Bases de Gröbner en Cryptographie Post-Quantique

Ludovic Perret

École Doctorale Informatique, Télécommunications et Électronique (EDITE) de Paris

Gröbner Bases Techniques in Quantum-Safe Cryptography

THÈSE

présentée et soutenue publiquement le

pour l'obtention d'une

Habilitation de Sorbonne Universités

(mention informatique)

par

Ludovic Perret

Composition du jury

Rapporteurs :	Steven Galbraith (Professor)	University of Auckland
	Henri Gilbert, Expert (HDR)	Head of the cryptographic team at ANSSI
	Igor Shparlinski (Professor)	The University of New South Wales
Examinateurs :	Jean-Charles Faugère	Research director, Head of PolSys, INRIA Paris
	Pierre-Alain Fouque (Professor)	IUF and University of Rennes 1
	Jacques Patarin (Professor)	University of Versailles Saint Quentin en Yvelines
	Mohab Safey El Din (Professor)	IUF and UPMC
	Gilles Zémor (Professor)	University of Bordeaux

Laboratoire d'Informatique de Paris 6 (LIP6) - UMR 7606

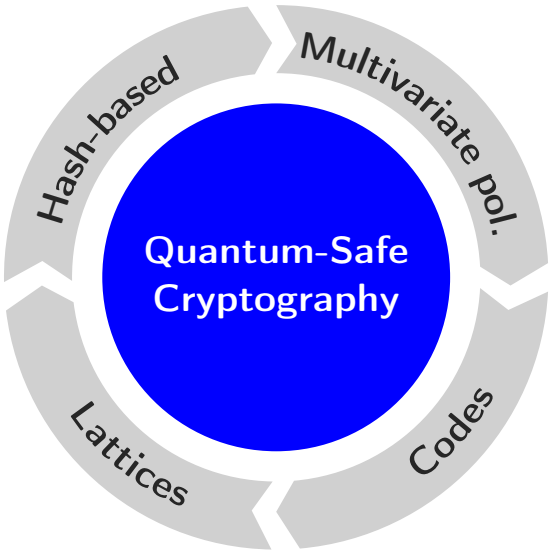
Mis en page avec la classe thesul.

CONTENTS

GENERAL CONTEXT – THE QUANTUM-SAFE REVOLUTION

The goal of *Quantum-Safe* (or *post-quantum*) cryptography[1] is to design cryptographic primitives which are secure against a classical and quantum adversary. This is a well established academic topic mainly motivated by Shor's milestone quantum algorithm [258]. Indeed, although no classical polynomial-time algorithm has been found for the number theoretic cryptographic problems used in practice – such as integer factorization (FACT), e.g., in RSA, and discrete logarithm (DLOG), e.g., in the Diffie-Hellman key-exchange – Shor's algorithm allows to solve DLOG and FACT in polynomial-time on a quantum computer.

Quantum-Safe Cryptography (QSC) is an active cryptographic topic which started soon after Shor's algorithm. Today, it is commonly admitted that the most promising quantum-safe cryptosystems include [43], [242]: *Quantum-Key Distribution* (QKD, [36]), *code-based* cryptosystems [229], *hash-based* cryptosystems [71], [186], *isogeny-based* cryptography [165], [188], *lattice-based* [219] cryptosystems and finally *multivariate-based* cryptosystems [109]. Among this list, QKD is different from the others techniques. It is not based on the hardness of an algorithmic problem but rather on a physical assumption. It can be also mentioned that finding isogenies between supersingular curves [165], [188] is the youngest algorithmic problem introduced in QSC; the first paper [188] by De Feo and Jao dates of 2011.



The status of quantum-safe cryptography is currently completely changing. It is quickly moving from a purely academic theme to a topic of major industrial interest. As such, this can be considered as a revolution. This new industrial interest is mainly driven by the fact that quantum-safe cryptography has received recently much attention from the standardization and policy spectra. The trigger event is the announcement in August 2015 by the National Security Agency (NSA) of preliminary plans for a transition to quantum resistant algorithms[2]:

[1] We adopt quantum-safe cryptography in this document.

[2] https://www.nsa.gov/ia/programs/suiteb_cryptography/

> "*Currently, Suite B cryptographic algorithms are specified by the National Institute of Standards and Technology (*NIST*) and are used by* NSA*'s Information Assurance Directorate in solutions approved for protecting classified and unclassified National Security Systems (*NSS*). Below, we announce preliminary plans for transitioning to* **quantum resistant** *algorithms.*"

This was quickly followed by an announcement of NIST which detailed the transition process. NIST, which has the authority to establish the security standards of the US government, released in January 2016 a call to select standards for quantum-safe public-key cryptosystems: public-key exchange, signature and public-key encryption [78], [226]. With historical perspective, for example with the *Advanced Encryption Standard*, it seems likely that the quantum-safe standards derived from this process will be widely endorsed around the world.
In parallel of the US process, Europe is also at the forefront of quantum-safe standardization with an industry specification group (ISG) on QKD[3] and a more recent ISG on quantum-safe cryptography[4]; the latter focusing more on algorithmic techniques. The International Standardization Organization (ISO SC 27/WG 2) is currently in a study period on quantum-safe cryptography. Industry-based think-tank such as the Cloud Security Alliance is also contributing to raise awareness on QSC with an industry group dedicated[5] to quantum-safe security. To complete the world tour, we mention that Asia recently started a dedicated forum on quantum-safe cryptography[6].

Gröbner Bases Techniques in Quantum-Safe Cryptography

It is clear that the effort to develop quantum-safe cryptosystems is now intensifying. Still, a key issue for a wide adoption of future quantum-safe standards is our confidence in their security. There is therefore a great need to develop cryptanalysis against quantum-safe cryptosystems. Cryptanalysis is of course a much needed tool to filter out the weakest primitives. However, it is also the only reliable technique to set the parameters of any cryptosystem and a fundamental element in the design of future standards.
Although quantum-safe cryptosystems are based on different hardness assumptions, a goal of this document is to show that *algebraic cryptanalysis* provides a general framework to analyse these primitives. The principle of such cryptanalysis is to model a cryptographic primitive by a set of algebraic equations. The system of equations is constructed in a way to have a correspondence between the zeroes and a secret information of the cryptographic primitive considered (for instance, the secret-key of an encryption scheme).
Algebraic cryptanalysis reduces the security analysis of a cryptographic primitive to the problem $\mathtt{PoSSo}_q$ which consists of solving a system of non-linear equations over a finite field $\mathbb{F}_q$ (with $q = p^s$, p prime and $s > 0$). This classical NP-hard problem [176] is defined as follows:

Polynomial System Solving over a Finite Field ($\mathtt{PoSSo}_q$)
Input. $p_1(x_1, \ldots, x_n), \ldots, p_m(x_1, \ldots, x_n) \in \mathbb{F}_q[x_1, \ldots, x_n]$.
Goal. Find – if any – a vector $(z_1, \ldots, z_n) \in \mathbb{F}_q^n$ such that:

$$p_1(z_1, \ldots, z_n) = 0, \ldots, p_m(z_1, \ldots, z_n) = 0.$$

[3] https://portal.etsi.org/tb.aspx?tbid=723&SubTB=723
[4] https://portal.etsi.org/tb.aspx?tbid=836&SubTB=836
[5] https://cloudsecurityalliance.org/group/quantum-safe-security/
[6] http://cps.cqu.edu.cn/

An algebraic attack has then typically two steps : *modeling* the cryptosystem as a set of algebraic equations and then *solving* the non-linear equations to recover a secret of this cryptosystem (secret-key, message corresponding to a ciphertext,...). The NP-hardness of $\texttt{PoSSo}_q$ guarantees that a modeling is always possible. This makes the approach very general. However, the NP-hardness of $\texttt{PoSSo}_q$ also implies that the technique cannot be always efficient, i.e. polynomial-time, for all cryptographic primitives.
Almost always, there are many possible modelings. The difficultly of algebraic attacks, and the core of my research activity, is to find a "suitable" modeling. In our context, a modeling is meaningful if the set of generated equations can be solved efficiently: that is, in polynomial-time, sub-exponential time or solved in practice for real-life parameters of the cryptosystems considered. The chosen modeling also depends on the technique used for solving $\texttt{PoSSo}_q$ (exhaustive search [61], Gröbner bases [67]–[69], [120], [121], [173], [174], SAT-solvers [260], characteristic set algorithms [175], Agreeing-Gluing algorithms for sparse polynomials [243], [244], [253]–[256], ...
A feature of my research activity is the use of Gröbner bases [68], [69]. This is fundamentally motivated by the fact that efficient tools, the F_4 and F_5 algorithms [120], [121], and softwares [59], [147] permit today to perform large scale experiments. We also have a rich set of tools from computer algebra/algebraic geometry that allow to understand the complexity of computing such bases. Finally, Gröbner bases algorithms turn to be quite flexible for taking advantage of *structured systems* that naturally appear in algebraic cryptanalysis.
In this document, we shall see that the efficiency of our algebraic attacks depends on our ability to capture algebraically the intrinsic structure of the cryptosystems considered. The notion of *semi-regularity* ([25], [29], Definition 1.1.5) defines, in some sense, the minimal algebraic structure that we can expect. This is the sole structure used (Chapter 2) for an algebraic cryptanalysis of a central problem in lattice-based cryptography : the *Learning With Errors problem* (LWE, [245]).
Another weak structure is naturally induced by solving over finite fields. As a matter of fact, we can enumerate elements. This simple idea leads to the *hybrid approach* for $\texttt{PoSSo}_q$ (Section 1.2). The complexity of solving $\texttt{PoSSo}_q$ for semi-regular sequences is exponential in the number of variables. The constant depends on the ratio between the number of equations and the number of variables as well as the size of the field. The hybrid approach allows to decrease this constant and permits to improve the complexity of solving $\texttt{PoSSo}_q$ with respect to a direct Gröbner basis computation.
More sophisticated structures will appear in multivariate cryptography (Chapter 3) and code-based cryptography (Chapter 4). In both cases, we can model the cryptographic primitives with algebraic equations having a *multi-homogeneous structure* [104], [200], [211], i.e. the equations are homogeneous with respect to distinct blocks of variables (Definition 1.1.7).
There is no general polynomial system solving algorithm taking full advantage of the multi-homogeneous structure. However, significant progress have been reported on the complexity of solving *structured systems* which are particular sub-families of multi-homogeneous systems. Typically, [160], [261] presented a precise complexity analysis of *bi-linear systems* (Definition 1.1.8), i.e. each equation is the product of two linear forms on distinct blocks of variables. In the bi-linear case, the complexity of solving $\texttt{PoSSo}_q$ with Gröbner bases is exponential in the size of the smaller block of variables (Proposition 1), and not exponential in the total number of variables as for semi-regular sequences. We have identified a class of systems than can be solved more efficiently than semi-regular sequences. In particular, if the smaller block is constant then the complexity of solving $\texttt{PoSSo}_q$ is polynomial.
A natural strategy is then to find modeling with as close as possible to bi-linear structure. However, the structures that will appear in multivariate and code-based cryptography are in some

sense less general than the multi-homogenous one but not as simple as the bi-linear one. In Chapter 3, we can in fact consider an alternative modeling with a *determinantal structure* [149], [151], i.e. equations correspond to the minors of a symbolic matrix, for which we have precise complexity bounds. This determinantal structure turned to be well suited for the problems considered in Chapter 3.
In Chapter 4, we show that key-recovery against the McEliece cryptosystem [214] reduces to solving a particular multi-homogeneous system. It is probably one of the most important open question in algebraic cryptanalysis to efficiently solve this structured algebraic system. An intermediate step is to consider a sub-system which has a *quasi bi-linear structure*, i.e. equations defined over an extension which are bi-linear when viewed on the base field. In fact, McEliece is a typical case where the structure on the algebraic systems will permit to discover new structural properties on the cryptosystem itself.

Organization of the Document & Main Results

Since my PhD [236], together with collaborators, I have applied and developed algebraic cryptanalysis in various contexts : ranging from block-ciphers [4], [5], [140], [144], hash functions [47], [264], stream-ciphers [259], elliptic curves [143] and quantum-safe cryptography [3], [4], [6]–[8], [10]–[12], [15], [45], [46], [51], [52], [80], [122]–[126], [128]–[132], [134]–[139], [141], [142], [158], [159].
We provide in this document a selected overview of the results obtained in the latter topic. This choice is mainly motivated by the fact that algebraic cryptanalysis turned to be more successful in public-key cryptography, and especially against quantum-safe cryptosystems. Also, quantum-safe cryptography is currently experimenting a rapid transition from academia to industry and algebraic cryptanalysis should play an important role in this quantum-safe transition.

Chapter 1. Polynomial System Solving Over Finite Fields

$\texttt{PoSSo}_q$ being NP-Hard [176], any algorithm for $\texttt{PoSSo}_q$ should be exponential in the number of variables, i.e. any algorithm for $\texttt{PoSSo}_q$ has a complexity of the form:

$$O(2^{c\,n}), \tag{1}$$

with n being the number of variables and $c > 0$ a constant.
Due to the numerous applications of $\texttt{PoSSo}_q$ – including cryptology and coding theory – it is important to minimize the value of c in the complexity (1). In Section 1.1, we recall classical results about Gröbner bases [68], [69] and the complexity of computing such bases. In particular, we can precisely determine the value of c under some *genericity assumption* (Theorem 1.1.5) for modern Gröbner bases algorithms such as Faugère's F_4 and F_5 algorithms [120], [121].
In Section 1.2, we describe and analyze a hybrid approach for solving $\texttt{PoSSo}_q$. Section 1.2 is based, in particular, on:

Hybrid Approach

[49] L. Bettale, J.-C. Faugère, and L. Perret, "Hybrid approach for solving multivariate systems over finite fields", *Journal of Mathematical Cryptology*, vol. 3, no. 3, pp. 177–197, 2010. [Online]. Available: `http://www-salsa.lip6.fr/~jcf/Papers/JMC2.pdf`.

[52] L. Bettale, J.-C. Faugère, and L. Perret, "Solving polynomial systems over finite fields: improved analysis of the hybrid approach", in *Proceedings of the 37th International Symposium on Symbolic and Algebraic Computation*, ser. ISSAC '12, Grenoble, France: ACM, 2012, pp. 67–74. [Online]. Available: `http://www-polsys.lip6.fr/~jcf/Papers/FBP12.pdf`.

The hybrid approach is a polynomial system solving method dedicated to finite fields. The goal is to decrease the value of c by taking advantage of the finite field structure. To do so, we combine exhaustive search with Gröbner bases. The efficiency of the hybrid approach is related to the choice of a *trade-off* between the two methods. Whilst the principle of the hybrid approach is simple, its careful analysis leads to rather surprising and somehow unexpected results.
All the complexity results for the hybrid approach are obtained assuming a natural algebraic hypothesis (Assumption 1). Under this assumption, the asymptotic complexity (Theorem 1.2.1) of the hybrid approach for solving quadratic instances of $\texttt{PoSSo}_q$ is

$$O\left(2^{\left(3.27-3.5\,\log_2(q)^{-1}\right)n}\right), \text{assuming in particular that } \log(q) \ll n.$$

This is to date, the best complexity for solving $\texttt{PoSSo}_q$ when $q > 2$ [51]. This can be compared compared with a recent work [204] where the authors presented a new technique for solving $\texttt{PoSSo}_q$. We summarize below the complexity provided in [204]:

- We can solve quadratic instances of $\texttt{PoSSo}_2$ in $O^*(2^{0.8765\,n})$ (the notation O^* omits polynomial factor),
- We can solve degree-d instances of $\texttt{PoSSo}_q$ in $O^*(q^{n(1-\frac{1}{5d})}n^{3d})$ when $p = 2$ (but $q > 2$ or $d > 2$),
- We can solve degree-d instances of $\texttt{PoSSo}_q$ in $O^*(q^{n(1-\frac{1}{200d})}n^{3qd})$ when $p > 2$ and $\log(p) < 4\mathrm{e}d$, with e the Napier's constant,
- We can solve degree-d instances of $\texttt{PoSSo}_q$ in $O^*\left(q^n\left(\frac{\mathrm{e}kd}{\log(q)}\right)^{dn}\right)$ when $p > 2$ and $\log(p) \geqslant 4\mathrm{e}d$.

Note that these complexities do not rely on any assumption. For $q = 2$, we also mention that [61] describes a fast exhaustive search for solving $\texttt{PoSSo}_2$ and provide the exact cost of their approach : $4\,\log_2(n)\,2^n$ binary operations. The best method for solving $\texttt{PoSSo}_2$ is due to [30] where the authors proposed an algorithm – called `BooleanSolve` – inspired but different from the hybrid approach described in Section 1.2. When $m = n$, the deterministic variant of `BooleanSolve` has complexity bounded by $O(2^{0.841n})$, while a Las-Vegas variant has expected complexity $O(2^{0.792n})$. We summarize below the best complexities (dominant part) known for solving $\texttt{PoSSo}_q$. For $q = 2$, it is due to [30]. For the others fields, the results are obtained with the hybrid approach (Theorem 1.2.1). For completeness, we also added in the last row the results that can be obtained using [204].

q	2 [30]	2^2	2^3	2^4	2^5	2^6	2^8	2^{16}
	$2^{0.792n}$	$2^{1.5n}$	$2^{2.08n}$	$2^{2.38n}$	$2^{2.56n}$	$2^{2.67n}$	$2^{2.82n}$	2^{3n}
[204]	$2^{0.8765n}$	$2^{1.8n}$	$2^{2.7n}$	$2^{3.6n}$	$2^{4.5n}$	$2^{5.39n}$	$2^{7.2n}$	$2^{14.4n}$

We can also quantify the gain provided by the hybrid approach compared to a direct Gröbner basis method. For quadratic systems, we show – again assuming a natural algebraic assumption (Assumption 1) – that this gain is exponential in the number of variables. Asymptotically, the gain (Theorem (1.2.2)) is $2^{1.49\,n}$ when both n and q grow to infinity and $\log(q) \ll n$.
All in all, Chapter 1 presents a rather simple approach that allows to take advantage of finite fields, as well as semi-regularity. This is the minimal structure that we can expect for polynomial systems occurring in algebraic cryptanalysis. This anyway permits to minimize the complexity of solving $\texttt{PoSSo}_q$.

