is a $p \times q$ matrix (with $q \leq p$) whose entries are polynomials of degree D in $\mathbb{K}[x_1, \dots, x_n]$.

   Result. *Under genericity assumptions on M, the arithmetic complexity of computing a lexicographical Gröbner basis of the ideal I generated by the minors of size $r + 1$ of M is bounded by*

$$O\left(\binom{p}{r+1}\binom{q}{r+1}\binom{\mathrm{d_{reg}}(I)+n}{n}^{\omega}+n\,(\mathsf{DEG}(I))^3\right),$$

   where $2 \leq \omega \leq 3$ is a feasible exponent for the matrix multiplication and

   - *if $n = (p - r)(q - r)$, then*

$$\begin{array}{rcl}\mathrm{d_{reg}}(I) & \leq & Dr(q-r)+(D-1)n+1,\\ \mathsf{DEG}(I) & \leq & D^{(p-r)(q-r)}\prod_{i=0}^{q-r-1}\frac{i!(p+i)!}{(q-1-i)!(p-r+i)!}.\end{array}$$

   - *if $n < (p - r)(q - r)$, then assuming that a conjecture is true (Conjecture 1.53, page 40),*

$$\begin{array}{rcl}\mathrm{d_{reg}}(I) & \leq & \deg(P(t))+1,\\ \mathsf{DEG}(I) & \leq & P(1)\end{array}$$

     where $P(t)$ is the polynomial obtained by truncating the series

$$(1-t^D)^{(p-r)(q-r)}\frac{\det A_r^{p,q}(t^D)}{t^{D\binom{r}{2}}(1-t)^n}$$

     at its first non-positive coefficient, and where $A_r^{p,q}(t)$ is the $r \times r$ matrix whose (i, j)-entry is $\sum_k \binom{p-i}{k}\binom{q-j}{k}t^k$.

These complexity results allow us to identify sub-families of generalized MinRank problems for which the complexity is polynomial in the size of the output. For instance, in the case of maximal minors (i.e. $r = q-1$), or when D (or p) is the only variable parameter, the complexity of the computation is polynomial in the degree of the ideal. Also, one of the main feature of the complexity bound is that, if $D = 1$, then the degree of regularity does not depend on the number of variables n. For given values of (p, q, r, D, n), we report in Table 1, the number of equations and the degree of the equations of the determinantal system, and then we give the degree and the degree of regularity of the ideal. This gives an idea of the size and the complexity of the systems that can be solved.

(p,q,r,D,n)	nb. eq.	deg. eq.	DEG	d_{reg}
(6,4,3,1,3)	15	4	20	4
(5,4,2,2,6)	40	6	3200	15
(4,4,2,3,4)	16	9	1620	21
(11,11,8,1,9)	3025	9	259545	25

Table 1: Sizes of determinantal systems; Degree and degree of regularity

4. (**Critical point systems.**) We investigate the problem of finding critical points of the projection $\pi_1 : (x_1, \ldots, x_n) \mapsto x_1$ restricted to the zero set V of a family of polynomials $f_1, \ldots, f_p \in \mathbb{K}[x_1, \ldots, x_n]$ of degree $D \in \mathbb{N}$. We show that, under genericity assumptions, the arithmetic complexity of computing a lexicographical Gröbner basis of the ideal I_{crit} vanishing on the critical points is uniformly polynomial in the number of critical points:

   Result. *For $D \geq 3$, $p \geq 2$ and $n \geq 2$, there exists a non-empty Zariski open subset $\mathscr{O} \subset \overline{\mathbb{K}}[X]_D^p$, such that, for $\mathbf{F} \in \mathscr{O} \cap \mathbb{K}[X]^p$, the arithmetic complexity of computing a lexicographical Gröbner basis of I_{crit} is bounded by*

   $$O\left(\mathsf{DEG}\,(I_{\mathsf{crit}})^{4.03\omega}\right).$$

   We also prove that if $D = 2$, the complexity is polynomial in n and exponential in p:

   Result. *If $D = 2$, then there exists a non-empty Zariski open subset $\mathscr{O} \subset \overline{\mathbb{K}}[X]_2^p$, such that for all $\mathbf{F} \in \mathscr{O} \cap \mathbb{K}[X]^p$, the arithmetic complexity of computing a lexicographical Gröbner basis of I_{crit} is bounded by*

   $$O\left(\left(p + \binom{n-1}{p}\right)\binom{n+2p}{2p}^{\omega} + n2^{3p}\binom{n-1}{p-1}^3\right).$$

   Moreover, if p is constant and $D = 2$, the arithmetic complexity is bounded by $O\left(n^{p(2\omega+1)}\right)$.

   We also generalize these complexity results to the *mixed* case where all polynomials $f_1, \ldots, f_p$ do not share the same degree: we show that the complexity of the computation is polynomial in the generic number of critical points when the degrees of the polynomials $f_1, \ldots, f_p$ are bounded above by a constant $D \in \mathbb{N}$.

5. (**Boolean systems.**) Under algebraic assumptions on the input system, we give an algorithm for solving quadratic boolean systems with n unknowns and n equations whose asymptotic complexity is bounded by $O(2^{0.841n})$ in a deterministic variant and by $O(2^{0.792n})$ in a probabilistic Las Vegas variant. More generally, for quadratic boolean systems of $\lceil \alpha n \rceil$ equations in n unknowns with $\alpha \geq 1$, we give estimates of the complexity:

(n_x, n_y)	Nb. useful red. (Buch./F_4)	Nb red. to 0 (Buch./F_4)	Nb red. to 0 (F_5)	Nb red. to 0 (**F_5 with new criterion**)
$(5, 6)$	1484	13063	495	**0**
$(6, 7)$	5866	64093	2002	**0**
$(4, 9)$	2869	31737	1794	**0**
$(3, 10)$	1212	13156	1300	**0**
$(3, 12)$	2123	27295	3018	**0**

Table 2: Experimental number of reductions to zero

Result. *Let $S = (f_1, \dots, f_m)$ be a system of quadratic polynomials in $\mathsf{GF}_2[x_1, \dots, x_n]$, with $m = \lceil \alpha n \rceil$ and $\alpha \geq 1$. Then, under precise algebraic assumptions, Algorithm BooleanSolve finds all its roots in GF_2^n with a number of arithmetic operations in GF_2 that is*

- *$O(2^{(1-0.159\alpha)n})$ with a deterministic variant;*
- *of expectation $O(2^{(1-0.208\alpha)n})$ with a Las Vegas probabilistic variant.*

The algorithm relies on a combination of efficient sparse linear algebra on the Macaulay matrix and exhaustive search. This complexity can be compared with the best *worst case complexity bound*: $4 \log_2(n) 2^n$ bit operations with a modified exhaustive search [BCC+10].

Structural results

1. (**Bilinear systems.**) If $(f_1, \dots, f_m) \in \mathbb{K}[x_0, \dots, x_{n_x}, y_0, \dots, y_{n_y}]^m$ (with $m \leq n_x + n_y$) is a generic bilinear family of polynomials, the Hilbert series can be extended to the Hilbert *bi-series*

$$\mathsf{mHS}_{\mathbb{K}[x_0,\dots,x_{n_x},y_0,\dots,y_{n_y}]/I}(t_1, t_2) = \sum_{d_1, d_2 \in \mathbb{N}} \dim_{\mathbb{K}}(\mathbb{K}[X, Y]_{d_1,d_2} / I_{d_1,d_2}) t_1^{d_1} t_2^{d_2},$$

where $\mathbb{K}[X, Y]_{d_1,d_2}$ (resp. I_{d_1,d_2}) denotes the vector-space of bi-homogeneous polynomials of bi-degree (d_1, d_2) in $\mathbb{K}[x_0, \dots, x_{n_x}, y_0, \dots, y_{n_y}]$ (resp. $\langle f_1, \dots, f_m \rangle$). We show that it is given by the formula:

$$\mathsf{mHS}_{\mathbb{K}[x_0,\dots,x_{n_x},y_0,\dots,y_{n_y}]/I}(t_1, t_2) = \frac{(1 - t_1 t_2)^m + N_m(t_1, t_2)}{(1 - t_1)^{n_x+1} (1 - t_2)^{n_y+1}},$$

$$\begin{aligned} N_m(t_1,t_2) = \; & \textstyle\sum_{\ell=1}^{m-(n_y+1)} (1-t_1t_2)^{m-(n_y+1)-\ell} t_1 t_2 (1-t_2)^{n_y+1} \big[1 - (1-t_1)^{\ell} \sum_{k=1}^{n_y+1} t_1^{n_y+1-k} \binom{\ell+n_y-k}{n_y+1-k}\big] + \\ & \textstyle\sum_{\ell=1}^{m-(n_x+1)} (1-t_1t_2)^{m-(n_x+1)-\ell} t_1 t_2 (1-t_1)^{n_x+1} \big[1 - (1-t_2)^{\ell} \sum_{k=1}^{n_x+1} t_2^{n_x+1-k} \binom{\ell+n_x-k}{n_x+1-k}\big]. \end{aligned}$$

This formula is obtained by giving a complete description of the syzygy module of the system $(f_1, \dots, f_m)$ under genericity assumptions and by investigating its combinatorial properties. This description of the syzygy module also leads to an extension of the F_5 criterion to avoid all reductions to 0 (which are useless computations) when the input of the F_5 algorithm is a generic bilinear system. Table 2 compares the number of reductions to 0 with the number of useful reductions for different Gröbner algorithms when the input is a random bilinear system of $n_x + n_y$ equations in $\mathbb{K}[x_0, \dots, x_{n_x}, y_0, \dots, y_{n_y}]$ (for these experiments, $\mathbb{K} = \mathsf{GF}_{65521}$ and the bilinear systems are picked uniformly at random in $\mathbb{K}[x_0, \dots, x_{n_x}, y_0, \dots, y_{n_y}]$).

2. (**Determinantal systems.**) Let $f_{1,1}, \ldots, f_{p,q} \in \mathbb{K}[x_1, \ldots, x_n]$ be homogeneous polynomials of degree $D \in \mathbb{N}$ and M be the $p \times q$ matrix whose (i,j)-th entry is $f_{i,j}$. We let I be the ideal generated by the minors of size $(r+1) \in \mathbb{N}$ of M. If $n \geq (p-r)(q-r)$ and under genericity assumptions, the ideal I has dimension $n-(p-r)(q-r)$ and we show that its Hilbert series is given by the formula:

$$\mathsf{HS}_{\mathbb{K}[x_1,\ldots,x_n]/I}(t) = \frac{\det\left(A_r^{p,q}(t^D)\right)(1-t^D)^{(p-r)(q-r)}}{t^{D\binom{r}{2}}(1-t)^n},$$

where $A_r^{p,q}(t)$ is the $r \times r$ matrix whose (i,j)-entry is $\sum_k \binom{p-i}{k}\binom{q-j}{k} t^k$. In the 0-dimensional case (i.e. when $n = (p-r)(q-r)$), the degree of regularity and the degree of the ideal I can be deduced:

$$\begin{aligned}
\mathrm{d_{reg}}(I) &= Dr(q-r) + (D-1)n + 1 \\
\mathsf{DEG}(I) &= D^{(p-r)(q-r)} \prod_{i=0}^{q-r-1} \frac{i!(p+i)!}{(q-1-i)!(p-r+i)!}.
\end{aligned}$$

These results are also generalized to the over-determined case (i.e. when $n < (p-r)(q-r)$) by assuming a variant of the Fröberg's conjecture.

In the case of maximal minors of a linear matrix (i.e. $r = q-1$, $D = 1$), we prove that under genericity assumptions the reduced grevlex Gröbner basis of I is a linear combination of the maximal minors of M. This is a variant of the result in [BZ93, SZ93] which states that the set of maximal minors of a matrix whose entries are algebraically independent variables is a *universal Gröbner basis* (i.e. a Gröbner basis with respect to every admissible monomial ordering).

3. (**Critical point systems.**) If $f_1, \ldots, f_p \in \mathbb{K}[x_1, \ldots, x_n]$ are polynomials of degree D, the ideal I_{crit} vanishing on the critical points of the projection π_1 restricted to the variety V associated to $f_1, \ldots, f_p$ is generated by the polynomials $f_1, \ldots, f_p$ and by the maximal minors of the matrix

$$\begin{bmatrix} \frac{\partial f_1}{\partial x_2} & \cdots & \frac{\partial f_1}{\partial x_n} \\ \vdots & \vdots & \vdots \\ \frac{\partial f_p}{\partial x_2} & \cdots & \frac{\partial f_p}{\partial x_n} \end{bmatrix}$$

We show that under genericity assumptions on the polynomials $f_1, \ldots, f_p$, the Hilbert series of I_{crit} is

$$\mathsf{HS}_{\mathbb{K}[x_1,\ldots,x_n]/I_{\mathsf{crit}}}(t) = \frac{\det(A_{p-1}^{p,n-1}(t^{D-1}))}{t^{(D-1)\binom{p-1}{2}}} \frac{(1-t^D)^p(1-t^{D-1})^{n-p}}{(1-t)^n}.$$

This formula is obtained by considering the properties of the determinantal part of the ideal I_{crit}. By giving a free resolution of this determinantal component, we also extend the result to the mixed case, i.e. when the polynomials $f_1, \ldots, f_p$ do not share the same degree. In that case, we let d_i denote the degree of f_i and we obtain the following formula for $\mathsf{HS}_{\mathbb{K}[x_1,\ldots,x_n]/I_{\mathsf{crit}}}(t)$:

$$\frac{\prod_{1\leq i\leq p}(1-t^{d_i})(1-t^{d_i-1})^{n-1}\left(1-\left[\sum\limits_{0\leq k\leq n-p-1}\left[(-1)^k \sum\limits_{i_1+\ldots+i_p=k}\binom{n-1}{p+k} t^{\sum\limits_{1\leq j\leq p}(i_j+1)(d_j-1)}\right]\right]\right)}{(1-t)^n\prod_{1\leq i\leq p}(1-t^{d_i-1})^{n-1}}.$$

This is actually a polynomial, and its degree can be computed

$$\begin{array}{rcl} \mathrm{d}_{\mathrm{reg}}(I_{\mathsf{crit}}) & = & \deg(\mathsf{HS}_{\mathbb{K}[x_1,\dots,x_n]/I_{\mathsf{crit}}}) + 1 \\ & = & (n-p-1)\max(d_i) - 2n + 2 + 2\sum_{1\leq i\leq p} d_i. \end{array}$$

4. (**Boolean systems.**) Let $f_1,\dots,f_m \in \mathsf{GF}_2[x_1,\dots,x_n]$ be a quadratic boolean system. Under algebraic assumptions that are satisfied for a large class of systems, we give an explicit formula for the Hilbert series of the ideal $I \subset \mathsf{GF}_2[x_1,\dots,x_n,h]$ generated by the homogeneous polynomials $h^2 f_1(x_1/h,\dots,x_n/h),\dots,h^2 f_m(x_1/h,\dots,x_n/h), x_1^2 - x_1 h,\dots,x_n^2 - x_n h$: it is the polynomial obtained by truncating the power series expansion of

$$\frac{(1+t)^n}{(1-t)(1+t^2)^m}$$

at its first nonpositive coefficient. A consequence of this formula is an asymptotic analysis of the degree of regularity of quadratic boolean system, which leads to complexity estimates.

Applications to Cryptology

The complexity estimates for solving polynomial systems can be used to evaluate the security of several multivariate cryptosystems.

MinRank authentication scheme. In [Cou01], N. Courtois proposes a zero-knowledge authentication scheme, whose security is based on the difficulty of the so-called *MinRank* problem. A modeling of the problem yields a determinantal system to solve. Using the complexity results for solving determinantal systems, we identify families of parameters for which this cryptosystem can be broken in polynomial time. From a more practical viewpoint, we give precise estimates of the computing time needed to solve a challenge proposed in [Cou01] which was considered untractable so far (see Section 8.2).

QUAD. The QUAD streamcipher [BGP09, BGP06] is a cryptosystem whose the security is proven to be related to the difficulty of solving quadratic systems of boolean equations. Therefore, a straightforward consequence of the complexity results for boolean systems is a reevaluation of the parameters of the QUAD cryptosystem in order to keep the same level of security (see Section 8.3).

The Algebraic Surface Cryptosystem. The Algebraic Surface Cryptosystem (ASC) is an asymmetric encryption scheme whose security relies on the so-called *Section Finding Problem*, which is rather unusual in multivariate cryptology. The main advantage of this construction is that it provides very short keys (linear in the security level). We show that by using algebraic techniques from computer algebra (Gröbner bases computations, decompositions of ideals, ...), the encryption process can be inverted in polynomial time with respect to all security parameters. We give an algorithm for this task, and an implementation in the computer algebra system `Magma` allowed us to recover plaintext messages in less than 0.05 seconds on a standard computer for recommended security parameters (see Section 8.1). This is actually faster than the legal decryption algorithm.

Structure	Complexity	Algorithms	Structural results	Applications
Bilinear	6.5	6.2.2, 6.5	6.3, 6.4	
Bihom. of bideg. $(D, 1)$	6.6	6.6		
Determinantal	4.6, 4.7		4.4, 4.5	8.2
Critical points	5.5, 5.7.3		5.3, 5.7.2	
Boolean	7.3	7.2	7.3.2	7.5

Table 3: Contributions and references of sections

Conclusion

In this thesis, we provide under genericity assumptions new complexity bounds for solving

1. affine bilinear systems;
2. affine bihomogeneous systems of bi-degree $(D, 1)$;
3. determinantal systems in the unmixed case: all polynomials in the matrix share the same degree;
4. mixed and unmixed critical point systems;
5. boolean quadratic systems.

We also give new algorithms for

1. computing Gröbner bases of bilinear systems without reductions to zero;
2. computing rational parametrizations of affine bi-homogeneous systems of bi-degree $(D, 1)$;
3. computing Gröbner bases of multi-homogeneous systems;
4. solving quadratic boolean systems.

We provide theoretical results:

1. an explicit form of the Hilbert bi-series of ideals generated by bilinear forms;
2. a description of the syzygy module of bilinear systems;
3. a formula for the Hilbert series of unmixed determinantal systems, mixed and unmixed critical point systems, and for homogenized boolean systems;
4. we identify families of determinantal systems and of critical point systems, for which the complexity of computing a lexicographical Gröbner basis is polynomial in the size of the output.

Finally, we give concrete applications in Cryptology:

1. precise estimates of the computing power needed to solve cryptographic challenges proposed in [Cou01] which are related to the MinRank;
2. an efficient and practical message-recovery attack on the Algebraic Surface Cryptosystem;
3. a reevaluation of the security parameters of the QUAD cryptosystem.

In Table 3, we report the sections of this thesis where the results are presented.

Further impact of these results

The results in this thesis have had impacts in other publications:

- In [BFP11, BFP12a], the authors obtain complexity estimates for algebraic attacks on the cryptosystem HFE (and variants). During this complexity analysis, the bounds on the degree of regularity of determinantal systems (Chapter 3) are used to get complexity estimates of Gröbner bases computations.
- In [FOPT10], the bound on the maximal degree reached during the computation of Gröbner bases of affine bilinear systems (Section 6.5.5) allows explaining the efficiency of the attack proposed on compact variants of the McEliece cryptosystem.

Perspectives.

Some points still need to be investigated. In this section, we report future possible developments of the results presented in this thesis and related open problems.

(**Multi-homogeneous systems.**) For bilinear systems, we give an explicit description of the syzygy module and we obtain from this a criterion to remove reductions to 0 during the F_5 algorithm. The next step is to generalize these results to multi-homogeneous systems. Similarly, obtaining an explicit formula for the generic multi-Hilbert series of multi-homogeneous system is still an open question. Before investigating general multi-homogeneous systems, a first step is to understand bihomogeneous systems.

(**Affine multi-homogeneous systems.**) In the case of affine bilinear systems, we observe that *degree falls* play an important role during the computation of Gröbner bases. The analysis is more difficult in that case than it is for homogeneous systems. The next step here is to develop a systematic approach to investigate affine systems which do not behave similarly to their homogeneous counterparts. Indeed, having sharp bounds on the maximal degree reached during the computation of Gröbner bases of affine multi-homogeneous systems is an open problem which is crucial to obtain practical bounds on the complexity of such computations.

(**Determinantal systems.**) We give in this thesis an analysis of the complexity and of the combinatorial structure of unmixed determinantal systems (i.e. all polynomials in the matrix share the same degree). The next step would be to understand how this structure can be used to design Gröbner basis algorithms dedicated to this family of systems. Also, investigating how the results in this thesis could be generalized to the mixed case is a natural follow-up of this work.

(**Critical point systems.**) Following the results in Section 5.7, the Eagon-Northcott complex yields a free resolution of the determinantal part of the ideal vanishing on the critical points. This could lead to an analysis of the syzygy module and yield a criterion to remove reductions to zero in the F_5 algorithm when the input is a critical point system. We plan to investigate this question in future works.

(**Implementation.**) In this thesis, several algorithms are proposed (solving boolean systems, computing rational parametrization of bihomogeneous systems of bidegree $(D, 1)$, computing Gröbner bases of multi-homogeneous systems). The next step is to implement these algorithms in a low-level language (C, C++,...) in order to solve larger structured polynomial systems.

(**Rational coefficients.**) In this thesis, we focus on the *arithmetic complexity*. This is a representative measure of the execution time when the base field is a finite field (this is the case in Cryptology). For applications in Geometry and Optimization, the ground field is often the field of rational numbers.

In that case, the arithmetic complexity is a first step but it would also be interesting to have estimates of the size of the coefficients in rational parametrizations in order to have a better understanding of the bit complexity. Such bounds are related to the *height* of the corresponding variety [KPS01], which can be bounded by using *Chow forms* in the sense of Philippon [Phi86].

(**Generalization of determinantal ideals.**) The entries of the Jacobian matrices of general multi-homogeneous systems are multi-homogeneous polynomials. Therefore, we plan to investigate the structure of determinantal systems when the entries of the matrix are themselves structured (for instance multi-homogeneous or boolean,...). In particular, this could lead to a better understanding of bihomogeneous ideals. Indeed, the structure of ideals generated by bihomogeneous polynomials is closely related to the combinatorial properties of the determinantal ideal generated by minors of the Jacobian matrices with respect to each block of variables. The entries of these matrices are also bihomogeneous. Therefore the next step to generalize the results on bilinear systems is to investigate the properties of these *bihomogeneous determinantal* ideals.

(**Related problems in Symbolic Computation.**) Determinantal ideals are basic objects of enumerative geometry and Schubert calculus. Consequently, we plan to investigate in future works how the results in this thesis can be extended to Schubert problems.

Organization of the thesis

In the first part of the thesis, we recall known facts about Gröbner bases. Chapter 1 is devoted to basic notions of Gröbner basis theory and commutative algebra that are used throughout the thesis. Then in Chapter 2, we give examples of applications in Engineering Sciences where structured algebraic systems naturally appear. Finally in Chapter 3, we recall known facts about determinantal and bihomogeneous systems.

The second part of the thesis is devoted to contributions. Most parts of Chapters and Sections are published or submitted articles. Therefore, these chapters are mostly self-contained and a few statements appear in different chapters. We list the references of these papers below (author names are in alphabetical order):

- Chapter 3 and Section 6.6: **On the Complexity of the Generalized MinRank Problem**. Jean-Charles Faugère, Mohab Safey El Din, Pierre-Jean Spaenlehauer. Submitted, `arXiv:1112.4411 [cs.SC]`.

- Chapter 5: **Critical Points and Gröbner Bases: the Unmixed Case**. Jean-Charles Faugère, Mohab Safey El Din, Pierre-Jean Spaenlehauer. Proceedings of the 37th International Symposium on Symbolic and Algebraic Computation (ISSAC 2012).

- Chapter 6: **Gröbner Bases of Bihomogeneous Ideals generated by Polynomials of Bidegree (1,1): Algorithms and Complexity**. Jean-Charles Faugère, Mohab Safey El Din, Pierre-Jean Spaenlehauer. Journal of Symbolic Computation, 46(4):406-437, 2011.

- Chapter 7: **On the Complexity of Solving Quadratic Boolean Systems**. Magali Bardet, Jean-Charles Faugère, Bruno Salvy, Pierre-Jean Spaenlehauer. Accepted for publication in Journal of Complexity, `arXiv:1112.6263 [cs.SC]`.

- Section 8.1: **Algebraic Cryptanalysis of the PKC'2009 Algebraic Surface Cryptosystem**. Jean-Charles Faugère, Pierre-Jean Spaenlehauer. Proceedings of the 13th International Conference on Practice and Theory in Public Key Cryptography (PKC 2010).

- Section 8.2: **Computing Loci of Rank Defects of Linear Matrices using Gröbner Bases and Applications to Cryptology**. Jean-Charles Faugère, Mohab Safey El Din, Pierre-Jean Spaenlehauer. Proceedings of the 35th International Symposium on Symbolic and Algebraic Computation (ISSAC 2010).

Acknowledgements

I am grateful to L. Perret and I. Emiris for their comments on the papers on bilinear systems and on the problem MinRank. I also wish to thank D. Bernstein, C. Diem, E. Kaltofen for valuable comments and pointers to important references on the topic of boolean systems.

Part I

Preliminaries

Chapter 1

Preliminaries on Gröbner Bases

In this chapter, we recall definitions, algorithms and properties of Gröbner bases algorithms that are used throughout this thesis. We refer the reader to [CLO97] for a more detailed exposition of Gröbner bases theory.

1.1 Polynomial Rings and Ideals

1.1.1 Definitions

Notations 1.1. *In the whole document, $\mathbb{K}$ is either a finite field or a field of characteristic 0 (and hence $\mathbb{K}$ is a perfect field). Its algebraic closure is denoted by $\overline{\mathbb{K}}$. The finite field of cardinality q is denoted by GF_q. The notation X stands for the set of variables $\{x_1, \ldots, x_n\}$. If R is a ring and $\mathbf{F} = \{r_1, \ldots, r_m\} \subset R$ is a family of elements of R, we let $\langle \mathbf{F} \rangle \subset R$ denote the ideal generated by $\mathbf{F}$. If I and J are ideals of $\mathbb{K}[X]$, then the following subsets of $\mathbb{K}[X]$ are also ideals of $\mathbb{K}[X]$:*

$$\begin{array}{lll} \textit{sum} & I + J & = \{f + g \mid f \in I, g \in J\}; \\ \textit{product} & IJ & = \{fg \mid f \in I, g \in J\}; \\ \textit{intersection} & I \cap J; & \\ \textit{radical} & \sqrt{I} & = \{f \in \mathbb{K}[X] \mid \exists k \in \mathbb{N} \textit{ s.t. } f^k \in I\}; \\ \textit{colon ideal} & I : J & = \{f \in \mathbb{K}[X] \mid fJ \subset I\}; \\ \textit{saturation} & I : J^\infty & = \{f \in \mathbb{K}[X] \mid \exists k \in \mathbb{N} \textit{ s.t. } fJ^k \subset I\}. \end{array}$$

In this thesis, we mainly focus on systems of polynomial equations that have a finite number of solutions in $\overline{\mathbb{K}}^n$: the polynomials generate a 0-dimensional ideal of $\mathbb{K}[X]$. Indeed, even when we study varieties of positive dimension, we will investigate subsets of points that are defined by 0-dimensional systems (for instance by computing the critical points of a projection restricted to the variety in Chapter 5).

Definition 1.2. *We call* dimension *of an ideal $I \subset \mathbb{K}[X]$ the Krull dimension of the quotient ring $\overline{\mathbb{K}}[X]/I$, i.e. the supremum of the number of strict inclusions in a chain of prime ideals of $\overline{\mathbb{K}}[X]/I$.*

This is a theoretical definition of the dimension. We give below in Proposition 1.43 a more algorithmic equivalent definition with the Hilbert series.

Example 1.3.
- *The dimension of the ideal $\langle (x_1 - 1)(x_1 - 2), x_2 + 3\rangle \subset \mathbb{K}[x_1, x_2]$ is 0 since the only prime ideals of $\overline{\mathbb{K}}[x_1, x_2]/\langle (x_1 - 1)(x_1 - 2), x_2 + 3\rangle$ are $\langle x_1 - 1\rangle$ and $\langle x_1 - 2\rangle$, and there are no inclusion relation between these two ideals (notice that in $\overline{\mathbb{K}}[x_1, x_2]/\langle (x_1 - 1)(x_1 - 2), x_2 + 3\rangle$, the ideal $\langle 0\rangle$ is not prime).*

- *The dimension of the ideal* $\langle x_1 + 1\rangle \subset \mathbb{K}[x_1, x_2]$ *is* 1 *since a longest chain of prime ideals of* $\overline{\mathbb{K}}[x_1, x_2]/\langle x_1 + 1\rangle$ *is* $\langle 0\rangle \subset \langle x_2\rangle$ *which has* 1 *inclusion.*

The *degree* is an important indicator of the "complexity" of a 0-dimensional ideal. It counts the number of solutions (with multiplicities) of the system of polynomial equations.

Definition – Proposition 1.4. *Let* $I \subset \mathbb{K}[X]$ *be a* 0*-dimensional ideal. Then* $\mathbb{K}[X]/I$ *is a* $\mathbb{K}$*-vector space of finite dimension. The dimension* $\dim_{\mathbb{K}}(\mathbb{K}[X]/I)$ *is called* degree *of* I *and is denoted by* $\mathsf{DEG}(I)$.

Proof. The proof that $\mathbb{K}[X]/I$ is a $\mathbb{K}$-vector space of finite dimension when I is 0-dimensional is postponed at the end of Section 1.1.3. □

The degree of an ideal can also be defined for ideals of positive dimension, but we will not need this notion in this thesis. As the dimension, the degree can be read off from the Hilbert series (Proposition 1.43).

The geometrical objects corresponding to ideals of $\mathbb{K}[X]$ are *affine varieties* of $\overline{\mathbb{K}}^n$ (also called *algebraic sets*). They are the sets of points where all polynomials in an ideal simultaneously vanish. Actually, if a family of polynomials simultaneously vanish on a subset $V \subset \overline{\mathbb{K}}^n$, then any algebraic combination of these polynomials also vanish on V. Therefore, the entire ideal $\langle F\rangle$ vanish on V.

Proposition 1.5. *Let* $I \subset \mathbb{K}[X]$ *be an ideal generated by a family* $\mathbf{F} = (f_1, \dots, f_m) \in \mathbb{K}[X]^m$*. Let* $Z(\mathbf{F})$ *(resp.* $Z(I)$*) denote the set* $\{\mathbf{x} \in \overline{\mathbb{K}}^\ell \mid f_1(\mathbf{x}) = \cdots = f_m(\mathbf{x}) = 0\}$ *(resp.* $\{\mathbf{x} \in \overline{\mathbb{K}}^\ell \mid \forall f \in I, f(\mathbf{x}) = 0\}$*). Then* $Z(\mathbf{F}) = Z(I)$*.*

Proof. Clearly $\mathbf{F} \subset I$, and hence $Z(I) \subset Z(\mathbf{F})$. Conversely, let $\mathbf{x} \in \overline{\mathbb{K}}^\ell$ be an element of $Z(\mathbf{F})$. For any polynomial $h \in I$, there exist $h_1, \dots, h_m \in \mathbb{K}[X]$ such that $h = \sum_{i=1}^m h_i f_i$. Therefore $h(\mathbf{x}) = \sum_{i=1}^m h_i(\mathbf{x}) f_i(\mathbf{x}) = 0$ and consequently $Z(\mathbf{F}) \subset Z(I)$. □

Notations 1.6. *If* S *is a subset of* $\overline{\mathbb{K}}^n$*, we let* $I(S) \subset \overline{\mathbb{K}}[X]$ *denote the ideal of the polynomials vanishing on all points of* S*. Notice that* $I(S)$ *is radical by Hilbert's Nullstellensatz [CLO97, Ch. 4, §1, Thm.2].*

An important property of algebraic sets of $\overline{\mathbb{K}}^n$ is that they define a topology on $\overline{\mathbb{K}}^n$:

Definition – Proposition 1.7 (Zariski topology)**.** *A subset* V *of* $\overline{\mathbb{K}}^n$ *is called* algebraic set *if there exists an ideal* $I \subset \overline{\mathbb{K}}[X]$ *such that* $V = Z(I)$*. Algebraic sets have the following properties:*

- *any intersection of algebraic sets is an algebraic set;*
- *any finite union of algebraic sets is an algebraic set;*
- $\overline{\mathbb{K}}^\ell$ *is an algebraic set;*
- $\emptyset$ *is an algebraic set.*

Therefore the algebraic sets are the closed sets of a topology, called the Zariski topology.

Proof.
- Let $\{V_\ell\}_{\ell \in L}$ be a family of algebraic sets. Then there exist families of polynomials $\{\mathbf{F}_\ell\}_{\ell \in L}$ such that $V_\ell = Z(\mathbf{F}_\ell)$. Therefore $\cap_{\ell \in L} V_\ell = Z(\langle \mathbf{F}_\ell\rangle_{\ell \in L})$, hence $\cap_{\ell \in L} V_\ell$ is an algebraic set;

- let $V_1 = Z(f_1, \ldots, f_s)$ and $V_2 = Z(h_1, \ldots, h_t)$ be two algebraic sets. Then $V_1 \cup V_2 = Z(\{\prod_{\substack{1 \leq i \leq s \\ 1 \leq j \leq t}} f_i h_j\})$ is also an algebraic set. By induction, any finite union of algebraic sets is an algebraic set;
- $\overline{\mathbb{K}}^{\ell} = Z(\langle 0 \rangle)$;
- $\emptyset = Z(\mathbb{K}[X])$.

□

The Zariski topology has several interesting properties. First, notice that any nonempty open subset of $\overline{\mathbb{K}}^n$ is dense. Also, finite intersections of nonempty open subsets are nonempty.

In particular, this topology will be useful for defining an algebraic notion of genericity for structured systems: a property of a family of systems $\mathscr{F} \subset \overline{\mathbb{K}}[X]$ which is a $\overline{\mathbb{K}}$-vector space of finite dimension is said to be *generic* if this property is satisfied on a nonempty Zariski open subset of $\mathscr{F}$ (which is thus dense in $\mathscr{F}$).

1.1.2 Modules, algebras and free resolutions

In this section, we recall definitions of tools of commutative algebra which will be useful in Chapter 5.7. Modules are among the main objects of study in commutative algebra. They are to commutative rings what vector spaces are to fields:

Definition 1.8 (Module)**.** *Let R be a commutative ring. A R-module is an abelian group $(M, +)$ and an operation $R \times M \to M$ such that*

- *(distributivity)* $\forall r, s \in R, \forall m, n \in M, r(m+n) = rm + rn$ *and* $(r+s)m = rm + sm$*;*
- *(associativity)* $\forall r, s \in R, \forall m \in M, (rs)m = r(sm)$*;*
- $\forall m \in M, 1_R m = m$.

Definition 1.9 (Free module)**.** *The* free module R^r *of rank r is the module of r-tuples of elements in R with component-wise addition.*

Two basic operations on modules are the direct sum and the tensor product. The tensor product $M \otimes_R N$ of two modules M and N can be seen as the smallest R-module such that we can express all R-bilinear maps from $M \times N$ to another module.

Definition 1.10 (Tensor product)**.** *Let M and N be two R-modules. The tensor product $M \otimes_R N$ (noted $M \otimes N$ when the ring is obvious) is the R-module with generators $\{m \otimes n \mid m \in M, n \in N\}$ and relations*

$$\forall r_1, r_2, s_1, s_2 \in R, \forall m_1, m_2 \in M, \forall n_1, n_2 \in N,$$
$$(r_1 m_1 + r_2 m_2) \otimes (s_1 n_1 + s_2 n_2) = r_1 s_1 m_1 \otimes n_1 + r_1 s_2 m_1 \otimes n_2 + r_2 s_1 m_2 \otimes n_1 + r_2 s_2 m_2 \otimes n_2.$$

The so-called *tensor algebra* is built by tensoring successively a module with itself.

Definition 1.11 (Tensor algebra)**.** *Let M be a R-module. The tensor algebra of M is defined as the direct sum*

$$T(M) = R \oplus M \oplus (M \otimes M) \oplus \ldots.$$

The product of two elements $x_1 \otimes \cdots \otimes x_m$ and $y_1 \otimes \cdots \otimes y_n$ is $x_1 \otimes \cdots \otimes x_m \otimes y_1 \otimes \cdots \otimes y_n$.

Finally, we need two more definitions, the symmetric and the exterior algebras of M, which are obtained by imposing commutativity (resp. skew-commutativity).

Definition 1.12 (Symmetric algebra)**.** *The symmetric algebra of the R-module M is the quotient of the algebra $T(M)$ by the ideal generated by the relations $x \otimes y - y \otimes x$ for all $x, y \in M$. It is denoted by $\mathsf{Sym}(M)$.*

Definition 1.13 (Exterior algebra)**.** *The exterior algebra of the R-module M is the quotient of the algebra $T(M)$ by the ideal generated by the relations $x \otimes x$ for all $x \in M$. It is denoted by $\wedge M$.*

Notice that the exterior algebra is skew-commutative since in $\wedge M$, $x \otimes y + y \otimes x = x \otimes x + y \otimes y + x \otimes y + y \otimes x = (x + y) \otimes (x + y) = 0$.

Resolutions are mathematical objects which yield information on the structure of polynomial ideals (and more generally commutative rings and modules). The following result is known as the *Hilbert syzygy theorem*. See [Eis95, Corollary 15.11] for a constructive proof.

Definition – Proposition 1.14 (Hilbert Syzygy Theorem)**.** *Let $I \subset \mathbb{K}[X]$ be a polynomial ideal. Then there exists a finite exact sequence of free $\mathbb{K}[X]$-modules*

$$\mathscr{F} : 0 \to F_r \xrightarrow{\varphi_n} \dots \xrightarrow{\varphi_2} F_1 \xrightarrow{\varphi_1} F_0$$

such that $\mathbb{K}[X]/I \cong F_0/\mathsf{Im}(\varphi_1)$ and $r \leq n$. Such a sequence is called a free resolution *of I.*

Free resolutions yield a good view of the structure of graded ideals. Many useful information can be read off from such objects: dimension, Hilbert series, Betti numbers, etc... We will use free resolutions to obtain information about the structure of mixed critical point systems in Chapter 5.7.

1.1.3 Primary decomposition and associated primes

Another useful tool for describing ideals (and varieties) is the decomposition into irreducible components. Indeed, from a geometrical viewpoint, affine varieties can be uniquely decomposed into irreducible varieties. An irreducible variety $V \subset \overline{\mathbb{K}}^n$ is an algebraic set verifying the following property: if $V_1, V_2 \subset \overline{\mathbb{K}}^n$ are algebraic sets such that $V = V_1 \cup V_2$, then $V_1 = V$ or $V_2 = V$.

Here, for simplicity, we will only consider decompositions over algebraically closed fields. But the definitions and properties can be extended for any field.

Theorem 1.15 (Irreducible decomposition of varieties)**.** *[CLO97, Ch. 4, §6, Thm.4] Let $V \subset \overline{\mathbb{K}}^n$ be an affine variety. Then there exists a unique finite set of algebraic sets $\{V_1, \dots, V_\ell\}$ such that $V = V_1 \cup \dots \cup V_\ell$ and for all $i, j \in \{1, \dots, \ell\}$, $V_i \not\subset V_j$.*

A proper ideal I is called *primary* if $fg \in I$ implies that either $f \in I$ or there exists $n \in \mathbb{N}$ such that $g^n \in I$. If I is primary, then its radical $\sqrt{I}$ is prime.

Similarly to Theorem 1.15, ideals can be decomposed into irreducible components. However, this decomposition is not necessarily unique.

Theorem 1.16 (Irreducible decomposition of ideals)**.** *[Eis95, Thm. 3.10] Let $I \subset \overline{\mathbb{K}}[X]$ be an ideal. Then there exists a* minimal primary decomposition *of I, i.e. a finite set of primary ideals $\{I_1, \dots, I_\ell\}$ such that $I = I_1 \cap \dots \cap I_\ell$ and for all i, j, $I_i \not\subset I_j$. This decomposition is not necessarily unique, but all minimal primary decompositions of I share the same cardinality.*

Although minimal primary decompositions are not uniquely defined, the radicals of the primary ideals are the same for any decomposition:

Definition – Proposition 1.17 (Associated primes)**.** *[Eis95, Ch. 3] Let $I \subset \overline{\mathbb{K}}[X]$ be an ideal. We let $\mathrm{Ass}(I)$ denote the set of prime ideals $P \supset I$ such that there exists $f \in \overline{\mathbb{K}}[X] \setminus I$ with $(I : f) = P$. The family $\mathrm{Ass}(I)$ satisfies the following properties:*

- $\mathrm{Ass}(I)$ *is finite;*
- *If $I_1 \cap \cdots \cap I_\ell$ is a minimal primary decomposition of I, then $\mathrm{Ass}(I) = \{\sqrt{I_1}, \ldots, \sqrt{I_\ell}\}$.*

The primes in $\mathrm{Ass}(I)$ are called primes associated *to I. Let $P_1 \in \mathrm{Ass}(I)$ be an associated prime of I. If there exists $P_2 \in \mathrm{Ass}(I)$ such that $P_2 \subset P_1$, then we say that P_1 is an* embedded prime *of I, else P is called an* isolated prime*. Moreover, the radical of I is the intersection of the isolated primes associated to I.*

These notions will be useful in the study of multi-homogeneous ideals (see Chapter 6). Decompositions of ideals will also be a crucial part of the attack on the cryptosystem ASC presented in Section 8.1.

We can now prove that if I is a 0-dimensional ideal, then $\mathbb{K}[X]/I$ is a vector space of finite dimension over $\mathbb{K}$.

Proof of Definition-Proposition 1.4. Let I be a zero dimensional ideal. Since $I \subset \sqrt{I}$, $\sqrt{I}$ is also 0-dimensional as an ideal of $\overline{\mathbb{K}}[X]$ and is included in all isolated primes (since $\sqrt{I}$ is equal to the intersection of isolated primes). By Krull's Theorem and by the definition of dimension, all associated primes are maximal ideals of $\overline{\mathbb{K}}[X]$. Any maximal ideal of $\overline{\mathbb{K}}[X]$ has the form $\langle x_1 - \alpha_1, \ldots, x_n - \alpha_n\rangle$, where $\alpha_i \in \overline{\mathbb{K}}$. Consequently, there exist $\alpha_1^{(1)}, \ldots, \alpha_n^{(\ell)} \in \overline{\mathbb{K}}$ such that

$$\sqrt{I} = \bigcap_{1\leq i\leq \ell} \langle x_1 - \alpha_1^{(i)}, \ldots, x_n - \alpha_n^{(i)}\rangle.$$

Next, notice that the elements $\alpha_1^{(1)}, \ldots, \alpha_1^{(\ell)}$ are algebraic over $\mathbb{K}$, therefore there exists a univariate polynomial $P_1 \in \mathbb{K}[x]$ which vanishes on $\alpha_1^{(1)}, \ldots, \alpha_1^{(\ell)}$. By the Hilbert's Nullstellensatz, $P_1(x_1) \in \sqrt{I}$, hence there exists a power Q_1 of P_1 such that $Q_1(x_1) \in I$. Similarly, there exist univariate polynomials $Q_2(x_2), \ldots, Q_n(x_n) \in I$. Therefore $\mathbb{K}[X]/I \subset \mathbb{K}[X]/\langle Q_1(x_1), \ldots, Q_n(x_n)\rangle$ which is a $\mathbb{K}$-vector space of finite dimension. □

1.2 Monomial orderings and Gröbner bases

In this section, we recall the basic definitions and some properties of monomial orderings and Gröbner bases. We also show how Gröbner bases preserve the graded structure in the context of homogeneous, quasi-homogeneous and multi-homogeneous ideals.

1.2.1 Definitions

A Gröbner basis of an ideal I is a set of generators of this ideal which has good properties. It generalizes Row Echelon bases for linear systems. For univariate systems, it corresponds to the greatest common divisor of the polynomials.

Gröbner bases are defined with respect to a total well-ordering on the monomials of $\mathbb{K}[X]$:

Definition 1.18. *[CLO97, Ch.2, §2, Def.1] A* monomial ordering *$\prec$ on $\mathbb{K}[X]$ is a relation on $\mathbb{N}^n$ (or on the monomials of $\mathbb{K}[X]$ by identification) satisfying:*

- *$\prec$ is a total ordering on $\mathbb{N}^n$;*

- *if $\alpha \prec \beta$ and $\gamma \in \mathbb{N}^n$, then $\alpha + \gamma \prec \beta + \gamma$;*
- *$\prec$ is a well-ordering: every nonempty subset of $\mathbb{N}^n$ has a smallest element.*

In this thesis, we focus mainly on the so-called *grevlex* (graded reverse lexicographical) and *lex* (lexicographical) orderings. The grevlex ordering is particularly well-suited for Gröbner bases computations, while the lex ordering yields a more explicit description of the solutions of a polynomial system.

The grevlex ordering is a graded ordering, i.e. monomials are first sorted by degree. This ordering also has other structural properties. For instance, in a homogeneous polynomial the first monomial is a monomial involving only the variable x_1, then come the monomials involving the variable x_1 and x_2, then the ones where the variables x_1, x_2 and x_3 appear, etc....

Definition 1.19 (Graded Reverse Lexicographical Ordering)**.** *[CLO97, Ch. 2, §2, Def. 6] Let $\alpha, \beta \in \mathbb{N}^n$; $\alpha \prec_{\mathsf{grevlex}} \beta$ if either:*

$$\textstyle\sum_{i=1}^n \alpha_i < \sum_{i=1}^n \beta_i$$

or

$$\textstyle\sum_{i=1}^n \alpha_i = \sum_{i=1}^n \beta_i \text{ and the rightmost entry of } \alpha - \beta \text{ is positive.}$$

Example 1.20.
- $x_1^2 \prec_{\mathsf{grevlex}} x_3^3$;
- $x_1^2 x_3 \prec_{\mathsf{grevlex}} x_2^3$.

For solving systems, the lex ordering is well-suited since it is a typical example of *elimination* ordering: if G is a *lex* Gröbner basis of an ideal $I \subset \mathbb{K}[X]$, then $G \cap \mathbb{K}[x_k, \dots, x_n]$ is a *lex* Gröbner basis of $I \cap \mathbb{K}[x_k, \dots, x_n]$ for any $i \in \{1, \dots, n\}$.

Definition 1.21 (Lexicographical Orderings)**.** *[CLO97, Ch. 2, §2, Def. 3] Let $\alpha, \beta \in \mathbb{N}^n$; $\alpha \prec_{\mathsf{lex}} \beta$ if the leftmost entry of $\alpha - \beta$ is negative.*

Example 1.22.
- $x_5^9 \prec_{\mathsf{lex}} x_1$;
- $x_1^2 x_3^2 \prec_{\mathsf{lex}} x_1^2 x_2$.

Definition 1.23 (Leading Monomial)**.** *[CLO97, Ch. 2, §2, Def. 7] Let $\prec$ be a monomial ordering, and $f = \sum_{\alpha \in \mathbb{N}^n} a_\alpha x^\alpha \in \mathbb{K}[X]$ (resp. $I \in \mathbb{K}[X]$) be a polynomial (resp. an ideal). Then its* leading monomial *(resp. its* leading monomial ideal*) with respect to $\prec$ is $\mathsf{LM}_\prec(f) = x^{\max_\prec\{\alpha \in \mathbb{N}^n : a_\alpha \neq 0\}}$ (resp. $\mathsf{LM}_\prec(I) = \langle \{\mathsf{LM}_\prec(f) : f \in I\} \rangle$).*

We can now give the definition of a Gröbner basis:

Definition 1.24 (Gröbner Basis)**.** *[CLO97, Ch. 2, §5, Def. 5] Let $\prec$ be a monomial ordering, and $I \subset \mathbb{K}[X]$ be an ideal. A Gröbner basis of I with respect to $\prec$ is a finite subset $G = \{g_1, \dots, g_\ell\} \subset I$ such that*

$$\mathsf{LM}_\prec(I) = \langle \mathsf{LM}_\prec(g_1), \dots, \mathsf{LM}_\prec(g_\ell) \rangle .$$

The following proposition shows a direct consequence of Definition 1.24: a Gröbner basis of an ideal generates it.

Proposition 1.25. *[CLO97, Ch. 2, §5, Cor. 6] Let $G = \{g_1, \dots, g_\ell\}$ be a Gröbner basis of an ideal $I \subset \mathbb{K}[X]$ with respect to a monomial ordering $\prec$. Then $\langle G \rangle = I$.*

Proof. Let $f_1 \neq 0$ be a polynomial in I. Therefore, $\mathsf{LM}_\prec(f_1) \in \mathsf{LM}_\prec(I)$ and hence there exists $t \in \{1, \dots, \ell\}$ such that $\mathsf{LM}_\prec(g_t)$ divides $\mathsf{LM}_\prec(f_1)$. Let f_2 be the polynomial $f_2 = f_1 - \frac{\mathsf{LM}_\prec(f_1)}{\mathsf{LM}_\prec(g_i)} g_i \in I$. Notice that $\mathsf{LM}_\prec(f_2) \prec \mathsf{LM}_\prec(f_1)$. By repeating this process, one can construct a sequence $f_1, f_2, \dots$ such that for every i, $f_i \in I$ and $\mathsf{LM}_\prec(f_{i+1}) \prec \mathsf{LM}_\prec(f_i)$. By Definition 1.18, $\prec$ is a well-ordering, and hence every strictly decreasing sequence of monomials terminates. Consequently, there exists $j \in \mathbb{N}$ such that $f_j = 0$. Finally, an induction on i from j to 1 shows that $f_1 \in \langle g_1, \dots, g_\ell \rangle$. □

Notice that in Definition 1.24, Gröbner bases are not uniquely defined: if G is a Gröbner basis of I and G' is a finite family such that $G \subset G' \subset I$, then G' is also a Gröbner basis of I.

Gröbner bases which are minimal for the inclusion are called *minimal Gröbner bases*. However, minimality is not sufficient for unicity: for instance, $G = \{x_1^2, x_2^2\}$ and $G' = \{x_1^2 + x_2^2, x_2^2\}$ are two minimal Gröbner bases of the same ideal for any monomial ordering.

Unicity can be obtained by considering *reduced Gröbner basis*.

Definition 1.26. *Let $I \subset \mathbb{K}[X]$ be an ideal and $\prec$ monomial ordering. A Gröbner basis G of I with respect to $\prec$ is called*

- minimal *if for all $g_i, g_j \in G$ such that $g_i \neq g_j$, $\mathsf{LM}(g_i)$ does not divide $\mathsf{LM}(g_j)$;*
- reduced *if the leading coefficient of all basis elements is 1 and for all $g = \sum a_\alpha x^\alpha \in G$ and all $\alpha \in \mathbb{N}^n$ such that $a_\alpha \neq 0$, $x^\alpha \notin \mathsf{LM}(\langle G \setminus \{g\} \rangle)$.*

Notice that a *reduced* Gröbner basis is *minimal* and is uniquely defined.

Once a Gröbner basis of an ideal I is computed, it can be used to compute a *normal form* with respect to I, which is a projection of $\mathbb{K}[X]$ whose kernel is I. This gives an algorithm to solve the Ideal Membership Problem since a polynomial f belongs to an ideal I if and only if the normal form of f with respect to I is 0.

Definition 1.27 (Normal form)**.** *Let $\prec$ be a monomial ordering, $I \subset \mathbb{K}[X]$ be an ideal and $f \in \mathbb{K}[X]$ be a polynomial. Then there exist unique polynomials $\tilde{f}$ and g such that:*

- $f = \tilde{f} + g$;
- $g \in I$;
- *no monomials appearing in $\tilde{f}$ are in $\mathsf{LM}_\prec(I)$.*

The polynomial $\tilde{f}$ is called the normal form *of f with respect to I and $\prec$ and is denoted by $\mathrm{NF}_{\prec,I}(f)$.*

Proof. **Unicity.** Let $\tilde{f}_1, \tilde{f}_2, g_1, g_2$ be such that

- $f = \tilde{f}_1 + g_1 = \tilde{f}_2 + g_2$;
- $g_1, g_2 \in I$;
- no monomials appearing in $\tilde{f}_1, \tilde{f}_2$ are in $\mathsf{LM}_\prec(I)$.

Then $\tilde{f}_1 - \tilde{f}_2 = g_2 - g_1 \in I$, hence $\mathsf{LM}_\prec(\tilde{f}_1 - \tilde{f}_2) \in \mathsf{LM}_\prec(I)$. Since no monomials appearing in $\tilde{f}_1, \tilde{f}_2$ are in $\mathsf{LM}_\prec(I)$, $\tilde{f}_1 - \tilde{f}_2 = 0$. Consequently, $\tilde{f}_1 = \tilde{f}_2$ and $g_1 = g_2$.

Existence. The existence of the normal form is ensured by the correctness and termination of Algorithm 3 below. □

Proposition 1.28. *Let $I \subset \mathbb{K}[X]$ be an ideal, and $f \in \mathbb{K}[X]$ be a polynomial. The following statements are equivalent:*

- *$f \in I$;*
- *For any monomial ordering $\prec$, $\mathrm{NF}_{\prec,I}(f) = 0$.*

Proof. From the unicity of g in Definition 1.27, if $f \in I$, then $g = f$ and hence $\mathrm{NF}_{\prec,I}(f) = 0$ for any monomial ordering $\prec$. Conversely, if $\mathrm{NF}_{\prec,I}(f) = 0$, then $f = f - \mathrm{NF}_{\prec,I}(f) \in I$. □

1.2.2 Homogeneous, quasi-homogeneous and multi-homogeneous gradings on $\mathbb{K}[X]$

Gröbner bases have a useful property: they preserve the gradation (or multi-gradation) of ideals. Indeed, the only arithmetic operations used in Buchberger's Algorithm [Buc65] and in F_4/F_5 Algorithms [Fau99, Fau02] are multiplication of polynomials by monomials and sum of polynomials with same leading monomials. Therefore, if the input system is homogeneous (resp. quasi-homogeneous, multi-homogeneous), then a Gröbner basis computed with any of these algorithms will be homogeneous (resp. quasi-homogeneous, multi-homogeneous). Moreover, the gradation allows us to decompose the analysis of the structure of polynomial ideals and to understand the combinatorial properties of structured systems.

In this section, we give definitions and properties of $\mathbb{N}^\ell$-graded ideals. The most common case is the classical homogeneous grading: all monomials in a homogeneous polynomial share the same total degree. This notion can be extended in two ways. In the quasi-homogeneous case, a weight is attached to each variable: all monomials in a quasi-homogeneous polynomial share the same weighted degree. In a multi-homogeneous polynomial, each variable belongs to a block of variables and all monomials share the same degrees with respect to each block of variables.

Definition 1.29. *An $\mathbb{N}^\ell$-graded ring R is a ring and a decomposition into a family of additive groups $\{R_{\mathbf{d}}\}_{\mathbf{d}\in\mathbb{N}^\ell}$ such that*

$$R = \bigoplus_{\mathbf{d}\in\mathbb{N}^\ell} R_{\mathbf{d}}, \text{ and}$$
$$\forall \mathbf{d}_1, \mathbf{d}_2 \in \mathbb{N}^\ell, R_{\mathbf{d}_1} R_{\mathbf{d}_2} \subset R_{\mathbf{d}_1+\mathbf{d}_2}.$$

The first example of graded polynomial ring is given by the classical *homogeneous* grading:

Definition 1.30. *The homogeneous grading on $\mathbb{K}[X]$ is given by the decomposition*

$$\mathbb{K}[X] = \bigoplus_{d\in\mathbb{N}} \mathbb{K}[X]_d,$$

where $\mathbb{K}[X]_d$ is the $\mathbb{K}$-vector space generated by all monomials of degree d. An element of $\mathbb{K}[X]_d$ is called homogeneous of degree d.

Example 1.31. *The polynomial $3x_1^2 + 5x_1x_2 + 8x_2^2 \in \mathbb{Q}[x_1, x_2]_2$ is homogeneous of degree 2.*

The homogeneous grading can be tweaked by adding a weight on variables, giving rise to the *quasi-homogeneous* grading:

Definition 1.32. *Let $\mathbf{w} = (w_1, \ldots, w_n) \in \mathbb{N}^n$. The weight degree of a monomial w.r.t. $\mathbf{w}$ is defined by*

$$\mathrm{wdeg}_{\mathbf{w}}(x_1^{i_1} \ldots x_n^{i_n}) = \sum_{j=1}^{n} w_j i_j.$$

The quasi-homogeneous grading on $\mathbb{K}[X]$ (w.r.t. $\mathbf{w}$) is given by the decomposition

$$\mathbb{K}[X] = \bigoplus_{d\in\mathbb{N}} \mathbb{K}[X]_d^{(\mathbf{w})},$$

where $\mathbb{K}[X]_d^{(\mathbf{w})}$ *is the* $\mathbb{K}$*-vector space generated by all monomials of weight degree* d*. An element of* $\mathbb{K}[X]_d^{(\mathbf{w})}$ *is called* quasi-homogeneous of weight degree d.

Example 1.33. *The polynomial* $2x_1^2x_3+4x_1^2x_2^2+8x_2^2x_3+9x_3^2 \in \mathbb{Q}[x_1,x_2,x_3]_4$ *is quasi-homogeneous of weight degree* 4*, with respect to the weight vector* $\mathbf{w}=(1,1,2)$.

The multi-homogeneous grading provides a more refined decomposition of the polynomial ring:

Definition 1.34. *Let* $X_1,\dots,X_\ell$ *be a partition of the set of variables. Since* $\mathbb{K}[X]=\bigotimes_{j=1}^{\ell}\mathbb{K}[X_j]$*, the multi-homogeneous grading on* $\mathbb{K}[X]$ *(w.r.t. the partition* $X=\cup_{i=1}^{\ell}X_i$*) is given by the decomposition*

$$\mathbb{K}[X]=\bigoplus_{(d_1,\dots,d_\ell)\in\mathbb{N}^\ell}\mathbb{K}[X]_{(d_1,\dots,d_\ell)},$$

where $\mathbb{K}[X]_{(d_1,\dots,d_\ell)}$ *is the* $\mathbb{K}$*-vector space* $\mathbb{K}[X_1]_{d_1}\otimes\cdots\otimes\mathbb{K}[X_\ell]_{d_\ell}$ *and where the tensor product is done over* $\mathbb{K}$*. An element of* $f\in\mathbb{K}[X]_{\mathbf{d}}$ *is called* multi-homogeneous of multi-degree $\mathbf{d}\in\mathbb{N}^\ell$ *(*$\mathrm{mdeg}(f)=\mathbf{d}$*).*

Example 1.35. *The polynomial* $2x_1^2y_1+7x_1^2y_2+x_1x_2y_1+4x_1x_2y_2+8x_2^2y_1+8x_2^2y_2 \in \mathbb{Q}[x_1,x_2,y_1,y_2]_{(2,1)}$ *is multi-homogeneous of multi-degree* $(2,1)$ *with respect to the partition* $\{x_1,x_2\}\cup\{y_1,y_2\}$.

Notice that the extreme case is the multi-homogeneous grading given by the partition $\{x_1\}\cup\dots\cup\{x_n\}$. In that case, we see $\mathbb{K}[X]$ as the direct sum of vector spaces of dimension 1 (each vector space being generated by a monomial).

By extension, we also use the degree, weighted degree and multi-degree for non-homogeneous polynomials. The degree $\deg(f)$ of a polynomial $f\in\mathbb{K}[X]$ is the usual total degree, and the weighted degree $\mathrm{wdeg}_{\mathbf{w}}(f)$ is the maximum of the weighted degrees of its monomials. The multi-degree $\mathrm{mdeg}(f)$ of f with respect to a partition of the variables $X=\cup_{i=1}^{\ell}X_i$ is the ℓ-tuple $(\deg_{X_1}(f),\dots,\deg_{X_\ell}(f))\in\mathbb{N}^\ell$ where $\deg_{X_i}(f)$ is the degree of f with respect to the variables in the block X_i.

These gradings of $\mathbb{K}[X]$ are important when we consider ideals compatible with this structure:

Notations 1.36. *Let* $I\subset\mathbb{K}[X]$ *be an ideal. The notations* I_d*,* $I_d^{(\mathbf{w})}$ *and* $I_{\mathbf{d}}$ *stand for*

$$\begin{array}{rcl} I_d &=& I\cap\mathbb{K}[X]_d;\\ I_d^{(\mathbf{w})} &=& I\cap\mathbb{K}[X]_d^{(\mathbf{w})};\\ I_{\mathbf{d}} &=& I\cap\mathbb{K}[X]_{\mathbf{d}}. \end{array}$$

Definition 1.37. *An ideal* $I\subset\mathbb{K}[X]$ *is called*

- homogeneous *if* $I=\oplus_{d=0}^{\infty}I_d$*;*
- quasi-homogeneous *w.r.t.* $\mathbf{w}\in\mathbb{N}^n$ *if* $I=\oplus_{d=0}^{\infty}I_d^{(\mathbf{w})}$*;*
- multi-homogeneous *w.r.t. a partition* $X=\cup_{i=1}^{\ell}X_i$ *if* $I=\oplus_{\mathbf{d}\in\mathbb{N}^\ell}I_{\mathbf{d}}$.

Proposition 1.38. *An ideal* $I\subset\mathbb{K}[X]$ *is homogeneous (resp. quasi-homogeneous, multi-homogeneous) if and only if there exists a set of homogeneous (resp. quasi-homogeneous, multi-homogeneous) generators.*

Proof. The proof is done here in the homogeneous context (it is similar for quasi-homogeneous and multi-homogeneous systems). Let $f_1, \dots, f_m$ be a set of homogeneous generators of I and $g = \sum \lambda_t t \in I$. Let $g^{(d)}$ be its homogeneous component of degree d (i.e. $g^{(d)} = \sum\limits_{\deg(t)=d} \lambda_t t$). Therefore $g = \sum\limits_{d \in \mathbb{N}} g^{(d)}$. Since $g \in I$, there exists $h_1, \dots h_p \in \mathbb{K}[X]$ such that

$$g = \sum_{i=1}^{m} h_i f_i.$$

Since products of homogeneous polynomials are also homogeneous, $g^{(d)} = \sum_{i=1}^m h_i^{(d-\deg(f_i))} f_i$ belongs to I_d (where $h_i^{(d-\deg(f_i))}$ denotes the homogeneous component of h_i of degree $d - \deg(f_i)$) and hence I is equal to $\bigoplus_{d \in \mathbb{N}} I_d$.

Conversely, let $I = \bigoplus_{d \in \mathbb{N}} I_d$ be a homogeneous ideal and let $\mathbf{F} = (f_1, \dots, f_m) \in \mathbb{K}[X]$ be a generating family. Then the homogeneous parts of $f_1, \dots, f_m$ are in I and thus yield a homogeneous family of generators of I. □

Notice that if $\langle \mathbf{F} \rangle$ is a homogeneous ideal, then its variety can be seen as a subvariety of the projective space $\mathbb{P}^{n-1}\overline{\mathbb{K}}$ since if $\mathbf{x} \in Z(\mathbf{F})$, then for any $\lambda \in \overline{\mathbb{K}}$, $\lambda \mathbf{x} \in Z(\mathbf{F})$. In that case, we use the notation $Z(\mathbf{F}, \mathbb{P}^{n-1}) \subset \mathbb{P}^{n-1}\overline{\mathbb{K}}$ to denote this projective variety.

Similarly, if $\langle \mathbf{F} \rangle$ is a multi-homogeneous ideal, we let $Z(\mathbf{F}, \mathbb{P}^{|X^{(1)}|-1} \times \dots \times \mathbb{P}^{|X^{(\ell)}|-1}) \subset \mathbb{P}^{|X^{(1)}|-1}\overline{\mathbb{K}} \times \dots \times \mathbb{P}^{|X^{(\ell)}|-1}\overline{\mathbb{K}}$ denote the associated multi-projective variety.

An another interesting object to study the combinatorial properties of graded ideals are the so-called *Hilbert function* and *Hilbert series* of their quotient rings:

Definition 1.39. *[Eis95, Ex. 10.11] Let $I \subset \mathbb{K}[X]$ be a homogeneous (resp. quasi-homogeneous, multi-homogeneous) ideal. Then the Hilbert function $\mathsf{HF}_{\mathbb{K}[X]/I} : \mathbb{N} \to \mathbb{N}$ (resp. the weighted Hilbert function $\mathsf{wHF}_{\mathbb{K}[X]/I} : \mathbb{N} \to \mathbb{N}$, the multi-Hilbert function $\mathsf{mHF}_{\mathbb{K}[X]/I} : \mathbb{N}^\ell \to \mathbb{N}$) and the Hilbert series $\mathsf{HS}_{\mathbb{K}[X]/I} \in \mathbb{N}[[t]]$ (resp. the weighted Hilbert series $\mathsf{wHS}_{\mathbb{K}[X]/I} \in \mathbb{N}[[t]]$, the multi-Hilbert series $\mathsf{mHS}_{\mathbb{K}[X]/I} \in \mathbb{N}[[t_1, \dots, t_\ell]]$) of the quotient ring $\mathbb{K}[X]/I$ are defined by:*

$$\mathsf{HF}_{\mathbb{K}[X]/I}(d) = \dim_{\mathbb{K}}(\mathbb{K}[X]_d / I_d); \qquad \mathsf{HS}_{\mathbb{K}[X]/I}(t) = \sum_{d=0}^{\infty} \mathsf{HF}_{\mathbb{K}[X]/I}(d) t^d;$$

$$\mathsf{wHF}_{\mathbb{K}[X]/I}(d) = \dim_{\mathbb{K}}(\mathbb{K}[X]_d^{(\mathbf{w})} / I_d); \qquad \mathsf{wHS}_{\mathbb{K}[X]/I}(t) = \sum_{d=0}^{\infty} \mathsf{wHF}_{\mathbb{K}[X]/I}(d) t^d;$$

$$\mathsf{mHF}_{\mathbb{K}[X]/I}(\mathbf{d}) = \dim_{\mathbb{K}}(\mathbb{K}[X]_{\mathbf{d}} / I_{\mathbf{d}}); \quad \mathsf{mHS}_{\mathbb{K}[X]/I}(t_1, \dots, t_\ell) = \sum_{\mathbf{d} \in \mathbb{N}^\ell} \left(\mathsf{mHF}_{\mathbb{K}[X]/I}(\mathbf{d}) \prod_{j=1}^{\ell} t_j^{d_j} \right).$$

Proposition 1.40. *The Hilbert series, weighted Hilbert series and multi-Hilbert series of $\mathbb{K}[X]$ are respectively*

- $\mathsf{HS}_{\mathbb{K}[X]}(t) = \dfrac{1}{(1-t)^n}$;
- $\mathsf{wHS}^{(\mathbf{w})}_{\mathbb{K}[X]}(t) = \dfrac{1}{\prod_{i=1}^n (1 - t^{w_i})}$;

- $\mathsf{mHS}_{\mathbb{K}[X]}(t_1,\dots,t_\ell)=\dfrac{1}{\prod_{i=1}^{n}(1-t_i)^{|X^{(i)}|}}.$

Proof. First, we prove the formula for the weighted Hilbert series of $\mathbb{K}[X]$ (the homogeneous case is obtained by choosing the weight vector $\mathbf{w}=(1,\dots,1)$). This is achieved by a combinatorial argument: $\mathsf{wHS}_{\mathbb{K}[X]}(t)$ is the generating series of $\mathbb{N}^n$ where the size of an element $\mathbf{d}\in\mathbb{N}^n$ is the dot product $\mathbf{d}\cdot\mathbf{w}$:

$$\mathsf{wHS}_{\mathbb{K}[X]}(t)=\sum_{\mathbf{d}\in\mathbb{N}^n} t^{\mathbf{d}\cdot\mathbf{w}}.$$

Therefore $\mathbb{N}^n$ is the combinatorial product of n copies of $\mathbb{N}$ where the size of an element d of the jth copy of $\mathbb{N}$ is dw_j. The generating series of the product of combinatorial classes is the product of their generating series:

$$\begin{aligned}\mathsf{wHS}_{\mathbb{K}[X]}(t) &= \prod_{j=1}^{n}\left(\sum_{d\in\mathbb{N}} t^{dw_j}\right)\\ &= \frac{1}{\prod_{j=1}^{n}(1-t^{w_j})}.\end{aligned}$$

Similarly for the multi-homogeneous case,

$$\begin{aligned}\mathsf{mHS}_{\mathbb{K}[X]}(t_1,\dots,t_\ell) &= \sum_{\substack{d\in\mathbb{N}\\ \mathfrak{m}\in\mathsf{Monomials}(\mathbb{K}[X],d)}} \mathbf{t}^{\mathrm{mdeg}(\mathfrak{m})}\\ &= \prod_{i=1}^{\ell}\mathsf{HS}_{\mathbb{K}[X^{(i)}]}(t_i)\\ &= \prod_{i=1}^{\ell}\frac{1}{(1-t_i)^{|X^{(i)}|}}.\end{aligned}$$

□

In the following proposition, we show that algebraic properties yield relations between Hilbert series. These relations will be often used in this thesis to obtain explicit formulas for the Hilbert series of structured ideals.

Proposition 1.41. *Let $I\subset\mathbb{K}[X]$ be a homogeneous ideal (resp. quasi-homogeneous, multi-homogeneous) and $f\in\mathbb{K}[X]_d$ be a homogeneous polynomial of degree $d\in\mathbb{N}$ (resp. $f\in\mathbb{K}[X]_d^{(\mathbf{w})}$ be a quasi-homogeneous polynomial of weight degree $d\in\mathbb{N}$, $f\in\mathbb{K}[X]_{\mathbf{d}}$ be a multi-homogeneous polynomial of multi-degree $\mathbf{d}\in\mathbb{N}^\ell$). If f does not divide 0 in the ring $\mathbb{K}[X]/I$, then*

$$\begin{aligned}\mathsf{HS}_{\mathbb{K}[X]/(I+\langle f\rangle)}(t) &= (1-t^d)\,\mathsf{HS}_{\mathbb{K}[X]/I}(t);\\ \mathsf{wHS}^{(\mathbf{w})}_{\mathbb{K}[X]/(I+\langle f\rangle)}(t) &= (1-t^d)\,\mathsf{wHS}^{(\mathbf{w})}_{\mathbb{K}[X]/I}(t);\\ \mathsf{mHS}_{\mathbb{K}[X]/(I+\langle f\rangle)}(t_1,\dots,t_\ell) &= (1-\textstyle\prod_{j=1}^{\ell}t_j^{d_j})\,\mathsf{mHS}_{\mathbb{K}[X]/I}(t_1,\dots,t_\ell).\end{aligned}$$

Proof. The proof is done here in the homogeneous context; the proofs for the quasi-homogeneous and multi-homogeneous gradings are similar. For every $\ell\in\mathbb{N}$, consider the following sequence of $\mathbb{K}$-vector spaces:

$$0\longrightarrow\mathbb{K}[X]_\ell/I_\ell\xrightarrow{\times f}\mathbb{K}[X]_{\ell+d}/I_{\ell+d}\xrightarrow{\pi}\mathbb{K}[X]_{\ell+d}/(I+\langle f\rangle)_{\ell+d}\longrightarrow 0,$$

where π is the canonical projection. Since f does not divide 0 in $\mathbb{K}[X]/I$, this sequence is exact. Therefore the alternate sum of the dimensions of these vector spaces is equal to 0. Consequently, $\mathsf{HF}_{\mathbb{K}[X]/I}(\ell) - \mathsf{HF}_{\mathbb{K}[X]/I}(\ell+d) + \mathsf{HF}_{\mathbb{K}[X]/(I+\langle f\rangle)}(\ell+d) = 0$, thus by multiplying this relation by $t^{d+\ell}$ and by summing over $\ell \in \mathbb{Z}$,

$$t^d\,\mathsf{HS}_{\mathbb{K}[X]/I}(t) - \mathsf{HS}_{\mathbb{K}[X]/I}(t) + \mathsf{HS}_{\mathbb{K}[X]/(I+\langle f\rangle)}(t) = 0.$$

□

A non-trivial property of Hilbert series of quotients $\mathbb{K}[X]/I$ is that they are always power series expansions of rational functions. This can be seen as a consequence of the Hilbert Syzygy Theorem (see Definition-Proposition 1.14). The numerator of this rational function is called the *K-polynomial* in [MS05].

Proposition 1.42. *[MS05, Theorem 8.20] Let $I \subset \mathbb{K}[X]$ be a homogeneous (resp. quasi-homogeneous, multi-homogeneous) ideal. Then there exists a polynomial $N(t) \in \mathbb{Z}[t]$ (resp. $N(t) \in \mathbb{Z}[t], N(t_1,\dots,t_\ell) \in \mathbb{Z}[t_1,\dots,t_\ell]$) such that:*

$$\begin{array}{rcl} \mathsf{HS}_{\mathbb{K}[X]/I}(t) & = & \dfrac{N(t)}{(1-t)^n}; \\ \mathsf{wHS}^{\mathbf{w}}_{\mathbb{K}[X]/I}(t) & = & \dfrac{N(t)}{\prod_{i=1}^n (1-t^{w_i})}; \\ \mathsf{mHS}_{\mathbb{K}[X]/I}(t_1,\dots,t_\ell) & = & \dfrac{N(t_1,\dots,t_\ell)}{\prod_{i=1}^n (1-t_i)^{|X^{(i)}|}}. \end{array}$$

Proof. The proof is done in the homogeneous context, but the proofs for quasi-homogeneous and multi-homogeneous ideals are exactly similar. By Hilbert Syzygy Theorem [Eis95, Thm. 1.13], I has a graded finite free resolution of length $r \le n$. Therefore, for any $d \in \mathbb{N}$, there is an exact sequence of $\mathbb{K}$-vector spaces

$$0 \to \bigoplus_{j=1}^{i_r} \mathbb{K}[X]_{d-d_{r,j}} \xrightarrow{\varphi_n} \dots \xrightarrow{\varphi_1} \bigoplus_{j=1}^{i_0} \mathbb{K}[X]_{d-d_{0,j}} \to \mathbb{K}[X]_d/I_d \to 0,$$

where $d_{i,j} \le d$ for all i,j. Since the alternate sum of the dimensions in an exact sequence of vector spaces is 0, we obtain that

$$\dim(\mathbb{K}[X]_d/I_d) = \sum_{k=0}^{r} (-1)^k \sum_{j=1}^{i_k} \dim(\mathbb{K}[X]_{d-d_{k,j}}).$$

Then, by letting $[t^d]S(t)$ denote the coefficient of t^d in a power series $S \in \mathbb{Z}[[t]]$, notice that

$$\begin{array}{rcl} \dim(\mathbb{K}[X]_{d-d_{k,j}}) & = & \dbinom{n+d-d_{k,j}-1}{d-d_{k,j}} \\ & = & [t^{d-d_{k,j}}]\dfrac{1}{(1-t)^n} \\ & = & [t^d]\dfrac{t^{d_{k,j}}}{(1-t)^n}. \end{array}$$

Therefore, by summing over d, we get

$$\begin{aligned}\mathsf{HS}_{\mathbb{K}[X]/I}(t) &= \sum_{d\geq 0}\dim(\mathbb{K}[X]_d/I_d)t^d \\ &= \sum_{k=0}^{r}(-1)^k\sum_{j=1}^{i_k}\frac{t^{d_{k,j}}}{(1-t)^n}.\end{aligned}$$

□

In the homogeneous case, the degree and the dimension can be read off from the Hilbert series.

Proposition 1.43. *Let $I \subset \mathbb{K}[X]$ be a proper homogeneous ideal and $\mathsf{HS}_{\mathbb{K}[X]/I}(t) = \frac{N(t)}{(1-t)^d}$ be the irreducible form of its Hilbert series (i.e. $N(1) \neq 0$). Then $\dim(I) = d$. Moreover, if $d = 0$, then $\mathsf{HS}_{\mathbb{K}[X]/I}(t)$ is a polynomial and $\mathsf{DEG}(I) = \mathsf{HS}_{\mathbb{K}[X]/I}(1)$.*

Proof. In [CLO97, Ch. 9, §3, Thm.11], it is proved that the degree of the Hilbert polynomial $\mathsf{HP}_{\mathbb{K}[X]/I}$ is equal to the projective dimension of I, which is $\dim(I) + 1$. Since a power series $\frac{N(t)}{(1-t)^d}$ is the generating series of the polynomial of degree $d-1$ if $\deg(N(t)) < d$ (see the proof of Definition - Proposition 1.65 for more details), we obtain $\dim(I) = d$. If $d = 0$, then $\mathsf{HS}_{\mathbb{K}[X]/I}(1) = \sum_{d\in\mathbb{N}}\dim_{\mathbb{K}}(\mathbb{K}[X]_d/I_d) = \dim_{\mathbb{K}}(\mathbb{K}[X]/I) = \mathsf{DEG}(I)$. □

1.2.3 Regular and Semi-regular Sequence

Regular and semi-regular sequences are important families of polynomial systems. Indeed, being regular is a *generic property*. Semi-regular sequences are also conjectured to be generic (see Conjecture 1.53 below).

Regular sequences

Definition 1.44. *A sequence of non-zero homogeneous polynomials $\mathbf{F} = (f_1, \dots, f_m) \in \mathbb{K}[X]^m$ is called* regular *if for all $i \in \{1, \dots, m-1\}$, f_{i+1} does not divide 0 in the ring $\mathbb{K}[X]/\langle f_1, \dots, f_i\rangle$.*

Proposition 1.45. *Let $\mathbf{F} = (f_1, \dots, f_m) \in \mathbb{K}[X]^m$ be a sequence of homogeneous polynomials. The following statements are equivalent:*

1. *$\mathbf{F}$ is a regular sequence;*
2. *for all $i \in \{1, \dots, m\}$, $\mathbf{F}_i = (f_1, \dots, f_i)$ is a regular sequence;*
3. *for all $i \in \{1, \dots, m-1\}$, $\langle \mathbf{F}_i\rangle : f_{i+1} = \langle \mathbf{F}_i\rangle$;*
4. *for all $i \in \{1, \dots, m-1\}$ and all $P \in \mathrm{Ass}(\langle f_1, \dots, f_i\rangle)$, $f_{i+1} \notin P$;*
5. *for all $i \in \{1, \dots, m\}$, $\dim(\langle \mathbf{F}_i\rangle) = n - i$.*

Proof. By definition of regular sequences, the statements (1), (2) and (3) are clearly equivalent. Let $\langle f_1, \dots, f_i\rangle = I_1 \cap \dots \cap I_\ell$ be a minimal primary decomposition of $\langle f_1, \dots, f_i\rangle$. Suppose first that f_{i+1} does divide 0 in $\mathbb{K}[X]/\langle f_1, \dots, f_i\rangle$. Thus there exists $g \notin \langle f_1, \dots, f_i\rangle$ such that $f_{i+1}g \in \langle f_1, \dots, f_i\rangle$. Since $g \notin \langle f_1, \dots, f_i\rangle$, there exists j such that $g \notin I_j$. Since I_j is primary, there exists $k \in \mathbb{N}$ such that $f_{i+1}^k \in I_j$, and consequently, $f_{i+1} \in \sqrt{I_j} \in \mathrm{Ass}(\langle f_1, \dots, f_i\rangle)$. Conversely, suppose that $f_{i+1} \in P$, with $P \in \mathrm{Ass}(\langle f_1, \dots, f_i\rangle)$. Then by definition of Ass, there exists $g \notin \langle f_1, \dots, f_i\rangle$ such that $gP \subset I$. Therefore $f_{i+1}g \in I$ and hence f_{i+1} divides 0 in $\mathbb{K}[X]/\langle f_1, \dots, f_i\rangle$.

The fact that (1) is equivalent to (5) will be proved in Theorem 1.48. □

Regular sequences have an interesting property: all algebraic relations in a regular sequence can be deduced from the commutativity of $\mathbb{K}[X]$, i.e. from the relations $f_i f_j - f_j f_i$. Intuitively, this means that the ideal generated by a sequence $\mathbf{F} = (f_1, \ldots, f_m)$ (with $m \leq n$) is "largest" when $\mathbf{F}$ is a regular sequence. This notion of algebraic relations is formalized in the following definition and proposition.

Definition 1.46 (Syzygy). *Let $\mathbf{F} = (f_1, \ldots, f_m) \in \mathbb{K}[X]^m$ be a sequence of polynomials. The* syzygy module *of $\mathbf{F}$ is the submodule $\mathsf{Syz}(\mathbf{F}) \subset \mathbb{K}[X]^m$ of all vectors $(s_1, \ldots, s_m) \in \mathbb{K}[X]^m$ such that $\sum_{i=1}^{m} s_i f_i = 0$. The degree of a syzygy $\mathbf{s} \in \mathsf{Syz}(\mathbf{F})$ is defined as* $\deg(\mathbf{s}) = \max_{1 \leq i \leq m}\{\deg(s_i) + \deg(f_i)\}$.

Proposition 1.47. *Let $\mathbf{F} = (f_1, \ldots, f_m) \in \mathbb{K}[X]^m$ be a polynomial family. The two following statements are equivalent:*

- $\mathbf{F}$ *is a regular sequence;*
- *the syzygy module of $\mathbf{F}$ is generated by the syzygies coming from the commutativity of $\mathbb{K}[X]$:*

$$\mathsf{Syz}(\mathbf{F}) = \langle f_i \mathbf{e}_j - f_j \mathbf{e}_i \rangle,$$

where $\mathbf{e}_i = (0, \ldots, 0, 1, 0, \ldots, 0) \in \mathbb{K}[X]^m$ is the vector whose only nonzero entry is at the i-th position.

Proof. We prove this proposition by induction on m. Let $\mathbf{s} \in \mathsf{Syz}(\mathbf{F})$ be a syzygy. If $m = 1$, then $\mathbf{F} = (f_1)$ with $f_1 \neq 0$ (since $\mathbf{F}$ is a regular sequence) and hence $\mathsf{Syz}(\mathbf{F}) = 0 \in \mathbb{K}[X]$. Now assume that $m > 1$. Then

$$\sum_{i=1}^{m} s_i f_i = 0.$$

Therefore s_m belongs to the colon ideal $\langle f_1, \ldots, f_{m-1} \rangle : f_m$. By Definition 1.44, $\langle f_1, \ldots, f_{m-1} \rangle = \langle f_1, \ldots, f_{m-1} \rangle : f_m$ and hence s_m can be written as

$$s_m = \sum_{i=1}^{m-1} f_i h_i.$$

Now, consider the syzygy $\mathbf{s}' = \mathbf{s} - h_i(f_i \mathbf{e}_m - f_m \mathbf{e}_i)$. Then $s'_m = 0$ and $\sum_{i=1}^{m-1} s'_i f_i = 0$. Therefore by the inductive hypothesis, $\mathbf{s}'$ is in the module generated by the syzygies $\langle f_i \mathbf{e}_j - f_j \mathbf{e}_i \rangle_{1 \leq i,j \leq m-1} \subset \langle f_i \mathbf{e}_j - f_j \mathbf{e}_i \rangle_{1 \leq i,j \leq m}$. Therefore $\mathbf{s} = \mathbf{s}' + h_i(f_i \mathbf{e}_m - f_m \mathbf{e}_i) \in \langle f_i \mathbf{e}_j - f_j \mathbf{e}_i \rangle$.

Conversely, for $m = 1$, it is clear that $\mathsf{Syz}(\mathbf{F}) = \mathbf{0}$ if and only if $f_1 \neq 0$. By induction, assume now that $(f_1, \ldots, f_{m-1})$ is a regular sequence. Let $s_m \in \langle f_1, \ldots, f_{m-1} \rangle : f_m$ be a polynomial. Then there exists $s_1, \ldots, s_{m-1} \in \mathbb{K}[X]$ such that $\mathbf{s} = (s_1, \ldots, s_m)$ is a syzygy. Since $\mathsf{Syz}(\mathbf{F}) = \langle f_i \mathbf{e}_j - f_j \mathbf{e}_i \rangle$, it follows that $s_m \in \langle f_1, \ldots, f_{m-1} \rangle$ and hence $\langle f_1, \ldots, f_{m-1} \rangle : f_m = \langle f_1, \ldots, f_{m-1} \rangle$. Consequently, $\mathbf{F}$ is a regular sequence. □

The Hilbert series of regular systems is a direct consequence of Proposition 1.41:

Theorem 1.48. *[Bar04, BFS04, BFSY04] Let $\mathbf{F} = (f_1, \ldots, f_m) \in \mathbb{K}[X]^m$ be a homogeneous system with $m \leq n$. The three following statements are equivalent:*

- $\mathbf{F}$ *is regular;*

- *the Hilbert series of* $\mathbb{K}[X]/\langle \mathbf{F} \rangle$ *is*

$$\mathsf{HS}_{\mathbb{K}[X]/\langle \mathbf{F} \rangle}(t) = \frac{\prod_{i=1}^{m}(1 - t^{\deg(f_i)})}{(1-t)^n};$$

- $\dim(\langle \mathbf{F} \rangle) = n - m.$

This notion of regularity is essential since the regular (and semi-regular) sequences correspond exactly to the systems such that there is no reduction to zero during the computation of a Gröbner basis with the F_5 Algorithm (see [Fau02]). Moreover, generic systems with less equations than variables are regular, as shown by the following theorem:

Theorem 1.49 (Genericity of homogeneous regular sequences)**.** *Let* $m \le n$ *and* $(d_1, \dots, d_m) \in \mathbb{N}^m$ *be a sequence of degrees. Then there exists a Zariski open subset* $O \subset \overline{\mathbb{K}}[X]_{d_1} \times \cdots \times \overline{\mathbb{K}}[X]_{d_m}$ *such that any* $\mathbf{F} \in O$ *is a regular sequence.*

Proof. See [Par10, Section 2]. □

Homogeneous semi-regular sequences

Semi-regular sequences extend the notion of regularity when there are more equations than variables. The Hilbert series of semi-regular sequences is known and is given below.

Notations 1.50. *Let* $S \in \mathbb{Z}[[t]]$ *be a power series. We let* $[S]_+ \in \mathbb{N}[[t]]$ *denote the series obtained by truncating* S *at its first non-positive coefficient. Notice that if there is a non-positive coefficient in* S*, then* $[S]_+$ *is a polynomial.*

Theorem 1.51. *[Bar04, BFS04, BFSY04, Die] Let* $\mathbf{F} = (f_1, \dots, f_m) \in \mathbb{K}[X]^m$ *be a homogeneous system with* $m \ge n$*. The three following statements are equivalent:*

- *the Hilbert series of* $\mathbb{K}[X]/\langle \mathbf{F} \rangle$ *is*

$$\mathsf{HS}_{\mathbb{K}[X]/\langle \mathbf{F} \rangle}(t) = \left[\frac{\prod_{i=1}^{m}(1 - t^{\deg(f_i)})}{(1-t)^n}\right]_+;$$

- *the ideal* $\langle \mathbf{F} \rangle$ *has dimension* 0 *and every syzygy of* $\mathbf{F}$ *of degree at most* $\deg(\mathsf{HS}_{\mathbb{K}[X]/\langle \mathbf{F} \rangle})$ *is in the module generated by the trivial syzygies* $\langle f_i \mathbf{e}_j - f_j \mathbf{e}_i \rangle$*.*

If the sequence $\mathbf{F}$ *verifies these properties, it is called* semi-regular.

Another genericity property is given below, it yields a sufficient condition for a sequence to be semi-regular:

Proposition 1.52. *Let* $\mathbf{F} = (f_1, \dots, f_m) \in \mathbb{K}[X]^m$ *be a sequence of homogeneous non-zero polynomials. If for all* $i \in \{1, \dots, m-1\}$*, and for all* $d \in \mathbb{N}$*, the linear map*

$$\mathbb{K}[X]_d/\langle f_1, \dots, f_i \rangle_d \xrightarrow{\times f_i} \mathbb{K}[X]_{d+\deg(f_i)}/\langle f_1, \dots, f_i \rangle_{d+\deg(f_i)}$$

is of maximal rank (i.e. it is either injective or surjective), then the sequence $\mathbf{F}$ *is semi-regular.*

The analysis of semi-regular systems is crucial for understanding the behavior of Gröbner basis algorithms, since it has been observed that in practice, random systems are semi-regular. However, this is not proved and is expressed by the famous *Fröberg's conjecture* (which is reformulated here, see [Fro85] for the original statement):

Conjecture 1.53 (Fröberg's Conjecture)**.** *[Fro85] Let $(d_1, \dots, d_m) \in \mathbb{N}^m$ be a sequence of integers and $\mathscr{S}$ be the $\overline{\mathbb{K}}$-vector space of homogeneous systems $\mathbf{F} = (f_1, \dots, f_m) \in \overline{\mathbb{K}}[X]^m$ such that for all $i \in \{1, \dots, m\}$, $\deg(f_i) = d_i$. Then there exists a non-empty Zariski open subset $O \subset \mathscr{S}$ such that every system $\mathbf{F} \in O$ is semi-regular.*

Although this conjecture remains an important open problem in commutative algebra, it has been proved in several cases:

- when $n \geq m$ (see Theorem 1.49);
- $n = 2$;
- $n = 3$ in characteristic 0;
- $m = n + 1$;
- when the system is quadratic and $n \leq 11$;
- when the system is cubic and $n \leq 8$.

We refer to [Bar04] for more details on semi-regular systems.

Affine semi-regular systems

In [Bar04, BFS04, BFSY04], the definition of semi-regular systems is extended to affine systems (since in applications, systems arising are usually affine) by considering the homogeneous part of highest degree.

Definition 1.54 (Affine Semi-Regular Sequence)**.** *Let $\mathbf{F} = (f_1, \dots, f_m) \in \mathbb{K}[X]^m$ be a polynomial system. $\mathbf{F}$ is called* semi-regular *if the system of homogeneous parts of highest degrees $\mathbf{F}^{(h)} = (f_1^{(h)}, \dots, f_m^{(h)})$ is semi-regular.*

We will use similar definitions and techniques for the analysis of overdetermined systems in Section 4.5 in the context of determinantal systems.

1.2.4 Boolean semi-regular systems.

When $\mathbb{K}$ is the boolean field GF_2 and when we want to find boolean solutions in GF_2, a standard strategy is to add the so-called *field equations* $x_i^2 - x_i = 0$. Systems with field equations are not semi-regular in the sense of Definition 1.54. Consequently, in [Bar04, BFSY04], a notion of *semi-regularity over* GF_2 is introduced which takes into account the relation $f^2 = f \mod \langle x_1^2 - x_1, \dots, x_n^2 - x_n\rangle$ for any polynomial $f \in \mathsf{GF}_2[X]$.

Definition 1.55 (semi-regular sequence over GF_2)**.** *[Bar04, BFSY04] Let $\mathbf{F} = (f_1, \dots, f_m) \in \mathsf{GF}_2[X]^m$ be a polynomial system. It is called* boolean semi-regular *if the system of homogeneous parts of highest degrees $\mathbf{F}^{(h)} = (f_1^{(h)}, \dots, f_m^{(h)})$ verifies the following property: for all $i \in \{1, \dots, m\}$ and for all $g \in \mathsf{GF}_2[X]$ such that $f_i^{(h)} g \in \langle f_1^{(h)}, \dots, f_{i-1}^{(h)}, x_1^2, \dots, x_n^2\rangle$ and $\deg(f_i^{(h)} g) \leq \mathsf{i}_{\mathsf{reg}}(\langle f_1^{(h)}, \dots, f_i^{(h)}, x_1^2, \dots, x_n^2\rangle)$ (where $\mathsf{i}_{\mathsf{reg}}$ is the* index of regularity*, see Definition-Proposition 1.65), then $g \in \langle f_1^{(h)}, \dots, f_i^{(h)}\rangle$.*

In Chapter 7, we define another notion of *boolean semi-regularity* by investigating properties of homogenized boolean systems with homogenized field equations $x_i^2 - x_i h = 0$.