Chapter 2. Algebraic Algorithms for LWE

The *Learning With Errors* problem (LWE, [245]) is a fundamental problem in lattice-based cryptography. LWE can be viewed as the problem of *decoding a random linear code* over $\mathbb{F}_q$ with *Gaussian noise*, or as the problem of solving an *erroneous* system of linear equations over $\mathbb{F}_q$ with Gaussian errors. There are at least two prominent categories of algorithms for solving LWE : lattice-based and combinatorial based [7], [57]. In [16], Arora and Ge introduced the first algebraic algorithm for solving LWE. Their approach reduces LWE to finding the common root of a multivariate system of high-degree, *error-free* polynomials of the following form:

$$p_1 = P\big(c_1 - \sum_{i=1}^{n} x_j g_{i,1}\big), \ldots, p_m = P\big(c_m - \sum_{i=1}^{n} x_j g_{i,m}\big) \tag{2}$$

with $\big(\mathbf{G} = (g_{i,j}) \in \mathcal{M}_{n\times m}(\mathbb{F}_q) \times \mathbb{F}_q^n$, and $P(X) = X\prod_{i=1}^{D}(X+i)(X-i)$, where $D > 1$ is a parameter. For Gaussian errors, the classical case [16], we have $D = \tilde{O}(n^{\epsilon})$ where $\epsilon, 0 \leqslant \epsilon \leqslant 1$ is a parameter related to the Gaussian distribution and m unbounded. For uniform errors, as in [218], $D = O(1)$ and the number of samples m is bounded. In both cases, $q \in \mathrm{poly}(n)$ is prime. In [16], the authors used linearization for solving the algebraic system (2). In Chapter 2, we propose to use Gröbner bases instead and derive new asymptotic results for LWE with Gaussian errors and LWE with binary errors (i.e. $D = 2$). This last variant, that we call BinaryErrorLWE, was introduced by Micianccio and Peikert in [218]. Chapter 2 is based, in particular, on

Algebraic Algorithms for LWE

[3] M. R. Albrecht, C. Cid, J.-C. Faugère, R. Fitzpatrick, and L. Perret, "Algebraic algorithms for LWE problems", *IACR Cryptology ePrint Archive*, vol. 2014, p. 1018, 2014. [Online]. Available: http://eprint.iacr.org/2014/1018.

[6] M. Albrecht, C. Cid, J.-C. Faugère, R. Fitzpatrick, and L. Perret, "On the complexity of the Arora-Ge algorithm against LWE", in *SCC '12: Proceedings of the 3nd International Conference on Symbolic Computation and Cryptography*, Castro-Urdiales (Spain), Jul. 2012, pp. 93–99.

We proved that the use of Gröbner basis techniques yields an exponential speed-up over the basic Arora-Ge algorithm. For typical parameters of LWE, the Gröbner basis algorithm has a complexity:

$$O(2^{6.69\,n}).$$

This is obtained under a *genericity assumption*, i.e. by assuming that the algebraic system derived from (2) behaves as a semi-regular sequence. We can save a logarithmic factor in the

exponent with respect to the linearization approach of Arora and Ge [16]. This places it in the same complexity class, namely $2^{O(n)}$, than the best algorithms for solving LWE [13], [185] albeit with a larger leading constant in the exponent.
This result can be improved for BinaryErrorLWE (Section 2.2). We show, for instance, that the algebraic system (2) can be solved in subexponential time given access to a quasi-linear number of samples m. We also give precise complexity bounds for BinaryErrorLWE in function of the number of samples (Theorem 2.2.1). This addresses an open question by the designers of BinaryErrorLWE.
To derive our complexity results, we use a genericity hypothesis. Informally, we state that the algebraic systems derived from (2) are not harder to solve than semi-regular sequences (Definition 1.1.5) of the same size. Experimental evidences seem to show that random systems of equations tend to be semi-regular (some experiments are presented in Table 1.1). Hence, our semi-regularity assumption essentially states that the algebraic system (2) is neither easier nor harder to solve than random systems of equations. The genericity assumption associated to (2) is essentially equivalent to prove that a system generated by the Dth powers of m generic linear forms in n variables is semi-regular. This is a fundamental question essentially equivalent to the well-known *Fröberg's conjecture* ([168], [169], Section 1.1.2) about the existence of semi-regular sequences; which is considered to be difficult in the general case. In Section 2.3, we report some results about this open question (and the validity of our genericty assumption) in restricted cases.
Gröbner bases are a new tool in the cryptanalytic toolbox dedicated to LWE. Chapter 2 shows that we can obtain non-trivial asymptotic results against LWE and BinaryErrorLWE with algebraic attacks. There results are obtained by only using a weak structure from a Gröbner basis point of view : semi-regularity. A challenge is improve the complexity exponents of our algorithms. Another merit of this application is to establish a (somewhat unexpected) connection with Fröberg's conjecture about the existence of semi-regular sequences (Section 1.1.2).

Chapter 3. Algebraic Cryptanalysis of Multivariate Public-Key Cryptosystems

Multivariate cryptography is usually defined as the set of cryptographic schemes using the computational hardness of $\mathtt{PoSSo}_q$, or more generally the hardness of computing a Gröbner basis of a polynomial ideal. This is a classical candidate in quantum-safe cryptography. $\mathtt{PoSSo}_q$ being NP-hard, it is unlikely that it can be solved in quantum polynomial-time [35].
The most active area in the design of multivariate schemes is public-key cryptography. This is a sub-area of multivariate cryptography – known as MPKC – which has been introduced and popularized by Matsumoto, Imai and Patarin [212], [231]. An important part of my activity is related to the design and security analysis of this family of schemes, e.g. [45], [46], [48]–[52], [60], [125], [127], [128], [136], [137], [145], [159]. Chapter 3 summarizes a subset of these results. It includes a small part of [46], but it is mostly based on:

Cryptanalysis of MPKC

[51] L. Bettale, J.-C. Faugère, and L. Perret, "Cryptanalysis of HFE, Multi-HFE and variants for odd and even characteristic", *Designs, Codes and Cryptography*, vol. 69, no. 1, pp. 1–52, 2013. [Online]. Available: http://hal.inria.fr/hal-00776072.

[127] J.-C. Faugère, F. Levy-dit-Vehel, and L. Perret, "Cryptanalysis of MinRank", in *Advances in Cryptology CRYPTO 2008*, D. Wagner, Ed., ser. Lecture Notes in Computer Science, vol. 5157, Santa Barbara, CA, USA: Springer-Verlag, Aug. 2008, pp. 280–296. [Online]. Available: `http://www-salsa.lip6.fr/~jcf/Papers/crypto08.pdf`.

[138] J.-C. Faugère and L. Perret, "On the security of UOV", in *First International Conference on Symbolic Computation and Cryptography, SCC 08*, ser. LMIB, Beijing, China, Apr. 2008, pp. 103–109. [Online]. Available: `http://www-salsa.lip6.fr/~jcf/Papers/SCC08a.pdf`.

In Section 3.1, we briefly describe the general design principle of MPKC and sketch two prominent multivariate schemes : HFE [231] and UOV ([192]). The latter acronym refers to a multivariate signature scheme proposed Goubin, Kipnis and Patarin in 1999. Since, its introduction no efficient attack has been reported against UOV. HFE is probably one of the most famous MPKC. It can be used as a public-key encryption scheme and as a signature scheme. In some sense, HFE popularized algebraic cryptanalysis with Gröbner bases since Faugère and Joux demonstrated [146], [155] that modern Gröbner bases could be indeed used to attack real-world parameters of HFE.

More generally, the hard problems underlying the security of MPKC can be naturally expressed in terms of algebraic equations and are then a natural target for algebraic attacks. Typically, the public-key of MPKC is given by a set of multivariate equations and a message-recovery reduces to the hardness of solving particular instances of $\mathtt{PoSSo}_q$. In Chapter 1, we described a method which is currently the best generic method for solving $\mathtt{PoSSo}_q$. In Section 3.2, we show that this hybrid approach can be used to forge efficiently ($\approx$ 3 hours) a signature for some parameters of UOV initially recommended by the designers [192]. This is not a structural attack, since the hybrid approach only uses weak structures (semi-regularity, and finite field structure). However, it validates the relevance of the hybrid approach and shows that the parameters should be more carefully selected. In Section 3.2, we apply the hybrid approach to derive a set of minimal parameters for multivariate signature schemes (Table 3.2).

In Section 3.3, we focus on the MinRank problem [73], [82]. This is basic problem from linear algebra originally introduced in [73] where the authors proved its NP-hardness. Later, it was reformulated by Courtois [82] in the cryptographic context who described a Zero-Knowledge scheme – that we will call ZKMR – based on MinRank.

We focus on this problem because MinRank is also fundamentally related to the security of many multivariate schemes.

MinRank

Input. A set of $k+1$ matrices $\mathbf{M}_0, \mathbf{M}_1, \ldots, \mathbf{M}_k \in \mathcal{M}_{N\times n}(\mathbb{F}_q)$ and an integer $r > 0$.

Question. Find – if any – a k-tuple $(\lambda_1, \ldots, \lambda_k) \in \mathbb{F}_q^k$ such that:

$$\mathrm{Rank}\left(\sum_{i=1}^{k} \lambda_i \mathbf{M}_i - \mathbf{M}_0\right) \leqslant r.$$

We summarize a first algorithmic contribution [156] on MinRank in Section 3.3. In particular, we present an improved algebraic attack against ZKMR. To do that, we have used an algebraic modeling due to Kipnis and Shamir [194] together with Gröbner bases. This permitted to break 3 (among 5) sets of parameters recommended by the designer of ZKMR [82]. The efficiency of the attack can be explained by the multi-homogeneous structure of the modeling used. In [149], [151], Faugère, Safey El Din and Spaenlehauer consider an alternative modeling of MinRank which use symbolic minors and provide precise complexity results about solving MinRank with

such minors. In particular, this leads to an even more efficient attack against ZKMR. We review this approach in Section 3.3.3. An important feature of the minors is that the complexity of solving MinRank is well mastered with this modeling (Proposition 8).
Finally, we explain in Section 3.4 how MinRank is connected to a key-recovery against HFE. This part is based on [50], [52]. The principle of reducing key-recovery in HFE to MinRank is due to Kipnis and Shamir [194]. In Section 3.4.1, we revisit this classical MinRank against HFE. This leads to an easier and arguably most natural formulation of the attack from [194]. It makes this attack practical for a wide range of parameter whereas the original attack was considered theoretical. Finally, we build on the theoretical results [149], [151] about minors (Proposition 8) to derive in Section 3.4.2 a complexity bound for our key-recovery in the case of HFE (Section 3.4.2). In [194], it was conjectured that the basic Kipnis-Shamir (KS) attack against HFE runs in expected polynomial-time (but the complexity turned to be incorrect [103]). Under a regularity assumption, we show that solving the MinRank instances of Theorem 3.4.1 is exponential in $r = \lceil \log_q(D) \rceil$, where D is the degree of the HFE polynomial (Definition 3.1.1).
Algebraic cryptanalysis is an intrinsic technique to evaluate the security of multivariate schemes. Globally, the results of Chapter 3 provide a set of reference tools to analyze the security of MPKC. We emphasize that after an intense period of cryptanalysis, it appears now that few schemes resisted to the test of time : UOV (1999, and variants of HFE (1995, [192]). An open problem in this area is to design good proposals for quantum-safe standards. In particular, Chapter 3 demonstrates that the complexity of the best attacks against HFE are all exponential in $O(\log_q(D))$. We have then only one parameter which allows to control the security and efficiency of this scheme. We are now in a better position to derive secure parameters for HFE and variants such that the *minus variant* (HFE-, [231]) and the *vinegar variant* (HFEv, [232]).

Chapter 4. Algebraic Techniques in Code-Based Cryptography

After almost forty years now, the McEliece cryptosystem still belongs to the very few public-key cryptosystems which remain unbroken [214]. The public-key in McEliece is given by the generator matrix of a particular linear code: a *binary Goppa code* (Definition 4.1.2). Its security (message-recovery) relies upon the intractability (i.e. NP-Hardness) of decoding linear codes [40]. Decoding a random linear code is a long-standing problem whose most effective algorithms, e.g. [41], [75], [198], [199], [213], [263], have all an exponential-time complexity in the classical [270] as well than in the quantum setting [42]. Although the complexity of the best decoding attack remains exponential, progress on the exact exponent have been continuously reported. The latest result from [213] brings down the complexity to $2^{0.097\,n}$ for decoding random binary linear codes of length n. The situation is rather different for the key-recovery problem. The reference attack used to be the so-called *Support Splitting Attack* (SSA) proposed by Sendrier and Loidreau in [203]. SSA is essentially an exhaustive search on a part of the secret-key (that is, the *Goppa polynomial*). SSA only weakly uses the structure of binary Goppa codes. Besides SSA, no significant breakthrough has been been reported during the past years regarding key-recovery in McEliece. In Chapter 4, we overview new algebraic key-recovery techniques on McEliece and some variants from

Algebraic Cryptanalysis of McEliece

[122] J.-C. Faugère, V. Gauthier-Umana, A. Otmani, L. Perret, and J.-P. Tillich, "A distinguisher for high rate Mceliece cryptosystems", *IEEE Transactions on Information Theory*,

vol. 59, no. 10, pp. 6830–6844, Jun. 2013. [Online]. Available: `http://hal.inria.fr/hal-00776068`.

[131] J.-C. Faugère, A. Otmani, L. Perret, F. De Portzamparc, and J.-P. Tillich, "Structural cryptanalysis of Mceliece schemes with compact keys", *Designs, Codes and Cryptography*, pp. 87–112, Jan. 2016. [Online]. Available: `https://hal.inria.fr/hal-00964265`.

[133] J.-C. Faugère, A. Otmani, L. Perret, and J.-P. Tillich, "Algebraic cryptanalysis of McEliece variants with compact keys", in *Proceedings of Eurocrypt 2010*, ser. Lecture Notes in Computer Science, vol. 6110, Monaco: Springer Verlag, 2010, pp. 279–298. [Online]. Available: `http://www-salsa.lip6.fr/~jcf/Papers/Eurocrypt2010.pdf`.

[135] J.-C. Faugère, A. Otmani, L. Perret, and J.-P. Tillich, "Algebraic Cryptanalysis of McEliece variants with compact keys – toward a complexity analysis", in *SCC '10: Proceedings of the 2nd International Conference on Symbolic Computation and Cryptography*, London (UK), Jun. 2010, pp. 45–55. [Online]. Available: `http://www-salsa.lip6.fr/~jcf/Papers/SCC2010a.pdf`.