1.3 Polynomial System Solving with Gröbner Bases

Historically, Gröbner bases were introduced by Buchberger [Buc65] in order to solve the Ideal Membership Problem, i.e. given polynomials $f, f_1, \dots, f_m \in \mathbb{K}[X]$, decide whether f belongs to the ideal $\langle f_1, \dots, f_m \rangle$ or not. Nowadays Gröbner basis is also one the most standard tools for solving symbolically algebraic systems of multivariate equations. In particular, Gröbner bases with respect to the lexicographical ordering have interesting properties for polynomial system solving.

But first, we need to define what "Polynomial System Solving" exactly means. Indeed, in this thesis, we focus on *explicit* and *exact solutions* of polynomial systems. When $\mathbb{K}$ is a finite field, a solving algorithm outputs the list of all solutions. These can be obtained from a *lex* Gröbner basis by solving a sequence of univariate polynomials (see Proposition 1.56 below).

When $\mathbb{K}$ has characteristic 0, we want an algebraic description of the solutions from which properties can be easily extracted (as well as certified approximations of the solutions). This can be achieved for instance with the *lextriangular* algorithm [Laz92] which takes as input a lexicographical Gröbner basis and outputs a decomposition in triangular sets.

Therefore, in the whole thesis, we will focus on computing *lex* Gröbner bases of polynomial systems. This is also motivated by the triangular structure of 0-dimensional lex Gröbner bases, which is described in the following proposition.

Proposition 1.56. *Let $I \subset \mathbb{K}[X]$ be a 0-dimensional ideal and $G = \{g_1, \dots, g_\ell\}$ be a minimal Gröbner basis of I with respect to $\prec_{\mathsf{lex}}$, such that $\mathsf{LM}(g_\ell) \prec_{\mathsf{lex}} \dots \prec_{\mathsf{lex}} \mathsf{LM}(g_1)$. Then $g_\ell \in \mathbb{K}[x_n]$ and there exists a strictly increasing sequence $1 = i_1 < i_2 < \dots < i_n = \ell$, such that for all $j \in \{1, \dots, n-1\}$ and all $k \in \{i_j, \dots, i_{j+1} - 1\}$, $g_k \in \mathbb{K}[x_j, \dots, x_n]$ and $g_k \notin \mathbb{K}[x_{j+1}, \dots, x_n]$:*

$$G = \left\{ \begin{array}{c} g_1(x_1, \dots, x_n) \\ \vdots \\ g_{i_2-1}(x_1, \dots, x_n) \\ g_{i_2}(x_2, \dots, x_n) \\ \vdots \\ g_{i_3}(x_3, \dots, x_n) \\ \vdots \\ g_{\ell-1}(x_{n-1}, x_n) \\ g_\ell(x_n) \end{array} \right\}$$

If $\mathbb{K}$ is a finite field, a possible strategy to obtain the solutions in $\mathbb{K}^n$ (or in a finite extension of $\mathbb{K}$) of a polynomial system of equations is to compute a lexicographical Gröbner basis of the ideal generated by the polynomials. Then by solving the univariate polynomial g_ℓ, we recover the possible values of the variable x_n. Substituting these values in the equations involving x_{n-1} and x_n, we can recover the possible values of x_{n-1}. By repeating this process, all solutions of the initial system can be recovered by solving a sequence of univariate equations.

Often, the lex Gröbner basis of a 0-dimensional ideal already yields a rational parametrization of the variety: under some assumptions that are satisfied generically, a lex Gröbner basis is in the so-called *shape position*.

Definition 1.57 (Shape position)**.** *Let* $\mathbf{F} = (f_1, \ldots, f_m) \in \mathbb{K}[X]^m$ *be a* 0*-dimensional polynomial system. The system* $\mathbf{F}$ *is said to be in* shape position *if the reduced lex Gröbner basis of* $\langle \mathbf{F} \rangle$ *has the following shape:*

$$G = \left\{ \begin{array}{c} x_1 - h_1(x_n) \\ \vdots \\ x_{n-1} - h_{n-1}(x_n) \\ h_n(x_n) \end{array} \right\},$$

where $h_1, \ldots, h_n$ *are univariate polynomials*

When the system is in shape position, then after computing a lex Gröbner basis, the solutions of the univariate polynomial h_n give an explicit description of the zeroes of the system. Moreover, it has been proved in [BMMT94] that, if the ideal $\langle \mathbf{F} \rangle$ is radical, then the probability that it becomes in shape position after a random linear change of coordinates is overwhelming (provided that the cardinality of the field $\mathbb{K}$ is large enough).

1.3.1 Gröbner basis Algorithms

In this section, we describe algorithmic tools used for computing Gröbner basis. Notice that these algorithms are simplified variants of what is implemented in practice (for instance in the FGb library[1]). These simplifications are made in order to make the complexity analysis easier. We refer the reader to the articles [Fau99, Fau02, FGLM93, FM11, FL10] and to references therein for a precise description of the state of art algorithms for computing Gröbner bases.

Definition 1.58. *Let* $\prec$ *be a monomial ordering,* $\mathbf{F} = (f_1, \ldots, f_m) \in \mathbb{K}[X]^m$ *be homogeneous polynomials of respective degrees* $(d_1, \ldots, d_m) \in \mathbb{N}^m$*. The Macaulay matrix of* $f_1, \ldots, f_m$ *in degree* D *is the matrix* $\mathsf{Mac}_{\prec,D}(\mathbf{F})$ *with entries in* $\mathbb{K}$ *such that:*

- *the number of rows is* $\sum_{i=1}^{m} \binom{n+D-d_i-1}{n}$*; a signature* (i, t) *is attached to each row, where* $i \in \{1, \ldots, m\}$ *and* $t \in \mathbb{K}[X]$ *is a monomial of degree* $D - d_i$*. The rows are sorted in decreasing order as follows:*

$$(i_1, t_1) > (i_2, t_2) \Leftrightarrow \begin{cases} i_1 < i_2 \text{ or} \\ (i_1 = i_2 \text{ and } t_2 \prec t_1); \end{cases}$$

- *the number of columns is* $\binom{n+D-1}{n}$*, a signature* (u) *is attached to each column, where* $u \in \mathbb{K}[X]$ *is a monomial of degree* D*. They are sorted in decreasing ordering with respect to* $\prec$*;*

- *the element in* $\mathsf{Mac}_{\prec,D}(\mathbf{F})$ *at the intersection of the row* (i, t) *and the column* (u) *is the coefficient of the monomial* u *in the polynomial* $t f_i$*.*

The Macaulay matrices up to some degree d can be used to compute a partial Gröbner basis, called d-Gröbner basis:

Definition 1.59 (d-Gröbner basis)**.** *Let* $\prec$ *be a monomial ordering,* $I \subset \mathbb{K}[X]$ *be a homogeneous ideal and* $d \in \mathbb{N}$*. A* d*-Gröbner basis of* I *with respect to* $\prec$ *is a finite subset* $G = \{g_1, \ldots, g_\ell\} \subset I$ *such that, for every polynomial* $f \in I$ *of degree at most* d*,*

$$\mathsf{LM}_\prec(f) \in \langle \mathsf{LM}_\prec(g_1), \ldots, \mathsf{LM}_\prec(g_\ell) \rangle \,.$$

[1] Available at http://www-calfor.lip6.fr/~jcf/Software/

This notion of d-Gröbner basis is motivated by the fact that for d large enough, d-Gröbner bases are actually Gröbner bases:

Proposition 1.60. *Let $\prec$ be a monomial ordering and $I \subset \mathbb{K}[X]$ be a homogeneous ideal. There exists an integer $d_0 \in \mathbb{N}$ such that for every $d \geq d_0$, every d-Gröbner basis of I is a Gröbner basis of I.*

Proof. Consider the following increasing sequence of ideals

$$\langle \mathsf{LM}_\prec(I_0)\rangle \subset \langle \mathsf{LM}_\prec(I_0) \cup \mathsf{LM}_\prec(I_1)\rangle \subset \cdots \subset \langle \cup_{\ell=0}^{d} \mathsf{LM}_\prec(I_\ell)\rangle \subset \ldots.$$

Since $\mathbb{K}[X]$ is Noetherian, there exists $d_0 \in \mathbb{N}$ such that this chain stabilizes

$$\forall d \geq d_0, \left\langle \bigcup_{\ell=0}^{d} \mathsf{LM}_\prec(I_\ell) \right\rangle = \left\langle \bigcup_{\ell=0}^{d_0} \mathsf{LM}_\prec(I_\ell) \right\rangle.$$

Notice that $\mathsf{LM}(I) = \langle \bigcup_{\ell=0}^{\infty} \mathsf{LM}_\prec(I_\ell)\rangle = \left\langle \bigcup_{\ell=0}^{d_0} \mathsf{LM}_\prec(I_\ell)\right\rangle$. Let $d \geq d_0$, $G = (g_1, \ldots, g_t)$ be a d-Gröbner basis of I and $t \in \mathsf{LM}_\prec(I)$ be a monomial. Then t belongs to $\left\langle \bigcup_{\ell=0}^{d_0} \mathsf{LM}_\prec(I_\ell)\right\rangle$. Consequently there exists a monomial $u \in \bigcup_{\ell=0}^{d_0} \mathsf{LM}_\prec(I_\ell)$ of degree at most d_0 which divides t. Therefore, $u \in \langle \mathsf{LM}_\prec(G)\rangle$, hence $t \in \langle \mathsf{LM}_\prec(G)\rangle$. Consequently, G is a Gröbner basis of I. $\square$

Algorithm 2 Homogeneous Lazard's algorithm

Input: $\mathbf{F} = (f_1, \ldots, f_m) \in \mathbb{K}[X]^m$ a homogeneous family of polynomials of degrees $(d_1, \ldots, d_m)$;
$\prec$ a monomial ordering;
an integer D.
Output: G a D-Gröbner basis of $\langle F\rangle$ w.r.t. $\prec$.

- $G \leftarrow \emptyset$.
- **for** d from 1 to D **do**
  - $\mathbf{M}_d \leftarrow \binom{n+d-1}{d} \times 1$ vector of monomials of degree d in $\mathbb{K}[X]$ sorted in decreasing ordering with respect to $\prec$.
  - $\mathsf{Mac}'_{\prec,d}(\mathbf{F}) \leftarrow \mathsf{RowEchelonForm}(\mathsf{Mac}_{\prec,d}(\mathbf{F}))$.
  - $\mathbf{R}_d \leftarrow \mathsf{Mac}'_{\prec,d}(\mathbf{F}) \cdot \mathbf{M}_d$.
  - $G \leftarrow G \cup \{h \in \mathbf{R} \mid \forall g \in G, \mathsf{LM}_\prec(g) \text{ does not divide } \mathsf{LM}_\prec(h)\}$.
- **end for**
- Return G.

Theorem 1.61. *Algorithm 2 terminates and returns a D-Gröbner basis of $\langle \mathbf{F}\rangle$.*

Proof. The termination is straightforward since the algorithm enters the main loop a fixed number of times and there is no recursive call. We prove now that the output is a D-Gröbner basis of $\langle \mathbf{F}\rangle$. Notice that for all d, the rows of $\mathsf{Mac}_{\prec,d}(\mathbf{F})$ generate the vector space $\langle \mathbf{F}\rangle_d$. Since $\mathsf{Mac}'_{\prec,d}(\mathbf{F})$ is a row echelon basis of $\langle \mathbf{F}\rangle_d$, we obtain

$$\{\mathsf{LM}_\prec(h) \mid h \in \langle \mathbf{F}\rangle_d\} = \{\mathsf{LM}_\prec(h) \mid h \in \mathbf{R}_d\}.$$

Therefore, for any $h \in \langle \mathbf{F}\rangle$ of degree at most D, there exists a polynomial h' in $\mathbf{R}_{\deg(\mathsf{LM}_\prec(h))}$ such that $\mathsf{LM}_\prec(h) = \mathsf{LM}_\prec(h')$, hence there exists $g \in G$ such that $\mathsf{LM}_\prec(g)$ divides $\mathsf{LM}_\prec(h)$. By construction of G, there cannot be two polynomials $g_1, g_2 \in G$ such that $\mathsf{LM}_\prec(g_1)$ divides $\mathsf{LM}_\prec(g_2)$. Consequently, G is a minimal D-Gröbner basis. $\square$

Notice that the Macaulay matrices $\mathsf{Mac}_{\prec,d}(\mathbf{F})$ usually have a huge rank defect. Therefore, during the row echelon form computation, a lot of rows will become zero. This corresponds to useless computations. The F_5 criterion identifies some of those useless rows:

Theorem 1.62 (F_5 criterion)**.** *[Fau02] Let (i,t) be the signature of a row of $\mathsf{Mac}_{\prec,d}(\mathbf{F})$. If $t \in \mathsf{LM}_\prec(\langle f_1,\dots,f_{i-1}\rangle)$, then the row (i,t) is a linear combination of the rows on top of it.*

Proof. Since $t \in \mathsf{LM}_\prec(\langle f_1,\dots,f_{i-1}\rangle)$, there exist homogeneous polynomials $\{h^{(j)} = \sum_{\mathfrak{m}\in\mathsf{Monomials}(\mathbb{K}[X],d-\deg(f_j))} h^{(j)}_{\mathfrak{m}}\mathfrak{m}\}_{j\in\{1\dots i\}}$ such that $h^{(j)}_{\mathfrak{m}} \in \mathbb{K}$ and

$$h^{(i)} = \sum_{\ell=1}^{i-1} f_\ell h^{(\ell)}, \text{ and } \mathsf{LM}(h^{(i)}) = t.$$

Consequently,

$$\begin{array}{rcl} tf_i & = & h^{(i)} f_i - (h^{(i)} - t) f_i \\ & = & \left(\sum_{\ell=1}^{i-1} f_\ell(f_i h^{(\ell)})\right) - (h^{(i)} - t) f_i \end{array}$$

Notice that the polynomials $f_\ell(f_i h^{(\ell)})$ are linear combination of the rows with signature (ℓ, t') with $\ell < i$ and that the polynomial $(h^{(i)} - t)f_i$ is a linear combination of the rows (i, t') with $t' \prec t$. All these rows are on top of the row (i,t) in $\mathsf{Mac}_{\prec,d}(\mathbf{F})$. □

When the input of Lazard's algorithm is 0-dimensional, then the parameter D is not needed: as termination criterion, we can detect when all monomials of degree d are in $\langle \mathsf{LM}(G)\rangle$, ensuring that G is a Gröbner basis.

Actually, in the F_4 algorithm, even if the system is not 0-dimensional and the parameter D is not given, the termination is ensured since the matrices are constructed from critical pairs (they are submatrices of the Macaulay matrix). Therefore, when the set of critical pairs becomes empty, the algorithms returns the Gröbner basis.

The normal forms (Definition 1.27) can be computed as soon as a Gröbner basis of the ideal is known, as shown in Algorithm 3.

Algorithm 3 Normal form

Input: $\prec$ a monomial ordering;
  G a Gröbner basis of an ideal $I \subset \mathbb{K}[X]$ w.r.t. $\prec$;
  $f \in \mathbb{K}[X]$ a polynomial.
Output: $\mathrm{NF}_{\prec,I}(f)$.
1: $\widetilde{f} \leftarrow f$.
2: **while** there exists a monomial t in $\widetilde{f}$ and a polynomial $g \in G$ such that $\mathsf{LM}_\prec(g)$ divides t **do**
3: $\quad \widetilde{f} \leftarrow \widetilde{f} - \dfrac{t}{\mathsf{LM}_\prec(g)} g$
4: **end while**
5: Return $\widetilde{f}$.

Proposition 1.63. *Algorithm 3 terminates and is correct.*

Proof. **Termination.** During the execution of Algorithm 3, a reducible monomial $\widetilde{f}$ is replaced by smaller monomials each time the loop is entered. Since there is no infinitely decreasing sequence of monomials (Definition 1.18), Algorithm 3 terminates.

Correction. At the end of Algorithm 3, there is no monomial in $\widetilde{f}$ which is in $\langle \mathsf{LM}_\prec(G)\rangle = \mathsf{LM}_\prec(I)$. Therefore, by definition of the normal form, $\widetilde{f} = \mathrm{NF}_{\prec,I}(f)$. □

1.3.2 Matrix F_5 Algorithm

In this section, we give a description of a variant of the F_5 Algorithm [Fau02, FR09], called Matrix F_5 Algorithm, which is suitable for the complexity analysis (see [BFS04, BFSY04, Bar04]).

Given a set of generators $(f_1, \dots, f_m)$ of a homogeneous polynomial ideal $I \subset \mathbb{K}[X]$, an integer D and a monomial ordering $\prec$, the Matrix F_5 Algorithm computes a D-Gröbner basis of I with respect to $\prec$. It performs incrementally by considering the ideals $I_i = \langle f_1, \dots, f_i \rangle$ for $1 \le i \le m$.

As in [Fau02] and [BFSY04], we use a definition of the row echelon form of a matrix which is slightly different from the usual definition: we call *row echelon form* the matrix obtained by applying the Gaussian elimination Algorithm *without permuting the rows*. The idea of the Matrix F_5 Algorithm (see Algorithm 5 below) is to calculate triangular bases of the vector spaces $I_i \cap \mathbb{K}[X]_d$ for $1 \le d \le D$ and $1 \le i \le m$ and to deduce from them a d-basis of I_{i+1}. These triangular bases are obtained by computing row echelon forms of the Macaulay matrices.

When the row echelon form of a Macaulay matrix is computed, the rows which are linear combinations of preceding rows are reduced to zero. Such computations are useless: removing these rows before computing the row echelon form will not modify the result but lead to significant practical improvements. The F_5 criterion (see [Fau02]) is used to detect these *reductions to zero* and is given below in its algorithmic form (see Theorem 1.62 for the theoretical statement of the criterion). In Algorithm 5, the matrices $\mathcal{M}_{d,i}$ are similar to Macaulay matrices: their rows and their columns are sorted with the same orderings and their rows span the same vector spaces. Moreover, if $(f_1, \dots, f_m)$ is a regular sequence, then the rows of their row echelon form $\widetilde{\mathcal{M}}_{d,i}$ are *bases* of the vector spaces $I_i \cap \mathbb{K}[X]_d$.

We give in Algorithms 4 and 5 a description of F_5 criterion and of the Matrix F_5 Algorithm.

Algorithm 4 MatrixF_5criterion - returns a boolean

Input: $\begin{cases} (t, f_i) \text{ the signature of a row;} \\ \text{A matrix } \mathcal{M} \text{ in row echelon form.} \end{cases}$

Output: A boolean.

1: If t is the leading monomial of a row of $\mathcal{M}$, then return `true`,

2: else return `false`.

The rows eliminated by the F_5 criterion correspond to the trivial syzygies, i.e. the syzygies $(s_1, \dots, s_m)$ such that for all $i \in \{1, \dots, m\}$, $s_i \in \langle f_1, \dots, f_{i-1}, f_{i+1}, \dots, f_m \rangle$. These particular syzygies come from the commutativity of $\mathbb{K}[X]$ (for all $1 \le i, j \le m$, $f_i f_j - f_j f_i = 0$). We recall that in the generic case when $m \le n$ (regular sequences), the syzygy module of a polynomial system is generated by the trivial syzygies (see Proposition 1.47).

1.3.3 FGLM Algorithm

Another useful algorithm is the so-called *FGLM algorithm*. This algorithm does not compute a Gröbner basis from a polynomial, but takes as input a Gröbner basis of a 0-dimensional ideal with respect to some ordering, and outputs a Gröbner basis with respect to another ordering.

Proposition 1.64. *[FGLM93] Let $I \subset \mathbb{K}[X]$ be a 0-dimensional ideal, $\prec_1$ and $\prec_2$ be two monomial orderings, and G be a Gröbner basis of I w.r.t. $\prec_1$. The FGLM computes a Gröbner basis of I w.r.t. $\prec_2$ from G with complexity $O(n\,\mathsf{DEG}(I)^3)$, where $\mathsf{DEG}(I)$ is the degree of the ideal I, i.e. the dimension of $\mathbb{K}[X]/I$ as a $\mathbb{K}$-vector space.*

This complexity analysis comes from the first version of the algorithm [FGLM93]. Recently, the authors of [FM11, FGHR12] have shown sharper complexity bounds when some assumptions are

Algorithm 5 Matrix F_5 Algorithm [FR09, BFSY04, Fau02]

Input: $\begin{cases} (f_1, \dots, f_m) \text{ homogeneous polynomials of degree } d_1 \leq d_2 \leq \dots \leq d_m; \\ D \text{ an integer;} \\ \text{a monomial ordering } \prec . \end{cases}$

Output: G is a D-Gröbner basis of $\langle f_1, \dots, f_m \rangle$ for $\prec$.

$G \leftarrow \{f_1, \dots, f_m\}$
for d from d_1 to D **do**
  $\widetilde{\mathcal{M}}_{d,0} \leftarrow$ matrix with 0 rows
  for i from 1 to m **do**
    Construct $\mathcal{M}_{d,i}$ by adding to $\widetilde{\mathcal{M}}_{d,i-1}$ the following rows:
    if $d_i = d$ **then**
      add the row f_i with signature $(1, f_i)$
    end if
    if $d > d_i$ **then**
      for all f from $\widetilde{\mathcal{M}}_{d-1,i}$ with signature (e, f_i), such that x_λ is the
      greatest variable of e, add the $n - \lambda + 1$ rows $x_\lambda f, x_{\lambda+1} f, \dots, x_n f$ with the
      signatures $(x_\lambda e, f_i), (x_{\lambda+1} e, f_i), \dots, (x_n e, f_i)$ except those which satisfy:
      MatrixF_5criterion $((x_{\lambda+k} e, f_i), \widetilde{\mathcal{M}}_{d-d_i,i-1})$=true
    end if
    Compute $\widetilde{\mathcal{M}}_{d,i}$ the row echelon form of $\mathcal{M}_{d,i}$
    Add to G the polynomials corresponding to rows of $\widetilde{\mathcal{M}}_{d,i}$ such that their
    leading monomial is different from the leading monomial of
    the row with same signature in $\mathcal{M}_{d,i}$
  end for
end for
return G

satisfied (for instance the exponent 3 can be replaced by ω when the system is in shape position).

0-dimensional solving strategy. As we mentioned in Section 1.3, we need a lex Gröbner basis in order to solve 0-dimensional systems. However, the grevlex ordering is usually more efficient for computing a Gröbner basis with the F_4/F_5 Algorithms. Therefore, an efficient solving strategy is to compute first a grevlex Gröbner basis with the F_5 Algorithm, and then compute a lex Gröbner basis by using the FGLM Algorithm.

1.4 Bounds on the Degree and Degree of Regularity

This section is devoted to bounds on the degree and on the degree of regularity of polynomial systems, which will be used in the next section in order to obtain complexity estimates of the Gröbner bases computations.

1.4.1 Definitions

Crucial indicators of the complexity of Gröbner basis algorithms are the degree of the ideal (Definition-Proposition 1.4) and the so-called *index of regularity*, since in the 0-dimensional homogeneous case, it bounds the maximal degree in a minimal Gröbner basis.

Definition – Proposition 1.65. *Let $I \subset \mathbb{K}[X]$ be a homogeneous ideal. There exists a polynomial $\mathsf{HP}_{\mathbb{K}[X]/I}(t) \in \mathbb{Z}[t]$ of degree $\dim(I) - 1$ (with the convention that the null polynomial has degree -1) and an integer $d_0 \in \mathbb{N}$ such that, for all $d \geq d_0$,*

$$\mathsf{HF}_{\mathbb{K}[X]/I}(d) = \mathsf{HP}_{\mathbb{K}[X]/I}(d).$$

The polynomial $\mathsf{HP}_{\mathbb{K}[X]/I}$ is called the Hilbert polynomial *of $\mathbb{K}[X]/I$ and the smallest integer d_0 verifying this property is called the* index of regularity *and is denoted by $\mathsf{i}_{\mathsf{reg}}(I)$.*

Proof. By Proposition 1.42 and Proposition 1.43, the Hilbert series of $\mathbb{K}[X]/I$ is a rational function $\mathsf{HS}_{\mathbb{K}[X]/I}(t) = \frac{N(t)}{(1-t)^{\dim(I)}}$, where $N(t) \in \mathbb{Z}[t]$ and $N(1) \neq 0$. Partial fraction expansion yields

$$\mathsf{HS}_{\mathbb{K}[X]/I}(t) = Q(t) + \sum_{i=1}^{\dim(I)} \frac{a_i}{(1-t)^i},$$

where $Q(t) \in \mathbb{Z}[t]$, $a_1, \ldots, a_{\dim(I)} \in \mathbb{Z}$ and $a_{\dim(I)} \neq 0$. Notice that $[t^d]\frac{a_i}{(1-t)^i} = a_i\binom{i+d-1}{i-1}$, which is polynomial in d of degree $i-1$. Consequently, for all $d > \deg(Q(t))$, $\mathsf{HF}_{\mathbb{K}[X]/I}(d)$ is a polynomial function $\mathsf{HP}_{\mathbb{K}[X]/I}(d) = \sum_{i=1}^{\dim(I)} a_i\binom{i+d-1}{i-1}$ of degree $\dim(I) - 1$. Moreover, $\mathsf{HF}_{\mathbb{K}[X]/I}(\deg(Q(t))) \neq \mathsf{HP}_{\mathbb{K}[X]/I}(\deg(Q(t)))$. Therefore, $\mathsf{i}_{\mathsf{reg}}(I) = \deg(Q(t)) + 1$. $\square$

In the 0-dimensional case, the index of regularity can be easily read off from the Hilbert series (which is a polynomial):

Corollary 1.66. *If $I \subset \mathbb{K}[X]$ is a 0-dimensional homogeneous ideal, then $\mathsf{i}_{\mathsf{reg}}(I) = \deg(\mathsf{HS}_{\mathbb{K}[X]/I}) + 1$. Moreover, for any monomial ordering $\mathsf{i}_{\mathsf{reg}}(I)$ bounds the degree of all polynomial in a minimal homogeneous Gröbner basis of I.*

Proof. When $I \subset \mathbb{K}[X]$ is a homogeneous 0-dimensional ideal, then $\mathsf{HP}_{\mathbb{K}[X]/I}(d) = 0$ and thus $\mathsf{i_{reg}}(I)$ is the first null coefficient of $\mathsf{HS}_{\mathbb{K}[X]/I}(t)$. Consequently, $\mathsf{i_{reg}}(I) = \deg(\mathsf{HS}_{\mathbb{K}[X]/I}) + 1$. Moreover, $I_{\mathsf{i_{reg}}(I)} = \mathbb{K}[X]_{\mathsf{i_{reg}}(I)}$: all monomials of degree $\mathsf{i_{reg}}(I)$ are in I. By contradiction, assume that there is a homogeneous polynomial f of degree larger than $\mathsf{i_{reg}}(I)$ in a minimal Gröbner basis G of I. Then there exists a monomial of degree $\mathsf{i_{reg}}(I)$ that divides $\mathsf{LM}(g)$. Since this monomial is in I, by definition of a Gröbner basis there exists $g \in G$ such that $\mathsf{LM}(g)$ divides $\mathsf{LM}(f)$, and hence G is not minimal. □

In the case of homogeneous and quasi-homogeneous regular sequences, explicit formulas for the Hilbert series can be computed.

Theorem 1.67. *Let $\mathbf{w} \in \mathbb{N}^n$ be a weight vector and $\mathbf{F} = (f_1, \ldots, f_n) \in \mathbb{K}[X]^n$ be a family of quasi-homogeneous polynomials of respective weight degrees $(d_1, \ldots, d_n) \in \mathbb{N}^n$, generating a 0-dimensional ideal $I = \langle f_1, \ldots, f_n \rangle$. Then, the weighted Hilbert series of the ring $\mathbb{K}[X]/I$ is*

$$\mathsf{wHS}_{\mathbb{K}[X]/I}(t) = \frac{\prod_{j=1}^{n}(1 - t^{d_j})}{\prod_{j=1}^{n}(1 - t^{w_j})}.$$

Proof. By Theorem 1.48, if $\mathbf{F} = (f_1, \ldots, f_n)$ generates a 0-dimensional ideal in the ring $\mathbb{K}[x_1, \ldots, x_n]$, then $\mathbf{F}$ is a regular sequence: for all $i \in \{2, \ldots, n\}$, f_i does not divide 0 in the ring $\mathbb{K}[X]/\langle f_1, \ldots, f_{i-1} \rangle$. Consequently, by Proposition 1.41,

$$\mathsf{wHS}_{\mathbb{K}[X]/\langle f_1,\ldots,f_i \rangle}(t) = (1 - t^{d_i})\, \mathsf{wHS}_{\mathbb{K}[X]/\langle f_1,\ldots,f_{i-1} \rangle}(t).$$

Therefore, by induction on i, the following holds

$$\begin{aligned} \mathsf{wHS}_{\mathbb{K}[X]/I}(t) &= \textstyle\prod_{j=1}^{n}(1 - t^{d_j})\, \mathsf{wHS}_{\mathbb{K}[X]}(t) \\ &= \frac{\prod_{j=1}^{n}(1 - t^{d_j})}{\prod_{j=1}^{n}(1 - t^{w_j})}. \end{aligned}$$

□

Notice that Theorem 1.67 also gives an explicit formula for the Hilbert series of homogeneous regular sequences by considering the weight vector $\mathbf{w} = (1, \ldots, 1)$.

From the Hilbert series, one can compute the value of the degree of an ideal, its dimension and its index of regularity.

Corollary 1.68. *Let $\mathbf{F} = (f_1, \ldots, f_n) \in \mathbb{K}[X]^n$ be a family of homogeneous polynomials of respective degrees $(d_1, \ldots, d_n) \in \mathbb{N}^n$, generating a 0-dimensional ideal $I = \langle f_1, \ldots, f_n \rangle$. Then*

- *(Bézout bound) the* degree *of I is* $\mathsf{DEG}(I) = \displaystyle\prod_{j=1}^{n} d_j$;
- *(Macaulay bound) the* index of regularity *of I is* $\mathsf{i_{reg}}(I) = 1 + \displaystyle\sum_{j=1}^{n}(d_j - 1)$.

Proof. By Theorem 1.67, the Hilbert series of $\mathbb{K}[X]/I$ is

$$\mathsf{HS}_{\mathbb{K}[X]/I}(t) = \frac{\prod_{j=1}^{n}(1 - t^{d_j})}{(1 - t)^n},$$

which is a polynomial, and thus

$$\begin{array}{rcl}\mathsf{DEG}(I) & = & \dim_{\mathbb{K}}(\mathbb{K}[X]/I)\\ & = & \mathsf{HS}_{\mathbb{K}[X]/I}(1)\\ & = & \prod_{j=1}^{n} d_j;\\ \mathsf{i}_{\mathsf{reg}}(I) & = & 1+\deg(\mathsf{HS}_{\mathbb{K}[X]/I}(t))\\ & = & 1+\sum_{j=1}^{n}(d_j-1).\end{array}$$

□

A bound similar to the Bézout bound exist for counting the number of isolated solutions of multi-homogeneous systems:

Theorem 1.69 (Multi-homogeneous Bézout number)**.** *[MS87] Let* $\mathbf{F} = (f_1, \dots, f_n) \in \mathbb{K}[X^{(1)}, \dots, X^{(\ell)}]^n$ *be a system (non-homogeneous) of multi-degrees* $\mathrm{mdeg}(f_i) = (d_{i,1}, \dots, d_{i,\ell})$*. Then the number of isolated zeroes of* $\mathbf{F}$ *is bounded above by the coefficient of* $\alpha_1^{|X^{(1)}|} \dots \alpha_\ell^{|X^{(\ell)}|}$ *in the polynomial*

$$(d_{1,1}\alpha_1 + \dots + d_{1,\ell}\alpha_\ell) \dots (d_{n,1}\alpha_1 + \dots + d_{n,\ell}\alpha_\ell).$$

1.4.2 Affine 0-dimensional systems – Degree of regularity

For affine polynomial systems, [Bar04] provides a generalization of the notion of index of regularity by considering the homogeneous components of highest degrees of the system:

Definition – Proposition 1.70. *Let* $\mathbf{F} = (f_1, \dots, f_m) \in \mathbb{K}[X]^m$ *be a polynomial system (non necessarily homogeneous). Let* $\mathbf{F}^{(h)} = (f_1^{(h)}, \dots, f_m^{(h)})$ *be the homogeneous components of highest degree of* $\mathbf{F}$*. If* $\dim(\langle \mathbf{F}^{(h)} \rangle) = 0$*, then* $\dim(\mathbf{F}) = 0$ *and we call* degree of regularity *of* $\mathbf{F}$ *(denoted by* $\mathrm{d_{reg}}(\mathbf{F})$*) the index of regularity of* $\langle \mathbf{F}^{(h)} \rangle$*.*

Notice that if $\mathbf{F}$ is homogeneous and 0-dimensional, then the notions of degree of regularity and index of regularity coincide: $\mathrm{d_{reg}}(\mathbf{F}) = \mathsf{i}_{\mathsf{reg}}(\langle \mathbf{F} \rangle)$ since $\mathbf{F}^{(h)} = \mathbf{F}$.

However, in the non-homogeneous case, the degree of regularity is not an invariant of the ideal: two families of polynomials generating the same ideal do not necessarily share the same degree of regularity (for instance $\langle x \rangle = \langle x^2 + x, x^2 \rangle$, but $\mathrm{d_{reg}}(x) = 1$ and $\mathrm{d_{reg}}(x^2 + x, x^2) = \mathsf{i}_{\mathsf{reg}}(\langle x^2 \rangle) = 2$).

By slight abuse of notation, in the next chapters, we will sometimes use $\mathrm{d_{reg}}$ for ideals when there is no possible confusion on the polynomial family generating it.

From the algorithmic viewpoint, Lazard's algorithm can be used in the affine context too, but using it directly has the drawback that we do not take profit of *degree falls*. Indeed, when dealing with affine systems, reductions of polynomials of degree d can give rise to polynomials of lower degrees. In that case, in order to speed-up the algorithm, it is efficient to restart the computations at a lower degree. This is handled in a general way by the so-called *normal strategy* in the F_4 algorithm [Fau99]: when a new polynomial is found during a reduction step and his degree is lower than the degree of the matrix, the algorithm F_4 continues by constructing matrices in lower degree in order to use this new information. Formally speaking, the *normal strategy* consists in reducing at each step the critical pairs with the smallest degree before proceeding to the next step.

In order to study the behavior of Gröbner basis algorithms when the input system is affine and when the *normal strategy* is used, we introduce the following notation.

	homogenous/affine systems	dimension	depends on the monomial ordering	
$\mathsf{i}_{\mathrm{reg}}$	homogeneous	any	no	Definition 1.65
$\mathrm{d}_{\mathrm{reg}}$	both	0	no	Definition 1.70
$\mathrm{d}_{\max\prec}$	both	any	yes	Page 50
$\mathrm{d}_{\mathrm{wit}}$	both	any	yes	Definition 7.4
degree max in a minimal Gröbner basis	homogeneous	any	yes	Section 4.6.1

Table 1.1: Different notions of regularity

Definition 1.71. *Let $\mathscr{V}$ denote the set of $\mathbb{K}$-vector subspaces of finite dimension of $\mathbb{K}[X]$. We let χ denote the application*

$$\begin{array}{rccl} \chi: & \mathbb{N}\times\mathscr{V} & \to & \mathscr{V} \\ & (d,V) & \mapsto & \{\textit{finite sums} \displaystyle\sum_{\substack{h\in\mathbb{K}[X]\\ f\in V\\ \deg(h)+\deg(f)\leq d}} hf\} \end{array}$$

Therefore, for several d, the F_4 algorithm computes bases of the successive vector spaces $S_i = \chi(d, S_{i-1})$ until a complete Gröbner basis is obtained.

We define $\mathrm{d}_{\max\prec}(\mathbf{F})$ as the highest degree reached during the computation of a Gröbner basis with the F_4 algorithm:

$$\begin{array}{rcl} \mathrm{d}_{\max\prec}(\mathbf{F}) & = & \min_{d\in\mathbb{N}}\{d \mid \exists \ell\in\mathbb{N}, \exists V_0,\dots,V_\ell\in\mathscr{V} s.t. \\ & & V_0=\mathbf{F},\ \text{for all } i, V_i=\chi(d,V_{i-1}) \text{ and} \\ & & V_\ell \text{ contains a Gröbner basis of } \langle\mathbf{F}\rangle \text{ with respect to } \prec\} \end{array}$$

This notion will be useful in Section 6.5.5 for estimating the complexity of computing Gröbner bases of affine bilinear systems.

1.4.3 Relations between notions of regularity

In this thesis, several notions of regularity are used. There are slight differences between them but they are all related with the highest degree occurring during the computation of a Gröbner basis. Consequently, they main role is to bound the complexity of Gröbner bases algorithms. All these notions are reported in Table 1.1. It is worth noticing that there exist other notions of regularity in the literature (e.g. the Castelnuovo-Mumford regularity [Eis95, Section 20.5]).

There exist relations between these degrees. First, recall that the index of regularity and the degree of regularity coincide for 0-dimensional homogeneous systems.

Let $\mathbf{F} = (f_1,\dots,f_m) \in \mathbb{K}[x_1,\dots,x_n]^m$ be an affine system, $\mathbf{F}^{(h)} = (f_1^{(h)},\dots,f_m^{(h)}) \in \mathbb{K}[x_1,\dots,x_n]^m$ be the system of its homogeneous components of highest degrees, and $\widetilde{\mathbf{F}} = (\widetilde{f_1},\dots,\widetilde{f_m}) \in \mathbb{K}[x_1,\dots,x_n,h]^m$ be its homogenized system (i.e. $\widetilde{f_i}(x_1,\dots,x_n,h) = h^{\deg(f_i)}f_i(x_1/h,\dots,x_n/h)$). Also, let G (resp. $G^{(h)}$, $\widetilde{G}$) be a minimal grevlex Gröbner basis of $\langle\mathbf{F}\rangle$ (resp. $\langle\mathbf{F}^{(h)}\rangle$, $\left\langle\widetilde{\mathbf{F}}\right\rangle$). Then the following equalities between the maximal degrees in these Gröbner bases hold:

$$\max(\deg(G)) \leq \max(\deg(G^{(h)})) \leq \max(\deg(\widetilde{G})).$$

These inequalities are a consequence of the following known fact: the specialization of h in a homogeneous grevlex Gröbner basis yields a grevlex Gröbner basis of the corresponding specialized ideal. Notice that $\mathbf{F}$ and $\mathbf{F}^{(h)}$ are respectively the specializations of $\widetilde{\mathbf{F}}$ at $h = 1$ and at $h = 0$. Moreover, the variable h divides a polynomial if and only if it divides its leading monomial with respect to the grevlex ordering. Therefore, the polynomials in $\widetilde{G}$ that are divisible by h become 0 with the specialization at $h = 0$, and they lead to polynomials of smaller degree when they are specialized at $h = 1$.

Another consequence of this specialization property is the inequality $\mathrm{d}_{\max \prec_{\mathsf{grevlex}}}(\mathbf{F}) \leq \max(\deg(\widetilde{G}))$. Indeed, let $g \in G$ be a polynomial in a minimal Gröbner basis of $\langle \mathbf{F} \rangle$. Then there exists a polynomial $\widetilde{g}$ in a minimal grevlex Gröbner basis of $\widetilde{\mathbf{F}}$ whose specialization at $h = 1$ is g. Hence there exist homogeneous polynomials $\widetilde{h_1}, \dots, \widetilde{h_m}$ such that $\widetilde{g} = \sum_{i=1}^{m} \widetilde{f_i}\widetilde{h_i}$. By dehomogenizing this relation, we see that g belongs to $\chi(\mathbf{F}, \max(\deg(\widetilde{G}))$, and hence the inequality $\mathrm{d}_{\max \prec_{\mathsf{grevlex}}}(\mathbf{F}) \leq \max(\deg(\widetilde{G}))$ holds.

The definition and the properties of the witness degree $\mathrm{d}_{\mathrm{wit}}$ (which is related to the degree of regularity of the homogenized system) are postponed to Chapter 7.

1.5 Complexity

1.5.1 Complexity model

In this thesis, unless otherwise said, we measure the *arithmetic complexity* of algorithms, i.e. the number of arithmetic operations $+, -, \times, \div$ in the base field $\mathbb{K}$. We also use the Landau notations:

- if $f : \mathbb{R}_+ \to \mathbb{R}_+$ is a positive function, we let $O(f)$ denote the class of functions $g : \mathbb{R}_+ \to \mathbb{R}_+$ such that there exists two positive numbers C, x_0 such that for all $x \geq x_0$, $g(x) \leq Cf(x)$. By abuse of notation, we write $g(x) = O(f(x))$ or $g(x) \leq O(f(x))$ when $g \in O(f)$;

- we let $\widetilde{O}(f)$ denote the class of functions $g : \mathbb{R}_+ \to \mathbb{R}_+$ such that there exists k such that $g \in O(f(x) \log^k(f(x)))$;

- we write $g \in \Omega(f)$ when $f \in O(g)$;

- by extension, we also use these notations for functions with several variables: if $f : \mathbb{R}_+^\ell \to \mathbb{R}_+$ is a positive function, we let $O(f)$ denote the class of functions $g : \mathbb{R}_+^\ell \to \mathbb{R}_+$ such that there exist two positive numbers C, A such that for all $\mathbf{x}$ with $x_i \geq A$ for all i, $g(x) \leq Cf(x)$. Sometimes, we explicitly fix some parameters. For instance, if m is fixed, $O(n^m)$ represents the class of functions of one variable such that $g(n) \leq Cn^m$ for n large enough. On the other hand, $O(n^m)$ with variables n, m represents the class of functions of two variables such that $g(n, m) \leq Cn^m$ for n and m large enough.

Also, in the whole thesis, the notation ω stands the exponent of the matrix multiplication, i.e. ω is the smallest positive number such that the product of two $N \times N$ matrices can be achieved in $O(N^\omega)$ arithmetic operations. Classical bounds for ω are:

- $\omega \leq 3$: schoolbook matrix multiplication;

- $\omega \leq 2.807$: Strassen's algorithm [Str69];

- $\omega \leq 2.376$: Coppersmith-Winograd's algorithm [CW90].

Recent improvements by [Sto10, Vas11] have decreased it to $\omega \leq 2.373$.

1.5.2 Complexity of Gröbner basis algorithms

Homogeneous systems

Homogeneous systems are usually easier to study since there are no degree falls during the execution of the F_4/F_5 algorithm. The following theorem bounds the complexity of the Lazard's algorithm by the cost of linear algebra on the Macaulay matrices. However, the following bound is general and not very precise: it does not take into account the rows eliminated by the F_5 criterion nor the structure of the Macaulay matrices. In the particular case of regular sequences, better complexity bounds are obtained in [Bar04] by performing a step by step analysis of the F_5 algorithm.

Theorem 1.72. *[Bar04, BFS04, BFSY04] Let $\mathbf{F} = (f_1, \dots, f_m) \in \mathbb{K}[X]^m$ be a family of homogeneous polynomials generating a 0-dimensional ideal. The complexity of computing a Gröbner basis (for any monomial ordering) of the ideal $\langle \mathbf{F} \rangle$ is bounded by*

$$O\left(\sum_{i=0}^{\mathrm{d_{reg}}(\mathbf{F})}\left[\binom{n+i-1}{i}\left(\sum_{j=1}^{m}\binom{n+i-\deg(f_j)-1}{i-\deg(f_j)}\right)\left(\binom{n+i-1}{i}-\mathsf{HF}_{\mathbb{K}[X]/\langle\mathbf{F}\rangle}(i)\right)^{\omega-2}\right]\right)$$
$$\leq\ O\left(m\binom{n+\mathrm{d_{reg}}(\mathbf{F})}{\mathrm{d_{reg}}(\mathbf{F})}^{\omega}\right).$$

Proof. In Lazard's Algorithm for 0-dimensional systems, the number of arithmetic operations corresponds to the cost of linear algebra. The algorithm stops when $d = \mathrm{d_{reg}}(\mathbf{F})$. When $d = i$, the number of rows, number of columns and rank of the Macaulay matrix $\mathsf{Mac}_{\prec,i}(\mathbf{F})$ are respectively

$$\begin{aligned}
\mathsf{nbrows} &= \sum_{j=1}^{m}\binom{n+i-\deg(f_j)-1}{i-\deg(f_j)};\\
\mathsf{nbcols} &= \binom{n+i-1}{i};\\
\mathsf{rank} &= \binom{n+i-1}{i} - \mathsf{HF}_{\mathbb{K}[X]/\langle\mathbf{F}\rangle}(i).
\end{aligned}$$

By [Sto00], the complexity of computing the row echelon form of a $\mathsf{nbrows} \times \mathsf{nbcols}$ is bounded by

$$O(\mathsf{nbrows}\cdot\mathsf{nbcols}\cdot\mathsf{rank}^{\omega-2}),$$

whence the complexity of Algorithm 2 up to the degree $\mathrm{d_{reg}}(\mathbf{F})$ is bounded by

$$O\left(\sum_{i=0}^{\mathrm{d_{reg}}(\mathbf{F})}\left[\binom{n+i-1}{i}\left(\sum_{j=1}^{m}\binom{n+i-\deg(f_j)-1}{i-\deg(f_j)}\right)\left(\binom{n+i-1}{i}-\mathsf{HF}_{\mathbb{K}[X]/\langle\mathbf{F}\rangle}(i)\right)^{\omega-2}\right]\right).$$

This is also bounded above by

$$O\left(\sum_{i=0}^{\mathrm{d_{reg}}(\mathbf{F})} m\binom{n+i-1}{i}^{\omega}\right).$$

Since $\omega > 1$, we obtain

$$\begin{aligned}
m\sum_{i=0}^{\mathrm{d_{reg}}(\mathbf{F})}\binom{n+i-1}{i}^{\omega} &\leq m\left(\sum_{i=0}^{\mathrm{d_{reg}}(\mathbf{F})}\binom{n+i-1}{i}\right)^{\omega}\\
&\leq m\binom{n+\mathrm{d_{reg}}(\mathbf{F})}{\mathrm{d_{reg}}(\mathbf{F})}^{\omega}.
\end{aligned}$$

□

1.5.3 Complexity of solving affine systems

For affine systems, it is more difficult to obtain complexity bounds because of the degree falls. However, when the homogeneous part of highest degree is 0-dimensional, it is possible to obtain similar bounds as in the homogeneous case. The following theorem is from a personal communication with J.-C. Faugère and it is a work in progress by M. Bardet, J.-C. Faugère and B. Salvy.

Theorem 1.73. *Let* $\mathbf{F} = (f_1, \dots, f_m) \in \mathbb{K}[X]^m$ *be a polynomial family and let* $\mathbf{F}^{(h)}$ *denote the family of homogeneous components of highest degree. If* $\langle \mathbf{F}^{(h)} \rangle$ *is* 0*-dimensional, then the complexity of computing a Gröbner basis of* $\mathbf{F}$ *(for any* graded *monomial ordering) is bounded by*

$$O\left(m \binom{n + \mathrm{d_{reg}}(\mathbf{F}^{(h)})}{\mathrm{d_{reg}}(\mathbf{F}^{(h)})}^{\omega} + n \,\mathsf{DEG}\left(\left\langle \mathbf{F}^{(h)} \right\rangle \right)^3 \right).$$

1.5.4 Relation between the complexity and the degree of the ideal

If I is a 0-dimensional ideal, the arithmetic size of a Gröbner basis of I (i.e. the number of coefficients in $\mathbb{K}$) is closely related to $\mathsf{DEG}(I)$, especially if the system generating the ideal is in *shape position*. Experimentally, Gröbner bases algorithms seem to be output-sensitive: their practical running time often depends on the degree of the ideal and on the size of the output. However, from a theoretical viewpoint, there are few families of systems for which it is proved that the complexity is related to $\mathsf{DEG}(I)$.

Proposition 1.74. *Let* $\mathbf{F} \in \mathbb{K}[X]^m$ *be a* 0*-dimensional ideal in shape position, and* G *be the reduced Gröbner basis of* $\langle \mathbf{F} \rangle$ *with respect to* $\prec_{\mathsf{lex}}$*. Then the number of monomials in* G *is bounded above by* $n \,\mathsf{DEG}(\langle \mathbf{F} \rangle)$.

Proof. The shape position states that

$$G = \left\{ \begin{array}{c} x_1 - h_1(x_n) \\ \vdots \\ x_{n-1} - h_{n-1}(x_n) \\ h_n(x_n) \end{array} \right\},$$

where $\deg(h_n) = \mathsf{DEG}(I)$ and for all $i \in \{1, \dots, n-1\}$, $\deg(h_i) < \mathsf{DEG}(I)$. Therefore the number of monomials in G is bounded by $n \,\mathsf{DEG}(I)$. □

Even when a 0-dimensional system is not in shape position, the size of the reduced Gröbner basis (for any monomial ordering) is still polynomial in the degree of the ideal:

Proposition 1.75. *[FGLM93] Let* $\mathbf{F} \in \mathbb{K}[X]^m$ *be a* 0*-dimensional system, and* G *be the reduced Gröbner basis of* $\langle \mathbf{F} \rangle$ *with respect to a monomial ordering* $\prec$*. Then the number of monomials in* G *is bounded above by* $n \,\mathsf{DEG}(\langle \mathbf{F} \rangle)(1 + \mathsf{DEG}(\langle \mathbf{F} \rangle))$.

Proof. In [FGLM93, Corollary 2.1], it is proven that the number of polynomials in a reduced Gröbner basis is bounded by $n \,\mathsf{DEG}(\langle \mathbf{F} \rangle)$. Each polynomial $g \in G$ has the form $g = \mathsf{LM}_{\prec}(g) + \sum_{\substack{\mathfrak{m} \text{ monomial} \\ \mathfrak{m} \notin \mathsf{LM}(\langle \mathbf{F} \rangle)}} a_{\mathfrak{m}} \mathfrak{m}$. Consequently, the number of monomials in g is bounded by $1 + \mathsf{DEG}(\langle \mathbf{F} \rangle)$. □

Therefore, it is interesting to identify families of systems for which the complexity is polynomial in the degree of the corresponding ideal. As far as we know, there are few such bounds for Gröbner bases algorithms.

One family of systems for which such a bound is reached are regular sequences of polynomials of the same degrees.

Proposition 1.76. *Let $\mathbf{F} = (f_1, \dots, f_n) \in \mathbb{K}[X]^n$ be a homogeneous regular sequence of polynomials of degree d. Then, for a fixed d, as $n \to \infty$, the complexity of Algorithm 2 up to the degree of regularity is bounded by*

$$\widetilde{O}\left(2^{n\omega(d\log_2(d)-(d-1)\log_2(d-1))}\right),$$

which is polynomial in $\mathsf{DEG}(\langle \mathbf{F}\rangle) = d^n$.

Proof. By Theorem 1.72, the complexity of Algorithm 2 is bounded by $O\left(n\binom{n+\mathsf{i}_{\mathsf{reg}}(\langle \mathbf{F}\rangle)}{\mathsf{i}_{\mathsf{reg}}(\langle \mathbf{F}\rangle)}^{\omega}\right)$. For regular sequences of degree d, the index of regularity is $\mathsf{i}_{\mathsf{reg}}(\langle \mathbf{F}\rangle) = (d-1)n+1$ (Corollary 1.68). Consequently,

$$\begin{aligned} n\binom{n+\mathsf{i}_{\mathsf{reg}}(\langle \mathbf{F}\rangle)}{\mathsf{i}_{\mathsf{reg}}(\langle \mathbf{F}\rangle)}^{\omega} &= n\binom{dn+1}{n}^{\omega} \\ &= \left(\frac{n(dn+1)}{(d-1)n+1}\binom{dn}{n}\right)^{\omega} \end{aligned}$$

By Stirling's formula, as n grows, we obtain

$$\binom{dn}{n} \underset{n\to\infty}{=} O\left(2^{n(d\log_2(d)-(d-1)\log_2(d-1))}\right).$$

Therefore the complexity of Algorithm 2 is bounded by

$$\widetilde{O}\left(2^{n\omega(d\log_2(d)-(d-1)\log_2(d-1))}\right).$$

This is polynomial in d^n since

$$2^{n\omega(d\log_2(d)-(d-1)\log_2(d-1))} = (d^n)^{\omega(d-(d-1)\log_2(d-1)/\log_2(d))}.$$

$\square$

We would like to point out that the complexity bound

$$O\left(m\binom{n+\mathrm{d}_{\mathrm{reg}}(\mathbf{F}^{(h)})}{\mathrm{d}_{\mathrm{reg}}(\mathbf{F}^{(h)})}^{\omega} + n\,\mathsf{DEG}\left(\left\langle \mathbf{F}^{(h)}\right\rangle\right)^3\right)$$

is not uniformly polynomial in the Bézout bound for regular sequences. For instance, consider the following family of sequences of degrees:

$$\mathbf{d}_i = (2^i, 2, 2, \dots, 2) \in \mathbb{N}^i.$$

Now let $\mathbf{F}_i \in \mathbb{K}[x_1, \dots, x_i]^i$ be a homogeneous regular sequence of degrees $\mathbf{d}_i \in \mathbb{N}^i$. The Bézout bound yields $\mathsf{DEG}(\langle \mathbf{F}_i\rangle) = 2^{2i-1}$ and the Macaulay bound gives $\mathrm{d}_{\mathrm{reg}}(\mathbf{F}_i) = 2^i + i$. Consequently $\binom{i+\mathrm{d}_{\mathrm{reg}}(\mathbf{F}^{(h)})}{\mathrm{d}_{\mathrm{reg}}(\mathbf{F}^{(h)})} = \binom{2^i+2i}{i} \geq \frac{2^{i^2}}{i^i} = 2^{i^2 - i\log_2 i}$ which is not polynomial in the Bézout bound 2^{2i-1}.

However, the complexity bound for Gröbner basis algorithms is only an upper bound and hence deciding if the complexity of Gröbner basis algorithms is uniformly polynomial in the Bézout bound for 0-dimensional regular sequences is still an open problem.

Chapter 2

Algebraic Systems in Applications

This thesis deals with polynomial system solving of "structured systems", i.e. systems whose structural properties can be exploited in order to obtain sharp complexity bounds or dedicated algorithms. This structure sometimes comes from the *shape* of the equations (multi-homogeneous, determinantal systems). It can also stem from the set of solutions that we are investigating (e.g. finding one *boolean* solution of a boolean system), even if the system itself is not structured.

This is motivated by applications in Engineering sciences. We focus here on systems coming from Cryptology, Coding Theory, Geometry and Optimization. We present in this chapter where these algebraic systems come from and what the current state of the art is.

We first describe cryptosystems whose security is directly related to the difficulty of solving structured systems. In particular, attacks on the MinRank authentication scheme and on the HFE cryptosystem can be modeled by a rank condition on a polynomial matrix. Finding points where this condition holds is the so-called *MinRank problem* and can be modeled by a multi-homogeneous and by a determinantal system. These kinds of systems also appear during the analysis of *rank-metric codes*, which are a special kind of linear codes where the distance between words is not the usual Hamming distance. We also describe a multi-homogeneous modeling of the McEliece cryptosystem, and we give a short description of the QUAD streamcipher, whose security relies on the difficulty of solving boolean quadratic systems.

Then we describe some fundamental problems in Real Geometry and Optimization: polynomial programs, quantifier elimination, roadmap computations, and computing at least one point by connected component in a real semi-algebraic set. Recent algorithms for solving these problems need to compute Gröbner bases of determinantal or multi-homogeneous systems as a central subroutine.

2.1 MinRank

2.1.1 Description of the MinRank problem

The MinRank problem is a classical problem from linear algebra which appears in several applications (Cryptology, Information theory, Optimization, Geometry,...), and which is related to determinantal and to multi-homogeneous systems.

MinRank. Given three integers $r, p, q \in \mathbb{N}$ such that $r \leq q \leq p$, and a family of $p \times q$ matrices $M_0, \ldots, M_n \in \mathbb{K}^{p\times q}$, find $\lambda_1, \ldots, \lambda_n \in \mathbb{K}$ (or in $\overline{\mathbb{K}}$, depending on the context) such that

$$\mathsf{Rank}(M_0 - \sum_{i=1}^{n} \lambda_i M_i) \leq r.$$

This problem is NP-hard as soon as $\mathbb{K}$ is a finite field [BFS99]. Random instances are also difficult to solve, and consequently this problem has been used to design cryptosystems whose security relies on its difficulty. The MinRank problem can also be seen as a multivariate generalization of the classical Eigenvalue problem. Indeed, if $n = 1$, $p = q$ and $r = q - 1$, then the MinRank problem can be reduced to the problem of finding the eigenvalues of a square matrix.

In Chapter 4, we study a generalization of the MinRank problem, where the dependency on the variables can be polynomial (and not necessarily linear as in the classical MinRank problem).

Generalized MinRank Problem: given a field $\mathbb{K}$, a $p\times q$ matrix $\mathscr{M}$ whose entries are polynomials of degree D over $\mathbb{K}[x_1,\dots,x_n]$, and $r < \min(p,q)$ an integer, compute the set of points at which the evaluation of the matrix has rank at most r.

2.1.2 Algebraic techniques for solving the MinRank problem

Minors modeling. The direct and most straightforward way to represent the MinRank as a system of polynomial equations is to consider the set of all minors of size $r+1$ of the matrix $M_0 - \sum_{i=1}^{n}\lambda_i M_i$. Indeed, those minors simultaneously vanish exactly at the solutions of the MinRank problem. However, the drawback of this modeling is the size of the polynomial system. For instance, if $n = 9$, $r = 9$, $p = q = 12$, there are 220 minors of size 10, and each one is a dense polynomial of degree 10 in 9 variables: each one is the sum of $\binom{19}{10} = 92378$ monomials.

Kipnis-Shamir modeling. The Kipnis-Shamir modeling was introduced in [KS99] and yields a way to represent MinRank problems as systems of bilinear equations. Roughly speaking, the idea to represent the locus of rank defect of the matrix M is to introduce fresh variables representing the kernel of the matrix. This is in the same spirit as Lagrange multipliers in optimization. It is done by looking for a triangular basis of the right kernel of the matrix $M_0 - \sum_{i=1}^{n}\lambda_i M_i$. Indeed, this matrix has rank at most r if and only if its right kernel has dimension at least $q - r$. We assume moreover that this right kernel is in systematic form, i.e. that the projection on the last coordinates

$$\begin{array}{rcl}\mathsf{Ker}_R(M_0 - \sum_{i=1}^{n}\lambda_i M_i) & \to & \mathbb{K}^{q-r}\\ (y_1,\dots,y_q) & \mapsto & (y_{r+1},\dots,y_q)\end{array}$$

is injective. This condition can be easily verified if the cardinality of $\mathbb{K}$ is large enough by performing first a random invertible linear change of coordinates on the variables λ_i. Then we introduce $(q-r)r$ new variables $y_1^{(1)},\dots,y_r^{(q-r)}$, and we look for solutions of the bilinear system obtained by considering the matrix relation

$$\left(M_0 - \sum_{i=1}^{n}\lambda_i M_i\right)\cdot\begin{bmatrix} y_1^{(1)} & y_1^{(2)} & \dots & y_1^{(q-r)}\\ \vdots & \vdots & \vdots & \vdots\\ y_r^{(1)} & y_r^{(2)} & \dots & y_r^{(q-r)}\\ 1 & 0 & \dots & 0\\ \vdots & \vdots & \vdots & \vdots\\ 0 & 0 & \dots & 1\end{bmatrix} = 0.$$

This yields an algebraic system of $p(q-r)$ bilinear equations in $n+r(q-r)$ variables. Notice that this system has as many equations as variables if $n = (p-r)(q-r)$. In fact, in that case, the system is 0-dimensional if genericity assumptions on the matrices M_i are verified (see Chapter 4). The main advantage of this representation is the size of the system. For $n = 9$, $r = 9$, $p = q = 12$, it is a system of 108 equations. Moreover, each of these equations is represented by only 100 monomials. Consequently, the system is much smaller than the system obtained by the minors modeling.

Algebraic properties of this bilinear system were investigated in [FLP08]. In this paper, the authors show that Gröbner basis algorithms have a specific behavior on these systems. Also, Challenges A and B from Table 2.1 were solved by using this modeling and Gröbner bases algorithms.

2.2 Cryptology and Information Theory

2.2.1 Courtois Authentication Scheme

In [Cou01], the author proposes a zero-knowledge authentication scheme, and proves that its security can be reduced to the difficulty of solving the MinRank problem. We give here a short description of this cryptosystem. For simplicity of notations, we restrict ourselves to the case where the matrices are square, but this can be generalized without any major modification (see [Cou01] for more details).

We recall that in a zero-knowledge authentication scheme, the prover knows a secret key (which proves his identity), and a protocol allows him to convince any verifier with overwhelming probability that he knows this secret without revealing any information about it.

Public key. A integer r and a set of $p \times p$ matrices $M_0, \ldots, M_n \in \mathsf{GF}_q^{p\times p}$ such that there exists $(x_1, \ldots, x_n) \in \mathsf{GF}_q^n$ satisfying $\mathsf{Rank}\left(M_0 - \sum_{i=1}^n x_i M_i\right) \leq r$.

Secret key. A vector $(\alpha_1, \ldots, \alpha_n) \in \mathsf{GF}_q^n$ such that $\mathsf{Rank}\left(M_0 - \sum_{i=1}^n \alpha_i M_i\right) = r$. We denote by M the matrix $M = M_0 - \sum_{i=1}^n \alpha_i M_i$.

For this zero-knowledge authentication scheme, we need a collision-resistant hash function H, i.e. a function which takes as input a finite sequence of elements in GF_q (the set of all finite sequences is denoted by $\mathsf{GF}_q^{(\mathbb{N})}$) and returns an element in a finite set S

$$H : \mathsf{GF}_q^{(\mathbb{N})} \longrightarrow S.$$

Collision-resistance means that it should be computationally infeasible to find a collision, i.e. two finite sequences $a, b \in \mathsf{GF}_q^{(\mathbb{N})}$ with $a \neq b$ and $H(a) = H(b)$. Usual hash-functions such as SHA-2 have this property (no such collisions have been found so far).

In the sequel, the owner of the secret key is called *the prover*, and the one who wants to verify its identity is called *the verifier*.

One round of authentication:

1. The prover chooses two $p \times p$ random invertible matrices $S, T \in \mathsf{GF}_q^{p\times p}$, and a random matrix $X \in \mathsf{GF}_q^{p\times p}$.

2. The prover chooses $\beta_1 = (\beta_{1,1}, \ldots, \beta_{1,n}) \in \mathsf{GF}_q^n$ at random. Let $\beta_2 = \beta_1 + \alpha \in \mathsf{GF}_q^n$, $N_1 = \sum_{i=1}^p \beta_{1,i} M_i \in \mathsf{GF}_q^{p\times p}$, and $N_2 = \sum_{i=1}^p \beta_{2,i} M_i \in \mathsf{GF}_q^{p\times p}$ (and hence $N_2 - N_1 = M_0 - M$).

3. The prover sends to the verifier
$$H(S \mid T \mid X), \quad H(T \cdot N_1 \cdot S + X), \quad H(T \cdot N_2 \cdot S + X - T \cdot M_0 \cdot S),$$
where $S \mid T \mid X$ denotes the concatenation of S, T and X.

4. The verifier chooses $Q \in \{0, 1, 2\}$ and sends it to the chooser.

5. If $Q = 0$, the prover reveals $(T \cdot N_1 \cdot S + X)$ and $(T \cdot N_2 \cdot S + X - T \cdot M_0 \cdot S)$. The verifier then checks that $H(T \cdot N_1 \cdot S + X)$ and $H(T \cdot N_2 \cdot S + X - T \cdot M_0 \cdot S)$ are correct, then he computes $(T \cdot N_1 \cdot S + X) - (T \cdot N_2 \cdot S + X - T \cdot M_0 \cdot S) = T \cdot M \cdot S$ and checks that it is indeed of rank r.

Parameter set	n	p	r	$\mathbb{K}$	Security bound
A	10	6	3	GF_{65521}	2^{106}
B	10	7	4	GF_{65521}	2^{122}
C	10	11	8	GF_{65521}	2^{138}
D	81	19	10	GF_2	2^{64}
E	121	21	10	GF_2	2^{81}
F	190	29	15	GF_2	2^{128}

Table 2.1: Courtois MinRank challenges

6. If $Q = 1$ or $Q = 2$, the prover reveals S, T, X and β_Q. The verifier checks that S, T are invertible and that $H(S \mid T \mid X)$ is correct. Then he computes $T \cdot N_Q \cdot S = \sum_{i=1}^{p} \beta_{Q,i} M_i$ and verifies that $H(T \cdot N_1 \cdot S + X)$ (if $Q = 1$) or $H(T \cdot N_2 \cdot S + X - T \cdot M_0 \cdot S)$ (if $Q = 2$) is correct.

In [Cou01], Courtois shows that any cheater (i.e. someone who does not know α) can be detected by the verifier with probability at least $1/3$. Therefore, if someone succeeds ℓ rounds of authentication, then the verifier knows that this person knows α with probability at least $1 - (2/3)^{\ell}$. More precisely, he shows that, assuming that H is collision-resistant, a false prover can answer the questions of the verifier with probability more than $2/3$ only if he knows a solution of the MinRank problem.

Therefore, the security of this scheme directly relies on the difficulty of the MinRank problem. Courtois proposed several sets of parameters for which the MinRank problem seemed untractable, yielding secure parameters for the authentication scheme. We report them in Table 2.1.

2.2.2 Rank metric codes

Rank metric codes are a class of linear error-correcting codes where the metric between words is different from the classical Hamming distance. They are defined over an extension GF_{q^e} of a finite field. The distance between two vectors $\mathbf{a} = (a_1, \ldots, a_n) \in \mathsf{GF}_{q^e}^n$ and $\mathbf{b} = (b_1, \ldots, b_n) \in \mathsf{GF}_{q^e}^n$ is given by the rank of the $e \times n$ matrix $(a_1 - b_1, \ldots, a_n - b_n) \in \mathsf{GF}_{q^e}^n$ where each element is seen as a vector in GF_q^e by identifying GF_{q^e} and GF_q^e as GF_q-vector spaces (by fixing a basis $(\beta_1, \ldots, \beta_e) \in \mathsf{GF}_{q^e}$ of linearly independent vectors over GF_q).

A rank-metric code is a vector subspace of dimension k of $\mathsf{GF}_{q^e}^n$ (which is seen as a GF_q-vector space). Given a set of generators $G_1, \ldots, G_k \in \mathsf{GF}_{q^e}^n$ (which can be represented by matrices in $\mathsf{GF}_q^{e \times n}$), and a received word $W \in \mathsf{GF}_{q^e}^n$, decoding W means finding the closest word in the code for the rank-metric.

This is a MinRank problem since it is equivalent to finding a vector $(x_1, \ldots x_n) \in \mathsf{GF}_q^n$ such that the rank of $W - \sum_{i=1}^{q} x_i G_i$ is minimal.

In particular, these rank-metric codes have been used to design cryptosystems. We refer the reader to [Gab85, OJ02, Ove05] for a detailed exposition.

2.2.3 Hidden Field Equations (HFE)

Hidden Field Equations (HFE for short) is an asymmetric encryption scheme proposed in [Pat96]. Its security against message recovery attacks relies on the difficulty of solving boolean systems. However, in [FJ03], the authors show that these boolean systems are actually structured and that this structure can be exploited during Gröbner basis computations. We give here a short and simplified description of HFE. We refer the reader to [KS99] for more details.

The main idea of HFE is that the secret key is a univariate polynomial over an extension field GF_{q^n}. This polynomial has a special shape:

$$P(x) = \sum_{0 \le i,j \le r} a_{i,j} x^{q^i + q^j}.$$

By choosing a basis of GF_{q^n} as a GF_q-vector space, the map $x \mapsto P(x)$ yields a map

$$S: \begin{array}{ccc} \mathsf{GF}_q^n & \to & \mathsf{GF}_q^n \\ \mathbf{x} & \mapsto & (f_1(\mathbf{x}), \ldots, f_n(\mathbf{x})) \end{array},$$

where each polynomial f_i is a quadratic polynomial (since each monomial of P has the form $x^{q^i+q^j}$). Therefore f_i can be represented by a $n \times n$ matrix. Then the structure is hidden by performing invertible linear transforms in GF_q^n on the variables and on the polynomials f_i.

In order to be able to decrypt, the polynomial P should have a relatively low degree so that it can be easily factored in $\mathsf{GF}_{q^n}[x]$. Therefore $r \ll n$, and each quadratic polynomial f_i is represented by a low rank matrix. These low rank matrices (or matrices corresponding to equivalent keys) can be recovered by solving a MinRank problem involving the polynomials of the public key [KS99, BFP11, BFP12a].

2.2.4 McEliece PKC.

The McEliece PKC is an asymmetric encryption scheme based on coding theory. In its original version, it is built upon Goppa codes, which are a family of codes which are easy to decode when it is known how they were constructed. However, knowing a generator matrix of this code is not sufficient to be able to decode efficiently.

The general framework of the McEliece PKC is described below.

Public key. A generator $k \times n$ matrix G of a linear code $\mathscr{C} \subset \mathsf{GF}_q^n$ and an integer $e \in \mathbb{N}$.

Private key. An efficient algorithm for decoding $\mathscr{C}$ up to e errors.

Encryption. To encrypt a vector $\mathbf{v} \in \mathsf{GF}_q^k$, compute $\mathbf{v} \cdot G$ and add e random errors.

Decryption. Use the decoding algorithm to recover $\mathbf{v} \cdot G$, then recover $\mathbf{v}$ by solving a linear system.

In the classical version of McEliece, Goppa codes are proposed as codes whose structure can be easily hidden by linear transforms. These codes are part of a larger family called *alternant codes*. The main specificity of these codes is that there exists a special *parity check matrix* (a matrix such that $H \cdot {}^tG = \mathbf{0}$) with the following shape:

$$H = \begin{bmatrix} y_0 & y_1 & \cdots & y_{n-1} \\ x_0 y_0 & x_1 y_1 & \cdots & x_{n-1} y_{n-1} \\ \vdots & \vdots & \vdots & \vdots \\ x_0^{\delta-1} y_0 & x_1^{\delta-1} y_1 & \cdots & x_{n-1}^{\delta-1} y_{n-1} \end{bmatrix},$$

where the x_i are pairwise distinct in an extension GF_{q^m}, and the y_i are nonzero elements of GF_{q^m}. Once the x_i and the y_i are known, there is an efficient algorithm to decode such codes (see e.g. [FJ98]).

Therefore, one way to attack the McEliece cryptosystem is to solve the algebraic system obtained by the relation $H \cdot {}^tG = \mathbf{0}$:

$$\left\{\begin{array}{rcl}\sum_{i=0}^{n-1} g_{1,i+1} y_i & = & 0,\\ \sum_{i=0}^{n-1} g_{2,i+1} x_i y_i & = & 0,\\ & \vdots & \\ \sum_{i=0}^{n-1} g_{\delta,i+1} x_i^{\delta-1} y_i & = & 0.\end{array}\right. \tag{2.1}$$

This system is overdetermined and bihomogeneous, and finding its solutions in GF_{q^m} would break the McEliece cryptosystem. However, for practical parameters this algebraic system seems to be untractable. Nevertheless some variants of the cryptosystem were recently proposed: in [BCGO09], the authors use quasi-cyclic alternant codes; in [MB09], dyadic Goppa codes are used. The goal is to reduce the sizes of the keys (which is the main drawback of the McEliece cryptosystem) by adding structure to the system. In [FOPT10], the authors show that this structure adds redundancy to the algebraic system and thus propose theoretical and practical attacks on these variants of McEliece by solving bilinear systems. More precisely, they show that the quasi-cyclic and the dyadic structures add linear equations to the modeling of the Mc Eliece cryptosystem. As a result, they can extract a subsystem of "quasi-bilinear equations" where the size of one block of variables is very small compared to the other block of variables.

In Chapter 6, we show that Gröbner bases of affine bilinear systems are easier to compute than Gröbner bases of general quadratic systems and that the maximal degree reached during the computation depends only on the smaller block of variables. This explains the efficiency of the attack proposed in [FOPT10].

2.2.5 QUAD

QUAD is a stream cipher proposed in [BGP09, BGP06]; its security relies on the difficulty of solving quadratic algebraic systems over finite fields. We give here a short and simplified description of the cipher over GF_2. We refer the reader to [BGP09, BGP06] for more details.

In QUAD, we consider a publicly known system $S = (f_1, \ldots, f_{2n})$ of $2n$ quadratic equations in n variables over GF_2. The internal state of the system $\mathbf{x} \in \mathsf{GF}_2^n$ is a vector of n bits. At each round, this internal state is updated as follows

$$\mathbf{x} \leftarrow (f_1(\mathbf{x}), \ldots, f_n(\mathbf{x})).$$

Then n bits of output are generated by computing $(f_{n+1}(\mathbf{x}), \ldots, f_{2n}(\mathbf{x}))$. This process is iterated in order to generate any number of bits.

The designers of this cryptosystem give in [BGP09] a proof that the security of QUAD is related to the difficulty of solving boolean systems. Therefore, in order to estimate secure parameters for QUAD (for instance the value of n), it is important to have good estimates of the complexity of solving boolean systems. This issue is investigated in Chapter 7: we provide an algorithm to exploit the fact that we are looking for solutions in GF_2^n (and not in the algebraic closure).

2.2.6 The Algebraic Surface Cryptosystem

The Algebraic Surface Cryptosystem is an algebraic asymmetric scheme proposed in [AGM09] (a previous version have been given in [AG04]); the design of this cryptosystem was partially supported by Toshiba. It is based on an unusual algebraic problem, the *Section Finding Problem*:

Section Finding Problem (SFP). Given an algebraic surface defined by the polynomial $X(x, y, t) \in \mathsf{GF}_p[x, y, t]$, find two polynomials $u_x(t), u_y(t) \in \mathsf{GF}_p[t]$ of degree d, such that $X(u_x(t), u_y(t), t) = 0$.

We give here a brief description of ASC (see Section 8.1.2 for more details). We consider the ring of polynomials $\mathsf{GF}_p[x, y, t]$ where p is a prime number. For any polynomial $P \in \mathsf{GF}_p[x, y, t]$, Λ_P denotes its support in $\mathsf{GF}_p(t)[x, y]$ (that is to say the set of couples $(i, j) \in \mathbb{N}^2$ such that $t^\ell x^i y^j$ is a monomial of P).

Secret key. A pair of polynomials $(u_x(t), u_y(t)) \in \mathsf{GF}_p[t]$ of degree d.

Public key. A surface described by an irreducible polynomial $X(x, y, t) \in \mathsf{GF}_p[x, y, t]$ such that $X(u_x(t), u_y(t), t) = 0$.

There are some additional technical conditions in order to be able to encrypt and decrypt.

Encryption. Consider a plaintext embedded into a polynomial

$$m(x, y, t) = \sum_{(i,j)\in\Lambda_m} m_{ij}(t)x^i y^j$$

where $\deg(m_{ij}(t)) = d_{ij}^{(m)}$. Choose a random *divisor polynomial*

$$f(x, y, t) = \sum f_{ij}(t)x^i y^j$$

where the degrees of the polynomials f_{ij} are given. Then select four random polynomials r_0, r_1, s_0, s_1 such that, for $\ell \in \{0, 1\}$, r_i has the same monomials as f (only the coefficients are different), and s_i has the same shape as X.

The ciphertext $(F_0(x, y, t), F_1(x, y, t))$ is equal to m masked by the polynomials f, r_i, s_i and X:

$$\begin{aligned} F_0(x, y, t) &= m(x, y, t) + f(x, y, t)s_0(x, y, t) + X(x, y, t)r_0(x, y, t), \\ F_1(x, y, t) &= m(x, y, t) + f(x, y, t)s_1(x, y, t) + X(x, y, t)r_1(x, y, t). \end{aligned}$$

Decryption. For $\ell \in \{0, 1\}$, consider $h_\ell(t) = F_\ell(u_x(t), u_y(t), t)$ and compute the difference $h_0(t) - h_1(t) = f(u_x(t), u_y(t), t)(s_0(u_x(t), u_y(t), t) - s_1(u_x(t), u_y(t), t))$. Next, find a factor of $h_0(t) - h_1(t)$ whose degree matches $\deg(f(u_x(t), u_y(t), t))$. Let $\widetilde{f}(t)$ denote this factor. Then compute $\widetilde{m}(u_x(t), u_y(t), t) = h_0(t) \mod \widetilde{f}(t)$. Finally, retrieve $\widetilde{m}(x, y, t)$ by solving the linear system:

$$\widetilde{m}(u_x(t), u_y(t), t) = \sum \widetilde{m}_{ijk} u_x(t)^i u_y(t)^j t^k.$$

We show in Section 8.1 how Gröbner bases techniques and algebraic tools (normal forms, decompositions of ideals, Gröbner basis computations) can be used to fully break this system: we propose an attack which recovers the plaintext message m in less than 0.05s for recommended parameters. The general principle of the attack in to replace the factorization process in the decryption algorithm by decomposition of ideals.

2.3 Real Solving and Optimization

The critical point method has recently been given a lot of attention for studying properties of real algebraic and semi-algebraic sets. In particular, this method is a subroutine in algorithms for solving optimization problems, for quantifier elimination [HS11], for answering connectivity queries [SS10], or for computing at least one point by connected component in semi-algebraic sets [BPR96, BPR98, GV88, HRS89, HRS93].

2.3.1 Problem statements

Real semi-algebraic sets are sets of points $\mathbf{x} \in \mathbb{R}^n$ satisfying equalities $f_1(\mathbf{x}) = \ldots = f_m(\mathbf{x}) = 0$ and inequalities $g_1(\mathbf{x}) > 0, \ldots, g_k(\mathbf{x}) > 0$ ($f_1, \ldots, f_m, g_1, \ldots, g_k \in \mathbb{R}[x_1, \ldots, x_n]$) or finite unions of such sets. Real algebraic sets appear frequently in engineer sciences (in fact, as soon as there are polynomial constraints). For instance, recent results show that geometrical results can yield important results in control theory [Hen08] or in game theory (see e.g. [HKL+11]).

Therefore, it is crucial to develop efficient tools for studying semi-algebraic sets. In particular, we give here three examples of central problems in effective real geometry.

Polynomial Program. Polynomial programs are families of hard optimization problems:
Given $P, f_1, \ldots, f_p \in \mathbb{Q}[X]$, find (if it exists) $\mathbf{x}_{\mathsf{min}} \in \mathbb{R}$ such that

$$P(\mathbf{x}_{\mathsf{min}}) = \min_{\mathbf{x} \in Z(f_1, \ldots, f_p) \cap \mathbb{R}^n} P(\mathbf{x}).$$

Polynomial programs are usually hard to tackle with numerical algorithms for many reasons. First, the feasible region $Z(f_1, \ldots, f_p) \cap \mathbb{R}^n$ is in general neither convex nor finite. Also there are usually a lot of local extrema, and consequently it is difficult to adapt iterative methods (for instance based on Newton iteration) in this context.

Quantifier elimination. In 1951, Tarski showed in [Tar51] that the theory of real closed fields admits quantifier elimination and is decidable. This means that, over a real closed field (for instance $\mathbb{R}$), any quantified formula of the form

$$\exists (y_1, \ldots, y_\ell) \in \mathbb{R}^\ell, (h_1(X, Y) \bowtie 0, \ldots, h_i(X, Y) \bowtie 0),$$

where $\bowtie$ is either $=$ or $>$ is equivalent to a disjunction of quantifier-free formulas of the form

$$f_1(X) = \ldots = f_m(X) = 0, g_1(X) > 0, \ldots, g_k(X) > 0.$$

An equivalent statement is that if φ is a first-order formula with n free variables, then the set of points $(x_1, \ldots, x_n) \in \mathbb{R}^n$ which satisfy φ is a semi-algebraic set. The corresponding computational problem is to compute from a quantified formula an equivalent quantifier-free formula. It was proved in [DH88] that the size of the quantifier-free formula can be doubly-exponential in the size of the quantified formula. While there exist doubly exponential algorithms solving this problem (for instance the *cylindrical algebraic decomposition* [Col75]), recent results show that in several applicative contexts, the output of the algorithm (i.e. the quantifier free formula) can be weakened but still contains all the useful information. For instance, in [HS11], the authors describe the stability region of the MacCormack scheme (which is the finite difference scheme used to study numerically hyperbolic partial differential equations) by the implementation of a singly exponential algorithm based on the critical point method and Gröbner bases algorithms.

Connectivity queries. Another fundamental problem in real algebraic geometry is to answer connectivity queries, i.e. given a semi-algebraic set $V \subset \mathbb{R}^n$ and two points $\mathbf{x}_1, \mathbf{x}_2 \in V$, decide whether $\mathbf{x}_1$ and $\mathbf{x}_2$ are in the same connected component of V. If so, we also want an algorithm which outputs a path from $\mathbf{x}_1$ to $\mathbf{x}_2$. In order to answer this question, Canny introduced the notion of *roadmap* in [Can88, Can93]. Roughly speaking, a roadmap of V is a 1-dimensional semi-algebraic subset of V which is connected inside each connected component of V. Once a roadmap is computed, it is used as a skeleton to answer connectivity queries. Algorithms for computing roadmaps rely on the critical point method and practical software computing them make intensive use of Gröbner bases computations on critical point systems (see e.g. [SS10] and references therein).

At least one point by connected component. In order to describe the topology of a semi-algebraic set $V \in \mathbb{R}^n$ given by a set of equations and inequalities (or even to decide whether V is empty or not), an important routine is to compute at least one point by connected component of V. Several methods exist to do this and in conjunction with roadmaps, they yield a description of the topology of V. Optimal complexity bounds were achieved by using infinitesimal transformations in [BPR96, BPR98], however these algorithms did not lead to software able to solve this problem in practice. Recently, algorithms based on polar varieties and critical point methods were proposed (see e.g. [SS03, SS04] and references therein). They have been implemented in the Maple RAGlib library and rely heavily on Gröbner bases computations of structured systems.

2.3.2 Algebraic Tools for Real Solving

Critical Point Method. The problem of polynomial optimization can be algebraically represented as follows: a local extrema $\mathbf{x}$ of P restricted to the real trace of $V = Z(f_1, \dots, f_p)$ is also a critical point of the restriction of P to V. Therefore, the evaluation at any local extrema $\mathbf{x}$ of the Jacobian matrix

$$\mathsf{jac}(P, f_1, \dots, f_p) = \begin{bmatrix} \frac{\partial P}{\partial x_1} & \cdots & \frac{\partial P}{\partial x_n} \\ \frac{\partial f_1}{\partial x_1} & \cdots & \frac{\partial f_1}{\partial x_n} \\ \vdots & \vdots & \vdots \\ \frac{\partial f_p}{\partial x_1} & \cdots & \frac{\partial f_p}{\partial x_n} \end{bmatrix}$$

is rank defective. Therefore, $\mathbf{x}_{\mathsf{min}}$ is a real zero of the system of polynomials $\{f_1, \dots, f_p\} \cup \mathsf{MaxMinors}(\mathsf{jac}(P, f_1, \dots, f_p))$.

Under mild genericity assumptions, this system is 0-dimensional (see Chapter 5), so Gröbner basis techniques can be used to obtain a rational parametrization of the real local extrema of the restriction of P to V.

Under these genericity assumptions, the system $\{f_1, \dots, f_p\} \cup \mathsf{MaxMinors}(\mathsf{jac}(P, f_1, \dots, f_p))$ is the union of a regular sequence $(f_1, \dots, f_p)$ and of a determinantal system $\mathsf{MaxMinors}(\mathsf{jac}(P, f_1, \dots, f_p))$. This kind of systems is studied in Chapter 5. Notice that it is also possible to express the rank condition of the Jacobian matrix by using *Lagrange multipliers*, i.e. a set of fresh variables modeling a vector in the kernel.

More generally, the critical point method can be used to study any semi-algebraic set V: the critical points of the projection on the first coordinate

$$\begin{array}{cccc} \pi_i : & V \cap \mathbb{R}^n & \longrightarrow & \mathbb{R}^i \\ & (x_1, \dots, x_n) & \longmapsto & (x_1, \dots, x_i) \end{array}$$

yield useful information on the geometry of V and their computation is a subroutine of several algorithms for real solving.

In [SS03], the authors show that the set of such critical points is a 0-dimensional variety under mild genericity assumptions on the polynomials $f_1, \dots, f_p$ defining the variety V. Algorithms for computing such points are given in [BGHM01, BGHM97, BGHP05, BGHP04, BGH$^+$10, SS03, ARS02, FMRS08]. The RAGlib maple package implements the algorithms given in [SS03, FMRS08] using Gröbner bases.

Most known complexity results for computing critical points are based on the complexity of geometric resolution [BGHM01, BGHM97, BGHP05, BGHP04]. However, in practice, it has been observed that Gröbner bases algorithms are also efficient for solving critical point systems and several challenges and open problems have been solved by using the RAGlib maple package with the

FGb Gröbner engine. We give in Chapter 5 first theoretical complexity estimates which explain this behavior.

Polar varieties. Polar varieties can be seen as generalizations of critical points of the projection π_1 and their computation is a subroutine of several algorithms for studying properties of real (semi-)algebraic sets.

Let $f_1, \dots, f_p \in \mathbb{Q}[X]$ be polynomials generating a radical ideal such that their zero set $V \subset \mathbb{C}^n$ is equidimensional of dimension d (i.e. all irreducible components of the variety have dimension d, see Theorem 1.15). Let π_i denote the restriction to $V \cap \mathbb{R}^n$ of the projection on the i first coordinates:

$$\begin{array}{cccc} \pi_i : & V \cap \mathbb{R}^n & \longrightarrow & \mathbb{R}^i \\ & (x_1, \dots, x_n) & \longmapsto & (x_1, \dots, x_i) \end{array}$$

Then for i from 1 to d, the $n - i + 1$-th polar variety is defined as the critical points of π_i, i.e. the points of $V \cap \mathbb{R}^n$ where the rank of the truncated Jacobian matrix

$$\mathsf{jac}(\mathbf{F}, i) = \begin{bmatrix} \frac{\partial f_1}{\partial x_{i+1}} & \cdots & \frac{\partial f_1}{\partial x_n} \\ \vdots & \vdots & \vdots \\ \frac{\partial f_p}{\partial x_{i+1}} & \cdots & \frac{\partial f_p}{\partial x_n} \end{bmatrix}$$

is less than $n - d$. Therefore the $n - i + 1$-th polar variety, denoted by W_i, is defined by the vanishing of the polynomials $f_1, \dots, f_p$ and of the minors of size $n - d$ of $\mathsf{jac}(\mathbf{F}, i)$: $Z(f_1, \dots, f_p, \mathsf{Minors}(\mathsf{jac}(\mathbf{F}, i), n - d))$. Following [SS03], under some properness assumptions that are satisfied generically, the dimension of W_i is equal to $i - 1$. These varieties play a central role in several algorithms in effective real algebra (see e.g. [BGHM01, BGHM97, BGHP05, BGHP04, SS03, BGH$^+$10, GS11]) Therefore it is important to be able to compute Gröbner bases of the ideal $\langle f_1, \dots, f_p, \mathsf{Minors}(\mathsf{jac}(\mathbf{F}, i), n - d)\rangle$ and to estimate the complexity of such computations.

Chapter 3

Preliminaries on Determinantal and Multi-homogeneous systems

As shown in Chapter 2, determinantal and multi-homogeneous ideals appear frequently in several areas. We recall in this chapter several known results on their structural properties.

3.1 Determinantal systems

In this section, we focus on the structure of the determinantal ideal $\mathscr{D}_r \subset \mathbb{K}[U]$ generated by the set of $(r+1)$-minors of the matrix

$$\mathscr{U} = \begin{bmatrix} u_{1,1} & \dots & u_{1,q} \\ \vdots & \vdots & \vdots \\ u_{p,1} & \dots & u_{p,q} \end{bmatrix}$$

Without loss of generality, we assume that $q \leq p$.

The ideal $\mathscr{D}_r$ has been extensively studied during last decades. In particular, explicit formulas for its degree and for its Hilbert series are known (see e.g. [Ful97, Example 14.4.14] and [CH94]), as well as structural properties such as Cohen-Macaulayness and primality [HE70, HE71].

Notations 3.1. *We let U denote the set of variables $\{u_{1,1}, \dots, u_{p,q}\}$. The notation $A_r^{p,q}(t) \in \mathbb{Z}[t]^{r\times r}$ stands for the $r \times r$-matrix whose (i,j)-entry is $\sum_{\ell\in\mathbb{N}} \binom{p-i}{\ell}\binom{q-j}{\ell}$.*

The formula for the Hilbert series of $\mathbb{K}[U]/\mathscr{D}_r$ is related to combinatorial properties of the ideal $\mathscr{D}_r$. In particular, in [CH94], the authors show a relation between this series and the combinatorial structure of a class of non-intersecting path; [Kra93, Kul96] enumerates such paths, and these formulas are used in [CH94] to obtain the Hilbert series of $\mathbb{K}[U]/\mathscr{D}_r$.

Theorem 3.2. *[Abh88, CH94, Kra93, Kul96] The Hilbert series of the ring $\mathbb{K}[U]/\mathscr{D}_r$ is*

$$\mathsf{HS}_{\mathbb{K}[U]/\mathscr{D}_r}(t) = \frac{\det\left(A_r^{p,q}(t)\right)}{t^{\binom{r}{2}}(1-t)^{(p+q-r)r}}.$$

By Proposition 1.43, the dimension and the degree can be read off from the Hilbert series:

Corollary 3.3. *The dimension and the degree of the ideal $\mathscr{D}_r \subset \mathbb{K}[U]$ are respectively*

- $\dim(\mathscr{D}_r) = (p+q-r)r$*;*

- $\mathsf{DEG}(\mathscr{D}_r) = \prod_{i=0}^{q-r-1} \frac{i!(p+i)!}{(q-1+i)!(p-r+i)!}$.

Proof. For the formula for the degree, we refer the reader to [Ful97, Example A.9.4, Example 14.4.14]. In particular, [Ful97, Example A.9.4] shows that $\det(A_r^{p,q}(1)) \neq 0$. Therefore, the dimension of $\mathscr{D}_r \subset \mathbb{K}[U]$ is equal to the exponent of $(1-t)$ in the denominator of the rational function given in Theorem 3.2, namely $\dim(\mathscr{D}_r) = (p+q-r)r$. □

An interesting feature of determinantal ideal is that they provide a class of non-trivial Cohen-Macaulay domains.

Proposition 3.4. *[BV88] The ring $\mathbb{K}[U]/\mathscr{D}_r$ is a* Cohen-Macaulay *domain: if $h_1, \ldots, h_\ell \in \mathbb{K}[U]$ are homogeneous polynomials such that $\dim(\mathscr{D} + \langle h_1, \ldots, h_\ell\rangle) = \dim(\mathscr{D}_r) - \ell$ then $(h_1, \ldots, h_\ell)$ is a $\mathbb{K}[U]/\mathscr{D}_r$-regular sequence, i.e. for all $i \in \{1, \ldots, \ell\}$, h_i does not divide 0 in the ring $\mathbb{K}[U]/\left(\mathscr{D}_r + \langle h_1, \ldots, h_{i-1}\rangle\right)$.*

In Chapters 4 and 5, the results above are the cornerstones of the proofs for analyzing the structure of ideals corresponding to Generalized MinRank problems and critical point systems.

3.2 Structure of multi-homogeneous ideals

In this section, we recall several known results on multi-homogeneous ideals. For simplicity of notations, most results are only stated for bilinear systems but they can be easily extended to multi-homogeneous systems.

When $f_1, \ldots, f_m$ is a bi-homogeneous system in $\mathbb{K}[x_0, \ldots, x_{n_x}, y_0, \ldots, y_{n_y}]$ (with degree at least 1 with respect to each block of variables), there is a set of trivial solutions that we need to take into account: the varieties $Z(x_0, \ldots, x_{n_x})$ and $Z(y_0, \ldots, y_{n_y})$ are necessarily subsets of $Z(f_1, \ldots, f_m)$. The corresponding ideals $\langle x_0, \ldots, x_{n_x}\rangle$ and $\langle y_0, \ldots, y_{n_y}\rangle$ are called the *irrelevant ideals*. Contrary to the classical homogeneous case, these irrelevant ideals are not maximal, and this fact has several consequences: for instance, as soon as $m \geq n_x + 1$, regular bi-homogeneous sequences of size m do not exist.

In the section, we recall tools to transpose some results on homogeneous ideals in this context. Roughly speaking, the objective is to show that there exists a generic property of bi-homogeneous systems which is similar to *regularity* for homogeneous systems.

We use the following notations:

Notations 3.5.
- *$\mathscr{BL}_{\mathbb{K}}(n_x, n_y)$ the $\mathbb{K}$-vector space of bilinear forms in $\mathbb{K}[X, Y] = \mathbb{K}[x_0, \ldots, x_{n_x}, y_0, \ldots, y_{n_y}]$;*

- *X (resp. Y) is the ideal $\langle x_0, \ldots, x_{n_x}\rangle$ (resp. $\langle y_0, \ldots, y_{n_y}\rangle$);*

- *An ideal is called* bihomogeneous *if it admits a set of bihomogeneous generators.*

- *If $(f_1, \ldots, f_m) \in \mathscr{BL}_{\mathbb{K}}(n_x, n_y)^m$ is a family of bilinear forms, I_i denotes the ideal $\langle f_1, \ldots, f_i\rangle$ and J_i denotes the saturated ideal $I_i : (X \cap Y)^\infty$;*

- *Given a polynomial sequence $\mathbf{F} = (f_1, \ldots, f_m)$, we denote by $\mathsf{Syz}_{\mathsf{triv}}(\mathbf{F})$ the module of trivial syzygies, i.e. the set of all syzygies $(s_1, \ldots, s_m)$ such that for all $i \in \{1, \ldots, m\}$, $s_i \in \langle f_1, \ldots, f_{i-1}, f_{i+1}, \ldots, f_m\rangle$;*

- *A primary ideal $P \subset \mathbb{K}[X, Y]$ is called* admissible *if $X \not\subset \sqrt{P}$ and $Y \not\subset \sqrt{P}$;*

Lemma 3.6. *[ST06] Let $f_1, \ldots, f_m \in \mathbb{K}[X,Y]$ be polynomials, $I_m = \cap P_{\ell \in L}$ be a minimal primary decomposition of I_m and let $\mathsf{Adm} = \{P_\ell \mid X \not\subset \sqrt{P_\ell}$ and $Y \not\subset \sqrt{P_\ell}\}$ be the set of the admissible ideals of the decomposition. Then $J_m = \cap_{P \in \mathsf{Adm}} P$.*

Proof. Let $h \in J_m$ be a polynomial. Since $J_m = I_m : (X \cap Y)^\infty$, there exists an integer $k \in \mathbb{N}$ such that $h(X \cap Y)^k \subset I_m$. Consequently, for all $\ell \in L$, $h(X \cap Y)^k \subset P_\ell$. The ideal P_ℓ is primary, therefore if $h \notin P_\ell$, then there exists an integer $k' \in \mathbb{N}$ such that $(X \cap Y)^{k'} \subset P_\ell$. Hence $(X \cap Y) \subset \sqrt{P_\ell}$ and thus, since $\sqrt{P_\ell}$ is prime, $X \subset \sqrt{P_\ell}$ or $Y \subset \sqrt{P_\ell}$. It follows that P_ℓ is not an admissible ideal. Conversely, let $h \in \cap_{P \in \mathsf{Adm}} P$ be a polynomial. Let P_ℓ be a non-admissible primary ideal of the decomposition. Therefore there exists an integer $k_\ell \in \mathbb{N}$ such that $(X \cap Y)^{k_\ell} \subset P_\ell$. Let $k' \in \mathbb{N}$ be the integer defined by

$$k' = \max_{\substack{\ell \in L \\ (X \cap Y) \subset \sqrt{P_\ell}}} \{k_\ell\}.$$

Then notice that $h(X \cap Y)^{k'}$ is a subset of all ideals in the primary decomposition of I_m, and hence $h(X \cap Y)^{k'} \subset I_m$. Consequently, $h \in J_m$. □

The polynomial f_m always divides 0 in the ring $\mathbb{K}[X,Y]/\langle f_1, \ldots, f_{m-1}\rangle$ if the polynomials $f_1, \ldots, f_m$ are bilinear and $m \geq \min n_x, n_y + 1$. However, this is due to the irrelevant ideals. Therefore, we have to consider the ideals after saturation by these irrelevant ideals:

Proposition 3.7. *let $f_1, \ldots, f_m \in \mathbb{K}[X,Y]$ be polynomials with $m \leq n_x + n_y$, and $\mathrm{Ass}(I_{i-1})$ be the set of prime ideals associated to I_{i-1}. The following assertions are equivalent:*

1. *for all $i \in \{2, \ldots, m\}$, f_i is not a divisor of 0 in $\mathbb{K}[X,Y]/J_{i-1}$.*
2. *for all $i \in \{2, \ldots, m\}$, $(f_i \in P, P \in \mathrm{Ass}(I_{i-1})) \Rightarrow P$ is non-admissible.*

Proof. It is a straightforward consequence of Lemma 3.6. □

In the following, let $\mathfrak{a}$ be the set

$$\mathfrak{a} = \{\mathfrak{a}_{j,k}^{(i)} \mid 1 \leq i \leq m, 0 \leq j \leq n_x, 0 \leq k \leq n_y\}.$$

We consider generic polynomials $\mathfrak{f}_1, \ldots, \mathfrak{f}_m$ in $\mathbb{K}(\mathfrak{a})[x_0, \ldots, x_{n_x}, y_0, \ldots, y_{n_y}]$:

$$\mathfrak{f}_i = \sum \mathfrak{a}_{j,k}^{(i)} x_j y_k$$

and we denote by $I \subset \mathbb{K}(\mathfrak{a})[x_0, \ldots, x_{n_x}, y_0, \ldots, y_{n_y}]$ the ideal they generate.

Lemma 3.8. *Let P be an admissible prime ideal of $\mathbb{K}[X,Y]$. The set of bilinear polynomials $f \in \mathscr{BL}_{\overline{\mathbb{K}}}(n_x, n_y)$ such that $f \notin P$ contains a non-empty Zariski open set of $\mathscr{BL}_{\overline{\mathbb{K}}}(n_x, n_y)$ (which is seen as a $\overline{\mathbb{K}}$-vector space of dimension $(n_x+1)(n_y+1)$.*

Proof. Let $\mathfrak{f}$ be the generic bilinear polynomial

$$\mathfrak{f} = \sum_{j,k} \mathfrak{a}_{j,k} x_j y_k$$

in $\mathbb{K}(\{\mathfrak{a}_{j,k}\}_{0 \leq j \leq n_x, 0 \leq k \leq n_y})[x_0, \ldots, x_{n_x}, y_0, \ldots, y_{n_y}]$. Since P is admissible, there exists $x_{j_0} y_{k_0}$ such that $x_{j_0} y_{k_0} \notin P$ (this shows the non-emptiness). Let $\prec$ be an admissible order. Then consider the normal form for this order

$$\mathrm{NF}_{\prec, P}(\mathfrak{f}) = \sum_{t \text{ monomial}} h_t(\mathfrak{a}_{0,0} \ldots, \mathfrak{a}_{n_x, n_y}) t.$$

By multiplying by the least common multiple of the denominators, we can assume without loss of generality that for each t, h_t is a polynomial. Thus, if a bilinear polynomial is in P, then its coefficients are solutions of the polynomial equation $h_t(\mathfrak{a}_{0,0}, \dots, \mathfrak{a}_{n_x,n_y}) = 0$ for any monomial t. $\square$

Lemma 3.9. *For all $i \in \{1, \dots, m-1\}$, $\mathfrak{f}_{i+1}$ does not divide 0 in $\mathbb{K}(\mathfrak{a})[X, Y]/(\langle \mathfrak{f}_1, \dots, \mathfrak{f}_i\rangle : (X \cap Y)^\infty)$.*

Proof. Let P be an admissible prime associated to $\langle \mathfrak{f}_1, \dots, \mathfrak{f}_i\rangle$. Then there exists a family of generators of P which involves only the parameters $\mathfrak{a}_{j,k}^{(\ell)}$ with $\ell \le i$. By an argument similar to the proof of Lemma 3.8, $\mathrm{NF}_P(\mathfrak{f}_{i+1}) \neq 0$. Since this is true for every admissible prime in $\mathrm{Ass}(\langle \mathfrak{f}_1, \dots, \mathfrak{f}_i\rangle)$, $\mathfrak{f}_{i+1}$ does not divide 0 in $\mathbb{K}(\mathfrak{a})[X, Y]/(\langle \mathfrak{f}_1, \dots, \mathfrak{f}_i\rangle : (X \cap Y)^\infty)$. $\square$

We can now define a property similar to *regularity* for bi-homogeneous systems (and by extension for multi-homogeneous systems):

Proposition 3.10. *Let $m \le n_x + n_y$ and $f_1, \dots, f_m$ be bilinear polynomials such that for all $i \in \{1, \dots, m-1\}$, f_{i+1} is not a divisor of 0 in $\mathbb{K}[X, Y]/J_i$. Then for all $i \in \{1, \dots, m\}$, the ideal J_i is equidimensional and its codimension is i.*

Proof. We prove the Proposition by induction on m.

- $J_1 = I_1$ is equidimensional and $\mathsf{codim}(I_1) = 1$;
- Suppose that J_{i-1} is equidimensional of codimension $i-1$. Then $J_i = (J_{i-1} + f_i) : (X \cap Y)^\infty$. f_i does not divide 0 in $\mathbb{K}[X, Y]/J_{i-1}$, thus $J_{i-1} + f_i$ is equidimensional of codimension i. The saturation does not decrease the dimension of any primary component of $J_{i-1} + \langle f_i\rangle$. Therefore, J_i is equidimensional and its codimension is i.

$\square$

Corollary 3.11. *If $m \le n_x + n_y$ then for all $i \in \{2, \dots, m\}$, the ideals $\langle \mathfrak{f}_1, \dots, \mathfrak{f}_i\rangle : (X \cap Y)^\infty$ and $\langle \mathfrak{f}_1, \dots, \mathfrak{f}_{i-1}\rangle : (X \cap Y)^\infty + \langle \mathfrak{f}_i\rangle$ are equidimensional of codimension i.*

Proof. This is a direct consequence of Lemma 3.9 and of Proposition 3.10. $\square$

Theorem 3.12 states that the property of non-divisibility of zero with respect to the saturated ideal is a *generic* property, similarly to regularity for homogeneous systems.

Theorem 3.12. *Let $m, n_x, n_y \in \mathbb{N}$ such that $m \le n_x + n_y$. Then there exists a non-empty Zariski open subset $O \subset \mathscr{BL}_{\overline{\mathbb{K}}}(n_x, n_y)^m$ such that, for all bilinear system $(f_1, \dots, f_m) \in O \cap \mathscr{BL}_{\mathbb{K}}(n_x, n_y)^m$, f_{i+1} does not divide 0 in $\mathbb{K}[X, Y]/J_i$.*

Proof. There exists an algorithm which computes the equidimensional decomposition of a polynomial ideal by using only arithmetic operations on the coefficients of the polynomial system [Lec03]. During the computation of the equidimensional decompositions of all ideals $\langle \mathfrak{f}_1, \dots, \mathfrak{f}_i\rangle$ with this algorithm, a finite number of rational functions in $\mathbb{K}(\mathfrak{a})$ appear. Therefore there exists a non-empty Zariski open subset $O \subset \mathscr{BL}_{\overline{\mathbb{K}}}(n_x, n_y)^m$ such that, for all bilinear system $(f_1, \dots, f_m) \in O \cap \mathscr{BL}_{\mathbb{K}}(n_x, n_y)^m$, the numerators and the denominators of these rational functions do not vanish, and hence the equidimensional decomposition of $\langle f_1, \dots, f_i\rangle$ is equal to the specialization of that of $\langle \mathfrak{f}_1, \dots, \mathfrak{f}_i\rangle$. Therefore $\mathsf{codim}\,(\langle f_1, \dots, f_i\rangle : (X \cap Y)^\infty) = i$ and $\mathsf{codim}\,(\langle f_1, \dots, f_i\rangle : (X \cap Y)^\infty + \langle f_{i+1}\rangle) = i + 1$. Consequently f_{i+1} does not divide 0 in $\mathbb{K}[X, Y]/J_i$. $\square$

3.3 Ideals generated by generic affine bilinear systems

In this section, we focus on structural properties of ideals generated by affine bilinear systems (these results are not true for general multi-homogeneous systems). In particular, we show that the projection of the variety of an affine bilinear system on the space defined by one block of variables is exactly the zero set of a determinantal system. This establishes a correspondence between determinantal systems (where the entries of the matrix are linear) and bilinear systems.

We assume here that $\mathbb{K}$ is a field of characteristic 0: this is needed in the proof of Lemma 3.14 to use an algebraic version of Sard's Theorem (however there exist variants of Sard's Theorem in positive characteristic, see e.g. [Eis95, Corollary 16.23]). Let $m = n_x + n_y$ denote the number of equations, and $\mathfrak{a}$ be the set

$$\mathfrak{a} = \{\mathfrak{a}^{(i)}_{j,k} \mid 1 \le i \le m, 0 \le j \le n_x, 0 \le k \le n_y\}.$$

We consider generic polynomials $f_1, \dots, f_m$ in $\mathbb{K}(\mathfrak{a})[x_0, \dots, x_{n_x}, y_0, \dots, y_{n_y}]$:

$$f_i = \sum \mathfrak{a}^{(i)}_{j,k} x_j y_k$$

and we denote by $I \subset \mathbb{K}(\mathfrak{a})[x_0, \dots, x_{n_x}, y_0, \dots, y_{n_y}]$ the ideal they generate. In the sequel of this section, ϑ denotes the dehomogenization morphism:

$$\begin{array}{ccc} \mathbb{K}[x_0, \dots, x_{n_x}, y_0, \dots, y_{n_y}] & \longrightarrow & \mathbb{K}[x_0, \dots, x_{n_x-1}, y_0, \dots, y_{n_y-1}] \\ f(x_0, \dots, x_{n_x}, y_0, \dots, y_{n_y}) & \longmapsto & f(x_0, \dots, x_{n_x-1}, 1, y_0, \dots, y_{n_y-1}, 1) \end{array}.$$

For $\mathbf{a} \in \mathbb{K}^{m(n_x+1)(n_y+1)}$, $\varphi_{\mathbf{a}}$ stands for the specialization:

$$\begin{array}{cccc} \varphi_{\mathbf{a}}: & \mathbb{K}(\mathfrak{a})[x_0, \dots, x_{n_x}, y_0, \dots, y_{n_y}] & \to & \mathbb{K}[x_0, \dots, x_{n_x}, y_0, \dots, y_{n_y}] \\ & f(\mathfrak{a})(x_0, \dots, x_{n_x}, y_0, \dots, y_{n_y}) & \mapsto & f(\mathbf{a})(x_0, \dots, x_{n_x}, y_0, \dots, y_{n_y}) \end{array}$$

Also $\varphi_{\mathbf{a}}(I)$ denotes the ideal $\langle \varphi_{\mathbf{a}}(f_1), \dots, \varphi_{\mathbf{a}}(f_m) \rangle \subset \mathbb{K}[X, Y]$ and $Z(\varphi_{\mathbf{a}}(I), \mathbb{P}^{n_x} \times \mathbb{P}^{n_y}) \subset \mathbb{P}^{n_x}\overline{\mathbb{K}} \times \mathbb{P}^{n_y}\overline{\mathbb{K}}$ (resp. $Z(\vartheta \circ \varphi_{\mathbf{a}}(I)) \subset \overline{\mathbb{K}}^{n_x+n_y}$) denotes the variety of $\varphi_{\mathbf{a}}(I)$ (resp. $\vartheta \circ \varphi_{\mathbf{a}}(I)$).

First, we recall that generically all isolated solutions of a bilinear system are located on an affine chart.

Lemma 3.13. *There exists a nonempty Zariski open set $O_1 \subset \overline{\mathbb{K}}^{m(n_x+1)(n_y+1)}$ such that if $\mathbf{a} \in O_1 \cap \mathbb{K}^{m(n_x+1)(n_y+1)}$, then for all $(\alpha_0, \dots, \alpha_{n_x}, \beta_0, \dots, \beta_{n_y}) \in Z(\varphi_{\mathbf{a}}(I), \mathbb{P}^{n_x} \times \mathbb{P}^{n_y})$, $\alpha_{n_x} \neq 0$ and $\beta_{n_y} \neq 0$. This implies that the application*

$$\begin{array}{ccc} Z(\vartheta \circ \varphi_{\mathbf{a}}(I)) & \longrightarrow & Z(\varphi_{\mathbf{a}}(I), \mathbb{P}^{n_x} \times \mathbb{P}^{n_y}) \\ (\alpha_0, \dots, \alpha_{n_x-1}, \beta_0, \dots, \beta_{n_y-1}) & \longmapsto & ((\alpha_0 : \dots : \alpha_{n_x-1} : 1), (\beta_0 : \dots : \beta_{n_y-1} : 1)) \end{array}$$

is a bijection.

Proof. See [Van29, page 751]. □

Lemma 3.14. *There exists a nonempty Zariski open set $O_2 \subset \overline{\mathbb{K}}^{m(n_x+1)(n_y+1)}$, such that if $\mathbf{a} \in O_2 \cap \mathbb{K}^{m(n_x+1)(n_y+1)}$, then the ideal $\vartheta \circ \varphi_{\mathbf{a}}(I)$ is radical.*

Proof. Denote by $\mathbf{F}$ the polynomial family $(f_1, \dots, f_m) \in \mathbb{K}[\mathfrak{a}, X, Y]^m$. Let $J \subset \mathbb{K}[\mathfrak{a}]$ be the ideal $\big(I + \langle \det(\mathsf{jac}_{X,Y}(\mathbf{F}))\rangle\big) \cap \mathbb{K}[\mathfrak{a}]$ and $Z(J)$ be its associated algebraic variety. By the Jacobian Criterion (see e.g. [Eis95, Theorem 16.19]), if $\mathbf{a}$ does not belong to $Z(J)$, then $\vartheta \circ \varphi_{\mathbf{a}}(I)$ is radical. Thus, it is sufficient to prove that $\mathbb{K}^{m(n_x+1)(n_y+1))} \setminus Z(J)$ is non-empty.

To do that, we prove that for all $\mathbf{a} \in \overline{\mathbb{K}}^{m(n_x+1)(n_y+1)}$, there exists $(\varepsilon_1, \ldots, \varepsilon_m)$ such that the ideal $\langle \vartheta \circ \varphi_{\mathbf{a}}(f_1) + \varepsilon_1, \ldots, \vartheta \circ \varphi_{\mathbf{a}}(f_m) + \varepsilon_m \rangle$ is radical. For $i \in \{1, \ldots, m\}$, let g_i denote the polynomial $\vartheta \circ \varphi_{\mathbf{a}}(f_i)$ and consider the mapping Ψ

$$x \in \overline{\mathbb{K}}^m \to (g_1(x), \ldots, g_m(x)) \in \overline{\mathbb{K}}^m.$$

Suppose first that $\Psi(\overline{\mathbb{K}}^m)$ is not dense in $\overline{\mathbb{K}}^m$. Since $\Psi(\mathbb{K}^m)$ is a constructible set, it is contained in a Zariski-closed subset of $\overline{\mathbb{K}}^m$ and there exists $(\varepsilon_1, \ldots, \varepsilon_m)$ such that the algebraic variety defined by $g_1 - \varepsilon_1 = \cdots = g_m - \varepsilon_m = 0$ is empty. Since there exists $\mathbf{a}'$ such that $g_i - \varepsilon_i = \vartheta \circ \varphi_{\mathbf{a}'}(f_i)$, we conclude that $\vartheta \circ \varphi_{\mathbf{a}'}(I) = \langle 1 \rangle$. This implies that $\mathbf{a}' \notin Z(J)$.

Suppose now that $\Psi(\overline{\mathbb{K}}^m)$ is dense in $\overline{\mathbb{K}}^m$. By Sard's Theorem [Sha94, Chap. 2, Section 6.2, Theorem 2], there exists $(\varepsilon_1, \ldots, \varepsilon_m) \in \overline{\mathbb{K}}^m$ which does not lie in the set of critical values of Ψ. This implies that at any point of the algebraic variety defined by $g_1 - \varepsilon_1 = \cdots = g_m - \varepsilon_m = 0$, $\vartheta \circ \varphi_{\mathbf{a}}(\det(\mathsf{jac}_{X,Y}(\mathbf{F})))$ does not vanish. Remark now that there exists $\mathbf{a}'$ such that $g_i - \varepsilon_i = \vartheta \circ \varphi_{\mathbf{a}'}(f_i)$. We conclude that $\mathbf{a}' \in \mathbb{K}^{m(n_x+1)(n_y+1)} \setminus Z(J)$, which ends the proof. □

The following lemma establishes the relation between the solutions of bilinear systems and the locus of rank defect of linear matrices.

Lemma 3.15. *There exists a nonempty Zariski open set* $O_3 \subset \overline{\mathbb{K}}^{m(n_x+1)(n_y+1)}$*, such that if* $\mathbf{a} \in O_3 \cap \mathbb{K}^{m(n_x+1)(n_y+1)}$,

$$\sqrt{\langle \mathsf{MaxMinors}(\vartheta \circ \varphi_{\mathbf{a}}(\mathsf{jac}_{\mathbf{y}}(\mathbf{F}))) \rangle} = \langle \vartheta \circ \varphi_{\mathbf{a}}(f_1), \ldots, \vartheta \circ \varphi_{\mathbf{a}}(f_m) \rangle \cap \mathbb{K}[x_0, \ldots, x_{n_x-1}].$$

Proof. Let $\mathbf{a}$ be an element in O_2 (as defined in Lemma 3.14). Thus $\vartheta \circ \varphi_{\mathbf{a}}(I)$ is radical. Now let $(v_0, \ldots, v_{n_x-1}, w_0, \ldots, w_{n_y-1}) \in Z(\vartheta \circ \varphi_{\mathbf{a}}(I))$ be an element of the variety. Then

$$\left(\vartheta \circ \varphi_{\mathbf{a}}(\mathsf{jac}_{\mathbf{y}}(\mathbf{F}))_{x_i = v_i}\right) \cdot \begin{pmatrix} w_0 \\ \vdots \\ w_{n_y-1} \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}.$$

This implies that $\mathsf{Rank}(\vartheta \circ \varphi_{\mathbf{a}}(\mathsf{jac}_{\mathbf{y}}(\mathbf{F}))_{x_i = v_i}) < n_y + 1$, and therefore

$$(v_0, \ldots, v_{n_x-1}) \in Z(\langle \mathsf{MaxMinors}(\vartheta \circ \varphi_{\mathbf{a}}(\mathsf{jac}_{\mathbf{y}}(\mathbf{F}))) \rangle).$$

Conversely, let $(v_0, \ldots, v_{n_x-1}) \in Z(\langle \mathsf{MaxMinors}(\vartheta \circ \varphi_{\mathbf{a}}(\mathsf{jac}_{\mathbf{y}}(\mathbf{F}))) \rangle)$. Thus there exists a non trivial vector $(w_0, \ldots, w_{n_y})$ in the right kernel $\mathsf{Ker}(\vartheta \circ \varphi_{\mathbf{a}}(\mathsf{jac}_{\mathbf{y}}(\mathbf{F}))_{x_i = v_i})$. This means that $(v_0, \ldots, v_{n_x-1}, 1, w_0, \ldots, w_{n_y})$ is in the variety of $\varphi_{\mathbf{a}}(I)$:

$$(v_0, \ldots, v_{n_x-1}, 1, w_0, \ldots, w_{n_y}) \in Z(\varphi_{\mathbf{a}}\left(\mathsf{jac}_{\mathbf{y}}(\mathbf{F})\right) \cdot \begin{pmatrix} y_0 \\ \vdots \\ y_{n_y} \end{pmatrix})$$

From Lemma 3.13, there exists a nonempty Zariski open set O_1 such that, if $\mathbf{F} \in O_1$, $w_{n_y} \neq 0$. Hence

$$(v_0, \ldots, v_{n_x-1}, \frac{w_0}{w_{n_y}}, \ldots, \frac{w_{n_y-1}}{w_{n_y}}) \in Z(\vartheta \circ \varphi_{\mathbf{a}}(I)).$$

Finally, we have

$$Z(\langle \mathsf{MaxMinors}(\vartheta \circ \varphi_{\mathbf{a}}(\mathsf{jac}_{\mathbf{y}}(\mathbf{F}))) \rangle) = Z(\langle \vartheta \circ \varphi_{\mathbf{a}}(f_1), \ldots, \vartheta \circ \varphi_{\mathbf{a}}(f_m) \rangle \cap \mathbb{K}[x_0, \ldots, x_{n_x-1}])$$

and $\vartheta \circ \varphi_{\mathbf{a}}(I)$ is radical (Lemma 3.14). The Nullstellensatz concludes the proof. □

Corollary 3.16. *There exists a nonempty Zariski open set* $O_4 \subset \overline{\mathbb{K}}^{m(n_x+1)(n_y+1)}$*, such that if* $\mathbf{a} \in O_4 \cap \mathbb{K}^{m(n_x+1)(n_y+1)}$*, the set* $Z(\vartheta \circ \varphi_{\mathbf{a}}(I))$ *is finite and its cardinality is*

$$\operatorname{card}(Z(\vartheta \circ \varphi_{\mathbf{a}}(I))) = \mathsf{DEG}(\vartheta \circ \varphi_{\mathbf{a}}(I)) = \binom{n_x + n_y}{n_x}$$

Proof. According to Lemmas 3.14 and 3.13, if $\mathbf{a} \in O_1 \cap O_2$, then $\deg(\vartheta \circ \varphi_{\mathbf{a}}(I)) = \operatorname{card}(Z(\vartheta \circ \varphi_{\mathbf{a}}(I)) = \operatorname{card}(Z(\varphi_{\mathbf{a}}(I)))$. This value is the so-called multihomogeneous Bézout number of $\varphi_{\mathbf{a}}(I)$, i.e. the coefficient of $z_1^{n_x} z_2^{n_y}$ in $(z_1 + z_2)^{n_x+n_y}$ (see e.g. [MS87]), namely $\binom{n_x+n_y}{n_x}$. □

Part II

Contributions

Chapter 4

Determinantal Systems

The results presented in this chapter are joint work with J.-C. Faugère and M. Safey El Din and are in the preprint [FSS11b] (in submission).

In this chapter, we study the complexity of solving the *generalized MinRank problem*, i.e. computing the set of points where the evaluation of a polynomial matrix has rank at most r. A natural algebraic representation of this problem gives rise to a *determinantal ideal*: the ideal generated by all minors of size $r + 1$ of the matrix. Under genericity assumptions on the input matrix, we give new complexity bounds for solving this problem using Gröbner bases algorithms. In particular, these complexity bounds allow us to identify families of generalized MinRank problems for which the arithmetic complexity of the solving process is polynomial in the number of solutions.

4.1 Introduction

We focus in this chapter on a problem which admits a natural algebraic formulation as a structured system of polynomial equations:

Generalized MinRank Problem: given a field $\mathbb{K}$, a $p \times q$ matrix $\mathscr{M}$ whose entries are polynomials of degree D in $\mathbb{K}[x_1, \ldots, x_n]$, and $r < \min(p, q)$ an integer, compute the set of points at which the evaluation of the matrix has rank at most r.

Being able to estimate precisely the complexity of this problem (which is known to be NP-complete when $\mathbb{K}$ is a finite field [BFS99]) is of first importance for applications. In Cryptology, the security of several multivariate cryptosystems relies on the difficulty of solving the classical MinRank problem (i.e. when the entries of the matrix are linear [KS99, FLP08], see Section 8.2). In coding theory, rank-metric codes can be decoded by computing the set of points where a polynomial matrix has rank less than a given value [OJ02, FLP08]. Also, in Geometry and Optimization the critical points of a map are defined by the rank defect of its Jacobian matrix (see Chapter 5). Moreover, this problem also underlies other problems from Symbolic Computation (for instance solving multi-homogeneous systems, see e.g. [FSS11a]).

To study the Generalized MinRank problem, we consider the algebraic system of all $(r + 1)$-minors of the input matrix. Indeed, these minors simultaneously vanish on the locus of rank defect and hence give rise to a *determinantal ideal*.

Several tools can be used to solve this algebraic system by taking profit of the underlying structure. For instance, the geometric resolution [GLS01] can use the fact that these systems can be evaluated efficiently. Also, recent works on homotopy methods show that numerical algorithms can solve determinantal problems efficiently [Ver99]. The goal of this chapter is to show that Gröbner bases algorithms also greatly benefit from the combinatorial structure underlying determinantal ideals.

In this chapter, an algebraic representation of the locus of rank defect is obtained by computing a lexicographical Gröbner basis with the algorithms F_5 [Fau02] and FGLM [FGLM93]. Indeed, experiments suggest that these algorithms take profit of the determinantal structure. The aim of this work is to give an explanation of this behavior from the viewpoint of asymptotic complexity analysis.

Main results

The goal of this chapter is to obtain complexity bounds for Gröbner bases algorithms when the input system is the set of $(r+1)$-minors of a $p \times q$ matrix $\mathscr{M}$, whose entries are polynomials of degree D with generic coefficients. By *generic*, we mean that there exists a non-identically null multivariate polynomial h such that the complexity results holds when this polynomial does not vanish on the coefficients of the polynomials in the matrix. Therefore, from a practical viewpoint, the complexity bounds can be used in applications where the cardinality of the base field $\mathbb{K}$ is large enough: in that case, the probability that the coefficients of $\mathscr{M}$ do not belong to the zero set of h is close to 1.

We start by studying the homogeneous generalized MinRank problem (i.e. when the entries of $\mathscr{M}$ are homogeneous polynomials) and by proving an explicit formula for the Hilbert series of the ideal $\mathscr{I}_r$ generated by the $(r+1)$-minors of the matrix $\mathscr{M}$. The general framework of the proofs is the following: we consider the ideal $\mathscr{D}_r \subset \mathbb{K}[U]$ generated by the $(r+1)$-minors of a matrix $\mathscr{U} = (u_{i,j})$ whose entries are variables. Then we consider the ideal $\widetilde{\mathscr{D}_r} = \mathscr{D}_r + \langle g_1, \dots, g_{pq}\rangle \subset \mathbb{K}[U, X]$, where the polynomials g_i are quasi-homogeneous forms that are the sum of a linear form in $\mathbb{K}[U]$ and a homogeneous polynomial of degree D in $\mathbb{K}[X]$. If some conditions on the g_i are verified, by performing a linear combination of the generators there exist $f_{1,1}, \dots, f_{p,q} \in \mathbb{K}[X]$ such that

$$\widetilde{\mathscr{D}_r} = \mathscr{D}_r + \langle u_{1,1} - f_{1,1}, \dots, u_{p,q} - f_{p,q}\rangle.$$

Then we use the fact that $(\mathscr{D}_r + \langle u_{1,1} - f_{1,1}, \dots, u_{p,q} - f_{p,q}\rangle) \cap \mathbb{K}[X] = \mathscr{I}_r$ to prove that properties of $\mathscr{D}_r$ transfer to $\mathscr{I}_r$ when the entries of the matrix $\mathscr{M}$ are generic. This allows us to use results known about the ideal $\mathscr{D}_r$ to study the algebraic structure of $\mathscr{I}_r$.

We study separately three different cases:

- $n > (p-r)(q-r)$. Under genericity assumptions on the input, the dimension of the set of solutions of the generalized MinRank problem is positive.
- $n = (p-r)(q-r)$. This is the 0-dimensional case, where the problem has finitely-many solutions under genericity assumptions.
- $n < (p-r)(q-r)$. In the over-determined case, we need to assume that a variant of Fröberg's Conjecture holds in order to generalize the results in [FSS10].

In particular, when $n \geq (p-r)(q-r)$, we prove that the Hilbert series of the quotient ring $\mathbb{K}[X]/\mathscr{I}_r$ is the power series expansion of the rational function

$$\mathsf{HS}_{\mathbb{K}[X]/\mathscr{I}_r}(t) = \frac{\det A_r^{p,q}(t^D)(1-t^D)^{(p-r)(q-r)}}{t^{D\binom{r}{2}}(1-t)^n},$$

where $A_r^{p,q}(t)$ is the $r \times r$ matrix whose (i,j)-entry is $\sum_k \binom{p-i}{k}\binom{q-j}{k} t^k$. Assuming w.l.o.g. that $q \leq p$, we also prove that the degree of $\mathscr{I}_r$ is equal to

$$\mathsf{DEG}(\mathscr{I}_r) = D^{(p-r)(q-r)} \prod_{i=0}^{q-r-1} \frac{i!(p+i)!}{(q-1-i)!(p-r+i)!}.$$

From these explicit formulas, complexity bounds can be deduced. Indeed, one way to get a representation of the solutions of the problem in the 0-dimensional case is to compute a *lexicographical* Gröbner basis of the ideal generated by the minors. As shown in Chapter 1, this can be achieved by using first the F_5 algorithm [Fau02] to compute a Gröbner basis for the so-called *grevlex* ordering and then use the FGLM algorithm [FGLM93] to convert it into a *lexicographical* Gröbner basis. The complexities of these algorithms are respectively governed by the degree of regularity and by the degree of the ideal.

Therefore the theoretical results on the structure of $\mathscr{I}_r$ yield bounds on the complexity of solving the generalized MinRank problem with Gröbner bases algorithms. More precisely, when $n = (p-r)(q-r)$ and under genericity assumptions on the input polynomial matrix, we prove that the arithmetic complexity for computing a lexicographical Gröbner basis of $\mathscr{I}_r$ is bounded above by

$$O\left(\binom{p}{r+1}\binom{q}{r+1}\binom{\mathrm{d_{reg}}(\mathscr{I}_r)+n}{n}^{\omega} + n\left(\mathsf{DEG}\left(\mathscr{I}_r\right)\right)^3\right),$$

where $2 \le \omega \le 3$ is a feasible exponent for the matrix multiplication, and

$$\mathrm{d_{reg}}(\mathscr{I}_r) = Dr(q-r) + (D-1)n + 1.$$

This complexity bound allows us to identify families of Generalized MinRank problems for which the number of arithmetic operations during the Gröbner basis computations is polynomial in the number of solutions.

In the over-determined case (i.e. $n < (p-r)(q-r)$), we obtain similar complexity results, by assuming a variant of Fröberg's Conjecture which is supported by experiments.

Organization of the chapter

Section 4.2 provides notations used throughout this chapter and preliminary results. In Section 4.3, we show how properties of the ideal $\mathscr{D}_r$ generated by the $(r+1)$-minors of $\mathscr{U}$ transfer to the ideal $\mathscr{I}_r$. Then, the case when the homogeneous Generalized MinRank Problem has non-trivial solutions (under genericity assumptions) is studied in Section 4.4. Section 4.5 is devoted to the study of the over-determined MinRank Problem (i.e. when $n < (p-r)(q-r)$). Then, the complexity analysis is performed in Section 4.6. Some consequences of this complexity analysis are drawn in Section 4.7. Experimental results are given in Section 4.7.4

4.2 Notations and preliminaries

In the sequel, p, q, r and n and D are positive integers with $r < q \le p$. For $d \in \mathbb{N}$, $\mathsf{Monomials}(\mathbb{K}[X], d)$ denotes the set of monomials of degree d in the polynomial ring $\mathbb{K}[X] = \mathbb{K}[x_1, \ldots, x_n]$. Its cardinality is $\#\,\mathsf{Monomials}(\mathbb{K}[X], d) = \binom{d-1+n}{d}$.

We denote by $\mathfrak{a}$ the set of parameters $\{\mathfrak{a}_t^{(i,j)} : 1 \le i \le p, 1 \le j \le q, t \in \mathsf{Monomials}(\mathbb{K}[X], d)\}$. The set of variables $\{u_{i,j} : 1 \le i \le p, 1 \le j \le q\}$ is denoted by U.

For $1 \le i \le p, 1 \le j \le q$, we denote by $f_{i,j} \in \mathbb{K}(\mathfrak{a})[X]$ the generic form of degree D

$$f_{i,j} = \sum_{t \in \mathsf{Monomials}(\mathbb{K}[X], D)} \mathfrak{a}_t^{(i,j)} t.$$

Let $\mathscr{I}_r \subset \mathbb{K}(\mathfrak{a})[X]$ be the ideal generated by the $(r+1)$-minors of the $p \times q$ matrix

$$\mathscr{M} = \begin{pmatrix} f_{1,1} & \cdots & f_{1,q} \\ \vdots & \ddots & \vdots \\ f_{p,1} & \cdots & f_{p,q} \end{pmatrix},$$

and $\mathscr{D}_r \subset \mathbb{K}(\mathfrak{a})[U, X]$ be the determinantal ideal generated by the $(r+1)$-minors of the matrix

$$\mathscr{U} = \begin{pmatrix} u_{1,1} & \dots & u_{1,q} \\ \vdots & \ddots & \vdots \\ u_{p,1} & \dots & u_{p,q} \end{pmatrix}.$$

We define $\widetilde{\mathscr{I}_r}$ as the quasi-homogeneous ideal $\mathscr{D}_r + \langle u_{i,j} - f_{i,j}\rangle_{1\leq i\leq p, 1\leq j\leq q} \subset \mathbb{K}(\mathfrak{a})[U, X]$: the quasi-homogeneous grading is given by $\mathrm{wdeg}(x_i) = 1$, $\mathrm{wdeg}(u_{i,j}) = D$. Notice that $\widetilde{\mathscr{I}_r} = \mathscr{I}_r + \langle u_{i,j} - f_{i,j}\rangle_{1\leq i\leq p, 1\leq j\leq q} \subset \mathbb{K}(\mathfrak{a})[U, X]$. Therefore, $\mathscr{I}_r = \widetilde{\mathscr{I}_r} \cap \mathbb{K}(\mathfrak{a})[X]$.

4.3 Transferring determinantal properties

In this section, we prove that generic structural properties (such as the dimension, the structure of the leading monomial ideal,...) of the ideal $\widetilde{\mathscr{I}_r}$ are the same as properties of the ideal $\mathscr{D}_r$ where several generic forms have been added. Hence several classical properties of the determinantal ideal $\mathscr{D}_r$ transfer to the ideal $\widetilde{\mathscr{I}_r}$. In particular, this technique permits to obtain an explicit formula of the Hilbert series of the ideal $\widetilde{\mathscr{I}_r}$.

In the following, we let $\mathfrak{b}$ and $\mathfrak{c}$ denote the following sets of parameters:

$$\begin{array}{rcl} \mathfrak{b} & = & \{\mathfrak{b}_t^{(\ell)} \mid t \in \mathsf{Monomials}(\mathbb{K}[X], D), 1 \leq \ell \leq pq\}; \\ \mathfrak{c} & = & \{\mathfrak{c}_{i,j}^{(\ell)} \mid 1 \leq i \leq p, 1 \leq j \leq q, 1 \leq \ell \leq pq\}. \end{array}$$

Also, $g_1, \dots, g_{pq} \in \mathbb{K}(\mathfrak{b}, \mathfrak{c})[U, X]$ are generic quasi-homogeneous forms of type $(D, 1)$ and of weight degree D:

$$g_\ell = \sum_{t\in\mathsf{Monomials}(\mathbb{K}[X],D)} \mathfrak{b}_t^{(\ell)} t + \sum_{\substack{1\leq i\leq p\\1\leq j\leq q}} \mathfrak{c}_{i,j}^{(\ell)} u_{i,j}.$$

We let $\widetilde{\mathscr{D}_r}$ denote the ideal $\mathscr{D}_r + \langle g_1, \dots, g_{pq}\rangle \subset \mathbb{K}(\mathfrak{b}, \mathfrak{c})[U, X]$. For $\mathbf{a} \in \overline{\mathbb{K}}^{pq\binom{D-1+n}{D}}$, we denote by $\varphi_{\mathbf{a}}$ the following evaluation morphism:

$$\begin{array}{rccc} \varphi_{\mathbf{a}}: & \mathbb{K}[\mathfrak{a}] & \longrightarrow & \overline{\mathbb{K}} \\ & f(\mathfrak{a}) & \longmapsto & f(\mathbf{a}) \end{array}$$

Also, for $(\mathbf{b}, \mathbf{c}) \in \overline{\mathbb{K}}^{pq(\binom{D-1+n}{D}+pq)}$, we denote by $\psi_{\mathbf{b},\mathbf{c}}$ the evaluation morphism:

$$\begin{array}{rccc} \psi_{\mathbf{b},\mathbf{c}}: & \mathbb{K}[\mathfrak{b}, \mathfrak{c}] & \longrightarrow & \overline{\mathbb{K}} \\ & f(\mathfrak{b}, \mathfrak{c}) & \longmapsto & f(\mathbf{b}, \mathbf{c}) \end{array}$$

By abuse of notation, we let $\varphi_{\mathbf{a}}(\widetilde{\mathscr{I}_r})$ (resp. $\psi_{\mathbf{b},\mathbf{c}}(\widetilde{\mathscr{D}_r})$) denote the ideal $\mathscr{D}_r + \langle u_{i,j} - \varphi_{\mathbf{a}}(f_{i,j})\rangle \subset \overline{\mathbb{K}}[U, X]$ (resp. $\mathscr{D}_r + \langle \psi_{\mathbf{b},\mathbf{c}}(g_1), \dots, \psi_{\mathbf{b},\mathbf{c}}(g_{pq})\rangle \subset \overline{\mathbb{K}}[U, X]$).

We call *property* a map from the set of ideals of $\overline{\mathbb{K}}[U, X]$ to $\{\mathtt{true}, \mathtt{false}\}$:

$$\mathscr{P}: \quad \mathsf{Ideals}(\overline{\mathbb{K}}[U, X]) \quad \to \quad \{\mathtt{true}, \mathtt{false}\}.$$

Definition 4.1. *Let $\mathscr{P}$ be a property. We say that $\mathscr{P}$ is*

- $\widetilde{\mathscr{I}_r}$*-generic if there exists a non-empty Zariski open subset* $O \subset \overline{\mathbb{K}}^{pq\binom{D-1+n}{D}}$ *such that*

$$\mathbf{a} \in O \Rightarrow \mathscr{P}\left(\varphi_{\mathbf{a}}\left(\widetilde{\mathscr{I}_r}\right)\right) = \texttt{true};$$

- $\widetilde{\mathscr{D}_r}$*-generic if there exists a non-empty Zariski open subset* $O \subset \overline{\mathbb{K}}^{pq\left(\binom{D-1+n}{D}+pq\right)}$ *such that*

$$(\mathbf{b}, \mathbf{c}) \in O \Rightarrow \mathscr{P}\left(\psi_{\mathbf{b},\mathbf{c}}\left(\widetilde{\mathscr{D}_r}\right)\right) = \texttt{true}.$$

The following lemma is the main result of this section:

Lemma 4.2. *A property* $\mathscr{P}$ *is* $\widetilde{\mathscr{I}_r}$*-generic if and only if it is* $\widetilde{\mathscr{D}_r}$*-generic.*

Proof. To obtain a representation of $\varphi_{\mathbf{a}}\left(\widetilde{\mathscr{D}_r}\right)$ for a generic $\mathbf{a}$ as a specialization of $\widetilde{\mathscr{I}_r}$, it is sufficient to perform a linear combination of the generators. This proof shows that genericity is preserved during this linear transform.

In the sequel we denote by $\mathfrak{A}$, $\mathfrak{B}$ and $\mathfrak{C}$ the following matrices (of respective sizes $pq \times \binom{D-1+n}{D}$, $pq \times \binom{D-1+n}{D}$ and $pq \times pq$):

$$\begin{aligned}
\mathfrak{A} &= \begin{pmatrix} \mathfrak{a}^{(1)}_{x_1^D} & \mathfrak{a}^{(1)}_{x_1^{D-1}x_2} & \cdots & \mathfrak{a}^{(1)}_{x_n^D} \\ \vdots & \vdots & \vdots & \vdots \\ \mathfrak{a}^{(pq)}_{x_1^D} & \mathfrak{a}^{(pq)}_{x_1^{D-1}x_2} & \cdots & \mathfrak{a}^{(pq)}_{x_n^D} \end{pmatrix} \\
\mathfrak{B} &= \begin{pmatrix} \mathfrak{b}^{(1)}_{x_1^D} & \mathfrak{b}^{(1)}_{x_1^{D-1}x_2} & \cdots & \mathfrak{b}^{(1)}_{x_n^D} \\ \vdots & \vdots & \vdots & \vdots \\ \mathfrak{b}^{(pq)}_{x_1^D} & \mathfrak{b}^{(pq)}_{x_1^{D-1}x_2} & \cdots & \mathfrak{b}^{(pq)}_{x_n^D} \end{pmatrix} \\
\mathfrak{C} &= \begin{pmatrix} \mathfrak{c}^{(1)}_{1,1} & \cdots & \mathfrak{c}^{(1)}_{p,q} \\ \vdots & \vdots & \vdots \\ \mathfrak{c}^{(pq)}_{1,1} & \cdots & \mathfrak{c}^{(pq)}_{p,q} \end{pmatrix}.
\end{aligned}$$

Therefore, we have

$$\begin{aligned}
\begin{pmatrix} u_{1,1} - f_{1,1} \\ \vdots \\ u_{p,q} - f_{p,q} \end{pmatrix} &= \mathsf{Id}_{pq} \cdot \begin{pmatrix} u_{1,1} \\ \vdots \\ u_{p,q} \end{pmatrix} - \mathfrak{A} \cdot \begin{pmatrix} x_1^D \\ x_1^{D-1}x_2 \\ \vdots \\ x_n^D \end{pmatrix} \\
\begin{pmatrix} g_1 \\ \vdots \\ g_{pq} \end{pmatrix} &= \mathfrak{C} \cdot \begin{pmatrix} u_{1,1} \\ \vdots \\ u_{p,q} \end{pmatrix} + \mathfrak{B} \cdot \begin{pmatrix} x_1^D \\ x_1^{D-1}x_2 \\ \vdots \\ x_n^D \end{pmatrix}
\end{aligned}$$

In this proof, for $\mathbf{a} \in \mathbb{K}^{pq\binom{D-1+n}{D}}$ (resp. $\mathbf{b} \in \mathbb{K}^{pq\binom{D-1+n}{D}}$, $\mathbf{c} \in \mathbb{K}^{p^2q^2}$), the notation $\mathbf{A}$ (resp. $\mathbf{B}, \mathbf{C}$) stands for the evaluation of the matrix $\mathfrak{A}$ (resp. $\mathfrak{B}, \mathfrak{C}$) at $\mathbf{a}$ (resp. $\mathbf{b}, \mathbf{c}$). Also, we implicitly identify $\mathbf{A}$ with $\mathbf{a}$ (resp. $\mathbf{B}$ with $\mathbf{b}$, $\mathbf{C}$ with $\mathbf{c}$, $\mathfrak{A}$ with $\mathfrak{a}$, $\mathfrak{B}$ with $\mathfrak{b}$, $\mathfrak{C}$ with $\mathfrak{c}$).

- Let $\mathscr{P}$ be a $\widetilde{\mathscr{I}_r}$-generic property. Thus there exists a non-zero polynomial $h_1(\mathfrak{A}) \in \overline{\mathbb{K}}[\mathfrak{a}]$ such that if $h_1(\mathbf{A}) \neq 0$ then $\mathscr{P}\left(\varphi_{\mathbf{a}}(\widetilde{\mathscr{I}_r})\right) = \mathtt{true}$.

  Let $\mathsf{adj}(\mathfrak{C})$ denote the adjugate of $\mathfrak{C}$ (i.e. $\mathsf{adj}(\mathfrak{C}) = \det(\mathfrak{C}) \cdot \mathfrak{C}^{-1}$ in $\mathbb{K}(\mathfrak{c})$). Consider the polynomial $\widetilde{h_1}$ defined by $\widetilde{h_1}(\mathfrak{B}, \mathfrak{C}) = h_1(-\mathsf{adj}(\mathfrak{C}) \cdot \mathfrak{B}) \in \overline{\mathbb{K}}[\mathfrak{b}, \mathfrak{c}]$. The polynomial inequality $\det(\mathfrak{C})\widetilde{h_1}(\mathfrak{B}, \mathfrak{C}) \neq 0$ defines a non-empty Zariski open subset $O \subset \overline{\mathbb{K}}^{pq(\binom{D-1+n}{D}+pq)}$. Let $(\mathbf{B}, \mathbf{C}) \in O$ be an element in this set, then $\mathbf{C}$ is invertible since $\det(\mathbf{C}) \neq 0$. Let $\widetilde{\mathbf{A}}$ be the matrix $\widetilde{\mathbf{A}} = -\mathsf{adj}(\mathbf{C}) \cdot \mathbf{B}$. Therefore the generators of the ideal $\varphi_{\widetilde{\mathbf{a}}}\left(\widetilde{\mathscr{I}_r}\right)$ are an invertible linear combination of the generators of $\psi_{\mathbf{b},\mathbf{c}}\left(\widetilde{\mathscr{D}_r}\right)$. Consequently, $\varphi_{\widetilde{\mathbf{a}}}\left(\widetilde{\mathscr{I}_r}\right) = \psi_{\mathbf{b},\mathbf{c}}\left(\widetilde{\mathscr{D}_r}\right)$. Moreover, $h_1(\widetilde{\mathbf{A}}) = \widetilde{h_1}(\mathbf{B}, \mathbf{C}) \neq 0$ implies that the polynomial $\widetilde{h_1}$ is not identically 0. Therefore,

$$\forall (\mathbf{b}, \mathbf{c}) \in O, \mathscr{P}\left(\psi_{\mathbf{b},\mathbf{c}}\left(\widetilde{\mathscr{D}_r}\right)\right) = \mathscr{P}\left(\varphi_{\widetilde{\mathbf{a}}}\left(\widetilde{\mathscr{I}_r}\right)\right) = \mathtt{true},$$

  and hence $\mathscr{P}$ is a $\widetilde{\mathscr{D}_r}$-generic property.

- Conversely, consider a $\widetilde{\mathscr{D}_r}$-generic property $\mathscr{P}$. Thus, there exists a non-zero polynomial $h_2(\mathfrak{B}, \mathfrak{C}) \in \overline{\mathbb{K}}[\mathfrak{b}, \mathfrak{c}]$ such that if $h_2(\mathbf{b}, \mathbf{c}) \neq 0$ then $\mathscr{P}\left(\psi_{\mathbf{b},\mathbf{c}}(\widetilde{\mathscr{D}_r})\right) = \mathtt{true}$. Since $\mathscr{P}$ is $\widetilde{\mathscr{D}_r}$-generic, there exists $(\mathbf{b}, \mathbf{c})$ such that $h_2(\mathbf{b}, \mathbf{c}) \det(\mathbf{c}) \neq 0$. Let $\widetilde{h_2}$ be the polynomial $\widetilde{h_2}(\mathfrak{b}) = h_2(-\mathbf{C} \cdot \mathfrak{B}, \mathbf{C})$.

  Since $\det(\mathbf{C}) \neq 0$, the matrix $\mathbf{C}$ is invertible and $\widetilde{h_2}(-\mathbf{C}^{-1} \cdot \mathbf{B}) = h_2(\mathbf{B}, \mathbf{C}) \neq 0$ and hence the polynomial $\widetilde{h_2}$ is not identically 0. Moreover, if $\mathbf{a} \in \mathbb{K}^{pq\binom{D-1+n}{D}}$ is such that $\widetilde{h_2}(\mathbf{A}) \neq 0$, then $h_2(-\mathbf{C} \cdot \mathbf{A}, \mathbf{C}) \neq 0$ and thus $\mathscr{P}\left(\psi_{-\mathbf{C}\cdot\mathbf{A},\mathbf{C}}(\widetilde{\mathscr{D}_r})\right) = \mathtt{true}$. Finally, $\psi_{-\mathbf{C}\cdot\mathbf{A},\mathbf{C}}(\widetilde{\mathscr{D}_r}) = \varphi_{\mathbf{A}}(\widetilde{\mathscr{I}_r})$ since the generators of $\psi_{-\mathbf{C}\cdot\mathbf{A},\mathbf{C}}(\widetilde{\mathscr{D}_r})$ are an invertible linear combination of that of $\varphi_{\mathbf{a}}(\widetilde{\mathscr{I}_r})$ (the linear transformation is given by the invertible matrix $\mathbf{C}$) and hence they generate the same ideal. Therefore, the property $\mathscr{P}$ is $\widetilde{\mathscr{I}_r}$-generic.

$\square$

In the sequel, $\prec$ is an admissible monomial ordering (see Definition 1.18). If I is an ideal of $\mathbb{K}[U, X]$, $\mathbb{K}(\mathfrak{a})[U, X]$, or $\mathbb{K}(\mathfrak{b}, \mathfrak{c})[U, X]$, we let $\mathsf{LM}_\prec(I)$ denote the ideal generated by the leading monomials of the polynomials.

By slight abuse of notation, if I_1 and I_2 are ideals of $\mathbb{K}[U, X]$, $\mathbb{K}(\mathfrak{a})[U, X]$, or $\mathbb{K}(\mathfrak{b}, \mathfrak{c})[U, X]$ (I_1 and I_2 are not necessarily ideals of the same ring), we write $\mathsf{LM}_\prec(I_1) = \mathsf{LM}_\prec(I_2)$ if the sets $\{\mathsf{LM}_\prec(f) \mid f \in I_1\}$ and $\{\mathsf{LM}_\prec(f) \mid f \in I_2\}$ are equal.

Lemma 4.3. *Let $\mathscr{P}_{\widetilde{\mathscr{I}_r}}$ and $\mathscr{P}_{\widetilde{\mathscr{D}_r}}$ be the properties defined by*

$$\mathscr{P}_{\widetilde{\mathscr{I}_r}}(I) = \begin{cases} \mathtt{true} \textit{ if } \mathsf{LM}_\prec(I) = \mathsf{LM}_\prec\left(\widetilde{\mathscr{I}_r}\right); \\ \mathtt{false} \textit{ otherwise.} \end{cases}$$

$$\mathscr{P}_{\widetilde{\mathscr{D}_r}}(I) = \begin{cases} \mathtt{true} \textit{ if } \mathsf{LM}_\prec(I) = \mathsf{LM}_\prec\left(\widetilde{\mathscr{D}_r}\right); \\ \mathtt{false} \textit{ otherwise.} \end{cases}$$

Then $\mathscr{P}_{\widetilde{\mathscr{I}_r}}$ (resp. $\mathscr{P}_{\widetilde{\mathscr{D}_r}}$) is a $\widetilde{\mathscr{I}_r}$-generic (resp. $\widetilde{\mathscr{D}_r}$-generic) property.

Proof. We prove here that $\mathscr{P}_{\widetilde{\mathscr{I}_r}}$ is $\widetilde{\mathscr{I}_r}$-generic (the proof for $\mathscr{P}_{\widetilde{\mathscr{D}_r}}$ is similar).

The outline of this proof is the following: during the computation of a Gröbner basis G of $\widetilde{\mathscr{I}_r}$ in $\mathbb{K}(\mathfrak{a})[U, X]$ (for instance with Buchberger's algorithm), a finite number of polynomials are constructed. Let $\varphi_{\mathbf{a}}$ be a specialization. If the images by $\varphi_{\mathbf{a}}$ of the leading coefficients of all non-zero polynomials arising during the computation do not vanish, then $\varphi_{\mathbf{a}}(G) \subset \varphi_{\mathbf{a}}(\widetilde{\mathscr{I}_r})$ is a Gröbner basis of the ideal it generates. It remains to prove that $\varphi_{\mathbf{a}}(G)$ is a Gröbner basis of $\varphi_{\mathbf{a}}(\widetilde{\mathscr{I}_r})$. This is achieved by showing that generically, the normal form (with respect to $\varphi_{\mathbf{a}}(G)$) of the generators of $\varphi_{\mathbf{a}}(\widetilde{\mathscr{I}_r})$ is equal to zero.

For polynomials f_1, f_2, we let $\mathsf{LC}(f_1)$ (resp. $\mathsf{LC}(f_2)$) denote the leading coefficient of f_1 (resp. f_2) and $\mathsf{Spol}(f_1, f_2) = \frac{\mathrm{LCM}(\mathsf{LM}_\prec(f_1), \mathsf{LM}_\prec(f_2))}{\mathsf{LC}(f_1)\,\mathsf{LM}_\prec(f_1)} f_1 - \frac{\mathrm{LCM}(\mathsf{LM}_\prec(f_1), \mathsf{LM}_\prec(f_2))}{\mathsf{LC}(f_2)\,\mathsf{LM}_\prec(f_2)} f_2$ denote the *S-polynomial* of f_1 and f_2.

We prove first that there exists a non-empty Zariski open subset $O_1 \subset \overline{\mathbb{K}}^{pq\binom{D-1+n}{D}}$ such that

$$\mathbf{a} \in O_1 \Rightarrow \mathsf{LM}_\prec(\varphi_{\mathbf{a}}(\widetilde{\mathscr{I}_r})) = \mathsf{LM}_\prec(\widetilde{\mathscr{I}_r}).$$

To do so, consider a Gröbner basis $G \subset \mathbb{K}(\mathfrak{a})[U, X]$ of $\widetilde{\mathscr{I}_r}$ such that each polynomial g can be written as a combination $g = \sum h_\ell f_\ell$, where the f_ℓ's range over the set of minors of size $r+1$ of $\mathscr{U}$ and the polynomials $u_{i,j} - f_{i,j}$, and $h_\ell \in \mathbb{K}[\mathfrak{a}][U, X]$. Buchberger's criterion states that S-polynomials of polynomials in a Gröbner basis reduce to zero [CLO97, Chapter 2, §6, Theorem 6]. Thus each S-polynomial of $g_i, g_j \in G$ can be rewritten as an algebraic combination

$$\mathsf{Spol}(g_i, g_j) = \sum_{\ell=1}^{t} h'_\ell g_\ell,$$

where the polynomials h'_ℓ belong to $\mathbb{K}(\mathfrak{a})[U, X]$ and such that $\{g_1, \ldots, g_t\} \subset G$ and for each $1 \le s \le t$, $\mathsf{LM}_\prec(g_s)$ divides $\mathsf{LM}_\prec(\mathsf{Spol}(g, g') - \sum_{\ell=1}^{s-1} h'_\ell g_\ell)$. Next, consider:

- the product $Q_1(\mathfrak{a}) = \prod_{g\in G} \mathsf{LC}(g)$ of the leading coefficients of the polynomials in the Gröbner basis;
- for all $(g_i, g_j) \in G^2$ such that $\mathsf{Spol}(g_i, g_j) \neq 0$, the product $Q_2(\mathfrak{a})$ of the numerators and denominators of the leading coefficients arising during the reduction of $\mathsf{Spol}(g_i, g_j)$.

These coefficients belong to $\mathbb{K}[\mathfrak{a}]$. Let $Q(\mathfrak{a}) = Q_1(\mathfrak{a})Q_2(\mathfrak{a}) \in \mathbb{K}[\mathfrak{a}]$ denote their product. The inequality $Q(\mathfrak{a}) \neq 0$ defines a non-empty Zariski open subset $O_1 \subset \overline{\mathbb{K}}^{pq\binom{D-1+n}{D}}$. If $\mathbf{a} \in O_1$, then

$$\varphi_{\mathbf{a}}(\mathsf{Spol}(g, g')) = \sum_{\ell=1}^{t} \varphi_{\mathbf{a}}(h'_\ell)\varphi_{\mathbf{a}}(g_\ell),$$

and for each $1 \le i \le t$, $\mathsf{LM}_\prec(\varphi_{\mathbf{a}}(g_i))$ divides $\mathsf{LM}_\prec(\varphi_{\mathbf{a}}(\mathsf{Spol}(g, g')) - \sum_{\ell=1}^{i-1} \varphi_{\mathbf{a}}(h'_\ell)\varphi_{\mathbf{a}}(g_\ell))$. Thus $\varphi_{\mathbf{a}}(G)$ is a Gröbner basis of the ideal it spans. Moreover, $\langle \varphi_{\mathbf{a}}(G)\rangle \subset \varphi_{\mathbf{a}}(\widetilde{\mathscr{I}_r})$.

We prove now that there exists a non-empty Zariski open set where the other inclusion $\varphi_{\mathbf{a}}(\widetilde{\mathscr{I}_r}) \subset \langle \varphi_{\mathbf{a}}(G)\rangle$ holds. Let $\mathrm{NF}_G(\cdot)$ be the normal form associated to this Gröbner basis (as defined as the *remainder of the division by* G in [CLO97, Chapter 2, §6, Proposition 1]). For each generator f of $\widetilde{\mathscr{I}_r}$ (i.e. either a maximal minor of the matrix $\mathscr{U}$, or a polynomial $u_{i,j} - f_{i,j}$), we have $\mathrm{NF}_G(f) = 0$. During the computation of $\mathrm{NF}_G(f)$ by using the division Algorithm in [CLO97, Chapter 2, §3], a finite set of polynomials (in $\mathbb{K}(\mathfrak{a})[U,X]$) is constructed. Let $Q_3^{(f)} \in \mathbb{K}[\mathfrak{a}]$ denote the product of the numerators and denominators of all their nonzero coefficients in $\mathbb{K}(\mathfrak{a})$. Consequently, if $Q_3^{(f)}(\mathbf{a}) \neq 0$, then $\mathrm{NF}_{\varphi_{\mathbf{a}}(G)}(\varphi_{\mathbf{a}}(f)) = 0$ and hence $\varphi_{\mathbf{a}}(f) \in \langle \varphi_{\mathbf{a}}(G)\rangle$. Repeating this operation for all the generators of $\widetilde{\mathscr{I}_r}$ yields a finite set of non-identically null polynomials $Q_3^{(f)} \in \mathbb{K}[\mathfrak{a}]$. Let $Q_4 \in \mathbb{K}[\mathfrak{a}]$ denote their product. Therefore, if $Q_4(\mathbf{a}) \neq 0$, then $\varphi_{\mathbf{a}}(\widetilde{\mathscr{I}_r}) \subset \langle \varphi_{\mathbf{a}}(G)\rangle$.

Finally, consider the non-empty Zariski open subset $O \subset \mathbb{K}^{pq\binom{D+n-1}{D}}$ defined by the inequality $Q_1 \cdot Q_2 \cdot Q_4 \neq 0$. For all $\mathbf{a} \in O$, we have $\varphi_{\mathbf{a}}(\widetilde{\mathscr{I}_r}) = \langle \varphi_{\mathbf{a}}(G)\rangle$.

□

Corollary 4.4. *The leading monomials of $\widetilde{\mathscr{I}_r}$ are the same as that of $\widetilde{\mathscr{D}_r}$:*

$$\mathsf{LM}_{\prec}\left(\widetilde{\mathscr{I}_r}\right) = \mathsf{LM}_{\prec}\left(\widetilde{\mathscr{D}_r}\right).$$

Proof. By Lemmas 4.2 and 4.3, the property $\mathscr{P}_{\widetilde{\mathscr{I}_r}}$ (resp. $\mathscr{P}_{\widetilde{\mathscr{D}_r}}$) is $\widetilde{\mathscr{I}_r}$-generic and $\widetilde{\mathscr{D}_r}$-generic. Since $\mathscr{P}_{\widetilde{\mathscr{D}_r}}$ (resp. $\mathscr{P}_{\widetilde{\mathscr{I}_r}}$) is $\widetilde{\mathscr{D}_r}$-generic, there exists a non-empty Zariski open subset $O_1 \subset \overline{\mathbb{K}}^{pq(\binom{D-1+n}{D})+pq)}$ (resp. $O_2 \subset \overline{\mathbb{K}}^{pq(\binom{D-1+n}{D})+pq)}$) such that, for $(\mathbf{b},\mathbf{c}) \in O_1$ (resp. O_2), $\mathsf{LM}_{\prec}\left(\psi_{(\mathbf{b},\mathbf{c})}(\widetilde{\mathscr{D}_r})\right) = \mathsf{LM}_{\prec}\left(\widetilde{\mathscr{D}_r}\right)$ (resp. $\mathsf{LM}_{\prec}\left(\psi_{(\mathbf{b},\mathbf{c})}(\widetilde{\mathscr{D}_r})\right) = \mathsf{LM}_{\prec}\left(\widetilde{\mathscr{I}_r}\right)$).

Notice that $O_1 \cap O_2$ is not empty, since for the Zariski topology, the intersection of finitely-many non-empty open subsets is non-empty. Let $(\mathbf{b},\mathbf{c})$ be an element of $O_1 \cap O_2$. Then

$$\mathsf{LM}_{\prec}\left(\widetilde{\mathscr{I}_r}\right) = \mathsf{LM}_{\prec}\left(\psi_{(\mathbf{b},\mathbf{c})}(\widetilde{\mathscr{D}_r})\right) = \mathsf{LM}_{\prec}\left(\widetilde{\mathscr{D}_r}\right).$$

□

Corollary 4.5. *The weighted Hilbert series of $\widetilde{\mathscr{I}_r}$ is the same as that of $\widetilde{\mathscr{D}_r}$.*

Proof. It is well-known that, for any positively graded ideal I and for any monomial ordering, $\mathsf{wHS}_I(t) = \mathsf{wHS}_{\mathsf{LM}_{\prec}(I)}(t)$ (see e.g. the proof of [CLO97, Chapter 9, §3, Proposition 9] which works similarly in the case of quasi-homogeneous ideals). By Corollary 4.4, $\mathsf{LM}_{\prec}\left(\widetilde{\mathscr{I}_r}\right) = \mathsf{LM}_{\prec}\left(\widetilde{\mathscr{D}_r}\right)$, which implies that

$$\mathsf{wHS}_{\mathsf{LM}_{\prec}\left(\widetilde{\mathscr{I}_r}\right)}(t) = \mathsf{wHS}_{\mathsf{LM}_{\prec}\left(\widetilde{\mathscr{D}_r}\right)}(t),$$

and hence $\mathsf{wHS}_{\widetilde{\mathscr{I}_r}}(t) = \mathsf{wHS}_{\widetilde{\mathscr{D}_r}}(t)$. □

4.4 The case $n \geq (p-r)(q-r)$

As we will see in the sequel, the Krull dimension of the ring $\mathbb{K}(\mathfrak{a})[X]/\mathscr{I}_r$ is equal to $\max(n-(p-r)(q-r),0)$. This section is devoted to the study of the case $n \geq (p-r)(q-r)$.

We recall that the polynomials g_ℓ are defined by

$$g_\ell = \sum_{t \in \mathsf{Monomials}(\mathbb{K}[X],D)} \mathfrak{b}_t^{(\ell)} t + \sum_{\substack{1\leq i\leq p\\ 1\leq j\leq q}} \mathfrak{c}_{i,j}^{(\ell)} u_{i,j}.$$

Lemma 4.6. *Let $1 \leq \ell \leq pq$ be an integer. If the polynomial g_ℓ divides zero in the ring $\mathbb{K}(\mathfrak{b},\mathfrak{c})[U,X]/(\mathscr{D}_r + \langle g_1,\ldots,g_{\ell-1}\rangle)$, then there exists a prime ideal P associated to $\mathscr{D}_r + \langle g_1,\ldots,g_{\ell-1}\rangle$ such that* $\dim(P) = 0$.

Proof. If g_ℓ divides zero in $\mathbb{K}(\mathfrak{b},\mathfrak{c})[U,X]/\left(\mathscr{D}_r + \langle g_1,\dots,g_{\ell-1}\rangle\right)$, then there exists a prime ideal P associated to $\mathscr{D}_r + \langle g_1,\dots,g_{\ell-1}\rangle$ such that $g_\ell \in P$. For $\ell \leq pq$, let $\mathfrak{b}^{(\leq \ell)}$ and $\mathfrak{c}^{(\leq \ell)}$ denote the sets of parameters

$$\begin{aligned}\mathfrak{b}^{(\leq \ell)} &= \{\mathfrak{b}_t^{(s)} \mid t \in \mathsf{Monomials}(\mathbb{K}[X], D), 1 \leq s \leq \ell\}\\ \mathfrak{c}^{(\leq \ell)} &= \{\mathfrak{c}_{i,j}^{(s)} \mid 1 \leq i \leq p, 1 \leq j \leq q, 1 \leq s \leq \ell\}.\end{aligned}$$

Since $(\mathscr{D}_r + \langle g_1,\dots,g_{\ell-1}\rangle)$ is an ideal of $\mathbb{K}(\mathfrak{b}^{(\leq \ell-1)},\mathfrak{c}^{(\leq \ell-1)})[U,X]$, and P is an associated prime, there exists a Gröbner basis G_P of P (for any monomial ordering $\prec$) which is a finite subset of $\mathbb{K}(\mathfrak{b}^{(\leq \ell-1)},\mathfrak{c}^{(\leq \ell-1)})[U,X]$.

Let $\mathrm{NF}_{P,\prec}(\cdot)$ denote the normal form associated to this Gröbner basis (as defined as the *remainder of the division by* G_P in [CLO97, Chapter 2, §6, Proposition 1]).

Since $g_\ell \in P$, we have $\mathrm{NF}_{P,\prec}(g_\ell) = 0$. By linearity of $\mathrm{NF}_{P,\prec}(\cdot)$, we obtain

$$\sum_{t \in \mathsf{Monomials}(\mathbb{K}[X],D)} \mathfrak{b}_t^{(\ell)} \,\mathrm{NF}_{P,\prec}(t) + \sum_{\substack{1\leq i\leq p\\ 1\leq j\leq q}} \mathfrak{c}_{i,j}^{(\ell)} \,\mathrm{NF}_{P,\prec}(u_{i,j}) = 0.$$

Since $G_p \subset \mathbb{K}(\mathfrak{b}^{(\leq \ell-1)},\mathfrak{c}^{(\leq \ell-1)})[U,X]$, we can deduce that for any monomial t, $\mathrm{NF}_{P,\prec}(t) \in \mathbb{K}(\mathfrak{b}^{(\leq \ell-1)},\mathfrak{c}^{(\leq \ell-1)})[U,X]$. Therefore, by algebraic independence of the parameters, the following properties hold: for all $t \in \mathsf{Monomials}(\mathbb{K}[X],D)$, $\mathrm{NF}_{P,\prec}(t) = 0$, and for all i,j, $\mathrm{NF}_{P,\prec}(u_{i,j}) = 0$. Consequently, all monomials of weight degree D in $\mathbb{K}(\mathfrak{b},\mathfrak{c})[U,X]$ are in P, and hence P has dimension 0. $\square$

Lemma 4.7. *For all $\ell \in \{2,\dots,pq\}$, the polynomial g_ℓ does not divide zero in the ring $\mathbb{K}(\mathfrak{b},\mathfrak{c})[U,X]/(\mathscr{D}_r + \langle g_1,\dots,g_{\ell-1}\rangle)$ and* $\dim(\mathscr{D}_r + \langle g_1,\dots,g_\ell\rangle) = n + (p+q-r)r - \ell$.

Proof. We prove the Lemma by induction on ℓ. According to [HE70, Corollary 2 of Theorem 1], the ring $\mathbb{K}(\mathfrak{b},\mathfrak{c})[U,X]/\mathscr{D}_r$ is Cohen-Macaulay and purely equidimensional. First, notice that the dimension is equal to $n + (p+q-r)r$ for $\ell = 0$ since the dimension of the ideal $\mathscr{D}_r \subset \mathbb{K}[U]$ is $(p+q-r)r$ (see e.g. [CH94] and references therein). Now, suppose that the dimension of the ideal $\mathscr{D}_r + \langle g_1,\dots,g_{\ell-1}\rangle \subset \mathbb{K}(\mathfrak{b},\mathfrak{c})[U,X]$ is $n+(p+q-r)r-\ell+1$. Since the ring $\mathbb{K}(\mathfrak{b},\mathfrak{c})[U,X]/\mathscr{D}_r$ is Cohen-Macaulay and $\langle g_1,\dots,g_{\ell-1}\rangle$ has co-dimension $\ell-1$ in $\mathbb{K}(\mathfrak{b},\mathfrak{c})[U,X]/\mathscr{D}_r$, the Macaulay unmixedness Theorem [Eis95, Corollary 18.14] implies that $\mathscr{D}_r + \langle g_1,\dots,g_{\ell-1}\rangle$ as an ideal in $\mathbb{K}(\mathfrak{b},\mathfrak{c})[U,X]$ has no embedded component and is equidimensional. By contradiction, suppose now that g_ℓ divides zero in $\mathbb{K}(\mathfrak{b},\mathfrak{c})[U,X]/(\mathscr{D}_r + \langle g_1,\dots,g_{\ell-1}\rangle)$. By Lemma 4.6, there exists a prime P associated to $\mathscr{D}_r + \langle g_1,\dots,g_{\ell-1}\rangle$ such that $\dim(P) = 0$, which contradicts the fact that $\mathscr{D}_r + \langle g_1,\dots,g_{\ell-1}\rangle$ is purely equidimensional of dimension $n + (p+q-r)r - \ell + 1 > 0$. $\square$

Lemma 4.8. *The Hilbert series $\mathsf{HS}_{\mathbb{K}(\mathfrak{a})[X]/\mathscr{I}_r}(t)$ is equal to the weighted Hilbert series $\mathsf{wHS}_{\mathbb{K}(\mathfrak{a})[X,U]/\widetilde{\mathscr{I}_r}}(t)$.*

Proof. Let $\prec_{\mathsf{lex}}$ denote a lexicographical ordering on $\mathbb{K}(\mathfrak{a})[X,U]$ such that $x_k \prec_{\mathsf{lex}} u_{i,j}$ for all k,i,j. By [CLO97, Section 9.3, Proposition 9], $\mathsf{HS}_{\mathbb{K}(\mathfrak{a})[X]/\mathscr{I}_r}(t) = \mathsf{HS}_{\mathbb{K}(\mathfrak{a})[X]/\mathsf{LM}_{\prec_{\mathsf{lex}}}(\mathscr{I}_r)}(t)$ and $\mathsf{wHS}_{\mathbb{K}(\mathfrak{a})[U,X]/\widetilde{\mathscr{I}_r}}(t) = \mathsf{wHS}_{\mathbb{K}(\mathfrak{a})[U,X]/\mathsf{LM}_{\prec_{\mathsf{lex}}}(\widetilde{\mathscr{I}_r})}(t)$. Since $\mathsf{LM}_{\prec_{\mathsf{lex}}}(u_{i,j} - f_{i,j}) = u_{i,j}$, we deduce that all monomials which are multiples of a variable $u_{i,j}$ are in $\mathsf{LM}_{\prec_{\mathsf{lex}}}(\widetilde{\mathscr{I}_r})$. Therefore, the remaining monomials in $\mathsf{LM}_{\prec_{\mathsf{lex}}}(\widetilde{\mathscr{I}_r})$ are in $\mathbb{K}(\mathfrak{a})[X]$:

$$\begin{aligned}\mathsf{LM}_{\prec_{\mathsf{lex}}}(\widetilde{\mathscr{I}_r}) &= \left\langle \{u_{i,j}\} \cup \mathsf{LM}_{\prec_{\mathsf{lex}}}(\widetilde{\mathscr{I}_r} \cap \mathbb{K}(\mathfrak{a})[X]) \right\rangle\\ &= \left\langle \{u_{i,j}\} \cup \mathsf{LM}_{\prec_{\mathsf{lex}}}(\mathscr{I}_r) \right\rangle.\end{aligned}$$

Therefore, $\frac{\mathbb{K}(\mathfrak{a})[U,X]}{\mathsf{LM}_{\prec_{\mathsf{lex}}}(\widetilde{\mathscr{I}_r})}$ is isomorphic (as a graded $\mathbb{K}(\mathfrak{a})$-algebra) to $\frac{\mathbb{K}(\mathfrak{a})[X]}{\mathsf{LM}_{\prec_{\mathsf{lex}}}(\mathscr{I}_r)}$. Thus

$$\mathsf{HS}_{\mathbb{K}(\mathfrak{a})[X]/\,\mathsf{LM}_{\prec_{\mathsf{lex}}}(\mathscr{I}_r)}(t) = \mathsf{wHS}_{\mathbb{K}(\mathfrak{a})[U,X]/\,\mathsf{LM}_{\prec_{\mathsf{lex}}}(\widetilde{\mathscr{I}_r})}(t),$$

and hence

$$\mathsf{HS}_{\mathbb{K}(\mathfrak{a})[X]/\mathscr{I}_r}(t) = \mathsf{wHS}_{\mathbb{K}(\mathfrak{a})[U,X]/\widetilde{\mathscr{I}_r}}(t).$$

□

In the sequel, $A_r^{p,q}(t)$ denotes the $r \times r$ matrix whose (i,j)-entry is $\sum_k \binom{p-i}{k}\binom{q-j}{k} t^k$. The following theorem is the main result of this section:

Theorem 4.9. *The dimension of the ideal $\mathscr{I}_r$ is $n-(p-r)(q-r)$ and its Hilbert series is*

$$\mathsf{HS}_{\mathbb{K}(\mathfrak{a})[X]/\mathscr{I}_r}(t) = \frac{\det\left(A_r^{p,q}(t^D)\right)(1-t^D)^{(p-r)(q-r)}}{t^{D\binom{r}{2}}(1-t)^n}.$$

Proof. According to [CH94, Corollary 1] (and references therein), the ideal $\mathscr{D}_r$ seen as an ideal of $\mathbb{K}[U]$ has dimension $(p+q-r)r$ and its Hilbert series (for the standard gradation: $\deg(u_{i,j})=1$) is the power series expansion of

$$\mathsf{HS}_{\mathbb{K}[U]/\mathscr{D}_r}(t) = \frac{\det A_r^{p,q}(t)}{t^{\binom{r}{2}}(1-t)^{(p+q-r)r}}.$$

By putting a weight D on each variable $u_{i,j}$ (i.e. $\deg(u_{i,j}) = D$), the weighted Hilbert series of $\mathscr{D}_r \subset \mathbb{K}[U]$ is

$$\mathsf{wHS}_{\mathbb{K}[U]/\mathscr{D}_r}(t) = \frac{\det A_r^{p,q}(t^D)}{t^{D\binom{r}{2}}(1-t^D)^{(p+q-r)r}}.$$

By considering $\mathscr{D}_r$ as an ideal of $\mathbb{K}(\mathfrak{b},\mathfrak{c})[U,X]$, the dimension becomes $n+(p+q-r)r$ and its weighted Hilbert series is

$$\mathsf{wHS}_{\mathbb{K}(\mathfrak{b},\mathfrak{c})[U,X]/\mathscr{D}_r}(t) = \frac{\det A_r^{p,q}(t^D)}{t^{D\binom{r}{2}}(1-t)^n(1-t^D)^{(p+q-r)r}}.$$

According to Lemma 4.7, for each $\ell \le pq$, the polynomial g_ℓ does not divide zero in the ring $\mathbb{K}(\mathfrak{b},\mathfrak{c})[U,X]/(\mathscr{D}_r + \langle g_1,\dots,g_{\ell-1}\rangle)$. This implies the following relations:

$$\begin{array}{rcl} \dim(\mathscr{D}_r + \langle g_1,\dots,g_\ell\rangle) &=& \dim(\mathscr{D}_r + \langle g_1,\dots,g_{\ell-1}\rangle) - 1 \\ \mathsf{wHS}_{\mathbb{K}(\mathfrak{b},\mathfrak{c})[U,X]/(\mathscr{D}_r+\langle g_1,\dots,g_\ell\rangle)}(t) &=& (1-t^D)\,\mathsf{wHS}_{\mathbb{K}(\mathfrak{b},\mathfrak{c})/(\mathscr{D}_r+\langle g_1,\dots,g_{\ell-1}\rangle)}(t). \end{array}$$

Therefore the dimension of $\widetilde{\mathscr{D}_r}$ is $n-(p-r)(q-r)$ and its weighted Hilbert series is

$$\mathsf{wHS}_{\mathbb{K}(\mathfrak{b},\mathfrak{c})[U,X]/\widetilde{\mathscr{D}_r}}(t) = \frac{\det\left(A_r^{p,q}(t^D)\right)}{t^{D\binom{r}{2}}(1-t)^n(1-t^D)^{(p+q-r)r-pq}} = \frac{\det\left(A_r^{p,q}(t^D)\right)(1-t^D)^{(p-r)(q-r)}}{t^{D\binom{r}{2}}(1-t)^n}.$$

By Corollary 4.5, the ideal $\widetilde{\mathscr{I}_r}$ has the same weighted Hilbert series. Finally, by Lemma 4.8, the Hilbert series of $\mathscr{I}_r = \widetilde{\mathscr{I}_r} \cap \mathbb{K}(\mathfrak{a})[X]$ is the same as that of $\widetilde{\mathscr{I}_r}$. □

Corollary 4.10. *The degree of the ideal $\mathscr{I}_r$ is:*

$$\begin{aligned}\mathsf{DEG}(\mathscr{I}_r) &= D^{(p-r)(q-r)} \prod_{i=0}^{q-r-1} \frac{i!(p+i)!}{(q-1-i)!(p-r+i)!} \\ &= D^{(p-r)(q-r)} \prod_{i=0}^{q-r-1} \frac{\binom{p+q-r-1}{r+i}}{\binom{p+q-r-1}{i}}.\end{aligned}$$

Proof. From [Ful97, Example 14.4.14], the degree of the ideal $\mathscr{D}_r$ is

$$\prod_{i=0}^{q-r-1} \frac{i!(p+i)!}{(q-1-i)!(p-r+i)!}.$$

Since the degree is equal to the numerator of the Hilbert series of $\mathscr{D}_r$ evaluated at $t=1$,

$$\det A_r^{p,q}(1) = \prod_{i=0}^{q-r-1} \frac{i!(p+i)!}{(q-1-i)!(p-r+i)!}.$$

By Theorem 4.9, the Hilbert series of $\mathscr{I}_r$ is

$$\begin{aligned}\mathsf{HS}_{\mathbb{K}(\mathfrak{a})[X]/\mathscr{I}_r}(t) &= \frac{\det\left(A_r^{p,q}(t^D)\right)(1-t^D)^{(p-r)(q-r)}}{t^{D\binom{r}{2}}(1-t)^n} \\ &= \frac{\det A_r^{p,q}(t^D)(1+t+\cdots+t^{D-1})^{(p-r)(q-r)}}{t^{D\binom{r}{2}}(1-t)^{n-(p-r)(q-r)}}.\end{aligned}$$

Thus, the evaluation of the numerator at $t=1$ yields

$$\mathsf{DEG}(\mathscr{I}_r) = D^{(p-r)(q-r)} \prod_{i=0}^{q-r-1} \frac{i!(p+i)!}{(q-1-i)!(p-r+i)!}.$$

To prove the second equality, notice that

$$\prod_{i=0}^{q-r-1} \frac{\binom{p+q-r-1}{r+i}}{\binom{p+q-r-1}{i}} = \prod_{i=0}^{q-r-1} \frac{i!(p+q-r-i-1)!}{(r+i)!(p+q-2r-i-1)!}.$$

By substituting i by $q-r-1-i$, we obtain that

$$\begin{aligned}\prod_{i=0}^{q-r-1}(p+q-r-i-1)! &= \prod_{i=0}^{q-r-1}(p+i)! \\ \prod_{i=0}^{q-r-1}(r+i)! &= \prod_{i=0}^{q-r-1}(q-i-1)! \\ \prod_{i=0}^{q-r-1}(p+q-2r-i-1)! &= \prod_{i=0}^{q-r-1}(p-r+i)!.\end{aligned}$$

Consequently,

$$\prod_{i=0}^{q-r-1} \frac{i!(p+i)!}{(q-1-i)!(p-r+i)!} = \prod_{i=0}^{q-r-1} \frac{\binom{p+q-r-1}{r+i}}{\binom{p+q-r-1}{i}}.$$

□

4.5 The over-determined case

To study the over-determined case ($n < (p-r)(q-r)$), we need to assume that a determinantal variant of Fröberg's Conjecture holds [Fro85]:

Conjecture 4.11. *Let $\mathscr{D}_{\ell,i}$ denote the vector space of quasi-homogeneous polynomials of weight degree i in $\mathscr{D}_r + \langle g_1, \ldots, g_\ell \rangle$. Then the linear map*

$$\begin{array}{ccc} \mathbb{K}(\mathfrak{b},\mathfrak{c})[U,X]_i/\mathscr{D}_{\ell,i} & \longrightarrow & \mathbb{K}(\mathfrak{b},\mathfrak{c})[U,X]_{i+D}/\mathscr{D}_{\ell,i+D} \\ f & \longmapsto & f g_{\ell+1} \end{array}$$

has maximal rank, i.e. it is either injective or onto.

Remark 4.12. *If $n + (p+q-r)r - \ell > 0$, then Conjecture 4.11 is proved by Lemma 4.7: $g_{\ell+1}$ does not divide zero in $\mathbb{K}(\mathfrak{b},\mathfrak{c})[U,X]/(\mathscr{D}_r + \langle g_1, \ldots, g_\ell \rangle)$ and hence the linear map is injective for all $i \in \mathbb{N}$.*

Notation. Given a power series $S(t) \in \mathbb{Z}[[t]]$, we let $[t^i]S(t)$ denote the coefficient of t^i and $[S(t)]_+$ denote the power series obtained by truncating $S(t)$ at its first non positive coefficient.

Lemma 4.13. *If Conjecture 4.11 is true, then the Hilbert series of $\mathscr{D}_r + \langle g_1, \ldots, g_{\ell+1} \rangle$ is*

$$\mathsf{wHS}_{\mathbb{K}(\mathfrak{b},\mathfrak{c})[U,X]/(\mathscr{D}_r+\langle g_1,\ldots,g_{\ell+1}\rangle)}(t) = \left[(1-t^D)\,\mathsf{wHS}_{\mathbb{K}(\mathfrak{b},\mathfrak{c})[U,X]/(\mathscr{D}_r+\langle g_1,\ldots,g_\ell\rangle)}(t)\right]_+ .$$

Proof. In this proof, for simplicity of notation, we let R denote the ring $\mathbb{K}(\mathfrak{b},\mathfrak{c})[U,X]$. If $S(t) = \sum_{i\in\mathbb{N}} s_i t^i \in \mathbb{Z}[[t]]$ is a power series, $[S(t)]_{\geq 0}$ denotes the series

$$[S(t)]_{\geq 0} = \sum_{i\in\mathbb{N}} \max(s_i, 0) t^i.$$

Let $\operatorname{ann}(g_{\ell+1})$ be the ideal $\{f \in R \mid f g_{\ell+1} \in \mathscr{D}_r + \langle g_1, \ldots, g_\ell \rangle\}$. For $i \in \mathbb{N}$, consider the following exact sequence:

$$\begin{array}{l} 0 \to \operatorname{ann}(g_{\ell+1})_i \to R_i/\mathscr{D}_{\ell,i} \xrightarrow{\times g_{\ell+1}} R_{i+D}/\mathscr{D}_{\ell,i+D} \to \\ \qquad\qquad\qquad\qquad\quad \to R_{i+D}/\mathscr{D}_{\ell+1,i+D} \to 0. \end{array}$$

Conjecture 4.11 states that

$$\dim(\operatorname{ann}(g_{\ell+1})_i) = \max(0, \dim(R_i/\mathscr{D}_{\ell,i}) - \dim(R_{i+D}/\mathscr{D}_{\ell,i+D})).$$

Since the alternate sum of the dimensions of the vector spaces occurring in an exact sequence is zero, it follows that

$$\begin{array}{rcl} \dim(R_{i+D}/\mathscr{D}_{\ell+1,i+D}) & = & \dim(R_{i+D}/\mathscr{D}_{\ell,i+D}) - \dim(R_i/\mathscr{D}_{\ell,i}) + \\ & & \max(0, \dim(R_i/\mathscr{D}_{\ell,i}) - \dim(R_{i+D}/\mathscr{D}_{\ell,i+D})) \\ & = & \max(0, \dim(R_{i+D}/\mathscr{D}_{\ell,i+D}) - \dim(R_i/\mathscr{D}_{\ell,i})). \end{array}$$

Multiplying this equation by t^{i+D} yields

$$\begin{array}{rcl} \left[t^{i+D}\right] \mathsf{wHS}_{\mathbb{K}(\mathfrak{b},\mathfrak{c})[U,X]/(\mathscr{D}_r+\langle g_1,\ldots,g_{\ell+1}\rangle)}(t) & = & \dim\left(R_{i+D}/\mathscr{D}_{\ell+1,i+D})\right) \\ & = & \max\left(0, \dim(R_{i+D}/\mathscr{D}_{\ell,i+D}) - \dim(R_i/\mathscr{D}_{\ell,i})\right) \\ & = & \max\left(0, [t^{i+D}](1-t^D)\,\mathsf{wHS}_{\mathbb{K}(\mathfrak{b},\mathfrak{c})[U,X]/(\mathscr{D}_r+\langle g_1,\ldots,g_\ell\rangle)}(t)\right) \\ & = & [t^{i+D}]\left[(1-t^D)\,\mathsf{wHS}_{\mathbb{K}(\mathfrak{b},\mathfrak{c})[U,X]/(\mathscr{D}_r+\langle g_1,\ldots,g_\ell\rangle)}(t)\right]_{\geq 0}. \end{array}$$

Since any monomial in $\mathbb{K}(\mathfrak{a})[X,U]$ of weight degree greater that D is a multiple of a monomial of weight degree D, we deduce that if there exists $i_0 \geq D$ such that

$$\left[t^{i_0}\right] \mathsf{wHS}_{\mathbb{K}(\mathfrak{b},\mathfrak{c})[U,X]/(\mathscr{D}_r+\langle g_1,\ldots,g_{\ell+1}\rangle)}(t) = 0,$$

then for all $i > i_0$, $\left[t^{i}\right] \mathsf{wHS}_{\mathbb{K}(\mathfrak{b},\mathfrak{c})[U,X]/(\mathscr{D}_r+\langle g_1,\ldots,g_{\ell+1}\rangle)}(t) = 0$. Therefore

$$\left[t^{i+D}\right] \mathsf{wHS}_{\mathbb{K}(\mathfrak{b},\mathfrak{c})[U,X]/(\mathscr{D}_r+\langle g_1,\ldots,g_{\ell+1}\rangle)}(t) = [t^{i+D}] \left[(1-t^D)\, \mathsf{wHS}_{\mathbb{K}(\mathfrak{b},\mathfrak{c})[U,X]/(\mathscr{D}_r+\langle g_1,\ldots,g_{\ell}\rangle)}(t)\right]_+,$$

Finally, by summing over i, we get

$$\mathsf{wHS}_{\mathbb{K}(\mathfrak{b},\mathfrak{c})[U,X]/(\mathscr{D}_r+\langle g_1,\ldots,g_{\ell+1}\rangle)}(t) = \left[(1-t^D)\, \mathsf{HS}_{\mathbb{K}(\mathfrak{b},\mathfrak{c})[U,X]/(\mathscr{D}_r+\langle g_1,\ldots,g_{\ell}\rangle)}(t)\right]_+.$$

□

Theorem 4.14. *If Conjecture 4.11 is true, then the Hilbert series of $\mathscr{I}_r$ is*

$$\mathsf{HS}_{\mathbb{K}(\mathfrak{a})[X]/\mathscr{I}_r}(t) = \left[(1-t^D)^{(p-r)(q-r)} \frac{\det\left(A_r^{p,q}(t^D)\right)}{t^{D\binom{r}{2}}(1-t)^n}\right]_+,$$

where $A_r^{p,q}(t)$ is the $r\times r$ matrix whose (i,j)-entry is $\displaystyle\sum_{k=0}^{\min(p-i,q-j)} \binom{p-i}{k}\binom{q-j}{k} t^k.$

Proof. By applying pq times Lemma 4.13, we obtain that

$$\mathsf{wHS}_{\mathbb{K}(\mathfrak{b},\mathfrak{c})[U,X]/\widetilde{\mathscr{D}_r}}(t) = \left[(1-t^D)\left[(1-t^D)\ldots\left[(1-t^D)\frac{\det A_r^{p,q}(t^D)}{t^{D\binom{r}{2}}(1-t)^n(1-t^D)^{(p+q-r)r}}\right]_+\ldots\right]_+\right]_+.$$

Let $S = \sum_{0\leq i} a_i t^i \in \mathbb{Z}[[t]]$ be a power series such that $a_0 > 0$, and let $i_0 \in \mathbb{N}\cup\{\infty\}$ be defined as

$$i_0 = \begin{cases} \infty \text{ if for all } i\geq 0, a_i > 0;\\ \min(\{i \mid a_i \leq 0\}) \text{ otherwise.}\end{cases}$$

Therefore, $[S(t)]_+ = \sum_{0\leq i<i_0} a_i t^i$. By convention, for $i<0$, we put $a_i = 0$. Then

$$\begin{array}{rcl}(1-t^D)S(t) & = & \sum_{0\leq i}(a_i - a_{i-D})t^i\\ (1-t^D)\,[S(t)]_+ & = & \sum_{0\leq i<i_0}(a_i - a_{i-D})t^i\end{array}.$$

Consequently, the coefficients of $(1-t^D)S(t)$ and of $(1-t^D)\,[S(t)]_+$ are equal up to the index i_0.