More precisely, Section 4.2 summarizes [131], [133], [135] and describes a new structural attack against against McEliece-like cryptosystems; introducing algebraic cryptanalysis in code-based cryptography. The very definition of a binary Goppa code – as *alternant code* (Definition 4.1.1) – implies that key-recovery in McEliece reduces to solve the following algebraic system :

$$\mathsf{A}_{\mathbf{X},\mathbf{Y}} = \bigcup_{\ell=0}^{t-1} \left\{ \sum_{j=0}^{n-1} g_{i,j} X_j^{\ell} Y_j \quad | \quad 0 \leq i \leq k-1 \right\}, \text{ where } \mathbf{G} = (g_{i,j}) \in \mathcal{M}_{k\times n}(\mathbb{F}_q). \tag{3}$$

The solutions of this system lie in $\mathbb{F}_{q^m}$ and k is an integer which is at least equal to $n - tm$ (t is the degree of the Goppa polynomial). We can remark that the system $\mathsf{A}_{\mathbf{X},\mathbf{Y}}$ is very structured. Indeed, the only monomials occurring are of the form $Y_j X_j^{\ell}$ with $\ell, 0 \leq \ell \leq t-1$, i.e. each equation are *bi-homogeneous of bi-degree* $(1, \ell)$. The number of unknowns $2n$ and the maximum degree t of the equations can be extremely high when cryptographic parameters are considered. However, the system can also be largely over-determined.
The basic algebraic modeling (3) can actually be refined by using to several properties of binary Goppa codes; leading to systems with more equations and less variables (Section 4.2). Still, it is not clear whether an efficient algebraic attack can be mounted in general. However, the approach can be very efficient as soon as the code used as an additional structure : typically for variants such as Wild McEliece (Incognito) [141] (which uses special Goppa polynomials) or compact variants [133], [142], [158] (which uses structured block matrices as a public-key). The key point was to take into account the additional structure of these McEliece variants to refine the modelings. This permitted to drastically improve the solving step. For instance, we have been able to break practically all the parameters proposed in [39] for a variant of McEliece whose public-key as a *quasi-cyclic* structure. In Section 4.2.2, we explain how the algebraic cryptanalysis allows in fact to unveil a fundamental weakness of all known compact variants of McEliece [33], [39], [220], [237].
In Section 4.3, we consider the *Goppa Code Distinguishing* (`GD`) problem. This decision problem appeared first in [87] and aims at recognizing a generator matrix of a binary Goppa code from a randomly drawn binary matrix. Before our algebraic distinguisher [122], it was assumed that no polynomial time algorithm exists that distinguishes a generator matrix of a (binary) Goppa code (or more generally *alternant codes* that includes Goppa codes) from a randomly picked

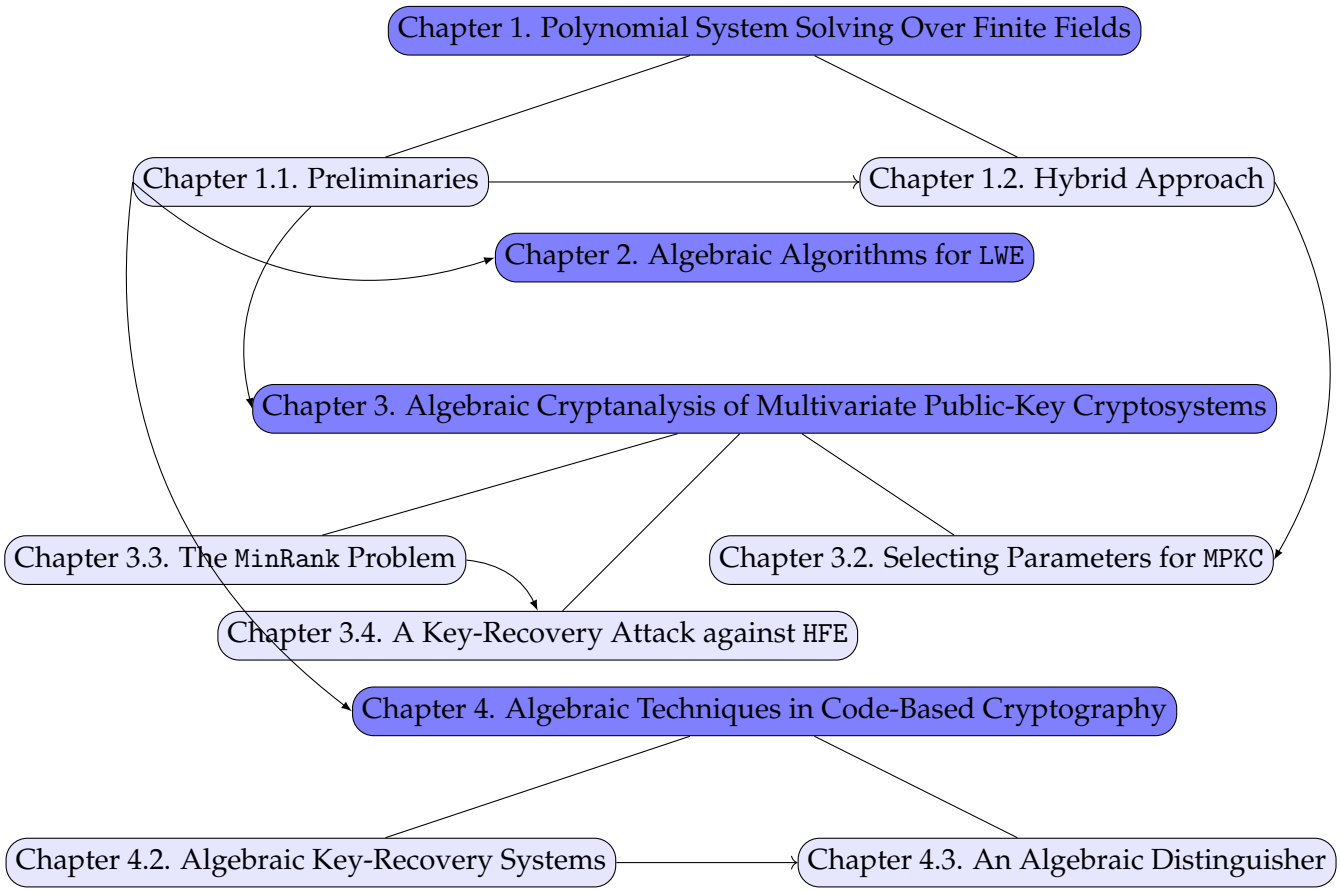
generator matrix. All in all, the hardness of GD was a classical belief in code-based cryptography, and as a consequence, a *de facto* assumption widely used in code-based cryptography. In Section 4.3, we present a deterministic polynomial-time solving GD for codes whose rate $R = \frac{k}{n}$ is close to 1. This includes in particular codes encountered with the McEliece signature scheme CFS [87], [167]. Our distinguisher is based on the algebraic modeling explained in Section 4.2. The fundamental idea of the distinguisher is to study the behavior of a Gröbner basis computation on the algebraic system (3) where $\mathbf{G} \in \mathbb{F}_q^{k \times n}$ is a random matrix or a McEliece's public-key. The very (specific) structure of the linear codes used in McEliece will induce a specific behavior at the first step of a Gröbner basis computation. This behavior is captured by the rank of a particular linear system $\mathcal{L}_{\mathbf{P}}$ defined in Section 4.3. We derive precise formula on the rank of $\mathcal{L}_{\mathbf{P}}$ when $\mathbf{G}$ is a random matrix (Theorem 4.3.1) of a McEliece's public-key (Theorems 4.3.3 and 4.3.2). It appears that the distinguisher is valid in a certain range of parameters; typically depending on the code rate. Let $n = q^m$ and assume that $q \in O(1)$. Theorem 4.3.4 states that when m tends to infinity, we can efficiently distinguish all codes with a rate bigger than :

$$1 - \sqrt{\frac{2\,m\,\log_2 q}{q^m \log_2 m}}\big(1 + o(1)\big).$$

In Chapter 4, we present a rather wide variety of algebraic techniques related to key-recovery against McEliece cryptosystems. Before the introduction of algebraic cryptanalysis in code-based cryptography [133], the only technique for key-recovery was essentially a partial exhaustive search on the secret-key [203]. All in all, the results of Chapter 4 demonstrated that algebraic cryptanalysis is a new emerging technique to assess the security of McEliece's public-key cryptosystems; much more powerful than previously known key-recovery attacks.

Reading Roadmap

The picture below is a roadmap of this document. We provide the logical dependencies between the 4 chapters summarizing my contributions.



CHAPTER 1

POLYNOMIAL SYSTEM SOLVING OVER FINITE FIELDS

This chapter is devoted to the polynomial system solving problem over $\mathbb{F}_q$ ($\texttt{PoSSo}_q$). In Section 1.1, we introduction basic notions and tools used throughout this document. This includes Gröbner bases, varieties and various tools allowing to evaluate the complexity of solving $\texttt{PoSSo}_q$ with Gröbner bases. For a more detailed introduction to these topics, we refer to classical textbooks such as [93], [177] for instance. In Section 1.2, we present an improved method for solving $\texttt{PoSSo}_q$: the so-called hybrid approach. This last part is a summary of the results in [49], [51].

1.1 Preliminaries

The general problem of polynomial system solving over a finite filed $\mathbb{F}_q$ is defined as follows:

Polynomial System Solving over a Finite Field ($\texttt{PoSSo}_q$)
Input. $p_1(x_1, \ldots, x_n), \ldots, p_m(x_1, \ldots, x_n) \in \mathbb{F}_q[x_1, \ldots, x_n]$.
Goal. Find – if any – a vector $(z_1, \ldots, z_n) \in \mathbb{F}_q^n$ such that:

$$p_1(z_1, \ldots, z_n) = 0, \ldots, p_m(z_1, \ldots, z_n) = 0.$$

From a theoretical-complexity point of view, $\texttt{PoSSo}_q$ is NP-Hard independently on the size q [176]. The proof is by reduction from 3-SAT to $\texttt{PoSSo}_q$. We construct a set of cubic polynomials from an instance, i.e. a set of logical expressions, of 3-SAT. The problem remains NP-Hard if we restrict $\texttt{PoSSo}_q$ to equations of degree smaller than 2. To do so, we partially linearize – i.e. add a new variable $y_{i,j}$ to each product $x_i x_j$ – the cubic polynomials obtained by reduction from an instance of 3-SAT. This maps cubic polynomials to quadratic polynomials.

These reductions already suggest two algorithmic techniques for solving $\texttt{PoSSo}_q$. Full linearization, i.e. we add a new variable to each monomial of the equations, seems of course very appealing. However, this approach only works if the number of equations is much bigger than the total number of variables introduced to linearize the equations. Typically, for dense quadratic equations, we have to introduce $O(n^2)$ new variables and the approach only works if $m \in \Omega(n^2)$.

3-SAT being NP-Complete, $\texttt{PoSSo}_q$ is equivalent to 3-SAT, i.e. there exists also a poly-time reduction from $\texttt{PoSSo}_q$ to 3-SAT. For $q = 2$, the reduction from $\texttt{PoSSo}_2$ to 3-SAT is explicit. As a consequence, $\texttt{PoSSo}_2$ can be solved via reduction to 3-SAT. The rationale is that SAT-solvers [118],

[221], whilst using rather simple algorithmic tools such as guessing and back-tracking, can be extremely efficient. The domain is also very active with a SAT-solver competition organized on a very regular basis [7].This connection was first used in [24], [85]. Since them, SAT-solvers have been used in various attacks against symmetric ciphers, for instance [228], [247], [260], and a dedicated SAT-solver to cryptography actually exists : CRYPTOMINISAT [8]. $\texttt{PoSSo}_q$ is a fundamental problem and many others solving techniques have been proposed. Over finite fields, these include : *optimized exhaustive search* for $\texttt{PoSSo}_2$ [61], a recent *dimension reduction* approach combined with exhaustive search in [204], *characteristic set algorithms* in [175], and the so-called *Agreeing-Gluing* approach developed in a series of papers [243], [244], [253]–[256].
Most of these techniques are either specific to $\texttt{PoSSo}_2$ ([61], SAT-solvers) or dedicated to sparse systems (Agreeing-Gluing family of algorithms). In particular, the former technique can solve [243] very *sparse instances* of $\texttt{PoSSo}_q$ in

$$q^{\frac{n}{5.7883}+O(\log n)} = q^{0.17n+O(\log n)}.$$

This complexity is obtained for instances of $\texttt{PoSSo}_q$ such that $m = n$ and where each equation contains at most 3 distinct variables.
In contrast, the recent dimension reduction technique from [204] is general and can solve :

- quadratic instances of $\texttt{PoSSo}_2$ in $O^*(2^{0.8765\,n})$ (the notation O^* omits polynomial factor),
- degree-d instances of $\texttt{PoSSo}_q$ in $O^*(q^{n(1-\frac{1}{5d})}n^{3d})$ when $p = 2$ (but $q > 2$ or $d > 2$),
- degree-d instances of $\texttt{PoSSo}_q$ in $O^*(q^{n(1-\frac{1}{200d})}n^{3qd})$ when $p > 2$ and $\log(p) < 4\mathrm{e}d$, with e the Napier's constant,
- degree-d instances of $\texttt{PoSSo}_q$ in $O^*\left(q^n\left(\frac{\mathrm{e}kd}{\log(q)}\right)^{dn}\right)$ when $p > 2$ and $\log(p) \geqslant 4\mathrm{e}d$.

The characteristic set method [175] is another interesting general technique for $\texttt{PoSSo}_q$. However, its complexity in the worst-case is not better than exhaustive search. As demonstrated in [175], the technique is however relevant and can perform much better than exhaustive search on some particular instances of $\texttt{PoSSo}_q$ (typically, systems arising in the algebraic cryptanalysis of stream ciphers). It is still an open problem to derive sharper bounds for the complexity of such method [175].
In this document, we consider another tool for solving $\texttt{PoSSo}_q$: *Gröbner bases* [68], [69]. These bases allows to solve any instance of $\texttt{PoSSo}_q$. We have also efficient algorithms, F_4 and F_5, due to Faugère [120], [121] for computing Gröbner bases as well as efficient softwares [59], [63], [147]. Gröbner bases also come with a wider variety of theoretical tools for the complexity analysis of Gröbner bases algorithms. Also, the Gröbner basis framework turns to be quiet flexible for taking advantage of *structures instances* of $\texttt{PoSSo}_q$ both in practice and in the complexity analysis. We detail these points now.
Before that, we first give the definition of a Gröbner basis. To do so, we need to recall that a *monomial* in $\mathbb{F}_q[x_1, \ldots, x_n]$ is a power product of the variables, i.e. an element of the form $x_1^{\alpha_1} \cdots x_n^{\alpha_n}$. The *leading monomial* of $p \in \mathbb{F}_q[x_1, \ldots, x_n]$ – denoted by $\mathrm{LM}(g, \prec)$ – is the largest monomial w.r.t. some admissible monomial ordering $\prec$ among the monomials of f.

[7] http://www.satcompetition.org/
[8] https://www.msoos.org/cryptominisat4/

Definition 1.1.1 (Gröbner bases). *Let $\mathcal{I} = \langle p_1, \ldots, p_m \rangle \subset \mathbb{F}_q[x_1, \ldots, x_n]$ be a polynomial ideal. A subset $G \subset \mathcal{I}$ is a Gröbner basis – w.r.t. an admissible monomial ordering $\prec$ – of $\mathcal{I}$ if:*

$$\forall p \in \mathcal{I}, \text{ there exists } g \in G \text{ such that } \mathrm{LM}(g, \prec) \text{ divides } \mathrm{LM}(p, \prec).$$

Gröbner bases as defined above are not unique. If G is a Gröbner basis of $\mathcal{I} \subset \mathbb{F}_q[x_1, \ldots, x_n]$, then we can multiply each polynomial of $g \in G$ by any non-zero constant to get another Gröbner basis G' of $\mathcal{I}$. We can define a *reduced Gröbner basis* [93, Definition 5] so that each ideal $\mathcal{I} \subset \mathbb{F}_q[x_1, \ldots, x_n]$ has a unique reduced Gröbner basis. These reduced Gröbner basis can be computed efficiently from any Gröbner basis.
It is clear from Definition 1.1.1, that the notion of Gröbner bases depends on a admissible monomial ordering. Gröbner bases have different computational and algorithmic properties with respect to the monomial ordering considered. We consider here mainly two monomial orderings : LEX and DRL. They are defined as follows. Let $\alpha = (\alpha_1, \ldots, \alpha_n)$ and $\beta = (\beta_1, \ldots, \beta_n) \in \mathbb{N}^n$, then:

- $x_1^{\alpha_1} \cdots x_n^{\alpha_n} \prec_{\mathrm{LEX}} x_1^{\beta_1} \cdots x_n^{\beta_n}$ if the first left-most nonzero entry of $\beta - \alpha$ is positive.
- $x_1^{\alpha_1} \cdots x_n^{\alpha_n} \prec_{\mathrm{DRL}} x_1^{\beta_1} \cdots x_n^{\beta_n}$ if $\sum_{i=1}^n \alpha_i < \sum_{i=1}^n \beta_i$, or $\sum_{i=1}^n \alpha_i = \sum_{i=1}^n \beta_i$ and the right-most nonzero entry of $\beta - \alpha$ is negative.

Typically, LEX-Gröbner bases (i.e. Gröbner bases w.r.t. the lexicographical ordering) allow to eliminate variables whilst DRL is a more computational-friendly ordering as we will see.
The historical method for computing Gröbner bases – known as Buchberger's algorithm – has been introduced by Buchberger in his PhD thesis [68], [69]. Many improvements on Buchberger's algorithm have been done leading – in particular – to more efficient algorithms such as the F_4 and F_5 algorithms of Faugère [120], [121]. The F_4 algorithm, for example, is the default algorithm for computing Gröbner bases in the computer algebra software MAGMA [59]. The F_5 algorithm, which is available through the FGb [147] software [9], provides today the state-of-the-art method for computing Gröbner bases. In this document, we report our experimental results by using the F_4 algorithm implemented in MAGMA or the F_5 algorithm implemented in FGb. We refer, for instance, to [25], [28], [29], [120], [121], [261], [265] for a detailed description of theses algorithms.
Besides F_4 and F_5, there is large literature of algorithms computing Gröbner bases. We mention for instance `PolyBory` [67] which is a general framework to compute Gröbner basis in $\mathbb{F}_2[x_1, \ldots, x_n]/\langle x_i^2 - x_i \rangle_{1 \leq i \leq n}$. It uses a specific data structure – dedicated to the Boolean ring – for computing Gröbner basis on top of a tweaked Buchberger's algorithm[10] .
Another technique proposed in cryptography is the `XL` algorithm [88]. It is now clearly established that `XL` is a special case of Gröbner basis algorithm [17]. More recently, a zoo of algorithms such as `G2V` [173], `GVW` [174], ..., flourished building on the core ideas of F_4 and F_5. This literature is vast and we refer to [117] for a recent survey of these algorithms. In fact, [117] introduced a new general algorithmic framework – called `RB` – that includes as a specialized versions many algorithms such as F_5, `G2V`, `GVW`, ...
Despite this important algorithmic literature, if is fair to say that MAGMA and FGb remain the references software over finite fields.