- If $i_0 = \infty$, then $(1-t^D)S(t) = (1-t^D)\,[S(t)]_+$ and hence

$$\left[(1-t^D)S(t)\right]_+ = \left[(1-t^D)\,[S(t)]_+\right]_+;$$

- if $i_0 < \infty$, then a_{i_0-D} is positive and thus $a_{i_0} - a_{i_0-D}$ is negative. Let i_1 be the index of the first non-positive coefficient of $(1-t^D)S(t)$. Then $i_1 < i_0$, and hence $\left[(1-t^D)S(t)\right]_+ = \left[(1-t^D)\,[S(t)]_+\right]_+$.

Therefore, for all power series $S \in \mathbb{Z}[[t]]$ such that $S(0) > 0$, we have

$$\left[(1-t^D)\,[S]_+\right]_+ = \left[(1-t^D)S\right]_+ .$$

Consequently, an induction shows that

$$\mathsf{wHS}_{\mathbb{K}(\mathfrak{b},\mathfrak{c})[U,X]/\widetilde{\mathscr{D}_r}}(t) = \left[(1-t^D)^{(p-r)(q-r)}\frac{\det A(t^D)}{t^{D\binom{r}{2}}(1-t)^n}\right]_+ .$$

Then, by Corollary 4.5, $\mathsf{wHS}_{\mathbb{K}(\mathfrak{b},\mathfrak{c})[U,X]/\widetilde{\mathscr{D}_r}}(t) = \mathsf{wHS}_{\mathbb{K}(\mathfrak{a})[U,X]/\widetilde{\mathscr{I}_r}}(t)$. Finally, by Lemma 4.8, we conclude that $\mathsf{HS}_{\mathbb{K}(\mathfrak{a})[X]/\mathscr{I}_r}(t) = \mathsf{wHS}_{\mathbb{K}(\mathfrak{a})[U,X]/\widetilde{\mathscr{I}_r}}(t)$. □

4.6 Complexity analysis

Using the previous results on the Hilbert series of $\mathscr{I}_r$, we analyze now the arithmetic complexity of solving the generalized MinRank problem with Gröbner bases algorithms. In the first part of this section (until Section 4.6.2), we consider the homogeneous MinRank problem (i.e. the polynomials $f_{i,j}$ are homogeneous).

Computing a Gröbner basis of the ideal $\varphi_{\mathbf{a}}(\mathscr{I}_r)$ for the lexicographical ordering yields an explicit description of the set of points V such that the matrix

$$\varphi_{\mathbf{a}}(\mathscr{M}) = \begin{pmatrix} \varphi_{\mathbf{a}}(f_{1,1}) & \dots & \varphi_{\mathbf{a}}(f_{1,q}) \\ \vdots & \ddots & \vdots \\ \varphi_{\mathbf{a}}(f_{p,1}) & \dots & \varphi_{\mathbf{a}}(f_{p,q}) \end{pmatrix}$$

has rank less than $r+1$. In this section, we study the complexity of this computation when $\mathbf{a} \in \mathbb{K}^{pq}$ is generic (i.e. $\mathbf{a}$ belongs to a given non-empty Zariski open subset of $\overline{\mathbb{K}}^{pq}$) by using the theoretical results from Sections 4.4 and 4.5. We focus on the 0-dimensional cases $k = (p-r)(q-r)$ and $n < (p-r)(q-r)$ (over-determined case). Therefore, the set of points where the evaluation of the matrix $\varphi_{\mathbf{a}}(\mathscr{M})$ has rank less than $r+1$ is finite.

In order to compute this set of points, we use the following strategy:

- compute a Gröbner basis of $\varphi_{\mathbf{a}}(\mathscr{I}_r)$ for the *grevlex* (graded reverse lexicographical) ordering with the F_5 algorithm [Fau02];
- convert it into a lexicographical Gröbner basis of $\varphi_{\mathbf{a}}(\mathscr{I}_r)$ by using the FGLM algorithm [FGLM93, FM11].

First, we recall some results about the complexity of the algorithms F_5 and FGLM. The two quantities which allow us to estimate their complexity are respectively the *degree of regularity* and the *degree* of the ideal. If I is a homogeneous 0-dimensional ideal, the degree of regularity of a homogeneous ideal I is an upper bound on the maximum degree in a minimal Gröbner basis. It is also the smallest integer d such that all monomials of degree d are in I and it is independent on the monomial ordering. Moreover, in the 0-dimensional case, the Hilbert series is a polynomial from which the degree of regularity can be read off: $\mathrm{d_{reg}}(I) = \deg(\mathsf{HS}_{\mathbb{K}[X]/I}(t)) + 1$.

Lemma 4.15. *If $n = (p-r)(q-r)$, then the degree of regularity of $\mathscr{I}_r$ is*

$$\mathrm{d_{reg}}\,(\mathscr{I}_r) = Dr(q-r) + (D-1)n + 1.$$

Proof. According to Theorem 4.9, the Hilbert series of $\mathscr{I}_r$ is

$$\mathsf{HS}_{\mathbb{K}(\mathfrak{a})[X]/\mathscr{I}_r}(t) = \frac{\det A_r^{p,q}(t^D)(1-t^D)^{(p-r)(q-r)}}{t^{D\binom{r}{2}}(1-t)^n}.$$

By definition of the matrix $A_r^{p,q}(t)$, the highest degree on each row is reached on the diagonal. Thus, the degree of $\det(A_r^{p,q}(t))$ is the degree of the product of its diagonal elements:

$$\deg(\det(A_r^{p,q}(t))) = \sum_{i=1}^{r}(\min(p,q)-i) = rq - \binom{r+1}{2}.$$

Therefore, we can compute the degree of the Hilbert series which is a polynomial since the ideal is 0-dimensional:

$$\begin{array}{rcl} \mathrm{d_{reg}}\,(\mathscr{I}_r) & = & \deg(\mathsf{HS}_{\mathbb{K}(\mathfrak{a})[X]/\mathscr{I}_r}(t)) + 1 \\ & = & \deg(\det(A_r^{p,q}(t^D))) + D\,(p-r)\,(q-r) - D\binom{r}{2} - n + 1 \\ & = & D(rq - \binom{r+1}{2} + pq - (p+q-r)r - \binom{r}{2}) - n + 1 \\ & = & Dr(q-r) + (D-1)n + 1. \end{array}$$

$\square$

Corollary 4.16. *If $n = (p-r)(q-r)$, then there exists a non-empty Zariski open subset $O \subset \overline{\mathbb{K}}^{pq\binom{D-1+n}{D}}$ such that for all $\mathbf{a} \in O$, the degree of regularity of $\varphi_{\mathbf{a}}(\mathscr{I}_r)$ is*

$$\mathrm{d_{reg}}\,(\varphi_{\mathbf{a}}(\mathscr{I}_r)) = Dr(q-r) + (D-1)n + 1.$$

Proof. According to Lemma 4.3, there exists a Zariski open subset O such that for all $\mathbf{a} \in O$, $\mathsf{LM}\,(\mathscr{I}_r) = \mathsf{LM}\,(\varphi_{\mathbf{a}}(\mathscr{I}_r))$. Consequently, the Hilbert series of both ideals are equal, and hence $\mathrm{d_{reg}}\,(\varphi_{\mathbf{a}}(\mathscr{I}_r)) = \mathrm{d_{reg}}\,(\mathscr{I}_r)$. Lemma 4.15 concludes the proof. $\square$

The degree of regularity governs the complexity of the Gröbner basis computation with respect to the grevlex ordering. The complexity of the algorithm FGLM is bounded by $O(n \cdot \mathsf{DEG}(I)^3)$ which is polynomial in the degree of the ideal [FGLM93, FM11].

Consequently, we can now state the main complexity result:

Theorem 4.17. *There exists a non-empty Zariski open subset $O \subset \overline{\mathbb{K}}^{pq\binom{D-1+n}{D}}$ such that for any $\mathbf{a} \in O$, the arithmetic complexity of computing a lexicographical Gröbner basis of the ideal generated by the $(r+1) \times (r+1)$-minors of the matrix $\varphi_{\mathbf{a}}(\mathscr{M})$ is bounded by*

$$O\left(\binom{p}{r+1}\binom{q}{r+1}\binom{\mathrm{d_{reg}}+n}{n}^{\omega} + n\,(\mathsf{DEG}\,(\varphi_{\mathbf{a}}(\mathscr{I}_r)))^3\right),$$

where $2 \leq \omega \leq 3$ is a feasible exponent for the matrix multiplication, and

- *if $n = (p-r)(q-r)$, then*

$$\mathrm{d_{reg}} = \deg(\mathsf{HS}_{\mathbb{K}[X]/\varphi_{\mathbf{a}}(\mathscr{I}_r)}(t)) + 1 = Dr(q-r) + (D-1)n + 1$$

and $\mathsf{DEG}(\varphi_{\mathbf{a}}(\mathscr{I}_r)) = \mathsf{HS}_{\mathbb{K}[X]/\varphi_{\mathbf{a}}(\mathscr{I}_r)}(1) = D^{(p-r)(q-r)} \prod_{i=0}^{q-r-1} \frac{i!(p+i)!}{(q-1-i)!(p-r+i)!}$.

- *if $n < (p-r)(q-r)$, then assuming that Conjecture 1.53 is true,*

$$\mathrm{d_{reg}} = \deg(\mathsf{HS}_{\mathbb{K}[X]/\varphi_{\mathbf{a}}(\mathscr{I}_r)}(t)) + 1$$

and $\mathsf{DEG}(\varphi_{\mathbf{a}}(\mathscr{I}_r)) = \mathsf{HS}_{\mathbb{K}[X]/\varphi_{\mathbf{a}}(\mathscr{I}_r)}(1)$ *where*

$$\mathsf{HS}_{\mathbb{K}[X]/\varphi_{\mathbf{a}}(\mathscr{I}_r)}(t) = \left[(1-t^D)^{(p-r)(q-r)} \frac{\det A_r^{p,q}(t^D)}{t^{D\binom{r}{2}}(1-t)^n}\right]_+ .$$

Proof. The number of $(r+1)$-minors of the matrix $\varphi_{\mathbf{a}}(\mathscr{M})$ is $\binom{p}{r+1}\binom{q}{r+1}$. Consequently, the theorem is a straightforward consequence of the bounds on the complexity of the F_5 algorithm (Theorem 1.72) and of the FGLM algorithm [FGLM93, FM11], together with the formulas for the degree of regularity (Corollary 4.16) and for the degree (Corollary 4.10). □

4.6.1 Positive dimension

When $n > (p-r)(q-r)$, the dimension of the ideal $\mathscr{I}_r$ is positive. To obtain complexity bounds in that case, we need upper bounds on the maximal degree in the reduced Gröbner basis of $\mathscr{I}_r$.

Lemma 4.18. *If $n > (p-r)(q-r)$, then the maximal degree in a reduced Gröbner basis of $\mathscr{I}_r$ is*

$$Dr(q-r) + (D-1)(p-r)(q-r) + 1.$$

Proof. Consider the ideal J obtained by specializing the last $n-(p-r)(q-r)$ to zero in $\mathscr{I}_r$. We prove now that $\mathsf{LM}(\mathscr{I}_r) = \mathsf{LM}(J)$. First, notice that for the grevlex ordering, $\mathsf{LM}(J) \subset \mathsf{LM}(\mathscr{I}_r)$. According to Theorem 4.9, the Hilbert series of the ideal $J \cap \mathbb{K}(\mathfrak{a})[x_1, \ldots, x_{(p-r)(q-r)}]$ is equal to

$$\frac{\det A_r^{p,q}(t^D)(1-t^D)^{(p-r)(q-r)}}{t^{D\binom{r}{2}}(1-t)^{(p-r)(q-r)}}.$$

Consequently the Hilbert series of J as an ideal of $\mathbb{K}(\mathfrak{a})[x_1, \ldots, x_n]$ is equal to

$$\frac{\det A_r^{p,q}(t^D)(1-t^D)^{(p-r)(q-r)}}{t^{D\binom{r}{2}}(1-t)^n},$$

which is equal to the Hilbert series of $\mathscr{I}_r$.

Since $\mathsf{HS}_J(t) = \mathsf{HS}_{\mathscr{I}_r}(t)$ and $\mathsf{LM}(J) \subset \mathsf{LM}(\mathscr{I}_r)$, we can deduce that $\mathsf{LM}(J) = \mathsf{LM}(\mathscr{I}_r)$.

Consequently, the leading monomials in the reduced Gröbner bases of J and $\mathscr{I}_r$ are the same. Hence, the polynomials in both Gröbner bases have the same degrees since they are homogeneous.

Finally, notice that the Gröbner basis of the ideal J is the same as that of the ideal $J \cap \mathbb{K}(\mathfrak{a})[x_1, \ldots, x_{(p-r)(q-r)}]$ which, by Lemma 4.15, is a zero-dimensional ideal whose degree of regularity is $Dr(q-r) + (D-1)(p-r)(q-r) + 1$. Therefore the maximal degree of the polynomials in the minimal reduced Gröbner basis of $\mathscr{I}_r$ is $Dr(q-r) + (D-1)(p-r)(q-r) + 1$. □

Using a proof similar to that of Corollary 4.16, we deduce that

Corollary 4.19. *If $n > (p-r)(q-r)$, then there exists a non-empty Zariski open subset $O \subset \mathbb{K}^{pq\binom{D-1+n}{D}}$ such that, for $\mathbf{a} \in O$, the maximal degree of the polynomials in a minimal grevlex Gröbner basis of $\varphi_{\mathbf{a}}(\mathscr{I}_r)$ is*

$$Dr(q-r) + (D-1)(p-r)(q-r) + 1.$$

Theorem 4.20. *If $n > (p-r)(q-r)$, then there exists a non-empty Zariski open subset $O \subset \mathbb{K}^{pq\binom{D-1+n}{D}}$ such that for any $\mathbf{a} \in O$, the arithmetic complexity of computing a grevlex Gröbner basis of $\varphi_{\mathbf{a}}(\mathscr{I}_r)$ is bounded by*

$$O\left(\binom{p}{r+1}\binom{q}{r+1}\binom{Dr(q-r)+(D-1)(p-r)(q-r)+1+n}{n}^{\omega}\right).$$

Proof. This is a consequence of Corollary 4.19. The complexity bound is obtained by bounding the complexity of linear algebra on the Macaulay matrices up to the maximal degree in the Gröbner basis (see Theorem 1.72 for similar complexity estimates for 0-dimensional systems). □

4.6.2 The 0-dimensional affine case

For practical applications, the affine case (i.e. when the entries of the input matrix $\mathscr{M}$ are affine polynomials of degree D) is more often encountered than the homogeneous one. In this case, the matrix $\mathscr{M}$ is defined as follows

$$\mathscr{M} = \begin{pmatrix} f_{1,1} & \dots & f_{1,q} \\ \vdots & \ddots & \vdots \\ f_{p,1} & \dots & f_{p,q} \end{pmatrix} \qquad\qquad f_{i,j} = \sum_{\ell=0}^{D} \sum_{t \in \mathsf{Monomials}(\mathbb{K}[X],\ell)} \mathfrak{a}_t^{(i,j)} t.$$

We show in this section that the complexity results (Theorems 4.17 and 4.20) still hold in the affine case. This is achieved by considering the homogenized system:

Definition 4.21. *[CLO97, Chapter 8, §2, Proposition 7] Let $(s_1, \dots, s_\ell) \in \mathbb{K}[x_1, \dots, x_n]^\ell$ be an affine 0-dimensional polynomial system. We let $(\widetilde{s_1}, \dots, \widetilde{s_\ell}) \in \mathbb{K}[x_1, \dots, x_n, x_{n+1}]^\ell$ denote its* homogenized system *defined by*

$$\forall i, \text{ s.t. } 1 \leq i \leq \ell, \widetilde{s_i}(x_1, \dots, x_n, x_{n+1}) = x_{n+1}^{\deg(s_i)} s_i\left(\frac{x_1}{x_{n+1}}, \dots, \frac{x_n}{x_{n+1}}\right).$$

Notice that if an affine polynomial system has solutions, then the dimension of the ideal generated by its homogenized system is positive.

The study of the homogenized system is motivated by the fact that, for the grevlex ordering, the dehomogenization of a Gröbner basis of $\langle \widetilde{s_1}, \dots, \widetilde{s_\ell} \rangle$ is a Gröbner basis of $\langle s_1, \dots, s_\ell \rangle$. Therefore, in order to compute a Gröbner basis of the affine system, it is sufficient to compute a Gröbner basis of the homogenized system (for which we have complexity estimates by Theorems 4.17 and 4.20).

To estimate the complexity of the change of ordering, we need bounds on the degree of the ideal in the affine case:

Lemma 4.22. *The degree of the ideal $\langle s_1, \dots, s_\ell \rangle$ is bounded above by that of $\langle \widetilde{s_1}, \dots, \widetilde{s_\ell} \rangle$.*

Proof. The rings $\mathbb{K}[x_1, \dots, x_n]/\langle s_1, \dots, s_\ell \rangle$ and $\mathbb{K}[x_1, \dots, x_n, x_{n+1}]/\langle \widetilde{s_1}, \dots, \widetilde{s_\ell}, x_{n+1} - 1 \rangle$ are isomorphic. Therefore the degrees of the ideals $\langle s_1, \dots, s_\ell \rangle$ and $\langle \widetilde{s_1}, \dots, \widetilde{s_\ell}, x_{n+1} - 1 \rangle$ are equal. Since $\deg(x_{n+1} - 1) = 1$, we obtain:

$$\begin{aligned} \mathsf{DEG}\left(\langle s_1, \dots, s_\ell \rangle\right) &= \mathsf{DEG}\left(\langle \widetilde{s_1}, \dots, \widetilde{s_\ell}, x_{n+1} - 1 \rangle\right) \\ &\leq \mathsf{DEG}\left(\langle \widetilde{s_1}, \dots, \widetilde{s_\ell} \rangle\right). \end{aligned}$$

□

Lemma 4.23. *The degree of regularity with respect to the grevlex ordering of the system $(s_1, \dots, s_\ell)$ is bounded above by the maximal degree in a reduced Gröbner basis of $\langle \widetilde{s_1}, \dots, \widetilde{s_\ell} \rangle$.*

Proof. Let θ denote the dehomogenization morphism:

$$\begin{array}{rccc}\theta: & \mathbb{K}[x_1,\dots,x_{n+1}] & \longrightarrow & \mathbb{K}[x_1,\dots,x_n] \\ & f(x_1,\dots,x_n,x_{n+1}) & \longmapsto & f(x_1,\dots,x_n,1)\end{array}$$

If G is a grevlex Gröbner basis of $\langle \widetilde{s_1},\dots,\widetilde{s_\ell}\rangle$, then $\theta(G)$ is a grevlex Gröbner basis of $\langle s_1,\dots,s_\ell\rangle$ (this is a consequence of the following property of the grevlex ordering: for all $f\in\mathbb{K}[x_1,\dots,x_{n+1}]$ homogeneous, $\mathsf{LM}(\theta(f))=\theta(\mathsf{LM}(f))$). Also, notice that for each $g\in G$, any relation $g=\sum_{i=1}^{\ell}\widetilde{s_i}h_i$ gives a relation $\theta(g)=\sum_{i=1}^{\ell}s_i\theta(h_i)$ of lower degree since

$$\deg(\theta(\widetilde{s_i})\theta(h_i))\le\deg(\widetilde{s_i}h_i).$$

Consequently, a Gröbner basis of $\langle s_1,\dots,s_\ell\rangle$ can be obtained by computing the row echelon form of the Macaulay matrix of $(s_1,\dots,s_\ell)$ in degree $\mathrm{d_{reg}}(\langle\widetilde{s_1},\dots,\widetilde{s_\ell}\rangle)$. Therefore, the degree of regularity with respect to the grevlex ordering of the system $(s_1,\dots,s_\ell)$ is bounded above by the maximal degree in a reduced Gröbner basis of $\langle\widetilde{s_1},\dots,\widetilde{s_\ell}\rangle$. □

We can now state the main complexity result for the affine generalized MinRank problem:

Theorem 4.24. *Suppose that the matrix $\mathscr{M}$ contains generic affine polynomials of degree D:*

$$\mathscr{M}=\begin{pmatrix} f_{1,1} & \dots & f_{1,q}\\ \vdots & \ddots & \vdots\\ f_{p,1} & \dots & f_{p,q}\end{pmatrix}\qquad\qquad f_{i,j}=\sum_{\ell=0}^{D}\sum_{t\in\mathsf{Monomials}(\mathbb{K}[X],\ell)}\mathfrak{a}_t^{(i,j)}t.$$

There exists a non identically null polynomial $h\in\mathbb{K}[\mathfrak{a}]$ such that for any $\mathbf{a}\in\mathbb{K}^{pq\binom{D+n}{D}}$ such that $h(\mathbf{a})\neq 0$, the overall arithmetic complexity of computing the set of points such that the matrix $\varphi_{\mathbf{a}}(\mathscr{M})$ has rank at most r with Gröbner basis algorithms is bounded by

$$O\left(\binom{p}{r+1}\binom{q}{r+1}\binom{\mathrm{d_{reg}}+n}{n}^{\omega}+n\left(\mathsf{DEG}(\varphi_{\mathbf{a}}(\mathscr{I}_r))\right)^3\right),$$

where $2\le\omega\le3$ is a feasible exponent for the matrix multiplication and

- *if $n=(p-r)(q-r)$, then*

$$\mathrm{d_{reg}}\le Dr(q-r)+(D-1)n+1,$$

$$\mathsf{DEG}(\varphi_{\mathbf{a}}(\mathscr{I}_r))\le D^{(p-r)(q-r)}\prod_{i=0}^{q-r-1}\frac{i!(p+i)!}{(q-1-i)!(p-r+i)!}.$$

- *if $n<(p-r)(q-r)$, then assuming that Conjecture 1.53 is true,*

$$\mathrm{d_{reg}}\le\deg(P(t))+1,$$

and $\mathsf{DEG}(\varphi_{\mathbf{a}}(\mathscr{I}_r))\le P(1)$ where

$$P(t)=\left[(1-t^D)^{(p-r)(q-r)}\frac{\det A_r^{p,q}(t^D)}{t^{D\binom{r}{2}}(1-t)^n}\right]_+.$$

Proof. This is a direct consequence of Theorem 1.72, Lemma 4.22, Lemma 4.23 and the complexity of the FGLM algorithm [FGLM93, FM11] ($O(n\,\mathsf{DEG}(\varphi_{\mathbf{a}}(\mathscr{I}_r)^3))$). □

4.7 Case studies

The aim of this section is to compare the complexity of the grevlex Gröbner basis computation with the degree of the ideal in the 0-dimensional case (i.e. the number of solutions of the MinRank problem counted with multiplicities). Since the "arithmetic size" (i.e. the number of monomials) of the lexicographical Gröbner basis is close to the degree of the ideal in the 0-dimensional case (Propositions 1.74 and 1.75), it is interesting to identify families of parameters for which the arithmetic complexity of the computation is polynomial in this degree under genericity assumptions.

Throughout this section, we focus on the 0-dimensional case: $n = (p-r)(q-r)$. Under genericity assumptions, we recall that, by Corollary 4.10 and Lemma 4.15,

$$\begin{aligned} \mathrm{d}_{\mathrm{reg}} &= Dr(q-r) + (D-1)n + 1 \\ \mathsf{DEG} &= D^{(p-r)(q-r)} \prod_{i=0}^{q-r-1} \frac{i!(p+i)!}{(q-1-i)!(p-r+i)!}. \end{aligned}$$

According to Theorem 4.24, the complexity of the computation of the grevlex Gröbner basis is then bounded by

$$O\left(\binom{p}{r+1}\binom{q}{r+1}\binom{Dr(q-r)+(D-1)n+1}{n}^{\omega} + n\left(\mathsf{DEG}\left(\mathscr{I}_r\right)\right)^3\right).$$

Since the complexity of FGLM is polynomial in the degree of the ideal, we focus on the complexity of the F_5 algorithm. To this end, we introduce the notation

$$\mathsf{Compl} = \binom{p}{r+1}\binom{q}{r+1}\binom{Dr(q-r)+(D-1)n+1}{n}^{\omega}.$$

The goal here is to prove that the complexity bound of the F_5 algorithm is polynomial in the degree of the ideal $\mathscr{I}$ for subfamilies of generalized MinRank problems. This is done by showing that the ratio $\log(\mathsf{Compl})/\log(DEG(\mathscr{I})$ is bounded by a constant. As proved and verified experimentally below, this is true for several subfamilies of problems.

However, Figure 4.3 seems to show experimentally that this is not always the case. This fact can have two different explanation: the actual complexity of the F_5 algorithm may not be polynomial in the degree of the ideal or it is also possible that the complexity bound used for the analysis is not precise enough for these families of parameters. This remains an important open question.

4.7.1 D grows, p, q, r are fixed

We first study the case where p, q and r are fixed (and thus $n = (p-r)(q-r)$ is constant too), and D grows. In that case, the arithmetic complexity of the grevlex Gröbner basis computation is bounded by $O(D^{n\omega})$, and the degree of $\mathscr{I}_r$ is lower bounded by $\Omega(D^n)$. Therefore the arithmetic complexity of the Gröbner basis computation is polynomial in the degree of the ideal for these parameters.

4.7.2 p grows, q, r, D are fixed

This section is devoted to the study of the subfamilies of Generalized MinRank problems when the parameters q, r and D are constant values and p grows. Let ℓ denote the constant value $\ell = q - r$. First, we assume that $D = 1$. When p grows, by Corollary 4.10 we have

$$\begin{aligned} \log(\mathsf{DEG}) &= \log\left(\prod_{i=0}^{\ell-1} \frac{\binom{p+\ell-1}{r+i}}{\binom{p+\ell-1}{i}}\right) \\ &\underset{p\to\infty}{\sim} r\ell\log(p) \end{aligned}$$

On the other hand,

$$\begin{array}{rcl}\log(\mathsf{Compl}) & = & \omega\log\binom{(p-r)\ell+r\ell+1}{(p-r)\ell}+\log\binom{p}{r+1}+\log\binom{q}{r+1}\\ & = & \omega\log\binom{p\ell+1}{r\ell+1}+\log\binom{p}{r+1}+\log\binom{q}{r+1}\\ & \underset{p\to\infty}{\sim} & (\omega(r\ell+1)+r+1)\log(p).\end{array}$$

Therefore, $\log(\mathsf{Compl})/\log(\mathsf{DEG}) \underset{p\to\infty}{\sim} \dfrac{\omega(r\ell+1)+r+1}{r\ell}$ and hence the number of arithmetic operations is polynomial in the degree of the ideal.

Also, if $D \geq 2$ is constant, a similar analysis yields

$$\begin{array}{rcl}\log(\mathsf{DEG}) & = & (p-r)\ell\log(D)+\log\left(\prod_{i=0}^{\ell-1}\frac{\binom{p+\ell-1}{r+i}}{\binom{p+\ell-1}{i}}\right)\\ & \underset{n\to\infty}{\sim} & \log(D)\ell p.\\ \log(\mathsf{Compl}) & = & \omega\log\binom{n+Dr\ell+(D-1)n+1}{n}+\log\binom{p}{r+1}+\log\binom{q}{r+1}\\ & = & \omega\log\binom{Dp\ell+1}{(p-r)\ell}+\log\binom{p}{r+1}+\log\binom{q}{r+1}\\ & \underset{p\to\infty}{\sim} & \omega\log\binom{Dp\ell}{p\ell}.\end{array}$$

Then, using the fact that $\log\binom{\alpha p}{\beta p} \underset{p\to\infty}{\sim} p\left(\alpha\log(\alpha)-\beta\log(\beta)-(\alpha-\beta)\log(\alpha-\beta)\right)$, we obtain that

$$\log(\mathsf{Compl}) \underset{p\to\infty}{\sim} p\omega\ell(D\log(D)-(D-1)\log(D-1)).$$

Therefore, $\log(\mathsf{Compl})/\log(\mathsf{DEG})$ is bounded above by a constant value and hence the arithmetic complexity of the Gröbner basis computation is also polynomial in the degree of the ideal for this subclass of Generalized MinRank problems under genericity assumptions.

4.7.3 The case $r = q - 1$

The case $r = q - 1$ is a special case of the setting studied in Section 4.7.2 which arises in several applications, since it is the problem of finding the points where the evaluation of a polynomial matrix is rank defective. In this setting, the formulas in Theorem 4.24 are much simpler:

- the 0-dimensional condition yields $n = p - q + 1$;
- $\mathrm{d}_{\mathrm{reg}} \leq Dp - (p - q)$;
- $\mathsf{DEG} \leq D^{p-q+1}\binom{p}{q-1}$.

Therefore, the arithmetic complexity of the Gröbner basis computation is $\mathsf{Compl} = O(\binom{Dp+1}{p-q+1}^{\omega})$.

If $D > 1$ and q are fixed, $\log\left(\binom{Dp+1}{p-q+1}^{\omega}\right) \underset{p\to\infty}{\sim} \omega\log\binom{Dp}{p}$ and a direct application of Stirling's formula shows that

$$\omega\log\binom{Dp}{p} \underset{p\to\infty}{\sim} \omega(D\log D-(D-1)\log(D-1))p.$$

(p,q,D,r,n)	degree	d_{reg}	F_4 time(Magma)	FGLM time(Magma)	F_5 time/nb.ops(FGb)	FGLM time(FGb)
(6,5,2,4,2)	60	11	0.001s	0.001s	0.00s/$2^{13.32}$	0.00s
(6,5,3,4,2)	135	17	0.002s	0.019s	0.00s/$2^{15.29}$	0.00s
(6,5,4,4,2)	240	23	0.004s	0.09s	0.01s/$2^{16.79}$	0.01s
(5,5,2,3,4)	800	17	0.25s	6.3s	0.24s/$2^{25.56}$	0.19s
(8,5,2,4,4)	1120	13	0.7s	20s	0.43s/$2^{26.71}$	0.58s
(5,5,3,3,4)	4050	27	6.7s	567s	5.43s/$2^{30.68}$	3s
(6,5,2,3,6)	11200	19	479s	17703s	94.85s/$2^{35.7}$	203s

Table 4.1: Experimental results

On the other hand, $\log(\mathsf{DEG}) \underset{p\to\infty}{\sim} p \log D$. Therefore, $\log(\mathsf{Compl})/\log(\mathsf{DEG})$ has a finite limit when p grows and q is fixed, showing that in this setting the arithmetic complexity is polynomial in the degree of the ideal.

4.7.4 Experimental results

In this section, we present experimental results obtained by using the Gröbner bases package FGb and the implementation of the F_4 algorithm in the MAGMA computer algebra system [BCP97]. All instances were constructed as uniformly random 0-dimensional MinRank problems (i.e. $(p-r)(q-r)=n$) over the finite field GF_{65521}. All experiments were conducted on a 2.93 GHz Intel Xeon with 132 GB RAM.

Useful information can be read from Table 4.1. First, the experimental values of the degree of regularity and of the degree match exactly the theoretical values given in Lemma 4.15 and in Corollary 4.10. Also, it can be noted that the most relevant indicator of the complexity of the Gröbner basis computation seems to be the degree of the ideal.

The comparison between the complexity bound and the degree of the ideal is illustrated in Figures 4.1 and 4.2. First, Figure 4.1 shows that the bound on the complexity of the Gröbner computation is polynomial in the degree of the ideal when D grows ($p = q = 20$, $r = 10$ fixed), since $\log(\mathsf{Compl})/\log(\mathsf{DEG})$ is bounded by 5. This is in accordance with the analysis performed in Section 4.7.1.

Then Figure 4.2 shows empirically that if $q = \lfloor \beta p \rfloor$ and $r = \lfloor \alpha p \rfloor - 1$ (with $0 < \alpha \leq \beta \leq 1$) and p grows, then the complexity bound is also polynomial in the degree of the ideal.

However, there also exist families of generalized MinRank problem such that the complexity bound for the Gröbner basis computation is *not* polynomial in the degree of ideal. For instance, taking $p = q$ and fixing the values of r and D yields such a family. The experimental behavior of $\log(\mathsf{Compl})/\log(\mathsf{DEG})$ with this setting is plotted in Figure 4.3. We would like to point out that this does not necessarily mean that the complexity of the Gröbner basis computation is not polynomial in the degree of the ideal. Indeed, the complexity bound $O\left(\binom{p}{r+1}\binom{q}{r+1}\binom{n+d_{\mathrm{reg}}}{n}^{\omega}\right)$ is only an upper bound and the figure only indicates that this bound is not polynomial.

The problem of showing whether the actual arithmetic complexity of the F_5 algorithm is polynomial or not in the degree of the ideal for all families of parameters of the generalized MinRank problem remains an open question.

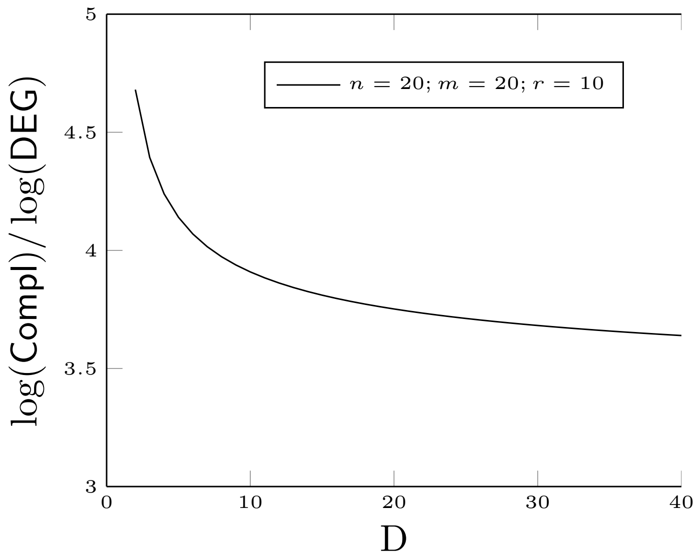

Figure 4.1: Numerical values of $\log(\mathsf{Compl})/\log(\mathsf{DEG})$, for $p = q = 20, r = 10, n = (p-r)(q-r)$.

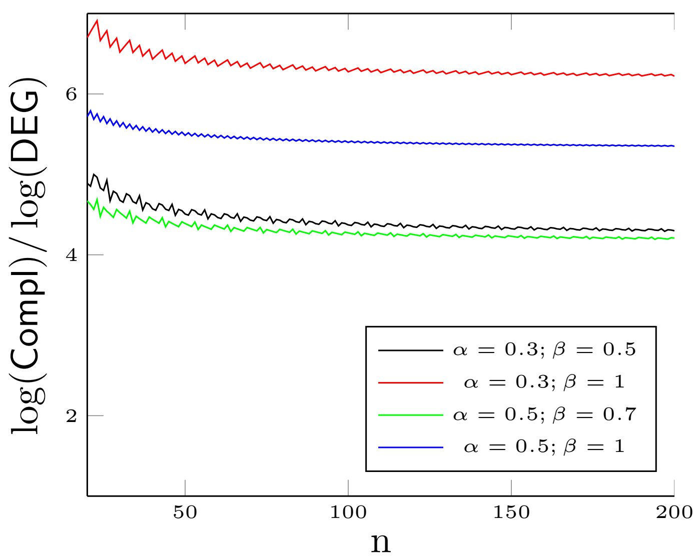

Figure 4.2: Numerical values of $\log(\mathsf{Compl})/\log(\mathsf{DEG})$, for $q = \lfloor \beta p \rfloor, r = \lfloor \alpha p \rfloor - 1, D = 1, n = (p-r)(q-r)$.

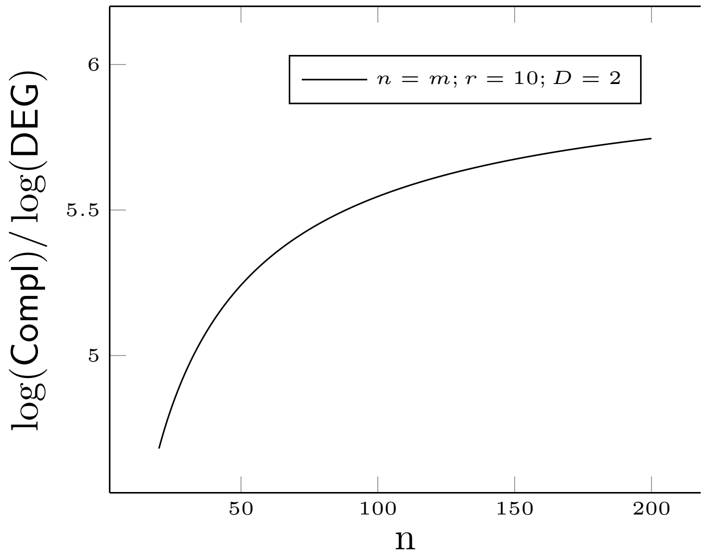

Figure 4.3: Numerical values of $\log(\mathsf{Compl})/\log(\mathsf{DEG})$, for $q = \lfloor \beta p \rfloor, r = \lfloor \alpha p \rfloor - 1, D = 1, n = (p - r)(q - r)$.

Chapter 5

Critical Point Systems

The results presented in this chapter are joint work with J.-C. Faugère and M. Safey El Din. Sections 5.1 to 5.6 are published in [FSS12].

In this chapter, we consider the problem of computing critical points of the restriction of a polynomial map to an algebraic variety. This is of first importance since the global minimum of such a map is reached at a critical point. Thus, these points appear naturally in non-convex polynomial optimization which occurs in a wide range of scientific applications (control theory, chemistry, economics,...).

Critical points also play a central role in recent algorithms of effective real algebraic geometry. Experimentally, it has been observed that Gröbner basis algorithms are efficient to compute such points. Therefore, recent software based on the so-called Critical Point Method are built on Gröbner bases engines.

Let $f_1, \dots, f_p$ be polynomials in $\mathbb{Q}[x_1, \dots, x_n]$ of degree D, $V \subset \mathbb{C}^n$ be their complex variety and π_1 be the projection map $(x_1, \dots, x_n) \mapsto x_1$. The critical points of the restriction of π_1 to V are defined by the vanishing of $f_1, \dots, f_p$ and some maximal minors of the Jacobian matrix associated to $f_1, \dots, f_p$. Such a system is algebraically structured: the ideal it generates is the sum of a determinantal ideal and the ideal generated by $f_1, \dots, f_p$.

We provide the first complexity estimates on the computation of Gröbner bases of such systems defining critical points. We prove that under genericity assumptions on $f_1, \dots, f_p$, the complexity is polynomial in the generic number of critical points, i.e. $D^p(D-1)^{n-p}\binom{n-1}{p-1}$. More particularly, in the quadratic case $D = 2$, the complexity of such a Gröbner basis computation is polynomial in the number of variables n and exponential in p. We also give experimental evidence supporting these theoretical results.

5.1 Introduction

Motivations and problem statement. The local extrema of the restriction of a polynomial map to a real algebraic variety are reached at the critical points of the map under consideration. Hence, computing these critical points is of first importance for polynomial optimization which arises in a wide range of applications in engineering sciences (control theory, chemistry, economics, etc.).

Computing critical points is also the cornerstone of algorithms for asymptotically optimal algorithms for polynomial system solving over the reals (singly exponential in the number of variables). Indeed, for computing sample points in each connected component of a semi-algebraic set, the algorithms based on the so-called critical point method rely on a reduction of the initial problem to polynomial optimization problems. In [BPR96, BPR98] (see also [GV88, HRS89, HRS93]), the best complexity bounds are obtained using infinitesimal deformation techniques of semi-algebraic geometry, nevertheless obtaining efficient implementations of these algorithms remains an issue.

Tremendous efforts have been made to obtain fast implementations relying on the critical point method (see [SS03, ELLS09, Saf07, HS11, HS09, FMRS08, SS04]). This is achieved with techniques based on algebraic elimination and complex algebraic geometry. For instance, when the input polynomial system $(\mathbf{F}) : f_1 = \cdots = f_p = 0$ in $\mathbb{Q}[x_1, \ldots, x_n]$ satisfies genericity assumptions, one is led to compute the set of critical points of the restriction of the projection on the first coordinate $\pi_1 : (x_1, \ldots, x_n) \mapsto x_1$ to the algebraic variety $Z(\mathbf{F}) \subset \mathbb{C}^n$ defined by $\mathbf{F}$; this set is denoted by $\mathsf{crit}(\pi_1, Z(\mathbf{F}))$.

The set $\mathsf{crit}(\pi_1, Z(\mathbf{F}))$ is defined by $\mathbf{F}$ and the vanishing of the maximal minors of the truncated Jacobian matrix of $\mathbf{F}$ obtained by removing the partial derivatives with respect to x_1. This system is highly-structured: algebraically, we are considering the sum of a determinantal ideal with the ideal $\langle f_1, \ldots, f_p\rangle$.

In practice, we compute a rational parametrization of this set through Gröbner bases computations which are fast in practice. We have observed that the behavior of Gröbner bases on these systems is specific: the highest degree reached during the computations is unexpectedly small. In the particular case of quadratic equations, the complexity of the computation seems to be polynomial in n and exponential in p which meets the best complexity known bound for the quadratic minimization problem: an approximation algorithm with such a complexity is given in [Bar93] (see also [GP05] for general polynomial algorithms in optimization). Understanding the complexity of these computations is a first step towards the design of dedicated Gröbner bases algorithms, so we focus on the following important open problems:

- **(A)** Can we provide *complexity estimates* for the computation of Gröbner bases of ideals defined by such *structured algebraic systems*?
- **(B)** Is this computation *polynomial in the generic number of critical points*?
- **(C)** In the *quadratic case*, is this computation *polynomial in the number of variables* (and exponential in the codimension)?

Under genericity assumptions, we actually provide affirmative answers to all these questions.

Computational methodology and related complexity issues. Gröbner bases are computed using multi-modular arithmetics and we will focus only on arithmetic complexity results; so we may consider systems defining critical points with coefficients not only in $\mathbb{Q}$ but also in a prime field.

Let $\mathbb{K}$ be a field, $\overline{\mathbb{K}}$ be its algebraic closure and $\mathbf{F} = (f_1, \ldots, f_p)$ be a family of polynomials in $\mathbb{K}[x_1, \ldots, x_n]$ of degree D and $Z(\mathbf{F})$ be their set of common zeroes in $\overline{\mathbb{K}}^n$.

We denote the Jacobian matrix

$$\begin{bmatrix} \frac{\partial f_1}{\partial x_1} & \cdots & \frac{\partial f_1}{\partial x_n} \\ \vdots & & \vdots \\ \frac{\partial f_p}{\partial x_1} & \cdots & \frac{\partial f_p}{\partial x_n} \end{bmatrix}$$

by $\mathsf{jac}(\mathbf{F})$ and the submatrix obtained by removing the first i columns by $\mathsf{jac}(\mathbf{F}, i)$. The set of maximal minors of a given rectangular matrix M will be denoted by $\mathsf{MaxMinors}(\mathsf{M})$.

Finally, let $\mathbf{I}(\mathbf{F}, 1)$ be the ideal $\langle \mathbf{F}\rangle + \langle \mathsf{MaxMinors}(\mathsf{jac}(\mathbf{F}, 1))\rangle$. When $\mathbf{F}$ is a reduced regular sequence and $Z(\mathbf{F})$ is smooth, the algebraic variety associated to $\mathbf{I}(\mathbf{F}, 1)$ is exactly $\mathsf{crit}(\pi_1, Z(\mathbf{F}))$.

So, to compute a rational parametrization of $\mathsf{crit}(\pi_1, Z(\mathbf{F}))$, we use the classical solving strategy which proceeds in two steps:

(i) compute a Gröbner basis for a *grevlex* ordering of $\mathbf{I}(\mathbf{F}, 1)$ using the F_5 algorithm (see [Fau02]);

(ii) use the FGLM algorithm [FGLM93, FM11] to obtain a Gröbner basis of $\mathbf{I}(\mathbf{F}, 1)$ for the lexicographical ordering or a rational parametrization of $\sqrt{\mathbf{I}(\mathbf{F}, 1)}$.

Algorithm F_5 (Step *(i)*) computes Gröbner bases by row-echelon form reductions of submatrices of the Macaulay matrix up to a given degree. This latter degree is called *degree of regularity*. When the input satisfies regularity properties, this complexity of this step can be analyzed by estimating the degree of regularity.

FGLM algorithm [FGLM93] (Step *(ii)*) and its recent efficient variant [FM11] are based on computations of characteristic polynomials of linear endomorphisms in $\mathbb{K}[x_1, \ldots, x_n]/\mathbf{I}(\mathbf{F}, 1)$. This is done by performing linear algebra operations of size the *degree of* $\mathbf{I}(\mathbf{F}, 1)$ (which is the number of solutions counted with multiplicities).

Thus, we are faced to the following problems:

(1) estimate the degree of regularity of the ideal generated by the homogeneous components of highest degree of the set of generators $\mathbf{F}, \mathsf{MaxMinors}(\mathsf{jac}(\mathbf{F}, 1))$ and bound the complexity of computing a *grevlex* Gröbner basis of $\mathbf{I}(\mathbf{F}, 1)$;

(2) provide sharp bounds on the degree of the ideal $\mathbf{I}(\mathbf{F}, 1)$.

As far as we know, no results are known for problem *(1)*. Problem *(2)* has already been investigated in the literature: see [NR09] where some bounds are given on the cardinality of $\mathsf{crit}(\pi_1, Z(\mathbf{F}))$. We give here a new algebraic proof of these bounds.

Main results. Let $\mathbb{K}[x_1, \ldots, x_n]_D$ denote $\{f \in \mathbb{K}[x_1, \ldots, x_n] \mid \deg(f) = D\}$ and note that it is a finite-dimensional vector space. In the following, we solve the three aforementioned problems under a *genericity* assumption on $\mathbf{F}$: we actually prove that there exists a non-empty Zariski open set $\mathscr{O} \subset \overline{\mathbb{K}}[x_1, \ldots, x_n]_D^p$ such that for all $\mathbf{F} \in \mathscr{O}$:

(1) the degree of regularity of the ideal generated by the homogeneous components of largest degree of $\mathbf{F}, \mathsf{MaxMinors}(\mathsf{jac}(\mathbf{F}, 1))$ is $\mathrm{d}_{\mathrm{reg}} = D(p-1) + (D-2)n + 2$ (see Theorem 5.5);

(2) the degree of $\mathbf{I}(\mathbf{F}, 1)$ is $\leq \delta = D^p(D-1)^{n-p}\binom{n-1}{p-1}$.

The degree of regularity given in *(1)* is obtained thanks to an explicit formula for the Hilbert series of the homogeneous ideal under consideration (see Proposition 5.7). This is obtained by taking into account the determinantal structure of some of the generators of the ideal we consider. The above estimates are the key results which enable us to provide positive answers to questions **A**, **B** and **C** under genericity assumptions.

Before stating complexity results on the computation of critical points with Gröbner bases, we need to introduce a standard notation. Let ω be a real number such that a row echelon form of a $n \times n$-matrix with entries in $\mathbb{K}$ is computed within $O(n^\omega)$ arithmetic operations in $\mathbb{K}$.

We prove that there exists a non-empty Zariski open set $\mathscr{O} \subset \overline{\mathbb{K}}[x_1, \ldots, x_n]_D^p$ such that for all $\mathbf{F} \in \mathscr{O} \cap \mathbb{K}[x_1, \ldots, x_n]^p$:

(A) computing a *grevlex* Gröbner basis of $\mathbf{I}(\mathbf{F}, 1)$ can be done within $O\left(\left(p + \binom{n-1}{p}\right)\binom{n+\mathrm{d}_{\mathrm{reg}}}{n}^\omega\right)$ arithmetic operations in $\mathbb{K}$ (see Theorem 5.15);

(B) computing a rational parametrization of $\mathsf{crit}(\pi_1, Z(\mathbf{F}))$ using Gröbner bases can be done within $O\left(\delta^{4.03\omega}\right)$ arithmetic operations in $\mathbb{K}$ (see Corollary 5.18);

(C) when $D = 2$ (quadratic case), a rational parametrization of $\mathsf{crit}(\pi_1, Z(\mathbf{F}))$ using Gröbner bases can be computed within $O\left(\binom{n+2p}{2p}^\omega + n2^{3p}\binom{n-1}{p-1}^3\right)$ arithmetic operations in $\mathbb{K}$, this is polynomial in n and exponential in p (see Corollary 5.16).

We also provide more accurate complexity results. The uniform complexity bound given for answering question **(B)** is rather pessimistic. The exponent 4.03ω being obtained after majorations which are not sharp; numerical experiments are given to support this (see Section 5.6). Moreover, under the above genericity assumption, we prove that, when p and D are fixed, computing a rational parametrization of $\mathsf{crit}(\pi_1, Z(\mathbf{F}))$ using Gröbner bases is done within $O(D^{3.57n})$ arithmetic operations in $\mathbb{K}$ (see Corollary 5.17).

We also give timings for computing grevlex and lex Gröbner bases of $\mathbf{I}(\mathbf{F}, 1)$ with the MAGMA computational algebra system and with the FGb library when $\mathbb{K} = \mathsf{GF}(65521)$. These experiments show that the theoretical bounds on the degree of regularity and on the degree of $\mathbf{I}(\mathbf{F}, 1)$ (Theorem 5.12) are sharp. They also provide some indication on the size of problems that can be tackled in practice: e.g. when $D = 2$ and $p = 3$ (resp. $D = 3$ and $p = 1$), random dense systems with $n \leq 21$ (resp. $n \leq 14$) can be tackled (see Section 5.6).

Related works. As far as we know, dedicated complexity analysis of Gröbner bases on ideals defining critical points has not been investigated before. However, as we already mentioned, the determinantal structure of the system defining $\mathsf{crit}(\pi_1, Z(\mathbf{F}))$ plays a central role in this chapter.

In [FSS10], we provided complexity estimates for the computation of Gröbner bases of ideals generated by minors of a linear matrix. This is generalized in [FSS11b] for matrices with entries of degree D. Nevertheless, the analysis which is done here differs significantly from these previous works. Indeed, in [FSS10, FSS11b] a genericity assumption is done on the entries of the considered matrix. We cannot follow the same reasonings since $\mathsf{MaxMinors}(\mathsf{jac}(\mathbf{F}, 1))$ depends on $\mathbf{F}$. Nevertheless, it is worthwhile to note that, as in [FSS10, FSS11b], we use properties of determinantal ideals given in [CH94].

Bounds on the number of critical points (under genericity assumptions) are given in [NR09] using the Giambelli-Thom-Porteous degree bounds on determinantal varieties (see [Ful97, Ex. 14.4.14]).

In [Bar93], the first polynomial time algorithms in n for deciding emptiness of a quadratic system of equations over the reals is given. Further complexity results in the quadratic case for effective real algebraic geometry have been given in [GP05]. In the general case, algorithms based on the so-called critical point method are given in [BPR96, BPR98, GV88, HRS89, HRS93]. Critical points defined by systems $\mathbf{F}, \mathsf{MaxMinors}(\mathsf{jac}(\mathbf{F}, 1))$ are computed in algorithms given in [BGHM01, BGHM97, BGHP05, BGHP04, BGH$^+$10, SS03, ARS02, FMRS08]. The RAGlib maple package implements the algorithms given in [SS03, FMRS08] using Gröbner bases.

The systems $\mathbf{F}, \mathsf{MaxMinors}(\mathsf{jac}(\mathbf{F}, 1))$ define polar varieties: indeed, this notion coincides with critical points in the regular case. In [BGHM01, BGHM97, BGHP05, BGHP04, BGH$^+$10], rational parametrizations are obtained using the geometric resolution algorithm [GLS01] and a local description of these polar varieties. This leads to algorithms computing critical points running in probabilistic time polynomial in $D^p(p(D-1))^{n-p}$. Note that this bound for $D = 2$ and $p = n/2$ is not satisfactory. In this chapter, we also provide complexity estimations for computing critical points but using Gröbner bases, which is the engine we use in practice. Our results provide an explanation of the good practical behavior we have observed.

We would like to mention that other dedicated algebraic techniques exist for elimination in determinantal varieties. In particular, the *determinantal resultant* introduced and studied in [Bus04] can be used for this task. It is implemented in the Macaulay2 package `Resultants`[1].

Organization of the chapter. Section 5.2 recalls well-known properties of generic polynomial systems. Problem *(1)* mentioned above is tackled in Sections 5.3 and 5.4. Problem *(2)* is solved at the end of Section 5.4. Complexity results are derived in Section 5.5. Experimental results supporting the theoretical results are given in Section 5.6. In Section 5.7, we extend the give a formula for the Hilbert series in the mixed case and we generalize the complexity results.

[1] written by L. Busé, N. Botbol and M. Dubinsky

5.2 Preliminaries

Notations 5.1. *The set of variables* $\{x_1,\ldots,x_n\}$ *is denoted by* X. *For* $d\in\mathbb{N}$, $\mathsf{Monomials}(d)$ *denotes the set of monomials of degree* d *in the polynomial ring* $\mathbb{K}[X]$ *(where* $\mathbb{K}$ *is a field, its algebraic closure being denoted by* $\overline{\mathbb{K}}$*). We let* $\mathfrak{a}$ *denote the finite set of parameters* $\{\mathfrak{a}_{\mathfrak{m}}^{(i)}: 1\le i\le p, \mathfrak{m}\in\bigcup_{0\le d\le D}\mathsf{Monomials}(d)\}$.

We also introduce the following generic systems:

- $\mathfrak{F}=(\mathfrak{f}_1,\ldots,\mathfrak{f}_p)\in\mathbb{K}(\mathfrak{a})[X]^p$ *is the generic polynomial system of degree* D*:*

$$\mathfrak{f}_i=\sum_{\substack{\mathfrak{m}\ \textit{monomial}\\ \deg(\mathfrak{m})\le D}}\mathfrak{a}_{\mathfrak{m}}^{(i)}\mathfrak{m};$$

- $\mathfrak{F}^h=(\mathfrak{f}_1^h,\ldots,\mathfrak{f}_p^h)\in\mathbb{K}(\mathfrak{a})[X]^p$ *is the generic* homogeneous *polynomial system of degree* D*:*

$$\mathfrak{f}_i=\sum_{\substack{\mathfrak{m}\ \textit{monomial}\\ \deg(\mathfrak{m})= D}}\mathfrak{a}_{\mathfrak{m}}^{(i)}\mathfrak{m}.$$

We let $Z(\mathbf{F})\subset\overline{\mathbb{K}}^n$ *denote the variety of* $\mathbf{F}=(f_1,\ldots,f_p)$. *The projective variety of a homogeneous family of polynomials* $\mathbf{F}^h$ *is denoted in this chapter by* $W(\mathbf{F}^h)$. *The projection on the first coordinate is denoted by* π_1*, and the critical points of the restriction of* π_1 *to* $Z(\mathbf{F})$ *are denoted by* $\mathsf{crit}(\pi_1,Z(\mathbf{F}))\subset Z(\mathbf{F})$. *Also,* $\mathbf{I}(\mathbf{F},1)$ *denotes the ideal generated by* $\mathbf{F}$ *and by the maximal minors of the truncated Jacobian matrix* $\mathsf{jac}(\mathbf{F},1)$.

The goal of this section is to prove that the ideal $\mathbf{I}(\mathfrak{F}^h,1)$ is 0-dimensional. This will be done in Lemma 5.4 below; to do that we will use geometric statements of Sard's Theorem which require $\mathbb{K}$ to have characteristic 0. This latter assumption can be weakened using algebraic equivalents of Sard's Theorem (see [Eis95, Corollary 16.23]).

Lemma 5.2. *Let* $\mathbf{I}(\mathfrak{F},0)$ *be the ideal generated by* $\mathfrak{F}$ *and by the maximal minors of its Jacobian matrix. Its variety* $Z(\mathbf{I}(\mathfrak{F},0))\subset\overline{\mathbb{K}(\mathfrak{a})}^n$ *is empty and hence* $Z(\mathfrak{F})$ *is* smooth.

Proof. To simplify notations hereafter, we denote by $h_1,\ldots,h_p$ the polynomials obtained from $\mathfrak{f}_1,\ldots,\mathfrak{f}_p$ by removing their respective constant terms $\mathfrak{a}_1^{(1)},\ldots,\mathfrak{a}_1^{(p)}$. We will also denote by $\mathscr{A}$ the remaining parameters in $h_1,\ldots,h_p$. Let ψ denote the mapping

$$\begin{array}{rcl}\psi:\overline{\mathbb{K}(\mathscr{A})}^n & \longrightarrow & \overline{\mathbb{K}(\mathscr{A})}^p\\ \mathbf{c} & \longmapsto & (h_1(\mathbf{c}),\ldots,h_p(\mathbf{c}))\end{array}$$

Suppose first that $\psi(\overline{\mathbb{K}(\mathscr{A})}^n)$ is not dense (for the Zariski topology) in $\overline{\mathbb{K}(\mathscr{A})}^p$. Since the image $\psi(\overline{\mathbb{K}(\mathscr{A})}^n)$ is a constructible set, it is contained in a proper Zariski closed subset $\mathscr{W}\subset\overline{\mathbb{K}(\mathscr{A})}^p$. Since there is no algebraic relation between $\mathfrak{a}_1^{(1)},\ldots,\mathfrak{a}_1^{(p)}$ and the parameters in $\mathscr{A}$, this implies that the variety defined by $h_1+\mathfrak{a}_1^{(1)}=\cdots=h_p+\mathfrak{a}_1^{(p)}=0$ is empty and consequently smooth. Since $h_i+\mathfrak{a}_i^{(1)}=\mathfrak{f}_i$, our statement follows.

Suppose now that $\psi(\overline{\mathbb{K}(\mathscr{A})}^n)$ is dense in $\overline{\mathbb{K}(\mathscr{A})}^p$. Let $K_0\subset\overline{\mathbb{K}(\mathscr{A})}^p$ be the set of critical values of ψ. By Sard's Theorem [Sha94, Chap. 2, Sec. 6.2, Thm 2], K_0 is contained in a proper closed subset of $\overline{\mathbb{K}(\mathscr{A})}^p$. Again, there is no algebraic relation between $\mathfrak{a}_1^{(1)},\ldots,\mathfrak{a}_1^{(p)}$ and the parameters in $\mathscr{A}$. Consequently, the variety associated to the ideal generated by the system $\mathfrak{f}_1,\ldots,\mathfrak{f}_p$ and by the maximal minors of $\mathsf{jac}(\mathfrak{F})$ is empty. $\square$

Corollary 5.3. *Let $\mathbf{I}(\mathfrak{F}^h, 0)$ be the ideal generated by $\mathfrak{F}^h$ and by the maximal minors of its Jacobian matrix. Then the associated projective variety $W(\mathbf{I}(\mathfrak{F}^h, 0)) \subset \mathbb{P}^{n-1}\overline{\mathbb{K}(\mathfrak{a})}$ is empty.*

Proof. For $1 \leq i \leq n$, we denote by O_i the set

$$\{(c_1 : \ldots : c_n) \mid c_i \neq 0\} \subset \mathbb{P}^{n-1}\overline{\mathbb{K}(\mathfrak{a})}$$

and we consider the canonical open covering of $\mathbb{P}^{n-1}\overline{\mathbb{K}(\mathfrak{a})}$:

$$\mathbb{P}^{n-1}\overline{\mathbb{K}(\mathfrak{a})} = \bigcup_{1 \leq i \leq n} O_i.$$

Therefore $W(\mathbf{I}(\mathfrak{F}^h, 0)) = \bigcup_{1 \leq i \leq n} (W(\mathbf{I}(\mathfrak{F}^h, 0)) \cap O_i)$. Denote by $\mathfrak{F}_i$ the system obtained by substituting the variable x_i by 1 in $\mathfrak{F}^h$. According to Lemma 5.2 applied to $\mathfrak{F}_i$, the variety $Z(\mathbf{I}(\mathfrak{F}_i, 0))$ is empty. Therefore, the set $W(\mathbf{I}(\mathfrak{F}^h, 0)) \cap O_i$ is also empty. Consequently, $W(\mathbf{I}(\mathfrak{F}^h, 0)) = \emptyset$. $\square$

We can now deduce the following result.

Lemma 5.4. *The projective variety $W(\mathbf{I}(\mathfrak{F}^h, 1)) \subset \mathbb{P}^{n-1}\overline{\mathbb{K}(\mathfrak{a})}$ is empty, and hence* $\dim(\mathbf{I}(\mathfrak{F}^h, 1)) = 0$.

Proof. We let φ_0 and φ_1 denote the two following morphisms:

$$\begin{array}{rrcl} \varphi_0 : & \mathbb{K}(\mathfrak{a})[x_1, \ldots, x_n] & \to & \mathbb{K}(\mathfrak{a})[x_2, \ldots, x_n] \\ & g(x_1, \ldots, x_n) & \mapsto & g(0, x_2, \ldots, x_n) \end{array}$$

$$\begin{array}{rrcl} \varphi_1 : & \mathbb{K}(\mathfrak{a})[x_1, \ldots, x_n] & \to & \mathbb{K}(\mathfrak{a})[x_2, \ldots, x_n] \\ & g(x_1, \ldots, x_n) & \mapsto & g(1, x_2, \ldots, x_n) \end{array}$$

Then $W(\mathbf{I}(\mathfrak{F}^h, 1))$ can be identified with the disjoint union of the variety $Z(\varphi_1(\mathbf{I}(\mathfrak{F}^h, 1))) \subset \overline{\mathbb{K}(\mathfrak{a})}^{n-1}$ and the projective variety $W(\varphi_0(\mathbf{I}(\mathfrak{F}^h, 1))) \subset \mathbb{P}^{n-2}\overline{\mathbb{K}(\mathfrak{a})}$.

- Notice that $\varphi_1(\mathbf{I}(\mathfrak{F}^h, 1)) = \mathbf{I}(\varphi_1(\mathfrak{F}^h), 0)$. Therefore, the ideal $\varphi_1(\mathbf{I}(\mathfrak{F}^h, 1)) \subset \overline{\mathbb{K}(\mathfrak{a})}[x_2, \ldots, x_n]$ is spanned by $\varphi_1(\mathbf{F}^h)$ (which is a generic system of degree D in $n-1$ variables) and by the maximal minors of its Jacobian matrix. According to Lemma 5.2, the variety $Z(\varphi_1(\mathbf{I}(\mathfrak{F}^h, 1)))$ is empty.
- Similarly, $\varphi_0(\mathbf{I}(\mathfrak{F}^h, 1)) = \mathbf{I}(\varphi_0(\mathfrak{F}^h), 0) \subset \mathbb{K}(\mathfrak{a})[x_2, \ldots, x_n]$ is generated by the homogeneous polynomials $\varphi_0(\mathfrak{F}^h)$ and by the maximal minors of the Jacobian matrix $\mathsf{jac}(\varphi_0(\mathfrak{F}^h))$. Thus, according to Corollary 5.3, the variety $W(\varphi_0(\mathbf{I}(\mathfrak{F}^h, 1)))$ is also empty.

$\square$

5.3 The homogeneous case

In this section, our goal is to estimate the degree of regularity of the ideal $\mathbf{I}(\mathfrak{F}^h, 1) \subset \mathbb{K}(\mathfrak{a})[X]$ which is a homogeneous ideal generated by $\mathfrak{F}^h$ and $\mathsf{MaxMinors}(\mathfrak{F}^h, 1)$ (see Notations 5.1). Recall that the degree of regularity $\mathrm{d}_{\mathrm{reg}}(I)$ of a 0-dimensional homogeneous ideal I is the smallest positive integer such that all monomials of degree $\mathrm{d}_{\mathrm{reg}}(I)$ are in I. Notice that $\mathrm{d}_{\mathrm{reg}}(I)$ is an upper bound on the degrees of the polynomials in a minimal Gröbner basis of I with respect to the grevlex ordering.

Theorem 5.5. *The degree of regularity of the ideal* $\mathbf{I}(\mathfrak{F}^h, 1)$ *is*

$$\mathrm{d_{reg}}(\mathbf{I}(\mathfrak{F}^h, 1)) = D(p-1) + (D-2)n + 2.$$

Notations 5.6. *To prove Theorem 5.5, we need to introduce a few more objects and notations.*

- *A set of new variables* $\{u_{i,j} : 1 \leq i \leq p, 2 \leq j \leq n\}$ *which is denoted by* U*;*
- *the determinantal ideal* $\mathscr{D} \subset \mathbb{K}[U]$ *generated by the maximal minors of the matrix*

$$\begin{bmatrix} u_{1,2} & \dots & u_{1,n} \\ \vdots & \vdots & \vdots \\ u_{p,2} & \dots & u_{p,n} \end{bmatrix}.$$

- $\mathfrak{g}_1, \dots, \mathfrak{g}_{p(n-1)}$ *which denote the polynomials* $u_{i,j} - \frac{\partial \mathfrak{f}_i^h}{\partial x_j}$*, for* $1 \leq i \leq p, 2 \leq j \leq n$ *and* $\mathfrak{g}_{p(n-1)+1}, \dots, \mathfrak{g}_{pn}$ *which denote the polynomials* $\mathfrak{f}_1^h, \dots, \mathfrak{f}_p^h$*;*
- *the ideals* $\mathfrak{I}_{(\ell)} = \mathscr{D} + \langle \mathfrak{g}_1, \dots, \mathfrak{g}_\ell \rangle \subset \mathbb{K}(\mathfrak{a})[U, X]$*;*
- *if* $g \in \mathbb{K}[X]$ *(resp.* $I \subset \mathbb{K}[X]$*) is a polynomial and* $\prec$ *is a monomial ordering (see e.g. [CLO97, Ch. 2, §2, Def. 1]),* $\mathsf{LM}_\prec(g)$ *(resp.* $\mathsf{LM}_\prec(I)$*) denotes its leading monomial (resp. the ideal generated by the leading monomials of the polynomials in* I*);*
- *a* degree ordering *is a monomial ordering* $\prec$ *such that for all pair of monomials* $m_1, m_2 \in \mathbb{K}[X]$*,* $\deg(m_1) < \deg(m_2)$ *implies* $m_1 \prec m_2$*.*

Obviously the polynomials $\mathfrak{g}_k$ for $1 \leq k \leq p(n-1)$ will be used to mimic the process of substituting the new variables $u_{i,j}$ by $\frac{\partial \mathfrak{f}_i^h}{\partial x_j}$; indeed we have $\mathfrak{I}_{(pn)} \cap \mathbb{K}[X] = \mathbf{I}(\mathfrak{F}^h, 1)$.

Our strategy to prove Theorem 5.5 will be to deduce the degree of regularity of $\mathbf{I}(\mathfrak{F}^h, 1)$ from an explicit form of its *Hilbert series*.

Proposition 5.7. *The Hilbert series of the homogeneous ideal* $\mathbf{I}(\mathfrak{F}^h, 1) \subset \mathbb{K}(\mathfrak{a})[X]$ *is*

$$\mathsf{HS}_{\mathbb{K}(\mathfrak{a})[X]/\mathbf{I}(\mathfrak{F}^h,1)}(t) = \frac{\det(A(t^{D-1}))}{t^{(D-1)\binom{p-1}{2}}} \frac{(1-t^D)^p(1-t^{D-1})^{n-p}}{(1-t)^n},$$

where $A(t)$ *is the* $(p-1) \times (p-1)$ *matrix whose* (i,j)*-entry is* $\sum_k \binom{p-i}{k}\binom{n-1-j}{k} t^k$*.*

The proof of Proposition 5.7 is postponed to Section 5.3.2.

Proof of Theorem 5.5. By definition, the Hilbert series of a zero-dimensional homogeneous ideal is a polynomial of degree $\mathrm{d_{reg}} - 1$. By Lemma 5.4, $\mathbf{I}(\mathfrak{F}^h, 1)$ has dimension 0. Thus, using Proposition 5.7, we deduce that:

$$\mathrm{d_{reg}}(\mathbf{I}(\mathfrak{F}^h, 1)) = 1 + \deg\left(\frac{\det(A(t^{D-1}))}{t^{(D-1)\binom{p-1}{2}}} \frac{(1-t^D)^p(1-t^{D-1})^{n-p}}{(1-t)^n}\right).$$

The highest degree on each row of $A(t)$ is reached on the diagonal. Thus $\deg(\det A(t)) = \frac{p(p-1)}{2}$ and a direct degree computation yields $\mathrm{d_{reg}}(\mathbf{I}(\mathfrak{F}^h, 1)) = D(p-1) + (D-2)n + 2$. $\square$

From Proposition 5.7, one can also deduce the degree of $\mathbf{I}(\mathfrak{F}^h, 1)$; this provides an alternate proof of [NR09, Theorem 2.2].

Corollary 5.8. *The degree of the ideal* $\mathbf{I}(\mathfrak{F}^h, 1)$ *is*

$$\mathsf{DEG}(\mathbf{I}(\mathfrak{F}^h, 1)) = \binom{n-1}{p-1} D^p (D-1)^{n-p}.$$

Proof. By definition of the Hilbert series, the degree of the 0-dimensional homogeneous ideal $\mathbf{I}(\mathfrak{F}^h, 1)$ is equal to $\mathsf{HS}_{\mathbb{K}(\mathfrak{a})[X]/\mathbf{I}(\mathfrak{F}^h,1)}(1)$. By Proposition 5.7, direct computations show that $\mathsf{HS}_{\mathbb{K}(\mathfrak{a})[X]/\mathbf{I}(\mathfrak{F}^h,1)}(1) = \det(A(1)) D^p (D-1)^{n-p}$. The determinant of the matrix $A(1)$ can be evaluated by using Vandermonde's identity and a formula by Harris-Tu (see e.g. [Ful97, Example 14.4.14, Example A.9.4]). We deduce that $\det(A(1)) = \binom{n-1}{p-1}$ and hence $\mathsf{HS}_{\mathbb{K}(\mathfrak{a})[X]/\mathbf{I}(\mathfrak{F}^h,1)}(1) = \binom{n-1}{p-1} D^p (D-1)^{n-p}$. □

It remains to prove Proposition 5.7. This is done in the next sections following several steps:

- provide an explicit form of the Hilbert series of the ideal $\mathscr{D}$; this is actually already done in [CH94]; we recall the statement of this result in Lemma 5.9;
- deduce from it an explicit form of Hilbert series of the ideal $\mathfrak{I}_{(pn)}$ using genericity properties satisfied by the polynomials $\mathfrak{g}_k$ and properties of quasi-homogeneous ideals; this is done in Lemma 5.10;
- deduce from it the Hilbert series associated to $\mathbf{I}(\mathfrak{F}^h, 1)$.

5.3.1 Auxiliary results

We start by restating a special case of [CH94, Cor. 1].

Lemma 5.9 ([CH94, Corollary 1])**.** *The Hilbert series of the ideal* $\mathscr{D} \subset \mathbb{K}[U]$ *is* $\mathsf{HS}_{\mathbb{K}[U]/\mathscr{D}}(t) = \frac{\det A(t)}{t^{\binom{p-1}{2}}(1-t)^{n(p-1)}}$.

Lemma 5.10. *For each* $2 \leq \ell \leq np$*,* $\mathfrak{g}_\ell$ *does not divide* 0 *in* $\mathbb{K}(\mathfrak{a})[U, X]/\mathfrak{I}_{(\ell-1)}$.

Proof. According to [HE70, Thm. 2][HE71], the ring $\mathbb{K}(\mathfrak{a})[U]/\mathscr{D}$ is a Cohen-Macaulay domain of Krull dimension $(n-1+p-(p-1))(p-1) = n(p-1)$. Therefore, the ring $\mathbb{K}(\mathfrak{a})[U, X]/\mathscr{D}$ is also a Cohen-Macaulay domain, and has dimension np.

Consider now the ideal $\langle \mathfrak{g}_1, \ldots, \mathfrak{g}_{np} \rangle \subset (\mathbb{K}(\mathfrak{a})[U]/\mathscr{D})[X]$. According to Lemma 5.4, the ideal $\mathbf{I}(\mathfrak{F}^h, 1) = (\mathscr{D} + \langle \mathfrak{g}_1, \ldots, \mathfrak{g}_{n(p-1)} \rangle) \cap \mathbb{K}(\mathfrak{a})[X]$ is zero-dimensional. Let $\prec$ denote a lexicographical monomial ordering such that for all i, j, k, $u_{i,j} \succ x_k$. Since the variables U can be expressed as functions of X ($u_{i,j} - \frac{\partial f_i}{\partial x_j} \in \mathfrak{I}_{(pn)}$), we have $\mathsf{LM}_\prec(\mathscr{D} + \langle \mathfrak{g}_1, \ldots, \mathfrak{g}_{np} \rangle) = \langle u_{i,j} \rangle + \mathsf{LM}_\prec(\mathbf{I}(\mathfrak{F}^h, 1))$ which is zero-dimensional. Therefore, the ideal $\mathscr{D} + \langle \mathfrak{g}_1, \ldots, \mathfrak{g}_{np} \rangle \subset \mathbb{K}(\mathfrak{a})[U, X]$ is zero-dimensional and hence so is $\langle \mathfrak{g}_1, \ldots, \mathfrak{g}_{np} \rangle \subset \mathbb{K}(\mathfrak{a})[U, X]/\mathscr{D}$. Now suppose by contradiction that there exists ℓ such that $\mathfrak{g}_\ell$ divides 0 in $\mathbb{K}(\mathfrak{a})[U, X]/\mathfrak{I}_{(\ell-1)}$. Let ℓ_0 be the smallest integer satisfying this property. Since $\mathscr{D}$ is equidimensional and for all $\ell < \ell_0$, $\mathfrak{g}_\ell$ does not divide 0 in $\mathbb{K}(\mathfrak{a})[U, X]/\mathfrak{I}_{(\ell-1)}$, the ideal $\langle \mathfrak{g}_1, \ldots, \mathfrak{g}_{\ell_0-1} \rangle \subset \mathbb{K}(\mathfrak{a})[U, X]/\mathscr{D}$ is equidimensional, has codimension $\ell_0 - 1$, and thus has no embedded components by the unmixedness Theorem [Eis95, Corollary 18.14]. Since $\mathfrak{g}_{\ell_0}$ divides 0 in the ring $\mathbb{K}(\mathfrak{a})[U, X]/(\mathscr{D} + \langle \mathfrak{g}_1, \ldots, \mathfrak{g}_{\ell_0-1} \rangle)$, the ideal $\langle \mathfrak{g}_1, \ldots, \mathfrak{g}_{\ell_0} \rangle \subset \mathbb{K}(\mathfrak{a})[U, X]/\mathscr{D}$ has also codimension $\ell_0 - 1$. Therefore the codimension of $\langle \mathfrak{g}_1, \ldots, \mathfrak{g}_{np} \rangle \subset \mathbb{K}(\mathfrak{a})[U, X]/\mathscr{D}$ is strictly less than np, which leads to a contradiction since we have proved that the dimension of this ideal is 0. □

The degrees in the matrix whose entries are the variables $u_{i,j}$ have to be balanced with $D-1$, the degree of the partial derivatives. This is done by changing the gradation by putting a *weight* on the variables $u_{i,j}$, giving rise to *quasi-homogeneous* polynomials.

The following lemma and its proof are similar to Lemma 4.8.

Lemma 5.11. *The Hilbert series of* $\mathbf{I}(\mathfrak{F}^h, 1) \subset \mathbb{K}(\mathfrak{a})[X]$ *and the weighted Hilbert series of* $\mathfrak{I}_{(pn)} \subset \mathbb{K}(\mathfrak{a})[U, X]$ *are equal.*

Proof. Let $\prec_{\mathsf{lex}}$ be a lex ordering on the variables of the polynomial ring $\mathbb{K}(\mathfrak{a})[U, X]$ such that $x_k \prec_{\mathsf{lex}} u_{i,j}$ for all k, i, j. By [CLO97, Sec. 6.3, Prop. 9], $\mathsf{HS}_{\mathbb{K}(\mathfrak{a})[X]/\mathbf{I}(\mathfrak{F}^h,1)}(t) = \mathsf{HS}_{\mathbb{K}(\mathfrak{a})[X]/\,\mathsf{LM}_{\prec_{\mathsf{lex}}}(\mathbf{I}(\mathfrak{F}^h,1))}(t)$ and $\mathsf{wHS}_{\mathbb{K}(\mathfrak{a})[U,X]/\mathfrak{I}_{(p(n-1))}}(t) = \mathsf{wHS}_{\mathbb{K}(\mathfrak{a})[U,X]/\,\mathsf{LM}_{\prec_{\mathsf{lex}}}(\mathfrak{I}_{(p(n-1))})}(t)$. Since $\mathsf{LM}_{\prec_{\mathsf{lex}}}(u_{i,j} - f_{i,j}) = u_{i,j}$ and $\mathfrak{I}_{(pn)} \cap \mathbb{K}[X] = \mathbf{I}(\mathfrak{F}^h, 1)$, we deduce that

$$\begin{aligned} \mathsf{LM}_{\prec_{\mathsf{lex}}}(\mathfrak{I}_{(pn)}) &= \left\langle \{u_{i,j}\} \cup \mathsf{LM}_{\prec_{\mathsf{lex}}}(\mathfrak{I}_{(pn)} \cap \mathbb{K}(\mathfrak{a})[X]) \right\rangle \\ &= \left\langle \{u_{i,j}\} \cup \mathsf{LM}_{\prec_{\mathsf{lex}}}(\mathbf{I}(\mathfrak{F}^h, 1)) \right\rangle. \end{aligned}$$

Therefore, $\frac{\mathbb{K}(\mathfrak{a})[U,X]}{\mathsf{LM}_{\prec_{\mathsf{lex}}}(\mathfrak{I}_{(pn)})}$ is isomorphic (as a graded $\mathbb{K}(\mathfrak{a})$-algebra) to $\frac{\mathbb{K}(\mathfrak{a})[X]}{\mathsf{LM}_{\prec_{\mathsf{lex}}}(\mathbf{I}(\mathfrak{F}^h,1))}$.

Thus, $\mathsf{HS}_{\mathbb{K}(\mathfrak{a})[X]/\,\mathsf{LM}_{\prec_{\mathsf{lex}}}(\mathbf{I}(\mathfrak{F}^h,1))}(t) = \mathsf{wHS}_{\mathbb{K}(\mathfrak{a})[U,X]/\,\mathsf{LM}_{\prec_{\mathsf{lex}}}(\mathfrak{I}_{(pn)})}(t)$, and hence $\mathsf{HS}_{\mathbb{K}(\mathfrak{a})[X]/\mathbf{I}(\mathfrak{F}^h,1)}(t) = \mathsf{wHS}_{\mathbb{K}(\mathfrak{a})[U,X]/\mathfrak{I}_{(pn)}}(t)$. □

5.3.2 Proof of Proposition 5.7

We reuse Notations 5.6: $\mathbf{I}(\mathfrak{F}^h, 1) = (\mathscr{D} + \langle \mathfrak{g}_1, \ldots, \mathfrak{g}_{pn} \rangle) \cap \mathbb{K}(\mathfrak{a})[X]$. According to Lemma 5.9 and by putting a weight $D-1$ on the variables U, the weighted Hilbert series of $\mathscr{D} \subset \mathbb{K}(\mathfrak{a})[U]$ is

$$\mathsf{wHS}_{\mathbb{K}(\mathfrak{a})[U]/\mathscr{D}}(t) = \frac{\det A(t^{D-1})}{t^{(D-1)\binom{p-1}{2}}(1-t^{D-1})^{n(p-1)}}.$$

Considering $\mathscr{D}$ as an ideal of $\mathbb{K}(\mathfrak{a})[U, X]$, we obtain

$$\mathsf{wHS}_{\mathbb{K}(\mathfrak{a})[U,X]/\mathscr{D}}(t) = \frac{1}{(1-t)^n}\,\mathsf{wHS}_{\mathbb{K}(\mathfrak{a})[U]/\mathscr{D}}(t).$$

If $I \subset \mathbb{K}(\mathfrak{a})[U, X]$ is a quasi-homogeneous ideal and if g is a quasi-homogeneous polynomial of weight degree d which does not divide 0 in the quotient ring $\mathbb{K}(\mathfrak{a})[U, X]/I$, then the Hilbert series of the ideal $I + \langle g \rangle$ is equal to $(1 - t^d)$ multiplied by the Hilbert series of I (Proposition 1.41).

Notice that the polynomials $\mathfrak{g}_1, \ldots, \mathfrak{g}_{p(n-1)}$ are quasi-homogeneous of weight degree $D-1$ (these polynomials have the form $u_{i,j} - \frac{\partial \mathfrak{f}_i}{\partial x_j}$) and the polynomials $\mathfrak{g}_{p(n-1)+1}, \ldots, \mathfrak{g}_{pn}$ are quasi-homogeneous of weight degree D (these polynomials are $\mathfrak{f}_1, \ldots, \mathfrak{f}_p$). Since $\mathfrak{g}_\ell$ does not divide 0 in $\mathbb{K}(\mathfrak{a})[U, X]/\mathfrak{I}_{(\ell-1)}$ (Lemma 5.10), the Hilbert series of the ideal $\mathfrak{I}_{(pn)} \subset \mathbb{K}(\mathfrak{a})[U, X]$ is

$$\mathsf{wHS}_{\mathbb{K}(\mathfrak{a})[U,X]/\mathfrak{I}_{(pn)}}(t) = \frac{\det A(t^{D-1})}{t^{(D-1)\binom{p-1}{2}}} \frac{(1-t^D)^p(1-t^{D-1})^{n-p}}{(1-t)^n}.$$

Finally, by Lemma 5.11, $\mathsf{wHS}_{\mathbb{K}(\mathfrak{a})[U,X]/\mathfrak{I}_{(pn)}}(t) = \mathsf{HS}_{\mathbb{K}(\mathfrak{a})[X]/\mathbf{I}(\mathfrak{F}^h,1)}(t)$.

5.4 The affine case

The degree of regularity of a polynomial system is the highest degree reached during the computation of a Gröbner basis with respect to the grevlex ordering with the F_5 algorithm. Therefore, it is a crucial indicator of the complexity of the Gröbner basis computation. On the other hand, the complexity of the FGLM algorithm depends on the degree of the ideal $\mathbf{I}(\mathbf{F}, 1)$ since this value is equal to $\dim_{\mathbb{K}} (\mathbb{K}[X]/\mathbf{I}(\mathbf{F}, 1))$.

In this section, we show that the bounds on the degree and the degree of regularity of the ideal $\mathbf{I}(\mathfrak{F}^h, 1)$ are also valid for (not necessarily homogeneous) polynomial families in $\mathbb{K}[X]$ under genericity assumptions.

Theorem 5.12. *There exists a non-empty Zariski open subset $\mathscr{O} \subset \overline{\mathbb{K}}[X]_D^p$ such that, for any $\mathbf{F}$ in $\mathscr{O} \cap \mathbb{K}[X]^p$,*

$$\begin{array}{rcl} \mathrm{d_{reg}}(\mathbf{I}(\mathbf{F}, 1)) & \leq & D(p-1) + (D-2)n + 2, \\ \mathsf{DEG}(\mathbf{I}(\mathbf{F}, 1)) & \leq & \binom{n-1}{p-1} D^p (D-1)^{n-p}. \end{array}$$

In the sequel, $\overline{\mathbb{K}}[X]_D$ denotes $\{f \in \overline{\mathbb{K}}[X] \mid \deg(f) = D\}$, and $\overline{\mathbb{K}}[X]_{D,\mathsf{hom}}$ denotes the homogeneous polynomials in $\overline{\mathbb{K}}[X]_D$. In order to prove Theorem 5.12 (the proof is postponed to the end of this section), we first need two technical lemmas.

Lemma 5.13. *There exists a non-empty Zariski open subset $\mathscr{O} \subset \overline{\mathbb{K}}[X]_{D,\mathsf{hom}}^p$ such that for all $\mathbf{F}^h \in \mathscr{O} \cap \mathbb{K}[X]^p$, $\mathsf{LM}_\prec(\mathbf{I}(\mathbf{F}^h, 1)) = \mathsf{LM}_\prec(\mathbf{I}(\mathfrak{F}^h, 1))$.*

Proof. See e.g. [FSS11b, Proof of Lemma 2] for a similar proof. □

Lemma 5.14. *Let $G = (g_1, \dots, g_m)$ be a polynomial family and let $G^h = (g_1^h, \dots, g_m^h)$ denote the family of homogeneous components of highest degree of G. If the dimension of the ideal $\langle G^h \rangle$ is 0, then $\mathsf{DEG}(\langle G \rangle) \leq \mathsf{DEG}(\langle G^h \rangle)$.*

Proof. Let $\prec$ be an admissible degree monomial ordering. Let $\mathsf{LM}_\prec(h)$ denote the leading monomial of a polynomial h with respect to $\prec$. Let $m \in \mathsf{LM}_\prec(\langle G^h \rangle)$ be a monomial. Then there exist polynomials $s_1, \dots, s_m$ such that $\mathsf{LM}_\prec\left(\sum_{i=1}^m s_i g_i^h\right) = m$. Since $\prec$ is a degree ordering, $\mathsf{LM}_\prec\left(\sum_{i=1}^m s_i g_i\right) = m$. Therefore $\mathsf{LM}_\prec(\langle G^h \rangle) \subset \mathsf{LM}_\prec(\langle G \rangle)$. If the ideal $\langle G^h \rangle$ is 0-dimensional, then so is $\langle G \rangle$ and $\mathsf{DEG}(\mathsf{LM}_\prec(\langle G \rangle)) \leq \mathsf{DEG}(\mathsf{LM}_\prec(\langle G \rangle))$. Since $\mathsf{DEG}(I) = \mathsf{DEG}(\mathsf{LM}_\prec(I))$, we obtain $\mathsf{DEG}(\langle G \rangle) \leq \mathsf{DEG}(\langle G^h \rangle)$.

□

Proof of Theorem 5.12. Let $\prec$ be a degree monomial ordering, and $\mathbf{F}^h = (f_1^h, \dots, f_p^h) \in \overline{\mathbb{K}}[X]_{D,\mathsf{hom}}^p$ denote the homogeneous system where f_i^h is the homogeneous component of highest degree of f_i. By Lemma 5.13, there exists a non-empty Zariski subset $\mathscr{O} \subset \overline{\mathbb{K}}[X]_D^p$ such that, for any $\mathbf{F}$ in $\mathscr{O} \cap \mathbb{K}[X]^p$, $\mathsf{LM}_\prec(\mathbf{I}(\mathbf{F}^h, 1)) = \mathsf{LM}_\prec(\mathbf{I}(\mathfrak{F}^h, 1))$. By [CLO97, Ch.9, §3, Prop.9], the Hilbert series (and thus the dimension, the degree, and the degree of regularity) of a homogeneous ideal is the same as that of its leading monomial ideal. Hence, by Lemma 5.4,

$$\begin{aligned} \dim(\mathbf{I}(\mathbf{F}^h, 1)) = \dim(\mathsf{LM}_\prec(\mathbf{I}(\mathbf{F}^h, 1))) &= \dim(\mathsf{LM}_\prec(\mathbf{I}(\mathfrak{F}^h, 1))) \\ &= \dim(\mathbf{I}(\mathfrak{F}^h, 1)) = 0. \end{aligned}$$

Similarly, by Theorem 5.5,

$$\mathrm{d_{reg}}(\mathbf{I}(\mathbf{F}^h, 1)) = \mathrm{d_{reg}}(\mathbf{I}(\mathfrak{F}^h, 1)) = D(p-1) + (D-2)n + 2.$$

The highest degree reached during the F_5 Algorithm is bounded above by the degree of regularity of the ideal generated by the homogeneous components of highest degree of the generators when this homogeneous ideal has dimension 0 (see e.g. [BFSY04] and references therein). Therefore, the highest degree reached during the computation of a Gröbner basis of $\mathbf{I}(\mathbf{F}, 1)$ with the F_5 Algorithm with respect to a degree ordering is bounded above by

$$\mathrm{d}_{\mathrm{reg}} \leq D(p-1) + (D-2)n + 2.$$

The bound on the degree is obtained by Corollary 5.8 and Lemma 5.14,

$$\begin{aligned} \mathsf{DEG}(\mathbf{I}(\mathbf{F}, 1)) \leq \mathsf{DEG}(\mathbf{I}(\mathbf{F}^h, 1)) &\leq \mathsf{DEG}(\mathsf{LM}_{\prec}(\mathbf{I}(\mathfrak{F}^h, 1))) \\ &\leq \tbinom{n-1}{p-1} D^p (D-1)^{n-p}. \end{aligned}$$

□

5.5 Complexity

In the sequel, ω is a real number such that there exists an algorithm which computes the row echelon form of $n \times n$ matrix in $O(n^\omega)$ arithmetic operations (the best known value is $\omega \approx 2.376$ by using the Coppersmith-Winograd algorithm, see [Sto00]).

Theorem 5.15. *There exists a non-empty Zariski open subset $\mathscr{O} \subset \overline{\mathbb{K}}[X]_D^p$, such that, for all $\mathbf{F} \in \mathscr{O} \cap \mathbb{K}[X]^p$, the arithmetic complexity of computing a lexicographical Gröbner basis of $\mathbf{I}(\mathbf{F}, 1)$ is bounded by*

$$O\left(\left(p + \binom{n-1}{p}\right)\binom{D(p-1) + (D-1)n + 2}{D(p-1) + (D-2)n + 2}^{\omega} + n\binom{n-1}{p-1}^{3} D^{3p}(D-1)^{3(n-p)}\right).$$

Proof. According to [BFS04, BFSY04], the complexity of computing a Gröbner basis with the F_5 Algorithm with respect to the grevlex ordering of a zero-dimensional ideal is bounded by $O\left(m\binom{n+\mathrm{d}_{\mathrm{reg}}}{\mathrm{d}_{\mathrm{reg}}}^{\omega}\right)$ where $\mathrm{d}_{\mathrm{reg}}$ is the highest degree reached during the computation and m is the number of polynomials generating the ideal. In order to obtain a lexicographical Gröbner basis, one can use the FGLM algorithm [FGLM93]. Its complexity is $O\left(n\,\mathsf{DEG}(\mathbf{I}(\mathbf{F}, 1))^3\right)$ (better complexity bounds are known in specific cases, see [FM11]).

According to Theorem 5.12, there exists a non-empty Zariski open subset $\mathscr{O} \subset \overline{\mathbb{K}}[X]_D^p$ such that, for all $\mathbf{F}$ in $\mathscr{O} \cap \mathbb{K}[X]^p$,

$$\begin{array}{rcl} \mathrm{d}_{\mathrm{reg}}(\mathbf{I}(\mathbf{F}, 1)) & \leq & D(p-1) + (D-2)n + 2, \\ \mathsf{DEG}(\mathbf{I}(\mathbf{F}, 1)) & \leq & \binom{n-1}{p-1} D^p (D-1)^{n-p}. \end{array}$$

Therefore, for all $\mathbf{F}$ in $\mathscr{O} \cap \mathbb{K}[X]^p$, the total complexity of computing a lexicographical Gröbner basis of $\mathbf{I}(\mathbf{F}, 1)$:

$$O\left(\left(p + \binom{n-1}{p}\right)\binom{D(p-1) + (D-1)n + 2}{D(p-1) + (D-2)n + 2}^{\omega} + n\binom{n-1}{p-1}^{3} D^{3p}(D-1)^{3(n-p)}\right).$$

□

Corollary 5.16. *If $D = 2$, then there exists a non-empty Zariski open subset $\mathscr{O} \subset \overline{\mathbb{K}}[X]_2^p$, such that for all $\mathbf{F} \in \mathscr{O} \cap \mathbb{K}[X]^p$, the arithmetic complexity of computing a lexicographical Gröbner basis of $\mathbf{I}(\mathbf{F}, 1)$ is bounded by*

$$O\left(\left(p + \binom{n-1}{p}\right)\binom{n+2p}{2p}^{\omega} + n2^{3p}\binom{n-1}{p-1}^3\right).$$

Moreover, if p is constant and $D = 2$, the arithmetic complexity is bounded by $O\left(n^{p(2\omega+1)}\right)$.

Proof. This complexity is obtained by putting $D = 2$ in the formula from Theorem 5.15. □

In the sequel, the binary entropy function is denoted by h_2:

$$\forall x \in [0,1], h_2(x) = -x\log_2(x) - (1-x)\log_2(1-x).$$

Corollary 5.17. *Let $D > 2$ and $p \in \mathbb{N}$ be constant. There exists a non-empty Zariski open subset $\mathscr{O} \subset \overline{\mathbb{K}}[X]_D^p$, such that, for all $\mathbf{F} \in \mathscr{O} \cap \mathbb{K}[X]^p$, the arithmetic complexity of computing a lexicographical Gröbner basis of $\mathbf{I}(\mathbf{F}, 1)$ is bounded by*

$$O\left(\frac{n^p}{\sqrt{n}} 2^{(D-1)h_2\left(\frac{1}{D-1}\right)n\omega}\right) = \widetilde{O}\left((D-1)^{3.57n}\right).$$

Proof. Let x be a real number in $[0,1]$. Then by applying Stirling's Formula, we obtain that $\binom{n}{xn} = O\left(\frac{1}{\sqrt{n}}2^{h_2(x)n}\right)$. Therefore,

$$\begin{aligned}\binom{(D-1)n}{n} &= O\left(\tfrac{1}{\sqrt{n}}2^{(D-1)h_2\left(\frac{1}{D-1}\right)n}\right)\\ &= O\left(\tfrac{1}{\sqrt{n}}((D-1)e)^n\right).\end{aligned}$$

Let C denote the constant $D(p-1)+2$. Then

$$\begin{aligned}\binom{D(p-1)+(D-1)n+2}{D(p-1)+(D-2)n+2} &= \binom{(D-1)n+C}{n} = O\left(\binom{(D-1)n}{n}\right)\\ &= O\left(\tfrac{1}{\sqrt{n}}2^{(D-1)h_2\left(\frac{1}{D-1}\right)n}\right).\end{aligned}$$

The right summand in the complexity formula given in Theorem 5.15 is $O\left(n^{3p}(D-1)^{3n}\right)$ when p and D are constants; this is bounded by $O\left(\frac{1}{\sqrt{n}}2^{(D-1)h_2\left(\frac{1}{D-1}\right)n\omega}\right)$. Let $\mathscr{O}$ be the non-empty Zariski open subset defined in Theorem 5.15. For all $\mathbf{F} \in \mathscr{O} \cap \mathbb{K}[X]^p$, the arithmetic complexity of computing a grevlex Gröbner basis of $\mathbf{F}$ is bounded by

$$\begin{aligned}O\left(\tfrac{n^p}{\sqrt{n}}2^{(D-1)h_2\left(\frac{1}{D-1}\right)n\omega}\right) &= O\left(\tfrac{n^p}{\sqrt{n}}((D-1)e)^{n\omega}\right)\\ &= \widetilde{O}\left((D-1)^{(1+1/\log(D-1))n\omega}\right)\\ &= \widetilde{O}\left((D-1)^{3.57n}\right),\end{aligned}$$

since $D \geq 3$ and $\omega \leq 2.376$ with the Coppersmith-Winograd algorithm. On the other hand the asymptotic complexity of the FGLM part of the solving process is

$$O\left(n^{3(p-1)+1}(D-1)^{3n}\right) = \widetilde{O}\left((D-1)^{3n}\right),$$

which is bounded above by the complexity of the grevlex Gröbner basis computation. □

The following corollary shows that the arithmetic complexity is polynomial in the number of critical points.

Corollary 5.18. *For $D \geq 3$, $p \geq 2$ and $n \geq 2$, there exists a non-empty Zariski open subset $\mathscr{O} \subset \overline{\mathbb{K}}[X]_D^p$, such that, for $\mathbf{F} \in \mathscr{O} \cap \mathbb{K}[X]^p$, the arithmetic complexity of computing a lexicographical Gröbner basis of $\mathbf{I}(\mathbf{F}, 1)$ is bounded by*

$$\widetilde{O}\left(\mathsf{DEG}\left(\mathbf{I}(\mathbf{F},1)\right)^{\max\left(\frac{\log(2eD)}{\log(D-1)}\omega, 4\right)}\right) \leq O\left(\mathsf{DEG}\left(\mathbf{I}(\mathbf{F},1)\right)^{4.03\omega}\right).$$

Proof. Let $\mathscr{O} \subset \overline{\mathbb{K}}[X]_D^p$ be the non-empty Zariski open subset defined in Theorem 5.12, and $\mathbf{F} \in \mathscr{O} \cap \mathbb{K}[X]_D^p$ be a polynomial family. First, notice that, since $p \geq 2$ and $n \geq 2$,

$$\begin{array}{rcl} \mathsf{DEG}\left(\mathbf{I}(\mathbf{F},1)\right) & = & \binom{n-1}{p-1}(D-1)^{n-p}D^p \\ & \geq & n \end{array}$$

Therefore the complexity of the FGLM algorithm is bounded by $O\left(n\,\mathsf{DEG}\left(\mathbf{I}(\mathbf{F},1)\right)^3\right) \leq O\left(\mathsf{DEG}\left(\mathbf{I}(\mathbf{F},1)\right)^4\right)$. The complexity of computing a grevlex Gröbner basis of $\mathbf{I}(\mathbf{F}, 1)$ is bounded by

$$\begin{aligned} \mathsf{GREVLEX}(p,n,D) &= O\left(\binom{n-1}{p}\binom{D(p-1)+(D-1)n+2}{n}^{\omega}\right) \\ &\leq O\left(\binom{n-1}{p}\binom{2Dn}{n}^{\omega}\right). \end{aligned}$$

Notice that $\binom{2Dn}{n} \leq (2D)^n \frac{n^n}{n!}$. By Stirling's formula, there exists C_0 such that $\frac{n^n}{n!} \leq C_0 e^n$. Hence $\mathsf{GREVLEX}(p,n,D) = \widetilde{O}\left((2De)^n\right)$.

Since $D \geq 3$ and $n \leq \log(\mathsf{DEG}(\mathbf{I}(\mathbf{F},1)))/\log(D-1)$, we obtain

$$\begin{array}{rcl} O\left((2De)^{n\omega}\right) & \leq & O\left(D^{\frac{\log(2eD)}{\log D}n\omega}\right) \\ & \leq & O\left(\mathsf{DEG}\left(\mathbf{I}(\mathbf{F},1)\right)^{\frac{\log(2eD)}{\log(D-1)}\omega}\right). \end{array}$$

The function $D \mapsto \frac{\log(2eD)}{\log(D-1)}$ is decreasing, and hence its maximum is reached for $D = 3$, and $\frac{\log(6e)}{\log(2)} \leq 4.03$. $\square$

Notice that in the complexity formula in Corollary 5.18, the exponent $\frac{\log(2eD)}{\log(D-1)}\omega$ tends towards ω when D grows. Therefore, when D is large, the complexity of the grevlex Gröbner basis computation is close to the cost of linear algebra $O\left(\mathsf{DEG}(\mathbf{I}(\mathbf{F},1))^{\omega}\right)$. Also, we would like to point out that the bound in Corollary 5.18 is not sharp since the formula $O\left(m\binom{n+\mathrm{d_{reg}}}{n}^{\omega}\right)$ for the complexity of the F_5 algorithm is pessimistic, and the majorations performed in the proof of Corollary 5.18 are not tight.

5.6 Experimental Results

In this section, we report experimental results supporting the theoretical complexity results in the previous sections. Since our complexity results concern the arithmetic complexity, we run experiments where $\mathbb{K}$ is the finite field $\mathsf{GF}(65521)$ (Tables 5.1 and 5.2), so that the timings represent the arithmetic complexity. In that case, systems are chosen uniformly at random in $\mathsf{GF}(65521)[X]_D$.

n	p	D	$\mathrm{d_{reg}}$	DEG	F_4 time	FGLM time
9	4	2	8	896	3.12s	18.5s
11	4	2	8	1920	61s	202s
13	4	2	8	3520	369s	1372s
15	4	2	8	5824	2280s	7027s
17	4	2	8	8960	10905s	>1d
30	2	2	4	116	3.00s	0.14s
35	2	2	4	136	7.5s	0.36s
40	2	2	4	156	13.3s	0.64s
6	4	3	17	3240	16s	400s
8	4	3	19	45360	35593s	>1d
7	2	3	12	1728	9.9s	91s
8	2	3	13	4032	121s	1169s
9	2	3	14	9216	736s	>1d

Table 5.1: Timings using MAGMA and $\mathbb{K} = \mathsf{GF}(65521)$.

n	p	D	$\mathsf{DEG}(\mathbf{I}(\mathbf{F}, 1))$	F_5 time	FGLM time	matrix density
16	3	2	840	2.20s	0.03s	36.91%
18	3	2	1088	4.62s	0.12s	37.00%
20	3	2	1368	9.54s	0.10s	37.07%
15	4	2	5824	131.65	10.66s	33.53%
17	4	2	8960	480.9s	68.9s	34.00%
19	4	2	13056	1600.1s	215.1s	34.35%
21	4	2	18240	10371.7s	590.3s	34.62%
10	1	3	1536	1.5s	0.15s	20.84%
12	1	3	6144	19.6s	2.46s	19.32%
14	1	3	24576	1759s	587s	18.08%
7	2	3	1728	1.4s	0.14s	20.73%
9	2	3	9216	105s	37s	19.47%
10	2	3	20736	909s	504s	19.08%
7	3	3	6480	31.3s	3.81s	17.39%
8	4	3	45360	5126.9s	3833.9s	15.15%
8	2	4	81648	21362.6s	19349.4s	13.26%
7	3	4	77760	13856.8s	16003s	11.83%

Table 5.2: Timings using the FGb library and $\mathbb{K} = \mathsf{GF}(65521)$.

n	p	D	$\log \binom{n+\mathrm{d_{reg}}}{n} / \log(\mathsf{DEG})$
5	4	3	1.53
10	4	3	1.36
100	4	3	1.73
10000	4	3	1.99
10000	9999	3	2.28
30000	29999	3	2.28
1000	500	3	1.32
20000	2	3	2.00
500	250	1000	1.09
500	2	10000	1.11

Table 5.3: Numerical values: $\log \binom{n+\mathrm{d_{reg}}}{n} / \log\left(\mathsf{DEG}(\mathbf{I}(\mathbf{F},1))\right)$.

We give experiments by using respectively the implementation of F_4 and FGLM algorithms in the MAGMA Computer Algebra Software, and by using the F_5 and FGLM implementations from the FGb package.

Experiments were conducted on a 2.93GHz Intel Xeon E7220 with 128 GB RAM.

Interpretation of the results. Notice that the degree of regularity and the degree match exactly the bounds given in Theorem 5.12. In Tables 5.1 and 5.2, we can see a different behavior when $D = 2$ or $D = 3$. In the case $D = 2$, since the complexity is polynomial in n (Corollary 5.16), the computations can be performed even when n is large (close to 20). Moreover, notice that for $D = 2$ or $D = 3$, there is a strong correlation between the degree of the ideal and the timings, showing that, in accordance with Corollary 5.18, this degree is a good indicator of the complexity.

Also, in Table 5.2, we give the proportion of non-zero entries in the multiplication matrices. This proportion plays an important role in the complexity of FGLM, since recent versions of FGLM take advantage of this sparsity [FM11]. We can notice that the sparsity of the multiplication matrices increases as D grows.

Numerical estimates of the complexity. Corollary 5.18 states that the complexity of the grevlex Gröbner basis computation is bounded by $O\left(\mathsf{DEG}(\mathbf{I}(\mathbf{F},1))^{4.03\omega}\right)$ when $D \geq 3$, $p \geq 2$, $n \geq 2$. However, the value 4.03 is rather pessimistic. In Table 5.3, we report numerical values of the ratio $\log \binom{n+\mathrm{d_{reg}}}{n} / \log\left(\mathsf{DEG}(\mathbf{I}(\mathbf{F},1))\right)$ which show the difference between 4.03 and experimental values.

Notice that all ratios are smaller than 4.03, as predicted by Corollary 5.18. Experimentally, the ratio decreases and tends towards 1 when D grows, in accordance with the complexity formula

$$O\left(\mathsf{DEG}\left(\mathbf{I}(\mathbf{F},1)\right)^{\frac{\log(2eD)}{\log(D-1)}\omega}\right)$$

for the grevlex Gröbner basis computation. Also, when $D \geq 3$, the worst ratio seems to be reached when $p = n-1$, $D = 3$ and n grows, and experiments in Table 5.3 tend to show that it is bounded from above by 2.28.

Systems with rational coefficients. In applications, the critical points appearing are most often with rational coefficients. However, by using a multi-modular approach, the bit complexity of the lexicographical Gröbner basis computation will be quasi-linear in the heights of these coefficients. Therefore, the whole bit complexity will still be polynomial in the bit size of the output (the lex Gröbner basis). For instance, with the FGb library, the lex Gröbner basis of a critical point system with $p = 1$, $D = 4$ and $n = 7$ and integer coefficients between -99 and 99 was computed in 45 minutes.

Nevertheless, it is still an interesting question to obtain good theoretical bounds on the heights of the polynomials in the lex Gröbner basis of critical point systems – in particular in order to know if the bit complexity is still polynomial in the number variables in the case $D = 2$. We plan to investigate these issues in future works.

5.7 Mixed systems

This section is devoted to the study of mixed critical point systems: the polynomials f_i do not necessarily share the same degree. The general strategy to obtain complexity results is similar to the one followed in the unmixed case: the main tool is the Hilbert series of the ideal $\mathbf{I}(\mathbf{F}, 1)$ vanishing on the critical points. However, since the degrees of the f_i's are different, we cannot directly use the combinatorial properties of the determinantal ideals, nor the explicit formula for the Hilbert series of the ideal $\mathscr{D}$. To avoid this problem, we use the fact that a free resolution of the ideal $\mathscr{D}$ is given by the so-called *Eagon-Northcott complex*. From this free resolution, we can read off an explicit formula for the degree of regularity of $\mathbf{I}(\mathbf{F}, 1)$ and obtain complexity bounds for the Gröbner basis computation.

5.7.1 Preliminaries on the Eagon-Northcott complex

The Eagon-Northcott complex is an explicit complex which gives a minimal free resolution of ideals generated by maximal minors of polynomial matrices, under some assumptions which are satisfied generically.

Consider the following example, where $n = 5$ and $p = 2$. We want a free resolution of the ideal $\mathscr{D}$ generated by the maximal minors of the matrix

$$\begin{pmatrix} u_{1,2} & \cdots & u_{1,5} \\ u_{2,2} & \cdots & u_{2,5} \end{pmatrix}$$

In this case, the Eagon-Northcott complex is

$$\mathsf{EN}: \quad 0 \to R^3 \xrightarrow{\delta_3} R^8 \xrightarrow{\delta_2} R^6 \xrightarrow{\delta_1} R\ ,$$

where the letter R stands for the ring $\mathbb{K}[U]$, and where the morphisms δ_i are given by the following matrices:

$$\begin{aligned} \delta_1 = \ & (u_{1,3}u_{2,5} - u_{2,3}u_{1,5}, \quad u_{1,4}u_{2,3} - u_{1,3}u_{2,4}, \quad u_{1,2}u_{2,4} - u_{1,4}u_{2,2}, \\ & u_{1,5}u_{2,2} - u_{2,5}u_{1,2}, \quad u_{1,5}u_{2,4} - u_{2,5}u_{1,4}, \quad u_{1,3}u_{2,2} - u_{2,3}u_{1,2}) \end{aligned}$$

$$\delta_2 = \begin{pmatrix} -u_{1,4} & -u_{2,4} & 0 & 0 & u_{1,2} & u_{2,2} & 0 & 0 \\ -u_{1,5} & -u_{2,5} & 0 & 0 & 0 & 0 & u_{1,2} & u_{2,2} \\ 0 & 0 & -u_{1,5} & -u_{2,5} & 0 & 0 & u_{1,3} & u_{2,3} \\ 0 & 0 & -u_{1,4} & -u_{2,4} & u_{1,3} & u_{2,3} & 0 & 0 \\ -u_{1,3} & -u_{2,3} & u_{1,2} & u_{2,2} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & -u_{1,5} & -u_{2,5} & u_{1,4} & u_{2,4} \end{pmatrix} \qquad \delta_3 = \begin{pmatrix} u_{1,2} & u_{2,2} & 0 \\ 0 & u_{1,2} & u_{2,2} \\ u_{1,3} & u_{2,3} & 0 \\ 0 & u_{1,3} & u_{2,3} \\ u_{1,4} & u_{2,4} & 0 \\ 0 & u_{1,4} & u_{2,4} \\ u_{1,5} & u_{2,5} & 0 \\ 0 & u_{1,5} & u_{2,5} \end{pmatrix}$$

Direct computations show that this is a complex (since for all i, $\delta_{i-1} \circ \delta_i = 0$) and it is clear that $R/\mathsf{Im}(\delta_1)$ is isomorphic to $R/\mathscr{D}$. The fact that this complex is exact (i.e. for all i, $\mathsf{Im}(\delta_i) = \mathsf{Ker}(\delta_{i-1})$) is more difficult to prove, and we refer the reader to [Eis01, Appendix A2H] for a more detailed presentation of the properties of this free resolution.

The next step is to take into account the quasi-homogeneous grading $\mathrm{wdeg}(u_{i,j}) = d_i - 1$. We use here the classical notation $R[-s]$ to denote the ring R where the grading has been shifted by s. For instance, if $\eta : R \to R$ is a morphism of degree s, we write

$$R[-s] \xrightarrow{\eta} R[0]$$

to express the fact that η maps an element of degree d to an element of degree $d + s$. The Eagon-Northcott complex can then be rewritten as

$$\begin{aligned}\mathsf{EN}:\quad & 0 \to R[-3d_1-d_2-4] \oplus R[-2d_1-2d_2-4] \oplus R[-d_1-3d_2-4] \xrightarrow{\delta_4} \\ & (R[-2d_1-d_2-3] \oplus R[-d_1-2d_2-3])^4 \xrightarrow{\delta_3} R[-d_1-d_2-2]^6 \xrightarrow{\delta_2} R[0] \xrightarrow{\delta_1} R/\mathscr{D} \to 0.\end{aligned}$$

This approach can be generalized as follows. Let R be a ring. Following the notations in [Eis01, Appendix A2H], we write $F = R^f$ and $G = R^g$, where f and g are two integers such that $g < f$. For a $g \times f$ matrix whose entries are in R, we let $\alpha : F \to G$ denote the corresponding morphism of modules. The Eagon-Northcott complex is then defined by:

$$\begin{aligned}\mathsf{EN}(\alpha):\quad & 0 \to (\mathsf{Sym}_{f-g}G)^* \otimes \textstyle\bigwedge^f F \xrightarrow{\delta_{f-g-1}} (\mathsf{Sym}_{f-g-1}G)^* \otimes \textstyle\bigwedge^{f-1} F \xrightarrow{\delta_{f-g}} \\ & \cdots \to (\mathsf{Sym}_2 G)^* \otimes \textstyle\bigwedge^{g+2} F \xrightarrow{\delta_3} G^* \otimes \textstyle\bigwedge^{g+1} F \xrightarrow{\delta_2} \textstyle\bigwedge^g F \xrightarrow{\wedge^g \alpha} \textstyle\bigwedge^g G,\end{aligned}$$

where $\mathsf{Sym}_i G$ is the R-module of elements of order i in the symmetric algebra $\mathsf{Sym}(G)$.

First, notice that as a R-module, $\mathsf{Sym}_i G$, (and hence also its dual $(\mathsf{Sym}_i G)^*$) is a free module isomorphic to $R^{\binom{i+g-1}{i}}$. Similarly, $\bigwedge^i F$ is isomorphic to $R^{\binom{f}{i}}$.

For a detailed description of the maps δ_i, we refer the reader to [Eis01, Appendix A2H],[Eis95, Appendix A2.6]. In the context of this chapter, $f = n-1$, $g = p$, $R = \mathbb{K}[U]$ with the gradation given by $\mathrm{wdeg}(u_{i,j}) = d_i - 1$, and the morphism α corresponds to the matrix

$$\mathscr{U} = \begin{pmatrix} u_{1,2} & \dots & u_{1,n} \\ \vdots & \vdots & \vdots \\ u_{p,2} & \dots & u_{p,n} \end{pmatrix}.$$

Using the notation $s = \sum_{1 \le i \le p}(d_i - 1)$, and taking into account the gradation of $\mathbb{K}[U]$, the complex can be rewritten as

$$\begin{aligned}\mathsf{EN}(\alpha): 0 \to & \bigoplus_{\substack{i_1+\ldots+i_p \\ = \\ n-p-1}} R\left[-s - \sum_{1\le j\le p} i_j(d_j-1)\right] \xrightarrow{\delta_{f-g-1}} \bigoplus_{\substack{i_1+\ldots+i_p \\ = \\ n-p-2}} R\left[-s - \sum_{1\le j\le p} i_j(d_j-1)\right]^{\binom{n-1}{n-2}} \xrightarrow{\delta_{f-g}} \\ \to & \bigoplus_{\substack{i_1+\ldots+i_p \\ = 2}} R\left[-s - \sum_{1\le j\le p} i_j(d_j-1)\right]^{\binom{n-1}{p+2}} \xrightarrow{\delta_3} \bigoplus_{1\le i\le p} R\left[-s-(d_i-1)\right]^{\binom{n-1}{p+1}} \xrightarrow{\delta_2} \\ \to & R[-s]^{\binom{n-1}{p}} \xrightarrow{\wedge^g \alpha} R[0].\end{aligned}$$

5.7.2 Hilbert series and degree of regularity in the mixed case

In this section, we use the Eagon-Northcott complex to obtain an explicit formula for the Hilbert series and the degree of regularity of the ideal $\mathbf{I}(\mathbf{F}, 1)$ in the mixed case. Indeed, the Hilbert series of an ideal I can be computed when a free resolution is known, since the Hilbert series of I is equal to the alternate sum of the Hilbert series of the modules occurring in the resolution (see e.g. [Eis01, Theorem 1.11] for more details).

This is a generalization of the results in the unmixed case. However, in the unmixed case, we were able to use the combinatorial properties of the ideal $\mathscr{D}$ in order to obtain a compact formula for the numerator of the rational function $\mathsf{wHS}_{\mathbb{K}[U]/\mathscr{D}}(t)$ (where the determinantal structure appears as the determinant of the matrix $A_r^{p,q}(t)$).

In the mixed case, the analysis is more complicated, but it also leads to an explicit formula for the degree of regularity of the ideal $\mathbf{I}(\mathbf{F}, 1)$.

Proposition 5.19. *The weighted Hilbert series of the ideal $\mathscr{D} \subset \mathbb{K}[U]$ generated by the maximal minors of the matrix $\mathscr{U}$ with $\operatorname{wdeg}(u_{i,j}) = d_i - 1$ is the power series expansion of the rational function*

$$\mathsf{wHS}_{\mathbb{K}[U]/\mathscr{D}}(t) = \frac{1 - \left[\sum_{0\leq k\leq n-p-1}\left[(-1)^k \sum_{i_1+\ldots+i_p=k}\binom{n-1}{p+k} t^{\sum_{1\leq j\leq p}(i_j+1)(d_j-1)}\right]\right]}{\prod_{1\leq i\leq p}(1-t^{d_i-1})^{n-1}}.$$

Proof. According to [Eis01, Theorem 1.11], the Hilbert series of a graded ideal can be computed from a minimal free resolution: it is equal to the alternate sum of the Hilbert series of the free modules occurring in the resolution. For $i, j \in \mathbb{N}$, the Hilbert series of $R[-i]^j$ is equal to

$$\mathsf{wHS}_{R[-i]^j}(t) = \frac{jt^i}{\prod_{1\leq i\leq p}(1-t^{d_i-1})^{n-1}}.$$

Moreover, the Hilbert series of a direct sum of modules is equal to the sum of their Hilbert series. Therefore, by using the Eagon-Northcott complex which is a free resolution of $\mathscr{D}$, direct computations yield the formula for the weighted Hilbert series of $\mathscr{D}$. $\square$

Corollary 5.20. *Let $\mathfrak{F}^h$ is a generic system of homogeneous polynomial equations of degrees $(d_1, \ldots, d_p)$, with $d_i \geq 2$ for all i. The Hilbert series of $\mathbf{I}(\mathfrak{F}^h, 1) \subset \mathbb{K}(\mathfrak{a})[X]$ is*

$$\mathsf{HS}_{\mathbb{K}(\mathfrak{a})[X]/\mathbf{I}(\mathfrak{F}^h,1)}(t) = \mathsf{wHS}_{\mathbb{K}[U]/\mathscr{D}}(t) \cdot \frac{\prod_{1\leq i\leq p}(1-t^{d_i})(1-t^{d_i-1})^{n-1}}{(1-t)^n}.$$

Proof. Let $\widetilde{\mathscr{D}}$ denote the ideal $\mathscr{D} \cdot \mathbb{K}(\mathfrak{a})[X, U] \subset \mathbb{K}(\mathfrak{a})[X, U]$ (where the quasi-homogeneous grading of $\mathbb{K}(\mathfrak{a})[X, U]$ is given by $\operatorname{wdeg}(x_i) = 1$, $\operatorname{wdeg}(u_{i,j}) = d_i - 1$). Therefore, the Hilbert series of $\widetilde{\mathscr{D}}$ is

$$\mathsf{wHS}_{\mathbb{K}(\mathfrak{a})[U,X]/\widetilde{\mathscr{D}}}(t) = \frac{\mathsf{wHS}_{\mathbb{K}(\mathfrak{a})[U]/\mathscr{D}}(t)}{(1-t)^n}.$$

We recall that, with the notations of Lemma 5.10,

$$\mathbf{I}(\mathfrak{F}^h, 1) = \left(\widetilde{\mathscr{D}} + \langle \mathfrak{g}_1, \ldots, \mathfrak{g}_{np}\rangle\right) \cap \mathbb{K}(\mathfrak{a})[X].$$

According to Lemma 5.10, for each $2 \leq \ell \leq np$, $\mathfrak{g}_\ell$ does not divide 0 in $R/\mathfrak{I}_{\ell-1}$. Therefore, a proof similar to that of Proposition 5.7 shows that the Hilbert series of $\mathbf{I}(\mathfrak{F}^h, 1)$ is equal to

$$\begin{aligned}\mathsf{HS}_{\mathbb{K}(\mathfrak{a})[X]/\mathbf{I}(\mathfrak{F}^h,1)}(t) &= \mathsf{wHS}_{\mathbb{K}(\mathfrak{a})[U,X]/\widetilde{\mathscr{D}}}(t)\prod_{1\leq i\leq p}(1-t^{d_i})(1-t^{d_i-1})^{n-1}\\ &= \mathsf{wHS}_{\mathbb{K}(\mathfrak{a})[U]/\mathscr{D}}(t)\cdot\frac{\prod_{1\leq i\leq p}(1-t^{d_i})(1-t^{d_i-1})^{n-1}}{(1-t)^n}.\end{aligned}$$

$\square$

Corollary 5.21. *The degree of regularity of $\mathbf{I}(\mathfrak{F}^h, 1)$ is*

$$d_{\mathrm{reg}}(\mathbf{I}(\mathfrak{F}^h, 1)) = (n-p-1)\max\{d_i - 1\} - n - p + 1 + 2\sum_{1\leq i\leq p} d_i.$$

Proof. According to Lemma 5.4, the ideal $\mathbf{I}(\mathfrak{F}^h, 1)$ is 0-dimensional. Consequently, $\mathsf{HS}_{\mathbb{K}(\mathfrak{a})[X]/\mathbf{I}(\mathfrak{F}^h,1)}$ is a polynomial and $\mathrm{d_{reg}} = \deg(\mathsf{HS}_{\mathbb{K}(\mathfrak{a})[X]/\mathbf{I}(\mathfrak{F}^h,1)}) + 1$. Let j_0 be the index of one of the maximal degrees of the polynomials f_j: $\deg(f_{j_0}) = \max\{\deg(f_j)\}$. In the sums in the numerator of the formula given in Proposition 5.19, the maximal degree is reached when $k = n - p - 1$, $i_{j_0} = k$, and $i_j = 0$ for $j \neq j_0$. Therefore the degree of the numerator of $\mathsf{wHS}_{\mathbb{K}(\mathfrak{a})[U]/\mathscr{D}}(t)$ is equal to

$$\deg\left(1 - \left[\sum_{0\le k\le n-p-1}\left[(-1)^k \sum_{i_1+\ldots+i_p=k}\binom{n-1}{p+k} t^{\sum\limits_{1\le j\le p}(i_j+1)(d_j-1)}\right]\right]\right)$$
$$= (n-p-1)(\max\{d_j\}-1) + \textstyle\sum_{1\le j\le p}(d_j-1).$$

On the other hand, we have

$$\begin{cases}\deg\left(\prod\limits_{1\le i\le p}(1-t^{d_i-1})^{n-1}\right) = (n-1)\sum\limits_{1\le i\le p}(d_i-1);\\ \deg\left(\prod\limits_{1\le i\le p}(1-t^{d_i})(1-t^{d_i-1})^{n-1}\right) = \sum\limits_{1\le i\le p} d_i + (n-1)\sum\limits_{1\le i\le p}(d_i-1);\\ \deg\left((1-t)^n\right) = n.\end{cases}$$

Therefore, using the formula in Corollary 5.20, we obtain

$$\begin{aligned}\deg(\mathsf{HS}_{\mathbb{K}(\mathfrak{a})[X]/\mathbf{I}(\mathfrak{F}^h,1)}) &= (n-p-1)(\max\{d_j\}-1) + \sum_{1\le j\le p}(d_j-1) - (n-1)\sum_{1\le i\le p}(d_i-1) +\\ &\quad \sum_{1\le i\le p} d_i + (n-1)\sum_{1\le i\le p}(d_i-1) - n\\ &= (n-p-1)\max\{d_i-1\} - n - p + 2\sum_{1\le i\le p} d_i,\end{aligned}$$

and hence $\mathrm{d_{reg}} = (n-p-1)\max\{d_i-1\} - n - p + 1 + 2\sum\limits_{1\le i\le p} d_i$. □

5.7.3 Complexity

In this section, we show that under genericity assumptions Gröbner bases of mixed critical point systems can be computed with a complexity which is polynomial in the generic number of critical points in two cases:

1. when p is a constant;
2. when all degrees are bounded above by a constant $D \in \mathbb{N}$.

Theorem 5.22. *Let $p \in \mathbb{N}$ be a fixed integer and $d_i \ge 2$ for all $i \in \{1, \ldots, p\}$, then there exists a nonempty Zariski open subset $\mathscr{O} \subset \overline{\mathbb{K}}[X]_{d_1} \times \cdots \times \overline{\mathbb{K}}[X]_{d_p}$, such that for all $\mathbf{F} \in \mathscr{O} \cap \mathbb{K}[X]^p$, the complexity of computing a lexicographical Gröbner basis of $\mathbf{I}(\mathbf{F}, 1)$ is polynomial in the number of critical points:*

$\forall p \in \mathbb{N}^*$, $\exists b > 0$, $\exists c > 0$, $\forall (d_1, \ldots, d_p) \in \{2, 3, \ldots\}^p$, $\exists$ *a nonempty Zariski open subset* $\mathscr{O} \subset \overline{\mathbb{K}}[X]_{d_1} \times \cdots \times \overline{\mathbb{K}}[X]_{d_p}$, *s.t.*

$$\mathbf{F} \in \mathscr{O} \cap \mathbb{K}[X]^p \Rightarrow (\mathit{Compl} \le b \cdot \#\,\mathsf{crit}^c),$$

where Compl denotes the number of arithmetic operations during the computation of a lex Gröbner basis of $\mathbf{I}(\mathbf{F}, 1)$ with the algorithms F_5 and FGLM, and $\#\,\mathsf{crit} = \left(\prod_{1\le i\le p} d_i\right)\sum_{i_1+\cdots+i_p=n-p}(d_1-1)^{i_1}\dots(d_p-1)^{i_p}$ is the generic number of critical points.

Proof. Since the algorithm FGLM is polynomial in the degree of the ideal, it is sufficient to prove that computing a grevlex Gröbner basis of $\mathbf{I}(\mathbf{F}, 1)$ is also polynomial in $\#\,\mathsf{crit}$. If all polynomials $f_1, \dots, f_p$ are quadratic, then Corollary 5.18 concludes the proof. Therefore, we assume in the sequel that $\max(d_i) \ge 3$. According to [NR09],

$$\#\,\mathsf{crit} = \left(\prod_{1\le i\le p} d_i\right)\sum_{i_1+\cdots+i_p=n-p}(d_1-1)^{i_1}\dots(d_p-1)^{i_p}.$$

Let A (resp. G) be the arithmetic (resp. geometric) average of the set

$$\begin{aligned}
&(d_1-1,\dots,d_p-1,\underbrace{\max\{d_i-1\},\dots,\max\{d_i-1\}}_{n-p}),\\
A &= \frac{1}{n}((n-p)\max\{d_i-1\} + \sum_{1\le i\le p}(d_i-1))\\
G &= (\max\{d_i-1\}^{n-p}\prod_{1\le i\le p}(d_i-1))^{1/n}.
\end{aligned}$$

Consequently, $\#\,\mathsf{crit} > G^n$ and the complexity is bounded above by

$$\begin{aligned}
\binom{n-1}{p}\binom{n+\mathrm{d_{reg}}}{n}^{\omega} &\le \binom{n-1}{p}\binom{2An}{n}^{\omega}\\
&\le n^p\left(\frac{(2An)^n}{n!}\right)^{\omega}\\
&\le O(n^p(2Ae)^{n\omega}).
\end{aligned}$$

Also, notice that $\#\,\mathsf{crit}$ is bounded below by $2^p\binom{n-1}{p}$ and hence $\log(n^p)/\log(\#\,\mathsf{crit}) = O(1)$ since p is constant. The next step is to notice that $A \le \max\{d_i - 1\}$ and $G \ge \max\{d_i-1\}^{(n-p)/n} \ge 2$. Consequently, $\frac{\log A}{\log G}$ is bounded above by $\frac{n}{n-p} \le p+1$, and

$$\begin{aligned}
\frac{\log(\binom{n+\mathrm{d_{reg}}}{n}^{\omega})}{\log(\#\,\mathsf{crit})} &\le \frac{\log(\binom{n+\mathrm{d_{reg}}}{n}^{\omega})}{n\log(G)}\\
&\le \omega\log(2Ae)/\log G\\
&\le \omega\,(\log A/\log G + \log(2e)/\log G)\\
&\le \omega(p+1+\log(2e)/\log(2)),
\end{aligned}$$

and hence $\log(\mathsf{Compl})/\log(\#\,\mathsf{crit})$ is bounded by a constant. □

Theorem 5.23. *Let $D \in \mathbb{N}$ be an integer. Then for all $p \in \mathbb{N}^*$ and for all $(d_1, \dots, d_p) \in \mathbb{N}^p$ with $2 \le d_i \le D$, there exists a nonempty Zariski open subset $\mathscr{O} \subset \overline{\mathbb{K}}[X]_{d_1} \times \cdots \times \overline{\mathbb{K}}[X]_{d_p}$, such that for all $\mathbf{F} \in \mathscr{O} \cap \mathbb{K}[X]^p$, the complexity of computing a lexicographical Gröbner basis of $\mathbf{I}(\mathbf{F}, 1)$ is polynomial in the number of critical points:*

> *$\forall D \in \mathbb{N}^*$, $\exists b > 0$, $\exists c > 0$, $\forall p \in \mathbb{N}^*$, $\forall (d_1, \dots, d_p) \in \{2, 3, \dots, D\}^p$, $\exists$ a nonempty Zariski open subset $\mathscr{O} \subset \overline{\mathbb{K}}[X]_{d_1} \times \cdots \times \overline{\mathbb{K}}[X]_{d_p}$, s.t.*
>
> $$\mathbf{F} \in \mathscr{O} \cap \mathbb{K}[X]^p \Rightarrow (\mathsf{Compl} \le b \cdot \#\,\mathsf{crit}^c),$$

where Compl *denotes the number of arithmetic operations during the computation of a lex Gröbner basis of* $\mathbf{I}(\mathbf{F}, 1)$ *with the algorithms* F_5 *and FGLM, and* $\#\,\mathsf{crit} = \left(\prod_{1\le i\le p} d_i\right) \sum_{i_1+\cdots+i_p=n-p} (d_1 - 1)^{i_1} \dots (d_p - 1)^{i_p}$ *is the generic number of critical points.*

Proof. In this proof, we use the same notations as in the proof of Theorem 5.22. In this case, $A \le D$ and $G \ge 2$. Therefore $\frac{\log A}{\log G}$ is bounded above by $\log_2(D)$. Then a proof similar to that of Theorem 5.22 shows that

$$\frac{\log\left(\binom{n+\mathrm{d_{reg}}}{n}^{\omega}\right)}{\log(\#\,\mathsf{crit})} \le \omega \log_2(2De).$$

As a consequence, $\log(\mathsf{Compl})/\log(\#\,\mathsf{crit})$ is bounded above by a constant. $\square$

Chapter 6

Multi-Homogeneous Systems

The results presented in this chapter are joint work with J.-C. Faugère and M. Safey El Din. Sections 6.1 to 6.5 come from the article [FSS11a]. Compared to the published version, the section 6.5.5 on the complexity of solving affine bilinear systems has been improved. Section 6.6 comes from the preprint [FSS11b] (in submission).

6.1 Introduction

In this chapter, we consider multi-homogeneous systems, which are not regular sequences. Such systems can appear in cryptography [FLP08], in coding theory [OJ02] or in effective geometry [SS03, ST06].

A multi-homogeneous polynomial is defined with respect to a partition of the unknowns, and is homogeneous with respect to each subset of variables. The finite sequence of degrees is called the *multi-degree* of the polynomial. For instance, a bi-homogeneous polynomial f of bidegree (d_1, d_2) in $\mathbb{K}[x_0, \dots, x_{n_x}, y_0, \dots, y_{n_y}]$ ($\mathbb{K}[X, Y]$ for short) is a polynomial such that

$$\forall \lambda, \mu, f(\lambda x_0, \dots, \lambda x_{n_x}, \mu y_0, \dots, \mu y_{n_y}) = \lambda^{d_1} \mu^{d_2} f(x_0, \dots, x_{n_x}, y_0, \dots, y_{n_y}).$$

In general, multi-homogeneous systems are not regular. Consequently, the F_5 criterion does not remove all the reductions to zero. Our goal is to understand the underlying structure of these multi-homogeneous algebraic systems, and then use it to speed up the computation of a Gröbner basis in the context of the F_5 Algorithm. In this chapter, we focus on bi-homogeneous ideals generated by polynomials of bidegree $(1, 1)$.

Main results

Let $\mathbb{K}$ be a field, $f_1, \dots f_m \in \mathbb{K}[x_0, \dots, x_{n_x}, y_0, \dots, y_{n_y}]$ be bilinear polynomials. We denote by $\mathbf{F}_i$ the polynomial family $(f_1, \dots, f_i)$ and by I_i the ideal $\langle \mathbf{F}_i \rangle$. We start by describing the algorithmic results of this chapter, obtained by exploiting the algebraic structure of bilinear systems.

In order to understand this structure, we study properties of the Jacobian matrices with respect to the two subsets of variables $x_0, \dots, x_{n_x}$ and $y_0, \dots, y_{n_y}$:

$$\mathsf{jac}_{\mathbf{x}}(\mathbf{F}_i) = \begin{bmatrix} \frac{\partial f_1}{\partial x_0} & \cdots & \frac{\partial f_1}{\partial x_{n_x}} \\ \vdots & \vdots & \vdots \\ \frac{\partial f_i}{\partial x_0} & \cdots & \frac{\partial f_i}{\partial x_{n_x}} \end{bmatrix} \qquad \mathsf{jac}_{\mathbf{y}}(\mathbf{F}_i) = \begin{bmatrix} \frac{\partial f_1}{\partial y_0} & \cdots & \frac{\partial f_1}{\partial y_{n_y}} \\ \vdots & \vdots & \vdots \\ \frac{\partial f_i}{\partial y_0} & \cdots & \frac{\partial f_i}{\partial y_{n_y}} \end{bmatrix}$$

We show that the kernels of those matrices (whose entries are linear forms) correspond to the reductions to zero not detected by the classical F_5 criterion. In general, all the elements in these kernels are vectors of maximal minors of the Jacobian matrices (Lemma 6.1). For instance, if $n_x = n_y = 2$ and $m = 4$, consider

$$\mathsf{v} = (\mathsf{minor}(\mathsf{jac}_{\mathbf{x}}(\mathbf{F}_4), 1), -\mathsf{minor}(\mathsf{jac}_{\mathbf{x}}(\mathbf{F}_4), 2), \mathsf{minor}(\mathsf{jac}_{\mathbf{x}}(\mathbf{F}_4), 3), -\mathsf{minor}(\mathsf{jac}_{\mathbf{x}}(\mathbf{F}_4), 4))$$

and

$$\mathsf{w} = (\mathsf{minor}(\mathsf{jac}_{\mathbf{y}}(\mathbf{F}_4), 1), -\mathsf{minor}(\mathsf{jac}_{\mathbf{y}}(\mathbf{F}_4), 2), \mathsf{minor}(\mathsf{jac}_{\mathbf{y}}(\mathbf{F}_4), 3), -\mathsf{minor}(\mathsf{jac}_{\mathbf{y}}(\mathbf{F}_4), 4)),$$

where $\mathsf{minor}(\mathsf{jac}_{\mathbf{x}}(\mathbf{F}_4), k)$ (resp. $\mathsf{minor}(\mathsf{jac}_{\mathbf{y}}(\mathbf{F}_4), k)$) denotes the determinant of the matrix obtained from $\mathsf{jac}_{\mathbf{x}}(\mathbf{F}_4)$ (resp. $\mathsf{jac}_{\mathbf{y}}(\mathbf{F}_4)$) by removing the k-th row. The generic *syzygies* corresponding to reductions to zero which are not detected by the classical F_5 criterion are

$$\mathsf{v} \in \mathsf{Ker}_L(\mathsf{jac}_{\mathbf{x}}(\mathbf{F}_4)) \text{ and } \mathsf{w} \in \mathsf{Ker}_L(\mathsf{jac}_{\mathbf{y}}(\mathbf{F}_4)).$$

We show (Corollary 6.17) that, in general, the ideal $I_{i-1} : f_i$ is spanned by I_{i-1} and by the maximal minors of $\mathsf{jac}_{\mathbf{x}}(\mathbf{F}_{i-1})$ (if $i > n_x + 1$) and $\mathsf{jac}_{\mathbf{y}}(\mathbf{F}_{i-1})$ (if $i > n_y + 1$). The leading monomial ideal of $I_{i-1} : f_i$ describes the reductions to zero associated to f_i. Thus we need results about ideals generated by maximal minors of matrices whose entries are linear forms in order to get a description of the syzygy module. In particular, we prove that, in general, *grevlex* Gröbner bases of those ideals are linear combinations of the generators (Theorem 6.5). Based on this result, one can compute efficiently a Gröbner basis of $I_{i-1} : f_i$ once a Gröbner basis of I_{i-1} is known.

This allows us to design an Algorithm (Algorithm 7) dedicated to bilinear systems, which yields an extension of the classical F_5 criterion. This subroutine, when merged within a matrix version of the F_5 Algorithm (Algorithm 5), eliminates all the reductions to zero during the computation of a Gröbner basis of a generic bilinear system. For instance, during the computation of a grevlex Gröbner basis of a system of 12 generic bilinear equations over $\mathbb{K}[x_0, \dots, x_6, y_0, \dots, y_6]$, the new criterion detects 990 reductions to zero which are not found by the usual F_5 criterion. Even if this new criterion seems more complicated than the usual F_5 criterion (some precomputations have to be performed), we prove that the cost induced by those precomputations is negligible compared to the cost of the whole computation.

Next, we introduce a notion of *bi-regularity* which describes the structure of generic bilinear systems. When the input of Algorithm 7 is a bi-regular system, then it returns all the reductions to zero not found by the F_5 criterion. We also give a complete description of the syzygy module of such systems, up to a conjecture (Conjecture 6.7) on a linear algebra problem over rings. This conjecture is supported by practical experiments. We also prove that there are no reduction to zero with the classical F_5 criterion for affine bilinear systems (Proposition 6.26), which is important for practical applications.

We describe now the main complexity results of the chapter. For bi-regular bilinear system, we give an explicit form of these series (Theorem 6.22):

$$\mathsf{mHS}_{\mathbb{K}[x_1,\dots,x_{n_x},y_1,\dots,y_{n_y}]/I}(t_1, t_2) = \frac{(1 - t_1 t_2)^m + N_m(t_1, t_2) + N_m(t_2, t_1)}{(1 - t_1)^{n_x+1}(1 - t_2)^{n_y+1}},$$

$$N_m(t_1, t_2) = \sum_{\ell=1}^{m-(n_y+1)} (1 - t_1 t_2)^{m-(n_y+1)-\ell} t_1 t_2 (1 - t_2)^{n_y+1} \left[1 - (1 - t_1)^\ell \sum_{k=1}^{n_y+1} t_1^{n_y+1-k} \binom{\ell + n_y - k}{n_y + 1 - k} \right]$$

We propose a variant of the Matrix F_5 Algorithm dedicated to multi-homogeneous systems. The key idea is to decompose the Macaulay matrices into a set of smaller matrices whose row echelon forms can be computed independently. We provide some experimental results of an implementation of this algorithm in `Magma2.15`. This multi-homogeneous variant can be more than 20 times faster for bi-homogeneous systems than our `Magma` implementation of the classical Matrix F_5 Algorithm. We perform a theoretical complexity analysis based on the Hilbert series in the case of bilinear systems, which provides an explanation of this gap. Indeed, the coefficients of the Hilbert series provide the sizes of all matrices occurring during the execution of the F_5 algorithm.

We also establish a sharp upper bound on the highest degree during the F_5 algorithm for 0-dimensional affine bilinear systems (Proposition 6.29). Let $\mathbf{F} = (f_1, \dots, f_{n_x+n_y})$ be an affine bilinear system of $\mathbb{K}[x_0, \dots, x_{n_x-1}, y_0, \dots, y_{n_y-1}]$, then the maximal degree reached during the computation of a Gröbner basis with respect to the grevlex ordering is bounded above by:

$$\mathrm{d}_{\max \prec_{\mathsf{grevlex}}}(\mathbf{F}) \leq \min\left(n_x + 2, n_y + 2\right).$$

This bound permits to derive complexity estimates for solving bilinear systems (Corollary 6.30) which can be applied to practical problems (see for instance [FSS10] for an application to the MinRank problem).

Finally, we give an algorithm to compute a rational parametrization of the solutions of an affine system of bidegree $(D, 1)$. Its complexity is strongly related to the complexity of solving an underlying Generalized MinRank Problem.

Related works

The complexity analysis that we perform by proving properties on the Hilbert bi-series of bilinear ideals follows a path which is similar to the one used to analyze the complexity of the F_5 algorithm in the case of homogeneous regular sequences (see [BFSY04]). In [KRHV02], the properties of Buchberger's Algorithm are investigated in the context of multi-graded rings. [CDS07] gives an analysis of the structure of the syzygy module in the case of three bi-homogeneous equations with no common solution in the biprojective space.

The algorithmic use of multi-homogeneous structures has been investigated mostly in the framework of multivariate resultants (see [DE03, EM09] and references therein for the most recent results) following the line of work initiated by [McC33]. In the context of solving polynomial systems by using straight-line programs as data-structures, [JS07] provides an alternative way to compute resultant formula for multi-homogeneous systems.

As we have seen in the description of the main results, the knowledge of Gröbner bases of ideals generated by maximal minors of linear matrices plays a crucial role. Theorem 6.5 which states that such Gröbner bases are obtained by a single row echelon form computation is a variant of the main results in [SZ93] and [BZ93].

More generally, the theory of multi-homogeneous elimination is investigated in [Rém01a, Rém01b] providing tools to generalize some well-known notions (e.g. Chow forms, resultant formula, heights) in the homogeneous case to multi-homogeneous situations. Such works are initiated in [Van29] where the Hilbert bi-series of bi-homogeneous ideals is introduced.

Organization of the chapter

This chapter is articulated as follows. In Section 6.2.2, we investigate the case of bilinear systems and propose an algorithm to remove all the reductions to zero during the Gröbner basis computation. Then we prove its correctness and explain why it is efficient for *generic* bilinear systems in Section 6.3. To continue our study of the structure of bilinear ideals, we give in Section 6.4 the explicit form of the Hilbert bi-series of generic bilinear ideals. We prove in Section 6.5 a new bound on the maximal degree reached during the computation of a grevlex Gröbner basis of generic affine bilinear systems and we use it to derive new complexity bounds. Finally, we generalize some of these results in Section 6.6 to affine systems of bi-degree $(D, 1)$.

6.2 Computing Gröbner bases of bilinear systems

6.2.1 Overview

Let $\mathbf{F} = (f_1, \dots, f_4) \in \mathbb{K}[X]^4$ be four bilinear polynomials in $\mathbb{K}[x_0, x_1, x_2, y_0, y_1, y_2]$, I be the ideal generated by $\mathbf{F}$ and $Z(\mathbf{F}) \subset \mathbb{C}^6$ be its associated algebraic variety. As above, I_i denotes the ideal $\langle f_1, \dots, f_i\rangle$, and we consider the grevlex ordering with $x_0 \succ_{\mathsf{grevlex}} \cdots \succ_{\mathsf{grevlex}} x_{n_x} \succ_{\mathsf{grevlex}} y_0 \succ_{\mathsf{grevlex}} \cdots \succ_{\mathsf{grevlex}} y_{n_y}$. Since $f_1, \dots, f_4$ are bilinear, for all $(a_0, a_1, a_2) \in \mathbb{K}^3$ and $1 \le i \le 4$, $f_i(a_0, a_1, a_2, 0, 0, 0) = 0$. Hence, $Z(\mathbf{F})$ contains the linear affine subspace defined by $y_0 = y_1 = y_2 = 0$ which has dimension 3. We conclude that $Z(\mathbf{F})$ has dimension at least 3.

Consequently, the sequence (f_1, f_2, f_3, f_4) is not regular (since the codimension of an ideal generated by a regular sequence is equal to the length of the sequence). Hence, there are reductions to zero during the computation of a Gröbner basis with the F_5 Algorithm (see [Fau02]).

When the four polynomials are chosen randomly, one remarks experimentally that these reductions correspond to the rows with signatures (x_0^3, f_4) and (y_0^3, f_4). This experimental observation can be explained as follows.

Consider the Jacobian matrices

$$\mathsf{jac}_{\mathbf{x}}(\mathbf{F}) = \begin{bmatrix} \frac{\partial f_1}{\partial x_0} & \frac{\partial f_1}{\partial x_1} & \frac{\partial f_1}{\partial x_2} \\ \vdots & \vdots & \vdots \\ \frac{\partial f_4}{\partial x_0} & \frac{\partial f_4}{\partial x_1} & \frac{\partial f_4}{\partial x_2} \end{bmatrix} \quad \text{and} \quad \mathsf{jac}_{\mathbf{y}}(\mathbf{F}) = \begin{bmatrix} \frac{\partial f_1}{\partial y_0} & \frac{\partial f_1}{\partial y_1} & \frac{\partial f_1}{\partial y_2} \\ \vdots & \vdots & \vdots \\ \frac{\partial f_4}{\partial y_0} & \frac{\partial f_4}{\partial y_1} & \frac{\partial f_4}{\partial y_2} \end{bmatrix}$$

and the vectors of variables $\mathbf{X}$ and $\mathbf{Y}$. By Euler's formula, it is immediate that for any sequence of polynomials (q_1, q_2, q_3, q_4),

$$(q_1, \dots, q_4).\mathsf{jac}_{\mathbf{x}}(\mathbf{F}).\mathbf{X} = \sum_{i=1}^{4} q_i f_i \quad \text{and} \quad (q_1, \dots, q_4).\mathsf{jac}_{\mathbf{y}}(\mathbf{F}).\mathbf{Y} = \sum_{i=1}^{4} q_i f_i \tag{6.1}$$

Let $\mathsf{Ker}_L(\mathsf{jac}_{\mathbf{x}}(\mathbf{F}))$ (resp. $\mathsf{Ker}_L(\mathsf{jac}_{\mathbf{y}}(\mathbf{F}))$) denote the left kernel of $\mathsf{jac}_{\mathbf{x}}(\mathbf{F})$ (resp. $\mathsf{jac}_{\mathbf{y}}(\mathbf{F})$).

Therefore, if $(q_1, \dots, q_4)$ belongs to $\mathsf{Ker}_L(\mathsf{jac}_{\mathbf{x}}(\mathbf{F}))$ (resp. $\mathsf{Ker}_L(\mathsf{jac}_{\mathbf{y}}(\mathbf{F}))$), then the relation (6.1) implies that $(q_1, \dots, q_4)$ belongs to the syzygy module of $\mathbf{F}$.

Given a $(k+1, k)$-matrix M, denote by $\mathsf{minor}(\mathsf{M}, j)$ the minor obtained by removing the j-th row from M. Consider

$$\mathsf{v} = (\mathsf{minor}(\mathsf{jac}_{\mathbf{x}}(\mathbf{F}), 1), -\mathsf{minor}(\mathsf{jac}_{\mathbf{x}}(\mathbf{F}), 2), \mathsf{minor}(\mathsf{jac}_{\mathbf{x}}(\mathbf{F}), 3), -\mathsf{minor}(\mathsf{jac}_{\mathbf{x}}(\mathbf{F}), 4)).$$

By Cramer's rule, v belongs to $\mathsf{Ker}_L(\mathsf{jac}_{\mathbf{x}}(\mathbf{F}))$. A symmetric statement can be made for $\mathsf{jac}_{\mathbf{y}}(\mathbf{F})$. From this observation, one deduces that $\mathsf{minor}(\mathsf{jac}_{\mathbf{x}}(\mathbf{F}), 4) f_4$ (resp. $\mathsf{minor}(\mathsf{jac}_{\mathbf{y}}(\mathbf{F}), 4) f_4$) belongs to $I_3 = \langle f_1, f_2, f_3\rangle$.

We conclude that the rows with signature

$$(\mathsf{LM}(\mathsf{minor}(\mathsf{jac}_{\mathbf{x}}(\mathbf{F}), 4)), f_4) \text{ and } (\mathsf{LM}(\mathsf{minor}(\mathsf{jac}_{\mathbf{y}}(\mathbf{F}), 4)), f_4)$$

are reduced to zero when performing the Matrix F_5 Algorithm described in the previous section. A straightforward computation shows that if $\mathbf{F}$ contains polynomials which are chosen randomly, $\mathsf{LM}(\mathsf{minor}(\mathsf{jac}_{\mathbf{x}}(\mathbf{F}), 4)) = y_0^3$ and $\mathsf{LM}(\mathsf{minor}(\mathsf{jac}_{\mathbf{y}}(\mathbf{F}), 4)) = x_0^3$.

In this section, we generalize this approach to sequences of bilinear polynomials of arbitrary length. Hence, the Jacobian matrices have a number of rows which is not the number of columns incremented by 1. But, even in this more general setting, we exhibit a relationship between the left kernels of the Jacobian matrices and the syzygy module of the sequence $\mathbf{F}$. This allows us to prove a new F_5-criterion dedicated to bilinear systems. On the one hand, when plugged into the Matrix F_5 Algorithm (Algorithm 5), this criterion detects reductions to zero which are not detected by the classical criterion. On the other hand, we prove that a D-Gröbner basis is still computed by the Matrix F_5 Algorithm when it uses the new criterion.

6.2.2 Jacobian matrices of bilinear systems and syzygies

From now on, we use the following notations:

- $R = \mathbb{K}[x_0, \ldots, x_{n_x}, y_0, \ldots, y_{n_y}]$;
- $\mathbf{F} = (f_1, \ldots, f_m) \subset R^m$ is a sequence of bilinear polynomials and $\mathbf{F}_i = (f_1, \ldots, f_i)$ for $1 \leq i \leq m$;
- I is the ideal generated by $\mathbf{F}$ and I_i is the ideal generated by $\mathbf{F}_i$;
- Let M be a $\ell \times c$ matrix, with $\ell > c$. We call *maximal minors* of M the determinants of the $c \times c$ sub-matrices of M;
- $\mathsf{jac}_{\mathbf{x}}(\mathbf{F}_i)$ and $\mathsf{jac}_{\mathbf{y}}(\mathbf{F}_i)$ are respectively the Jacobian matrices

$$\begin{bmatrix} \frac{\partial f_1}{\partial x_0} & \cdots & \frac{\partial f_1}{\partial x_{n_x}} \\ \vdots & \vdots & \vdots \\ \frac{\partial f_i}{\partial x_0} & \cdots & \frac{\partial f_i}{\partial x_{n_x}} \end{bmatrix} \text{ and } \begin{bmatrix} \frac{\partial f_1}{\partial y_0} & \cdots & \frac{\partial f_1}{\partial y_{n_y}} \\ \vdots & \vdots & \vdots \\ \frac{\partial f_i}{\partial y_0} & \cdots & \frac{\partial f_i}{\partial y_{n_y}} \end{bmatrix};$$

- Given a matrix M, we let $\mathsf{Ker}_L(\mathsf{M})$ denote the left kernel of M;
- $\mathbf{X}$ is the vector $[x_0, \ldots, x_{n_x}]^t$ and $\mathbf{Y}$ is the vector $[y_0, \ldots, y_{n_y}]^t$;

Lemma 6.1. *Let $i > n_x + 1$ (resp. $i > n_y + 1$), and let $\mathfrak{s}$ be a maximal minor of* $\mathsf{jac}_{\mathbf{x}}(\mathbf{F}_{i-1})$ *(resp.* $\mathsf{jac}_{\mathbf{y}}(\mathbf{F}_{i-1})$*). Then there exists a vector $(s_1, \ldots, s_{i-1}, \mathfrak{s})$ in* $\mathsf{Ker}_L(\mathsf{jac}_{\mathbf{x}}(\mathbf{F}_i))$ *(resp.* $\mathsf{Ker}_L(\mathsf{jac}_{\mathbf{y}}(\mathbf{F}_i))$*).*

Proof. The proof is done when considering $\mathfrak{s}$ as a maximal minor of $\mathsf{jac}_{\mathbf{x}}(\mathbf{F}_{i-1})$ with $i > n_x + 1$. The case where $\mathfrak{s}$ is a maximal minor of $\mathsf{jac}_{\mathbf{y}}(\mathbf{F}_{i-1})$ with $i > n_y + 1$ is proved similarly.

Notice that $\mathsf{jac}_{\mathbf{x}}(\mathbf{F}_{i-1})$ is a matrix with $i - 1$ rows and $n_x + 1$ columns and $i - 1 \geq n_x + 1$. Denote by $(j_1, \ldots, j_{i-n_x-2})$ the rows deleted from $\mathsf{jac}_{\mathbf{x}}(\mathbf{F}_{i-1})$ to construct its submatrix J whose determinant is $\mathfrak{s}$.

Consider now the $i \times (i - n_x - 2)$-matrix T such that its (ℓ, k) entry is 1 if and only if $\ell = j_k$, else it is 0. N denotes the following $i \times (i - 1)$ matrix:

$$\mathsf{N} = \left[\begin{array}{c|c} \mathsf{jac}_{\mathbf{x}}(\mathbf{F}_i) & \mathsf{T} \end{array} \right].$$

A straightforward use of Cramer's rule shows that

$$(\mathsf{minor}(\mathsf{N}, 1), -\mathsf{minor}(\mathsf{N}, 2), \ldots, (-1)^{i+1}\,\mathsf{minor}(\mathsf{N}, i)) \in \mathsf{Ker}_L(\mathsf{N}).$$

Notice that this implies

$$(\mathsf{minor}(\mathsf{N}, 1), -\mathsf{minor}(\mathsf{N}, 2), \ldots, (-1)^{i+1}\,\mathsf{minor}(\mathsf{N}, i)) \in \mathsf{Ker}_L(\mathsf{jac}_{\mathbf{x}}(\mathbf{F}_i)).$$

Computing $\mathsf{minor}(\mathsf{N}, i)$ by going across the last columns of N shows that $\mathsf{minor}(\mathsf{N}, i) = \pm\mathfrak{s}$. $\square$

Theorem 6.2. *Let $i > n_x + 1$ (resp. $i > n_y + 1$) and let s be a linear combination of maximal minors of $\mathsf{jac}_{\mathbf{x}}(\mathbf{F}_{i-1})$ (resp. $\mathsf{jac}_{\mathbf{y}}(\mathbf{F}_{i-1})$). Then $s \in I_{i-1} : f_i$.*

Proof. By assumption, $s = \sum_\ell a_\ell \, \mathfrak{s}_\ell$ where each $\mathfrak{s}_\ell$ is a maximal minor of $\mathsf{jac}_{\mathbf{x}}(\mathbf{F}_{i-1})$. According to Lemma 6.1, for each minor $\mathfrak{s}_\ell$ there exists $(s_1^{(\ell)}, \ldots, s_{i-1}^{(\ell)})$ such that

$$(s_1^{(\ell)}, \ldots, s_{i-1}^{(\ell)}, \mathfrak{s}_\ell) \in \mathsf{Ker}_L(\mathsf{jac}_{\mathbf{x}}(\mathbf{F}_i))$$

Thus, by summation over ℓ, one obtains

$$\left(\sum_\ell a_\ell s_1^{(\ell)}, \ldots, \sum_\ell a_\ell s_{i-1}^{(\ell)}, s\right) \in \mathsf{Ker}_L(\mathsf{jac}_{\mathbf{x}}(\mathbf{F}_i)). \tag{6.2}$$

Moreover, by Euler's formula

$$\left(\sum_\ell a_\ell s_1^{(\ell)}, \ldots, \sum_\ell a_\ell s_{i-1}^{(\ell)}, s\right) \mathsf{jac}_{\mathbf{x}}(\mathbf{F}_i)\mathbf{X} = s\, f_i + \sum_{j=1}^{i-1}\left(\sum_\ell a_\ell s_j^{(\ell)}\right) f_j.$$

By Relation (6.2), $s\, f_i + \sum_{j=1}^{i-1}\left(\sum_\ell a_\ell s_j^{(\ell)}\right) f_j = 0$, which implies that $s \in I_{i-1} : f_i$. $\square$

Corollary 6.3. *Let $i > n_x + 1$ (resp. $i > n_y + 1$), $M_{\mathbf{x}}^{(i)}$ (resp. $M_{\mathbf{y}}^{(i)}$) be the ideal generated by the maximal minors of $\mathsf{jac}_{\mathbf{x}}(\mathbf{F}_i)$ (resp. $\mathsf{jac}_{\mathbf{y}}(\mathbf{F}_i)$). Then $M_{\mathbf{x}}^{(i-1)} \subset I_{i-1} : f_i$ (resp. $M_{\mathbf{y}}^{(i-1)} \subset I_{i-1} : f_i$).*

Proof. By Theorem 6.2, all minors of $\mathsf{jac}_{\mathbf{x}}(\mathbf{F}_{i-1})$ (resp. $\mathsf{jac}_{\mathbf{y}}(\mathbf{F}_{i-1})$) are elements of $I_{i-1} : f_i$. Thus, $I_{i-1} : f_i$ contains a set of generators of $M_{\mathbf{x}}^{(i-1)}$ (resp. $M_{\mathbf{y}}^{(i-1)}$). $\square$

Example 6.4. *Consider the following bilinear system in $\mathsf{GF}_7[x_0, x_1, x_2, y_0, y_1, y_2, y_3]$:*

$$\begin{aligned}
f_1 &= x_0y_0 + 5x_1y_0 + 4x_2y_0 + 5x_0y_1 + 3x_1y_1 + x_0y_2 + 4x_1y_2 + 5x_2y_2 + 5x_0y_3 + x_1y_3 + 2x_2y_3,\\
f_2 &= 2x_0y_0 + 4x_1y_0 + 6x_2y_0 + 2x_0y_1 + 5x_1y_1 + 6x_0y_2 + 4x_2y_2 + 3x_0y_3 + 2x_1y_3 + 4x_2y_3,\\
f_3 &= 5x_0y_0 + 5x_1y_0 + 2x_2y_0 + 4x_0y_1 + 6x_1y_1 + 4x_2y_1 + 6x_1y_2 + 4x_2y_2 + x_0y_3 + x_1y_3 + 5x_2y_3,\\
f_4 &= 6x_0y_0 + 5x_2y_0 + 4x_0y_1 + 5x_1y_1 + x_2y_1 + x_0y_2 + x_1y_2 + 6x_2y_2 + 2x_0y_3 + 4x_1y_3 + 5x_2y_3,\\
f_5 &= 6x_0y_0 + 3x_1y_0 + 6x_2y_0 + 3x_0y_1 + 5x_2y_1 + 2x_0y_2 + 4x_1y_2 + 5x_2y_2 + 2x_0y_3 + 4x_1y_3 + 5x_2y_3.
\end{aligned}$$

Its Jacobian matrices $\mathsf{jac}_{\mathbf{x}}(\mathbf{F}_4)$ and $\mathsf{jac}_{\mathbf{y}}(\mathbf{F}_4)$ are:

$$\mathsf{jac}_{\mathbf{x}}(\mathbf{F}_4) = \begin{pmatrix} y_0 + 5y_1 + y_2 + 5y_3 & 5y_0 + 3y_1 + 4y_2 + y_3 & 4y_0 + 5y_2 + 2y_3 \\ 2y_0 + 2y_1 + 6y_2 + 3y_3 & 4y_0 + 5y_1 + 2y_3 & 6y_0 + 4y_2 + 4y_3 \\ 5y_0 + 4y_1 + y_3 & 5y_0 + 6y_1 + 6y_2 + y_3 & 2y_0 + 4y_1 + 4y_2 + 5y_3 \\ 6y_0 + 4y_1 + y_2 + 2y_3 & 5y_1 + y_2 + 4y_3 & 5y_0 + y_1 + 6y_2 + 5y_3 \end{pmatrix}.$$

$$\mathsf{jac}_{\mathbf{y}}(\mathbf{F}_4) = \begin{pmatrix} x_0 + 5x_1 + 4x_2 & 5x_0 + 3x_1 & x_0 + 4x_1 + 5x_2 & 5x_0 + x_1 + 2x_2 \\ 2x_0 + 4x_1 + 6x_2 & 2x_0 + 5x_1 & 6x_0 + 4x_2 & 3x_0 + 2x_1 + 4x_2 \\ 5x_0 + 5x_1 + 2x_2 & 4x_0 + 6x_1 + 4x_2 & 6x_1 + 4x_2 & x_0 + x_1 + 5x_2 \\ 6x_0 + 5x_2 & 4x_0 + 5x_1 + x_2 & x_0 + x_1 + 6x_2 & 2x_0 + 4x_1 + 5x_2 \end{pmatrix}.$$

An straightforward computation shows that the maximal minors of the matrix $\mathsf{jac}_{\mathbf{x}}(\mathbf{F}_4)$ and $\mathsf{jac}_{\mathbf{y}}(\mathbf{F}_4)$ are in $\langle f_1, f_2, f_3, f_4\rangle : f_5$, in accordance with Corollary 6.3. An example of a corresponding syzygy is obtained by the vanishing of the determinant

$$\begin{aligned}\det\left[\mathsf{jac}_{\mathbf{x}}(\mathbf{F}_5)|T|\mathbf{F}_5\right] = \det &\begin{pmatrix} y_0 + 5y_1 + y_2 + 5y_3 & 5y_0 + 3y_1 + 4y_2 + y_3 & 4y_0 + 5y_2 + 2y_3 & 1 & f_1 \\ 2y_0 + 2y_1 + 6y_2 + 3y_3 & 4y_0 + 5y_1 + 2y_3 & 6y_0 + 4y_2 + 4y_3 & 0 & f_2 \\ 5y_0 + 4y_1 + y_3 & 5y_0 + 6y_1 + 6y_2 + y_3 & 2y_0 + 4y_1 + 4y_2 + 5y_3 & 0 & f_3 \\ 6y_0 + 4y_1 + y_2 + 2y_3 & 5y_1 + y_2 + 4y_3 & 5y_0 + y_1 + 6y_2 + 5y_3 & 0 & f_4 \\ 6y_0 + 3y_1 + 2y_2 + 2y_3 & 3y_0 + 4y_2 + 4y_3 & 6y_0 + 5y_1 + 5y_2 + 5y_3 & 0 & f_5 \end{pmatrix} \\ = 0.&\end{aligned}$$

The above results imply that for all $g \in M_{\mathbf{x}}^{(i-1)}$ (resp. $g \in M_{\mathbf{y}}^{(i-1)}$), the rows of signature $(\mathsf{LM}(g), f_i)$ are reduced to zero during the Matrix F_5 Algorithm. In order to remove these rows, it is crucial to compute a Gröbner basis of the ideals $M_{\mathbf{x}}^{(i-1)}$ and $M_{\mathbf{y}}^{(i-1)}$. These ideals are generated by the maximal minors of matrices whose entries are linear forms. The goal of the following section is to understand the structure of such ideals and how Gröbner bases can be efficiently computed in that case.

6.2.3 Gröbner bases and maximal minors of matrices with linear entries

Let $\mathscr{L}$ be the set of homogeneous linear forms in the ring $R_{\mathbf{X}} = \mathbb{K}[x_0, \ldots, x_{n_x}]$, $\prec$ be the $grevlex$ ordering on $R_{\mathbf{X}}$ (with $x_0 \succ \cdots \succ x_{n_x}$) and $\mathsf{Mat}_{\mathscr{L}}(p,q)$ be the set of $p \times q$ matrices with entries in $\mathscr{L}$ with $p \geq q$ and $n_x \geq p-q$. Note that $\mathsf{Mat}_{\mathscr{L}}(p,q)$ is a $\mathbb{K}$-vector space of finite dimension.

Given $\mathsf{M} \in \mathsf{Mat}_{\mathscr{L}}(p,q)$, we denote by $\mathsf{MaxMinors}(\mathsf{M})$ the set of maximal minors of M. We recall that $\mathsf{Mac}_{\prec,q}(\mathsf{MaxMinors}(\mathsf{M}))$ denote the Macaulay matrix in degree q associated to $\mathsf{MaxMinors}(\mathsf{M})$ and to the ordering $\prec$ (each row represents a polynomial of $\mathsf{MaxMinors}(\mathsf{M})$ and the columns represent the monomials of degree q in $\mathbb{K}[x_0, \ldots, x_{n_x}]$ sorted by $\prec$, see Definition 1.58).

The main result of this paragraph lies in the following theorem: it states that, in general, a Gröbner basis of $\langle \mathsf{MaxMinors}(\mathsf{M}) \rangle$ is a *linear* combination of the generators.

Theorem 6.5. *There exists a nonempty Zariski-open set O in $\mathsf{Mat}_{\mathscr{L}}(p,q)$ such that for all $\mathsf{M} \in O$, a $grevlex$ Gröbner basis of $\langle \mathsf{MaxMinors}(\mathsf{M}) \rangle$ with respect to $\prec$ is obtained by computing the row echelon form of $\mathsf{Mac}_{\prec,q}(\mathsf{MaxMinors}(\mathsf{M}))$.*

This theorem is related with a result from Sturmfels, Bernstein and Zelevinsky [BZ93, SZ93], which states that the ideal generated by the maximal minors of a matrix whose entries are variables is a universal Gröbner Basis. We tried without success to use this result in order to prove Theorem 6.5.

In [FSS11a], we gave an ad-hoc proof of Theorem 6.5. In this thesis, we provide a short proof based on the results on determinantal ideals (Chapter 4).

Proof of Theorem 6.5. By Lemma 4.18 (with $D = 1$ and $r = q-1$)), there exists a nonempty Zariski-open set O in $\mathsf{Mat}_{\mathscr{L}}(p,q)$ such that for all $\mathsf{M} \in O$, the maximal degree in a reduced grevlex Gröbner basis of M is q. Since the maximal minors of M have degree q, a *grevlex* Gröbner basis of $\langle \mathsf{MaxMinors}(\mathsf{M}) \rangle$ with respect to $\prec$ is obtained by computing the row echelon form of $\mathsf{Mac}_{\prec,q}(\mathsf{MaxMinors}(\mathsf{M}))$. $\square$

In [FSS11a], we gave an explicit example of linear matrix in order to prove that the Zariski open set in Theorem 6.5 is nonempty. Although this explicit example is not necessary here (the proof above implies that the Zariski open set is nonempty), we still report it for its combinatorial properties:

Proof that O in Theorem 6.5 is nonempty. In order to prove that the Zariski open set O is nonempty, we exhibit an explicit element. Consider the matrix M of $\mathsf{Mat}_{\mathscr{L}}(p,q)$ whose (i,j)-entry is x_{i+j-2} if $0 \leq i+j-2 \leq p-q$ and $i \geq j$, else it is 0.

$$\mathsf{M} = \begin{pmatrix} x_0 & 0 & \ldots & 0 \\ x_1 & x_0 & \ddots & 0 \\ \vdots & x_1 & \ddots & \vdots \\ x_{p-q} & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & x_{p-q-1} \\ 0 & 0 & \ldots & x_{p-q} \end{pmatrix}.$$

Notice that $\mathsf{MaxMinors}(\mathsf{M}) \subset \mathbb{K}[x_0, \ldots, x_{p-q}]$. We prove in the sequel that the leading monomials of the maximal minors of M are exactly $\mathsf{Monomials}_{p-q}(q)$

A first observation is that the cardinality of $\mathsf{MaxMinors}(\mathsf{M})$ equals the cardinality of $\mathsf{Monomials}_{p-q}(q)$. Let m be a maximal minor of M. Thus m is the determinant of a $q \times q$ submatrix M' obtained by removing $p - q$ rows from M. Let $i_1, \ldots, i_{p-q}$ be the indices of these rows (with $i_1 < \ldots < i_{p-q}$). Denote by $\star$ the product coefficient by coefficient of two matrices (i.e. the *Hadamard product*) and let $\mathfrak{S}_q$ be the set of $q \times q$ permutation matrices. Thus $m = \sum_{\sigma \in \mathfrak{S}_q} (-1)^{\mathsf{sgn}(\sigma)} \det(\sigma \star \mathsf{M}')$.

Since for all $\sigma \in \mathfrak{S}_q$, $\det(\sigma \star \mathsf{M}')$ is a monomial, there exists $\sigma^0 \in \mathfrak{S}_q$ such that $\mathsf{LM}(m) = \pm \det(\sigma^0 \star \mathsf{M}')$

We prove now that $\sigma^0 = \mathsf{id}$. Suppose by contradiction that $\sigma^0 \neq \mathsf{id}$. In the sequel, we denote by

- $\mathsf{M}'[i, j]$ the (i, j)-entry of M'.
- $\mathbf{e}_i$ the $q \times 1$ unit vector whose i-th coordinate is 1 and all its other coordinates are 0;
- σ^0_j is the integer i such that $\sigma^0 \mathbf{e}_j = \mathbf{e}_i$.

Since, by assumption, $\sigma^0 \neq \mathsf{id}$, there exists $1 \leq i < j \leq q$ such that $\sigma^0_j > \sigma^0_i$. Because of the structure of M, we know that for the *grevlex* ordering $x_0 \succ \cdots \succ x_{n_x}$,

$$\mathsf{M}'[i, \sigma^0_j]\mathsf{M}'[j, \sigma^0_i] \succ \mathsf{M}'[i, \sigma^0_i]\mathsf{M}'[j, \sigma^0_j].$$

Let σ' be defined by

$$\sigma'_k = \begin{cases} \sigma^0_k \text{ if } k \neq i \text{ and } k \neq j \\ \sigma^0_j \text{ if } k = i \\ \sigma^0_i \text{ if } k = j \end{cases}$$

Then $\det(\sigma' \star \mathsf{M}') \succ \det(\sigma^0 \star \mathsf{M}')$ and by induction $\det(\mathsf{id} \star \mathsf{M}') \succ \det(\sigma^0 \star \mathsf{M}')$. This also proves that the coefficient of $\det(\mathsf{id} \star \mathsf{M}')$ in $\mathsf{MaxMinors}(\mathsf{M})$ is 1 and contradicts the fact that $\mathsf{LM}(m) = \pm \det(\sigma^0 \star \mathsf{M}')$.

This proves that $\mathsf{LM}(m) = |\det(\mathsf{id} \star \mathsf{M}')|$. Consequently,

$$\det(\mathsf{id} \star \mathsf{M}') = x_0^{i_1-1} x_1^{i_2-i_1-1} x_2^{i_3-i_2-1} \ldots x_{p-q}^{p-i_{p-q}-1}.$$

Thus if m_1, m_2 are distinct elements in $\mathsf{MaxMinors}(\mathsf{M})$, then $\mathsf{LM}(m_1) \neq \mathsf{LM}(m_2)$. Since for all m in $\mathsf{MaxMinors}(\mathsf{M})$, $\mathsf{LM}(m) \in \mathsf{Monomials}_{p-q}(q)$, and $\mathsf{MaxMinors}(\mathsf{M})$ has the same cardinality as $\mathsf{Monomials}_{p-q}(q)$, we can deduce that $\mathsf{LM}(\mathsf{MaxMinors}(\mathsf{M})) = \mathsf{Monomials}_{p-q}(q)$. □

6.2.4 An extension of the F_5 criterion for bilinear systems

We can now present the main algorithm of this section. Given a sequence of homogeneous bilinear forms $\mathbf{F} = (f_1, \ldots, f_m) \in R^m$ generating an ideal $I \subset R$ and $\prec$ a monomial ordering, it returns a set of pairs (g, f_i) such that $g \in I_{i-1} : f_i$ and $g \notin I_{i-1}$ (for $i > \min(n_x + 1, n_y + 1)$). Following Theorem 6.2 and 6.5, this is done by considering the matrices $\mathsf{jac}_{\mathbf{x}}(\mathbf{F}_i)$ (resp. $\mathsf{jac}_{\mathbf{y}}(\mathbf{F}_i)$) for $i > n_x + 1$ (resp. $i > n_y + 1$) and performing a row echelon form on $\mathsf{Mac}_{\prec, n_x+1}(\mathsf{MaxMinors}(\mathsf{jac}_{\mathbf{x}}(\mathbf{F}_i)))$ (resp. $\mathsf{Mac}_{\prec, n_y+1}(\mathsf{MaxMinors}(\mathsf{jac}_{\mathbf{y}}(\mathbf{F}_i)))$).

First we describe the subroutine **Reduce** (Algorithm 6) which reduces a set of homogeneous polynomials of the same degree:

The main algorithm uses this subroutine in order to compute a row echelon form of $\mathsf{Mac}_{\prec, n_x+1}(\mathsf{MaxMinors}(\mathsf{jac}_{\mathbf{x}}(\mathbf{F}_i)))$ (resp. $\mathsf{Mac}_{\prec, n_y+1}(\mathsf{MaxMinors}(\mathsf{jac}_{\mathbf{y}}(\mathbf{F}_i)))$):

The following proposition explains how the output of Algorithm 7 is related to reductions to zero occurring during the Matrix F_5 Algorithm.

Algorithm 6 Reduce

Input: $\prec$ a monomial ordering and (S, q) where S is a set of homogeneous polynomials of degree q.
Output: T is a reduced set of homogeneous polynomials of degree q.

$\mathsf{M} \leftarrow \mathsf{Mac}_{\prec,q}(S)$.
$\mathsf{M} \leftarrow \mathsf{RowEchelonForm}(\mathsf{M})$.
Return T the set of polynomials corresponding to the rows of M.

Algorithm 7 BLcriterion

Input: $\begin{cases} m \text{ bilinear polynomials } f_1, \dots, f_m \text{ such that } m \leq n_x + n_y. \\ \prec \text{ a monomial ordering over } \mathbb{K}[x_0, \dots, x_{n_x}, y_0, \dots, y_{n_y}] \end{cases}$
Output: V a set of pairs (h, f_i) such that $h \in I_{i-1} : f_i$ and $h \notin I_{i-1}$.

$V \leftarrow \emptyset$
for i from 2 to m **do**
    if $i > n_y + 1$ **then**
        $T \leftarrow \mathbf{Reduce}(\mathsf{MaxMinors}(\mathsf{jac}_{\mathbf{y}}(\mathbf{F}_{i-1})), n_y + 1)$.
        for h in T **do**
            $V \leftarrow V \cup \{(h, f_i)\}$
        end for
    end if
    if $i > n_x + 1$ **then**
        $T' \leftarrow \mathbf{Reduce}(\mathsf{MaxMinors}(\mathsf{jac}_{\mathbf{x}}(\mathbf{F}_{i-1})), n_x + 1)$.
        for h in T' **do**
            $V \leftarrow V \cup \{(h, f_i)\}$
        end for
    end if
end for
Return V

Proposition 6.6 (Extended F_5 criterion for bilinear systems)**.** *Let $f_1, \ldots, f_m$ be bilinear polynomials and $\prec$ be a monomial ordering. Let (t, f_i) be the signature of a row during the Matrix F_5 Algorithm and let V be the output of Algorithm* BLCRITERION. *Then if there exists (h, f_i) in V such that $\mathsf{LM}(h) = t$, then the row with signature (t, f_i) will be reduced to zero.*

Proof. According to Theorem 6.2, $hf_i \in I_{i-1}$. Therefore

$$tf_i = (h-t)f_i + \sum_{j=1}^{i-1} g_j f_j.$$

This implies that the row with signature (t, f_i) is a linear combination of preceding rows in $\mathsf{Mac}_{\prec,\deg(tf_i)}(\mathbf{F}_i)$. Hence this row will be reduced to zero. □

Now we can merge this extended criterion with the Matrix F_5 Algorithm. To do so, we denote by V the output of BLCRITERION (V has to be computed at the beginning of Matrix F_5 Algorithm), and we replace in Algorithm 5 the F_5CRITERION by the following BILINF_5CRITERION:

BILINF_5CRITERION - returns a boolean

Input: $\begin{cases} (t, f_i) \text{ the signature of a row} \\ \text{A matrix } \mathcal{M} \text{ in row echelon form} \end{cases}$

1: Return `true` if $\begin{cases} t \text{ is the leading monomial of a row of } \mathcal{M} \textbf{ or} \\ \exists (h, f_i) \in V \text{ such that } \mathsf{LM}(h) = t \end{cases}$

6.3 F_5 without reduction to zero for generic bilinear systems

6.3.1 Main results

The goal of this part of the chapter is to show that Algorithm 7 finds all reductions to zero for generic bilinear systems. In order to describe the structure of ideals generated by generic bilinear systems, we define a notion of *bi-regularity* (Definition 6.9). For bi-regular systems, we give a complete description of the syzygy module (Proposition 6.15 and Corollary 6.17). Finally, we show that, for such systems, Algorithm 7 finds all reductions to zero and that generic bilinear systems are bi-regular (Theorem 6.18), assuming a conjecture about the kernel of generic matrices whose entries are linear forms (Conjecture 6.7).

6.3.2 Kernel of matrices whose entries are linear forms

Consider a monomial ordering $\prec$ such that its restriction to $\mathbb{K}[x_0, \ldots, x_{n_x}]$ (resp. $\mathbb{K}[y_0, \ldots, y_{n_y}]$) is the $grevlex$ ordering (for instance the usual *grevlex* ordering with $x_0 \succ_{\mathsf{grevlex}} x_1 \succ_{\mathsf{grevlex}} \cdots \succ_{\mathsf{grevlex}} y_0 \succ_{\mathsf{grevlex}} \cdots \succ_{\mathsf{grevlex}} y_{n_y}$).

Let ℓ, c, n_x be integers such that $c < \ell \leq n_x + c - 1$. Let $\mathcal{M}$ be the set of matrices $\ell \times c$ whose coefficients are linear forms in $\mathbb{K}[x_0, \ldots, x_{n_x}]$. Let $\mathcal{T}$ be the set of $\ell \times (\ell - c - 1)$ matrices T such that:

- each column of T has exactly one 1 and the rest of the coefficients are 0;
- each row of T has at most one 1 and all the other coefficients are 0;

- $(\mathsf{T}[i_1, j_1] = \mathsf{T}[i_2, j_2] = 1$ and $i_1 < i_2) \Rightarrow j_1 < j_2$.

If $\mathsf{T} \in \mathcal{T}$ and $\mathsf{M} \in \mathcal{M}$, we denote by M_T the $\ell \times (\ell - 1)$ matrix obtained by adding to M the columns of T. According to the proof of Lemma 6.1, some elements of the left kernel of a matrix M can be expressed as vectors of maximal minors:

$$\forall \mathsf{T} \in \mathcal{T}, \begin{pmatrix} \mathsf{minor}(\mathsf{M}_\mathsf{T}, 1) \\ -\mathsf{minor}(\mathsf{M}_\mathsf{T}, 2) \\ \vdots \\ (-1)^{m+1}\mathsf{minor}(\mathsf{M}_\mathsf{T}, m) \end{pmatrix} \in \mathsf{Ker}_L(\mathsf{M}).$$

Actually, we observed experimentally that kernels of random matrices $\mathsf{M} \in \mathcal{M}$ are generated by those vectors of minors. This leads to the formulation of the following conjecture:

Conjecture 6.7. *The set of matrices* $\mathsf{M} \in \mathcal{M}$ *such that*

$$\mathsf{Ker}_L(\mathsf{M}) = \left\langle \left\{ \begin{pmatrix} \mathsf{minor}(\mathsf{M}_\mathsf{T}, 1) \\ -\mathsf{minor}(\mathsf{M}_\mathsf{T}, 2) \\ \vdots \\ (-1)^{m+1}\mathsf{minor}(\mathsf{M}_\mathsf{T}, m) \end{pmatrix} \right\}_{\mathsf{T} \in \mathcal{T}} \right\rangle$$

contains a nonempty Zariski open subset of $\mathcal{M}$.

This conjecture is proved when the matrix M contains independent variables (see e.g. [Onn94]). In future works, we intend to study how the results in [Onn94] can be applied when the matrix M contains generic polynomials; this could lead to a proof of Conjecture 6.7.

6.3.3 Structure of generic bilinear systems

With the following definition, we give an analog of regular sequences for bilinear systems. This definition is closely related to the generic behavior of Algorithm 7.

Remark 6.8. *In the following,* $\mathsf{Monomials}^{\mathbf{x}}_n(d)$ *(resp.* $\mathsf{Monomials}^{\mathbf{y}}_n(d)$*) denotes the set of monomials of degree d in* $\mathbb{K}[x_0, \dots, x_n]$ *(resp.* $\mathbb{K}[y_0, \dots, y_n]$*). If $n < 0$, we use the convention* $\mathsf{Monomials}^{\mathbf{x}}_n(d) = \mathsf{Monomials}^{\mathbf{y}}_n(d) = \emptyset$.

Definition 6.9. *Let $\prec$ be a monomial ordering such that its restriction to* $\mathbb{K}[x_0, \dots, x_{n_x}]$ *(resp.* $\mathbb{K}[y_0, \dots, y_{n_y}]$*) is the grevlex ordering. Let $m \leq n_x + n_y$ and $f_1, \dots, f_m$ be bilinear polynomials of R. We say that the polynomial sequence $(f_1, \dots, f_m)$ is a* bi-regular sequence *if $m = 1$ or if $(f_1, \dots, f_{m-1})$ is a bi-regular sequence and*

$$\begin{aligned} \mathsf{LM}(I_{m-1} : f_m) &= \langle \mathsf{Monomials}^{\mathbf{x}}_{m-n_y-2}(n_y + 1) \rangle \\ &\quad + \langle \mathsf{Monomials}^{\mathbf{y}}_{m-n_x-2}(n_x + 1) \rangle \\ &\quad + \mathsf{LM}(I_{m-1}) \end{aligned}$$

In the following, we use the notations:

- $\mathscr{BL}_{\mathbb{K}}(n_x, n_y)$ the $\mathbb{K}$-vector space of bilinear polynomials in $\mathbb{K}[x_0, \dots, x_{n_x}, y_0, \dots, y_{n_y}]$;
- X (resp. Y) is the ideal $\langle x_0, \dots, x_{n_x} \rangle$ (resp. $\langle y_0, \dots, y_{n_y} \rangle$);
- An ideal is called *bihomogeneous* if it admits a set of bihomogeneous generators;

- J_i denotes the saturated ideal $I_i : (X \cap Y)^\infty$;
- Given a polynomial sequence $\mathbf{F} = (f_1, \ldots, f_m)$, we denote by $\mathsf{Syz}_{\mathsf{triv}}(\mathbf{F})$ the module of trivial syzygies, i.e. the set of all syzygies $(s_1, \ldots, s_m)$ such that
$$\forall i, s_i \in \langle f_1, \ldots, f_{i-1}, f_{i+1}, \ldots, f_m \rangle;$$
- A primary ideal $P \subset R$ is called *admissible* if $X \not\subset \sqrt{P}$ and $Y \not\subset \sqrt{P}$;
- Let E be a $\mathbb{K}$-vector space such that $\dim(E) < \infty$. We say that a property $\mathcal{P}$ is *generic* if it is satisfied on a nonempty open subset of E (for the Zariski topology), i.e. $\exists h \in \mathbb{K}[\mathfrak{a}_1, \ldots, \mathfrak{a}_{\dim(E)}], h \neq 0$, such that
$$\mathcal{P} \text{ does not hold on } (a_1, \ldots, a_{\dim(E)}) \Rightarrow h(a_1, \ldots, a_{\dim(E)}) = 0.$$

Without loss of generality, we suppose in the sequel that $n_x \leq n_y$.

Lemma 6.10. *Let I_m be an ideal spanned by m generic bilinear equations $f_1, \ldots, f_m$ and $I_m = \cap_{P \in \mathcal{P}} P$ be a minimal primary decomposition.*

- *If $m < n_x + 1$, then all components of I_m are admissible.*
- *If $n_x + 1 \leq m < n_y + 1$ and $P_0 \in \mathcal{P}$ is a primary non-admissible component, then $Y \not\subset \sqrt{P_0}$.*

Proof. We prove that if $m < n_x + 1$ (resp. $m < n_y + 1$) and P_0 is a primary non-admissible component, then $X \not\subset \sqrt{P_0}$ (resp. $Y \not\subset \sqrt{P_0}$). Lemma 6.10 is a consequence of this fact.