[9] `http://www-polsys.lip6.fr/~jcf/FGb/index.html`
[10] `http://polybori.sourceforge.net`

The fundamental conceptual breakthrough that leads from the historical Buchberger's algorithm to Faugère's algorithms is the intensive use of linear algebra. The bridge has been established by Lazard [197] who proved that computing a Gröbner basis for a system of homogeneous polynomials $f_1 \ldots, f_m$ is equivalent to perform Gaussian elimination on the *Macaulay matrices* $\mathcal{M}_{d,m}^{\text{acaulay}}$ for d, where $\min\big(\deg(f_1), \ldots, \deg(f_m)\big) \leq d \leq D$ for some integer D. The Macaulay matrix [208] $\mathcal{M}_{d,m}^{\text{acaulay}}$ for a set of homogeneous polynomials $f_1 \ldots, f_m$ is defined as the coefficient matrix of $(t_{i,j} \cdot f_i)$ where $1 \leq i \leq m$ and $t_{i,j}$ runs through all monomials of degree $d - \deg(f_i)$.

Theorem 1.1.1 ([197]). *Let $\mathbf{p} = (p_1, \ldots, p_m) \in \mathbb{F}_q[x_1, \ldots, x_n]^m$ be homogeneous polynomials and $\prec$ be an admissible monomial ordering. There exists a positive integer D for which a row echelon computation on all $\mathcal{M}_{d,m}^{\text{acaulay}}(p_1, \ldots, p_m)$ matrices for $d, 1 \leq d \leq D$ computes a Gröbner basis of $\langle p_1, \ldots, p_m \rangle$ w.r.t. to $\prec$.*

Let $\texttt{Monomials}_q(n, d)$ be the set of all monomials in the variables $x_1, \ldots, x_n$ of total degree equals to d in $\mathbb{F}_q[x_1, \ldots, x_n]$. We have $\#\texttt{Monomials}_q(n, d) = \binom{n+d-1}{d}$ if $q > 2$ and we set $\#\texttt{Monomials}_2(n, d) = \binom{n}{d}$ for $q = 2$ (the number of square-free monomials). The number of columns of $\mathcal{M}_{d,m}^{\text{acaulay}}(p_1, \ldots, p_m)$ is $\texttt{Monomials}_q(n, d)$.
In algebraic cryptanalysis, we are usually not interested by Gröbner bases but rather on varieties:

Definition 1.1.2. *Let $\mathbb{F}_q \subset \mathbb{L}$ and $\mathcal{I} = \langle p_1, \ldots, p_m \rangle \subset \mathbb{F}_q[x_1, \ldots, x_n]$ be an ideal. We denote by*

$$V_{\mathbb{L}}(\mathcal{I}) = V_{\mathbb{L}}(p_1, \ldots, p_m) = \big\{\mathbf{z} = (z_1, \ldots, z_n) \in \mathbb{L}^n \mid p_i(\mathbf{z}) = 0, \forall i, 1 \leq i \leq m\big\},$$

the $\mathbb{L}$-variety associated to $\mathcal{I}$, i.e. the common zeroes over $\mathbb{L}^n$ of $p_1, \ldots, p_m$. When $\mathbb{L} = \overline{\mathbb{F}_q}$, we simply denote $V(\mathcal{I}) = V_{\overline{\mathbb{F}_q}}(\mathcal{I})$.

Gröbner bases provide convenient tools for computing with varieties. For instance:

Property 1.1.1. *Let $\mathcal{I} = \langle p_1, \ldots, p_m \rangle \subset \mathbb{F}_q[x_1, \ldots, x_n]$ be a polynomial ideal. If $\#V(\mathcal{I}) = 1$, then – for any admissible monomial ordering – the (reduced) Gröbner basis G of $\mathcal{I}$ is as follows:*

$$\{x_1 - a_1, \ldots, x_n - a_n\}, \text{ with } (a_1, \ldots, a_n) \in (\overline{\mathbb{F}_q})^n.$$

In the case of unique solution, the variety can be read directly from a Gröbner basis. This particular property of Gröbner bases is independent on the monomial ordering chosen. This is not always the case, and we need a LEX-Gröbner basis in general to compute nice representations of varieties.

Theorem 1.1.2 (Elimination theorem). *Let $\mathcal{I} = \langle p_1, \ldots, p_m \rangle \subset \mathbb{F}_q[x_1, \ldots, x_n]$ be a polynomial ideal and $G \subset \mathcal{I}$ be a* LEX*-Gröbner basis of $\mathcal{I}$. It holds that:*

$$\forall \ell, 0 \leq \ell < n, G_\ell = G \cap \mathbb{F}_q[x_{\ell+1}, \ldots, x_n],$$

is a LEX*-Gröbner basis of $\mathcal{I} \cap \mathbb{F}_q[x_{\ell+1}, \ldots, x_n]$.*

In the cryptographic context, we usually want to find $V_{\mathbb{F}_q}$, i.e. the solutions over the base field. This simply requires to add the field equations:

$$V_{\mathbb{F}_q}(p_1, \ldots, p_m) = V(p_1, \ldots, p_m, x_1^q - x_1, \ldots, x_n^q - x_n) = V(p_1, \ldots, p_m) \cap \mathbb{F}_q^n.$$

It is clear that $V_{\mathbb{F}_q}$ has always a finite number solutions.
We can deduce from the elimination property (Theorem 1.1.2) that a LEX-Gröbner basis of a zero-dimensional ideal (i.e. the variety has a finite number of solutions) is always as follows :

$$\begin{cases} f_1(x_n) & = & 0, \\ f_2(x_n, x_{n-1}) & = & 0, \\ & \dots & \\ f_{k_2}(x_n, x_{n-1}) & = & 0, \\ f_{k_2+1}(x_{n-1}, x_{n-2}, x_{n-3}) & = & 0, \\ & \vdots & \end{cases}$$

To compute the variety from a LEX-Gröbner basis, we then simply have to successively eliminate variables by computing zeroes of univariate polynomials and back-substituting the results.

1.1.1 Zero-Dimensional Solving

From a practical point of view, computing (directly) a LEX-Gröbner basis is usually slower that computing a Gröbner basis w.r.t. another monomial ordering. On the other hand, it is known that computing Gröbner bases w.r.t. to a degree reverse lexicographical (DRL-Gröbner bases) is much faster in practice. The FLGM algorithm [153] permits – in the zero-dimensional case – to efficiently solve this issue. This algorithm use the knowledge of a Gröbner basis computed for a given order to construct a Gröbner for another order. As many of the algorithms presented in this document, its complexity will involved the so-called *linear algebra constant* [177].

Definition 1.1.3. *The linear algebra constant is the smallest constant $\omega, 2 < \omega \leq 3$ such that two matrices of size $N \times N$ over a field $\mathbb{K}$ can be multiplied in $O(N^\omega)$ arithmetic operations over $\mathbb{K}$. The best current bound for the linear algebra constant is $\omega < 2.3728639$ [172].*

We recall below the complexity of FLGM.

Theorem 1.1.3. *Let $\mathcal{I} \subset \mathbb{F}_q[x_1, \dots, x_n]$ be a zero-dimensional ideal and $G_{\prec_{\text{old}}}$ be a $G_{\prec_{\text{old}}}$-Gröbner basis of $\mathcal{I}$ (w.r.t. to an admissible monomial ordering $\prec_{\text{old}}$).* FGLM *[153] permits to compute a $\prec_{\text{old}}$-Gröbner basis $G_{\prec_{\text{new}}}$ of $\mathcal{I}$ knowing $G_{\prec_{\text{old}}}$ in $\mathcal{O}(n \cdot D^\omega)$, with D being the number of zeroes of $\mathcal{I}$ counted with their multiplicities.*

The complexity of the version described in [153] can be improved using sparse linear algebra techniques [152], [157]. In any case, the complexity of FGLM is polynomial in the number of solutions of the considered ideal. This suggests the following strategy for computing the solutions of a zero-dimensional system $p_1 = 0, \dots, p_m = 0$.

1. Compute a DRL-Gröbner basis G_{DRL} of $\mathcal{I} = \langle p_1, \dots, p_m \rangle \subset \mathbb{F}_q[x_1, \dots, x_n]$.
2. Compute a LEX-Gröbner basis of $\mathcal{I}$ using FGLM on G_{DRL}.

This approach is sometimes called *zero-dimensional solving* and is widely used in practice. For instance, this is the default strategy used in MAGMA when calling the function `Variety`[11].

[11] `http://magma.maths.usyd.edu.au/magma/handbook/text/1218#13584`

1.1.2 Complexity of Gröbner Bases

We review below basic results about the complexity of computing Gröbner bases.

Definition 1.1.4. *Let $p_1, \ldots, p_m \in \mathbb{F}_q[x_1, \ldots, x_n]$ be homogeneous polynomials. We shall call degree of regularity of $p_1, \ldots, p_m$, denoted by $D_{\mathrm{reg}}(p_1, \ldots, p_m)$, the smallest integer $D_0 \geq 0$ such that the polynomials of degree D_0 in $\mathcal{I} = \langle p_1, \ldots, p_m \rangle$ generate – as a $\mathbb{K}$ vector space – the set of all monomials of degree D_0 in n variables, i.e.*

$$D_{\mathrm{reg}}(p_1, \ldots, p_m) = \min \left\{ D_0 \geq 0 \mid \dim_{\mathbb{F}_q}(\{p \in \mathcal{I} \mid \deg(p) = D_0\}) = \texttt{\#Monomials}_q(n, D_0) \right\}.$$

For a set of polynomials $\mathbf{p} = (p_1, \ldots, p_m) \in \mathbb{F}_q[x_1, \ldots, x_n]^m$, the degree of regularity is defined [25], [29] from the homogeneous components of highest degree $\mathbf{p}^{\mathrm{H}} = (p_1^{\mathrm{H}}, \ldots, p_m^{\mathrm{H}}) \in \mathbb{K}[x_1, \ldots, x_n]^m$ of the polynomials of $\mathbf{p}$. We then define $D_{\mathrm{reg}}(p_1, \ldots, p_m) = D_{\mathrm{reg}}(p_1^{\mathrm{H}}, \ldots, p_m^{\mathrm{H}})$.

Once this notion fixed, we can rather easily establish an upper bound on the cost of computing a Gröbner basis [25], [28], [29], [197], [208].

Theorem 1.1.4. *Let $\omega, 2 \leq \omega \leq 3$ be the linear algebra constant, and $p_1, \ldots, p_m \in \mathbb{F}_q[x_1, \ldots, x_n]$. We assume that $\mathbf{p}^{\mathrm{H}} = (p_1^{\mathrm{H}}, \ldots, p_m^{\mathrm{H}})$ is a zero dimensional system. Let then $D_{\mathrm{reg}} = D_{\mathrm{reg}}(p_1, \ldots, p_m)$ and $\prec$ be a total degree monomial ordering. We can compute a Gröbner basis of $\langle p_1, \ldots, p_m \rangle$ with respect to $\prec$ in*

$$O\left(m \cdot \binom{n + D_{\mathrm{reg}}}{D_{\mathrm{reg}}}^{\omega} \right) \text{ arithmetic operations over } \mathbb{F}_q. \tag{1.1}$$

More precise statements about the number of arithmetic operations performed in F_5 can be found in [25], [28]. In any case, the complexity of computing a Gröbner basis is exponential in the degree of regularity. Unfortunately, this degree of regularity is difficult to compute in general; as difficult as computing the Gröbner basis. Fortunately, there is a particular class of systems for which this degree can be computed efficiently : (*regular* and) *semi-regular sequences*. The notion of regular sequences is classical [208], [209] but holds only for $m \leq n$. Semi-regular sequences, that we recall below, have been introduced in [25], [29] for over-defined systems. By essence, the algebraic systems encountered in the cryptographic context are naturally over-defined due to the field equations motivating then such a notion.

Definition 1.1.5. *We assume that $m > n$ and $q > 2$. Let $p_1, \ldots, p_m \in \mathbb{F}_q[x_1, \ldots, x_n]$ be homogeneous polynomials of degrees $d_1, \ldots, d_m$ respectively. This sequence is semi-regular if:*

1. *$\langle p_1, \ldots, p_m \rangle \neq \mathbb{F}_q[x_1, \ldots, x_n]$,*
2. *for all $i, 1 \leq i \leq m$ and $g \in \mathbb{F}_q[x_1, \ldots, x_n]$:*

$$\deg(g \cdot p_i) < D_{\mathrm{reg}}(p_1, \ldots, p_m) \text{ and } g \cdot p_i \in \langle p_1, \ldots, p_{i-1} \rangle \Rightarrow g \in \langle p_1, \ldots, p_{i-1} \rangle.$$

Now, let $p_1, \ldots, p_m \in \mathbb{F}_q[x_1, \ldots, x_n]$ be polynomials of degrees $d_1, \ldots, d_m$ respectively. We shall say the sequence $p_1, \ldots, p_m$ is semi-regular if the sequence $p_1^{\mathrm{H}}, \ldots, p_m^{\mathrm{H}}$ semi-regular.

Regular sequences are defined almost as Definition 1.1.5. The only difference is that $m \leq n$ and the second condition is simply: $g \cdot p_i \in \langle p_1, \ldots, p_{i-1} \rangle \Rightarrow g \in \langle p_1, \ldots, p_{i-1} \rangle$. Definition 1.1.5 of semi-regular sequences requires to be adapted for the Boolean polynomial ring $\mathbb{F}_2[x_1, \ldots, x_n]/\langle x_i^2 - x_i \rangle_{1 \leq i \leq n}$ [25], [29] to take into account the field equations. Note that semi-regular sequences can be also defined from a more algorithmic point of view. Semi-regular

sequences can be defined as the sequences such that all the matrices generated during a Gröbner basis computation with F_5 are of maximal possible rank [121].
The degree of regularity can be computed explicitly for semi-regular sequences [25], [27]. It is derived from a particular coefficient in a power series.

Definition 1.1.6. *Let $S(z) = \sum_{k\geq 0} c_k z^k \in \mathbb{N}[[z]]$ be a formal power series. We define the index $\mathrm{Ind}(S)$ as the first k such that $c_k \leq 0$ (if such t not exist, then $\mathrm{Ind}(S) = \infty$). We will denote by:*

$$[S(z)]_+ = \sum_{k=0}^{\mathrm{Ind}(S)-1} c_k z^k, \text{ the series truncated at } \mathrm{Ind}(S).$$

We have then:

Property 1.1.2. *A sequence $p_1, \ldots, p_m \in \mathbb{F}_q[x_1, \ldots, x_n]$ of respective degrees $d_1, \ldots, d_m$ is semi-regular if and only if its Hilbert series is given by:*

$$\mathrm{HS}_q(t) = \left[\frac{\prod_{i=1}^{m}(1-z^{d_i})}{(1-z)^n}\right]_+, \text{ if } m > n \text{ and } q > 2, \text{ and } \mathrm{HS}_2(t) = \left[\frac{(1+z)^n}{\prod_{i=1}^{m}(1+z^{d_i})}\right]_+, \text{ if } q = 2.$$

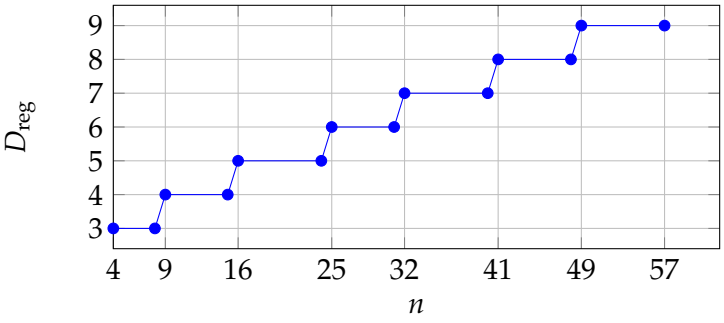
The degree of regularity of a semi-regular sequence $p_1, \ldots, p_m \in \mathbb{F}_q[x_1, \ldots, x_n]$ is given by $1 + \deg\left(\mathrm{HS}_q(t)\right)$.

For regular sequences ($m \leq n$), the degree of regularity is given by the so-called *Macaulay bound* [197], [208]:

$$\sum_{i=1}^{m}(d_i - 1) + 1. \tag{1.2}$$

For quadratic polynomials ($d_1, \ldots, d_m$), the bound is $n + 1$.
We can compute explicitly the degree of regularity of semi-regular sequences by expanding the power series (1.1.2) for specific values of $m, n, d_1, \ldots, d_m$. For example, we provide below the degree of regularity of a semi-regular system of n boolean equations in n variables.

n	D_{reg}
$4 \leq n \leq 8$	3
$9 \leq n \leq 15$	4
$16 \leq n \leq 24$	5
$25 \leq n \leq 31$	6
$32 \leq n \leq 40$	7
$41 \leq n \leq 48$	8
$49 \leq n \leq 57$	9



Property 1.1.2 and Theorem 1.1.4 allow to have a very precise knowledge about the cost of computing a Gröbner basis for semi-regular systems. We can also have asymptotic information about the trend of the regularity. For instance, it holds that [25]–[27], [29].

Theorem 1.1.5. *Let $q > 2, \alpha > 1$ be integers and $m = \lceil \alpha \cdot n \rceil$. If $p_1, \ldots, p_m \in \mathbb{F}_q[x_1, \ldots, x_n]$ is a semi-regular system of equations, then its degree of regularity behaves asymptotically as*

$$\left(\alpha - \frac{1}{2} - \sqrt{\alpha(\alpha-1)}\right) n + O(n^{1/3}). \tag{1.3}$$

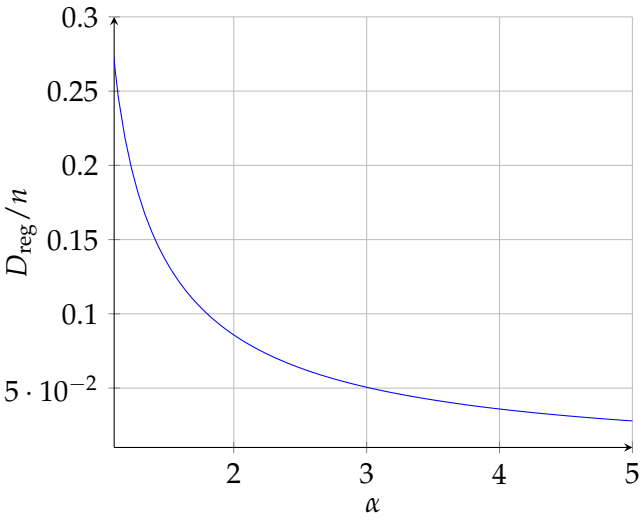
Also, if $p_1, \ldots, p_n \in \mathbb{F}_2[x_1, \ldots, x_n]$ is semi-regular, then its degree of regularity behaves asymptotically as

$$0.09\, n + O(n^{1/3}).$$

Note that we have a more simple formula in the case of a sequence of $n+1$ semi-regular polynomials $\in \mathbb{F}_q[x_1, \ldots, x_n]$ ($q > 2$). It is proved in [266] that the degree of regularity is:

$$\frac{\sum_{i=1}^{n}(d_i - 1) + 1}{2}.$$

For quadratic polynomials, the bound is $\frac{n+1}{2}$. We have divided by two the degree of regularity of 1.2 by having one more equation than the number of variables. More generally, we plotted below the constant of n in (1.3) for $\alpha \in \{1.1, \ldots, 5\}$. We can clearly observe that the degree of regularity decreases as the ratio α increases.



Fröberg's conjecture. A fundamental question in algebraic geometry is whether semi-regular sequences as defined in Definition 1.1.5 indeed exists. For regular sequences ($m \leq n$), the question is solved and well understood [168]. In the semi-regular case ($m > n$), the question remains vastly open. A famous conjecture of algebraic geometry is then attached to the existence of semi-regular sequences : the Fröberg conjecture [168]. This conjecture is classically stated over a field $\mathbb{K}$ of characteristic zero. The *Zariski topology* on $\mathbb{K}^n$ is the standard topology in algebraic geometry. In Zariski's topology, *closed sets* are the algebraic sets. The conjecture states that semi-regular sequences form a dense subset among the set of all sequences. This is equivalent to prove that there exists a *non-constant* polynomial F that vanishes the coefficients of non semi-regular sequences. For semi-regular sequences, it is not difficult to find such polynomial. However, the delicate point is to prove that the polynomial is not zero on the coefficients of a least one sequence of m polynomials. To prove Fröberg's conjecture, it is then sufficient to demonstrate that one particular family of $m > n$ polynomials in $\mathbb{K}[x_1, \ldots, x_n]$ is semi-regular for any sufficiently big n and any $m > n$. In finite fields, Zariski's topology is meaningless since all sets are algebraic. However, the proof strategy is essentially similar. Given the non-constant polynomial F, we can use Schwartz-Zippel-DeMillo-Lipton lemma [102], [252], [276] to upper bound the probability that F vanishes; that is the probability that a random sequence is non semi-regular.

From an experimental point of view, the conjecture seems indeed to hold. For instance, we report below some experimental results where we fixed the number of equations $m = \lceil n \log_2(n) \rceil$. We then randomly sampled 1000 quadratic systems in $\mathbb{F}_q[x_1, \ldots, x_n]$ where $q = 3$, q is a next prime bigger than n and q is a next prime bigger than n^2. The choice for these parameters are motivated by Chapter 2 where we consider in particular systems of equations with similar parameters. We then computed the Hilbert series of these systems and compared with the generic Hilbert series (Proposition 1.1.2). We reported the proportion of systems whose Hilbert series is exactly equal to the generic Hilbert series.

Table 1.1: Experimental results about semi-regularity.

n	Proportion (NextPrime(n))	Proportion (NextPrime(n^2))
20	1	1
26	1	1
35	1	1

In [267], the authors performed similar experiments on $\mathbb{F}_2[x_1,\ldots,x_n]/\langle x_i^2 - x_i\rangle_{1\leq i\leq n}$. The situation is not as clean than for bigger q as the authors in [267] managed to find non semi-regular sequences by such random walk. Still, a majority of the sequences turned to be also semi-regular. Fröberg's conjecture has been proven in a restricted number of cases: over a sufficiently big finite field, $n = 2, 3$, $m = n+1$, m polynomials of degree 2 with $n \leq 11$, and m polynomials of degree 3 with $n \leq 8$ [14], [168], [169]. More recently, [225] proves the conjecture for $n \geq 4$, equations of degree $d \geq 2$ and when

$$\binom{n+d-1}{d} - n \leq m \leq \binom{n+d-1}{d}.$$

For $d = 2$, the bound is for example $\binom{n-1}{2} \leq m \leq \binom{n+1}{2}$.
In Chapter 2, we consider a particular family of systems arising from an algebraic modeling of LWE, i.e. power of random affine forms. Note that the homogeneous parts of highest degree of such systems are powers of random linear forms. To derive complexity results, we need to assume that such systems behave as semi-regular sequences (Assumption 2). As a support to this assumption, it is interesting to remark that power of generic linear forms have been already investigated in the literature as potential candidates for semi-regularity. Fröberg and Hollman [169] demonstrated that the square of $n+1$ linear forms in n variables is sufficiently generic, i.e. its Hilbert series is equal to the Hilbert series of a semi-regular sequence (Property 1.1.2). This proves that semi-regular sequences of $n+1$ equations in n variables exist.

Structured systems. A common theme that appears all over this document is the necessity to exploit the structure of the polynomial systems arising in the various algebraic cryptanalysis. A rich structure which is common between multivariate cryptography (Chapter 3) and code-based cryptography (Chapter 4) is the multi-homogeneity defined [104], [200], [211] below:

Definition 1.1.7. *Let $\mathbf{p} = (p_1,\ldots,p_m) \in \mathbb{F}_q[x_1,\ldots,x_n]^m$ be homogeneous polynomials. We also consider $\{\mathbf{X}^{(1)},\ldots,\mathbf{X}^{(k)}\}$ a partition of $\mathbf{X} = \{x_1,\ldots,x_n\}$ such that :*

$$\mathbf{X}^{(j)} = \{x_{j_1},\ldots,x_{j_{k_j}}\}.$$

We shall say that $\mathbf{p}$ is multi-homogeneous if the polynomials p_i are homogenous w.r.t. the $\mathbf{X}^{(j)}$'s.

Many progresses have been reported on the complexity of solving structured systems which are sub families of multi-homogeneous systems. In particular :

Definition 1.1.8 ([160], [261])**.** *Let n_1 and n_2 be positive integers. Let also $\mathbf{X} = [x_1,\ldots,x_{n_1}]$ and $\mathbf{Y} = [y_1,\ldots,y_{n_2}]$. We shall say that $f \in \mathbb{F}_q[\mathbf{X},\mathbf{Y}]$ is bi-homogeneous of bi-degree (d_1,d_2) if:*

$$\forall \alpha,\mu \in \mathbb{F}_q,\ f(\alpha\mathbf{X},\mu\mathbf{Y}) = \alpha^{d_1}\mu^{d_2} f(\mathbf{X},\mathbf{Y}).$$

$f \in \mathbb{F}_q[\mathbf{X}, \mathbf{Y}]$ is bi-linear if it is of bi-degree $(1, 1)$. f is affine bi-linear if it is bi-linear up to a the constant term.
$f \in \mathbb{F}_{q^m}[\mathbf{X}, \mathbf{Y}]$ is quasi bi-linear if it is of bi-degree $(2^u, 2^{u'})$ for $0 \leqslant u, u' \leqslant m-1$. Finally, f is affine quasi bi-linear if it is quasi bi-linear up to a the constant term.

In particular, we have the following complexity result about affine bi-linear systems [160], [261].

Proposition 1. *Let $n_1, n_2 \in \mathbb{N}$, $\mathbf{X} = [x_1, \ldots, x_{n_1}]$ and $\mathbf{Y} = [y_1, \ldots, y_{n_2}]$ be blocks of variables. Let also $\mathbf{p} = (p_1, \ldots, p_m) \in \mathbb{F}_q[\mathbf{X}, \mathbf{Y}]$ be a set of affine bi-linear equations. If $m \leq n_1 + n_2$, and assuming a genericity condition, then it holds that:*

$$D_{\text{reg}}(p_1, \ldots, p_m) \leq \min(n_1, n_2) + 2.$$

This has to be compared with the degree of regularity D_{reg} of a regular sequence (1.2) which is $\leq n_1 + n_2$. In the affine bi-linear case, the degree of regularity D_{reg} only depends on the size of the smallest block. If $\min(n_1, n_2)$ is constant, then we can compute a Gröbner basis in polynomial-time (Theorem 1.1.4). The structure allows here to change the complexity class. Further structures have been considered in the literature : *weighted homogeneity* [150], *invariance by the action of some groups* [163], [265], *fewnomials* [162] and *determinantal systems* [149], [151]. The latter structure will be used and explained in Chapter 3.
In [161], [265], the authors introduced the concept of sparse Gröbner bases. This is an analogous to classical Gröbner bases which allows one to take into account the monomial structure (such as affine-bilinear, bi-homogeneous) – if any – of the polynomials. This permits a general treatment of structured systems as [161] proposed a complete suite of dedicated algorithms for computing efficiently such sparse Gröbner bases, i.e. `sparse-F5` and `sparse-FGLM`, as well as the combinatorial tools to study the complexity of these algorithms. The concept allows us to handle overdetermined systems of non-linear equations with a given monomial structure that were not treated in [160], [261]. Note that [161], [265] deal with the *unmixed case*, i.e. each equation of the system must have the same monomial structure (this excludes general multi-homogeneous systems).

1.2 Hybrid Approach

We describe in this part a hybrid approach for solving $\texttt{PoSSo}_q$. This section is, in particular, based on [49], [51]. The hybrid approach is a technique that combines exhaustive search and Gröbner bases (Section 1.1). The principle is to compute several Gröbner bases of smaller systems instead of one. We solve several systems obtained by fixing k variables. In what follows, we shall refer to k as the *trade-off*. The complete set of solutions is recovered from the computation of q^k varieties.
The principle is rather natural and a similar approach using the XL algorithm [88] – the so-called FXL algorithm – has been already proposed. It has been further studied in [274]. Even if XL can be considered as a special case of Gröbner basis algorithm [17], the analysis we provide here [49], [51] is tighter than [274]. We also give concrete asymptotic estimates of the complexity of hybrid approach.

The rationale of the hybrid approach is that the cost of computing a Gröbner basis decreases when the ratio m/n between the number of equations m and number of variables n increases (Theorem 1.1.5). Thus, the gain obtained by working on systems with less variables may overcome the loss due to the exhaustive search on the fixed variables.

The general hybrid approach depending on the trade-off parameter $k \in \mathbb{N}$ is given below.

Algorithm 1 GenHybridSolving (zero-dimensional)

Input. $p_1, \ldots, p_m \in \mathbb{F}_q[x_1, \ldots, x_n]$ and $k \in \mathbb{N}$
Output. $\mathcal{S} = \{(z_1, \ldots, z_n) \in \mathbb{F}_q^n \mid p_i(z_1, \ldots, z_n) = 0, \forall i, 1 \leqslant i \leqslant m\}$.
  $\mathcal{S} := \emptyset$
  for all $\mathbf{v} = (v_1, \ldots, v_k) \in \mathbb{F}_q^k$ **do**
    Find the set of solutions $\mathcal{S}_{\mathbf{v}} \subset \mathbb{F}_q^{n-k}$ of
    $p_1(x_1, \ldots, x_{n-k}, v_1, \ldots, v_k) = 0, \ldots, p_m(x_1, \ldots, x_{n-k}, v_1, \ldots, v_k) = 0$
    using the zero-dim solving strategy (Sec. 1.1.1)
    $\mathcal{S} := \mathcal{S} \cup \{(z_1, \ldots, z_{n-k}, v_1, \ldots, v_k) \mid (z_1, \ldots, z_{n-k}) \in \mathcal{S}_{\mathbf{v}}\}$
  end for
  return $\mathcal{S}$

Let $C_{\mathrm{GB}}\left(n, m, D_{\mathrm{reg}}\right)$ be the complexity of computing a DRL-Gröbner basis of a system of m equations of $\mathbb{F}_q[x_1, \ldots, x_n]$ in n variables (Theorem 1.1.4). The hybrid approach has complexity:

Proposition 2. *Let $p_1, \ldots, p_m \in \mathbb{F}_q[x_1, \ldots, x_n]$ be algebraic equations of respective degrees $d_1 \geqslant \cdots \geqslant d_m$. Let k be a non-negative integer and $\mathrm{D}_{\mathrm{reg}}^{max}(k)$ (resp. $\mathrm{D}^{max}(k)$) be the maximum degree of regularity (resp. maximum number of solutions in the algebraic closure of $\mathbb{F}_q$ counted with multiplicities) of all the systems:*

$$\left\{p_1(x_1, \ldots, x_{n-k}, v_1, \ldots, v_k), \ldots, p_m(x_1, \ldots, x_{n-k}, v_1, \ldots, v_k)\right\}, \text{ for any } (v_1, \ldots, v_k) \in \mathbb{F}_q^k.$$

We define:

$$C_{Hyb}(k) = q^k\Bigg(\underbrace{C_{\mathrm{GB}}\left(n-k, m, \mathrm{D}_{\mathrm{reg}}^{max}(k)\right)}_{\text{Gröbner basis}} + \underbrace{O\big((n-k)\,\mathrm{D}^{max}(k)^{\omega}\big)}_{\text{change of ordering}}\Bigg). \tag{1.4}$$

The complexity of the hybrid approach is dominated by:

$$\min_{0 \leqslant k \leqslant n}\left(C_{Hyb}(k)\right). \tag{1.5}$$

This is the complexity of computing q^k DRL-Gröbner bases of polynomial systems having m equations, $n-k$ variables, respective degrees $d_1 \geqslant \cdots \geqslant d_m$, plus the cost of performing a change of ordering with FGLM (Theorem 1.1.3). In order to study the asymptotic behavior of the hybrid approach, we assume a regularity condition about the sub-systems arising during Algorithm 1.

Assumption 1. *Let $p_1, \ldots, p_m \in \mathbb{F}_q[x_1, \ldots, x_n]$ be random algebraic equations of respective degrees $d_1 \geqslant \cdots \geqslant d_m$. Let $\beta_{\min}, 0 < \beta_{\min} < 1$ be a value that will be specified later. Then, for any $k, 0 \leqslant k \leqslant \lceil \beta_{\min} n \rceil$, and for each vector $(v_1, \ldots, v_k) \in \mathbb{F}_q^k$, the system:*

$$\{p_1(x_1, \ldots, x_{n-k}, v_1, \ldots, v_k), \ldots, p_m(x_1, \ldots, x_{n-k}, v_1, \ldots, v_k)\}$$

is semi-regular for n large enough.

We emphasize that Assumption 1 has been experimentally verified [30] for a large amount of random quadratic binary systems. In [49], such assumption has been verified for large q on algebraic systems coming coming from multivariate signature schemes such as `UOV` ([192],

Section 3.1.2). However, such systems are naturally under-defined. Thus, the total number of variables to be fixed ($m - n$ variables to have a square system plus k variables due to the hybrid approach) is sufficiently large to assume that the algebraic systems obtained after specialization behave as a semi-regular system. Note also that we performed some experiments to check this assumption for random systems of equations. We experimentally verified that Assumption 1 holds for random square systems with various values of $n, 6 \leq n \leq 16$, and with parameters $q > 2, \beta_{\min}$ as in Table 1.2. The assumption has also been investigated in [30] and turns to hold with high probability. However, whilst we never found a counter-example for Assumption 1 when $q > 2$ the situation is slightly different for the binary field in the sense that counter-examples can be found from time to time; according to [30], [267].
Under Assumption 1, all the sub-systems solved during the hybrid approach have – for a fixed k – the same degree of regularity. We denote this regularity by $\mathrm{D}_{\mathrm{reg}}(k)$, i.e. $\mathrm{D}_{\mathrm{reg}}^{\max}(k) = \mathrm{D}_{\mathrm{reg}}(k)$. Furthermore, the number of solutions of an over-determined semi-regular system of equations is always 0 or 1 (i.e. $0 \leq \mathrm{D}^{\max}(k) \leq 1$ as soon as $k > 0$). This allows to neglect the cost of FGLM in the complexity (1.4).

1.2.1 Analysis of the Hybrid Approach

In what follows, we always assume that $q > 2$. Our approach can be also applied when $q = 2$, but will be less efficient than the `BooleanSolve` algorithm from [30] in this special case.

Best Trade-Off for Quadratic Systems. To perform the asymptotic analysis, we need to assume – a priori – what is the global trend of the trade-off k. At first glance, it seems (rather) natural to believe that k is going to be small and should be then a constant. This is what was initially assumed in [49]. Surprisingly enough, we proved in [52] that the best trade-off is obtained asymptotically by fixing $\beta_0\, n$ variables, where β_0 is independent of n.