Consider the field $\mathbb{K}' = \mathbb{K}(y_0, \ldots, y_{n_y})$ and the canonical inclusion
$$\psi : R \to \mathbb{K}'[x_0, \ldots, x_{n_x}].$$
$\psi(I_m)$ is an ideal of $\mathbb{K}'[x_0, \ldots, x_{n_x}]$ spanned by m polynomials in $\mathbb{K}'[x_0, \ldots, x_{n_x}]$. Thus there exists an polynomial $f \in X$ (homogeneous in the x_is) such that $\psi(f)$ is not a divisor of 0 in $\mathbb{K}'[x_0, \ldots, x_{n_x}]/\psi(I_m)$. This means that $\psi(I_m) : \psi(f) = \psi(I_m)$. Suppose the assertion of Lemma 6.10 is false. Then $X \subset \sqrt{P_0}$ and hence, $f \in \sqrt{P_0}$. Therefore there exists $g \in \mathbb{K}[y_0, \ldots, y_{n_y}]$ such that, in R, $gf \in \sqrt{I_m}$ (take g in $(\cap_{P \in \mathcal{P} \setminus \{P_0\}} \sqrt{P}) \setminus \{\sqrt{P_0}\}$ which is nonempty). Thus $\psi(f) \in \sqrt{\psi(I_m)}$ (since $\psi(g)$ is invertible in $\mathbb{K}'$), which is impossible since $\psi(I_m) : \psi(f) = \psi(I_m)$. $\square$

Lemma 6.11.
- *If $m \leq n_x$ there exists a nonempty Zariski-open set $\mathcal{O} \subset \mathscr{BL}_{\overline{\mathbb{K}}}(n_x, n_y)^m$ such that $(f_1, \ldots, f_m) \in \mathcal{O} \cap \mathscr{BL}_{\mathbb{K}}(n_x, n_y)^m$ implies that I_m has codimension m and all the components of a minimal primary decomposition of I_m are admissible;*
- *if $n_x + 1 \leq m$, then there exists a nonempty Zariski-open set $\mathcal{O} \subset \mathscr{BL}_{\overline{\mathbb{K}}}(n_x, n_y)^m$ such that $(f_1, \ldots, f_m) \in \mathcal{O} \cap \mathscr{BL}_{\mathbb{K}}(n_x, n_y)^m$ implies that X is a prime associated to $\sqrt{I_m}$;*
- *if $n_y + 1 \leq m$, then there exists a nonempty Zariski-open set $\mathcal{O} \subset \mathscr{BL}_{\overline{\mathbb{K}}}(n_x, n_y)^m$ such that $(f_1, \ldots, f_m) \in \mathcal{O} \cap \mathscr{BL}_{\mathbb{K}}(n_x, n_y)^m$ implies that Y is a prime associated to $\sqrt{I_m}$.*

Proof.
- If $m \leq n_x$, then by Lemma 6.10, $J_m = I_m$. Then according to Theorem 3.12, there exists a nonempty Zariski-open set $\mathcal{O} \subset \mathscr{BL}_{\overline{\mathbb{K}}}(n_x, n_y)^m$ such that $(f_1, \ldots, f_m) \in \mathcal{O}$ implies that $(f_1, \ldots, f_m)$ is a regular sequence. Therefore, I_m has codimension m and all the components of a minimal primary decomposition of I_m are admissible.

- If $n_x + 1 \leq m$, then according to Proposition 3.10, $J_m = (I_m : Y^\infty) : X^\infty$ is equidimensional of codimension m. Let V_x be the set $\{(0, \ldots, 0, a_0, \ldots, a_{n_y}) | a_i \in \mathbb{K}\}$. Since $V_x \subset Z(I_m : Y^\infty)$ and $\mathsf{codim}(V_x) = n_x + 1$, it can be deduced that $V_x \not\subset Z(J_m)$ and $Z(I_m : Y^\infty) = Z(J_m) \cup V_x$. This means that $\sqrt{I_m : Y^\infty} = \sqrt{J_m} \cap X$ and $\sqrt{J_m} \not\subset X$. Thus X is a prime associated to $\sqrt{I_m : Y^\infty}$. Since Y is not a subset of X, X is also a prime ideal associated to $\sqrt{I_m}$.

- Similar proof in the case $n_y + 1 \leq m$.

□

Lemma 6.12. *Suppose that the local ring R_X/I_X (resp. R_Y/I_Y) is regular and that X (resp. Y) is a prime ideal associated to $\sqrt{I}$ and let Q be an isolated primary component of a minimal primary decomposition of I containing X (resp. Y). Then $Q = X$ (resp. $Q = Y$).*

Proof. By assumption, X is a prime ideal associated to $\sqrt{I}$. Then, there exists an isolated primary component of a minimal primary decomposition of I which contains a power of X and does not meet $R \setminus X$. This proves that I_X does not contain a unit in R_X.

By assumption R_X/I_X is regular and local, then R_X/I_X is an integral ring (see e.g. [Eis95, Corollary 10.14]) which implies that I_X is prime and does not contain a unit in R_X.

Let $I = Q_1 \cap \cdots \cap Q_s$ be a minimal primary decomposition of I. In the sequel, Q_{i_X} denotes the localization of Q_i by X. Suppose first that there exists $1 \leq i \leq s$ such that $I_X = Q_{i_X}$ with Q_i non-admissible which does not meet the multiplicatively closed part $R \setminus X$. Then Q_{i_X} is obviously prime which implies that Q_i itself is prime [AM69, Proposition 3.11 (iv)]. Our claim follows.

It remains to prove that $I_X = Q_{i_X}$ for some $1 \leq i \leq s$. Suppose that the Q_i's are numbered such that Q_j meets the multiplicatively closed set $R \setminus X$ for $r + 1 \leq j \leq s$ but not $Q_1, \ldots, Q_r$. $I_X = Q_{1_X} \cap \cdots \cap Q_{r_X}$ and it is a minimal primary decomposition [AM69, Proposition 4.9]. Hence, since I_X is prime, $r = 1$ and Q_1 is the isolated minimal primary component containing X. □

Proposition 6.13. *Let $\mathbb{K}$ be a field of characteristic 0. There exists a nonempty Zariski-open set $\mathcal{O} \subset \mathscr{BL}_{\overline{\mathbb{K}}}(n_x, n_y)^m$ such that for all $(f_1, \ldots, f_m) \in \mathcal{O} \cap \mathscr{BL}_{\mathbb{K}}(n_x, n_y)^m$ the non-admissible components of a minimal primary decomposition of $\langle f_1, \ldots, f_m \rangle$ are either X or Y.*

Proof. Suppose that $n_x + 1 \leq m$. Then, by Lemma 6.11, there exists a nonempty Zariski-open set O_1 such that X is an associated prime to $\sqrt{I}$. Note also that this implies that I_X has codimension $n_x + 1$. Thus, by Lemma 6.12, it is sufficient to prove that there exists a nonempty Zariski-open set O_2 such that for all $(f_1, \ldots, f_m) \in O_1 \cap O_2$, R_X/I_X is a regular local ring.

From the Jacobian Criterion (see e.g. [Eis95], Theorem 16.19), the local ring R_X/I_X is regular if and only if $\mathrm{jac}(f_1, \ldots, f_m)$ taken modulo X has codimension $n_x + 1$. Since the generators of I are bilinear, the latter condition is equivalent to saying that the matrix

$$J_X = \begin{bmatrix} \frac{\partial f_1}{\partial x_0} & \cdots & \frac{\partial f_1}{\partial x_{n_x}} \\ \vdots & \cdots & \vdots \\ \frac{\partial f_m}{\partial x_0} & \cdots & \frac{\partial f_m}{\partial x_{n_x}} \end{bmatrix}$$

has rank $n_x + 1$. We prove below that there exists a nonempty Zariski-open set O_3 such that for all $(f_1, \ldots, f_m) \in O_3$, J_X has rank $n_x + 1$.

Let $\mathfrak{c}_1, \ldots, \mathfrak{c}_m$ be vectors of coordinates of $\mathscr{BL}_{\overline{\mathbb{K}}}(n_x, n_y)^m$, $\mathfrak{M}$ be the vector of all bilinear monomials in R and $\mathfrak{K}$ be the field of rational functions $\mathbb{K}(\mathfrak{c}_1, \ldots, \mathfrak{c}_m)$. Consider the polynomials

$\mathfrak{g}_i = \mathfrak{M}.\mathfrak{c}_i^T$ for $1 \leq i \leq m$ and the Zariski-open set O_3 in $\mathscr{BL}_{\overline{\mathbb{K}}}(n_x, n_y)^m$ defined by the non-vanishing of all the coefficients of the maximal minors of the generic matrix

$$\mathfrak{J}_X = \begin{bmatrix} \frac{\partial \mathfrak{g}_1}{\partial x_0} & \dots & \frac{\partial \mathfrak{g}_1}{\partial x_{n_x}} \\ \vdots & \dots & \vdots \\ \frac{\partial \mathfrak{g}_m}{\partial x_0} & \dots & \frac{\partial \mathfrak{g}_m}{\partial x_{n_x}} \end{bmatrix}.$$

Then $(f_1, \dots, f_m) \in O_3$ implies that J_X has rank $n_x + 1$; our claim follows.

In the case where $n_y \leq m$, the proof follows the same pattern using Lemmas 6.11 and 6.12 and the Jacobian criterion. The only difference is that one has to prove that there exists a nonempty Zariski-open set O_4 such that for all $(f_1, \dots, f_m) \in O_4$ the matrix

$$J_Y = \begin{bmatrix} \frac{\partial f_1}{\partial y_0} & \dots & \frac{\partial f_1}{\partial y_{n_x}} \\ \vdots & \dots & \vdots \\ \frac{\partial f_m}{\partial y_0} & \dots & \frac{\partial f_m}{\partial y_{n_y}} \end{bmatrix}$$

has rank $n_y + 1$, which is done as above. □

Remark 6.14. *The proof of Proposition 6.13 relies on the use of the Jacobian Criterion. From [Eis95, Theorem 16.19], it remains valid if the characteristic of $\mathbb{K}$ is large enough so that the residue class field of X (resp. Y) is separable.*

The two following propositions explain why the rows reduced to zero in the generic case during the F_5 Algorithm have a signature (t, f_i) such that $t \in \mathbb{K}[x_0, \dots, x_{n_x}]$ or $t \in \mathbb{K}[y_0, \dots, y_{n_y}]$.

Proposition 6.15. *Let m be an integer such that $m \leq n_x + n_y$. Then there exists a nonempty Zariski open subset $O \subset \mathscr{BL}_{\overline{\mathbb{K}}}(n_x, n_y)^m$ such that for all bilinear system $\mathbf{F} = (f_1, \dots, f_m) \in O \cap \mathscr{BL}_{\mathbb{K}}(n_x, n_y)^m$, $\mathsf{Syz}(\mathbf{F}) = \langle (\mathsf{Syz}(\mathbf{F}) \cap \mathbb{K}[x_0, \dots, x_{n_x}]^m) \cup (\mathsf{Syz}(\mathbf{F}) \cap \mathbb{K}[y_0, \dots, y_{n_y}]^m) \cup \mathsf{Syz}_{\mathsf{triv}}(\mathbf{F}) \rangle$.*

Proof. Let $s = (s_1, \dots, s_m)$ be a syzygy. Thus, s_m is in $I_{m-1} : f_m$. We can suppose without loss of generality that the s_i are bihomogeneous of same bidegree (Proposition 1.38). According to Theorem 3.12, there exists a nonempty Zariski open set $O_1 \subset \mathscr{BL}_{\overline{\mathbb{K}}}(n_x, n_y)^m$, such that if $(f_1, \dots, f_m) \in O_1$, then f_m is not a divisor of 0 in R/J_{m-1}. We deduce from this observation that $s_m \in J_{m-1}$. So either $s_m \in I_{m-1}$ or there exists P a non-admissible primary component of I_{m-1} such that $s_m \notin P$. Assume that $s_m \notin I_{m-1}$. From Proposition 6.13, there exists a nonempty Zariski open set $O_2 \subset \mathscr{BL}_{\overline{\mathbb{K}}}(n_x, n_y)^m$, such that if $(f_1, \dots, f_m) \in O_2$, then $\langle x_0, \dots, x_{n_x} \rangle = P$ (or $\langle y_0, \dots, y_{n_y} \rangle = P$), which implies that $s_m \in \mathbb{K}[y_0, \dots, y_{n_y}]$ (or $s_m \in \mathbb{K}[x_0, \dots, x_{n_x}]$).

Finally, we see that, if $(f_1, \dots, f_m) \in O_1 \cap O_2$, then $s_m \in I_{m-1} \cup \mathbb{K}[y_0, \dots, y_{n_y}] \cup \mathbb{K}[x_0, \dots, x_{n_x}]$. Since the syzygy module of a bihomogeneous system is generated by bihomogeneous syzygies, it can be deduced that $\mathsf{Syz}(\mathbf{F}) = \langle (\mathsf{Syz}(\mathbf{F}) \cap \mathbb{K}[x_0, \dots, x_{n_x}]^m) \cup (\mathsf{Syz}(\mathbf{F}) \cap \mathbb{K}[y_0, \dots, y_{n_y}]^m) \cup \mathsf{Syz}_{\mathsf{triv}}(\mathbf{F}) \rangle$. □

Proposition 6.16. *Let V be the output of Algorithm* BLcriterion *and let (h, f_i) be an element of V. Then*

- *if $h \in \mathbb{K}[x_0, \dots, x_{n_x}]$, then for all $j, y_j h \in I_{i-1}$.*
- *if $h \in \mathbb{K}[y_0, \dots, y_{n_y}]$, then for all $j, x_j h \in I_{i-1}$.*

Proof. Suppose that $h \in \mathbb{K}[x_0, \ldots, x_{n_x}]$ is a maximal minor of $\mathsf{jac}_{\mathbf{y}}(F_{i-1})$ (the proof is similar if $h \in \mathbb{K}[y_0, \ldots, y_{n_y}]$). Consider the matrix $\mathsf{jac}_{\mathbf{y}}(F_{i-1})$ as defined in Algorithm 7. Then there exists an $(i-1) \times (i-1)$ extension M_T of $\mathsf{jac}_{\mathbf{y}}(F_{i-1})$ such that $\det(\mathsf{M}_T) = h$ (similarly to the proof of Lemma 6.1). Let $0 \leq j \leq n_y$ be an integer. Consider the polynomials $h_1, \ldots, h_{i-1}$, where h_k is the determinant of the $(i-2) \times (i-2)$ matrix obtained by removing the $(j+1)$-th column and the k-th row from M_T.

Then we can remark that

$$\begin{pmatrix} h_1 & -h_2 & \ldots & (-1)^i h_{i-1} \end{pmatrix} \cdot \mathsf{M}_T = \begin{pmatrix} 0 & \ldots & 0 & (-1)^j \det(\mathsf{M}_T) & 0 & \ldots & 0 \end{pmatrix}$$

where the only non-zero component is in the $(j+1)$th column. Keeping only the $n_y + 1$ first columns of M_T, we obtain

$$\begin{pmatrix} h_1 & -h_2 & \ldots & (-1)^i h_{i-1} \end{pmatrix} \cdot \mathsf{jac}_{\mathbf{y}}(F_{i-1}) = \begin{pmatrix} 0 & \ldots & 0 & (-1)^j \det(\mathsf{M}_T) & 0 & \ldots & 0 \end{pmatrix}$$

Since $\mathsf{jac}_{\mathbf{y}}(F_{i-1}) \cdot \begin{pmatrix} y_0 \\ \vdots \\ y_{n_y} \end{pmatrix} = \begin{pmatrix} f_1 \\ \vdots \\ f_{i-1} \end{pmatrix}$, the following equality holds

$$\begin{pmatrix} h_1 & -h_2 & \ldots & (-1)^{i-1} h_{i-2} & (-1)^i h_{i-1} \end{pmatrix} \cdot \begin{pmatrix} f_1 \\ \vdots \\ f_{i-1} \end{pmatrix} = y_j \det(\mathsf{M}_T) = y_j h.$$

This implies that $y_j h \in I_{i-1}$. □

Corollary 6.17. *Let m be an integer such that $m \leq n_x + n_y$ and let $\mathbf{F} = (f_1, \ldots, f_m)$ be a sequence of bilinear polynomials. Let V be the output of Algorithm* BLCRITERION*. Assume that*

$$(I_{m-1} : f_m) \cap \mathbb{K}[x_0, \ldots, x_{n_x}] = \langle \{h \in \mathbb{K}[x_0, \ldots, x_{n_x}] \ : \ (h, f_m) \in V\} \rangle.$$

$$(I_{m-1} : f_m) \cap \mathbb{K}[y_0, \ldots, y_{n_y}] = \langle \{h \in \mathbb{K}[y_0, \ldots, y_{n_y}] \ : \ (h, f_m) \in V\} \rangle.$$

Let G_x (resp G_y) be a Gröbner basis of $(I_{m-1} : f_m) \cap \mathbb{K}[x_0, \ldots, x_{n_x}]$ (resp. $(I_{m-1} : f_m) \cap \mathbb{K}[y_0, \ldots, y_{n_y}]$) and let G_{m-1} be a Gröbner basis of I_{m-1}. If $\mathsf{Syz}(\mathbf{F}) = \langle (\mathsf{Syz}(\mathbf{F}) \cap \mathbb{K}[x_0, \ldots, x_{n_x}]^m) \cup (\mathsf{Syz}(\mathbf{F}) \cap \mathbb{K}[y_0, \ldots, y_{n_y}]^m) \cup \mathsf{Syz}_{\mathsf{triv}}(\mathbf{F}) \rangle$, then $G_x \cup G_y \cup G_{m-1}$ is a Gröbner basis of $I_{m-1} : f_m$.

Proof. Let $f \in I_{m-1} : f_m$ be a polynomial. Thus there exist $s_1, \ldots, s_{m-1}$ such that $(s_1, \ldots, s_{m-1}, f) \in \mathsf{Syz}(\mathbf{F})$. Since I_{m-1} and f_m are bihomogeneous, we can suppose without loss of generality that f is bihomogeneous (Proposition 1.38). Let (d_1, d_2) denote its bidegree.

- If $d_2 = 0$ (resp. $d_1 = 0$), then $f \in \langle G_x \rangle$ (resp. $f \in \langle G_y \rangle$).
- Let $G_x = \{g_i^{(x)}\}_{1 \leq i \leq \mathsf{card}(G_x)}$ and $G_y = \{g_i^{(y)}\}_{1 \leq i \leq \mathsf{card}(G_y)}$. If $d_1 \neq 0$ and $d_2 \neq 0$ then, since $\mathsf{Syz}(\mathbf{F}) = \langle (\mathsf{Syz}(\mathbf{F}) \cap \mathbb{K}[x_0, \ldots, x_{n_x}]^m) \cup (\mathsf{Syz}(\mathbf{F}) \cap \mathbb{K}[y_0, \ldots, y_{n_y}]^m) \cup \mathsf{Syz}_{\mathsf{triv}}(\mathbf{F}) \rangle$,

$$f = \sum_{1 \leq i \leq \mathsf{card}(G_x)} q_i g_i^{(x)} + \sum_{1 \leq i \leq \mathsf{card}(G_y)} q_i' g_i^{(y)} + t$$

where $t \in I_{m-1}$ is a bihomogeneous polynomial and the q_i and q_i' are also bihomogeneous. Since $d_2 \neq 0$ and $g_i^{(x)} \in \mathbb{K}[x_0, \ldots, x_{n_x}]$, q_i must be in $\langle y_0, \ldots, y_{n_y} \rangle$. According to Proposition 6.16, for all i, $q_i g_i^{(x)} \in I_{m-1}$. By a similar argument, for all i, $q_i' g_i^{(y)} \in I_{m-1}$. Finally, $f \in I_{m-1}$.

We just proved that $I_{m-1} : f_m \subset I_{m-1} \cup \langle G_x \rangle \cup \langle G_y \rangle$. By construction, we also have the other inclusion $I_{m-1} \cup \langle G_x \rangle \cup \langle G_y \rangle \subset I_{m-1} : f_m$. Thus, $G_x \cup G_y \cup G_{m-1}$ is a Gröbner basis of $I_{m-1} : f_m$. □

Corollary 6.17 shows that, when a bilinear system is bi-regular, it is possible to find a Gröbner basis of $I_{m-1} : f_m$ (which yields the monomials t such that the row (t, f_m) reduces to zero) as soon as we know the three Gröbner bases G_x, G_y, and G_{m-1}. In fact, we only need G_x and G_y since the reductions to zero corresponding to G_{m-1} are eliminated by the usual F_5 criterion. Fortunately, we can obtain G_x and G_y just by performing linear algebra over the maximal minors of a matrix (Theorem 6.5).

We now present the main result of this section. If we suppose that Conjecture 6.7 is true, then the following Theorem shows that generic bilinear systems are bi-regular.

Theorem 6.18. *Let $m, n_x, n_y \in \mathbb{N}$ such that $m < n_x + n_y$. If Conjecture 6.7 is true, then there exists a nonempty Zariski open subset $\mathcal{O} \subset \mathscr{BL}_{\overline{\mathbb{K}}}(n_x, n_y)^m$ such that every sequence $\mathbf{F} = (f_1, \ldots, f_m) \in \mathcal{O} \cap \mathscr{BL}_{\mathbb{K}}(n_x, n_y)^m$ is bi-regular. Moreover, if $(f_1, \ldots, f_m)$ is a bi-regular sequence, then there are no reductions to zero with the extended F_5 criterion.*

Proof. Let G_m be a minimal Gröbner basis of $I_{m-1} : f_m$. The reductions to zero (t, f_m) which are not detected by the usual F_5 criterion are exactly those such that $t \in \mathsf{LM}(G_m)$ and $t \notin \mathsf{LM}(I_{m-1})$. We showed that there exists a nonempty Zariski open subset O_1 of $\mathscr{BL}_{\overline{\mathbb{K}}}(n_x, n_y)$ such that if $f_m \in O_1$, then $t \in \mathsf{LM}(I_{m-1} : f_m \cap \mathbb{K}[x_0, \ldots, x_{n_x}])$ or $t \in \mathsf{LM}(I_{m-1} : f_m \cap \mathbb{K}[y_0, \ldots, y_{n_y}])$ (Proposition 6.15). If we suppose that Conjecture 6.7 is true, then there exists a nonempty Zariski open subset O_2 of $\mathscr{BL}_{\overline{\mathbb{K}}}(n_x, n_y)$ such that if $f_m \in O_2$, $I_{m-1} : f_m \cap \mathbb{K}[x_0, \ldots, x_{n_x}]$ (resp. $I_{m-1} : f_m \cap \mathbb{K}[y_0, \ldots, y_{n_y}]$) is spanned by the maximal minors of $\mathsf{jac}_{\mathbf{x}}(F_{m-1})$ (resp. $\mathsf{jac}_{\mathbf{y}}(F_{m-1})$). Thus, by Theorem 6.5, there exists a nonempty Zariski open subset O_3 of $\mathscr{BL}_{\overline{\mathbb{K}}}(n_x, n_y)$ such that if $f_m \in O_3$, $\mathsf{LM}(I_{m-1} : f_m \cap \mathbb{K}[x_0, \ldots, x_{n_x}]) = \mathsf{Monomials}^{\mathbf{x}}_{m-n_y-2}(n_y + 1)\rangle$ (resp. $\mathsf{LM}(I_{m-1} : f_m \cap \mathbb{K}[y_0, \ldots, y_{n_y}]) = \mathsf{Monomials}^{\mathbf{y}}_{m-n_x-2}(n_x + 1)\rangle$). Suppose that $f_m \in O_1 \cap O_2 \cap O_3$ (which is a nonempty Zariski open subset) and that (t, f_m) is a reduction to zero such that $t \notin \mathsf{LM}(I_{m-1})$. Then

$$t \in \langle \mathsf{Monomials}^{\mathbf{x}}_{m-n_y-2}(n_y + 1)\rangle$$
$$\text{or}$$
$$t \in \langle \mathsf{Monomials}^{\mathbf{y}}_{m-n_x-2}(n_x + 1)\rangle.$$

By Theorem 6.5, t is a leading monomial of a linear combination of the maximal minors of $\mathsf{jac}_{\mathbf{x}}(F_{m-1})$ (or $\mathsf{jac}_{\mathbf{y}}(F_{m-1})$). Consequently, the reduction to zero (t, f_m) is detected by the extended F_5 criterion. □

Remark 6.19. *Thanks to the analysis of Algorithm 7, we know exactly which reductions to zero can be avoided during the computation of a Gröbner basis of a bilinear system. If a bilinear system is bi-regular, then Algorithm 7 finds all reductions to zero. Indeed, this algorithm detects reductions to zero coming from linear combinations of maximal minors of the matrices $\mathsf{jac}_{\mathbf{x}}(F_i)$ and $\mathsf{jac}_{\mathbf{y}}(F_i)$. According to Theorem 6.18, there are no other reductions to zero for bi-regular systems.*

Example 6.20. *The system $f_1, \ldots, f_5$ given in Example 6.4 is bi-regular and there are no reduction to zero during the computation of a Gröbner basis with the extended F_5 criterion.*

6.4 Hilbert bi-series of bilinear systems

An important tool to describe ideals spanned by bilinear equations is the so-called *Hilbert series*. In the homogeneous case, complexity results for F_5 were obtained with this tool (see e.g. [BFSY04, Bar04]). In this section, we provide an explicit form of the Hilbert bi-series – a bihomogeneous analog of the Hilbert series – for ideals spanned by generic bilinear systems. To find this bi-series, we use the combinatorics of the syzygy module of bi-regular systems. With this tool, we will be able to do a complexity analysis of a special version of the F_5 which will be presented in the next section.

The following notation will be used throughout this chapter: the vector space of bihomogeneous polynomials of bidegree (α, β) will be denoted by $R_{\alpha,\beta}$. If I is a bihomogeneous ideal, then $I_{\alpha,\beta}$ will denote the vector space $I \cap R_{\alpha,\beta}$.

Let I be a bihomogeneous ideal of R. We recall that the Hilbert bi-series is defined by

$$\mathsf{mHS}_{R/I}(t_1, t_2) = \sum_{(\alpha,\beta)\in\mathbb{N}^2} \dim(R_{\alpha,\beta}/I_{\alpha,\beta}) t_1^\alpha t_2^\beta.$$

Remark 6.21. *The usual univariate Hilbert series for homogeneous ideals can easily be deduced from the Hilbert bi-series by putting $t_1 = t_2$ (see [ST06]).*

We can now present the main result of this section: an explicit form of the bi-series for bi-regular bilinear systems.

Theorem 6.22. *Let $f_1, \dots, f_m \in R$ be a bi-regular bilinear sequence, with $m \le n_x + n_y$. Then its Hilbert bi-series is*

$$\mathsf{mHS}_{\mathbb{K}[x_1,\dots,x_{n_x},y_1,\dots,y_{n_y}]/I}(t_1, t_2) = \frac{(1 - t_1 t_2)^m + N_m(t_1, t_2) + N_m(t_2, t_1)}{(1 - t_1)^{n_x+1}(1 - t_2)^{n_y+1}},$$

$$N_m(t_1, t_2) = \sum_{\ell=1}^{m-(n_y+1)} (1 - t_1 t_2)^{m-(n_y+1)-\ell} t_1 t_2 (1 - t_2)^{n_y+1} \left[1 - (1 - t_1)^\ell \sum_{k=1}^{n_y+1} t_1^{n_y+1-k} \binom{\ell + n_y - k}{n_y + 1 - k}\right]$$

We decompose the proof of this theorem into a sequence of lemmas.

If I is an ideal of R and f is a polynomial, we denote by $\bar{f}$ the equivalence class of f in R/I and

$$\mathrm{ann}_{R/I}(f) = \{v \in R/I : v\bar{f} = 0\},$$

$$\mathrm{ann}_{R/I}(f)_{\alpha,\beta} = \{v \in R/I \text{ of bidegree } (\alpha, \beta) : v\bar{f} = 0\}.$$

If I is a bihomogeneous ideal and f is a bihomogeneous polynomial, we use the following notation:

$$G_{I,f}(t_1, t_2) = \sum_{(\alpha,\beta)\in\mathbb{N}^2} \dim(\mathrm{ann}_{R/I}(f)_{\alpha,\beta}) t_1^\alpha t_2^\beta.$$

Lemma 6.23. *Let $f_1, \dots, f_m \in R$ be bihomogeneous polynomials, with $1 < m \le n_x + n_y$. Let (d_1, d_2) be the bidegree of f_m. Then*

$$\mathsf{mHS}_{R/I_m}(t_1, t_2) = (1 - t_1^{d_1} t_2^{d_2}) \,\mathsf{mHS}_{R/I_{m-1}} + t_1^{d_1} t_2^{d_2} G_{I_{m-1},f}(t_1, t_2).$$

Proof. We have the following exact sequence:

$$0 \to \mathrm{ann}_{R/I_{m-1}}(f) \xrightarrow{\varphi_1} R/I_{m-1} \xrightarrow{\varphi_2} R/I_{m-1} \xrightarrow{\varphi_3} R/I_m \to 0.$$

where φ_1 and φ_3 are the canonical inclusion and projection, and φ_2 is the multiplication by f_m.

From this exact sequence of ideals, we can deduce an exact sequence of vector spaces:

$$0 \to (\mathrm{ann}_{R/I_{m-1}}(f))_{\alpha,\beta} \xrightarrow{\varphi_1} \left(\frac{R}{I_{m-1}}\right)_{\alpha,\beta} \xrightarrow{\varphi_2} \left(\frac{R}{I_{m-1}}\right)_{\alpha+d_1,\beta+d_2} \xrightarrow{\varphi_3} \left(\frac{R}{I_m}\right)_{\alpha+d_1,\beta+d_2} \to 0.$$

Thus the alternate sum of the dimensions of vector spaces of an exact sequence is 0:

$$\begin{aligned}&\dim((\mathrm{ann}_{R/I_{m-1}}(f))_{\alpha,\beta}) - \dim\left(\left(\tfrac{R}{I_{m-1}}\right)_{\alpha,\beta}\right) + \\ &\dim\left(\left(\tfrac{R}{I_{m-1}}\right)_{\alpha+d_1,\beta+d_2}\right) - \dim\left(\left(\tfrac{R}{I_m}\right)_{\alpha+d_1,\beta+d_2}\right) = 0.\end{aligned}$$

By multiplying this relation by $t_1^{\alpha}t_2^{\beta}$ and by summing over (α,β), we obtain the claimed relation:

$$\mathsf{mHS}_{R/I_m}(t_1,t_2) = (1-t_1^{d_1}t_2^{d_2})\,\mathsf{mHS}_{R/I_{m-1}} + t_1^{d_1}t_2^{d_2}G_{I_{m-1},f}(t_1,t_2).$$

□

Lemma 6.24. *Let $f_1,\dots,f_m \in R$ be a bi-regular bilinear sequence, with $m \le n_x+n_y$. Then, for all $2 \le i \le m$,*

$$G_{I_{i-1},f_i}(t_1,t_2) = g_x^{(i-1)}(t_1) + g_y^{(i-1)}(t_2),$$

where

$$g_x^{(i-1)}(t) = \begin{cases} 0 \text{ if } i \le n_y+1 \\ \frac{1}{(1-t)^{n_x+1}} - \sum_{1\le j\le n_y+1} \frac{\binom{i-1-j}{n_y+1-j}t^{n_y+1-j}}{(1-t)^{n_x+n_y-i+2}} \end{cases}.$$

$$g_y^{(i-1)}(t) = \begin{cases} 0 \text{ if } i \le n_x+1 \\ \frac{1}{(1-t)^{n_y+1}} - \sum_{1\le j\le n_x+1} \frac{\binom{i-1-j}{n_x+1-j}t^{n_x+1-j}}{(1-t)^{n_x+n_y-i+2}} \end{cases}.$$

Proof. Saying that $v \in \mathrm{ann}_{R/I_{i-1}}(f_i)$ is equivalent to saying that the row with signature $(\mathsf{LM}(v), f_i)$ is not detected by the classical F_5 criterion. According to Theorem 6.18, if the system is bi-regular, the reductions to zero corresponding to non-trivial syzygies are exactly:

$$\bigcup_{i=n_x+2}^{m} \{(t,f_i) : t \in \mathsf{Monomials}^{\mathbf{y}}_{i-n_x-2}(n_x+1)\} \bigcup_{i=n_y+2}^{m} \{(t,f_i) : t \in \mathsf{Monomials}^{\mathbf{x}}_{i-n_y-2}(n_y+1)\}.$$

By Proposition 6.16, we know that if $P \in \mathbb{K}[x_0,\dots,x_{n_x}] \cap (I_{i-1} : f_i)$ (resp. $\mathbb{K}[y_0,\dots,y_{n_y}] \cap (I_{i-1} : f_i)$), then $\forall j, y_jP \in I_{i-1}$ (resp. $x_jP \in I_{i-1}$). Thus $G_{I_{i-1},f_i}(t_1,t_2)$ is the generating bi-series of the monomials in $\mathbb{K}[x_0,\dots,x_{n_x}]$ which are a multiple of a monomial of degree n_y+1 in $x_0,\dots,x_{i-n_y-2}$ and of the monomials in $\mathbb{K}[y_0,\dots,y_{n_y}]$ which are a multiple of a monomial of degree n_x+1 in $y_0,\dots,y_{i-n_x-2}$. Denote by $g_x^{(i-1)}(t)$ (resp. $g_y^{(i-1)}(t)$) the generating series of the monomials in $\mathbb{K}[x_0,\dots,x_{n_x}]$ (resp. $\mathbb{K}[y_0,\dots,y_{n_y}]$) which are a multiple of a monomial of degree n_y+1 (resp. n_x+1) in $x_0,\dots,x_{i-n_y-2}$ (resp. $y_0,\dots,y_{i-n_x-2}$). Then we have

$$G_{I_{i-1},f_i}(t_1,t_2) = g_x^{(i-1)}(t_1) + g_y^{(i-1)}(t_2).$$

Next we use combinatorial techniques to give an explicit form of $g_x^{(i-1)}(t)$ and $g_y^{(i-1)}(t)$. Let $c(t)$ denote the generating series of the monomials in $\mathbb{K}[x_{i-n_y-1},\dots,x_{n_x}]$:

$$c(t) = \sum_{j=0}^{\infty} \binom{n_x+n_y-i+j+1}{j} t^j = \frac{1}{(1-t)^{n_x+n_y-i+2}}.$$

Let B_j denote the number of monomials in $\mathbb{K}[x_0, \ldots, x_{i-n_y-2}]$ of degree j. Then

$$\frac{1}{(1-t)^{n_x+n_y+2}} = c(t) + B_1 c(t) t + \cdots + B_{n_y} c(t) t^{n_y} + g_x^{(i-1)}(t).$$

Since $B_j = \binom{i-n_y-1+j}{j}$, we can conclude:

$$g_x^{(i-1)}(t) = \begin{cases} 0 \text{ if } i \leq n_y + 1 \\ \frac{1}{(1-t)^{n_x+1}} - \sum_{1 \leq j \leq n_y+1} \frac{\binom{i-1-j}{n_y+1-j} t^{n_y+1-j}}{(1-t)^{n_x+n_y-i+2}} \end{cases}.$$

□

Proof of Theorem 6.22. Since the polynomials are bilinear, by Lemma 6.23, we have

$$\mathsf{mHS}_{R/I_i}(t_1, t_2) = (1 - t_1 t_2)\, \mathsf{mHS}_{R/I_{i-1}} + t_1 t_2 G_{I_{i-1}, f_i}(t_1, t_2).$$

Lemma 6.24 gives the value of $G_{I_{i-1}, f_i}(t_1, t_2)$. To initiate the induction, we need

$$\mathsf{mHS}_{R/I_0}(t_1, t_2) = \mathsf{mHS}_{R/\langle 0 \rangle}(t_1, t_2) = \frac{1}{(1-t_1)^{n_x+1}(1-t_2)^{n_y+1}}.$$

Then we obtain the claimed form of the bi-series by induction:

$$\mathsf{mHS}_{R/I_i}(t_1, t_2) = \frac{(1 - t_1 t_2)^i + N_i(t_1, t_2)}{(1-t_1)^{n_x+1}(1-t_2)^{n_y+1}}$$

$$N_i(t_1, t_2) = \sum_{j=0}^{m-1} t_1 t_2 (1 - t_1 t_2)^j G_{I_j, f_{j+1}}(t_1, t_2).$$

□

Example 6.25. *The Hilbert bi-series of the ideal generated by the five polynomials of Example 6.4 is*

$$\mathsf{mHS}_{R/I_5}(t_1, t_2) = \frac{1}{(1-t_1)^3 (1-t_2)^4} ({t_1}^5 {t_2}^5 - 4{t_1}^5 {t_2}^4 + 6{t_1}^5 {t_2}^3 - 4{t_1}^5 {t_2}^2 + {t_1}^5 t_2 - 6{t_1}^3 {t_2}^5 + 15{t_1}^3 {t_2}^4 - 10{t_1}^3 {t_2}^3 + 8{t_1}^2 {t_2}^5 - 15{t_1}^2 {t_2}^4 + 10{t_1}^2 {t_2}^2 - 3t_1 {t_2}^5 + 5t_1 {t_2}^4 - 5t_1 t_2 + 1),$$

and is in accordance with the formula given in Theorem 6.22. Also, notice that the intermediate series $g_x(t)$ and $g_y(t)$ match the theoretical values. For instance:

$$g_y^{(3)}(t) = \frac{t^3}{(1-t)^4}.$$

6.5 Towards complexity results

6.5.1 A multihomogeneous F_5 Algorithm

We now describe how it is possible to use the multihomogeneous structure of the matrices arising in the Matrix F_5 Algorithm to speed-up the computation of a Gröbner basis. In order to have simple notations, the description is made in the context of bihomogeneous systems, but it can be easily transposed in the context of multihomogeneous systems.

Let $f_1, \ldots, f_m$ be a sequence of bihomogeneous polynomials. Consider the matrices M_d in degree d appearing during the Matrix F_5 Algorithm. Notice that each row represents a bihomogeneous polynomial. Let (d_1, d_2) be the bidegree of one row of this matrix. Then the only non-zero coefficients on this row are in columns which represent a monomial of bidegree (d_1, d_2). Therefore a possible strategy to use the bihomogeneous structure is the following:

					Multihomogeneous		Homogeneous		
n_x	n_y	m	bidegree	D	time	memory	time	memory	speed-up
3	4	7	$(1,1)$	6	**16.9s**	**30MB**	265.7s	280MB	**16**
3	4	7	$(1,1)$	7	**105s**	**92MB**	2018s	1317MB	**19**
4	4	8	$(1,1)$	7	**582s**	**275MB**	13670s	4210MB	**23**
5	4	9	$(1,1)$	7	**3343s**	**957MB**	66371s	12008MB	**20**
5	5	10	$(1,1)$	6	**645s**	**435MB**	10735s	4330MB	**17**
2	2	4	$(1,2)$	10	**11.4s**	**19MB**	397s	299MB	**35**
2	2	4	$(1,2)$	8	**1.7s**	**10MB**	16s	52MB	**9**
3	3	6	$(1,2)$	8	**67s**	**80MB**	1146s	983MB	**17**
4	4	8	$(1,2)$	8	**2222s**	**1031MB**	40830s	12319MB	**63**
2	2	4	$(2,2)$	11	**29s**	**27MB**	899s	553MB	**31**
3	3	6	$(2,2)$	8	**27s**	**47MB**	277s	452MB	**10**
3	3	6	$(2,2)$	9	**152s**	**154MB**	2380s	1939MB	**16**
3	4	7	$(2,2)$	9	**1034s**	**505MB**	18540s	7658MB	**18**
4	4	8	$(2,2)$	8	**690s**	**385MB**	7260s	4811MB	**11**
4	4	8	$(2,2)$	9	**6355s**	**2216MB**	—	>20000MB	—

Table 6.1: Execution time and memory usage of the multihomogeneous variant of F_5

- For each couple (d_1, d_2) such that $d_1 + d_2 = d$, construct the matrix M_{d_1,d_2}. The rows of this matrix represent the polynomials of M_d of bidegree (d_1, d_2) and the columns represent the monomials of R_{d_1,d_2}.
- Compute the row echelon form of the matrices M_{d_1,d_2}. This gives bases of I_{d_1,d_2}.
- The union of the bases gives a basis of I_d since $I_d = \bigoplus_{d_1+d_2=d} I_{d_1,d_2}$.

This way, instead of computing the row echelon form of a big matrix, we can decompose the problem and compute independently the row echelon form of smaller matrices. This strategy can be extended to multihomogeneous systems.

In Table 6.1, the execution time and the memory usage of this multihomogeneous variant of F_5 are compared to the classical homogeneous Matrix F_5 Algorithm for computing a D-Gröbner basis for random bihomogeneous systems (for the grevlex ordering). Both implementations are made in `Magma2.15-7` and follow the general framework of Algorithm 5. However, the row echelon form computation are performed with the naive algorithm (without taking advantage of the sparseness and of the structure of the Macaulay matrices). The experimental results have been obtained with a Xeon processor 2.50GHz cores and 20 GB of RAM. We are aware that we should compare efficient implementations (in a low-level language and with linear algebra routines adapted to the shape of the Macaulay matrices) of these two algorithms to have a more precise evaluation of the speed-up we can expect for practical applications. However, these experiments give a first estimation of that speed-up. Furthermore, we can also expect to save a lot of memory by decomposing the Macaulay matrix into smaller matrices. This is crucial for practical applications, since untractability is often due to the lack of memory.

6.5.2 Complexity estimates

In this section, we provide a theoretical explanation of the speed-up observed when using the bihomogeneous structure of bilinear systems. To estimate the complexity of the Matrix F_5 Algorithm, we consider that the cost is dominated by the cost of the reductions of the matrices with the highest degree. By using the new criterion described in Section 6.2.4, all the matrices appearing during the

n_x	n_y	m	D	*experimental speed−up*	$\mathbf{F}(n_x, n_y, m, D)$
3	4	7	6	**16**	**29**
3	4	7	7	**19**	**34**
4	4	8	7	**23**	**34**
5	4	9	7	**20**	**32**
5	5	10	6	**17**	**27**

Table 6.2: Decomposing the matrices: experimental speed-up

computations have full rank for generic inputs (these ranks are the dimensions of the $\mathbb{K}$-vector spaces I_{d_1,d_2}). We consider that the complexity of reducing a $r \times c$ matrix with Gauss elimination is $O(r^2c)$. Thus the complexity of computing a D-Gröbner basis with the usual Matrix F_5 Algorithm and the extended criterion for a bilinear system of m equations over $\mathbb{K}[x_0, \ldots, x_{n_x}, y_0, \ldots, y_{n_y}]$ is

$$T_{hom} = C_1 \left(\left(\binom{D + n_x + n_y + 1}{D} - [t^D]\,\mathsf{mHS}(t,t) \right)^2 \binom{D + n_x + n_y + 1}{D} \right).$$

When using the multihomogeneous structure, the complexity becomes:

$$T_{multihom} = C_2 \left(\sum_{\substack{d_1+d_2=D \\ 1 \le d_1, d_2 \le D-1}} \left(\dim(R_{d_1,d_2}) - [t_1^{d_1} t_2^{d_2}]\,\mathsf{mHS}(t_1,t_2) \right)^2 \dim(R_{d_1,d_2}) \right),$$

where $\dim(R_{d_1,d_2}) = \binom{d_1+n_x}{d_1}\binom{d_2+n_y}{d_2}$. Thus the theoretical speed-up that we expect is:

$$speedup_{th} = C_3 \mathbf{F}(n_x, n_y, m, D)$$

where $C_3 = \frac{C_1}{C_2}$ is a constant and

$$\mathbf{F}(n_x, n_y, m, D) = \left(\frac{\left(\binom{D+n_x+n_y+1}{D} - [t^D]\,\mathsf{mHS}(t,t) \right)^2 \binom{D+n_x+n_y+1}{D}}{\sum_{\substack{d_1+d_2=D \\ 1 \le d_1, d_2 \le D-1}} \left(\dim(R_{d_1,d_2}) - [t_1^{d_1} t_2^{d_2}]\,\mathsf{mHS}(t_1,t_2) \right)^2 \dim(R_{d_1,d_2})} \right).$$

Now let us compare this theoretical speed-up with the one observed in practice. We can see in Table 6.2 that experimental results match the theoretical complexity:

$$speedup \approx 0.6\mathbf{F}(n_x, n_y, m, D).$$

6.5.3 Number of reductions to zero removed by the extended F_5 criterion

Table 6.3 shows the number of reductions to zero during the execution of the Buchberger, F_4 and F_5 algorithm. The input systems are random bilinear systems of $n_x + n_y$ equations over

(n_x, n_y)	Nb. useful red. (Buch./F_4)	Nb red. to 0 (Buch./F_4)	Nb red. to 0 (F_5)
$(5,5)$	752	5772	240
$(5,6)$	1484	13063	495
$(6,6)$	3009	29298	990
$(6,7)$	5866	64093	2002
$(4,8)$	1912	19055	990
$(4,9)$	2869	31737	1794
$(3,10)$	1212	13156	1300
$(3,11)$	1665	19780	2016
$(3,12)$	2123	27295	3018

Table 6.3: Experimental number of reductions to zero

$\mathsf{GF}_{65521}[x_0, \dots, x_{n_x}, y_0, \dots, y_{n_y}]$. Experimentally, there is no reduction to zero when using the extended criterion (Algorithm 7). Notice that the number of reductions to zero which are not detected by the classical F_5 criterion matches the theoretical value for a bi-regular system (Definition 6.9):

$$\sum_{i=n_y+1}^{n_x+n_y-1} \binom{i}{n_y+1} + \sum_{i=n_x+1}^{n_x+n_y-1} \binom{i}{n_x+1}.$$

Although the number of reductions to zero removed by the extended criterion is not small compared to the number of useful reductions, they arise in low degree ($n_x + 1$ and $n_y + 1$). Hence, it is not clear what speed-up could be expected with an efficient implementation.

6.5.4 Structure of generic affine bilinear systems

In this part, m, n_x and n_y are three integers such that $m = n_x + n_y$. We consider affine systems of bilinear polynomials $\mathbf{F} = (f_1, \dots, f_m) \in \mathbb{K}[x_1, \dots, x_{n_x}, y_1, \dots, y_{n_y}]^m$. ϑ denotes the dehomogenization morphism:

$$\begin{array}{ccc} \mathbb{K}[x_0, \dots, x_{n_x}, y_0, \dots, y_{n_y}] & \longrightarrow & \mathbb{K}[x_1, \dots, x_{n_x}, y_1, \dots, y_{n_y}] \\ f(x_0, \dots, x_{n_x}, y_0, \dots, y_{n_y}) & \longmapsto & f(1, x_1, \dots, x_{n_x}, 1, y_1, \dots, y_{n_y}) \end{array}.$$

We denote by $\mathscr{BL}_{\mathbb{K}}^{a}(n_x, n_y) \subset \mathbb{K}[x_1, \dots, x_{n_x}, y_1, \dots, y_{n_y}]$ the set of *affine bilinear polynomials*, i.e. the image under ϑ of $\mathscr{BL}_{\mathbb{K}}(n_x, n_y)$ (ϑ is actually a bijection between $\mathscr{BL}_{\mathbb{K}}^{a}(n_x, n_y)$ and $\mathscr{BL}_{\mathbb{K}}(n_x, n_y)$).

We still assume without loss of generality that $n_x \leq n_y$. We also assume in this part of the chapter that the characteristic of $\mathbb{K}$ is 0 (although the results remain true when the characteristic is large enough).

First, we show that generic *affine* bilinear systems have a particular structure: they are regular (Definition 1.44). Consequently, the usual F_5 criterion removes all reductions to zero.

Proposition 6.26. *Let S be the set of affine bilinear systems over $\mathbb{K}[x_1, \dots, x_{n_x}, y_1, \dots, y_{n_y}]$ with $m \leq n_x + n_y$ equations. Then the subset*

$$\{(f_1, \dots, f_m) \in S \ : \ (f_1, \dots, f_m) \text{ is a regular sequence}\}$$

contains a Zariski nonempty open subset of S.

Proof. Let $(f_1,\dots,f_m)$ be a generic affine bilinear system. Assume that it is not regular. Then for some i, there exists $g\in R$ such that $g\notin I_{i-1}$ and $gf_i\in I_{i-1}$. Denote by g^h the bi-homogenization of g. Then $g^h\in\langle f_1^h,\dots,f_{i-1}^h\rangle : f_i^h$. $(f_1^h,\dots,f_m^h)$ is a generic bilinear system, hence it is bi-regular (Theorem 6.18). Thus $g^h\in\mathbb{K}[x_0,\dots,x_{n_x}]$ or $g^h\in\mathbb{K}[y_0,\dots,y_{n_y}]$. Let us suppose that $g^h\in\mathbb{K}[x_0,\dots,x_{n_x}]$ (the proof is similar if $g^h\in\mathbb{K}[y_0,\dots,y_{n_y}]$). Therefore $y_{n_y}g^h\in\langle f_1^h,\dots,f_{i-1}^h\rangle$ when the system is bi-regular (Proposition 6.16). By putting $x_{n_x}=1$ and $y_{n_y}=1$, we see that in this case, $g\in I_{i-1}$, which yields a contradiction. This shows that generic affine bilinear systems are regular. □

6.5.5 Maximal degree reached during the computation of affine bilinear systems

The goal of this section is to give an upper bound on the *maximal degree* reached during the computation of a grevlex Gröbner basis of a generic affine bilinear system with m equations and m variables. In the following, $\prec$ still denotes the grevlex ordering.

First we prove that all monomials in $\mathbb{K}[Y]$ of degree n_x+1 can be obtained by computing the row echelon form of the affine Macaulay matrix in degree n_x+2. To do this we use the notation χ introduced in Definition 1.71. We recall that $\chi(n_x+2,\mathbf{F})$ is the vector space of the polynomials that are algebraic combination of degree at most n_x+2 of $f_1,\dots,f_m$.

Lemma 6.27. *There exists a non-empty Zariski open subset $O\subset\mathscr{BL}_{\mathbb{K}}^a(n_x,n_y)^m$ such that, for any $\mathbf{F}\in O$, all monomials in $\mathbb{K}[Y]$ of degree n_x+1 are in the set* $\{\mathsf{LM}(f)\mid f\in\chi(n_x+2,\mathbf{F})\}$

Proof. For $\mathbf{F}=(f_1,\dots,f_m)\in\mathscr{BL}_{\mathbb{K}}^a(n_x,n_y)^m$, we let $A_{\mathbf{F}}$ be the $m\times(n_x+1)$-matrix defined by

$$A_{\mathbf{F}}=\begin{bmatrix}\frac{\partial f_1}{\partial x_1} & \dots & \frac{\partial f_1}{\partial x_{n_x}} & f_1(0,\dots,0,y_1,\dots,y_{n_y})\\ \vdots & \vdots & \vdots & \vdots\\ \frac{\partial f_m}{\partial x_1} & \dots & \frac{\partial f_m}{\partial x_{n_x}} & f_m(0,\dots,0,y_1,\dots,y_{n_y})\end{bmatrix}$$

We first prove that all maximal minors of $A_{\mathbf{F}}$ belong to $\chi(n_x+2,\mathbf{F})$. Let $1\le\ell_1<\dots<\ell_{n_x+1}\le m$ be a sequence of indices of rows of $A_{\mathbf{F}}$ and let $M\in\mathbb{K}[Y]$ be the determinant of the submatrix of $A_{\mathbf{F}}$ obtained by considering only the rows $\ell_1,\dots,\ell_{n_x+1}$. Now let $\mathbf{v}$ be the $1\times m$ vector defined by

$$v_i=\begin{cases}0 \text{ if } i\notin\{\ell_1,\dots,\ell_{n_x+1}\}\\ (-1)^k\,\mathsf{minor}(\mathsf{jac}_{\mathbf{x}}(\mathbf{F}),(\ell_1,\dots,\ell_{k-1},\ell_{k+1},\dots,\ell_{n_x+1})) \text{ if } \ell_k=i\end{cases}$$

By definition of $\mathbf{v}$, we have for each $i\in\{1,\dots,n_x\}$

$$\mathbf{v}\cdot\begin{bmatrix}\frac{\partial f_1}{\partial x_i}\\ \vdots\\ \frac{\partial f_m}{\partial x_i}\end{bmatrix}=\det\begin{bmatrix}\frac{\partial f_{\ell_1}}{\partial x_i} & \frac{\partial f_{\ell_1}}{\partial x_1} & \dots & \frac{\partial f_{\ell_1}}{\partial x_{n_x+1}}\\ \vdots & \vdots & \vdots & \vdots\\ \frac{\partial f_{\ell_{n_x+1}}}{\partial x_i} & \frac{\partial f_{\ell_{n_x+1}}}{\partial x_1} & \dots & \frac{\partial f_{\ell_{n_x+1}}}{\partial x_{n_x+1}}\end{bmatrix}=0.$$

Moreover,

$$\mathbf{v}\cdot\begin{bmatrix}f_1(0,\dots,0,y_1,\dots,y_{n_y})\\ \vdots\\ f_m(0,\dots,0,y_1,\dots,y_{n_y})\end{bmatrix}=\det\begin{bmatrix}f_{\ell_1}(0,\dots,0,y_1,\dots,y_{n_y}) & \frac{\partial f_{\ell_1}}{\partial x_1} & \dots & \frac{\partial f_{\ell_1}}{\partial x_{n_x+1}}\\ \vdots & \vdots & \vdots & \vdots\\ f_{\ell_{n_x+1}}(0,\dots,0,y_1,\dots,y_{n_y}) & \frac{\partial f_{\ell_{n_x+1}}}{\partial x_1} & \dots & \frac{\partial f_{\ell_{n_x+1}}}{\partial x_{n_x+1}}\end{bmatrix}=M.$$

Now notice that

$$\begin{bmatrix} f_1 \\ \vdots \\ f_m \end{bmatrix} = \mathsf{jac}_{\mathbf{x}}(\mathbf{F}) \cdot \begin{bmatrix} x_1 \\ \vdots \\ x_{n_x} \end{bmatrix} + \begin{bmatrix} f_1(0,\dots,0,y_1,\dots,y_{n_y}) \\ \vdots \\ f_m(0,\dots,0,y_1,\dots,y_{n_y}) \end{bmatrix}.$$

Consequently,

$$\mathbf{v} \cdot \begin{bmatrix} f_1 \\ \vdots \\ f_m \end{bmatrix} = M,$$

and the degree of this relation is $n_x + 2$, hence $M \in \chi(n_x + 2, \mathbf{F})$. This can be repeated for any maximal minor of $A_{\mathbf{F}}$.

Next, Theorem 6.5 states that there exists a non-empty Zariski open subset $O \subset \mathscr{BL}^a_{\mathbb{K}}(n_x, n_y)$ such that, if $\mathbf{F} \in O$ then the homogeneous part of highest degree of all maximal minors of $A_{\mathbf{F}}$ are a linear combination of a grevlex Gröbner basis. Therefore, all monomials in $\mathbb{K}[Y]$ of degree $n_x + 1$ are in $\{\mathsf{LM}(f) \mid f \in \chi(n_x + 2, \mathbf{F})\}$. □

Lemma 6.28. *There exists a non-empty Zariski open subset $O \subset \mathscr{BL}^a_{\mathbb{K}}(n_x, n_y)^m$ such that, for any $\mathbf{F} \in O$, for any $i \in \{1, \dots, n_x\}$ and for any monomial $\mathfrak{m}$ in $\mathbb{K}[Y]$ of degree at most n_x, there exists a polynomial $h \in \mathbb{K}[Y]$ of degree at most n_x such that $x_i\mathfrak{m} - h \in \chi(n_x + 2, \mathbf{F})$.*

Proof. Let $i \in \{1, \dots, n_x\}$. First, we show that for any maximal minor $M \in \mathbb{K}[Y]$ of $\mathsf{jac}_{\mathbf{x}}(\mathbf{F})$, there exists a polynomial $h_M \in \mathbb{K}[Y]$ of degree $n_x + 1$ such that $x_iM - h_M \in \chi(n_x + 2, \mathbf{F})$. This is essentially Cramer's rule: we are searching for a vector $\mathbf{v}$ such that

$$\mathbf{v} \cdot \left(\mathsf{jac}_{\mathbf{x}}(\mathbf{F}) \cdot \begin{bmatrix} x_1 \\ \vdots \\ x_{n_x} \end{bmatrix} + \begin{bmatrix} f_1(0,\dots,0,y_1,\dots,y_{n_y}) \\ \vdots \\ f_m(0,\dots,0,y_1,\dots,y_{n_y}) \end{bmatrix} \right) = x_iM - h'.$$

Let $\ell_1, \dots, \ell_{n_x}$ be the indices of the rows of the submatrix of $\mathsf{jac}_{\mathbf{x}}(\mathbf{F})$ whose determinant is M. The following vector $\mathbf{v}$ gives the wanted relation:

$$\mathbf{v}_j = \begin{cases} 0 \text{ if } j \notin \{\ell_1, \dots, \ell_{n_x}\} \\ (-1)^k \,\mathsf{minor}(A_i, \{\ell_1, \dots, \ell_{k-1}, \ell_{k+1}, \dots, \ell_{n_x}) \text{ if } \ell_k = j \end{cases}$$

where A_i is the matrix $\mathsf{jac}_{\mathbf{x}}(\mathbf{F})$ where the ith column has been removed. Consequently, direct computations show that $\mathbf{v} \cdot \mathsf{jac}_{\mathbf{x}}(\mathbf{F}) = x_iM$ and $h'_M = \mathbf{v} \cdot \begin{bmatrix} f_1(0,\dots,0,y_1,\dots,y_{n_y}) \\ \vdots \\ f_m(0,\dots,0,y_1,\dots,y_{n_y}) \end{bmatrix} \in \mathbb{K}[Y]$ is a polynomial of degree at most $n_y + 1$.

Notice that there are exactly $\binom{n_x+n_y}{n_x}$ maximal minors of $\mathsf{jac}_{\mathbf{x}}(\mathbf{F})$, which is equal to the number of monomials of degree at most n_x in $\mathbb{K}[Y]$. By Theorem 6.5, there exists a non-empty Zariski open subset $O_1 \subset \mathscr{BL}^a_{\mathbb{K}}(n_x, n_y)^m$, such that for $\mathbf{F} \in O_1$, these minors are linearly independent over $\mathbb{K}$. Therefore, by a linear combination of the polynomials $x_iM - h'_M \in \chi(n_x + 2, \mathbf{F})$ (recall that $\chi(n_x + 2, \mathbf{F})$ is a $\mathbb{K}$-vector space), we get polynomials $x_i\mathfrak{m} - h'_{\mathfrak{m}} \in \chi(n_x + 2, \mathbf{F})$, where the polynomials $h'_{\mathfrak{m}}$ have degree at most $n_x + 1$. By Lemma 6.27, there exists a non-empty Zariski open subset $O_2 \subset \mathscr{BL}^a_{\mathbb{K}}(n_x, n_y)^m$ such that, for any $\mathbf{F} \in O$, all monomials in $\mathbb{K}[Y]$ of degree $n_x + 1$ are in the set $\{\mathsf{LM}(f) \mid f \in \chi(n_x + 2, \mathbf{F})\}$. Therefore, if $\mathbf{F} \in O_1 \cap O_2$ and by reducing the polynomials $h'_{\mathfrak{m}}$, we obtain polynomials $h_{\mathfrak{m}}$ in $\mathbb{K}[Y]$ of degree at most n_x. □

In the following proposition, we use the notation S_0 to represent the $\mathbb{K}$-vector space generated by $\mathbf{F}$, and for $i \in \mathbb{N} \setminus \{0\}$, we let S_i denote the vector space $\chi(n_x+2, S_{i-1})$. Therefore, $S_0 \subset \cdots \subset S_{n_x}$.

Proposition 6.29. *There exists a non-empty Zariski open subset $O \subset \mathscr{BL}^a_{\mathbb{K}}(n_x, n_y)^m$ such that, for any $\mathbf{F} \in O$, all monomials in $\mathbb{K}[X, Y]$ are in $\{\mathsf{LM}(f) \mid f \in S_{n_x}\}$. Moreover, the minimal reduced Gröbner basis G of $\langle F \rangle$ is contained in S_{n_x}.*

Proof. By Lemmas 6.27 and 6.28, there exists a non-empty Zariski open subset $O_1 \subset \mathscr{BL}^a_{\mathbb{K}}(n_x, n_y)^m$ such that, for any $\mathbf{F} \in O_1$, all monomials in $\mathbb{K}[Y]$ of degree $n_x + 1$ are leading monomials in S_1 and for any $i \in \{1, \ldots, n_x\}$ and for any monomial $\mathfrak{m}$ in $\mathbb{K}[Y]$ of degree at most n_x, there exists a polynomial $h \in \mathbb{K}[Y]$ of degree at most n_x such that $x_i\mathfrak{m} - h \in S_1$.

Let $\mathbf{F}$ be a polynomial system in O_1. First, we prove by induction that for any $i \in \{1, \ldots, n_x+1\}$, for any monomial $\mathfrak{m}_{\mathbf{x}} \in \mathbb{K}[X]$ of degree i and for any monomial $\mathfrak{m}_{\mathbf{y}} \in \mathbb{K}[Y]$ of degree at most $n_x - i + 1$, there exists a polynomial $h \in \mathbb{K}[Y]$ of degree at most n_x such that $\mathfrak{m}_{\mathbf{x}}\mathfrak{m}_{\mathbf{y}} - h \in S_{i-1}$.

- **Initialization.** For $i = 1$, this is a direct consequence of Lemma 6.27.
- **Induction.** Let $\mathfrak{m}_{\mathbf{x}} \in \mathbb{K}[X]$ be a monomial of degree $i \geq 2$ and $\mathfrak{m}_{\mathbf{y}} \in \mathbb{K}[Y]$ be a monomial of degree at most $n_x - i + 1$. Let $j \in \{1, \ldots, n_x\}$ such that x_j divides $\mathfrak{m}_{\mathbf{x}}$. By induction, there exists $h' \in \mathbb{K}[Y]$ of degree at most n_x such that $\frac{\mathfrak{m}_{\mathbf{x}}\mathfrak{m}_{\mathbf{y}}}{x_j} - h' \in S_{i-2}$. Consequently, by multiplying by x_j,
$$\mathfrak{m}_{\mathbf{x}}\mathfrak{m}_{\mathbf{y}} - x_j h' \in \chi(n_x + 2, S_{i-2}) = S_{i-1}.$$
By Lemma 6.28, each monomial in $x_j h'$ can be reduced to a polynomial in $\mathbb{K}[Y]$ of degree at most n_x. Therefore, there exists a polynomial $h \in \mathbb{K}[Y]$ of degree at most n_x such that $\mathfrak{m}_{\mathbf{x}}\mathfrak{m}_{\mathbf{y}} - h \in S_{i-1}$.

Applying this result with $i = n_x$, since $S_0 \subset \cdots \subset S_{n_x}$, we obtain that for any monomial $\mathfrak{m} \in \mathbb{K}[X, Y]$ of degree $n_x + 1$ there exists $h \in \mathbb{K}[Y]$ of degree at most n_x such that $\mathfrak{m} - h \in S_{n_x}$. Since the grevlex ordering is a degree ordering, $\mathsf{LM}(\mathfrak{m} - h) = \mathfrak{m}$.

It remains to prove that S_{n_x} contains a Gröbner basis of $\langle \mathbf{F} \rangle$. By the multi-homogeneous Bézout bound (Theorem 1.69), there exists a non-empty Zariski open subset $O_2 \subset \mathscr{BL}^a_{\mathbb{K}}(n_x, n_y)^m$ such that, for any $\mathbf{F} \in O_2$, $\mathsf{DEG}(\langle \mathbf{F} \rangle) = \binom{n_x+n_y}{n_x}$. Let $\mathbf{F} \in O_1 \cap O_2$ be a bilinear system, and $\mathbb{K}[X, Y]_{\leq n_x+1}$ denote the $\mathbb{K}$-vector space of all polynomials of total degree at most $n_x + 1$. Then a basis of the $\mathbb{K}$-vector space $\mathbb{K}[X, Y]_{\leq n_x+1} / (\mathsf{LM}(S_{n_x}) \cap \mathbb{K}[X, Y]_{\leq n_x+1})$ is given by all monomials which are not in $\mathsf{LM}(S_{n_x})$, namely all monomials in $\mathbb{K}[Y]$ of degree at most n_x. Consequently, $\dim (\mathbb{K}[X, Y]_{\leq n_x+1} / (\mathsf{LM}(S_{n_x}) \cap \mathbb{K}[X, Y]_{\leq n_x+1})) = \binom{n_x+n_y}{n_x} = \mathsf{DEG}(\langle \mathbf{F} \rangle)$, and hence, S_{n_x} contains the minimal reduced Gröbner basis of $\langle \mathbf{F} \rangle$. □

With the notations introduced in Section 1.4.2, a direct consequence of Proposition 6.29 is that $\mathrm{d}_{\max}(\mathbf{F}) \leq n_x + 2$.

Corollary 6.30. *The arithmetic complexity of computing a Gröbner basis of a generic bilinear system $f_1, \ldots, f_{n_x+n_y} \in \mathbb{K}[x_0, \ldots, x_{n_x-1}, y_0, \ldots, y_{n_y-1}]$ with the F_4 Algorithm is bounded by*
$$O\left(\min(n_x, n_y)(n_x + n_y)\binom{n_x + n_y + \min(n_x + 2, n_y + 2)}{\min(n_x + 2, n_y + 2)}^{\omega}\right),$$
where $2 \leq \omega \leq 3$ is the linear algebra constant.

Proof. By Proposition 6.29, when $n_x \leq n_y$, we have to compute at most n_x times the row echelon form of a submatrix of the Macaulay matrix in degree $n_x + 2$ during the F_4 algorithm to obtain a Gröbner basis. Each of these computations costs $O\left((n_x + n_y)\binom{n_x+n_y+\min(n_x+2,n_y+2)}{\min(n_x+2,n_y+2)}^{\omega}\right)$ arithmetic operations in $\mathbb{K}$. □

n_x	n_y	nb. eq.	$\mathrm{d}_{\max}$	nb. reductions to 0
2	3	5	3	0
2	4	6	3	0
3	10	13	4	0
5	8	13	6	0
6	6	12	7	0
2	7	9	4	0

Table 6.4: Experimental results: degree maximal and reductions to zero for random affine bilinear systems

Remark 6.31. *This bound on $\mathrm{d}_{\max}(\mathbf{F})$ should be compared with the degree of regularity of a regular quadratic system with n equations and n variables. The Macaulay bound (see [Laz83]) says that the degree of regularity (and $\mathrm{d}_{\max}$) of such systems is $n+1$. The complexity of computing a Gröbner basis of a generic quadratic system of n equations in $\mathbb{K}[x_1, \ldots, x_n]$ is bounded by $O\left(\left(n\binom{2n}{n+1}\right)^{\omega}\right)$, which is larger than $O\left(\min(n_x, n_y)(n_x+n_y)\binom{n_x+n_y+\min(n_x+2,n_y+2)}{\min(n_x+2,n_y+2)}^{\omega}\right)$ when $n = n_x + n_y$. Notice also that if $\min(n_x, n_y)$ is constant, then the complexity of computing a Gröbner basis of a 0-dimensional generic affine bilinear system is polynomial in the number of unknowns $n = n_x + n_y$. Moreover, the inequality $\mathrm{d}_{\max}(\mathbf{F}) \leq \min(n_x+2, n_y+2)$ is sharp but often a bit pessimistic: for most values of n_x, n_y, $\mathrm{d}_{\max}(\mathbf{F})$ is equal to $\min(n_x+1, n_y+1)$ (see Table 6.4). However, there exist sets of parameters for which this bound is reached, e.g. $n_x = 2, n_y = 7$.*

6.6 Application to bi-homogeneous systems of bi-degree $(D, 1)$

In this section, we show that the complexity analysis made in Chapter 4 for the generalized MinRank problem can be used to obtain bounds on the complexity of solving bi-homogeneous systems of bi-degree $(D, 1)$ by using Gröbner bases algorithms. Under genericity assumptions, such systems have a finite number of solutions on the biprojective space $\mathbb{P}^{n_x} \times \mathbb{P}^{n_y}$. One way to compute them is to start by computing their projection on $\mathbb{P}^{n_x}$, and then lift them to $\mathbb{P}^{n_x} \times \mathbb{P}^{n_y}$ by solving linear systems (this can be done since the equations are linear with respect the variables $y_0, \ldots, y_{n_y}$).

The following proposition shows that computing the projection on $\mathbb{P}^{n_y}$ can be computed by solving a homogeneous generalized MinRank problem.

Proposition 6.32. *Let $f_1, \ldots, f_m \in \mathbb{K}[X, Y]$ be a bi-homogeneous system of bi-degree $(D, 1)$. If $m > n_y$, then $(x_0 : \ldots : x_{n_x}, y_0 : \ldots : y_{n_y}) \in \mathbb{P}^{n_x} \times \mathbb{P}^{n_y}$ is a zero of this system if and only if the matrix*

$$\mathsf{jac}_Y(x_0, \ldots, x_{n_x}) = \begin{pmatrix} \frac{\partial f_1}{\partial y_0} & \cdots & \frac{\partial f_1}{\partial y_{n_y}} \\ \vdots & \vdots & \vdots \\ \frac{\partial f_m}{\partial y_0} & \cdots & \frac{\partial f_m}{\partial y_{n_y}} \end{pmatrix}$$

is rank defective.

Proof. First, notice that

$$\begin{pmatrix} f_1 \\ \vdots \\ f_m \end{pmatrix} = \mathsf{jac}_Y(x_0, \ldots, x_{n_x}) \cdot \begin{pmatrix} y_0 \\ \vdots \\ y_{n_y} \end{pmatrix}.$$

Therefore, $(x_0 : \ldots : x_{n_x}, y_0 : \ldots : y_{n_y}) \in \mathbb{P}^{n_x} \times \mathbb{P}^{n_y}$ is a zero of the system if and only if $(y_0, \ldots, y_{n_y})$ belongs to the kernel of jac_Y. Since $m > n_y$, the number of rows is greater than or equal to the number of columns of jac_Y, and hence jac_Y is rank defective. □

In applications, most of bi-homogeneous systems occurring are *affine*: A polynomial $f \in \mathbb{K}[x_1, \ldots, x_{n_x}, y_1, \ldots, y_{n_y}]$ is called affine of bi-degree $(D, 1)$ if there exists a bi-homogeneous polynomial $f^h \in \mathbb{K}[x_0, \ldots, x_{n_x}, y_0, \ldots, y_{n_y}]$ of bi-degree $(D, 1)$ such that

$$f(x_1, \ldots, x_{n_x}, y_1, \ldots, y_{n_y}) = f^h(1, x_1, \ldots, x_{n_x}, 1, y_1, \ldots, y_{n_y}).$$

This means that each monomial occurring in f has bi-degree (i, j) with $i \leq D$ and $j \leq 1$. Notice that the polynomial f^h is uniquely defined and that Proposition 6.32 also holds in the affine context:

Proposition 6.33. *Let $f_1, \ldots, f_m \in \mathbb{K}[x_1, \ldots, x_{n_x}, y_1, \ldots, y_{n_y}]$ be an affine system of bi-degree $(D, 1)$. If $m > n_y$ and $(x_1, \ldots, x_{n_x}, y_1, \ldots, y_{n_y}) \in \mathbb{K}^{n_x} \times \mathbb{K}^{n_y}$ is a zero of the system, then the $m \times (n_y + 1)$ matrix*

$$\mathsf{jac}_Y^a(x_1, \ldots, x_{n_x}) = \begin{pmatrix} f_1(x_1, \ldots, x_{n_x}, 0, \ldots, 0) & \frac{\partial f_1}{\partial y_1} & \ldots & \frac{\partial f_1}{\partial y_{n_y}} \\ \vdots & \vdots & \vdots & \\ f_m(x_1, \ldots, x_{n_x}, 0, \ldots, 0) & \frac{\partial f_m}{\partial y_0} & \ldots & \frac{\partial f_m}{\partial y_{n_y}} \end{pmatrix}$$

is rank defective.

Proof. The proof is similar to that of Proposition 6.32 since

$$\begin{pmatrix} f_1 \\ \vdots \\ f_m \end{pmatrix} = \mathsf{jac}_Y^a(x_1, \ldots, x_{n_x}) \cdot \begin{pmatrix} 1 \\ y_1 \\ \vdots \\ y_{n_y} \end{pmatrix}.$$

Therefore, if $(x_1, \ldots, x_{n_x}, y_1, \ldots, y_{n_y})$ is a zero of the system then there is a non-zero vector in the kernel of jac_Y^a (however in the affine case, the converse is not true). Since $m > n_y$, the number of rows is greater than or equal to the number of columns of jac_Y^a, and hence jac_Y^a is rank defective. □

An algebraic description of the variety V of a 0-dimensional polynomial system can be obtained by computing a rational parametrization, i.e. a polynomial $g(u) \in \mathbb{K}[u]$ and a set of rational functions $g_1, \ldots, g_{n_x}, h_1, \ldots, h_{n_y} \in \mathbb{K}(u)$ such that

$$\begin{gathered} (x_1, \ldots, x_{n_x}, y_1, \ldots, y_{n_y}) \in V \\ \Updownarrow \\ \exists u \in \mathbb{K}, s.t. g(u) = 0, \forall i \in \{1, \ldots, n_x\}, x_i = g_i(u), \forall j \in \{1, \ldots, n_y\}, y_j = h_j(u). \end{gathered}$$

To obtain a rational parametrization, we need a separating element: a linear form which takes different values on all points of V. Therefore, a rational parametrization exists only if the cardinality of the field $\mathbb{K}$ is 0 or large enough.

Under the assumption that the field $\mathbb{K}$ is sufficiently large, Algorithm 8 uses the property described in Proposition 6.33 to find a rational parametrization of the zeroes of a radical and 0-dimensional system of $n_x + n_y$ affine polynomials of bi-degree $(D, 1)$. The algorithm proceeds by first computing a rational parametrization of the projection of the zero set on $\mathbb{K}^{n_x}$. This is done by computing a lexicographical Gröbner basis of a Generalized MinRank Problem. Then this parametrization is lifted to the whole space by solving a linear system (this can be done since the equations are linear with respect to the variables $y_1, \dots, y_{n_y}$).

The success of Algorithm 8 depends on the choice of the parameters α (a linear change of coordinates such that x_n is a separating element) and M. However, as we will see in Theorem 6.34, if the cardinality of $\mathbb{K}$ is infinite or large enough, then almost all choices of α and M are good. Therefore, these parameters can be chosen at random. If Algorithm 8 unluckily fails, then it can be restarted with the same algebraic system and different values of α and M.

We prove now that the complexity of Algorithm 8 is bounded by the complexity of the underlying generalized MinRank problem and that most choices of $(\alpha_1, \dots, \alpha_{n_x-1})$ and M do not fail.

Theorem 6.34. *Let $f_1, \dots, f_{n_x+n_y} \in \mathbb{K}[X, Y]$ be an affine system of bi-degree $(D, 1)$ such that the ideal $\langle f_1, \dots, f_{n_x+n_y}\rangle$ is radical and 0-dimensional. Then there exist non-identically null polynomials $h_1 \in \mathbb{K}[z_1, \dots, z_{n_x-1}]$ and $h_2 \in \mathbb{K}[z_1, \dots, z_{n_y,n_x+n_y}]$ such that, for any choice of $(\alpha_1, \dots, \alpha_{n_x-1})$ and $M = (m_{i,j}) \in \mathbb{K}^{n_y\times(n_x+n_y)}$ verifying:*

- *the matrix $\mathsf{jac}^{a}_{Y}(\widetilde{f_1}, \dots, \widetilde{f_{n_x+n_y}})$ verifies the conditions of Theorem 4.24;*
- *$h_1(\alpha_1, \dots, \alpha_{n_x-1})h_2(m_{1,1}, \dots, m_{n_y,n_x+n_y}) \neq 0$,*

Algorithm 8 returns a rational parametrization of the solutions of the system and its complexity is bounded by

$$O\left(\binom{n_x+n_y}{n_x-1}\binom{D(n_x+n_y)+1}{n_x}^{\omega} + n_x\left(D^{n_x}\binom{n_x+n_y}{n_x}\right)^3\right).$$

Proof. Let I denote the ideal generated by $f_1, \dots, f_{n_x+n_y}$. According to [Lak90, BMMT94], for any radical 0-dimensional ideal, there exists a polynomial h_1 such that if $h_1(\alpha_1, \dots, \alpha_{n_x-1}) \neq 0$, then the system is in shape position after the change of coordinates

$$x_{n_x} \mapsto x_{n_x} - \sum_{\ell=1}^{n_x-1} \alpha_\ell x_\ell.$$

The polynomial h_2 is chosen such that if $h_2(m_{i,j}) \neq 0$, then the linear system $\widehat{f_1} = \dots = \widehat{f_{n_y}} = 0$ in $\mathbb{K}(u)[Y]$ has rank exactly n_y. Therefore h_2 can be chosen as a nonzero coefficient in $\mathbb{K}[z_{1,1}, \dots, z_{n_y,n_x+n_y}]$ of a term u^β in the determinant of the following linear system in $\mathbb{K}[z_{1,1}, \dots, z_{n_y,n_x+n_y}, u][Y]$ (where the variables are $y_1, \dots, y_{n_y}$):

$$\begin{pmatrix} z_{1,1} & \dots & z_{1,n_x+n_y} \\ \vdots & \vdots & \vdots \\ z_{n_y,1} & \dots & z_{n_y,n_x+n_y} \end{pmatrix} \cdot \begin{pmatrix} \widetilde{f_1}(g_1(u), \dots, g_{n_x-1}(u), u, y_1, \dots, y_{n_y}) \mod g(u) \\ \vdots \\ \widetilde{f_{n_x+n_y}}(g_1(u), \dots, g_{n_x-1}(u), u, y_1, \dots, y_{n_y}) \mod g(u) \end{pmatrix} = 0.$$

Since the ideal generated by the input system $(f_1, \dots, f_{n_x+n_y})$ is 0-dimensional and proper, it has finitely-many solutions, this determinant (which lies in $\mathbb{K}[z_{1,1}, \dots, z_{n_y,n_x+n_y}, u]$) is not zero. Therefore the evaluation of this determinant is not null if and only if h_2 does not vanish.

- The complexity of the substitution for computing the polynomials $\widetilde{f_i}$ is bounded by $\widetilde{O}((n_x + n_y)Dn_xn_y)$.

Algorithm 8 Rational parametrization of systems of bi-degree $(D, 1)$

Input: $f_1, \ldots, f_{n_x+n_y} \in \mathbb{K}[X, Y]$ a system of affine polynomials of bi-degree $(D, 1)$ such that the ideal they generate is radical and 0-dimensional;
$(\alpha_1, \ldots, \alpha_{n_x-1}) \in \mathbb{K}^{n_x-1}$;
a full rank matrix $M = (m_{i,j}) \in \mathbb{K}^{n_y \times (n_x+n_y)}$.

Output: Returns a rational parametrization of the variety of the system or "fail".

1: Compute for each $i \in \{1, \ldots, n_x + n_y\}$,

$$\widetilde{f_i}(x_1, \ldots, x_{n_x-1}, u, y_1, \ldots, y_{n_y}) = f_i(x_1, \ldots, x_{n_x-1}, u - \sum_{\ell=1}^{n_x-1} \alpha_\ell x_\ell, y_1, \ldots, y_{n_y}).$$

2: Compute the matrix $\mathsf{jac}_Y^a(\widetilde{f_1}, \ldots, \widetilde{f_{n_x+n_y}})$.

3: Compute a lex Gröbner basis G of the ideal $I \subset \mathbb{K}[x_1, \ldots, x_{n_x-1}, u]$ generated by the maximal minors of the matrix $\mathsf{jac}_Y^a(\widetilde{f_1}, \ldots, \widetilde{f_{n_x+n_y}})$. If the Gröbner basis has the following shape (the *shape position*):

$$\begin{array}{r} x_1 - g_1(u) \\ x_2 - g_2(u) \\ \vdots \\ x_{n_x-1} - g_{n_x-1}(u) \\ g(u), \end{array}$$

then continue to Step 4, else return "fail".

4: Using M, compute a linear combination of the polynomials of the system evaluated at $(g_1(u), \ldots, g_{n_x-1}(u))$:

$$\begin{pmatrix} \widehat{f_1}(y_1, \ldots, y_{n_y}, u) \\ \vdots \\ \widehat{f_{n_y}}(y_1, \ldots, y_{n_y}, u) \end{pmatrix} = M \cdot \begin{pmatrix} \widetilde{f_1}(g_1(u), \ldots, g_{n_x-1}(u), u, y_1, \ldots, y_{n_y}) \mod g(u) \\ \vdots \\ \widetilde{f_{n_x+n_y}}(g_1(u), \ldots, g_{n_x-1}(u), u, y_1, \ldots, y_{n_y}) \mod g(u) \end{pmatrix}$$

5: If the linear system $\widehat{f_1} = \ldots = \widehat{f_{n_y}} = 0$ has rank n_y (as a linear system in $\mathbb{K}(u)[Y]$ where the variables are $y_1, \ldots, y_{n_y}$), continue to Step 6, else return "fail".

6: Using Cramer's rule, solve the system $\widehat{f_1} = \ldots = \widehat{f_{n_y}} = 0$ as a linear system in $\mathbb{K}(u)[Y]$. This yields rational functions $h_i(u) \in \mathbb{K}(u)$ such that, for $i \in \{1, \ldots, n_y\}$, $y_i - h_i(u) = 0$.

7: Return the rational parametrization

$$\begin{array}{rlrl} g(u) &= 0 & & \\ x_1 &= g_1(u) & y_1 &= h_1(u) \\ &\vdots & &\vdots \\ x_{n_x-1} &= g_{n_x-1}(u) & y_{n_y-1} &= h_{n_y-1}(u) \\ x_{n_x} &= u - \sum_{\ell=1}^{n_x-1} \alpha_\ell g_\ell(u) & y_{n_y} &= h_{n_y}(u) \end{array}$$

- By Theorem 4.24, the complexity of the Gröbner basis computation is bounded by

$$O\left(\binom{n_x+n_y}{n_x-1}\binom{D(n_x+n_y)+1}{n_x}^{\omega}+n_x\left(\mathsf{DEG}(I)\right)^3\right).$$

- Since $\deg(g_{n_x}) \leq \mathsf{DEG}(I)$, a monomial $u^{n_x}\prod_{i=1}^{n_x-1}x_i^{\alpha_i}$ of degree D can be evaluated in the univariate polynomials $(g_1(u),\ldots,g_{n_x-1}(u))$ modulo $g(u)$ in complexity $\widetilde{O}(D\,\mathsf{DEG}(I))$ by using a subproduct tree [BS05], quasi-linear multiplication of univariate polynomials and quasi-linear modular reduction. Since there are at most $(n_x+n_y)(n_y+1)\binom{n_x+D}{n_x}$ such monomials in the system $f_1,\ldots,f_{n_x+n_y}$, the Step 4 of the algorithm needs at most $\widetilde{O}\left((n_x+n_y)n_y\binom{n_x+D}{n_x}D\,\mathsf{DEG}(I)\right)$ arithmetic operations in $\mathbb{K}$.

  Notice that $n_x+n_y\leq\binom{n_x+n_y}{n_x-1}$ and $\mathsf{DEG}(I)\leq\binom{D(n_x+n_y)+1}{n_x}$.

  - If $D\geq 2$: for any $a,b,c\in\mathbb{N}$ such that $b<a$, we have $\binom{a}{b}c\leq\binom{a+c}{b}$. Therefore, $Dn_y\binom{n_x+D}{n_x}\leq\binom{n_x+n_y+2D}{n_x}$. Also, notice that for $D\geq 2$ and for any n_x,n_y such that $n_xn_y>1$, we obtain $n_x+n_y+2D\leq D(n_x+n_y)+1$. Consequently,

  $$\widetilde{O}\left((n_x+n_y)n_y\binom{n_x+D}{n_x}D\,\mathsf{DEG}(I)\right)\leq\widetilde{O}\left(\binom{n_x+n_y}{n_x-1}\binom{D(n_x+n_y)+1}{n_x}^2\right).$$

  - If $D=1$: $(n_x+n_y)n_y\binom{n_x+1}{n_x}=(n_x+n_y)n_yn_x$ is bounded by $\binom{n_x+n_y}{n_x-1}\binom{(n_x+n_y)+1}{n_x}$.

  Therefore, the complexity of the Step 4 of Algorithm 8 is bounded above by the complexity of the Gröbner basis computation: $O\left(\binom{n_x+n_y}{n_x-1}\binom{D(n_x+n_y)+1}{n_x}^{\omega}\right)$.

- To solve the linear system by using Cramer's rule, we need to compute n_x+1 determinants of $(n_x\times n_x)$-matrices whose entries are univariate polynomials of degree D. This can be achieved by using a fast evaluation-interpolation strategy with complexity $\widetilde{O}\left(Dn_x^{\omega+1}\right)$ (since multi-set evaluation and interpolation of univariate polynomials can be done in quasi-linear time, see e.g. [BS05]).

Since $\mathsf{DEG}(I)$ is bounded by $D^{n_x}\binom{n_x+n_y}{n_x}$, the sum of all these complexities is bounded by

$$O\left(\binom{n_x+n_y}{n_x-1}\binom{D(n_x+n_y)+1}{n_x}^{\omega}+n_x\left(D^{n_x}\binom{n_x+n_y}{n_x}\right)^3\right).$$

$\square$

Remark 6.35. *According to Corollary 3.16 and Lemma 3.14, if $D=1$, there exists a non-empty Zariski open subset O_1 of the set of systems of bi-degree $(1,1)$ which are 0-dimensional and radical. The proofs are similar for systems of bi-degree $(D,1)$ with $D\in\mathbb{N}$.*

Chapter 7

Boolean Systems

The results presented in this chapter are joint work with M. Bardet, J.-C. Faugère and B. Salvy and are accepted for publication in Journal of Complexity [BFSS12].

A fundamental problem in computer science is to find all the common zeroes of m quadratic polynomials in n unknowns over $\mathbb{F}_2$. The cryptanalysis of several modern ciphers reduces to this problem. Up to now, the best complexity bound was reached by an exhaustive search in $4\log_2 n\, 2^n$ operations. We give an algorithm that reduces the problem to a combination of exhaustive search and sparse linear algebra. This algorithm has several variants depending on the method used for the linear algebra step. Under precise algebraic assumptions, we show that the deterministic variant of our algorithm has complexity bounded by $O(2^{0.841n})$ when $m = n$, while a probabilistic variant of the Las Vegas type has expected complexity $O(2^{0.792n})$. Experiments on random systems show that the algebraic assumptions are satisfied with probability very close to 1. We also give a rough estimate for the actual threshold between our method and exhaustive search, which is as low as 200, and thus very relevant for cryptographic applications.

7.1 Introduction

Motivation and Problem Statement. Solving multivariate quadratic polynomial systems is a fundamental problem in Information Theory. Moreover, *random* instances seem difficult to solve. Consequently, the security of several multivariate cryptosystems relies on its hardness, either directly (e.g., HFE [Pat96], UOV [KPG99],. . .) or indirectly (e.g., McEliece [FOPT10]). In some cases, systems of special types have to be solved, but recent proposals like the new Polly Cracker type cryptosystem [AFFP11] rely on the hardness of solving *random* systems of equations. This motivates the study of the complexity of generic polynomial systems. A particularly important case for applications in cryptology is the Boolean case; in that case both the coefficients and the solutions of the system are over GF_2. The main problem to be solved is the following:

> **The Boolean Multivariate Quadratic Polynomial Problem (Boolean MQ)**
> **Input:** $(f_1, \dots, f_m) \in \mathsf{GF}_2[x_1, \dots, x_n]^m$ with $\deg(f_i) = 2$ for $i = 1, \dots, m$.
> **Question:** Find – if any – *all* $z \in \mathsf{GF}_2^n$ such that $f_1(z) = \cdots = f_m(z) = 0$.

Another related problem stems from the fact that in many cryptographic applications, it is sufficient to find *at least one* solution of the corresponding polynomial system (in that case a solution is the original clear message or is related to the secret key). For instance, the stream cipher QUAD [BGP06, BGP09] relies on the iteration of a set of multivariate quadratic polynomials over GF_2 so that the security of the keystream generation is related to the difficulty of finding at least one solution of the Boolean MQ problem. Thus, we also consider the following variant of the Boolean MQ problem:

The Boolean Multivariate Quadratic Polynomial Satisfiability Problem (Boolean MQ SAT)
Input: $(f_1, \dots, f_m) \in \mathsf{GF}_2[x_1, \dots, x_n]^m$ with $\deg(f_i) = 2$ for $i = 1, \dots, m$.
Question: Find – if any – *one* $z \in \mathsf{GF}_2^n$ such that $f_1(z) = \cdots = f_m(z) = 0$.

Testing for the existence of a solution is an NP-complete problem (it is plainly in NP and 3-SAT can be reduced to it [FY79]). Clearly, the Boolean MQ problem is at least as hard as Boolean MQ SAT, while an exponential complexity is achieved by exhaustive search.

Throughout this chapter, *random* means distributed according to the uniform distribution (given m and n, a random quadratic polynomial is uniformly distributed if all its coefficients are independently and uniformly distributed over GF_2). The relation between the difficulties of Boolean MQ and Boolean MQ SAT depends on the relative values of m and n. When $m > n$, the number of solutions of the algebraic system is 0 or 1 with large probability and thus finding one or all solutions is very similar, while when $m = n$, the probability that a random system has at least one solution over GF_2 tends to $1 - \frac{1}{e} \approx 0.63$ for large n [FB07]. Hence if we have to find at least one solution of a system with $m < n$ equations in n variables it is enough to specialize $n - m$ variables randomly in GF_2; the resulting system has at least one solution with limit probability 0.63 and is easier to solve (since the number of equations and variables is only m). Consequently, in the remainder of this article we restrict ourselves to the case $m \geq n$.

To the best of our knowledge, in the *worst case*, the best complexity bound to solve the Boolean MQ problem is obtained by a modified exhaustive search in $4\log_2(n)2^n$ operations [BCC⁺10]. Being able to decrease significantly this complexity is a long-standing open problem and is the main goal of this article. It is crucial for practical applications to have estimates of the asymptotic complexity: it is especially important in the cryptographic context where this value may have a strong impact on the sizes of the keys needed to reach a given level of security.

Main results. We describe a new algorithm BooleanSolve that solves Boolean MQ for determined or overdetermined systems ($m = \alpha n$ with $\alpha \geq 1$). We show how to adapt it to solve the Boolean MQ SAT problem. This algorithm has deterministic and Las Vegas variants, depending on the choice of some linear algebra subroutines. Our main result is:

Theorem 7.1. *The Boolean MQ Problem is solved by Algorithm* BooleanSolve*. If $m = n$ and the system fulfills algebraic assumptions detailed in Theorem 7.20, then this algorithm uses a number of arithmetic operations in GF_2 that is:*

- $O(2^{0.841n})$ *using the deterministic variant;*
- *of expectation $O(2^{0.792n})$ using the* Las Vegas *probabilistic variant.*

Recall that for a probabilistic algorithm of the Las Vegas type, the result is always correct, but the complexity is a random variable. Here its expectation is controlled well. Actually, the expectation of the complexity of the probabilistic variant behaves as the complexity of the deterministic algorithm where linear algebra would be performed in quadratic time.

Outline. Our algorithm is a variant of the hybrid approach by [BFP09, BFP12b]: we specialize the last k variables to all possible values, and check the consistency of the specialized overdetermined systems $(\tilde{f}_1, \dots, \tilde{f}_m)$ in the remaining variables $x_1, \dots, x_{n-k}$.

This consistency check is done by searching for polynomials $h_1, \dots, h_{m+n-k}$ in $x_1, \dots, x_{n-k}$ such that

$$h_1\tilde{f}_1 + \cdots + h_m\tilde{f}_m + h_{m+1}x_1(1 - x_1) + \cdots + h_{m+n-k}x_{n-k}(1 - x_{n-k}) = 1. \qquad (7.1)$$

If such polynomials exist then obviously the system is not consistent. Given a bound d on the degrees of the polynomials $h_i\tilde{f}_i$ and $h_{m+i}x_i(1-x_i)$, the existence of the h_i can be checked by linear algebra. The corresponding matrix is known as the Macaulay matrix in degree d. It is a matrix whose rows contain the coefficients of the polynomials $\tilde{f}_i$ and $x_i(1-x_i)$ multiplied by all monomials of degree at most $d-2$, each column corresponding to a monomial of degree at most d. Taking into account the special shape of the polynomials $x_i(1-x_i)$ leads to a more compact variant that we call the boolean Macaulay matrix (see Section 7.2).