Asymptotic Equivalent of the Regularity. From now on, we set $m = \lceil \alpha\, n \rceil$, with $\alpha \geq 1$ being a constant. According to the previous paragraph, the best trade-off is obtained for a k of the form $\beta \cdot n$. Thus, the hybrid approach considers sub-systems having $n' = (1-\beta)n$ variables and $m = \frac{\alpha}{1-\beta}\,(1-\beta)\,n = \theta\, n'$ equations. For such systems, Theorem 1.1.5 yields:

$$D_{\mathrm{reg}}(n', m) \sim \left(\theta - \frac{1}{2} - \sqrt{\theta\,(\theta-1)}\right) n' + O(n'^{1/3}).$$

Assuming a trade-off of the form $\beta \cdot n$, we get that any sub-system occurring in the hybrid approach has a degree of regularity asymptotically equivalent to $\gamma\, n' + O(n'^{1/3})$, with:

$$\gamma = \left(\alpha - \frac{1-\beta}{2} - \sqrt{\alpha\ (\alpha+\beta-1)}\right). \tag{1.6}$$

Implicit Form of the Best Trade-Off. The best trade-off at infinity $k_0 = \lceil \beta_0\, n \rceil$ can be obtained by solving an implicit equation. The idea is to derive an equivalent of the logarithmic derivative of C_{Hyb} using the degree of regularity (1.6).

Proposition 3. *Let $\mathbf{p} = (p_1, \ldots, p_m) \in \mathbb{F}_q[x_1, \ldots, x_n]^m$ be a system of quadratic equations verifying Assumption 1. Finally, let $\omega, 2 \leq \omega < 3$ be the linear algebra constant and $A_\infty(\beta) = \log(q) +$*

$$\omega \log(1-\beta) - \frac{\omega}{2}\left(1 + \sqrt{\frac{\alpha}{\alpha+\beta-1}}\right) \log\left(D_1(\alpha,\beta)\right) - \frac{\omega}{2}\left(1 - \sqrt{\frac{\alpha}{\alpha+\beta-1}}\right) \log\left(D_2(\alpha,\beta)\right),$$

with $D_1(\alpha,\beta) = \alpha + \frac{1-\beta}{2} - \sqrt{\alpha\ (\alpha+\beta-1)}$ *and* $D_2(\alpha,\beta) = \alpha - \frac{1-\beta}{2} - \sqrt{\alpha\ (\alpha+\beta-1)}$.
The best trade-off for solving $\mathbf{p}$ *with the hybrid approach is asymptotically to fix* $k_0 = \lceil \beta_0\, n \rceil$ *variables, where* $\beta_0, 0 < \beta_0 \leqslant 1$ *is a root of* A_∞. *The coefficient* β_0 *is independent on the number of variables* n.

A root β_0 of $A_\infty(\beta)$ can be computed numerically. In Table 1.2, we present the trade-off β_0 obtained for various values of α and q. We computed the roots of $A_\infty(\beta)$ with the MAPLE software [23].

Table 1.2: Sample values for β_0 depending on several values of α and q with $\omega = 2.3728639$. An entry is empty when there is no positive solution (i.e. best trade-off is $k = 0$).

q	2^2	2^3	2^4	2^5	2^6	2^8	2^{15}
$\beta_0\ (\alpha = 1)$	0.51	0.34	0.23	0.16	0.12	0.06	0.018
$\beta_0\ (\alpha = 1.1)$	0.46	0.28	0.16	0.08	0.03	–	–
$\beta_0\ (\alpha = 1.25)$	0.39	0.18	0.04	–	–	–	–
$\beta_0\ (\alpha = 1.5)$	0.27	0.02	–	–	–	–	–
$\beta_0\ (\alpha = 1.75)$	0.15	–	–	–	–	–	–
$\beta_0\ (\alpha = 2)$	0.03	–	–	–	–	–	–
$\beta_0\ (\alpha = 3)$	–	–	–	–	–	–	–

For square systems, i.e. $m = n$ and $\alpha = 1$, Proposition 3 can be refined as follows.

Proposition 4. *Let* $\mathbf{p} = (p_1, \ldots, p_m) \in \mathbb{F}_q[x_1, \ldots, x_n]^m$ *be a system of quadratic equations verifying Assumption 1. Let* $\omega, 2 \leq \omega < 3$ *be the linear algebra constant and* B_∞ *be defined as*

$$\begin{aligned} B_\infty(\nu) = \log(q) + \omega\, \log(2\nu+2) + \omega \log\left(\frac{\nu-1}{2\nu^2}\right) \\ - \frac{\omega}{2}\,(1+\nu)\,\log(3\nu+1) - \frac{\omega}{2}\,(1+\nu)\,\log\left(\frac{\nu-1}{2\nu^2}\right) \\ - \frac{\omega}{2}\,(1-\nu)\,\log(\nu-1) - \frac{\omega}{2}\,(1-\nu)\,\log\left(\frac{\nu-1}{2\nu^2}\right). \end{aligned}$$

The best trade-off for solving $\mathbf{p}$ *with the hybrid approach is asymptotically to fix* $k_0 = \left\lceil \frac{n}{\nu_0^2} \right\rceil$ *variables, where* $\nu_0, 0 < \nu_0 \leqslant 1$ *is a root of* $B_\infty(\nu)$. *The coefficient* $\beta_0 = \frac{1}{\nu_0^2}$ *is independent of* n.

We show in Table 1.3 the value of $\beta_0 = \frac{1}{\nu_0^2}$ with respect to several usual sizes of field q. We compare these values with the exact ratio β_0 when $n = 100$ and $n = 200$ (once the parameters are fixed, we can compute exact value β_0^{exact} minimizing the complexity of the hybrid approach). The table shows that our approximation matches well with the expected value.

Table 1.3: Sample values for β_0 for several field sizes with $\omega = 2.3728639$. We need less variables to reach the best trade-off when the field is bigger.

q	2^2	2^3	2^4	2^5	2^6	2^8	2^{15}
β_0	0.51	0.34	0.23	0.16	0.12	0.06	0.018
β_0^{exact}, $n = 100$	0.59	0.35	0.25	0.14	0.12	0.08	0.02
β_0^{exact}, $n = 200$	0.55	0.39	0.24	0.17	0.17	0.09	0.02

Note that the proportion of variables which needs to be fixed tends to 0 when the size of the field increases. This is consistent with the intuition that the exhaustive search becomes less interesting for too large fields.

Asymptotic Equivalent of the Best Trade-Off. From now on, we consider only the case $m = n$. Table 1.3 suggests that when q grows, $\beta_0 = \frac{1}{\nu_0^2}$ decreases. This means that $\nu_0 \to \infty$ when $q \to \infty$. This remark combined with Proposition 4 leads to the following result.

Proposition 5. *Let $\mathbf{p} = (p_1, \ldots, p_n) \in \mathbb{F}_q[x_1, \ldots, x_n]^n$ be a system of quadratic equations verifying Assumption 1. Asymptotically, the best trade-off for solving $\mathbf{p}$ with the hybrid approach is to fix $k_0 = \lceil n\,\beta_0 \rceil$ variables, with:*

$$\begin{aligned} \beta_0 &= \left(\frac{3\,\omega\,\log(3)}{6\,\log(q) + 6\,\omega\log(2) - 4\,\omega - 3\,\omega\,\log(3)} \right)^2, \\ &= \frac{10.86\,\omega^2}{(4.16\,\log_2(q) - 3.14\,\omega)^2}. \end{aligned} \tag{1.7}$$

Table 1.4: Sample values for β_0 for several field sizes with $\omega = 2.3728639$. We need less variables to reach the best trade-off when the field is bigger.

q	2^2	2^3	2^4	2^5	2^6	2^8	2^{15}
β_0	0.51	0.34	0.23	0.16	0.12	0.06	0.018

.

1.2.2 Complexity of the Hybrid Approach – An Asymptotic Equivalent

We are now in position to derive the (asymptotic) complexity of the hybrid approach. We use the value of β_0 provided in Proposition 5 together with the degree of regularity of (1.6). This leads to:

Theorem 1.2.1. *The complexity of the hybrid approach – using the trade-off $k_0 = \lceil \beta_0\, n \rceil$ of Proposition 5 – is asymptotically equivalent to*

$$O\left(2^{n\,\omega\left(1.38 - 0.63\,\omega\log_2(q)^{-1}\right)}\right), \text{ when } n \to \infty, q \to \infty \text{ and } \log(q) \ll n. \tag{1.8}$$

If $\omega = 2.3728639$ for instance, the complexity of the hybrid approach is:

$$2^{\left(3.27 - 3.5\,\log_2(q)^{-1}\right)n}.$$

Table 1.5: Sample values for (1.8) for several field sizes with $\omega = 2.3728639$.

q	2^2	2^3	2^4	2^5	2^6	2^8	2^{15}
	$O(2^{1.5n})$	$O(2^{2.08n})$	$O(2^{2.38n})$	$O(2^{2.56n})$	$O(2^{2.67n})$	$O(2^{2.82n})$	$O(2^{3n})$

Asymptotic Gain of the Hybrid Approach. We can quantify the gain of the hybrid approach with respect to a direct Gröbner basis approach.

Theorem 1.2.2. *Let $\mathbf{p} = (p_1, \ldots, p_m) \in \mathbb{F}_q[x_1, \ldots, x_n]^m$ be quadratic equations verifying Assumption 1. When $n \to \infty$, $q \to \infty$ and as long as $n \gg \log_2(q)$, the gain of the hybrid approach compared to a direct Gröbner basis approach is asymptotically $2^{0.62\,\omega\, n}$.*

Theorem 1.2.2 gives a trend of the asymptotic gain. It shows the overall efficiency of the hybrid approach compared to the simple Gröbner basis approach. For ω as before, we get a speed-up of $2^{1.47\,n}$ for instance. Asymptotically, the hybrid approach is then always better than a direct solving. Eventually, when q is too big (with respect to n), the cost of an exhaustive search, even in one single variable, will be too expensive compared to a Gröbner basis computation.

1.3 Final Remarks

We have presented a rather simple approach that allows to take advantage of finite fields to improve the solving of $\texttt{PoSSo}_q$. This approach turns to be the most efficient for $\texttt{PoSSo}_q$ when $q > 2$. We already mentioned that `BooleanSolve` is more efficient when $q = 2$. In fact, `BooleanSolve` [30] can be viewed as a hybrid approach allowing to use 2 for the linear algebra constant ω. This is not detailed here, but we could adapt the principle of `BooleanSolve` for larger fields. This would be essentially equivalent to set $\omega = 2$ in Theorem 1.2.1. Thus, we could get:

Table 1.6: Sample values for (1.8) for several field sizes with $\omega = 2$.

q	2^2	2^3	2^4	2^5	2^6	2^8	2^{15}
	$O(2^{1.49n})$	$O(2^{1.91n})$	$O(2^{2.12n})$	$O(2^{2.25n})$	$O(2^{2.33n})$	$O(2^{2.44n})$	$O(2^{2.58n})$

CHAPTER 2

ALGEBRAIC ALGORITHMS FOR LWE

*This chapter describes algebraic attacks against the Learning with Errors (*LWE*, [245], [246]) problem. It is based, in particular, on [3], [6]. We consider mainly here a variant of* LWE *with binary errors : the so-called* BinaryErrorLWE *[218]. We first describe (Section 2.1) these problems, and an algebraic algorithm due to Arora and Ge [16]. We then present a simple extension using Gröbner bases, and derive new asymptotic results about the complexity of solving* BinaryErrorLWE *with such extension (Section 2.2). To derive these results, we need to assume a genericity hypothesis (Assumption 2) that we discuss in Section 2.3.*

2.1 LWE and BinaryErrorLWE

We continue the document with the newest application of Gröbner bases in quantum-safe cryptography. We consider algebraic attacks against LWE and one of its variants using binary errors : BinaryErrorLWE [218].
Since its introduction, LWE has proven to be a rich and versatile source of many innovative cryptographic constructions, such as an oblivious transfer [235], a leakage-resilient cryptosystem [2], a traitor tracing scheme [201], a homomorphic encryption scheme [66], [178] and many more ... In addition, public-key schemes based on LWE, such as [58], [116] appeared to be a serious candidate for quantum-safe standards.

Definition 2.1.1 (LWE [245], [246])**.** *Let $m > n \geq 1$ be integers, q be an odd positive integer, χ be a probability distribution on $\mathbb{Z}_q$ and $\mathbf{s} \in \mathbb{Z}_q^n$ be a secret vector. We denote by $L_{\mathbf{s},\chi}^{(n)}$ the probability distribution on $\mathcal{M}_{n\times m}(\mathbb{Z}_q) \times \mathbb{Z}_q^m$ obtained by choosing $\mathbf{G} \in \mathcal{M}_{n\times m}(\mathbb{Z}_q)$ uniformly at random, sampling a vector $\mathbf{e} \in \mathbb{Z}_q^m$ according to χ^m, and returning $(\mathbf{G}, \mathbf{s}\cdot\mathbf{G} + \mathbf{e}) = (\mathbf{G}, \mathbf{c}) \in \mathcal{M}_{n\times m}(\mathbb{Z}_q) \times \mathbb{Z}_q^m$. (Search)* LWE *is the problem of finding $\mathbf{s} \in \mathbb{Z}_q^n$ from $(\mathbf{G}, \mathbf{s} \times \mathbf{G} + \mathbf{e})$ sampled according to $L_{\mathbf{s},\chi}^{(n)}$.*

Typically, $\chi_{\alpha,q}$ is a *discrete Gaussian distribution* over $\mathbb{Z}$ which returns an integer x with probability $\exp\left(-\pi\, x^2/s^2\right) / \sum_{y\in\mathbb{Z}} \exp\left(-\pi\, y^2/s^2\right)$, where $s = \alpha q$, considered modulo q. A typical setting is $\alpha q = O(n^\epsilon)$, with $\epsilon, 0 \leqslant \epsilon \leqslant 1$. It has been shown that as soon as $\epsilon > 1/2$, worst-case $\mathtt{GapSVP}_{\tilde{O}(n/\alpha)}$ reduces to average-case LWE [65], [234], [245], [246]. Thus, any algorithm solving LWE, for $\epsilon > 1/2$, can solve worst-case instances of $\mathtt{GapSVP}_{\tilde{O}(n/\alpha)}$. It is commonly admitted that only exponential, classical or quantum, algorithms exist for solving $\mathtt{GapSVP}_{\tilde{O}(n/\alpha)}$.

LWE with Binary Errors. We consider in this chapter a variant of LWE with errors values in $\{0,1\}$. This variant was introduced by Micianccio and Peikert in [218]. This generalizes an earlier result of Döttling and Müller-Quade [111] who first introduced a version of LWE with uniform errors whilst keeping a strong security reduction to lattice problems. These two works highlight an interest in studying variants of LWE with small errors. For instance, the current most efficient key-exhange scheme based on LWE from [58] considers errors with a very narrow Gaussian noise. From a practical perspective, these variants are interesting because they allow to forgo Gaussian sampling (with large parameters) which is often the most expensive step when implementing lattice-based cryptography. In this regard [218] represents a significant step forward as it allows to sample the error from a binary distribution while still regaining a reduction to GapSVP$_{\zeta}$ albeit with a severe limit on the number of samples m.

Theorem 2.1.1 (BinaryErrorLWE [218]). *Let $n, m = n\big(1 + o(1)\big)$ be integers, and $q \geq n^{O(1)}$ be a sufficiently large polynomially bounded (prime) modulus. Then, solving* LWE *with parameters n, m, q and independent uniformly random binary errors is at least as hard as approximating lattice problems in the worst-case on $\Theta\big(n/\log(n)\big)$-dimensional lattices within a factor $\tilde{\mathcal{O}}(\sqrt{n} \cdot q)$.*

Remark that the security reduction in BinaryErrorLWE no longer depends of the noise, as in LWE, but on the number of samples m. We denote by $\mathcal{U}(\mathbb{F}_2)$ the uniform distribution on $\{0,1\}^m$; so that BinaryErrorLWE is LWE with $\chi = \mathcal{U}(\mathbb{F}_2)$.