When linear algebra on the Macaulay matrix in degree d produces a solution of Equation (7.1), the corresponding h_i's give a certificate of inconsistency. Otherwise, our algorithm proceeds with an exhaustive search in the remaining variables. In summary, our algorithm is a partial exhaustive search where the Macaulay matrices permit to prune branches of the search tree. The correctness of the algorithm is clear.

The key point making the algorithm efficient is the choice of k and d. If d is large, then the cost of the linear algebra stage becomes high. If d is small, the matrices are small, but many branches with no solutions are not pruned and require an exhaustive search. This is where we use the relation between the Macaulay matrix and Gröbner bases. We define a *witness degree* $\mathrm{d_{wit}}$, which has the property that any polynomial in a minimal Gröbner basis of the system is obtained as a linear combination of the rows of the Macaulay matrix in degree $\mathrm{d_{wit}}$. Hilbert's Nullstellensatz states that the system has no solution *if and only if* 1 belongs to the ideal generated by the polynomials, which implies that 1 is a linear combination of the rows of the Macaulay matrix in degree $\mathrm{d_{wit}}$, making $\mathrm{d_{wit}}$ an upper bound for the choice of d in Equation (7.1).

Our complexity estimates rely on a good control of the witness degree. For a homogeneous polynomial ideal, the classical Hilbert function of the degree d is the dimension of the vector space obtained as the quotient of the polynomials of degree d by the polynomials of degree d in the ideal. The witness degree is bounded by the first degree where the Hilbert function of the ideal generated by the homogenized equations is 0 (i.e. it is bounded by the *degree of regularity* of the homogenized system). Under the algebraic assumption of boolean semi-regularity (see Definition 7.13), we obtain an explicit expression for the generating series of the Hilbert function, known as the Hilbert series of the ideal. From there, in Proposition 7.16, using the saddle-point method as in [BFS04, BFSY04, Bar04], we show that when $m = \alpha n$ and $n \to \infty$, the witness degree behaves like $\mathrm{d_{wit}} \leq c_\alpha n$ for a constant c_α that we determine explicitly. Informally, boolean semi-regularity amounts to demanding a "sufficient" independence of the equations. In the case of infinite fields, a classical conjecture by [Fro85] states that generic systems are semi-regular. In our context where the field is GF_2, we give strong experimental evidence (Section 7.4.1) that for n sufficiently large, boolean semi-regularity holds with probability very close to 1 for random systems. Thus, our complexity estimates for boolean semi-regular systems apply to a large class of systems in practice.

Once the witness degree is controlled, the size of the Macaulay matrix depends only on the choice of k and the optimal choice depends on the complexity of the linear algebra stage. In the Las Vegas version of Algorithm BooleanSolve, we exploit the sparsity of this matrix by using a variant of Wiedemann's algorithm [GLS98] (following [Wie86, KS91, Vil97]) for solving singular linear systems. One of the main feature of the algorithm in [GLS98] is that it yields a certificate of inconsistency when the linear system has no solution. In the deterministic version, we do not know of efficient ways to take advantage of the sparsity of the matrix, whence a slightly higher complexity bound. We can then draw conclusions and obtain a complexity estimate of the algorithm depending on k/n and n (Proposition 7.17). The optimal value for k is $\simeq 0.45\,n$ in the Las Vegas setting and $\simeq 0.59\,n$ in the deterministic variant, completing the proof of our main theorem.

The complexity analysis is especially important for practical applications in multivariate Cryptology based on the Boolean MQ problem, since it shows that in order to reach a security of 2^s (with s large), one has to construct systems of boolean quadratic equations with at least $s/0.7911 \simeq 1.264s$

variables.

Related works. Due to its practical importance, many algorithms have been designed to solve the MQ problem in a wide range of contexts. First, generic techniques for solving polynomial systems can be used. In particular, Gröbner basis algorithms (such as Buchberger's algorithm [Buc65], F_4 [Fau99], F_5 [Fau02], and FGLM [FGLM93]) are well suited for this task. For instance, the F_5 algorithm has broken several challenges of the HFE public-key cryptosystem [FJ03]. In the cryptanalysis context, the XL algorithm [KS99] (which can be seen as a variant of Gröbner basis algorithms [AFI$^+$04]) has given rise to a large family of variants. All these techniques are closely related to the Macaulay matrix, introduced by [Mac02] as a tool for elimination. In order to reduce the cost of linear algebra for the efficient computation of the resultant of multivariate polynomial systems, the idea of using Wiedemann's algorithm on the Macaulay matrix has been proposed by [CKY89]; however since the specificities of the Boolean case are not taken into account, the complexity of applying [CKY89] to quadratic equations is $O(2^{4n})$.

[YC04] propose a heuristic analysis of the FXL algorithm leading them to an upper bound $O(2^{0.875n})$ for the complexity of solving the MQ problem over GF_2. In particular, they give an explicit formula for the Hilbert series of the ideal generated by the polynomials. However, the exact assumptions that have to be verified by the input systems are unclear. Also, similar results have been announced in [YCC04, Section 2.2], but the analysis there relies on algorithmic assumptions (e.g., row echelon form of sparse matrices in quadratic complexity) that are not known to hold currently. Under these assumptions, the authors show that the best trade-off between exhaustive search and row echelon form computations in the FXL algorithm is obtained by specializing $0.45n$ variables. This is the same value we obtain and prove with our algorithm. Also, a limiting behavior of the cost of the hybrid approach is obtained in [BFP12b] when the size of the finite field is big enough; these results are not applicable over GF_2.

Other algorithms have been proposed when the system has additional structural properties. In particular, the Boolean MQ problem also arises in satisfiability problems, since boolean quadratic polynomials can be used for representing constraints. In these contexts, the systems are sparse and for such systems of higher degree the 2^n barrier has been broken [Sem08, Sem09]; similar results also exist for the k-SAT problem. Our algorithm does not exploit the extra structure induced by this type of sparsity and thus does not improve upon those results.

Organization of the chapter. The main algorithm and the algebraic tools that are used throughout the article are described in Section 7.2. Then a complexity analysis is performed in Section 7.3 by studying the asymptotic behavior of the witness degree and the sizes of the Macaulay matrices involved, under algebraic assumptions. In Section 7.4, we provide a conjecture and strong experimental evidence that these algebraic assumptions are verified with probability close to 1 for n sufficiently large. Finally, in Section 7.5 we propose an extension of the main algorithm that improves the quality of the linear filtering when n is small.

7.2 Algorithm

Notations. Let m and n be two positive integers and let R be the ring $\mathsf{GF}_2[x_1, \ldots, x_n]$. In the following, the notation $\mathsf{Monomials}(d)$ stands for the set of monomials in R of degree at most d.

Since we are looking for solutions of the system in GF_2 (and not in its algebraic closure), we have to take into account the relations $x_i^2 - x_i = 0$. Therefore, we consider the application φ mapping a monomial to its square-free part ($\varphi(\prod_{i=1}^n x_i^{a_i}) = \prod_{i=1}^n x_i^{\min(a_i,1)}$) and extended to R by linearity.

If $(f_1, \ldots, f_m) \in \mathsf{GF}_2[x_1, \ldots, x_n]^m$ is a system of polynomials, its homogenization is denoted by $(f_1^{(h)}, \ldots, f_m^{(h)}) \in \mathsf{GF}_2[x_1, \ldots, x_n, h]$ and is defined by

$$f_i^{(h)}(x_1, \ldots, x_n, h) = h^{\deg(f_i)} f_i\left(\frac{x_1}{h}, \ldots, \frac{x_n}{h}\right).$$

In the sequel, we consider the classical *grevlex* monomial ordering (**gr**aded **rev**erse **lex**icographical), as defined for instance in [CLO97, §2.2, Defn. 6]. Also, if f is a polynomial, $\mathsf{LM}(f)$ denotes its leading monomial for that order. If I is an ideal, then $\mathsf{LM}(I)$ denotes the ideal generated by the leading monomials of all polynomials in I.

7.2.1 Macaulay matrix

Definition 7.2. *Let $(f_1, \ldots, f_m)$ be polynomials in R. The* boolean Macaulay matrix *in degree d (denoted by $\mathsf{Macaulay}(d)$) is the matrix whose rows contain the coefficients of the polynomials $\{\varphi(tf_j)\}$ where $1 \le j \le m$, t is a squarefree monomial, and $\deg(tf_j) = d$. The columns correspond to the squarefree monomials in R of degree at most d and are ordered in descending order with respect to the grevlex ordering. The element in the row corresponding to $\varphi(tf_j)$ and the column corresponding to the monomial m is the coefficient of m in the polynomial $\varphi(tf_j)$.*

Note that the boolean Macaulay matrix can be obtained as a submatrix of the classical Macaulay matrix of the system $\langle f_1, \ldots, f_m, x_1^2 - x_1, \ldots, x_n^2 - x_n\rangle$ after Gaussian reduction by the rows corresponding to the polynomials $(x_1^2 - x_1, \ldots, x_n^2 - x_n)$.

Lemma 7.3. *Let M be the $r_{\mathsf{Mac}} \times c_{\mathsf{Mac}}$ boolean Macaulay matrix of the system $(f_1, \ldots, f_m)$ in degree d. Let $\mathbf{r}$ denote the $1 \times c_{\mathsf{Mac}}$ vector $\mathbf{r} = (0, \ldots, 0, 1)$. If the linear system $\mathbf{u} \cdot \mathsf{M} = \mathbf{r}$ has a solution, then the system $f_1 = \cdots = f_m = 0$ has no solution in GF_2^n.*

Proof. If the system $\mathbf{u} \cdot \mathsf{M} = \mathbf{r}$ has a solution, then there exists a linear combination of the rows of the Macaulay matrix which yields the constant polynomial 1. Therefore, $1 \in \langle f_1, \ldots, f_m, x_1^2 - x_1, \ldots, x_n^2 - x_n\rangle$. □

7.2.2 Witness degree

We consider an indicator of the complexity of affine polynomial systems: the *witness degree*. It has the property that a Gröbner basis of the ideal generated by the polynomials can be obtained by performing linear algebra on the Macaulay matrix in this degree. In particular, if the system has no solution, then the witness degree is closely related to the classical *effective Nullstellensatz* (see e.g., [Jel05]).

Definition 7.4. *Let $\mathbf{F} = (f_1, \ldots, f_m, x_1^2 - x_1, \ldots, x_n^2 - x_n)$ be a sequence of polynomials and $I = \langle \mathbf{F}\rangle$ the ideal it generates. Denote by $I_{\le d}$ and by $J_{\le d}$ the GF_2-vector spaces defined by*

$$\begin{array}{rcll} I_{\le d} & = & \{p \mid & p \in I, \deg(p) \le d\}, \\ J_{\le d} & = & \{p \mid & \exists h_1, \ldots, h_{m+n}, \forall i \in \{1, \ldots, m+n\}, \deg(h_i) \le d-2, \\ & & & p = \sum_{i=1}^m h_i f_i + \sum_{j=1}^n h_{m+j}(x_j^2 - x_j)\}. \end{array}$$

We call witness degree $(\mathrm{d_{wit}})$ *of $\mathbf{F}$ the smallest integer d_0 such that $I_{\le d_0} = J_{\le d_0}$ and $\langle\{\mathsf{LM}(f) \mid f \in I_{\le d_0}\}\rangle = \mathsf{LM}(I)$.*

Consider a row echelon form of the boolean Macaulay matrix in degree d of the system $(f_1, \ldots, f_m)$ of polynomials. Then the first nonzero element of each row corresponds to a leading monomial of an element of I, belonging to $\mathsf{LM}(I)$. For large enough d, Dickson's lemma [CLO97, §2.4, Thm. 5] implies that the collection of those monomials up to degree d generates $\mathsf{LM}(I)$ and thus the polynomials corresponding to those rows together with $\{x_1^2 - x_1, \ldots, x_n^2 - x_n\}$ form a Gröbner basis of I with respect to the grevlex ordering. Another interpretation of the *witness degree* is that it is precisely the smallest such d. Also, the witness degree is bounded from above by the degree of regularity of the corresponding homogenized system (see Proposition 7.10 below).

7.2.3 Algorithm

Algorithm 9 BooleanSolve

Input: $m, n, k \in \mathbb{N}$ such that $m \geq n > k$ and $f_1, \ldots, f_m$ quadratic polynomials in $\mathsf{GF}_2[x_1, \ldots, x_n]$.
Output: The set of boolean solutions of the system $f_1 = \cdots = f_m = 0$.

1: $S := \emptyset$.
2: $d_0 :=$ index of the first nonpositive coefficient in the series expansion at 0 of the rational function $\frac{(1+t)^{n-k}}{(1-t)(1+t^2)^m}$.
3: **for all** $(a_{n-k+1}, \ldots, a_n) \in \mathsf{GF}_2^k$ **do**
4:     **for** i from 1 to m **do**
5:         $\tilde{f}_i(x_1, \ldots, x_{n-k}) := f_i(x_1, \ldots, x_{n-k}, a_{n-k+1}, \ldots, a_n) \in \mathsf{GF}_2[x_1, \ldots, x_{n-k}]$.
6:     **end for**
7:     $\mathsf{M} :=$ boolean Macaulay matrix of $(\tilde{f}_1, \ldots, \tilde{f}_m)$ in degree d_0.
8:     **if** the system $\mathbf{u} \cdot \mathsf{M} = \mathbf{r}$ is inconsistent **then** $\triangleright$ $\mathbf{r}$ as defined in Lemma 7.3
9:         $T :=$ solutions of the system $(\tilde{f}_1 = \cdots = \tilde{f}_m = 0)$ (exhaustive search).
10:         **for all** $(t_1, \ldots, t_{n-k}) \in T$ **do**
11:             $S := S \cup \{(t_1, \ldots, t_{n-k}, a_{n-k+1}, \ldots, a_n)\}$.
12:         **end for**
13:     **end if**
14: **end for**
15: **return** S.

Our algorithm is given in Algorithm 9. The general principle is to perform an exhaustive search in two steps, using a test of consistency of the Macaulay matrix to prune most of the branches of the second step of the search.

When the system $\mathbf{u} \cdot \mathsf{M} = \mathbf{r}$ is consistent, the corresponding branch of the searching tree is not explored. In that case, by Lemma 7.3, any solution of the linear system $\mathbf{u} \cdot \mathsf{M} = \mathbf{r}$ can be used as a certificate that there exists no solution of the polynomial system $f_1 = \cdots = f_m = 0$ in this branch.

Proposition 7.5. *Algorithm* BooleanSolve *is correct and solves the Boolean MQ problem.*

Proof. By Lemma 7.3, if the test in line 8 finds the linear system to be consistent, then there can be no solution with the given values of $(a_{n-k+1}, \ldots, a_n)$. Otherwise, the exhaustive search proceeds and cannot miss a solution. It is important to note that the choice of the actual value d_0 does not have any impact on the correctness of the algorithm. $\square$

Algorithm BooleanSolve is easily be adapted to solve the Boolean MQ SAT problem by replacing lines 9-12 of the previous algorithm by:

9: $T :=$ at least one solution of the system $(\tilde{f}_1 = \cdots = \tilde{f}_m = 0)$ (modified exhaustive search).
10: **if** $T \neq \emptyset$ **then**
11:     **return** $\{(t_1, \ldots, t_{n-k}, a_{n-k+1}, \ldots, a_n) \mid (t_1, \ldots, t_{n-k}) \in T\}$
12: **end if**

7.2.4 Testing Consistency of Sparse Linear Systems

The choice of the algorithm to test whether the sparse linear system $\mathbf{u} \cdot \mathsf{M} = \mathbf{r}$ is consistent or not is crucial for the efficiency of Algorithm BooleanSolve. A simple deterministic algorithm consists in computing a row echelon form of the matrix: the linear system is consistent if and only if the last nonzero row of the row echelon form is equal to the vector $\mathbf{r}$. We show in Section 7.3 that this is sufficient to pass below the 2^n complexity barrier. We recall for future use the complexity of this method.

Proposition 7.6 ([Sto00], Proposition 2.11)**.** *The row echelon form of an $N \times M$ matrix over a field k can be computed in $O(NMr^{\omega-2})$ arithmetic operations in k, where r is the rank of the matrix and $\omega \leq 3$ is such that any two $n \times n$ matrices over k can be multiplied in $O(n^{\omega})$ arithmetic operations in k.*

Here, $\omega = 3$ is the cost of classical matrix multiplication and in this case a simple Gaussian reduction to row echelon form is sufficient. The best known value for ω has been 2.376 for a long time, by a result of [CW90]. Recent improvements of that method by [Sto10, Vas11] have decreased it to 2.3727, but this does not have a significant impact on our analysis.

This result does not exploit the sparsity of Macaulay matrices. We do not know of an efficient deterministic algorithm for row reduction that exploits this sparsity. Instead, we use an efficient Las Vegas variant of Wiedemann's algorithm due to [GLS98], whose specification is summarized in Algorithm TestConsistency. In this algorithm, the matrix A is given by two black boxes performing the operations $x \mapsto Ax$ and $u \mapsto A^t u$. The complexity is expressed in terms of the number of evaluations of these black boxes, which in our context will each have a cost bounded by the number of nonzero coefficients of Macaulay matrices. The algorithm is presented in [GLS98] for matrices with entries in an arbitrary field. We specialize it here in the case where the field is GF_2. The key ideas are a preconditioning of the matrix by multiplying it by random Toeplitz matrices and working in a suitable field extension to get access to sufficiently many points for picking random elements.

Algorithm 10 TestConsistency [GLS98]

Input:
- A black box for $\mathbf{x} \mapsto \mathsf{A} \cdot \mathbf{x}$, where $\mathsf{A} \in \mathbb{K}^{N \times N}$.
- A black box for $\mathbf{u} \mapsto \mathsf{A}^t \cdot \mathbf{u}$.
- $\mathbf{b} \in \mathbb{K}^{N \times 1}$.

Output:
- ("consistent",$\mathbf{x}$) with $\mathsf{A} \cdot \mathbf{x} = \mathbf{b}$ if the system has a solution
- ("inconsistent",$\mathbf{u}$) if the system does not have a solution, with $\mathbf{u}^t \cdot \mathsf{A} = 0$ and $\mathbf{u}^t \cdot \mathbf{b} \neq 0$, certifying the inconsistency.

Proposition 7.7 ([GLS98])**.** *Algorithm 10 determines the consistency of an $N \times N$ matrix with expected complexity $O(N \log N)$ evaluations of the black boxes and $O(N^2 \log^2 N \log\log N)$ additional operations in GF_2.*

Macaulay matrices are rectangular. We therefore first make them square by padding with zeroes. The complexity estimate is then used with N the maximum of the row and column dimensions of the matrices.

7.3 Complexity Analysis

Algorithm BooleanSolve deals with a large number of Macaulay matrices in degree d_0. We first obtain bounds on the row and column dimensions of Macaulay matrices, as well as their number of nonzero entries, in terms of the degree. We then bound the witness degree by d_0. The complexity analysis is concluded by optimizing the value of the ratio k/n that governs the number of variables evaluated in the first exhaustive search.

7.3.1 Sizes of Macaulay Matrices

Proposition 7.8. *Let $(f_1, \ldots, f_m)$ be quadratic polynomials in $\mathsf{GF}_2[x_1, \ldots, x_n]$. Denote by r_{Mac} (resp. c_{Mac}, s_{Mac}) the number of rows (resp. columns, number of nonzero entries) of the associated boolean Macaulay matrix in degree d. If $1 \le d < n/2$, then*

$$c_{\mathsf{Mac}} < \frac{1-x}{1-2x}\binom{n}{d}, \; r_{\mathsf{Mac}} < m\frac{x^2}{(1-2x)(1-x)}\binom{n}{d}, \; s_{\mathsf{Mac}} < mn^2\frac{x^2}{(1-2x)(1-x)}\binom{n}{d}, \quad (7.2)$$

where $x = d/n$.

Proof. The number of columns of the boolean Macaulay matrix is simply the number of squarefree monomials of degree at most d in n variables. The number of rows is that same number of monomials for degree $d-2$, multiplied by the number m of polynomials. Finally, each row corresponding to a polynomial f_i has a number of nonzero entries bounded by the number of squarefree monomials of degree at most 2 in n variables. Standard combinatorial counting thus gives

$$c_{\mathsf{Mac}} = \sum_{i=0}^{d}\binom{n}{i}, \qquad r_{\mathsf{Mac}} = m\sum_{i=0}^{d-2}\binom{n}{i}, \qquad s_{\mathsf{Mac}} \le \left(1 + n + \binom{n}{2}\right) r_{\mathsf{Mac}} \le n^2 r_{\mathsf{Mac}}, \quad (7.3)$$

where in the last inequality we use the fact that $n \ge 2$. Now, the bounds come from a well-known inequality on binomial coefficients: for $0 \le d < n/2$,

$$\sum_{i=0}^{d}\binom{n}{i} < \frac{1}{1-d/(n-d)}\binom{n}{d}.$$

Indeed, the sequence $\binom{n}{i}$ is increasing for $0 \le i \le n/2$. Factoring out $\binom{n}{d}$ leaves a sum that is bounded by the geometric series $1 + d/(n-d) + \cdots$. This gives the bound for c_{Mac}. The bound for r_{Mac} is obtained by evaluating this bound at $d-2$, writing $\binom{n}{d-2}$ as a rational function times $\binom{n}{d}$ and finally bounding $x(x-1/n)/((1-2x+4/n)((1-x)+1/n))$ by $x^2/((1-2x)(1-x))$. $\square$

7.3.2 Bound on the Witness Degree of Inconsistent Systems

First, we prove that the witness degree can be bounded above by the so-called *degree of regularity* of the homogenized system. Here and subsequently, we call *dimension* of an ideal $I \subset R$ the Krull dimension of the quotient ring R/I (see e.g., [Eis95, §8]).

Definition 7.9. *The* degree of regularity $\mathrm{d_{reg}}(I)$ *of a homogeneous ideal I of dimension 0 is defined as the smallest integer d such that all homogeneous polynomials of degree d are in I.*

Proposition 7.10. *Let $\mathbf{F} = \left(f_1, \ldots, f_m, x_1^2 - x_1, \ldots, x_n^2 - x_n\right)$ be a sequence of polynomials such that the system $\mathbf{F} = 0$ has no solution. Then the ideal generated by the homogenized system*

$$I^{(h)} = \left\langle f_1^{(h)}, \ldots, f_m^{(h)}, x_1^2 - x_1 h, \ldots, x_n^2 - x_n h \right\rangle$$

has dimension 0 and $\mathrm{d_{wit}}(\mathbf{F}) \le \mathrm{d_{reg}}(I^{(h)})$.

Proof. By Hilbert's Nullstellensatz, the ideal I generated by $\mathbf{F}$ contains 1 (hence 1 is a Gröbner basis of I). Therefore, there exists $\alpha \in \mathbb{N} \setminus \{0\}$ such that $h^\alpha \in I^{(h)}$. Consequently, for the grevlex ordering, $\langle x_1^2, \ldots, x_n^2, h^\alpha \rangle \subset \mathsf{LM}(I^{(h)})$ and thus the dimension of $\mathsf{LM}(I^{(h)})$ is 0. As a consequence (see [CLO97, §9.3, Prop. 9]), $\dim(I^{(h)}) = \dim(\mathsf{LM}(I^{(h)})) = 0$.

Let $G^{(h)}$ be a minimal Gröbner basis of the homogenized ideal $I^{(h)}$ for the grevlex ordering. By definition of the degree of regularity, there exist polynomials ℓ_i and ℓ'_j such that $h^{\mathrm{d_{reg}}(I^{(h)})} = \sum_{1\le i\le m} f_i^{(h)}\ell_i + \sum_{1\le j\le n}(x_j^2 - x_j h)\ell'_j$. The ideal $I^{(h)}$ being homogeneous, it is possible to find such a combination with $\deg(f_i^{(h)}\ell_i) \le \mathrm{d_{reg}}(I^{(h)}), \deg((x_j^2 - x_j h)\ell'_j) \le \mathrm{d_{reg}}(I^{(h)})$ for all i, j. Evaluating this identity at $h = 1$ shows that 1 belongs to the vector space generated by the rows of the boolean Macaulay matrix in degree $\mathrm{d_{reg}}(I^{(h)})$. □

The next step is to obtain information on the Hilbert series for a large class of systems. To this end, we consider the so-called *syzygy module*, which describes the algebraic relations between the polynomials of a system.

Definition 7.11. *Let $(g_1, \ldots, g_\ell) \in (R^{(h)})^\ell$ be a polynomial system. A* syzygy *of $(g_1, \ldots, g_\ell)$ is a ℓ-tuple $(s_1, \ldots, s_\ell) \in (R^{(h)})^\ell$ such that $\sum_{i=1}^{\ell} s_i g_i = 0$. The set of all syzygies of $(g_1, \ldots, g_\ell)$ is a submodule of $(R^{(h)})^\ell$. The degree of a syzygy $\mathbf{s} = (s_1, \ldots, s_\ell)$ is defined as $\deg(\mathbf{s}) = \max_{1\le i\le \ell} \deg(g_i s_i)$.*

Obviously, for any such polynomial system, commutativity induces syzygies of the type

$$g_i g_j - g_j g_i = 0. \tag{7.4}$$

Moreover, for any constant $a \in \mathsf{GF}_2$ we have $a^2 = a$, thus expanding the square of a polynomial $\sum_{\alpha\in\mathbb{N}^k} a_\alpha \mathbf{x}^\alpha \in \mathsf{GF}_2[x_1, \ldots, x_k]$ gives $\sum_{\alpha\in\mathbb{N}^k} a_\alpha \mathbf{x}^{2\alpha}$. As a consequence, for a homogeneous quadratic polynomial $f_i^{(h)} = \sum_{1\le j,k\le n} a_{j,k}\, x_j x_k + \sum_{1\le j\le n} b_j\, x_j h + c\, h^2 \in \mathsf{GF}_2[x_1, \ldots, x_n, h]$, we obtain the following syzygy of the system $(f_i^{(h)}, x_1^2 - x_1 h, \ldots, x_n^2 - x_n h)$:

$$(f_i^{(h)} - h^2) f_i^{(h)} + \sum_{1\le j,k\le n} a_{j,k}\left(x_k^2(x_j^2 - x_j h) + x_j h(x_k^2 - x_k h)\right) + \sum_{1\le j\le n} b_j h^2 (x_j^2 - x_j h) = 0. \tag{7.5}$$

Definition 7.12. *Let $\mathbf{F}^{(h)} = (f_1^{(h)}, \ldots, f_n^{(h)}, x_1^2 - x_1 h, \ldots, x_n^2 - x_n h)$ be a system of homogeneous quadratic polynomials over GF_2. We call* trivial syzygies *of $\mathbf{F}^{(h)}$ and note Syz_{triv} the module generated by the syzygies of types (7.4) and (7.5).*

Definition 7.13. *A boolean homogeneous system $(f_1^{(h)}, \ldots, f_m^{(h)})$ is called*

- boolean semi-regular *in degree D if any syzygy whose degree is less than D belongs to Syz_{triv};*
- boolean semi-regular *if it is boolean semi-regular in degree $\mathrm{d_{reg}}(\langle f_1^{(h)}, \ldots, f_m^{(h)}, x_1^2 - x_1 h, \ldots, x_n^2 - x_n h\rangle)$.*

(This notion is slightly different from the *semi-regularity over* GF_2 defined in [BFS04, BFSY04].)

In the sequel we use the following notations: if $S \in \mathbb{Z}[[t]]$ is a power series, then $[S]$ denotes the series obtained by truncating S just before the index of its first nonpositive coefficient. Also, $[t^d]S(t)$ denotes the coefficient of t^d in S.

Proposition 7.14. *Let $(f_1^{(h)}, \ldots, f_m^{(h)})$ be a boolean homogeneous system. Let D_0 denote the degree of regularity of the system $(f_1^{(h)}, \ldots, f_m^{(h)}, x_1^2 - x_1 h, \ldots, x_n^2 - x_n h)$. If the systems $(f_1^{(h)}, \ldots, f_{i-1}^{(h)}, f_i^{(h)} - h^2)$ and $(f_1^{(h)}, \ldots, f_{i-1}^{(h)}, f_i^{(h)})$ are $D_0 - 2$ (resp. D_0)-boolean semi-regular for each $i \in \{2, \ldots, m\}$, then the Hilbert series of the homogeneous ideal $\langle f_1^{(h)}, \ldots, f_m^{(h)}, x_1^2 - x_1 h, \ldots, x_n^2 - x_n h\rangle$ is*

$$\mathsf{HS}_{n,m}(t) := \left[\frac{(1+t)^n}{(1-t)(1+t^2)^m}\right].$$

Proof. Let S_i (resp. S'_i) denote the system $(f_1^{(h)}, \dots, f_i^{(h)}, x_1^2 - x_1 h, \dots, x_n^2 - x_n h)$ (resp. $(f_1^{(h)}, \dots, f_i^{(h)} - h^2, x_1^2 - x_1 h, \dots, x_n^2 - x_n h)$). The general framework of this proof is rather classical: we prove by induction on i and d that for all $i \leq m$ and $d < D_0$, $\mathsf{HF}_{R^{(h)}/\langle S_i \rangle}(d) = \mathsf{HF}_{R^{(h)}/\langle S'_i \rangle}(d) = [t^d] \frac{(1+t)^n}{(1-t)(1+t^2)^i}$.

First, notice that a basis of the GF_2-vector space $R/\langle x_1^2 - x_1 h, \dots, x_n^2 - x_n h \rangle$ is the set of monomials $\mathfrak{S} = \{x_1^{\delta_1} \cdots x_n^{\delta_n} h^\ell \mid \delta_1, \dots, \delta_n \in \{0,1\}, \ell \in \mathbb{N}\}$. The generating function of this set is

$$\sum_{\mathfrak{m} \in \mathfrak{S}} t^{\deg(\mathfrak{m})} = \frac{(1+t)^n}{(1-t)}.$$

Therefore, the initialization of the recurrence comes from the relations

$$\begin{cases} \mathsf{HF}_{R^{(h)}/\langle x_1^2 - x_1 h, \dots, x_n^2 - x_n h \rangle}(d) = [t^d] \frac{(1+t)^n}{(1-t)} \text{ for all } d \in \mathbb{N}; \\ \mathsf{HF}_{R^{(h)}/\langle S_i \rangle}(0) = \mathsf{HF}_{R^{(h)}/\langle S'_i \rangle}(0) = 1 \text{ and } \mathsf{HF}_{R^{(h)}/\langle S_i \rangle}(1) = \mathsf{HF}_{R^{(h)}/\langle S'_i \rangle}(1) = n+1 \text{ for all } i \leq m. \end{cases}$$

In the following, $2 \leq d < D_0$ and $1 \leq i \leq m$ are two integers, and we assume by induction that for all $(\ell, j) \in \mathbb{N}^2$ such that $\ell < d$ or ($\ell = d$ and $j < i$), we have

$$\mathsf{HF}_{R^{(h)}/\langle S_j \rangle}(\ell) = \mathsf{HF}_{R^{(h)}/\langle S'_j \rangle}(\ell) = [t^\ell] \frac{(1+t)^n}{(1-t)(1+t^2)^j}.$$

Consider the following sequences

$$\begin{array}{l} 0 \to R^{(h)}_{d-2}/(S_{i-1} + \langle f_i^{(h)} - h^2 \rangle)_{d-2} \xrightarrow{\times f_i^{(h)}} \quad R^{(h)}_d/(S_{i-1})_d \to R^{(h)}_d/(S_i)_d \to 0 \\ 0 \to R^{(h)}_{d-2}/(S_{i-1} + \langle f_i^{(h)} \rangle)_{d-2} \xrightarrow{\times (f_i^{(h)} - h^2)} \quad R^{(h)}_d/(S_{i-1})_d \to R^{(h)}_d/(S'_i)_d \to 0, \end{array}$$

where the last arrow of each sequence is the canonical projection. Let g be in the kernel of the application

$$R^{(h)}_{d-2}/(S_{i-1} + \langle f_i^{(h)} - h^2 \rangle)_{d-2} \xrightarrow{\times f_i^{(h)}} R^{(h)}_d/(S_{i-1})_d.$$

Then $g f_i^{(h)}$ belongs to $(S_{i-1})_d$, which implies that there exist polynomials $g_1, \dots, g_{i-1}$, $h_1, \dots, h_n$ such that $(g_1, \dots, g_{i-1}, g, h_1, \dots, h_n)$ is a syzygy of degree d of the system S_i. By the boolean semi-regularity assumption, this syzygy belongs to Syz_{triv}, and hence $g \in \langle S_{i-1} \rangle + \langle f_i^{(h)} - h^2 \rangle$. Therefore the application $\times f_i^{(h)}$ is injective and the first sequence is exact. One can prove similarly that the second sequence is also exact.

These exact sequences yield relations between the Hilbert functions:

$$\mathsf{HF}_{R^{(h)}/S'_i}(d-2) - \mathsf{HF}_{R^{(h)}/S_{i-1}}(d) + \mathsf{HF}_{R^{(h)}/S_i}(d) = 0, \tag{7.6}$$

$$\mathsf{HF}_{R^{(h)}/S_i}(d-2) - \mathsf{HF}_{R^{(h)}/S_{i-1}}(d) + \mathsf{HF}_{R^{(h)}/S'_i}(d) = 0. \tag{7.7}$$

Moreover, we have the relation

$$[t^\ell] \frac{(1+t)^n}{(1-t)(1+t^2)^j} = [t^\ell] \frac{(1+t)^n}{(1-t)(1+t^2)^{j-1}} - [t^{\ell-2}] \frac{(1+t)^n}{(1-t)(1+t^2)^j}. \tag{7.8}$$

Using Relations (7.6) and (7.7), and the induction hypothesis, we get the desired result.

The proof is completed by showing that D_0 is equal to the index of the first nonpositive coefficient of $\mathsf{HS}_{R^{(h)}/S_m}(t)$. First, by definition of the degree of regularity, the coefficients $[t^d]\,\mathsf{HS}_{R^{(h)}/S_m}(t)$ are zero for $d \geq D_0$. Next, that the coefficient $[t^{D_0}]\frac{(1+t)^n}{(1-t)(1+t^2)^m}$ is nonpositive follows from the following property (easily proved by induction on i, $0 \leq i \leq m$ using (7.7–7.8)):

$$[t^{D_0}]\frac{(1+t)^n}{(1-t)(1+t^2)^i} \leq \mathsf{HF}_{S_i}(D_0).$$

□

Putting everything together, we have obtained the following.

Corollary 7.15. *With the same notation as in Proposition 7.10, if the homogenized system verifies the conditions of Proposition 7.14, then the witness degree of the system*

$$(f_1, \ldots, f_m, x_1^2 - x_1, \ldots, x_n^2 - x_n)$$

is bounded by the degree of the polynomial $\mathsf{HS}_{n,m}(t)$.

At this stage, it might seem that choosing the degree of $\mathsf{HS}_{n-k,m}$ for d_0 in Algorithm Boolean-Solve amounts to making a very strong assumption on the nature of the systems obtained by specialization followed by homogenization. In Section 7.4, we discuss experiments showing that this assumption is actually quite reasonable.

Finally, in order to compute the asymptotic behavior of our complexity estimates in the next section, we need the following.

Proposition 7.16. *Let $\alpha \geq 1$ be a real number. Then, as $n \to \infty$,*

$$\deg\left(\mathsf{HS}_{n,\lceil \alpha n\rceil}(t)\right) \sim M(\alpha)n,$$
$$\textit{with} \qquad M(x) := -x + \frac{1}{2} + \frac{1}{2}\sqrt{2x^2 - 10x - 1 + 2(x+2)\sqrt{x(x+2)}}.$$

Proof. We follow the approach of [BFS04, BFSY04]. We start from a representation of the coefficient as a Cauchy integral:

$$[t^d]\frac{(1+t)^n}{(1-t)(1+t^2)^m} = \frac{1}{2\pi\imath}\oint \frac{(1+z)^n}{(1-z)(1+z^2)^{\lceil \alpha n\rceil}}\frac{1}{z^{d+1}}dz,$$

where the contour is a circle centered in 0 whose radius is smaller than 1. We are searching for a value of d where this integral vanishes, for large n. We first estimate the asymptotic behaviour of the integral for fixed d. The integrand has the form $\exp(nf(z))$ with

$$f(z) = \log(1+z) - \frac{\lceil \alpha n\rceil}{n}\log(1+z^2) - \frac{\log(1-z) + (d+1)\log(z)}{n}.$$

As n increases, the integral concentrates in the neighborhood of one or several saddle points, solutions to the saddle-point equation $zf' = 0$, which rewrites

$$\frac{d}{n} = \frac{z}{1+z} - \frac{2\frac{\lceil \alpha n\rceil}{n}z^2}{1+z^2} - \frac{1-2z}{n(1-z)} =: \phi(z) + O(1/n). \tag{7.9}$$

In [BFS04], it is shown that for the contributions of saddle points to cancel out, two of them must coalesce and give rise to a double saddle point, given by the smallest positive double real root of the saddle-point equation, which is therefore such that $(zf')' = 0$. When n grows, the solutions of this equation tend towards the roots of $\phi'(z) = 0$. Let z_0 be the smallest positive real root of this equation.

The saddle-point equation (7.9) then gives $d \sim \phi(z_0)n$. Finally, eliminating z_0 using $\phi'(z_0) = 0$ by a resultant computation yields

$$d \sim \left(-\alpha + \frac{1}{2} + \frac{1}{2}\sqrt{2\alpha^2 - 10\alpha - 1 + 2(\alpha+2)\sqrt{\alpha(\alpha+2)}}\right) n.$$

□

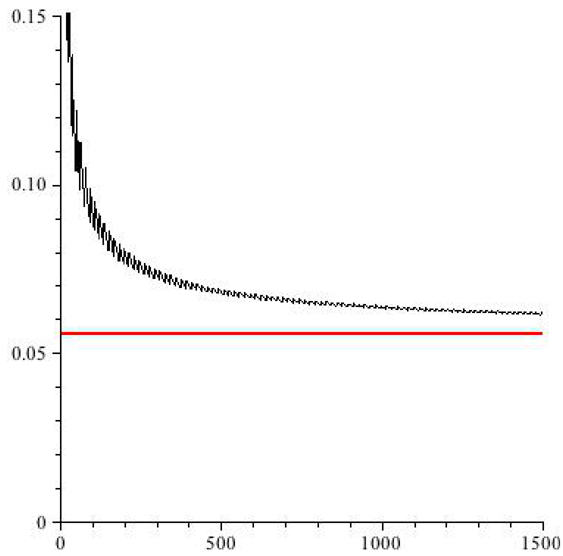

Figure 7.1: Comparison of $\deg(\mathsf{HS}_{n,\lceil n/.55\rceil})/n$ (black) with its limit (red).

Figure 7.1 shows the actual values of $\deg(\mathsf{HS}_{n,\lceil \alpha n\rceil})/n$ for $\alpha = 1/.55$. Notice that this sequence converges rather slowly. This is due to the fact that we only take into account the first term in the asymptotic expansion of $\deg(\mathsf{HS}_{n,\lceil \alpha n\rceil})$. It would be possible to obtain the full asymptotic expansion using techniques similar to those in [BFS04, BFSY04]. However, this would not change the asymptotic complexity of Algorithm 9.

7.3.3 Complexity

We now estimate the complexity of Algorithm BooleanSolve by going through its steps and making all necessary hypotheses explicit. We consider the case when the number of variables n and the number of polynomials m are related by $m \sim \alpha n$ for some $\alpha \geq 1$ and n is large. Also we assume that the ratio k/n is controlled by a parameter $\gamma \in [0, 1]$, i.e., $k = (1-\gamma)n$.

The first step (lines 4 to 6 in the algorithm) is to evaluate the polynomials $\tilde{f}_i$ from the polynomials f_i. With no arithmetic operations, the polynomials f_i can first be written as polynomials in $(x_1, \ldots, x_{n-k})$ with coefficients that are polynomials of degree at most 2 in $x_{n-k+1}, \ldots, x_n$ and at most 1 in each variable. Each such coefficient has at most $1 + k + \binom{k}{2}$ monomials, each of which can be evaluated with at most one arithmetic operation. The total number of these polynomial coefficients is at most $m(1 + n - k + \binom{n-k}{2})$. Thus the total cost of all the evaluations of the coefficients of the polynomials $\tilde{f}_i$ is at most $O(n^5 2^{(1-\gamma)n})$. This turns out to be asymptotically negligible compared to the next steps.

The next stage (line 8) of our algorithm consists in performing tests of inconsistency of the Macaulay matrices.

Proposition 7.17. *For any $\epsilon > 0$, $\alpha \geq 1$ and sufficiently large $m = \lceil \alpha n \rceil$, the complexity of all tests of consistency of Macaulay matrices in Algorithm* **BooleanSolve** *with parameters (m, n, k) is*

- $O(2^{(1-\gamma+\omega F_\alpha(\gamma)+\epsilon)n})$ *in the deterministic variant;*
- *of expectation $O(2^{(1-\gamma+2F_\alpha(\gamma)+\epsilon)n})$ in the probabilistic variant,*

where $\gamma = 1 - k/n$, $F_\alpha(\gamma) = -\gamma \log_2(D^D(1-D)^{1-D})$ with $D = M(\alpha/\gamma)$, the function M as in Proposition 7.16 and ω the complexity of linear algebra as in Proposition 7.6.

A feature of this result is that in terms of complexity, the probabilistic variant of our algorithm behaves as the deterministic one where the linear algebra would be performed in quadratic complexity (i.e., with $\omega = 2$).

Proof. We first estimate the size of the Macaulay matrices. By Proposition 7.16, the index d_0, which is $1 + \deg(\mathsf{HS}_{n-k,m})$ behaves asymptotically like $\gamma D n$. The function $M(x)$ is decreasing for $x \geq 1$, so that $D \leq M(1) < 1/2$. Thus, $d_0 < \gamma n/2$ for n sufficiently large and Proposition 7.8 applies with $d = d_0$, $m = \lceil \alpha n \rceil$ equations and $n - k = \gamma n$ variables. For n sufficiently large, the bound for r_{Mac} is larger than that for c_{Mac}, since the quotient of these two bounds is $m/(\frac{\gamma n}{d_0} - 1)^2$, which grows linearly with n.

Next, we turn to the tests of inconsistency. The previous bounds and Proposition 7.6 imply that the number of operations required for the computation of the row echelon form is $O(n\binom{\gamma n}{d_0}^\omega)$. Similarly, by Proposition 7.7, the complexity of checking the consistency of each matrix by the probabilistic method is $O(r_{\mathsf{Mac}} \log(r_{\mathsf{Mac}}) s_{\mathsf{Mac}}) = O(n^4 \binom{\gamma n}{d_0}^2 \log \binom{\gamma n}{d_0})$ and that bound dominates the cost of the additional operations in GF_2. Now, Stirling's formula implies that for any $0 < b < a$, $\log \binom{an}{bn} \sim n \log(a^a/(b^b(a-b)^{a-b}))$. Setting $a = \gamma$ and $b = \gamma D$ gives the result, the extra factor being due to the exhaustive search that performs this consistency check $2^{(1-\gamma)n}$ times. $\square$

In the cases where the linear system $\mathbf{u} \cdot \mathsf{M} = \mathbf{r}$ is found inconsistent, then the polynomial system itself may be consistent and the algorithm proceeds with an exhaustive search (line 9) in a system with γn unknowns. Each such search has cost $O(2^{(\gamma+\epsilon)n})$. As long as the number of these searches does not exceed $O(2^{(1-2\gamma+2F_\alpha(\gamma))n})$, the overall complexity of the algorithm is bounded by the complexity given in Proposition 7.17. There can be two causes for the inconsistency of the linear system that triggers such a search: the existence of an actual solution with $x_n = a_n, \ldots, x_{n-k+1} = a_{n-k+1}$; a witness degree of the specialized system larger than d_0 (e.g., if the homogenized specialized system is not boolean semi-regular). We now define a class of systems where this does not happen too much.

Definition 7.18. *Let $S = (f_1, \ldots, f_m)$ be quadratic polynomials in $\mathsf{GF}_2[x_1, \ldots, x_n]$, $0 \leq k = (1-\gamma)n < n$, $\alpha = m/n$ and $d_0 = 1 + \deg(\mathsf{HS}_{n-k,m})$. The system S is called γ-strong semi-regular if both the set of its solutions in GF_2^n and the set*

$$\big\{(a_{n-k+1}, \ldots, a_n) \in \mathsf{GF}_2^k \mid \\ \mathrm{d}_{\mathsf{wit}}(f_1(x_1, \ldots, x_{n-k}, a_{n-k+1}, \ldots, a_n), \ldots, f_m(x_1, \ldots, x_{n-k}, a_{n-k+1}, \ldots, a_n)) > d_0\big\}$$

have cardinality at most $2^{(1-2\gamma+2F_\alpha(\gamma))n}$, with F_α as in Proposition 7.17.

Note that since $1 - 2\gamma + 2F_\alpha(\gamma)$ is a decreasing function of γ, a γ-strong semi-regular system is also γ'-strong semi-regular for any $\gamma' < \gamma$.

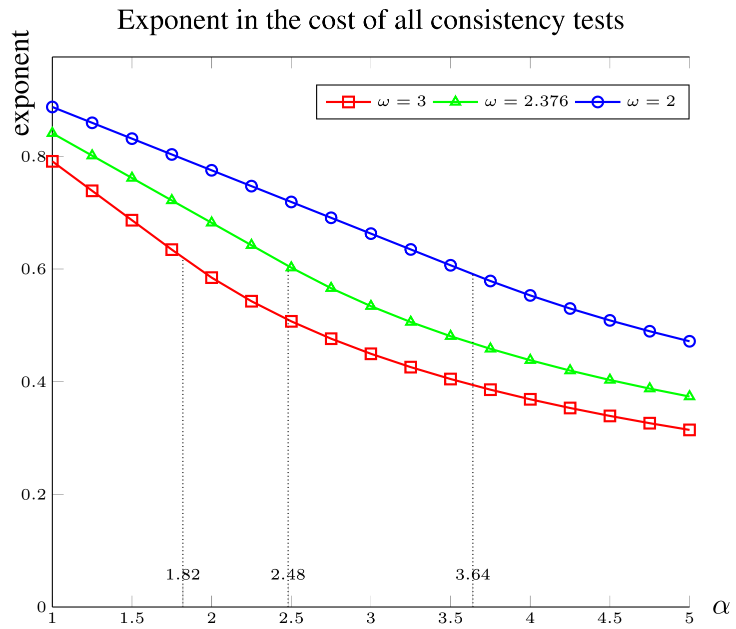

Figure 7.2: Exponent of the complexity for inconsistent systems in terms of the ratio α (see Thm. 7.20 and Cor. 7.19)

The first condition for a system to be γ-strong semi-regular concerns its number of solutions. For boolean systems drawn uniformly at random, it is known that the probability that the number of boolean solutions is s decreases more than exponentially with s [FB07], so that the first condition is fulfilled with large probability. The second condition is related to the proportion of boolean semi-regular systems. We discuss this condition in the next section and show that it is also of large probability experimentally. Under this assumption of γ-strong semi-regularity, we now state the complexity of the algorithm obtained by optimizing the choice of the number k of variables that are specialized.

We first discuss large values of γ. The function $1 - 2\gamma + 2F_\alpha(\gamma)$ is decreasing with α and negative when $\gamma = 1$. Thus, the first condition implies that a 1-strong semi-regular system has no solution. By continuity, this behavior persists for γ close to 1 and actually holds for $\gamma \in (0.824, 1)$. It also persists for smaller values of γ and larger α.

Corollary 7.19. *With the same notations as in Prop. 7.17, when a system is γ-strong semi-regular with α and γ such that $1 - 2\gamma + 2F_\alpha(\gamma) < 0$, then it is inconsistent and detected by Algorithm* BooleanSolve *with parameters $(m, n, 0)$ in $O(2^{(\omega F_\alpha(1)+\epsilon)n})$ operations.*

The value of the exponent $\omega F_\alpha(1)$ in terms of α is plotted in Figure 7.2 (it corresponds to the right part of the plots, i.e. $\alpha > 1.82$ for $\omega = 2$, $\alpha > 2.48$ for $\omega = 2.376$, $\alpha > 3.64$ for $\omega = 3$).

Proof. The hypothesis implies that the system has no solution and that its witness degree is bounded by d_0, so that its absence of solution is detected by the linear algebra step in degree d_0. In that case, no exhaustive search is needed. □

For smaller values of γ, the algorithm requires exhaustive searches. The optimal choice of k is obtained by an optimization on the complexity estimate. This leads to the following complexity estimates. In the next section, we argue that the required strong semi-regularities are very likely in practice, so that the only choice left to the user is that of the linear algebra routine.

Theorem 7.20. *Let $S = (f_1, \dots, f_m)$ be a system of quadratic polynomials in $\mathsf{GF}_2[x_1, \dots, x_n]$, with $m = \lceil \alpha n \rceil$ and $\alpha \geq 1$. Then Algorithm* BooleanSolve *finds all its roots in GF_2^n with a number of arithmetic operations in GF_2 that is*

- $O(2^{(1-0.112\alpha)n})$ *if S is $(.27\alpha)$-strong semi-regular using Gaussian elimination for the linear algebra step;*
- $O(2^{(1-0.159\alpha)n})$ *if S is $(.40\alpha)$-strong semi-regular using computation of the row echelon form with Coppersmith-Winograd multiplication;*
- *of expectation $O(2^{(1-0.208\alpha)n})$ if S is $(.55\alpha)$-strong semi-regular using the probabilistic Algorithm 10.*

In all cases, the value of k passed to the algorithm is $\lceil n(1-\gamma) \rceil$ with γ corresponding to the strong semi-regularity.

Proof. The correctness of the algorithm has already been proved in Proposition 7.5. Only the complexity remains to be proved.

By definition of strong semi-regularity, the number of exhaustive searches that need be performed in line 9 of the Algorithm is $O(2^{(1-2\gamma+2F_\alpha(\gamma))n})$, each of them using $O(2^{(\gamma+\epsilon)n})$ arithmetic operations for any $\epsilon > 0$. It follows that the overall cost of these exhaustive searches is $O(2^{(1-\gamma+2F_\alpha(\gamma)+\epsilon)n})$; it is bounded by the cost of the tests of inconsistency. We now choose γ in such a way as to minimize this cost, in terms of α. Direct computations lead to the following numerical results, that conclude the proof. □

Lemma 7.21. *With the same notation as in Proposition 7.17, the function $1-\gamma+\omega F_\alpha(\gamma)$ is bounded by*

- $1 - 0.112\alpha$ *when $\omega = 3$ and $\gamma = 0.27\alpha$;*
- $1 - 0.159\alpha$ *when $\omega = 2.376$ and $\gamma = 0.40\alpha$;*
- $1 - 0.208\alpha$ *when $\omega = 2$ and $\gamma = 0.55\alpha$.*

Proof. The function $1-\gamma+\omega F_\alpha(\gamma)$ has two parameters but its extrema can be found by reducing it to a one parameter function. Indeed, this function is maximal for $\alpha \geq 1$ and $\gamma \in [0,1]$ when $(-\gamma+\omega F_\alpha(\gamma))/\alpha$ is. Setting $\lambda = \gamma/\alpha$, this is exactly $-\lambda+\omega F_1(\lambda)$, with $\lambda \in [0, 1/\alpha]$. Direct computations lead to the optimal λ's: $\lambda = \min(1/\alpha, 0.27)$ when $\omega = 3$, $\lambda = \min(1/\alpha, 0.40)$ when $\omega = 2.376$, $\lambda = \min(1/\alpha, 0.55)$ when $\omega = 2$. □

7.4 Numerical Experiments on Random Systems

Probabilistic model. In this section, we study experimentally the behavior of Algorithm BooleanSolve of random quadratic systems where each coefficient is 0 or 1 with probability 1/2. These random boolean quadratic systems appear naturally in Cryptology since the security of several recent cryptosystems relies directly on the difficulty of solving such systems (see e.g., [BGP06, BGP09]).

7.4.1 γ-strong semi-regularity

The goal of this section is to give experimental evidence that the assumption of γ-strong semi-regularity is not a strong condition for random boolean systems. This is related to the notoriously difficult conjecture by [Fro85], which states that in characteristic 0, almost all systems are semi-regular (with the meaning of semi-regularity given in [BFSY04]), see also [MS03].

Consequently, we propose the following conjecture, which can be seen as a variant of Fröberg's conjecture for boolean systems:

Conjecture 7.22. *For any $\alpha \geq 1$ and $\gamma < 1$ such that $1-2\gamma+2F_\alpha(\gamma) > 0$, the proportion of γ-strong semi-regular systems of $\lceil \alpha n \rceil$ quadratic polynomials in $\mathsf{GF}_2[x_1, \ldots, x_n]$ tends to 1 when $n \to \infty$.*

The rest of this section is devoted to providing experiments supporting this conjecture.

In Figure 7.3, we show the relation between the value of the first nonpositive coefficient of the power series expansion of $\mathsf{HS}_{\lfloor \gamma n \rfloor, n}$ and γ-strong semi-regularity for small values of $n = m$ (i.e. $\alpha = 1$). For each n, the experiments are conducted on 1000 random quadratic boolean systems. For each of these systems, we compute the $2^{\lceil (1-\gamma)n \rceil}$ specialized systems and we count the number of specializations for which the filtering linear system is inconsistent.

Four curves are represented on each chart in Figure 7.3. The red (resp. green) one represents the average (resp. maximal) number of specializations for which the linear system (step 8 of Algorithm $\mathsf{BooleanSolve}$) is inconsistent. In contrast, the blue curve shows the upper bound on this number of specializations required to be γ-strong semi-regular (see Definition 7.18). The black curve shows the absolute value of the first nonpositive coefficient of the corresponding power series (i.e. $\mathsf{HS}_{\lfloor \gamma n \rfloor, n}$). The y-axis is represented in logarithmic scale. The value $\gamma = 0.1$ is never used in the complexity analysis (since in Theorem 7.20, $\gamma \geq .27$ for any value of $\alpha \geq 1$). However, it is still interesting to study the behavior of Algorithm 9 when almost all variables are specialized: the filtering remains very efficient in this case, and the branches which are explored during the second stage of the exhaustive search correspond to those containing solutions of the system.

Interpretation of Figure 7.3. First, notice that for $\gamma \leq 0.55$ the green curve is always below the blue one (except for the case $\gamma = .55, n = 23$), meaning that during our experiments, all randomly generated systems with those parameters were γ-strong semi-regular.

Next, in most curves (except $\gamma = 0.27$), the average (resp. maximal) number of points where the specialization leads to an inconsistent linear system is close to 1 (resp. 5). This can be explained by a simple Poisson model. Indeed, the number of solutions of a random boolean system with as many equations as unknowns follows a Poisson law with parameter 1 (see [FB07]). Therefore, the expectation of the number of solutions is 1. The expectation of the maximum of the number of solutions of 1000 random systems is then given as the maximum of 1000 iid random variables $P_1, \ldots, P_{1000}$ following a Poisson law of parameter 1:

$$\mathbf{E}(\max(P_1, \ldots, P_{1000})) = \sum_{k \geq 1} k \left((e^{-1} \sum_{i=0}^{k} \frac{1}{i!})^{1000} - (e^{-1} \sum_{i=0}^{k-1} \frac{1}{i!})^{1000} \right) \simeq 5.51,$$

which explains very well the observed behaviour.

This means that during Algorithm 9 with these parameters, almost all specializations giving rise to an inconsistent system correspond to a branch of the exhaustive search which contains an actual solution of the system. Therefore, the filtering is very efficient for those parameters.

Few specializations. In the case $\gamma = 0.9$, the blue curve has a negative slope. This is due to the fact that the quantity $1 - 2\gamma + 2F_\alpha(\gamma)$ (see Definition 7.18) is negative for $\alpha = 1$ and $\gamma > 0.82308$. Therefore, we cannot expect that a large proportion of boolean systems are γ-strong semi-regular in this setting. A limit case is investigated in the chart corresponding to $\gamma = 0.81$. There, $1 - 2\gamma + 2F_\alpha(\gamma) \approx 0.0102$ is positive but very close to zero. Experiments show that random boolean systems with these parameters and $10 \leq n \leq 24$ are γ-strong semi-regular with probability approximately equal to 0.75.

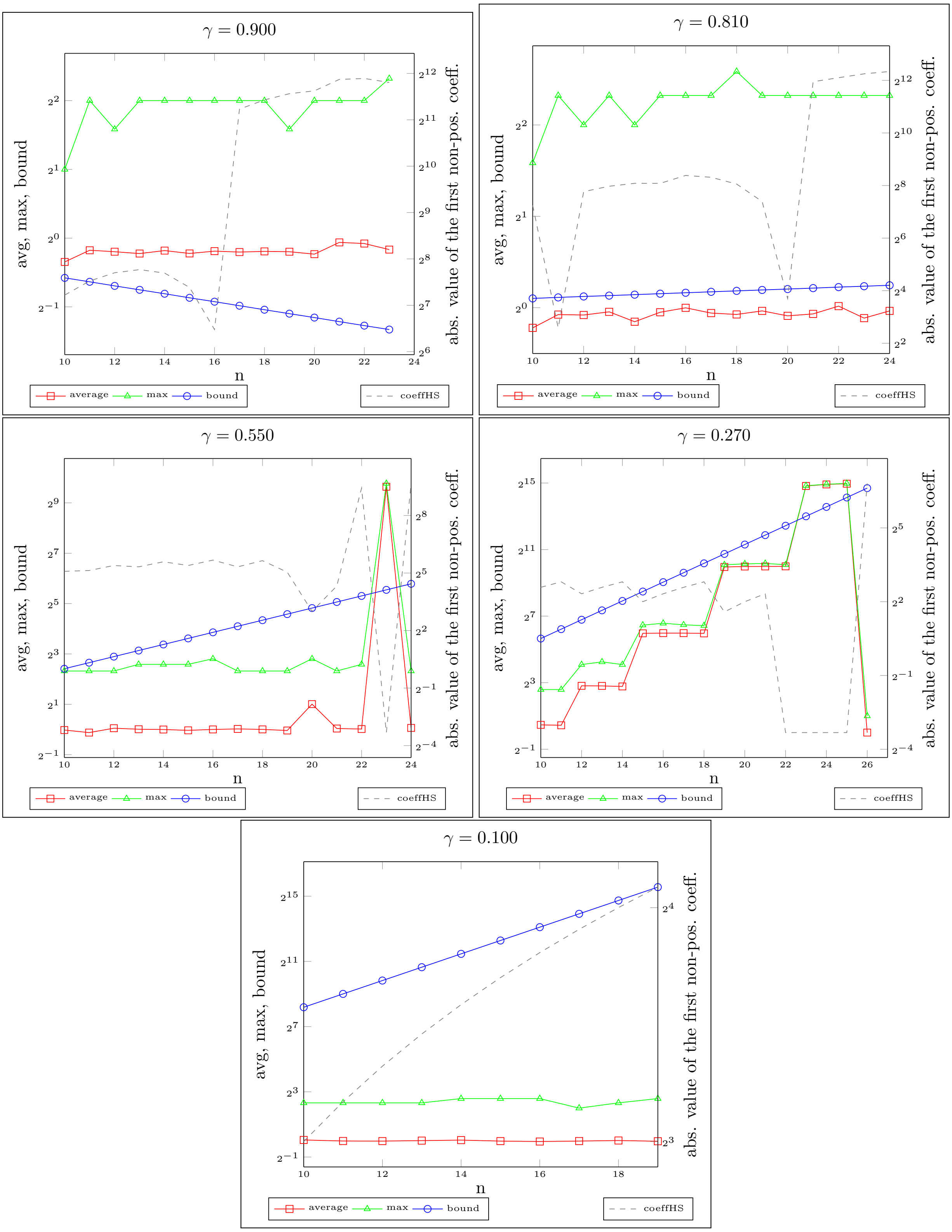

Figure 7.3: Relation between the quality of the filtering, the value of the first nonpositive coefficient of $\mathsf{HS}_{\lfloor \gamma n \rfloor, n}$, and γ-strong semi-regularity. In red (resp. green), the average (resp. maximum) number of specializations for which the linear system is inconsistent. In blue, the bound for γ-strong regularity. Dashed line: absolute value of the first non positive coefficient of $\mathsf{HS}_{\lfloor \gamma n \rfloor, n}$.

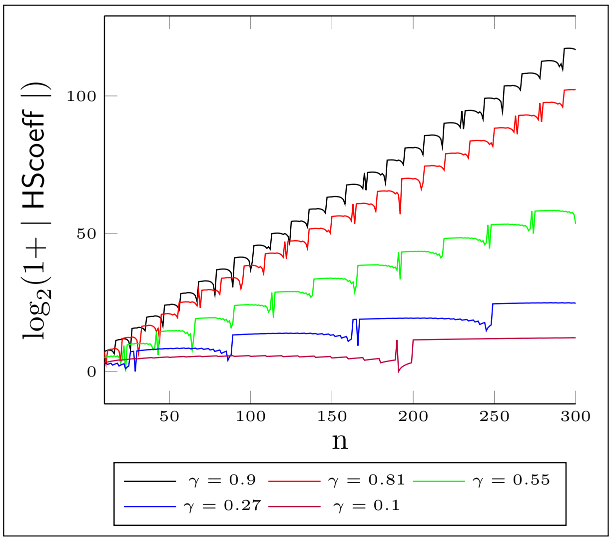

Figure 7.4: Evolution of the logarithm of the absolute value of the first nonpositive coefficient of $\mathsf{HS}_{\lfloor \gamma n \rfloor, n}$.

Absolute value of the first nonpositive coefficient of $\mathsf{HS}_{\lfloor \gamma n \rfloor, n}$ and γ-strong semi-regularity. Another interesting setting is $\gamma = .55, n = 23$. Here, no generated systems were γ-strong semi-regular (although all generated systems for $n \neq 23$ were γ-strong semi-regular). As explained in Section 7.5.1, this is due to the fact that the first nonpositive coefficient of the power series expansion of $\mathsf{HS}_{\lfloor \gamma n \rfloor, n}$ is equal to zero. In Section 7.5.2, we show that this phenomenon can be avoided by a simple variant of the algorithm.

A similar phenomenon happens for $\gamma = .27$: the first nonpositive coefficient of the power series has small absolute value. It is an accident due to the fact that this coefficient is close to zero for $n \leq 25$ (see Figure 7.4). On this chart, we can see clearly the relation between the absolute value of the first nonpositive coefficient of $\mathsf{HS}_{\lfloor \gamma n \rfloor, n}$ and the number of specializations for which the consistency test fails.

Indeed, experiments on 1000 random systems with $\gamma = .27$ and $n = 26$ were conducted and in this case the average number of specializations for which the linear system is inconsistent is 1.

These experiments justify the fact that the complexity analysis conducted in Section 7.3 is relevant for a large class of boolean systems. Also, it shows that the random systems for which the filtering may not be efficient can be detected *a priori* by looking at the absolute value of the first nonpositive coefficient in the power series. If this value is small, we show in Section 7.5.2 that the quality of the filtering can be improved at low cost by adding redundancy.

Figure 7.4 shows the evolution of the logarithm of the absolute value of the first nonpositive coefficient of $\mathsf{HS}_{\lfloor \gamma n \rfloor, n}$. This absolute value seems to grow exponentially with n for any given γ. Since the quality of the filtering is related to this absolute value, these experiments suggest that the proportion of γ-strong semi-regular systems tends towards 1 when n grows, as formulated in Conjecture 7.22.

7.4.2 Numerical estimates of the complexity

When $n = m$ and in the most favorable algorithmic case, our complexity estimate uses $\gamma = .55$. For this value, we display in Figure 7.1 (page 162) a comparison of the behaviour of $\deg(\mathsf{HS}_{n, \lceil \frac{n}{\gamma} \rceil})/n$ and its limit. This picture shows a relatively slow convergence. Thus, for a given number n of variables it is more interesting to optimize γ using the exact value of $\deg(\mathsf{HS}_{\lfloor \gamma n \rfloor, n})$ rather than a first order

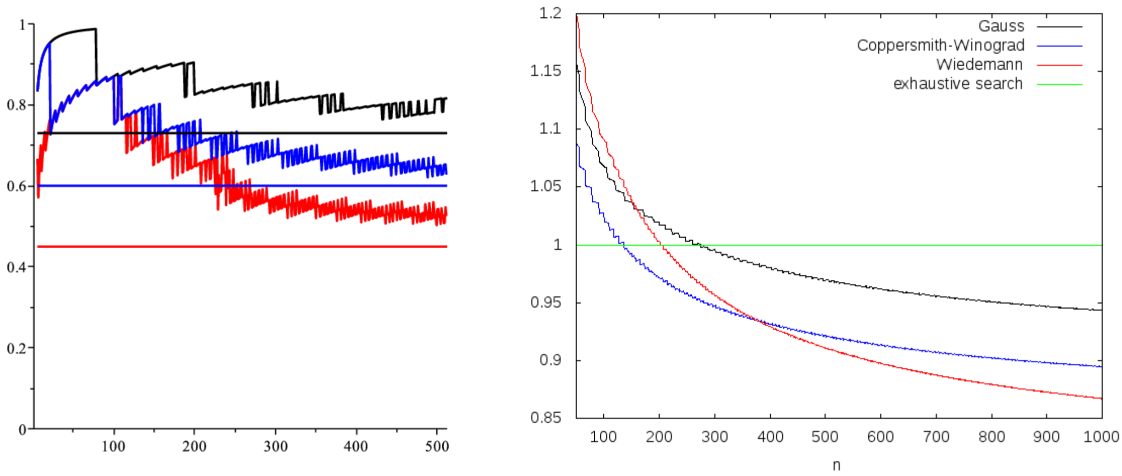

Figure 7.5: Left: optimal values of γ for the probabilistic variant (red), the deterministic variant with Gaussian elimination (black) and Coppersmith-Winograd matrix multiplication (blue), and their limits. Right: corresponding values of $\log_2 N/n$, with N given by Eq. (7.10). The green line corresponds to an exhaustive search.

asymptotic estimate. In the same spirit, one can also use the actual values given by Eq. (7.2) for the Macaulay matrix. Thus we seek to find γ that minimizes the following bounds on the number of operations:

$$\begin{aligned} & 2^{(1-\gamma)n} r_{\mathsf{Mac}} c_{\mathsf{Mac}} \min(r_{\mathsf{Mac}}, c_{\mathsf{Mac}})^{\omega-2}, \\ \text{resp.} \quad & 2^{(1-\gamma)n} \max(r_{\mathsf{Mac}}, c_{\mathsf{Mac}}) \log \max(r_{\mathsf{Mac}}, c_{\mathsf{Mac}}) s_{\mathsf{Mac}} \end{aligned} \tag{7.10}$$

in the deterministic (resp. probabilistic) variants, using Eq. (7.3) with n equations, $\lfloor \gamma n \rfloor$ variables and $d = \deg(\mathsf{HS}_{\lfloor \gamma n \rfloor, n})$. The corresponding values of γ are given in Figure 7.5, together with the corresponding values of the quantities in Eq. (7.10). Although these values do not take into account the constants hidden in the $O()$ estimates of the complexity, they suggest the relevance of these algorithms in the cryptographic sizes: the threshold between exhaustive search and our algorithm with Gaussian elimination is $n \simeq 280$, while the asymptotically faster Las Vegas variant starts being faster than exhaustive search for n larger than 200 and beats deterministic Gaussian elimination for n larger than 160.

7.5 Extensions and Applications

7.5.1 Adding Redundancy to Avoid Rank Defects

We showed in Section 7.4.1 that when the first nonpositive coefficient of $\mathsf{HS}_{n-k,n}$ is close to zero, then the linear filtering may not be as efficient as expected (for instance in the case $\gamma = .55$, $n = 23$ in Figure 7.3). Another case is shown in Figure 7.6. The curve $\delta = 0$ shows the behavior of Algorithm 9 on random square systems ($m = n$) where k is chosen as small as possible such that the witness degree is $\mathrm{d_{wit}} = 2$: this is obtained by choosing $k = \left\lceil 1/2 + n - \frac{\sqrt{-7+8n}}{2} \right\rceil$ (that is $d_0 = 2$).

First, we observe that specializing a uniformly distributed random quadratic polynomial $P \in \mathsf{GF}_2[x_1, \ldots, x_n]$ at a uniformly distributed random point in GF_2^k yields a random polynomial that is also uniformly distributed in $\mathsf{GF}_2[x_1, \ldots, x_{n-k}]$. We assume here that P is reduced modulo the field

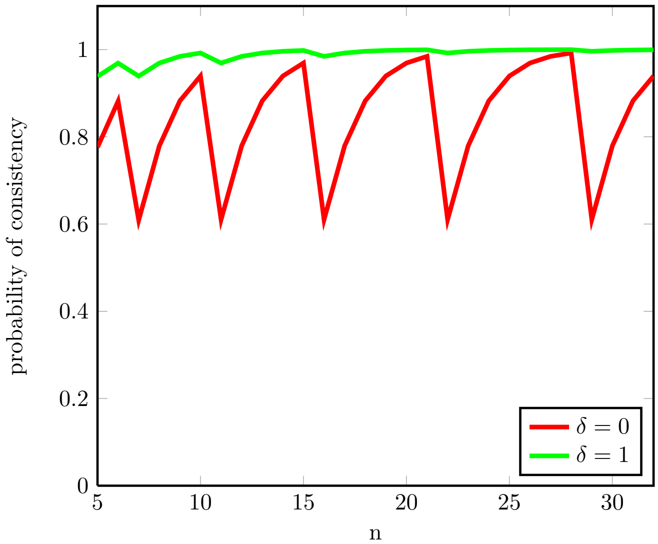

Figure 7.6: Proportion of specialized quadratic systems for which the linear system (line 9 of Algorithm 9) is consistent. Parameters: $k = \left\lceil 1/2 + n - \frac{\sqrt{-7+8n}}{2} \right\rceil$. In red, $\delta = 0$ (corresponding to Algorithm 9); in green, $\delta = 1$ (see Algorithm 11 of Section 7.5.2).

equations $\langle x_1^2 - x_1, \ldots, x_n^2 - x_n \rangle$. Let us assume first that $k = 1$. Then P can be rewritten as

$$P(x_1, \ldots, x_n) = x_n P_1(x_1, \ldots, x_{n-1}) + P_2(x_1, \ldots, x_{n-1}),$$

where P_1 (resp. P_2) is a random polynomial following a uniform distribution on the set of reduced boolean polynomials of degree 1 (resp. of degree 2) in $\mathsf{GF}_2[x_1, \ldots, x_{n-1}]$. Therefore, if $a \in \mathsf{GF}_2$ is a random variable, $P(x_1, \ldots, x_{n-1}, a) \in \mathsf{GF}_2[x_1, \ldots, x_{n-1}]$ is either P_1 or $P_1 + P_2$ and thus follows a uniform distribution on the set of reduced quadratic boolean polynomials. The extension to arbitrary $k < n$ follows by induction.

Consequently, in the special case $d_0 = 2$ of Figure 7.6 the boolean Macaulay matrix of a specialized system will be uniformly distributed among the boolean matrices with the same dimensions. Also, due to the choice of k, it will be roughly square. However, in GF_2, the probability that a random square matrix has full rank is not close to 1. An estimate of this probability can be obtained as follows.

The probability that a random $p \times q$ boolean matrix has rank r is (see [FA66, Sti87])

$$P(p, q, r) = 2^{-pq} \frac{\prod_{j=0}^{r-1}(2^p - 2^j) \prod_{j=0}^{r-1}(2^q - 2^j)}{\prod_{j=0}^{r-1}(2^r - 2^j)}.$$

Therefore, given a nonzero vector $\mathbf{v} \in \mathsf{GF}_2^p$ and a random boolean $p \times q$ matrix M, the probability that the linear system $\mathbf{u} \cdot \mathsf{M} = \mathbf{v}$ is consistent is

$$Q(p, q) = \sum_{i=1}^{p} P(p, q, i) \left(\frac{2^i - 1}{2^q - 1} \right).$$

Direct numerical computations show that for square matrices, $Q(p, p) \approx 0.61$ as soon as $p \geq 4$. This probability corresponds to the valleys of the curve $\delta = 0$ in Figure 7.6. Also, it can be noticed that $Q(p, q)$ grows quickly with $p - q$. For instance, $Q(p + 6, p) \approx 0.99$ when $p \geq 1$.

Consequently, it is interesting to specialize more variables than k in some cases (especially when the first nonpositive coefficient of $(1+t)^{n-k}/((1-t)(1+t^2)^m)$ has small absolute value): doing so increases the difference between the dimensions of the Macaulay matrices. This does not change the correctness of the algorithm (nor its asymptotic complexity), but increases the effectiveness of the filtering performed by linear algebra.

7.5.2 Improving the quality of the filtering for small values of n

In this section, we propose an extension of Algorithm BooleanSolve which takes an extra argument δ, in order to avoid the behavior of the algorithm shown in Section 7.5.1. The main idea is to specialize $k+\delta$ variables, but to take only k into account for the computation of d_0. Consequently, the difference between the number of columns and the rank of the Macaulay matrix is not too small, and hence the linear filtering performs better. The resulting algorithm is given in Algorithm 11.

Algorithm 11 improved BooleanSolve.

Input: $m, n, k, \delta \in \mathbb{N}$ such that $k+\delta < n \leq m$ and $f_1, \dots, f_m$ quadratic polynomials in $\mathsf{GF}_2[x_1, \dots, x_n]$.

Output: The set of boolean solutions of the system $f_1 = \cdots = f_m = 0$.

```
S := ∅.
d_0 := index of the first nonpositive coefficient in the series expansion of the rational function (1+t)^(n-k) / ((1-t)(1+t^2)^m).
for all (a_{n-k-δ+1}, ..., a_n) ∈ GF_2^{k+δ} do
    for i from 1 to m do
        f~_i(x_1, ..., x_{n-k-δ}) := f_i(x_1, ..., x_{n-k-δ}, a_{n-k-δ+1}, ..., a_n).
    end for
    M := boolean Macaulay matrix of (f~_1, ..., f~_m) in degree d_0.
    if the system u · M = r is inconsistent then          ▷ r as defined in Lemma 7.3
        T := solutions of the system (f~_1 = ··· = f~_m = 0) (exhaustive search).
        for all (t_1, ..., t_{n-k-δ}) ∈ T do
            S := S ∪ {(t_1, ..., t_{n-k-δ}, a_{n-k-δ+1}, ..., a_n)}.
        end for
    end if
end for
return S.
```

In Figure 7.6, we show the role of the parameter δ when k is chosen minimal such that $d_0 = 2$: adding redundancy by choosing a nonzero δ can greatly improve the quality of the filtering (in practice, choosing $\delta = 1$ is sufficient).

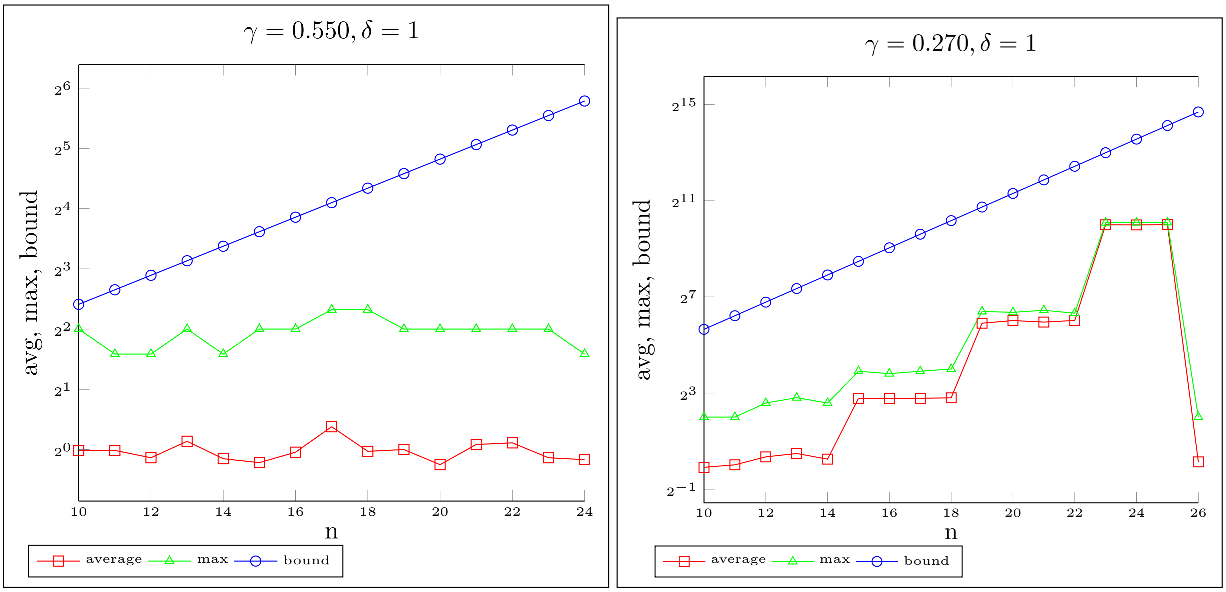

Figure 7.7: Quality of the filtering with $\delta = 1$.

Figure 7.7 shows further experimental evidence that adding redundancy by choosing $\delta = 1$ permits to avoid problems occurring when the first nonpositive coefficient of $\mathsf{HS}_{n-k,m}$ is close to zero. For instance, the peak at $\gamma = .55$, $n = 23$ that appeared in Figure 7.3 disappears when $\delta = 1$.

7.5.3 Cases with Low Degree of Regularity

In some cases, when the boolean system is *not random*, the choice of d_0 proposed in Algorithm BooleanSolve may be too large. This happens for instance for systems that have inner structure, which has an impact on the algebraic structure of the ideal generated by the polynomials. Examples of such structure can be found in Cryptology, for instance with boolean systems coming for the HFE cryptosystem [Pat96], as shown in [FJ03].

For these systems, the choice of d_0 as the index of the first non-positive coefficient of $\mathsf{HS}_{n,m}$ would be very pessimistic, since the Macaulay matrices in degree d_0 would be larger than necessary. However, if estimates of the witness degree are known (this is the case for HFE), then d_0 can be chosen accordingly as a parameter of the Algorithm BooleanSolve.

Chapter 8

Application to Cryptology

Section 8.1 is joint work with J.-C. Faugère and the results are published in [FS10]. Section 8.2 is joint work with J.-C. Faugère and M. Safey El Din and the results are published in [FSS10]. Section 8.3 is joint work with M. Bardet, J.-C. Faugère and B. Salvy and is a part of [BFSS12].

8.1 Cryptanalysis of the Algebraic Surface Cryptosystem

Notation for this section. In this section, to avoid any confusion between the symbols w and ω, we use the notation ϑ for the exponent in the complexity of linear algebra (i.e. two $n \times n$ matrices can be computed within $O(n^{\vartheta})$ arithmetic operations).

In this section, we propose an algebraic attack on the Algebraic Surface Cryptosystem (ASC for short) proposed at PKC'2009 [AGM09]. This cryptosystem is based on an unusual problem in multivariate cryptography: the *Section Finding Problem*. Given $w \in \mathbb{N}$ and an algebraic surface $X(x, y, t) \in \mathsf{GF}_p[x, y, t]$ such that $\deg_{xy} X(x, y, t) = w$, the problem is to find a pair of polynomials of degree d, $u_x(t)$ and $u_y(t)$, such that $X(u_x(t), u_y(t), t) = 0$. In ASC, the public key is the surface, and the secret key is the section. This asymmetric encryption scheme enjoys small sizes of the keys: for recommended parameters, the size of the secret key is only 102 bits and the size of the public key is 500 bits. We propose a message recovery attack whose complexity is quasi-linear in the size of the secret key. The main idea of this algebraic attack is to decompose ideals constructed from the ciphertext in order to avoid to solve the section finding problem. Experimental results show that we can break the cipher for recommended parameters (the security level is 2^{102}) in 0.05 seconds. Furthermore, the attack still applies even when the secret key is very large (more than 10000 bits). The complexity of the attack is $\widetilde{O}(w^{2\vartheta+1} d \log(p))$ which is polynomial with respect to all security parameters. In particular, it is quasi-linear in the size of the secret key which is $(2d + 2)\log(p)$. This result is rather surprising since the algebraic attack is often more efficient than the legal decryption algorithm.

8.1.1 Introduction

In 1994, Shor designed a quantum algorithm to compute efficiently discrete logarithm and factorization [Sho94]. Hence, if one could construct a quantum computer, a huge number of well-established public key cryptosystems – for instance, RSA or Elliptic Curve based systems – would be seriously threatened. Therefore, cryptographers are searching for post-quantum alternatives. The first step to design new cryptosystems is to identify hard problems to use as trapdoors. So far, most of the problems used in post-quantum cryptology can be classified into three main categories: Multivariate cryptography, Code-based cryptography and Lattice-based cryptography.

In this context, Akiyama, Goto, and Miyake propose a new multivariate public-key algorithm at PKC'2009: the Algebraic Surface Cryptosystem (*ASC* for short) [AGM09]. Interestingly, its security is based on an uncommon difficult problem which:

Section Finding Problem (SFP). Given an algebraic surface defined by the polynomial $X(x, y, t) \in \mathsf{GF}_p[x, y, t]$ (where GF_p denotes the finite field of cardinality p), find two polynomials $u_x(t), u_y(t) \in \mathsf{GF}_p[t]$ of degree d, such that $X(u_x(t), u_y(t), t) = 0$.

As stated in [AGM09], this problem is computationally hard: the only algorithm known so far induces to find roots of a huge multivariate polynomial system. Hence the idea of ASC is to use the surface as public key and the knowledge of a section of this surface as the trapdoor. In comparison to HFE [Pat96] or other multivariate systems, ASC has several interesting and unusual properties. In particular, the keys are very short. The security of multivariate systems is usually related to the difficulty of finding a zero of a system of low degree polynomials (often quadratic) in a huge number of variables. For instance, in the case of HFE, the size of the public key is precisely the size of the multivariate system: 265680 bits for a security of 2^{80}. In contrast with HFE, ASC enjoys a small public key of 500 bits for a security of 2^{102}. More generally, for a security level of 2^d, the size of the public key of HFE is $O(d^3)$. In comparison, the public key of ASC is a unique high degree polynomial in only three variables: its size is $O(d)$ bits for a security of 2^d. Actually, the designers explain that the keys of ASC are among the shortest of known post-quantum cryptosystems. More precisely, let w denote the degree of the public surface X in x and y. For a security level of p^{2d}, the size of the secret key is $2d \log(p)$ bits and the size of the public key is about $wd \log(p)$. The main observation is that the sizes of the keys are linear in $d \log(p)$, which is the logarithm of the security level.

Although a completely different version of ASC [AG04] has been attacked by Ivanov and Voloch [IV09], by Uchiyama and Tokunaga [UT07] and by Iwami [Iwa07], the new version of ASC, presented at PKC'2009, is resistant to all known attacks. We would like to mention that the decryption algorithm raises some questions. Indeed, one step of this algorithm is to recover some factors of given degree D of a univariate polynomial. In order to find those factors, the designers propose to recombine the irreducible factors of the polynomial by solving a knapsack. However, this problem is known to be NP-hard [GJ79]. Therefore, it is not clear if the cryptosystem remains practical for high security parameters.

Main results. We describe a message recovery attack which can break ASC in polynomial time. One important step of the legal decryption algorithm is the factorization of a univariate polynomial. The key idea of the algebraic attack is to perform this factorization step *implicitly* by decomposing ideals deduced from the ciphertext. Indeed, decomposition of ideals can be seen as a generalization of the standard factorization of polynomials. Hence, this technique allows us to bypass the Section Finding Problem, which is hard.

We present three versions of this attack. The Level 1 Attack is high-level, deterministic, offers a good view of the mechanisms involved, and can be implemented straightforwardly into a Computer Algebra System such as MAGMA (code given in Section 8.1.11). However, this version is not very efficient and cannot break ASC for the recommended parameters. The Level 2 Attack is based on the following observation: the polynomials occurring in ASC have a high degree in t and rather low degrees in x and y. Thus, it is natural to see expressions in t as *coefficients* instead of polynomials in t; in other words, in order to speed up the attack, we have to perform the computations in the ring $\mathsf{GF}_p(t)[x, y]$ (where $\mathsf{GF}_p(t)$ is the field of fractions of $\mathsf{GF}_p[t]$) instead of $\mathsf{GF}_p[x, y, t]$. In the Level 3 Attack, we replace the ground field $\mathsf{GF}_p(t)$ by a finite field $\mathsf{GF}_{p^D} \approx \mathsf{GF}_p[t]/(P(t))$ for a large enough D to avoid the swelling of the intermediate coefficients and to recover the initial message modulo $P(t)$. Even more efficiently, we can split $P(t)$ into several irreducible factors $P_i(t)$ of small degree; the Chinese Remainder Theorem is then used to recombine the congruences and retrieve the original message. In this third version of the attack, the size of the plaintext determines the number of congruences required as well as the size of the finite fields considered. Therefore, the complexity of the Level 3 Attack is expected to be *quasi-linear* in the size of the secret key. This behavior is confirmed by experimental results together with a complexity analysis. The binary complexity[1] of the Level 3 Attack is (Theorem 8.11):

$$\widetilde{O}(w^{2\vartheta}\,\mathrm{size}(m))$$

where $\mathrm{size}(m)$ denotes the binary size of the plaintext, w is the degree of X in the variables x and y and $\widetilde{O}()$ is the "soft Oh" notation (see e.g. [VZGG03, Definition 25.8]). Since the size of the secret key is smaller than $\mathrm{size}(m)$, the attack is also quasi-linear in the size of the secret key. In practice, $\mathrm{size}(m) \approx dw\log(p)$ (where d is the degree of the secret section). Thus the complexity of the attack is

$$\widetilde{O}(w^7 d\log(p)).$$

This can be compared with a lower bound on the binary complexity (see page 184) of the decryption algorithm:

$$\widetilde{O}(\log(p)(w^{\vartheta}d^{\vartheta} + dw\log(p))).$$

It can be noted that the decryption algorithm is cubic in the size of the secret key. Therefore, increasing the size of the secret key does not secure the system, since the cost of the decryption algorithm increases faster than the cost of the attack.

We implemented in MAGMA 2.15-7 the three variants. The Level 3 Attack can break ASC with parameters recommended in [AGM09] ($d = 50$, $p = 2$, $w = 5$) in only 0.05 seconds. Experiments confirm that increasing the size of the secret key with the parameters p and d does not really increase the security of the system. We are still able to break it in few seconds, even when the size of the secret key is more than 10000 bits! We also try to increase the parameter w (the degree in x and y of the public surface). For a reasonable size of the public key (less than 4000 bits), the message can be recovered in few hours. Finally, we try to figure out whether it is possible to secure the system by increasing the size of the support of the surface (the parameter k). However, as predicted by the complexity analysis, this parameter has very few effect on the complexity of the attack.

Structure of this section. After this introduction, this section is organized as follows. In Section 8.1.2, we give a short description of the ASC cryptosystem as it is presented in [AGM09]. Then, we explain the theoretical foundations of the attack. In Section 8.1.3, we describe the three variants of the attack and we show a concrete example by applying it to the toy example given in [AGM09]. We also perform a precise complexity analysis in Section 8.1.7. Finally, we give some experimental results showing that the attack is scalable.

8.1.2 Description of the cryptosystem

We give here a short description of ASC. For a more detailed presentation of this cryptosystem, we refer the reader to [AGM09]. We consider the ring of polynomials $\mathsf{GF}_p[x, y, t]$ where p is a prime number. In Section 8.1.10, a concrete example of encryption/decryption on a toy example is given. For any polynomial $P \in \mathsf{GF}_p[x, y, t]$, Λ_P denotes its support in $\mathsf{GF}_p(t)[x, y]$ (that is to say the set of couples $(i, j) \in \mathbb{N}^2$ such that $t^\ell x^i y^j$ is a monomial of P).

[1] The binary complexity is the number of arithmetic operations on bits, whereas the arithmetic complexity is the number of arithmetic operations in the base ring.

Parameters. The cryptosystem ASC has four parameters. The most important security parameters are p the cardinality of the ground field, and d the degree of the secret section. These two parameters are especially important for the security. They have a direct impact on the binary size of the secret key, which is $2d \log p$. Another parameter is w the degree in x and y of the public surface X. The last parameter is k, the cardinality of Λ_X (which is the support of X in $\mathsf{GF}_p(t)[x, y]$). The parameters w, d and p have an impact on the size of the public key which is approximatively $dw \log(p)$ bits.

Keys. The secret key is a pair of polynomials $(u_x(t), u_y(t)) \in \mathsf{GF}_p[t]$ of degree d.
The public key is given by:

- A surface described by an irreducible polynomial $X(x, y, t) \in \mathsf{GF}_p[x, y, t]$ such that $X(u_x(t), u_y(t), t) = 0$ and $\mathrm{card}(\Lambda_X) = k$.
- Λ_m the support of the plaintext polynomial and $\{d_{ij}^{(m)} \in \mathbb{N}\}_{(i,j)\in\Lambda_m}$ the degrees of the coefficients (in $\mathsf{GF}_p[t]$).
- Λ_f the support of the so-called *divisor polynomial* and $\{d_{ij}^{(f)} \in \mathbb{N}\}_{(i,j)\in\Lambda_f}$ the degrees of the coefficients (in $\mathsf{GF}_p[t]$).

For encryption/decryption it is required that:

$$\begin{gathered}\Lambda_m \subset \Lambda_f\Lambda_X = \{(i_1 + i_2, j_1 + j_2) : (i_1, j_1) \in \Lambda_f, (i_2, j_2) \in \Lambda_X\}.\\ \max\{i : (i, j) \in \Lambda_X\} < \max\{i : (i, j) \in \Lambda_m\} < \max\{i : (i, j) \in \Lambda_f\}.\\ \max\{j : (i, j) \in \Lambda_X\} < \max\{j : (i, j) \in \Lambda_m\} < \max\{j : (i, j) \in \Lambda_f\}.\\ \deg_t(X(x, y, t)) < \max\{d_{ij}^{(m)}\}_{(i,j)\in\Lambda_m} < \max\{d_{ij}^{(f)}\}_{(i,j)\in\Lambda_f}.\end{gathered}$$

Encryption. Consider a plaintext embedded into a polynomial

$$m(x, y, t) = \sum_{(i,j)\in\Lambda_m} m_{ij}(t)x^i y^j$$

where $\deg(m_{ij}(t)) = d_{ij}^{(m)}$. Choose a random *divisor polynomial*

$$f(x, y, t) = \sum_{(i,j)\in\Lambda_f} f_{ij}(t)x^i y^j$$

where $\deg(f_{ij}(t)) = d_{ij}^{(f)}$. Then select four random polynomials r_0, r_1, s_0, s_1 such that, for $\ell \in \{0, 1\}$,

$$r_\ell(x, y, t) = \sum_{(i,j)\in\Lambda_f} r_{ij}^{(\ell)}(t)x^i y^j, \quad s_\ell(x, y, t) = \sum_{(i,j)\in\Lambda_X} s_{ij}^{(\ell)}(t)x^i y^j$$

and for all i, j, $\deg(r_{ij}^{(\ell)}(t)) = \deg(f_{ij}(t)), \deg(s_{ij}^{(\ell)}(t)) = \deg(X_{ij}(t))$. Finally, construct the ciphertext $(F_0(x, y, t), F_1(x, y, t))$ where

$$\begin{aligned}F_0(x, y, t) &= m(x, y, t) + f(x, y, t)s_0(x, y, t) + X(x, y, t)r_0(x, y, t),\\ F_1(x, y, t) &= m(x, y, t) + f(x, y, t)s_1(x, y, t) + X(x, y, t)r_1(x, y, t).\end{aligned}$$

Decryption. For $\ell \in \{0, 1\}$, consider $h_\ell(t) = F_\ell(u_x(t), u_y(t), t)$ and compute the difference $h_0(t) - h_1(t) = f(u_x(t), u_y(t), t)(s_0(u_x(t), u_y(t), t) - s_1(u_x(t), u_y(t), t))$. Next, find a factor of $h_0(t) - h_1(t)$ whose degree matches $\deg(f(u_x(t), u_y(t), t))$. Let $\widetilde{f}(t)$ denote this factor. Then compute $\widetilde{m}(u_x(t), u_y(t), t) = h_0(t) \mod \widetilde{f}(t)$. Finally, retrieve $\widetilde{m}(x, y, t)$ by solving the linear system:

$$\widetilde{m}(u_x(t), u_y(t), t) = \sum \widetilde{m}_{ijk} u_x(t)^i u_y(t)^j t^k.$$

There are potentially several factors of $h_0(t) - h_1(t)$ whose degree is equal to $\deg(f(u_x(t), u_y(t), t))$. So, we have to verify that we picked the good one. To do so, the designers of ASC propose to use a Message Authentication Code (roughly speaking a cryptographic hash function with a key) to verify that $\widetilde{m}(x, y, t) = m(x, y, t)$. If the verification fails, we start again by considering another factor of $h_0(t) - h_1(t)$.