Arora's and Ge's algorithm. The problem of solving LWE has attracted a lot of attention in the literature. This can be noticed, for instance, from the surveys [13], [219] and the references therein. We can consider that there are at least two categories of algorithms : lattice-based and combinatorial based [13], [219]. In [16], Arora and Ge introduced the first algebraic algorithm for solving LWE. Their approach reduces LWE to finding the common root of a multivariate system of high-degree and error-free polynomials. The proposed algorithm recovers LWE secret in:

$$2^{\tilde{O}\left(n^{2\epsilon}\right)} \tag{2.1}$$

operations, hence being sub-exponential when $\epsilon < 1/2$. This shows that Regev's original reduction in [245], [246] is indeed tight. In more detail, let $(\mathbf{a}, c) \in \mathbb{Z}_q^n \times \mathbb{Z}_q$ be an LWE sample and write $p = c - \sum_{i=1}^{n} \mathbf{a}_i\, x_i$ where the x_i are variables. If we assume that the error e is in the interval $\{-T, \ldots, T\}$, then the polynomial

$$P(x_1, \ldots, x_n) = p \prod_{i=1}^{T} (p+i)(p-i), \tag{2.2}$$

of degree $2T+1$ evaluates to zero when $x_i = s_i$. Thus, if $T < \lfloor q/2 \rfloor$ then $F = 0$ is a constraint on the possible values for the secret vector $\mathbf{s} \in \mathbb{Z}_q^n$, and collecting many such equations and solving the resulting multivariate high-degree system of equations allows to recover the secret. In [16] these systems are solved by the linearization method, i.e. first replacing monomials with a new linearized variable and then by solving the resulting linear system of equations. This method requires $O(n^{2T+1})$ equations to succeed, which could be obtained by collecting more samples. However, since $\chi_{\alpha,q}$ is a discrete Gaussian distribution, requesting more samples also increases the probability that the noise of at least one sample falls outside of the chosen interval $\{-T, \ldots, T\}$ invalidating the constraint $F = 0$. Hence, as the number of samples grows so does the required value of T so that the polynomial system remains error-free. This on the other hand may require a further increase in the number of samples to linearize. This trade-off is analyzed

in [16] to obtain the complexity (2.1). We note however that the discussion above implies that the algorithm from [16] is not applicable if the number of available samples is smaller than $O(n^{2T+1})$. Indeed, the linearization approach described above is a special case of a Gröbner basis computation. In contrast to this special case, though, general Gröbner basis algorithms are also applicable if less than $O(n^{(2T+1)})$ polynomials of degree $2T+1$ are available at the cost of increased computational complexity. The rational is that that we can decrease the degree of the equation by decreasing the number of samples. But, this means that the cost of solving the resulting system will grow compared to that of linearization. The optimization target is then to find a trade-off allowing to improve upon linearization.
In [3], [6], we show that applying Gröbner basis algorithms gives an exponential speed-up over the algorithm of Arora and Ge [16]. This leads to an algorithm for solving LWE, where $s = \sqrt{n}$, whose complexity is:

$$O(2^{6.69\,n}). \tag{2.3}$$

This corresponds to the cost of solving a semi-regular sequence (Definition 1.1.5) in $\mathbb{Z}_q[x_1, \ldots, x_n]$ with $m = \mathrm{e}^{O(n)}$ equations of degree $O(n)$ and such that $q \in \mathrm{poly}(n)$ is prime. Note that the degree of the equations is non-constant; this is quite rare in algebraic cryptanalysis. The complexity (2.3) is obtained under a *genericity assumption*, i.e. by assuming that the algebraic system derived from (2.2) behaves as a semi-regular sequence. We can save a logarithmic factor in the exponent with respect to the linearization approach of Arora and Ge [16]. This places our approach in the same complexity class, namely $2^{O(n)}$, than the best algorithms for solving [13], [185] albeit with a larger leading constant in the exponent.
The complexity (2.3) can be improved as soon as we have an additional structure such as binary errors as illustrated in this chapter.

Related works. The most relevant technique for BinaryErrorLWE is due to Fouque and Kirchner [195] who presented a combinatorial algorithm solving BinaryErrorLWE in subexponential time if at least n samples are available. The complexity of this algorithm is

$$2^{\frac{\left(\log\left(2^{1/2}\right)+o(1)\right)n}{\log\log n}}. \tag{2.4}$$

In the next Section, we show how to improve Arora's and Ge's algorithm by using Gröbner bases (Section 2.2) in the context of BinaryErrorLWE and compare our results with the complexity (2.4) of [195]. Note that [70] recently proposed a combination of various techniques for solving BinaryErrorLWE. They can improve upon [195] for some parameters. Still, they do not improve the asymptotic complexity of [195]. In this part, we mainly focus on the asymptotic hardness of BinaryErrorLWE.

2.2 Gröbner Bases Techniques for BinaryErrorLWE

The approach of Arora-Ge (Section 2.1) can be easily adapted for BinaryErrorLWE. Let $P(X) = X(X-1)$ and $(\mathbf{G}, \mathbf{s} \times \mathbf{G} + \mathbf{e}) = (\mathbf{G}, \mathbf{c}) \in \mathcal{M}_{n\times m}(\mathbb{Z}_q) \times \mathbb{Z}_q^m$ be sampled according to $L^{(n)}_{\mathbf{s},\mathcal{U}(\mathbb{F}_2)}$. Then:

$$e_i = c_i - \sum_{j=1}^{n} s_j G_{j,i}, \text{ for } i, 1 \leqslant i \leqslant m.$$

It follows that the secret $\mathbf{s} \in \mathbb{Z}_q^n$ is a solution of the following algebraic system:

$$p_1 = P\big(c_1 - \sum_{j=1}^{n} x_j G_{j,1}\big) = 0, \ \ldots, \ p_m = P\big(c_n - \sum_{j=1}^{n} x_j G_{j,n}\big) = 0. \tag{2.5}$$

This is an algebraic system of m quadratic equations in $\mathbb{Z}_q[x_1, \ldots, x_n]$. As already pointed out in [16], [218], this system can be solved using linearization if $m = O(n^2)$. Still, there is a gap between $m = n\big(1 + o(1)\big)$, where the hardness of BinaryErrorLWE reduces to worst case GapSVP$_($ Theorem 2.1.1) and $m = O(n^2)$ where the problem can be solved in polynomial-time. Understanding the hardness of the problem for samples within this interval should be of great interest: applications in lattice-based cryptography typically require the provision of $n\big(1 + o(1)\big) < m < O(n^2)$ samples, e.g. $m = O(n)$ or $m = \tilde{O}(n)$. It is hence a natural open question how the security of BinaryErrorLWE degrades as more samples are made available. We address this problem of evaluating the complexity of solving the algebraic system (2.5) with an arbitrary number $m < O(n^2)$ of equations. Our analysis depends crucially on the following assumption about the algebraic systems considered:

Assumption 2. *Let $q \geq n^{O(1)}$ be a sufficiently large polynomially bounded (prime) modulus and $(\mathbf{G}, \mathbf{s} \cdot \mathbf{G} + \mathbf{e}) = (\mathbf{G}, \mathbf{c}) \in \mathcal{M}_{n \times m}(\mathbb{Z}_q) \times \mathbb{Z}_q^m$ be sampled according to $L^{(n)}_{\mathbf{s}, \mathcal{U}(\mathbb{F}_2)}$. Let also $P(X) = X(X-1)$, we define:*

$$p_1 = P\big(c_1 - \sum_{j=1}^{n} x_j G_{j,1}\big) = 0, \ldots, p_m = P\big(c_m - \sum_{j=1}^{n} x_j G_{j,m}\big) = 0. \tag{2.6}$$

It holds that the sequence $p_1, \ldots, p_m$ is semi-regular (Definition 1.1.5).

It is believed that random systems of equations are semi-regular. Hence, our semi-regularity assumptions essentially state that our systems are neither easier nor harder to solve than random systems of equations. It is interesting to remark that Fröberg and Hollman [169] as well as Nicklasson [225] already investigated semi-regularity of powers of generic linear forms which corresponds to the homogeneous part of the equations (2.6). Thus, the sequence (2.6) has already been considered as a plausible candidate for proving Fröberg's conjecture. For instance, [169] demonstrated that the assumption indeed Assumption 2 holds if $m = n + 1$.
If the systems considered in this work were easier than random systems this would imply that the analysis that will be presented could be much improved and lead to progress towards a sub-exponential classical algorithm for solving GapSVP. On the other hand, if these systems were harder to solve than random systems, this would reveal new algebraic dependencies among BinaryErrorLWE samples, which could likely also be used to improve (non-algebraic) solving strategies. Hence, the assumption that there is no special structure in our problem instances seems to be a reasonable one. Note also that in algebraic cryptanalysis, as illustrated in the next Chapters 3 and 4, we usually expect that the systems considered behave differently than random.
This assumption is motivated by the fact that the complexity of solving semi-regular systems is well mastered. In particular, the asymptotical results on the degree of regularity such as Theorem 1.1.5 and more generally from [25]–[27], [29] allow to classify the complexity of solving polynomial systems with respect to the ration between number of equations and number of variables. In Theorem 1.1.5, we have the asymptotic expansion for $m = O(n)$. We state below a more general version of this result:

Proposition 6. *Let $m = F(n)\,n$ with $F(n) \in \{n^\epsilon, \log^{1/\epsilon}(n), \log\log n\}$ where $\epsilon > 0$ is such that $\frac{m}{n} \to \infty$ and $\frac{m}{n^2} \to 0$. Then, the degree of regularity D_{reg} of a system of quadratic semi-regular equations $p_1, \ldots, p_m \in \mathbb{Z}_q[x_1, \ldots, x_n]$ behaves asymptotically as:*

$$\frac{n^2}{8m}\left(1 + o(1)\right).$$

A proof similar to this case of this result can be found in [26]. However, there is slight difference between [26] (binary fields) and our case (generic prime fields).

Under Assumption 2, we can classify below the hardness of a Gröbner basis approach for `BinaryErrorLWE` with various number of samples. The first one corresponds to the number of equations required in the security proof [218, Theorem 1.2]. We then consider a slightly larger number of equations than what is required in the security proof, i.e. $m = 2n$ equations. In addition we give the results for a quasi-linear number of equations.

Theorem 2.2.1. *Let $\omega, 2 \leqslant \omega < 3$, be the linear algebra constant, $\alpha > 0$, and $H_2(x) = -x\log_2(x) - (1-x)\log_2(1-x)$ with $x, 0 \leqslant x \leqslant 1$. Under Assumption 2, we have:*

- *If $m = n\left(1 + \frac{1}{\log(n)}\right)$, then there is an algorithm solving* `BinaryErrorLWE` *in $O(m \cdot 2^{1.37\,\omega n})$.*
- *If $m = 2 \cdot n$, then there is an algorithm solving* `BinaryErrorLWE` *in $O(m \cdot 2^{0.43\,\omega n})$.*
- *More generally, if $m = C \cdot n$, with $C > 1$ a constant, there is an algorithm solving* `BinaryErrorLWE` *in $O\left(m \cdot 2^{n\omega(1+\beta)H_2\left(\frac{\beta}{1+\beta}\right)}\right)$ with $\beta = \left(C - \frac{1}{2} - \sqrt{C(C-1)}\right)$.*
- *If $m = n^{1+\epsilon}$, for any $0 < \epsilon < 1$, then there is a subexponential algorithm solving* `BinaryErrorLWE` *with a time complexity $O\left(m \cdot 2^{n^{(1-\epsilon)}\epsilon\,\omega\log(n)}\right)$.*
- *If $m = \mathcal{O}\left(n\log\log n\right)$, then there is a subexponential algorithm solving* `BinaryErrorLWE` *in $O\left(m \cdot 2^{\frac{n\omega\log\log\log n}{8\log\log n}}\right)$.*
- *Finally, if $m = n \cdot \log^{1/\epsilon}(n)$, for any $\epsilon > 0$, then there is a subexponential algorithm solving* `BinaryErrorLWE` *in $O\left(m \cdot 2^{\frac{n\,\omega\,\log\left(\log^{1/\epsilon}(n)\right)}{8\log^{1/\epsilon}(n)}}\right)$.*

The proof is obtained by combining Theorem 1.1.4, Proposition 6, and Stirling's approximation of the binomial $\log_2\binom{n}{k} \approx n\,H_2\left(\frac{k}{n}\right)$. It can be noticed that the complexity (2.4) of [195] is already subexponential when $m = O(n)$. Our algorithm requires $m = \tilde{O}(n)$ to be subexponential. However, (2.4) does not see a gradual asymptotically improvement when the number of samples available increases as in our case. However, the complexity (2.4) does not rely on any assumption.

2.3 About the Genericity Hypothesis

Theorem 2.2.1 depends crucially on the assumption that all systems of equations occurring in this chapter are semi-regular (Assumption 2). We experimentally confirmed that our assumption holds for reasonably large parameters. To do so, we generated systems as in (2.6). We

took q as the next prime larger than n and $m = O(n \log_2(n))$ or $m = O(n)$. We then computed a Gröbner basis of the equations using MAGMA [59] (V2.21-6). We reported in the next table the maximal degree reached $D_{\max}$ in our experiments, the theoretical degree of regularity D_{reg} under Assumption 2 (Proposition 1.1.2), and the time T_{F_4} for computing a Gröbner bases.

Table 2.1: Experimental results about the degree of regularity.

n	m	D_{reg}	$D_{\max}$	T_{F_4}
$\in \{5, \ldots, 25\}$	$\lceil n \log_2(n) \rceil$	3	3	
$\in \{26, \ldots, 53\}$	$\lceil n \log_2(n) \rceil$	4	4	$\leqslant$ 48.92 h.
60	$\lceil 2\,n \log_2(n) \rceil$	3	3	1988 s.
100	$\cdot \lceil 4\,n \log_2(n) \rceil = 2658$	3	3	30.3 h.
35	$\lceil n \log_2(n) \rceil = 180$	4	4	250.8 s.
40	$\lceil n \log_2(n) \rceil = 213$	4	4	2030.7 s.
45	$\lceil n \log_2(n) \rceil = 248$	4	4	3.6 h.
50	$\lceil n \log_2(n) \rceil = 283$	4	4	19.37 h.
53	$\lceil n \log_2(n) \rceil = 283$	4	4	48.92 h.
23	$2\,n = 46$	5	5	1329.8 s.

Table 2.2 is the equivalent of Table 1.1 but for systems randomly generated as in (2.6). We also fixed $m = \lceil n \log_2(n) \rceil$ and then randomly sampled 1000 quadratic systems in $\mathbb{F}_q[x_1, \ldots, x_n]$ as in in (2.6) where $q = 3$, q is a next prime bigger than n and q is a next prime bigger than n^2. The choice for these parameters are motivated by Chapter 2 where we consider in particular systems of equations with similar parameters. We then computed the Hilbert series of these systems and compared with the generic Hilbert series (Proposition 1.1.2). We reported the proportion of systems whose Hilbert series is exactly equal to the generic Hilbert series.

Table 2.2: Experimental results about semi-regularity.

n	Proportion (NextPrime(n))	Proportion (NextPrime(n^2))
20	1	1
26	1	1
35	1	1

In addition of these experimental results, we provide in [3] formal proofs of the assumption in several restricted cases. Namely, we prove that the equations $p_1, \ldots, p_m \in \mathbb{Z}_q[x_1, \ldots, x_n]$ as in (2.6) are linearly independent with high probability (assuming $m \leqslant n(n+1)/2$). We also prove *semigenericity*, following the terminology of [169], of the sequence $p_1, \ldots, p_m \in \mathbb{Z}_q[x_1, \ldots, x_n]$ for $m \leqslant 3n/2$. This is equivalent to prove that the Macaulay matrix $\mathcal{M}_{3,m}^{\text{acaulay}}(p_1, \ldots, p_m)$ is of full rank. The proofs follow the same principle. We use Schwartz-Zippel-DeMillo-Lipton lemma [102], [252], [276] to lower bound the success probability. The difficult, and technical, point is to prove that the polynomial considered in Schwartz-Zippel-DeMillo-Lipton lemma is non-zero.

2.4 Final Remarks

We have introduced in this part a new application of Gröbner bases in LWE-based cryptography. This follows naturally from a linearization approach proposed by Arora-Ge [16]. For LWE with a Gaussian errors, Gröbner bases allows to improve upon [16]. The complexity (2.3) relies on a semi-regularity assumption on the algebraic systems derived from (2.2). This leads to an asymptotically faster algorithm than the basic Arora-Ge [16]. Its complexity is in the same class, namely $2^{O(n)}$, than the best algorithms for solving LWE [13], [185] albeit with a larger leading constant in the exponent.

It is interesting to observe that the complexity (2.3) can be much improved for a structured noise such as binary errors. Theorem 2.2.1 shows how the complexity changes from exponential, sub-exponential and polynomial in the case of BinaryErrorLWE. We emphasize that our results are obtained by exploiting – from a polynomial system solving point of view – the weakest possible algebraic structure : semi-regularity.

Binary errors (or uniform errors) is not the only form of structured noise that has been proposed in the litterature. LWE or BinaryErrorLWE could also be considered with *binary secrets* [65]. In [179], the authors proposed to use LWE with secrets in $\{-1, 0, 1\}$ and Gaussian errors with an upper bound on the Hamming weight (in [179], the Hamming weight is 64). It is plausible that these additional conditions could be used to decrease the constants in the complexity exponents. We mention also that lattice-based public-key exchange *Frodo* [58] uses Gaussian errors with small standard deviation, but only provides a limited number of samples (typically, $n + C$ equations with C a small constant). We are then in a situation close to BinaryErrorLWE but with slightly higher-degree equations.

Ring-LWE is a compact [207] and structured variant of LWE. The idea is to consider instances of LWE where the matrices are compacts (in particular, *anti-cyclic*). It is not known if this structure can be exploited in lattice-based cryptanalysis, including algebraic cryptanalysis. As a side remark, we remark that compact matrices have been also used in the design of code-based cryptosystems in so-called *compact variants* of McEliece. We address such compact variants in Section 4.2.2) and show that this additional structure can be used and allowed us to derive efficient key-recovery attacks.

All in all, algebraic cryptanalysis is a young tool in the cryptanalysis of LWE that has the potential to be improved. In particular, a challenge is to decrease the constant in the complexity exponents of our algorithms. The approach could be typically relevant to study the asymptotic complexity of LWE problems with small or structured errors and a limited number of samples.

Another merit of this application is to establish a connection with a classical assumption of algebraic geometry about the existence of semi-regular sequences (Section 1.1.2). Proving or disproving Assumption 2, for instance, is meaningful in any case. We are inclined to believe that Assumption 2 indeed holds. But, disproving Assumption 2 could lead to improved attacks against LWE.