To find factors of $h_0(t) - h_1(t)$ whose degree matches $\deg(f(u_x(t), u_y(t), t))$, the designers propose to factor $h_0(t) - h_1(t)$, then recombine its irreducible factors by solving a knapsack problem. However, the knapsack problem is NP-hard [GJ79]. Therefore, as pointed out in [AGM09], it is not clear if the decryption algorithm remains practicable when the security parameters are high.

Security of the system. The designers of the cryptosystem propose the following parameters:

- $p = 2$;
- d should be greater than 50;
- $w = \deg_{xy}(X) = \max\{i + j : (i, j) \in \Lambda_X\}$ should be greater than 5;
- The lower bound on k is 3.

The size of the secret key is around 100 bits and the size of the public key is close to 500 bits. According to the designers of ASC, there is so far no known attack faster than exhaustive search for these parameters. Therefore, the security level of ASC is expected to be the cost of exhaustive search of the secret key, namely p^{2d+2}.

8.1.3 Description of the attack

Overview of the attack. In this section, we propose a message recovery attack on the cryptosystem described above.

The main point of the attack is to decompose ideals, instead of factoring the univariate polynomial obtained by evaluating $F_0 - F_1$ in the section (u_x, u_y). This way, we can implicitly manipulate the so-called *divisor polynomial* f occurring in the decryption process. Consequently, we can avoid to solve the underlying Section Finding Problem, and we obtain an attack on ASC in polynomial complexity.

First, we present a high-level and deterministic version of the attack (Algorithm 12) based on two fundamental lemmas. Then, the algorithm is speeded-up by computing in the field of fractions $\mathsf{GF}_p(t)$ (Algorithm 13). Indeed, polynomials occurring in ASC have a high degree in t. Since the complexity of Gröbner bases algorithms is linear in the complexity of the arithmetic in the ground field, it is natural to compute in the field of fractions $\mathsf{GF}_p(t)$. Finally, we use a modular approach to implement efficiently the attack: we perform computations in some well-chosen finite fields $\mathsf{GF}_p[t]/(P)$ and recombine the results by using the Chinese Remainder Theorem (Algorithm 14). Doing this, the size of the coefficients of intermediate values are bounded (these coefficients can be huge when computations are performed in the field of fractions). This allows us to break bigger instances of ASC. In particular, we are able to break the system with recommended parameters in 0.05 seconds. Furthermore, this will allow us to perform a precise complexity analysis and to show that this attack is quasi-linear in the size of the secret key. Experimentally, we are able to break with this technique some instances where the size of the secret key is greater than 10000 bits.

Now we compare the efficiency of the three versions of the attack on a small example. For instance, we consider the following parameters $p = 11$, $d = 8$, $w = 5$ and $k = 3$ and we use our MAGMA implementation. The Level 1 Attack (code given in Section 8.1.11) recovers the plaintext in 136 seconds. As predicted, the Level 2 Attack is faster and can break the system in 74 seconds. Using the modular approach in the Level 3 Attack really speeds up the computations: it retrieves the plaintext in 0.05 seconds.

8.1.4 Level 1 Attack: decomposition of ideals.

The two following lemmas are the key elements of the attack.

Lemma 8.1. *Let I be the ideal $I = \langle F_0 - F_1, X\rangle \subset \mathsf{GF}_p[x, y, t]$. Then $I = I_1 \cap I_2$ where $I_1 = \langle f, X\rangle$ and $I_2 = \langle s_0 - s_1, X\rangle$. Generically, the ideals I_1 and I_2 are prime ideals of $\mathsf{GF}_p[x, y, t]$.*

Proof. $I = \langle F_0 - F_1, X\rangle = \langle f\,(s_0 - s_1), X\rangle = I_1 \cap I_2$. □

Lemma 8.1 shows that, once we managed to decompose the ideal $\langle F_0 - F_1, X\rangle = \langle f\,(s_0 - s_1), X\rangle$, we can manipulate implicitly the polynomial f through I_1.

Remark 8.2. *In order to decompose I, a strategy is to eliminate x from I by computing a Gröbner basis of $I \cap \mathsf{GF}_p[y, t]$. Generically, this Gröbner basis contains only one polynomial Q. If p is big enough, Q has in general two factors which depend on y and t (we do not consider the factors which are in $\mathsf{GF}_p[t]$). This fact is confirmed experimentally. The two factors correspond to I_1 and I_2. Then, we can construct I_1 (resp. I_2) by adding to I an appropriate factor of Q. Since $\deg_y(f) > \deg_y(s_1 - s_0)$, the factor of Q with the highest degree in y is the one corresponding to I_1. To factor efficiently the bivariate polynomial Q, we can use for instance the algorithm in [Lec10].*

Lemma 8.3. *Let J be the ideal of $\mathsf{GF}_p[x, y, t]$ generated by $J = \langle F_0, F_1, X\rangle + I_1$. Then $m(x, y, t) \in J$. Moreover, J is a zero-dimensional ideal.*

Proof. $J = \langle F_0, F_1, X\rangle + I_1 = \langle F_0, F_1, X, f\rangle = \langle m, f, X\rangle$. □

Remark 8.4. *Lemma 8.3 shows that we can compute explicitly a multivariate ideal which contains m. Since we know Λ_m, we can recover m by solving the following linear system:*

$$\mathrm{NF}_J(m) = \sum_{(i,j)\in\Lambda_m} \sum_{k=0}^{d_{ij}^{(m)}} m_{ijk}\,\mathrm{NF}_J(x^i y^j t^k) = 0$$

where NF_J denotes the normal form with respect to the ideal J for a chosen monomial ordering. Since $\lambda m \in J$ for all $\lambda \in \mathsf{GF}_p$, we retrieve m up to multiplication by a scalar.

Remark 8.5. *For efficiency purpose, we compute the Gröbner bases with respect to the* graded reverse lexicographical ordering *(Definition 1.19). Instead of computing the Gröbner basis of $\langle F_0 - F_1, X\rangle \cap \mathsf{GF}_p[y, t]$, it is also possible to compute a resultant to eliminate the variable x.*

Algorithm 12 Level 1 Attack.

1: Compute a Gröbner basis of the ideal $\langle F_0 - F_1, X\rangle \cap \mathsf{GF}_p[y,t]$. Generically this Gröbner basis contains only one polynomial $Q(y,t)$.
2: Factor $Q = \prod Q_i(y,t)$. Let $Q_0(y,t) \in \mathsf{GF}_p[y,t]$ denote an irreducible factor with highest degree with respect to y.
3: Compute a Gröbner basis of the ideal $J = \langle F_0, F_1, X, Q_0\rangle$.
4: To retrieve the plaintext (up to multiplication by a scalar in GF_p), solve the linear system over GF_p

$$\sum_{(i,j)\in\Lambda_m} \sum_{k=0}^{d_{ij}^{(m)}} m_{ijk}\,\mathrm{NF}_J(x^i y^j t^k) = 0.$$

If the system has no solution, go back to Step 2 and pick another factor of Q.

Remark 8.6. *The normal form* NF_J *is a linear application from* $\mathsf{GF}_p[x,y,t]$ *onto* $\mathsf{GF}_p[x,y,t]/J$*. In the last step of the attack, we are searching for the intersection of its kernel with the* GF_p*-linear subspace generated by* Γ_m *(where* Γ_m *denotes the support of* m *in* $\mathsf{GF}_p[x,y,t]$*). Therefore, the linear system has* $\mathrm{card}(\Gamma_m)$ *unknowns and* $\deg(J)$ *equations (*$\deg(J) = \dim(\mathsf{GF}_p[x,y,t]/J)$ *when* $\mathsf{GF}_p[x,y,t]/J$ *is seen as a* GF_p*-vector space). From the Bézout bound [Laz83],* $\deg(J) \approx \deg(m)\deg(X)\deg(f)$*. The decryption algorithm requires that* $\deg(m(u_x,u_y,t)) \geq \mathrm{card}(\Gamma_m)$ *(in order to solve the final linear system) and one can remark that* $\deg(X)\deg(f) > \deg(m(u_x,u_y,t)) \approx d\deg_{xy}(m) + \deg_t(m)$ *(since* $\deg_{xy}(f) > \deg_{xy}(m), \deg_t(f) > \deg_t(m)$ *and* $\deg(X) > d$*). Therefore, the linear system has more equations than unknowns:* $\mathrm{card}(\Gamma_m) \leq \deg(m(u_x,u_y,t)) \leq \deg(X)\deg(f) \leq \deg(J)$.

8.1.5 Level 2 Attack: computing in the field of fractions $\mathsf{GF}_p(t)$

Polynomials appearing in ASC have a high total degree, but their degree in the variables x and y is low. Hence, it is natural to consider these polynomials as bivariate polynomials in x and y over the field of fractions $\mathsf{GF}_p(t)$. Indeed, the degrees in x and y are completely independent of the security parameter d. In this section, we explain how to adapt the attack in this context. Doing that, we expect to have a lower complexity. Indeed, many operations on ideals – for instance Gröbner basis computations – are linear in the complexity of the arithmetic in the ground field.

From now on, $\mathbb{K}$ denotes the field of fractions $\mathsf{GF}_p(t)$.

First, we need to transpose the key lemmas in this new context. This can be done for Lemma 8.1 without any major modification:

Lemma 8.7. *Let* I *be the ideal* $I = \langle F_0 - F_1, X\rangle$ *(seen as an ideal of* $\mathbb{K}[x,y]$*). Then there exist* I_1 *and* I_2 *two proper ideals of* $\mathbb{K}[x,y]$ *such that* $I = I_1 \cap I_2$ *and* $\langle f, X\rangle \subset I_1$*.*

Unfortunately, Lemma 8.3 cannot be directly transposed in the context of the field of fractions. Indeed, the variety of the ideal $J = \langle F_0, F_1, X\rangle + I_1 = \langle m, f, X\rangle$ (seen as an ideal of $\mathbb{K}[x,y]$) is generically empty since it is generated by three independent equations. Therefore we have to introduce a new variable z if we want to keep the ideal zero-dimensional and strictly included in $\mathbb{K}[x,y,z]$. Roughly speaking, the role of z is to deform the ideal $\langle m, f, X\rangle$ in order to introduce new elements in the variety:

Lemma 8.8. *Let* $J \subset \mathbb{K}[x,y,z]$ *be the ideal* $J = \langle F_0 + z, F_1 + z, X\rangle + I_1$*. Then* $m(x,y,t) + z \in J$*. Moreover,* J *is a zero-dimensional ideal.*

Proof. $\langle F_0 + z, F_1 + z, X\rangle + I_1 = \langle F_0 + z, F_1 + z, X, f\rangle = \langle m + z, f, X\rangle$. □

Algorithm 13 Level 2 Attack: computing in the field of fractions $\mathbb{K} = \mathsf{GF}_p(t)$.

1: Compute the resultant $\mathsf{Res}_x(F_0 - F_1, X) \in \mathbb{K}[y]$.
2: Factor the resultant $\mathsf{Res}_x(F_0 - F_1, X) = \prod Q_i(y)$. Let $Q_0(y) \in \mathbb{K}[y]$ denote an irreducible factor of highest degree in y.
3: Compute a grevlex-Gröbner basis of the ideal $J = \langle F_0 + z, F_1 + z, X, Q_0\rangle \subset \mathbb{K}[x, y, z]$.
4: Consider the following linear system over $\mathbb{K}$:

$$\mathrm{NF}_J(z) + \sum_{(i,j)\in\Lambda_m} m_{ij}(t)\,\mathrm{NF}_J(x^i y^j) = 0.$$

If the system has no solution, then go back to Step 2 and choose another factor of the resultant.
5: Return $m = \sum_{(i,j)\in\Lambda_m} m_{ij}(t)x^i y^j$ where $(m_{ij}(t))$ is the unique solution of the linear system.

To be able to recover the plaintext, we need to solve a linear system with $\mathrm{card}(\Lambda_m)$ unknowns and $\deg(J)$ equations. In practice, there are more equations than unknowns. Thus, if we choose a wrong factor of the resultant (a factor which is not a divisor of f), then the linear system has generically no solution, and we just have to restart from Step 2 until we find an appropriate factor. In practice, the irreducible factor of the resultant with the highest degree in y is almost always a good choice.

Remark 8.9. *It is also possible to combine the two versions of the attack by computing a Gröbner basis of the elimination ideal and factoring it in* $\mathsf{GF}_p[x, y, t]$*, as in Level 1 attack (Steps* 1 *and* 2 *in Algorithm 12). Then, once we found* $Q_0 \in \mathsf{GF}_p[x, y, t]$*, we retrieve the message by computing a Gröbner Basis of* $J = \langle F_0 + z, F_1 + z, X, Q_0\rangle \subset \mathbb{K}[x, y, z]$ *in the field of fractions (Steps* $3, 4, 5$ *in Algorithm 13).*

8.1.6 Level 3 Attack: computing in finite fields GF_{p^m}

In this section, we study how to implement efficiently the attack in practice. In order to speed up the attack and to compute efficiently in the field of fractions, we perform all computations modulo polynomials of $\mathsf{GF}_p[t]$. Indeed, a bound on the degree of m with respect to t is known since $\deg_t(m) \leq \max\{d_{i,j}^{(m)}\}$.

We choose a constant C and $n = \deg_t(m)\log(p)/C$ irreducible polynomials $P_1, \dots, P_n$ of degree close to $C/\log(p)$ such that $\sum \deg(P_i) > \deg_t(m)$. Then for each P_i, we consider

$$\mathsf{GF}_p[t]/(P_i) = \mathsf{GF}_{p^{\deg(P_i)}}\ .$$

Considering all computations in $\mathbb{K} = \mathsf{GF}_p[t]/(P_i)$ instead of $\mathsf{GF}_p(t)$, the attack yields $m \mod P_i$. Finally we use the Chinese Remainder Theorem (CRT) to recover $m \mod \prod P_i$. Since $\deg(\prod P_i) > \deg_t(m)$, we retrieve the plaintext.

Remark 8.10. *The linear system at step 7 in Algorithm 14 has only* $\mathrm{card}(\Lambda_m)$ *unknowns and* $\deg(J) \approx \deg_{xy}(m)\deg_{xy}(f)\deg_{xy}(X)$ *equations. For practical parameters,* $\mathrm{card}(\Lambda_m) \approx k$ *is smaller than* $\deg(J)$*, thus the linear system is overdetermined and has in general only one solution. This fact is confirmed by experiments.*

The value $\sum \deg(P_i) \approx \deg_t(m)$ *is only dependent of the size of the plaintext. Therefore, the number of times we have to run the main loop of Algorithm 14 is linear in the size of the plaintext. Since the cost of arithmetic operations in* $\mathsf{GF}_p[t]/(P)$ *only depends on* C *(which is a constant chosen*

Algorithm 14 Level 3 Attack: computing in the finite fields $\mathbb{K} = \mathsf{GF}_p[t]/(P)$.

1: Choose $n \approx \deg_t(m)\log(p)/C$ irreducible polynomials of degree $\approx C/\log(p)$ such that $\sum \deg(P_i) > \deg_t(m)$.
2: **for** i from 1 to n **do**
3: Consider $\mathbb{K} = \mathsf{GF}_p[t]/(P_i)$.
4: Compute the resultant $\mathsf{Res}_x(F_0 - F_1, X) \in \mathbb{K}[y]$.
5: Factor the resultant $\mathsf{Res}_x(F_0 - F_1, X) = \prod Q_i(y)$. Let $Q_0(y) \in \mathbb{K}[y]$ denote an irreducible factor of highest degree in y.
6: Compute a grevlex-Gröbner basis of the ideal $J = \langle F_0 + z, F_1 + z, X, Q_0\rangle \subset \mathbb{K}[x, y, z]$.
7: Consider the following linear system over $\mathbb{K}$:

$$\mathrm{NF}_J(z) + \sum_{(i,j)\in\Lambda_m} m_{ij}(t)\,\mathrm{NF}_J(x^i y^j) = 0.$$

If the system has no solution, then go back to Step 2 and choose another factor of the resultant.
8: Retrieve a congruence $m \mod P_i = \sum_{(i,j)\in\Lambda_m} m_{ij}(t)x^i y^j$ where $(m_{ij}(t))$ is the solution of the linear system.
9: **end for**
10: Use the CRT to retrieve $m = m \mod \prod P_i$.

by the attacker), we expect this Level 3 Attack to be linear or quasi-linear in the size of the plaintext. This expectation will be confirmed by a complexity analysis and by experimental results. Besides, we would also like to mention that the main loop of Algorithm 14 can be easily computed in parallel.

A concrete example. We consider here the toy example given in [AGM09]. We have

- $p = 17$.
- The secret key is $(u_x, u_y) = (14t^3 + 12t^2 + 5t + 1, 11t^3 + 3t^2 + 5t + 4)$.
- The public surface is
  $X = (t+10)x^3y^2 + (16t^2 + 7t + 4)xy^2 + (3t^{16} + 8t^{15} + 13t^{14} + 8t^{13} + 3t^{12} + 12t^{11} + 4t^{10} + 8t^9 + 7t^8 + 4t^7 + 13t^6 + 2t^5 + 5t^4 + 4t^3 + 14t^2 + 9t + 14)$.
- The support of m and f are

$$\Lambda_m = \{(4,4),(0,0)\}, d^m_{00} = 17, d^m_{44} = 17,$$
$$\Lambda_f = \{(5,5),(1,2),(0,0)\}, d^f_{00} = 13, d^f_{12} = 11, d^f_{55} = 18.$$

Here we show how to recover the message m from the ciphertext (F_0, F_1) (given in [AGM09]) with the Level 3 Attack:

1. Since $\deg_t(m) = 17$, we choose (for instance) $P_1, P_2, P_3, P_4 \in \mathsf{GF}_p[t]$ irreducible such that $\sum \deg(P_i) \geq 18$. In particular,

$$P_1 = t^5 + t + 14,$$
$$P_2 = t^5 + 14t^4 + 4t^3 + 4t + 4,$$
$$P_3 = t^5 + 9t^4 + 15t^3 + 8t^2 + 4t + 8,$$
$$P_4 = t^5 + 11t^4 + 11t^3 + 8t^2 + 7t + 8.$$

   First, we consider the finite field $\mathbb{K} = \mathsf{GF}_p[t]/(P_1)$.

2. Compute the resultant in $\mathbb{K}[y]$:
$\mathsf{Res}_x(F_0 - F_1, X) = (9t^4+14t^3+4t^2+6t+13)y^{30} + (5t^4+t^3+14t^2+15t+8)y^{27} + (6t^4+9t^3+10t^2+7t+14)y^{26} + (7t^4+4t^3+8t^2+5t+8)y^{25} + (8t^4+4t^3+7t^2+7t+6)y^{24} + (12t^4+9t^3+8t^2+13t)y^{23} + (9t^4+4t^3+9t^2+15t+6)y^{22} + (3t^4+6t^3+10t^2+6t+6)y^{21} + (9t^4+9t^3+13t^2+15t+6)y^{20} + (4t^4+4t^3+15t^2)y^{19} + (2t^4+11t^3+2t^2+5t+2)y^{16}$.

3. Then factor it in $\mathbb{K}[y]$:
$\mathsf{Res}_x(F_0-F_1, X) = y^{16}(y+8t^4+3t^3+16t^2+8t+2)(y^2+2t^4+14t^3+14t^2+6t+10)(y^2+15t^4+3t^3+3t^2+11t+7)(y^2+(14t^4+7t^3+4t)y+13t^4+10t^3+7t^2+8t+1)(y^7+(12t^4+7t^3+t^2+5t+15)y^6+(t^4+5t^3+7t^2+12t+11)y^5+(9t^4+14t^3+5t^2+10t+10)y^4+(4t^4+7t^3+t^2+7t+14)y^3+(11t^4+13t^3+12t^2+8t+4)y^2+(15t^4+9t^3+16t^2+14t+14)y+14t^4+3t^3+9t^2+15t+8)$.

4. Consider Q_0 an irreducible factor with highest degree:
$Q_0 = y^7 + (12t^4+7t^3+t^2+5t+15)y^6 + (t^4+5t^3+7t^2+12t+11)y^5 + (9t^4+14t^3+5t^2+10t+10)y^4 + (4t^4+7t^3+t^2+7t+14)y^3 + (11t^4+13t^3+12t^2+8t+4)y^2 + (15t^4+9t^3+16t^2+14t+14)y + (14t^4+3t^3+9t^2+15t+8)$.

5. Compute a Gröbner basis G with respect to the grevlex ordering of the ideal $J = \langle F_0+z, F_1+z, X, Q_0\rangle \subset \mathbb{K}[x,y,z]$.

6. Since $\Lambda_m = \{(0,0),(4,4)\}$ compute $\mathrm{NF}_J(x^4y^4)$:
$\mathrm{NF}_J(x^4y^4) = N_1 z + N_2 = (15t^4+3t^3+t^2+13t+16)z + (5t^4+11t^2+t+7)$.

7. Solve the linear system $z + m_{44}\,\mathrm{NF}_J(x^4y^4) + m_{00} = 0$ over $\mathbb{K}$:
$$\begin{cases} m_{00} = N_2/N_1 \mod P_1 \\ m_{44} = -1/N_1 \mod P_1. \end{cases}$$

8. Recover a congruence: $m = m_{00} + m_{44}x^4y^4 \mod P_1$.

9. Repeat the process with P_2, P_3 and P_4.

10. Use the CRT to retrieve $m = m \mod \prod P_i$:
$m = (5t^{17}+15t^{16}+4t^{15}+9t^{14}+7t^{13}+2t^{12}+3t^{11}+8t^{10}+11t^9+6t^{17}+6t^8+3t^{16}+10t^7+11t^{15}+7t^6+t^5+t^{13}+14t^4+10t^{12}+3t^3+3t^{11}+12t^2+8t^{10}+11t+6t^9+2)x^4y^4 + (13t^8+2t^7+2t^6+10t^5+5t^4+2t^3+15t^2+3t+11)$.

8.1.7 Complexity analysis

In this part, we investigate the complexity of the Level 3 Attack. To simplify the notations, we suppose here that the complexity of multiplying two $n \times n$ matrices is $O(n^3)$. We note that C is a parameter chosen by the attacker. This parameter fixes the size of the finite fields considered. Indeed, we choose finite fields $\mathbb{K} = \mathsf{GF}_p/(P_i)$ with $\deg(P_i) \approx C/\log(p)$. Hence, $\log(\mathrm{card}(\mathbb{K})) \approx C$.

1. First, we estimate the complexity of the computation of the resultant with respect to x in $\mathbb{K}[x,y]$ (where $\mathbb{K} = \mathsf{GF}_p[t]/(P_i)$). According to [VZGG03] (Corollary 11.18), this can be done in $\widetilde{O}(w^{\vartheta})$ operations in $\mathbb{K}$, and the degree of the resultant is $O(w^2)$.

2. The probabilistic Cantor-Zassenhaus algorithm [VZGG03] factors a polynomial of degree n over a finite field GF_q in $\widetilde{O}(n^2+n\log(q))$ arithmetic operations in GF_q. Therefore the arithmetic complexity in $\mathbb{K}$ of the factorization of the resultant is
$$\widetilde{O}(w^4 + w^2\log(\mathrm{card}(\mathbb{K}))) = \widetilde{O}(w^4 + w^2 C).$$

3. The degree of regularity of an ideal is an important indicator of the complexity of computing its Gröbner basis with respect to the grevlex ordering: it is the highest degree of the polynomials occurring in the F_5 Algorithm. According to [Laz83, BFSY04, BFS04], if an ideal is spanned by m generic equations in n variables, then the complexity of computing a Gröbner basis is:

$$O\left(m^\vartheta \binom{\mathrm{d_{reg}} + n - 1}{n-1}^\vartheta\right).$$

   Since the ideal $J = \langle m + z, f, X\rangle$ is generated by three independent equations, its degree of regularity can be estimated from the Macaulay bound (see [Laz83]) as

$$\mathrm{d_{reg}}(J) = (\deg_{xy}(m+z) - 1) + (\deg_{xy}(f) - 1) + (\deg(X)_{xy} - 1) + 1.$$

   For practical parameters, $\deg_{xy}(m+z) \approx \deg_{xy}(f) \approx \deg(X)_{xy} \approx w$. Therefore, $\mathrm{d_{reg}} \approx 3w$. The arithmetic complexity in $\mathbb{K}$ of the Gröbner basis computation is then:

$$O\left(3^\vartheta \binom{\mathrm{d_{reg}}(J) + 2}{2}^\vartheta\right) = O(w^{2\vartheta}).$$

4. Finally we have a linear system to solve. The number of variables is $\mathrm{card}(\Lambda_m)$. For practical parameters, $\mathrm{card}(\Lambda_m) \approx k$, which is less than 1000 (the recommended parameter is $k = 3$). Hence, this step is negligible in practice compared to the Gröbner basis computation, since an overdetermined linear system with less than 1000 variables in a finite field can be easily solved. Furthermore, this step is analog to the linear system which is solved in the legal decryption algorithm. Therefore this step of the attack is faster than the decryption algorithm which has to be efficient for practical parameters.

The cost of an arithmetic operation in $\mathbb{K}$ is quasi-linear in $\log(\mathrm{card}(\mathbb{K})) \approx C$. The number of times we have to run the main loop of the attack is $\mathrm{size}(m)/C$. The complexity of the CRT is $\widetilde{O}(\mathrm{size}(m)\log(\mathrm{size}(m)))$ [VZGG03]. Putting all the steps together, we find the total complexity of the attack:

Theorem 8.11. *The total binary complexity of the Level 3 Attack is*

$$\underbrace{\widetilde{O}(\mathrm{size}(m)w^\vartheta)}_{\textit{resultant}} + \underbrace{\widetilde{O}(\mathrm{size}(m)(w^4 + w^2C))}_{\textit{factorization}} + \underbrace{\widetilde{O}(\mathrm{size}(m)w^{2\vartheta})}_{\textit{Gröbner}} + \underbrace{\widetilde{O}(\mathrm{size}(m))}_{\textit{CRT}}.$$

Hence, the total binary asymptotic complexity of the attack is bounded by

$$\widetilde{O}(w^{2\vartheta}\,\mathrm{size}(m)).$$

Corollary 8.12. *If we assume that* $\mathrm{size}(m) \approx wd\log(p)$ *(which is the case in practice), then the binary complexity of the attack is:* $\widetilde{O}(dw^{2\vartheta+1}\log(p))$.

Consequently, the attack is polynomial in all the security parameters and it is quasi-linear in the size of the secret key which is $2d\log(p)$. It can be noted that the parameter k has few effect on the complexity of the attack.

A lower bound on the complexity of the decryption algorithm.

The complexity of this attack has to be compared with a lower bound on the cost of the decryption process. During the decryption algorithm, one has to factor $(F_0 - F_1)(u_x(t), u_y(t), t)$ over $\mathsf{GF}_p[t]$. The degree of this polynomial is at least dw. To the best of our knowledge, the best probabilistic factorization algorithms have an arithmetic complexity of $\widetilde{O}(d^2w^2 + dw\log(p))$ [VZGG03]. Moreover, there is also a knapsack to solve after the factorization. The complexity of this step is difficult to estimate so we do not consider it here (remember that we try to establish a lower bound). The last step of the decryption process is the resolution of a linear system with $O(dw)$ variables: the arithmetic complexity of this step is $O(w^\vartheta d^\vartheta)$. Finally, the total binary complexity of the decryption algorithm is unsharply lower bounded by $\widetilde{O}(\log(p)(w^\vartheta d^\vartheta + dw\log(p)))$ which is cubic in the parameters d and w, and quadratic in $\log(p)$. In comparison, the attack is quasi-linear in d and $\log(p)$, and polynomial of degree $2\vartheta + 1$ in w.

8.1.8 Experimental results

Workstation.

The experimental results have been obtained with a Xeon bi-processor 3.2 GHz, with 64 GB of RAM. The instances of ASC have been generated with MAGMA2.15-7. To compute the Gröbner basis, we use the F_4 [Fau99] implementation in MAGMA.

To generate our instances, we pick $\ell, d \in \mathbb{N}$ and we consider the following parameters:

- $w = 2\ell + 5$.
- $\Lambda_m = \{(4+\ell, 4+\ell), (0,0)\}$.
- $\Lambda_X = \{(3+\ell, 2+\ell), (1+\ell, 2+\ell), (0,0)\}$.
- $\Lambda_f = \{(5+\ell, 5+\ell), (1+\ell, 2+\ell), (1,2), (0,0)\}$.
- $\forall (i,j) \in \Lambda_m, d_{ij}^{(m)} = (2\ell+5)d + 21$.
- $\forall (i,j) \in \Lambda_m, d_{ij}^{(f)} = (2\ell+5)d + 22$.

Construction of X, u_x and u_y.

$u_x, u_y \in \mathsf{GF}_p[t]$ are random polynomials of degree d.
To construct $X(x, y, t)$, we pick two random polynomials $R_1, R_2 \in \mathsf{GF}_p[t]$ of degree 20 and we consider

$$X = R_1(t)(x^{3+\ell}y^{2+\ell} - u_x(t)^{3+\ell}u_y(t)^{2+\ell}) + R_2(t)(x^{1+\ell}y^{2+\ell} - u_x(t)^{1+\ell}u_y(t)^{2+\ell}).$$

Then we verify that $X(x, y, t)$ is irreducible. If not, we restart by picking another R_1 and another R_2.

Table 8.1 shows the complexity of the Level 3 Attack for different values of p and d. Each entry in the table is obtained by considering the average results over 20 random instances of the cryptosystem.

p	d	w	k	size of public key	size of secret key	t_{res}	t_{fact}	t_{GB}	t_{total}	security bound
2	50	5	3	310 bits	102 bits	0.02s	0.02s	0.01s	0.05s	2^{102}
2	100	5	3	560 bits	202 bits	0.03s	0.02s	0.02s	0.07s	2^{202}
2	200	5	3	1060 bits	402 bits	0.05s	0.05s	0.05s	0.15s	2^{402}
2	400	5	3	2060 bits	802 bits	0.1s	0.1s	0.1s	0.30s	2^{802}
2	800	5	3	4060 bits	1602 bits	0.2s	0.2s	0.2s	0.65s	2^{1602}
2	1600	5	3	8060 bits	3202 bits	0.3s	0.3s	0.4s	1.0s	2^{3202}
2	2000	5	3	10060 bits	4002 bits	0.45s	0.4s	0.4s	1.3s	2^{4002}
2	5000	5	3	25060 bits	10002 bits	0.8s	1.3s	0.8s	3.0s	2^{10002}
17	50	5	3	1267 bits	409 bits	0.2s	2.4s	0.4s	3.0s	2^{409}
17	100	5	3	2289 bits	818 bits	0.3s	5.1s	0.6s	3.0s	2^{818}
17	400	5	3	8420 bits	3270 bits	1.45s	27.7s	3.9s	33.1s	2^{3270}
17	800	5	3	16595 bits	6500 bits	3.1s	70s	9.5s	83s	2^{6500}
10007	500	5	3	34019 bits	13289 bits	29s	217s	64s	310s	2^{13289}

Table 8.1: Level 3 Attack – Experimental results with $w = 5$

Table notations.

t_{res} denotes the time used for the computation of the resultant. t_{fact} is the time used by the factorization of the resultant, whereas t_{GB} denotes the cost of the Gröbner basis computation. The time for solving the linear system and for the recombination by the CRT is negligible and hence are not given in the table. According to [AGM09], there were no known attack better than exhaustive search when $d \geq 50$ and $w \geq 5$. Therefore the security bound is the cost of the exhaustive search of the secret section, namely p^{2d+2}.

Interpretation of the results.

It is worth remarking that the first line of Table 8.1 corresponds to the parameters recommended by the designers [AGM09] and are broken in 0.05 seconds. The major observation is that the complexity of the attack behaves as predicted by the complexity analysis: it is quasi-linear in the parameter d. We also ran some experiments to study the impact of the parameter k (the cardinality of the support of the surface X) on the complexity: as expected, increasing k has very few effect on the cost of the attack. To summarize, we see in Table 8.1 that trying to secure the system by increasing the size of the secret key (that is to say by increasing the parameters p and d) is pointless: even when the size of the secret key is bigger than 10000 bits, the system can be broken in few seconds.

The parameter w.

In order to secure the system, one can think of increasing the parameter w since the attack is in $O(w^{2\vartheta+1})$. However, we showed that the complexity decryption algorithm is lower bounded by $O(w^3)$. Consequently, the parameter w should not be too high if the owner of the secret key wants to be able to decrypt. Table 8.2 gives the experimental results of the attack when w increases.

p	d	w	k	size of public key	size of secret key	t_{res}	t_{fact}	t_{GB}	t_{LinSys}	t_{total}	security bound
2	50	5	3	310 bits	102 bits	0.02s	0.02s	0.01s	0.001s	0.05s	2^{102}
2	50	15	3	810 bits	102 bits	0.7s	0.3s	4.4s	0.03s	5.4s	2^{102}
2	50	25	3	1310 bits	102 bits	3s	1s	32s	0.2s	37s	2^{102}
2	50	35	3	1810 bits	102 bits	10s	3s	260s	1s	274s	2^{102}
2	50	45	3	2310 bits	102 bits	30s	7s	1352s	4s	1393s	2^{102}
2	50	55	3	2810 bits	102 bits	70s	12s	4619s	13s	4714s	2^{102}
2	50	65	3	3310 bits	102 bits	147s	22s	12408s	27s	12604s	2^{102}
2	50	75	3	3810 bits	102 bits	288s	38s	37900s	56s	38280s	2^{102}

Table 8.2: Level 3 Attack – Experimental results: increasing w

Interpretation of the results.

The main observation is that the complexity of the attack still behaves as predicted: when w is increased, the Gröbner basis computation is the most expensive step. Increasing w seems to be the best counter-measure against the attack. However, it should be noted that the attack is still feasible in practice, even when the public key is big.

8.1.9 Conclusion

In this section, we analyze the security of the PKC'2009 Algebraic Surface Cryptosystem. We provide three variants of a message recovery attack. We also estimate very precisely the complexity of the Level 3 Attack and we show that it is polynomial in all the parameters of the system. Furthermore, it is quasi-linear in the size of the secret key, whereas the decryption algorithm proposed in [AGM09] is cubic.

Experimental results confirm the theoretical analysis. We show that the attack can easily break ASC with recommended parameters. The best choice to try to secure ASC against the attack is to take p and d as small as possible ($p = 2$ and $d = 50$) and increase w. However our implementation is polynomial in w and can break the system in few hours, even when $w = 75$ (this value can be compared to the initial recommended $w = 5$).

Thereby, we consider that the system is fully broken, but we believe that the section finding problem is still an interesting problem for the design of cryptographic schemes; in this section, we have simply shown how to avoid to solve it in the context of ASC.

8.1.10 Toy example

We describe here the toy example given in [AGM09]:

- $\mathbb{K} = \mathbb{F}_{17}$.
- $w = 5$.
- $d = 3$.
- $k = 5$.

The public surface is
$X(x,y,t) = (t+10)x^3y^2 + (16t^2+7t+4)xy^2 + 3t^{16} + 8t^{15} + 13t^{14} + 8t^{13} + 3t^{12} + 12t^{11} + 4t^{10} + 8t^9 + 7t^8 + 4t^7 + 13t^6 + 2t^5 + 5t^4 + 4t^3 + 14t^2 + 9t + 14.$

and the secret keys are

$$u_x(t) = 14t^3 + 12t^2 + 5t + 1,$$

$$u_y(t) = 11t^3 + 3t^2 + 5t + 4.$$

The support of m and f are

$$\Lambda_m = \{(4,4),(0,0)\}, d^m_{00} = 17, d^m_{44} = 17,$$

$$\Lambda_f = \{(5,5),(1,2),(0,0)\}, d^f_{00} = 13, d^f_{12} = 11, d^f_{55} = 18.$$

Encryption

We consider the following plaintext: $m(x,y,t) = (5t^{17} + 15t^{16} + 4t^{15} + 9t^{14} + 7t^{13} + 2t^{12} + 3t^{11} + 8t^{10} + 11t^9 + 6t^8 + 10t^7 + 7t^6 + t^5 + 14t^4 + 3t^3 + 12t^2 + 11t + 2)x^4y^4 + 6t^{17} + 3t^{16} + 11t^{15} + t^{13} + 10t^{12} + 3t^{11} + 8t^{10} + 6t^9 + 13t^8 + 2t^7 + 2t^6 + 10t^5 + 5t^4 + 2t^3 + 15t^2 + 3t + 11.$

In order to encrypt, randomly pick f, s_1, s_2, r_1, r_2 with support fixed by Λ_f and Λ_X:
$f(x,y,t) = (t^{18} + 8t^{17} + 8t^{16} + 6t^{15} + 3t^{14} + 11t^{13} + 12t^{12} + 9t^{11} + 14t^{10} + 8t^9 + 11t^8 + 10t^7 + 7t^6 + 8t^5 + 16t^4 + 10t^3 + 12t^2 + 7t + 16)x^5y^5 + (7t^{11} + 2t^{10} + 16t^9 + 16t^8 + 2t^7 + 4t^6 + 4t^5 + 9t^4 + 9t^3 + t^2 + 7t + 14)xy^2 + 8t^{13} + 12t^{12} + 15t^{11} + 5t^9 + 12t^8 + 13t^7 + 6t^6 + 6t^5 + 2t^4 + 13t^3 + 14t^2 + 14t + 11.$

$s_0(x,y,t) = (4t+2)x^3y^2 + (16t^2 + 9t + 4)xy^2 + 8t^{16} + 4t^{15} + 11t^{14} + 7t^{13} + t^{12} + 11t^{10} + 8t^9 + 13t^8 + 12t^7 + 14t^6 + 16t^5 + 8t^4 + 13t^3 + 16t^2 + 14t + 4.$

$s_1(x,y,t) = (7t+11)x^3y^2 + (11t^2 + 3t + 3)xy^2 + t^{16} + 3t^{15} + 13t^{14} + t^{13} + 3t^{12} + 16t^{11} + 9t^{10} + 4t^9 + 12t^7 + t^6 + 7t^5 + t^4 + 4t^3 + 2t + 1.$

$r_0(x,y,t) = (10t^{18} + 3t^{17} + 7t^{16} + t^{15} + 10t^{14} + 10t^{13} + 5t^{12} + 7t^{11} + 15t^{10} + 10t^9 + 8t^8 + 2t^7 + 16t^6 + 4t^4 + t^3 + 3t^2 + 16t + 2)x^5y^5 + (t^{11} + 10t^{10} + 14t^9 + 10t^8 + 2t^7 + 4t^6 + 13t^5 + 6t^4 + 10t^3 + 10t^2 + 4t + 15)xy^2 + 5t^{13} + 16t^{12} + t^{11} + 8t^{10} + 8t^9 + 3t^8 + 3t^7 + 5t^6 + 3t^5 + 3t^4 + 9t^3 + 7t^2 + t + 15.$

$r_1(x,y,t) = (12t^{18} + 2t^{17} + 7t^{16} + 6t^{15} + 8t^{14} + 9t^{13} + 16t^{12} + 4t^{11} + 8t^8 + 8t^7 + 10t^6 + 13t^5 + 12t^4 + 11t^3 + 8t^2 + 4t + 16)x^5y^5 + (t^{11} + 8t^{10} + 2t^9 + t^8 + 4t^7 + 2t^6 + 8t^5 + 4t^4 + 13t^3 + 15t^2 + 2t + 8)xy^2 + 16t^{13} + 6t^{12} + t^{11} + 11t^{10} + 16t^9 + 4t^8 + 2t^7 + 14t^6 + 3t^5 + 7t^4 + 13t^3 + 13t^2 + 8t + 16.$

Then compute $F_i = m + s_i f + r_i X$:

$F_0(x,y,t) = (14t^{19} + t^{18} + 9t^{16} + 10t^{15} + 7t^{14} + 5t^{13} + 15t^{12} + 6t^{11} + 16t^{10} + 15t^9 + 8t^8 + 16t^7 + 2t^6 + 16t^5 + 11t^4 + 13t^3 + 13t^2 + 2t + 1)x^8y^7 + (6t^{20} + 3t^{18} + 5t^{17} + 6t^{16} + 2t^{15} + 7t^{13} + 16t^{12} + 5t^{11} + t^{10} + 11t^9 + 4t^8 + 11t^7 + 8t^6 + 6t^5 + 9t^4 + 14t^3 + 13t^2 + 12t + 4)x^6y^7 + (4t^{34} + 4t^{33} + 10t^{32} + 13t^{31} + 2t^{30} + 11t^{29} + 3t^{28} + 15t^{27} + 7t^{25} + 13t^{24} + 4t^{23} + 6t^{21} + 4t^{20} + t^{18} + 15t^{17} + 6t^{16} + 16t^{15} + 15t^{14} + 7t^{13} + 14t^{11} + 12t^{10} + 8t^9 + 9t^8 + 6t^7 + 6t^6 + 10t^5 + 14t^4 + 2t^3 + 4t^2 + t + 7)x^5y^5 + (5t^{17} + 15t^{16} + 4t^{15} + 9t^{14} + 7t^{13} + 14t^{12} + 11t^{11} + 3t^{10} + 2t^9 + 12t^8 + 3t^7 + 16t^6 + 11t^5 + 2t^4 + 16t^3 + 10t^2 + 10)x^4y^4 + (3t^{14} + 11t^{13} + 7t^{12} + 14t^{11} + 6t^{10} + 5t^9 + 7t^8 + 4t^6 + 2t^5 + 10t^4 + 9t^3 + 2t^2 + 12t + 2)x^3y^2 + (9t^{13} + 7t^{12} + 5t^{11} + 9t^{10} + 7t^9 + 9t^8 + 12t^7 + 8t^6 + 2t^5 + 13t^4 + 8t^3 + 4t^2 + 3t + 14)x^2y^4 + (8t^{27} + 14t^{26} + 8t^{25} + 16t^{24} + 16t^{23} + 13t^{22} + 6t^{21} + 13t^{20} + 10t^{19} + 4t^{18} + 10t^{17} + 10t^{16} + 13t^{15} + 11t^{14} + 14t^{13} + 14t^{12} + 15t^{11} + 4t^{10} + 11t^9 + 13t^8 + 5t^7 + 4t^6 + 10t^5 + 13t^4 + 3t^3 + 2t^2 + 16t + 13)xy^2 + 11t^{29} + 12t^{28} + 10t^{27} + t^{26} + 14t^{25} + 16t^{24} + 12t^{23} + 14t^{22} + 14t^{21} + 11t^{20} + 7t^{19} + 15t^{18} + 6t^{17} + 16t^{16} + 15t^{15} + 10t^{14} + 4t^{13} + 7t^{12} + 16t^{11} + 11t^{10} + 8t^9 + 2t^8 + 16t^7 + t^6 + 12t^5 + 3t^4 + 13t^3 + 12t^2 + 5t + 10.$

$F_1(x,y,t) = (2t^{19} + 2t^{18} + t^{17} + 2t^{16} + 2t^{15} + 12t^{14} + 5t^{13} + 2t^{12} + 16t^{11} + 6t^{10} + 3t^9 + 7t^8 + 11t^7 + 8t^6 + 2t^5 + 3t^4 + 6t^3 + 10t^2 + 7t + 13)x^8y^7 + (16t^{20} + 3t^{19} + 12t^{17} + t^{16} + 15t^{15} + 15t^{14} + 6t^{13} + 3t^{12} + 3t^{11} + 9t^{10} + 11t^9 + 14t^8 + 7t^7 + t^5 + 4t^4 + t^3 + 5t^2 + 10t + 10)x^6y^7 + (3t^{34} + 11t^{33} + 8t^{31} + 11t^{30} + 11t^{29} + 4t^{28} + 5t^{27} + t^{26} + 4t^{25} + 3t^{24} + 9t^{23} + 5t^{22} + 7t^{21} + 16t^{20} + 4t^{19} + 10t^{18} +$

$7t^{17}+9t^{16}+15t^{15}+13t^{14}+8t^{13}+9t^{12}+10t^{11}+10t^{10}+3t^{9}+14t^{7}+15t^{6}+4t^{5}+11t^{4}+2t^{3}+7t^{2}+t+2)x^{5}y^{5}+(5t^{17}+15t^{16}+4t^{15}+9t^{14}+7t^{13}+t^{12}+10t^{11}+3t^{10}+14t^{9}+6t^{8}+5t^{6}+5t^{5}+8t^{4}+16t^{3}+3t^{2}+10t+15)x^{4}y^{4}+(4t^{14}+15t^{13}+9t^{12}+16t^{11}+8t^{10}+14t^{9}+10t^{8}+15t^{7}+13t^{6}+15t^{5}+9t^{4}+10t^{3}+16t^{2}+4t+9)x^{3}y^{2}+(8t^{13}+8t^{12}+6t^{11}+3t^{10}+10t^{9}+9t^{8}+16t^{7}+13t^{6}+15t^{5}+4t^{4}+7t^{3}+6t^{2}+8t+6)x^{2}y^{4}+(10t^{27}+4t^{26}+9t^{25}+7t^{24}+3t^{23}+13t^{22}+16t^{21}+14t^{20}+t^{19}+t^{17}+6t^{16}+11t^{15}+9t^{14}+2t^{13}+16t^{12}+9t^{11}+16t^{10}+13t^{9}+2t^{7}+2t^{6}+14t^{5}+6t^{4}+15t^{3}+6t^{2}+14t+2)xy^{2}+5t^{29}+12t^{28}+6t^{27}+14t^{26}+5t^{25}+10t^{24}+12t^{23}+t^{22}+8t^{21}+2t^{20}+15t^{19}+3t^{18}+5t^{17}+14t^{15}+7t^{14}+5t^{13}+2t^{12}+9t^{11}+7t^{10}+11t^{9}+3t^{8}+10t^{7}+7t^{6}+14t^{4}+t^{3}+8t^{2}+6t+8.$

Decryption

To decrypt, first substitute the section into F_i:

$h_0(t)=F_0(u_x(t),u_y(t),t)=13t^{64}+8t^{63}+8t^{62}+13t^{61}+7t^{60}+16t^{58}+10t^{57}+13t^{56}+6t^{55}+3t^{54}+15t^{53}+3t^{52}+t^{51}+4t^{50}+2t^{49}+5t^{48}+12t^{47}+3t^{46}+8t^{44}+14t^{43}+9t^{42}+13t^{41}+14t^{40}+10t^{39}+8t^{38}+11t^{37}+12t^{36}+9t^{35}+7t^{33}+14t^{32}+12t^{31}+8t^{30}+4t^{28}+9t^{27}+15t^{26}+t^{25}+4t^{24}+8t^{23}+5t^{22}+14t^{21}+3t^{20}+7t^{19}+6t^{18}+7t^{17}+16t^{16}+9t^{15}+6t^{13}+3t^{12}+8t^{11}+11t^{10}+11t^{9}+14t^{8}+11t^{7}+15t^{6}+14t^{5}+2t^{4}+10t^{3}+10t^{2}+t+10.$

$h_1(t)=F_1(u_x(t),u_y(t),t)=14t^{64}+6t^{63}+6t^{62}+8t^{61}+7t^{60}+t^{59}+4t^{58}+t^{57}+7t^{56}+11t^{55}+10t^{54}+2t^{53}+13t^{52}+16t^{51}+14t^{50}+15t^{49}+3t^{48}+3t^{46}+t^{45}+11t^{44}+10t^{43}+13t^{42}+8t^{41}+6t^{40}+9t^{39}+4t^{38}+13t^{37}+16t^{36}+13t^{35}+12t^{34}+t^{33}+t^{32}+6t^{31}+15t^{30}+15t^{29}+16t^{28}+14t^{27}+2t^{26}+13t^{25}+16t^{24}+16t^{23}+3t^{22}+13t^{21}+4t^{20}+5t^{19}+15t^{18}+5t^{17}+4t^{16}+t^{15}+10t^{14}+15t^{13}+t^{11}+8t^{10}+6t^{9}+13t^{8}+15t^{6}+10t^{5}+4t^{4}+8t^{3}+11t^{2}+12t+2.$

Then factor $h_1(t)-h_0(t)$:

$h_1(t)-h_0(t)=16(t^3+3t^2+13t+3)(t^4+11t^3+15t^2+14t+13)(t^9+8t^8+11t^7+3t^5+4t^4+6t^3+14t^2+12t+13)(t^{17}+2t^{16}+14t^{15}+5t^{14}+5t^{13}+8t^{12}+9t^{11}+11t^{10}+3t^9+13t^8+10t^7+8t^6+15t^5+7t^4+12t^3+10t^2+3t+2)(t^5+13t^4+4t^3+2t^2+4t+13)(t^{16}+4t^{15}+11t^{14}+t^{13}+4t^{12}+13t^{11}+t^{10}+2t^9+t^8+2t^7+t^6+2t^4+15t^3+5t^2+11t+6)(t^6+4t^5+3t^4+10t^3+14t^2+2t+5)(t^4+4t^3+5t^2+16t+10).$

We know that $\deg(f(u_x(t),u_y(t),t))=48$, and that the irreducible divisors of $h_1(t)-h_0(t)$ have degrees $(3,4,4,5,6,9,16,17)$. The associate knapsack has four solutions, but only one corresponds to the real $f(u_x(t),u_y(t),t)$:

$f(u_x(t),u_y(t),t)=(t^3+3t^2+13t+3)(t^4+11t^3+15t^2+14t+13)(t^5+13t^4+4t^3+2t^2+4t+13)(t^6+4t^5+3t^4+10t^3+14t^2+2t+5)(t^9+8t^8+11t^7+3t^5+4t^4+6t^3+14t^2+12t+13)(t^{17}+2t^{16}+14t^{15}+5t^{14}+5t^{13}+8t^{12}+9t^{11}+11t^{10}+3t^9+13t^8+10t^7+8t^6+15t^5+7t^4+12t^3+10t^2+3t+2)(t^4+4t^3+5t^2+16t+10).$

From $f(u_x(t),u_y(t),t)$, we can deduce $m(u_x(t),u_y(t),t)$:

$m(u_x(t),u_y(t),t)=5t^{41}+10t^{40}+9t^{38}+9t^{36}+5t^{35}+12t^{34}+14t^{33}+9t^{31}+6t^{30}+t^{29}+t^{27}+7t^{26}+10t^{25}+3t^{24}+10t^{23}+13t^{22}+4t^{21}+10t^{20}+11t^{19}+6t^{18}+4t^{17}+5t^{16}+7t^{15}+14t^{14}+t^{13}+7t^{12}+11t^{11}+5t^{10}+2t^9+8t^8+14t^7+13t^6+12t^5+16t^4+13t^3+9t^2+13t+13.$

Finally, solve the linear system $m(u_x(t),u_y(t),t)=\sum m_{ijk}x^iy^jt^k$ and recover the plaintext.

8.1.11 MAGMA code for the Level 1 Attack

In the following piece of code, `p` and `d` are the parameters of the system. `deg_t` is the degree of m with respect to t and `Lambda_m` denotes the support of m (these values are public). `F0` and `F1` are the ciphertext, and `X` is the public surface.

```
R<x,y,t>:=PolynomialRing(GF(p),3,"grevlex");
Res:=Resultant(R!(F0-F1),R!X,x);  // Eliminate x
F:=Factorization(Res); // Factor the resultant
// Pick the irreducible factor of highest degree in y
maxdeg:=Max([Degree(R!f[1],R!y) : f in F]);
exists(Q0){f[1]:f in F| Degree(R!f[1],R!y) eq maxdeg};
J:=Ideal([R!Q0,R!X,R!F0,R!F1]);
Groebner(J);  // Compute the Gr\"obner basis of J
Coeffm:=PolynomialRing(GF(p),#Lambda_m*(deg_t+1));
R2<x,y,t>:=PolynomialRing(Coeffm,3);
// Construct the linear system
plaintext:=&+[Coeffm.((i-1)*(deg_t+1)+j)*
             R2!NormalForm(R!x^Lambda_m[i][1]*
             R!y^Lambda_m[i][2]*R!t^(j-1),J) :
             i in [1..#Lambda_m], j in [1..deg_t+1]];
// Solve the linear system:
V:=Variety(Ideal(Coefficients(plaintext)));
```

8.2 Cryptanalysis of MinRank

In this section, we show how Gröbner basis techniques can be used to study the security of the authentication scheme proposed in [Cou01] (see Section 2.2.1). In particular, we study the challenges A, B and C from Table 2.1. The security of the cryptosystem relies on the difficulty of a particular MinRank problem: it is defined in the finite field GF_{65521} and one solution of the problem lies in GF_{65521}^n.

In order to assess the security of the system against algebraic attacks, we focus on the minors modeling (see Section 2.1.2): we consider the set of minors of size $r+1$, which gives rise to a determinantal system.

Workstation. Experimental results have been obtained with 24 Xeon quadricore processors 3.2 GHz, with 64 GB of RAM.

8.2.1 Computing the minors

The minors modeling raises questions about how to generate the equations. It is not clear how to compute efficiently all minors of size $r+1$ of a big matrix. For a $p \times p$ matrix, there are $\binom{p}{r+1}^2$ such minors, and each is a polynomial of degree $r+1$ in n variables. For instance, for an affine problem with $\mathbb{K} = \mathsf{GF}_{65521}$, $p = 11$, $n = 9$ and $r = 8$, it took 14 days on one CPU (with Maple). Fortunately, this computation can be parallelized: with 120 processes running simultaneously on 24 CPU, the computation lasted 12 hours. The size of the resulting algebraic system is 3466 MB.

For this computation, we used naive algorithms (each determinant was computed independently) but we believe that there is room for improvement by using more sophisticated algorithms.

8.2.2 The well-defined case

Here, $n = (p-r)^2$ and the ground field is $\mathbb{K} = \mathsf{GF}_{65521}$.

Generation of the instances. For $(p, n, r) \in \mathbb{N}^3$, we generate a $p \times p$ matrix $M = (M_{i,j})$ where the $M_{i,j}$ are affine linear forms in n variables: $M_{i,j} = a_{i,j}^{(0)} + \sum_{\ell=1}^{n} a_{i,j}^{(\ell)} x_\ell$, where the $a_{i,j}^{(\ell)}$ are chosen uniformly at random in GF_{65521}.

Interpretation of the results. Table 8.3 describes experimental results, for different values of the triplet (p, n, r). In particular, we consider sets of parameters used in Cryptology for a MinRank-based

Chall.	**A**	**B**				**C**
	$(6, 9, 3)$	$(7, 9, 4)$	$(8, 9, 5)$	$(9, 9, 6)$	$(10, 9, 7)$	$(11, 9, 8)$
degree	**980**	**4116**	**14112**	**41580**	**108900**	259545
MH Bézout	8000	42875	175616	592704	1728000	4492125
	Minors					
F_5 time	**1.1s**	**37s**	**935s**	**18122s**	**229094s**	*2570396s*
F_5 mem	**488 MB**	**587 MB**	**1213 MB**	**5048 MB**	**25719MB**	
F_4 Magma	**4.6s**	**142.8s**	**3343.5s**	∞		
$\mathrm{d}_{\mathrm{reg}}$	**10**	**13**	**16**	**19**	**22**	*25*
Nb op.	**21.5**	**25.9**	**29.2**	**32.7**	**35.2**	*40.2*
FGLM time	**1.7s**	**97.2s**	∞			
	Kipnis-Shamir					
F_5 time	**30s**	**3795s**	**328233s**	∞		
F_5 mem	**407 MB**	**3113 MB**	**58587 MB**			
F_4 Magma	**300s**	**48745s**	∞			
$\mathrm{d}_{\mathrm{reg}}$	**5**	**6**	**7**			
Nb op.	**30.5**	**37.1**	**43.4**	*50.4*	*57.4*	*64.4*
FGLM time	**35s**	**2580s**	∞			

Table 8.3: Authentication scheme parameters

authentication scheme [Cou01]. The complexity of solving the MinRank problem is then directly related to the security of this cryptosystem. The values in italic font were not computed, but are estimates of the complexity based on the theoretical results from the previous section.

The row "degree" provides the degree of the ideal (i.e. the number of solutions in the algebraic closure) and can be compared with the multi-homogeneous Bézout bound ("MH Bézout"). The row "F_5 time" (resp. "F_5 mem") gives the time (resp. the memory) needed to compute the grevlex Gröbner basis of the ideal under consideration. The computation is done with the F_5 algorithm from the FGb package. We also give the time obtained for the same Gröbner basis computations with the implementation of F_4 in Magma2.16, so that experiments can be reproduced. "$\mathrm{d}_{\mathrm{reg}}$" gives the degree of regularity of the ideal. Finally "Nb op." indicates the logarithm (in base 2) of the exact number of arithmetic operations performed during the execution of the F_5 algorithm, and "FGLM time" provides the running time of FGLM (from the FGb package).

Note that the degree of regularity of the ideal generated by the minors matches the value given by Lemmas 4.23 and 4.15. Moreover, note that the degree of the ideal is equal to the value provided by Lemma 4.22 and Corollary 4.10.

The fact that the logarithm of the number of arithmetic operations seems to grow linearly (for both modeling) gives experimental evidence that the complexity of the Gröbner basis computation is polynomial in p when $p - r$ is fixed, as announced in [FLP08] and proved in Section 4.6) (see also [FSS10]).

We would like to emphasize that the FGLM step costs sometimes more than the grevlex Gröbner basis computation. In order to avoid this cost, a possible strategy is to combine the minors approach with an exhaustive search over some variables.

8.2.3 Solving the challenge C of the Courtois authentication scheme

Solving the challenge C requires to find one solution of a generic affine $(11, 9, 8)$-MinRank problem which has a particularity: it is known that there is a solution $(x_1, \ldots, x_9) \in \mathsf{GF}_{65521}^9$ in the ground field. Therefore we can combine the minors formulation with a partial exhaustive search. To this end, we specialize s variables and solve the corresponding over-determined $(11, 9 - s, 8)$-MinRank

		$(p = 11, n = 9 - s, r = 8)$			
	s	3	2	1	0
Minors	F_5 time	**79s**	**1594s**	**80255s**	*2570396s*
	F_5 mem	<**1000 MB**	**2400 MB**	**29929 MB**	
	$\mathrm{d_{reg}}$	**9**	**10**	**13**	*25*
	Nb op.	**73**	**60**	**49.1**	*40.2*
KS	F_5 FGb	**57000s**	∞		
	F_5 mem	**10539 MB**			
	$\mathrm{d_{reg}}$	**7**			
	Nb op.	**88.6**			

Table 8.4: Challenge C of the Courtois authentication scheme.

problem for all specializations of the s variables. The degree of regularity of the over-determined systems can be estimated with Theorem 4.17 and Lemma 4.23, so the complexity of the complete computation can be approximated. For these systems, the degree of the ideal is 0 or 1. Consequently, a grevlex Gröbner basis is also a lex Gröbner basis and the FGLM algorithm is no longer required.

Table 8.4 shows the experimental results for different values of s. The row "$\mathrm{d_{reg}}$" gives the degree of regularity obtained for each specialization of the s variables. The row "Nb op." gives an estimate of the logarithm in base 2 of total number of operations needed to solve the challenge C. It is equal to $\log_2(65521^s \mathsf{OpF}_5)$ where OpF_5 is the number of arithmetic operations used by the F_5 algorithm to solve one $(11, 9 - s, 8)$-MinRank problem. The values in italic font were not effectively computed but are given as estimates based on practical and theoretical results.

First of all, we want to emphasize the fact that the degree of regularity of the ideal generated by the minors matches the one deduced from the generic Hilbert series (Theorem 4.17) in the over-determined case.

According to Table 8.4, the best practical choice seems to be $s = 1$. In practice, the 65521 computations of the over-determined systems can be parallelized, and the total number of required arithmetic operations ($2^{49.1}$) is quite practical. We estimate to 238 days the time needed to effectively solve this challenge on 64 quadricore processors. Therefore, the authentication scheme cannot be considered secure anymore with the set of parameters $(p = 11, n = 9, r = 8)$.

Note that it may be possible to compute directly a Gröbner basis of the ideal generated by the minors ($s = 0$). By interpolating the practical results, we give a rough estimate of the complexity of this computation: it would take approximately 29 days (on one CPU). However, it is not clear how much memory would be required, and the FGLM step could be untractable since the degree of the ideal is 259545 (Corollary 4.10).

8.3 Analysis of QUAD

We estimate here the impact of the new algorithm $\mathsf{BooleanSolve}$ (Algorithm 9) from the point of view of a user in Cryptology. In other words, if the security of a cryptosystem relies on the hardness of solving a quadratic boolean polynomial system, by how much must the parameters be increased to keep the same level of security?

The stream cipher QUAD [BGP06, BGP09] enjoys a provable security argument to support its conjectured strength. It relies on the iteration of a set of overdetermined multivariate quadratic polynomials over GF_2 so that the security of the keystream generation is related, in the concrete security model, to the difficulty of solving the Boolean MQ SAT problem. A theoretical bound is used in [BGP09] to obtain secure parameters for a given security bound T and a given maximal length L of the keystream sequence that can be generated with a pair (key, IV): for instance (see [BGP09] p. 1711), for $T = 2^{80}, L = 2^{40}, k = 2$ and an advantage of more than $\varepsilon = 1/100$, the bound gives $n \geq 331$. We report in the following table various values of n depending on L, T and ε:

T	L	ε	n
2^{80}	2^{40}	1/100	331
2^{80}	2^{22}	1/100	253
2^{160}	2^{80}	1/100	613
2^{160}	2^{40}	1/100	445
2^{160}	2^{40}	1/1000	448
2^{160}	2^{40}	1/10000	467
2^{256}	2^{40}	1/100	584
2^{256}	2^{80}	1/100	758

Security parameters for the stream cipher QUAD [BGP09]

Now, the question is to achieve a security bound for $T = 2^{256}$; what are the minimal values of m and n ensuring that solving the Boolean MQ SAT requires at least T bit-operations? Using the complexity analysis of the BooleanSolve algorithm we can derive useful lower bounds for n when $m = n$ or $m = 2\,n$ ($m = 2\,n$ corresponds to the recommended parameters for QUAD). In the following table we report the corresponding values:

Security Bound T	2^{128}	2^{256}	2^{512}	2^{1024}
Minimal value of n when $m = n$	128	270	576	1202
Minimal value of n when $m = 2\,n$	145	335	738	1580

Comparing with exhaustive search we can see from this table that:

- our algorithm does not improve upon exhaustive search when n is small (for instance when $m = n$ and $T = 2^{128}$ that are the recommended parameters);
- by contrast, our algorithm can take advantage of the overdeterminedness of the algebraic systems: this explains why the values we recommend are larger than expected when n is large and/or $m/n > 1$.

Index

Index of Notations

Bibliography

[Abh88] S.S. Abhyankar. *Enumerative combinatorics of Young tableaux*. Marcel Dekker, 1988.

[AFFP11] M. Albrecht, J.-C. Faugère, P. Farshim, and L. Perret. Polly cracker, revisited. In *Proceedings of Asiacrypt 2011*, Lecture Notes in Computer Science, pages 1–14. Springer Verlag, 2011.

[AFI$^+$04] G. Ars, J.-C. Faugère, H. Imai, M. Kawazoe, and M. Sugita. Comparison between XL and Gröbner basis algorithms. In *Advances in Cryptology – AsiaCrypt 2004*, volume 3329/2004 of *LNCS*, pages 157–167, 2004.

[AG04] K. Akiyama and Y. Goto. An algebraic surface public-key cryptosystem. *IEIC Technical Report (Institute of Electronics, Information and Communication Engineers)*, 104(421):13–20, 2004.

[AGM09] K. Akiyama, Y. Goto, and H. Miyake. An algebraic surface cryptosystem. In *Proceedings of the 12th International Conference on Practice and Theory in Public Key Cryptography: PKC'09*, page 442. Springer, 2009.

[AM69] M.F. Atiyah and I.G. MacDonald. *Introduction to Commutative Algebra*. Addison-Wesley Publishing Company, 1969.

[ARS02] P. Aubry, F. Rouillier, and M. Safey El Din. Real solving for positive dimensional systems. *Journal of Symbolic Computation*, 34(6):543–560, 2002.

[Bar93] A.I. Barvinok. Feasibility testing for systems of real quadratic equations. *Discrete & Computational Geometry*, 10(1):1–13, 1993.

[Bar04] M. Bardet. *Étude des systèmes algébriques surdéterminés. Applications aux codes correcteurs et à la cryptographie*. PhD thesis, Université Paris 6, 2004.

[BCC$^+$10] C. Bouillaguet, H.-C. Chen, C.-M. Cheng, Tung Chou, R. Niederhagen, A. Shamir, and B.-Y. Yang. Fast exhaustive search for polynomial systems in F_2. In *Cryptographic Hardware and Embedded Systems, CHES 2010*, volume 6225 of *LNCS*, pages 203–218, 2010.

[BCGO09] T. Berger, P.L. Cayrel, P. Gaborit, and A. Otmani. Reducing key length of the McEliece cryptosystem. *Progress in Cryptology–AFRICACRYPT 2009*, pages 77–97, 2009.

[BCP97] 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.

[BFP09] L. Bettale, J.-C. Faugère, and L. Perret. Hybrid approach for solving multivariate systems over finite fields. *Journal of Mathematical Cryptology*, 3:177–197, 2009.

[BFP11] L. Bettale, J.-C. Faugère, and L. Perret. Cryptanalysis of multivariate and odd-characteristic HFE variants. In *Public Key Cryptography – PKC 2011*, volume 6571 of *Lecture Notes in Computer Science*, pages 441–458. Springer, 2011.

[BFP12a] L. Bettale, J.-C. Faugère, and L. Perret. Cryptanalysis of HFE, multi-HFE and variants for odd and even characteristic. *Design, Codes and Cryptography*, 2012.

[BFP12b] L. Bettale, J.-C. Faugère, and L. Perret. Solving polynomial systems over finite fields: Improved analysis of the hybrid approach. In *International Symposium on Symbolic and Algebraic Computation – ISSAC 2012*. ACM, 2012.

[BFS99] J.F. Buss, G.S. Frandsen, and J. Shallit. The computational complexity of some problems of linear algebra. *Journal of Computer and System Sciences*, 58(3):572–596, 1999.

[BFS04] M. Bardet, J.-C. Faugère, and B. Salvy. On the complexity of Gröbner basis computation of semi-regular overdetermined algebraic equations. In *Proceedings of the International Conference on Polynomial System Solving (ISCPP)*, pages 71–74, 2004.

[BFSS12] M. Bardet, J.-C. Faugère, B. Salvy, and P.-J. Spaenlehauer. On the complexity of solving quadratic boolean systems. *Journal of Complexity*, 2012. Accepted for publication.

[BFSY04] M. Bardet, J.-C. Faugère, B. Salvy, and B.-Y. Yang. Asymptotic expansion of the degree of regularity for semi-regular systems of equations. In *Effective Methods in Algebraic Geometry (MEGA)*, pages 71–74, 2004.

[BGH$^+$10] B. Bank, M. Giusti, J. Heintz, M. Safey El Din, and É. Schost. On the geometry of polar varieties. *Applicable Algebra in Engineering, Communication and Computing*, 21(1):33–83, 2010.

[BGHM97] B. Bank, M. Giusti, J. Heintz, and G.-M. Mbakop. Polar varieties and efficient real equation solving: the hypersurface case. *Journal of Complexity*, 13(1):5–27, 1997.

[BGHM01] B. Bank, M. Giusti, J. Heintz, and G.-M. Mbakop. Polar varieties and efficient real elimination. *Mathematische Zeitschrift*, 238(1):115–144, 2001.

[BGHP04] B. Bank, M. Giusti, J. Heintz, and L.-M. Pardo. Generalized polar varieties and efficient real elimination procedure. *Kybernetika*, 40(5):519–550, 2004.

[BGHP05] B. Bank, M. Giusti, J. Heintz, and L.-M. Pardo. Generalized polar varieties: Geometry and algorithms. *Journal of complexity*, 21(4):377–412, 2005.

[BGP06] C. Berbain, H. Gilbert, and J. Patarin. QUAD: A practical stream cipher with provable security. In Serge Vaudenay, editor, *Advances in Cryptology - EUROCRYPT 2006*, volume 4004 of *Lecture Notes in Computer Science*, pages 109–128. Springer Berlin / Heidelberg, 2006.

[BGP09] C. Berbain, H. Gilbert, and J. Patarin. QUAD: A multivariate stream cipher with provable security. *Journal of Symbolic Computation*, 44:1703–1723, December 2009.

[BL09] C. Beltrán and A. Leykin. Certified numerical homotopy tracking. Arxiv preprint arXiv:0912.0920, 2009.

[BMMT94] 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.

[BPR96] S. Basu, R. Pollack, and M.-F. Roy. On the combinatorial and algebraic complexity of quantifier elimination. *Journal of ACM*, 43(6):1002–1045, 1996.

[BPR98] S. Basu, R. Pollack, and M.-F. Roy. A new algorithm to find a point in every cell defined by a family of polynomials. In *Quantifier elimination and cylindrical algebraic decomposition*. Springer-Verlag, 1998.

[BS05] A. Bostan and É. Schost. Polynomial evaluation and interpolation on special sets of points. *Journal of Complexity*, 21(4):420–446, 2005.

[Buc65] B. Buchberger. *An algorithm for finding the basis elements of the residue class ring of a zero dimensional polynomial ideal*. PhD thesis, University of Innsbruck, 1965.

[Bus04] L. Busé. Resultants of determinantal varieties. *Journal of Pure and Applied Algebra*, 193(1-3):71–97, 2004.

[BV88] W. Bruns and U. Vetter. *Determinantal Rings*. Springer, 1988.

[BZ93] D. Bernstein and A. Zelevinsky. Combinatorics of maximal minors. *Journal of Algebraic Combinatorics*, 2(2):111–121, 1993.

[Can88] J.F. Canny. *Complexity of Robot Motion Planning*. PhD thesis, Massachusetts Institute of Technology, 1988.

[Can93] J.F. Canny. Computing roadmaps of general semi-algebraic sets. *The Computer Journal*, 36(5):504–514, 1993.

[CDS07] D. Cox, A. Dickenstein, and H. Schenck. A case study in bigraded commutative algebra. In I. Peeva, editor, *Syzygies and Hilbert functions*, Lecture Notes in Pure and Applied Mathematics. CRC Press, 2007.

[CH94] A. Conca and J. Herzog. On the Hilbert function of determinantal rings and their canonical module. *Proceedings of the American Mathematical Society*, 122(3):677–681, 1994.

[CKY89] J. F. Canny, E. Kaltofen, and L. Yagati. Solving systems of nonlinear polynomial equations faster. In *Proceedings of the ACM-SIGSAM 1989 International Symposium on Symbolic and Algebraic Computation*, ISSAC '89, pages 121–128, New York, NY, USA, 1989. ACM.

[CLO97] D. Cox, J. Little, and D. O'Shea. *Ideals, Varieties and Algorithms*. Springer, 3rd edition, 1997.

[Col75] G. Collins. Quantifier elimination for real closed fields by Cylindrical Algebraic Decomposition. In *Automata Theory and Formal Languages 2nd GI Conference Kaiserslautern, May 20–23, 1975*, pages 134–183. Springer, 1975.

[Cou01] N. Courtois. Efficient zero-knowledge authentication based on a linear algebra problem MinRank. In *Advances in Cryptology - ASIACRYPT 2001*, volume 2248 of *LNCS*, pages 402–421. Springer, 2001.

[CW90] D. Coppersmith and S. Winograd. Matrix multiplication via arithmetic progressions. *Journal of Symbolic Computation*, 9(3):251–280, 1990.

[DE03] A. Dickenstein and I. Emiris. Multihomogeneous resultant formulae by means of complexes. *Journal of Symbolic Computation*, 36(3-4):317–342, 2003.

[DH88] J.H. Davenport and J. Heintz. Real quantifier elimination is doubly exponential. *Journal of Symbolic Computation*, 5(1-2):29–35, February 1988.

[Die] C. Diem. Bounded regularity. 2012.

[Eis95] D. Eisenbud. *Commutative Algebra with a View Toward Algebraic Geometry*. Springer, 1995.

[Eis01] D. Eisenbud. *The geometry of syzygies*. Springer Verlag, 2001.

[ELLS09] H. Everett, D. Lazard, S. Lazard, and M. Safey El Din. The Voronoi diagram of three lines. *Discrete & Computational Geometry*, 42(1):94–130, 2009.

[EM09] I. Emiris and A. Mantzaflaris. Multihomogeneous resultant formulae for systems with scaled support. In *Proceedings of the 2009 International Symposium on Symbolic and Algebraic Computation*, pages 143–150. ACM, 2009.

[FA66] S.D. Fisher and M.N. Alexander. Matrices over a finite field. *The American Mathematical Monthly*, 73(6):639–641, 1966.

[Fau99] J.-C. Faugère. A new efficient algorithm for computing Gröbner bases (F4). *Journal of Pure and Applied Algebra*, 139(1–3):61–88, 1999.

[Fau02] J.-C. Faugère. A new efficient algorithm for computing Gröbner bases without reductions to zero (F5). In Teo Mora, editor, *Proceedings of the 2002 International Symposium on Symbolic and Algebraic Computation (ISSAC)*, pages 75–83. ACM Press, 2002.

[FB07] G. Fusco and E. Bach. Phase transition of multivariate polynomial systems. In *Theory and Applications of Models of Computation (TAMC 2007)*, volume 4484/2007 of *LNCS*, pages 632–645, 2007.

[FGHR12] J.-C. Faugère, P. Gaudry, L. Huot, and G. Renault. Fast change of ordering with exponent ω. Poster at the conference ISSAC2012, 2012.

[FGLM93] 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.

[FJ98] P. Fitzpatrick and S.M. Jennings. Comparison of two algorithms for decoding alternant codes. *Applicable Algebra In Engineering, Communication and Computing*, 9(3):211–220, 1998.

[FJ03] J.-C. Faugère and A. Joux. Algebraic cryptanalysis of Hidden Field Equation (HFE) cryptosystems using Gröbner bases. In *Advances in Cryptology - CRYPTO 2003*, volume 2729 of *LNCS*, pages 44–60. Springer, 2003.

[FL10] J.-C. Faugère and S. Lachartre. Parallel Gaussian elimination for Gröbner bases computations in finite fields. In *PASCO*, pages 89–97, 2010.

[FLP08] J.-C. Faugère, F. Lévy-dit-Vehel, and L. Perret. Cryptanalysis of MinRank. In *Advances in Cryptology - CRYPTO 2008*, volume 5157 of *LNCS*, pages 280–296. Springer, 2008.

[FM11] J.-C. Faugère and C. Mou. Fast algorithm for change of ordering of zero-dimensional Gröbner bases with sparse multiplication matrices. In *Proceedings of the 36th international symposium on Symbolic and algebraic computation (ISSAC '11*, pages 115–122. ACM, 2011.

[FMRS08] J.C. Faugère, G. Moroz, F. Rouillier, and M. Safey El Din. Classification of the perspective-three-point problem, discriminant variety and real solving polynomial systems of inequalities. In *Proceedings of the twenty-first international symposium on Symbolic and algebraic computation*, pages 79–86. ACM, 2008.

[FOPT10] J.-C. Faugère, A. Otmani, L. Perret, and J.-P. Tillich. Algebraic cryptanalysis of McEliece variants with compact keys. In *Proceedings of Eurocrypt 2010*, volume 6110 of *Lecture Notes in Computer Science*, pages 279–298. Springer Verlag, 2010.

[FP06] J.-C. Faugère and L. Perret. Polynomial equivalence problems: Algorithmic and theoretical aspects. In *Advances in Cryptology - EUROCRYPT 2006*, volume 4004 of *LNCS*, pages 30–47. Springer, 2006.

[FR09] J.-C. Faugère and S. Rahmany. Solving systems of polynomial equations with symmetries using SAGBI-Gröbner bases. In *Proceedings of the 2009 International Symposium on Symbolic and Algebraic Computation*, pages 151–158. ACM, 2009.

[Fro85] R. Froberg. An inequality for Hilbert series of graded algebras. *Mathematica Scandinavica*, 56:117–144, 1985.

[FS10] J.-C. Faugère and P.-J. Spaenlehauer. Algebraic cryptanalysis of the PKC'2009 Algebraic Surface Cryptosystem. In *Proceedings of the 13th International Conference on Practice and Theory in Public Key Cryptography*, volume 6056 of *LNCS*, pages 35–52. Springer, 2010.

[FSS10] J.-C. Faugère, M. Safey El Din, and P.-J. Spaenlehauer. Computing loci of rank defects of linear matrices using Gröbner bases and applications to cryptology. In Stephen M. Watt, editor, *Proceedings of the 2010 International Symposium on Symbolic and Algebraic Computation (ISSAC 2010)*, pages 257–264, 2010.

[FSS11a] J.-C. Faugère, M. Safey El Din, and P.-J. Spaenlehauer. Gröbner bases of bihomogeneous ideals generated by polynomials of bidegree (1,1): Algorithms and complexity. *Journal Of Symbolic Computation*, 46(4):406–437, 2011. Available online 4 November 2010.

[FSS11b] J.-C. Faugère, M. Safey El Din, and P.-J. Spaenlehauer. On the complexity of the Generalized MinRank Problem. *CoRR*, abs/1112.4411, 2011. submitted.

[FSS12] J.-C. Faugère, M. Safey El Din, and P.-J. Spaenlehauer. Critical points and Gröbner bases: the unmixed case. In *Proceedings of the 2012 International Symposium on Symbolic and Algebraic Computation (ISSAC 2012)*, pages 162–169, 2012.

[Ful97] W. Fulton. *Intersection Theory*. Springer, 2nd edition, 1997.

[FY79] A. S. Fraenkel and Y. Yesha. Complexity of problems in games, graphs and algebraic equations. *Discrete Appl. Math.*, 1(1-2):15–30, 1979.

[Gab85] È.M. Gabidulin. Theory of codes with maximum rank distance. *Problemy Peredachi Informatsii*, 21(1):3–16, 1985.

[GJ79] M.R. Garey and D.S. Johnson. *Computers and Intractability: A Guide to the Theory of NP-completeness*. wh freeman San Francisco, 1979.

[GLS98] M. Giesbrecht, A. Lobo, and D. Saunders. Certifying inconsistency of sparse linear systems. In *Proceedings of the 1998 International Symposium on Symbolic and Algebraic Computation (ISSAC 1998)*, pages 113–119, 1998.

[GLS01] M. Giusti, G. Lecerf, and B. Salvy. A gröbner free alternative for polynomial system solving. *Journal of Complexity*, 17(1):154–211, 2001.

[GP05] D. Grigoriev and D.V. Pasechnik. Polynomial-time computing over quadratic maps i: sampling in real algebraic sets. *Computational Complexity*, 14(1):20–52, April 2005.

[GS11] A. Greuet and M. Safey El Din. Deciding reachability of the infimum of a multivariate polynomial. In *ISSAC*, pages 131–138, 2011.

[GV88] D. Grigoriev and N. Vorobjov. Solving systems of polynomials inequalities in subexponential time. *Journal of Symbolic Computation*, 5:37–64, 1988.

[HE70] M. Hochster and J.A Eagon. A class of perfect determinantal ideals. *Bulletin of the American Mathematical Society*, 76(5):1026–1029, 1970.

[HE71] M. Hochster and J.A. Eagon. Cohen-Macaulay rings, invariant theory, and the generic perfection of determinantal loci. *American Journal of Mathematics*, 93(4):1020–1058, 1971.

[Hen08] D. Henrion. Polynomials and convex optimization for robust control. Habilitation Thesis, 2008.

[HKL+11] K.A. Hansen, M. Koucky, N. Lauritzen, P.B. Miltersen, and E.P. Tsigaridas. Exact algorithms for solving stochastic games. In *STOC 2011*, 2011.

[HRS89] J. Heintz, M.-F. Roy, and P. Solernò. On the complexity of semi-algebraic sets. In *Proceedings IFIP'89 San Francisco, North-Holland*, 1989.

[HRS93] J. Heintz, M.-F. Roy, and P. Solernò. On the theoretical and practical complexity of the existential theory of the reals. *The Computer Journal*, 36(5):427–431, 1993.

[HS09] H. Hong and M. Safey El Din. Variant real quantifier elimination: algorithm and application. In *Proceedings of the 2009 International Symposium on Symbolic and Algebraic Computation*, pages 183–190. ACM, 2009.

[HS10] J.D. Hauenstein and F. Sottile. alphaCertified: certifying solutions to polynomial systems. Arxiv preprint arXiv:1011.1091, 2010.

[HS11] H. Hong and M. Safey El Din. Variant quantifier elimination. *Journal of Symbolic Computation*, 2011.

[HSS98] B. Huber, F. Sottile, and B. Sturmfels. Numerical schubert calculus. *Journal of Symbolic Computation*, 26(6):767–788, 1998.

[IV09] P. Ivanov and J.F. Voloch. Breaking the Akiyama-Goto cryptosystem. *Arithmetic, Geometry, Cryptography and Coding Theory*, 487, 2009.

[Iwa07] M. Iwami. A reduction attack on Algebraic Surface public-key Cryptosystems. In *Workshop of Research Institute for Mathematical Sciences (RIMS) Kyoto University, New development of research on Computer Algebra, RIMS Kokyuroku*, volume 1572. Springer, 2007.

[Jel05] Z. Jelonek. On the effective Nullstellensatz. *Inventiones Mathematicae*, 162(1):1–17, 2005.

[JS07] G. Jeronimo and J. Sabia. Computing multihomogeneous resultants using straight-line programs. *Journal of Symbolic Computation*, 42(1-2):218–235, 2007.

[KPG99] A. Kipnis, J. Patarin, and L. Goubin. Unbalanced Oil and Vinegar signature schemes. In *Advances in Cryptology – Eurocrypt 99*, volume 1592/1999 of *LNCS*, pages 206–222, 1999.

[KPS01] T. Krick, L.M. Pardo, and M. Sombra. Sharp estimates for the arithmetic Nullstellensatz. *Duke Mathematical Journal*, 109(3):521–598, 2001.

[Kra93] C. Krattenthaler. Non-crossing two-rowed arrays and summations for Schur functions. In *Proc. of the 5th Conference on Formal Power Series and Algebraic Combinatorics, Florence*, pages 301–314. Citeseer, 1993.

[KRHV02] M. Kreuzer, L. Robbiano, J. Herzog, and V. Vulutescu. Basic tools for computing in multigraded rings. In *Commutative Algebra, Singularities and Computer Algebra, Proc. Conf. Sinaia*, pages 197–216, 2002.

[KS91] E. Kaltofen and D.B. Saunders. On Wiedemann's method of solving sparse linear systems. *Applied Algebra, Algebraic Algorithms and Error-Correcting Codes*, pages 29–38, 1991.

[KS99] A. Kipnis and A. Shamir. Cryptanalysis of the HFE public key cryptosystem by relinearization. In *Advances in Cryptology - CRYPTO' 99*, volume 1666 of *LNCS*, pages 19–30. Springer, 1999.

[Kul96] D.M. Kulkarni. Counting of paths and coefficients of the Hilbert polynomial of a determinantal ideal. *Discrete Mathematics*, 154(1):141–151, 1996.

[Lak90] 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.

[Laz83] D. Lazard. Gröbner bases, Gaussian elimination and resolution of systems of algebraic equations. In *Computer Algebra, EUROCAL'83*, volume 162 of *LNCS*, pages 146–156. Springer, 1983.

[Laz92] D. Lazard. Solving zero-dimensional algebraic systems. *Journal of symbolic computation*, 13(2):117–131, 1992.

[Lec03] G. Lecerf. Computing the equidimensional decomposition of an algebraic closed set by means of lifting fibers. *J. of Complexity*, 19(4):564–596, 2003.

[Lec10] G. Lecerf. New recombination algorithms for bivariate polynomial factorization based on Hensel lifting. *Applicable Algebra in Engineering, Communication and Computing*, 21(2):151–176, 2010.

[Mac02] F.S. Macaulay. Some formulæ in elimination. *Proceedings of the London Mathematical Society*, s1-35(1):3–38, 1902.

[MB09] R. Misoczki and P. Barreto. Compact McEliece keys from goppa codes. In *Selected Areas in Cryptography*, pages 376–392. Springer, 2009.

[McC33] N.H. McCoy. On the resultant of a system of forms homogeneous in each of several sets of variables. *Transactions of the American Mathematical Society*, pages 215–233, 1933.

[MS87] A. Morgan and A. Sommese. A homotopy for solving general polynomial systems that respects m-homogeneous structures. *Applied Mathematics and Computation*, 24(2):101–113, 1987.

[MS03] G. Moreno-Socías. Degrevlex Gröbner bases of generic complete intersections. *Journal of Pure and Applied Algebra*, 180(3):263–283, 2003.

[MS05] E. Miller and B. Sturmfels. *Combinatorial commutative algebra*, volume 227. Springer Verlag, 2005.

[NR09] J. Nie and K. Ranestad. Algebraic degree of polynomial optimization. *SIAM Journal on Optimization*, 20(1):485–502, 2009.

[OJ02] A.V. Ourivski and T. Johansson. New technique for decoding codes in the rank metric and its cryptography applications. *Problems of Information Transmission*, 38(3):237–246, 2002.

[Onn94] S. Onn. Hilbert series of group representations and Gröbner bases for generic modules. *Journal of Algebraic Combinatorics*, 3(2):187–206, 1994.

[Ove05] R. Overbeck. A new structural attack for GPT and variants. *Progress in Cryptology–Mycrypt 2005*, pages 50–63, 2005.

[Par10] K. Pardue. Generic sequences of polynomials. *Journal of Algebra*, 324(4):579–590, 2010.

[Pat96] J. Patarin. Hidden Fields Equations (HFE) and Isomorphisms of Polynomials (IP): Two new families of asymmetric algorithms. In *Advances in Cryptology - EUROCRYPT '96*, volume 1070 of *LNCS*, pages 33–48. Springer, 1996.

[Phi86] P. Philippon. Criteres pour l'indépendance algébrique. *Publications Mathématiques de l'IHÉS*, 64(1):5–52, 1986.

[Rém01a] G. Rémond. Elimination multihomogène. *Introduction to Algebraic Independence Theory. Lect. Notes Math*, 1752:53–81, 2001.

[Rém01b] G. Rémond. Géométrie diophantienne multiprojective, chapitre 7 de introduction to algebraic independence theory. *Lecture Notes in Math*, pages 95–131, 2001.

[Saf07] M. Safey El Din. Testing sign conditions on a multivariate polynomial and applications. *Mathematics in Computer Science*, 1(1):177–207, 2007.

[Sem08] I. Semaev. On solving sparse algebraic equations over finite fields. *Design, Codes and Cryptography*, 49(1-3):47–60, 2008.

[Sem09] I. Semaev. Sparse algebraic equations over finite fields. *SIAM Journal on Computing*, 39(2):388–409, 2009.

[Sha94] I.R. Shafarevich. *Basic Algebraic Geometry I*. Springer, second, revised and expanded edition, 1994.

[Sho94] P. W. Shor. Algorithms for quantum computation: discrete logarithms and factoring. In *SFCS '94: Proceedings of the 35th Annual Symposium on Foundations of Computer Science*, pages 124–134, Washington, DC, USA, 1994. IEEE Computer Society.

[SS03] M. Safey El Din and E. Schost. Polar varieties and computation of one point in each connected component of a smooth real algebraic set. In *Proceedings of the 2003 International Symposium on Symbolic and Algebraic Computation*, pages 224–231. ACM New York, NY, USA, 2003.

[SS04] M. Safey El Din and É. Schost. Properness defects of projections and computation of one point in each connected component of a real algebraic set. *Discrete and Computational Geometry*, 32(3):417–430, 2004.

[SS10] M. Safey El Din and É Schost. A baby steps/giant steps probabilistic algorithm for computing roadmaps in smooth bounded real hypersurface. *Discrete & Computational Geometry*, February 2010.

[ST06] M. Safey El Din and P. Trébuchet. Strong bi-homogeneous Bézout theorem and its use in effective real algebraic geometry. *Arxiv preprint cs/0610051*, 2006.

[Sti87] E.L. Stitzinger. The probability that a linear system is consistent. *Linear and Multilinear Algebra*, 21(4):367–371, 1987.

[Sto00] A. Storjohann. *Algorithms for Matrix Canonical Forms*. PhD thesis, University of Waterloo, 2000.

[Sto10] A. Stothers. *On the Complexity of Matrix Multiplication*. PhD thesis, University of Edinburgh, 2010.

[Str69] V. Strassen. Gaussian elimination is not optimal. *Numerische Mathematik*, 13(4):354–356, 1969.

[SZ93] B. Sturmfels and A. Zelevinsky. Maximal minors and their leading terms. *Advances in mathematics*, 98(1):65–112, 1993.

[Tar51] A. Tarski. A decision method for elementary algebra and geometry. *Bulletin of the American Society*, 59, 1951.

[UT07] S. Uchiyama and H. Tokunaga. On the security of the Algebraic Surface public-key Cryptosystems. In *Proceedings of SCIS*, 2007.

[Van29] B.L. Van der Waerden. On Hilbert's function, series of composition of ideals and a generalization of the theorem of Bezout. In *Proceedings of the Royal Academy of Sciences, Amsterdam*, volume 31, pages 749–770, 1929.

[Vas11] V. Vassilevska Williams. Breaking the Coppersmith-Winograd barrier. Technical report, UC Berkeley, 2011.

[Ver99] J. Verschelde. Polynomial homotopies for dense, sparse and determinantal systems, 1999. arXiv:math/9907060.

[Ver11] J. Verschelde. Polynomial homotopy continuation with PHCpack. *ACM Communications in Computer Algebra*, 44(3/4):217–220, 2011.

[Vil97] G. Villard. Further analysis of Coppersmith's block Wiedemann algorithm for the solution of sparse linear systems. In *Proceedings of the 1997 International Symposium on Symbolic and Algebraic Computation (ISSAC)*, pages 32–39. ACM, 1997.

[VZGG03] J. Von Zur Gathen and J. Gerhard. *Modern computer algebra*. Cambridge University Press, 2003.

[Wie86] D. Wiedemann. Solving sparse linear equations over finite fields. *IEEE Transactions on Information Theory*, 32(1):54–62, 1986.

[YC04] B.-Y. Yang and J.-M. Chen. Theoretical analysis of XL over small fields. In *Information Security and Privacy 2004*, volume 3108/2004 of *LNCS*, pages 277–288, 2004.

[YCC04] B.-Y. Yang, J.-M. Chen, and N.T. Courtois. On asymptotic security estimates in XL and Gröbner bases-related algebraic cryptanalysis. In *ICICS 2004, LNCS 3269*, pages 401–413. Springer-Verlag, 2004.