CHAPTER 3

CRYPTANALYSIS OF MULTIVARIATE PUBLIC-KEY CRYPTOSYSTEMS

*After a short introduction (Section 3.1) to multivariate public-key cryptography (*MPKC*), this chapter summarizes some contributions in the (algebraic) cryptanalysis of* MPKC*. This part is mainly based on : [46], [49], [138] for selecting minimal parameters for* MPKC *(Section 3.2), [156] for a cryptanalysis of* MinRank *(Section 3.3), and [50], [52] for a key-recovery against* HFE *(Section 3.4).*

3.1 Multivariate Public-Key Cryptography

3.1.1 General Principle

Multivariate cryptography is usually defined as the set of cryptographic schemes using the computational hardness of $\mathtt{PoSSo}_q$, or more generally the hardness of computing a Gröbner basis of a polynomial ideal. This is a classical candidate in quantum-safe cryptography [43], [242].

Most basic cryptographic primitives can be constructed in multivariate cryptography : hash-functions [56], stream-cipher [37], [38], Zero-Knowledge authentication scheme [182], [223], [248], [250], asymmetric encryption, e.g. [9], [31], [164], [212], [231], and signature [106], [187], [192], [231]. Few more advanced cryptographic primitives can also be constructed in multivariate cryptography. We mention for instance threshold ring signature [238], group signature [272], and a fully-homomorphic scheme that we proposed in [9], [12].

The most active area in the design of multivariate schemes is public-key cryptography. This is a sub-area of multivariate cryptography known as MPKC. Historically, the first multivariate scheme – known as $\mathtt{C}^*$ – has been proposed by Matsumoto and Imai [212]. $\mathtt{C}^*$ permits to perform public-key encryption as well as signature. However, this scheme has been completely broken by Patarin [230]. Still, the general principle inspired a whole generation of researchers that proposed improved variants of the Matsumoto-Imai (MI) principle, e.g. [84], [106], [192], [231], [232].

The generic description of a MI-like scheme is rather simple. Namely, we choose a particular system of algebraic equations $\mathbf{f} = (f_1, \ldots, f_m) \in \mathbb{F}_q[x_1, \ldots, x_n]^m$. Thanks to a well chosen structure, the corresponding system is easy to solve. That is, for all $(c_1, \ldots, c_m) \in \mathbb{F}_q^n$, we can

compute in polynomial-time:

$$V_{\mathbb{F}_q}(f_1 - c_1 = 0, \ldots, f_n - c_n). \tag{3.1}$$

We recall in Section 3.1.2 some known techniques to construct a polynomial system $\mathbf{f}$ with such property. The public-key is also a set of polynomials which is a hidden version $\mathbf{f}$. Namely, we randomly sample invertible matrices $(\mathbf{S}, \mathbf{T}) \in \mathrm{GL}_n\left(\mathbb{F}_q\right) \times \mathrm{GL}_m\left(\mathbb{F}_q\right)$ and construct the public-key as:

$$\mathbf{p} = (p_1, \ldots, p_m) = \left(f_1\left((x_1, \ldots, x_n)\,\mathbf{S}\right), \ldots, f_m\left((x_1, \ldots, x_n)\,\mathbf{S}\right)\right)\mathbf{T}. \tag{3.2}$$

In such schemes, the transformations $\mathbf{S}, \mathbf{T}$ and (usually) $\mathbf{f}$ are kept secret. We will refer to $\mathbf{f}$ as the secret-inner system. A notable exception was the $\mathtt{C}^*$ scheme where $\mathbf{f}$ is public. When $\mathbf{f}$ and $\mathbf{p}$ are known, the problem to recover the secret-key, i.e. $(\mathbf{S}, \mathbf{T}) \in \mathrm{GL}_n\left(\mathbb{F}_q\right) \times \mathrm{GL}_m\left(\mathbb{F}_q\right)$ is known as the *Isomorphism of Polynomials* (IP) problem [231]. Patarin also proposed a zero-knowledge authentication scheme based on IP [231]. IP has been also used to design a traitor tracing scheme [53], group signature scheme [272], ... From a theoretical point of view, IP is not NP-Hard unless the polynomial hierarchy collapses [136], [231]. We propose in [136] efficient algorithms for solving IP. We also consider variants such as the IP *problem with one secret* (IP1S) problem when the matrix $\mathbf{T}$ is the identity [45], [60], [80] or a generalization such as the *Functional Decomposition Problem* (FDP) [139], [145], [159].
To encrypt a message $\mathbf{m} = (m_1, \ldots, m_n) \in \mathbb{F}_q^n$ in the MI framework, we evaluate its components on the public-key, i.e.:

$$\mathbf{c} = \mathbf{p}(\mathbf{m}) = \left(c_1, \ldots, c_m\right) = \left(p_1(m_1, \ldots, m_n), \ldots, p_m(m_1, \ldots, m_n)\right) \in \mathbb{F}_q^m. \tag{3.3}$$

A direct message-recovery attack on such schemes reduces to compute:

$$V_{\mathbb{F}_q}(p_1 - c_1 = 0, \ldots, p_n - c_n). \tag{3.4}$$

Remark 1. *The general difficultly in the construction of* MI *schemes is to find a* $\mathbf{f} = (f_1, \ldots, f_m) \in \mathbb{F}_q[x_1, \ldots, x_n]^m$ *such that computing the variety* (3.1) *is easy whilst computing the variety of its hidden version* $\mathbf{p} = (p_1, \ldots, p_m) \in \mathbb{F}_q[x_1, \ldots, x_n]^m$ *should be hard.*

With the knowledge of the secret-key, we can decrypt by noticing that $\mathbf{c} \cdot \mathbf{T}^{-1} = \mathbf{f}(\mathbf{m} \cdot \mathbf{S})$. So, we need to find $\mathbf{z} \in \mathbb{F}_q^n$ such that $\mathbf{c} \cdot \mathbf{T}^{-1} = \mathbf{f}(\mathbf{z})$. This can be done efficiently thanks to the structure of $\mathbf{f}$. The message is then given by $\mathbf{m} = \mathbf{z} \cdot \mathbf{S}^{-1}$. In the case of encryption, we usually require that $m \geq n$ to have uniqueness of the decryption process.
The very same (secret-key,public-key) pair can be used to construct a digital signature. A signature $\mathbf{s} \in \mathbb{F}_q^n$ is valid for a digest (or hash) $\mathbf{d} \in \mathbb{F}_q^m$ of a message if $\mathbf{p}(\mathbf{s}) = \mathbf{d}$. A signature $\mathbf{s}$ is produced by applying the decryption process to $\mathbf{d}$. Namely, we find $\mathbf{z} \in \mathbb{F}_q^n$ such that:

$$\mathbf{d} \cdot \mathbf{T}^{-1} = \mathbf{f}(\mathbf{z}). \tag{3.5}$$

A valid signature is then given by $\mathbf{s} = \mathbf{z} \cdot \mathbf{S}^{-1}$. Note that any $\mathbf{z} \in \mathbb{F}_q^n$ solution of (3.5) leads to a valid signature. Thus, the number of equations m can be smaller than the number of variables n for multivariate signature schemes.

3.1.2 Trapdoors in the MI Family

The MI framework leads to the design of an important number of schemes, e.g. [84], [107], [192], [231], [232] that only differ in the method of constructing the secret-inner transformation $\mathbf{f}$. We recall below two prominent schemes.

Hidden Field Equations (HFE) and variants. The Hidden Field Equations (HFE) cryptosystem [231] is probably one of the most popular. In HFE, the secret-inner system $\mathbf{f} = (f_1, \ldots, f_m) \in \mathbb{F}_q[x_1, \ldots, x_n]^m$ is obtained from a particular univariate polynomial $F \in \mathbb{F}_{q^n}[X]$ which is defined below:

Definition 3.1.1. *Let $D > 0$ be an integer, and q be prime. A polynomial $F \in \mathbb{F}_{q^n}[X]$ has a HFE-shape if it has the following structure:*

$$F = \sum_{\substack{0 \leqslant i \leqslant j < n \\ q^i + q^j \leqslant D}} A_{i,j} X^{q^i + q^j} + \sum_{\substack{0 \leqslant i < n \\ q^i \leqslant D}} B_i X^{q^i} + C, \text{with } A_{i,j}, B_i, C \in \mathbb{F}_{q^n}, \forall i, j, 0 \leqslant i, j < n. \tag{3.6}$$

Remark 2. *The univariate polynomial as in* (3.6) *is also called a Dembrowsky-Ostrom [101] polynomial. In this part, we always assume that q is prime.*

The special structure of a HFE polynomial is chosen such that its *multivariate representation* over the base field $\mathbb{F}_q$ has only quadratic polynomials. Indeed, let $(\theta_1, \ldots, \theta_n) \in (\mathbb{F}_{q^n})^n$ be a basis of $\mathbb{F}_{q^n}$ over $\mathbb{F}_q$. We set $\varphi : V = \sum_{i=1}^n v_i \theta_i \in \mathbb{F}_{q^n} \longrightarrow \varphi(V) = (v_1, \ldots, v_n) \in \mathbb{F}_q^n$. The classical result below allows to make explicit this morphism between $\mathbb{F}_{q^n}$ and $\mathbb{F}_q^n$.

Proposition 7. *Let $(\theta_1, \ldots, \theta_n) \in (\mathbb{F}_{q^n})^n$ be a vector basis of $\mathbb{F}_{q^n}$ over $\mathbb{F}_q$. We set:*

$$\mathbf{M}_n = \begin{pmatrix} \theta_1 & \theta_1^q & \ldots & \theta_1^{q^{n-1}} \\ \theta_2 & \theta_2^q & & \vdots \\ \vdots & & \ddots & \vdots \\ \theta_n & \theta_n^q & \ldots & \theta_n^{q^{n-1}} \end{pmatrix} \in \mathrm{GL}_n\left(\mathbb{F}_{q^n}\right).$$

For all $V = \sum_{i=1}^n v_i \theta_i \in \mathbb{F}_{q^n}$, we have

$$(v_1, \ldots, v_n)\, \mathbf{M}_n = (V, V^q, \ldots, V^{q^{n-1}}).$$

We can know define a set of multivariate polynomials $\mathbf{f} = (f_1, \ldots, f_n) \in (\mathbb{F}_q[x_1, \ldots, x_n])^n$ derived from a HFE polynomial $F \in \mathbb{F}_{q^n}[X]$ as follows:

$$F(\varphi^{-1}(x_1, \ldots, x_n)) = \varphi^{-1}(f_1, \ldots, f_n)$$
$$F\left(\sum_{i=1}^n \theta_i x_i\right) = \sum_{i=1}^n \theta_i f_i\,.$$

The polynomials $f_1, \ldots, f_n \in \mathbb{F}_q[x_1, \ldots, x_n]^n$ are the *components* of F over $\mathbb{F}_q$.
As explained in Section 3.1.1, it is also possible to derive a signature scheme from HFE. The decryption (resp. signature generation) step is essentially equivalent to find the roots of a polynomial as in (3.6). Univariate root finding is then crucial for the efficiency of HFE. We recall below the classical complexity result for finding the roots of a univariate polynomial with the Cantor-Zassenhauss algorithm [177, Coro. 14.16]:

Theorem 3.1.1. *Let $F \in \mathbb{F}_p[X]$ be a univariate polynomial of degree $\leq D$. We can find all the roots of F using an expected number of $\tilde{O}\big(D \log(p)\big)$ operations over $\mathbb{F}_p$.*

Algorithm 2 Hidden Field Equations (HFE) Public-Key Encryption [231]

PARAMETERS. size of the field q, number of variables n, and degree of the univariate polynomial D.
Plaintext space: $\mathbb{F}_2^n$. Ciphertext space: $\mathbb{F}_2^n$.

KEYGEN. We randomly select a polynomial $F \in \mathbb{F}_{q^n}[X]$ of degree D with a HFE-shape as in (3.6) and $\mathbf{f} = (f_1, \ldots, f_n) \in (\mathbb{F}_q[x_1, \ldots, x_n])^n$ such that $F\left(\sum_{i=1}^n \theta_i x_i\right) = \sum_{i=1}^n \theta_i f_i$. We also randomly select $(\mathbf{S}, \mathbf{T}) \in \mathrm{GL}_n\left(\mathbb{F}_q\right) \times \mathrm{GL}_n\left(\mathbb{F}_q\right)$.

PRIVATE-KEY. $F \in \mathbb{F}_{q^n}[X]$ and $(\mathbf{S}, \mathbf{T}) \in \mathrm{GL}_n\left(\mathbb{F}_q\right) \times \mathrm{GL}_n\left(\mathbb{F}_q\right)$.
PUBLIC-KEY. It is given by:

$$\mathbf{p} = (p_1, \ldots, p_n) = \left(f_1\left((x_1, \ldots, x_n)\,\mathbf{S}\right), \ldots, f_n\left((x_1, \ldots, x_n)\,\mathbf{S}\right)\right)\mathbf{T}.$$

ENCRYPT.
1: Input $\mathbf{m} \in \mathbb{F}_q^n$
2: Output $\mathbf{c} = \mathbf{p}(\mathbf{m}) \in \mathbb{F}_q^n$.

DECRYPT.
1: Input $\mathbf{c} \in \mathbb{F}_q^n$
2: Compute $C' = \varphi^{-1}(\mathbf{c}'\,\mathbf{S}^{-1})$
3: Compute the roots $\underline{Z} \in \mathbb{F}_{q^n}$ of:

$$F(\underline{Z}) - C' = 0. \tag{3.7}$$

4: Output $\varphi(Z) \cdot \mathbf{S}^{-1}$

In HFE, we have that $p = q^n$. Assuming that q is a constant, the roots of $F \in \mathbb{F}_{q^n}[X]$ can be found in $\tilde{O}(n\,D)$. As a consequence, a HFE polynomial has to be chosen of moderate degree for being efficiently solvable. To fix the ideas on the degree that can be considered in practice, we provide below some timings of the roots finding function Roots of MAGMA applied on random HFE-polynomial $F \in \mathbb{F}_{2^n}[X]$ of degree D. We have performed the experiments on a MacBook Air, Intel dual-core i5 1.6 GHz with 4 GB of RAM. The timings are obtained by taking the average time of the MAGMA function Roots on 100 calls. We also report the average number of roots (#Roots).

Table 3.1: Experiments with the MAGMA function Roots.

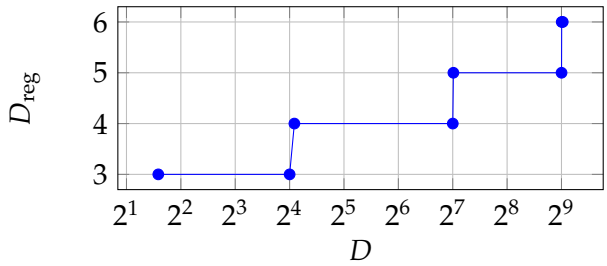
D	129	257	513	1025	2049	4097	8193	16385	32769
$n = 128$	0.1 s.	0.21 s.	0.55 s.	1.55 s.	7.36 s.	13.82 s.	30.28 s.	61.47 s	132.09 s.
#Roots	0.92	1.18	0.9	1.2	1.12	0.96	0.94	0.97	0.91

For the range of considered parameters, a HFE-shape has usually very few roots.
Over the years, the security of HFE has been thoroughly investigated, e.g. [62], [86], [100], [105], [114], [115], [146], [155], [183]. The first major result in the analysis of HFE appeared in [146], [155]. The authors demonstrated an efficient direct Gröbner basis attack and reported a practical break of the HFE Challenge[12] 1 : $m = n = 80, q = 2$ and $D = 96$. It appeared [146], [155] that inverting the public-key of the original HFE is much easier than expected, i.e. in comparison to a random system of the same size. For the basic HFE, with $q = 2$, the degree of regularity has been experimentally shown to be smaller than $\log_2(D)$. More precisely, the degree of regularity

[12] http://hfe.minrank.org/

D_{reg} for computing a Gröbner basis of a HFE public-key will vary in function of the degree of the secret univariate polynomial D [146], [155]. This is not the behavior expected for a semi-regular system of equations (Section 1.1.2). The degree of regularity of a semi-regular sequence should increase with the number of variables (Theorem 1.1.5). We report below the degree of regularity D_{reg} observed in practice for HFE as a function of the degree D of the secret univariate HFE polynomial. Note also that the bounds are valid for a sufficiently large n which is given in the first column. The reason is that the degree of regularity observed is not bigger than the degree of regularity of a semi-regular sequence (Property 1.1.2).

Minimal n	HFE(D)	D_{reg}
$\geqslant 4$	$3 \leq D \leq 16$	3
$\geqslant 9$	$17 \leq D \leq 128$	4
$\geqslant 16$	$129 \leq D \leq 512$	5
$\geqslant 25$	$513 \leq D \leq 1280$	6



The major observation is that the complexity of a direct Gröbner basis attack against HFE is not exponential in the number of variables, but exponential in $O\big(\log_2(D)\big)$. This phenomenon has been further studied and confirmed in a series of papers, e.g. [105], [114], [183]. In particular, the authors of [105] provided an upper bound on the degree of regularity for HFE systems defined over arbitrary q. The bound is given by:

$$\frac{(q-1)\big(\lceil \log_p(D-1)\rceil + 1\big)}{2} + 2. \tag{3.8}$$

We remark that this upper bound is linear in q. One could think that this would suggest that attacking HFE should be harder when q increases. We will see in Section 3.4 that this is not the case. We present a key-recover attack against HFE of complexity $\mathcal{O}\left(n^{(\log_q(D)+1)\,\omega}\right)$ (Proposition 10), with $\omega, 2 \leqslant \omega < 3$ being the linear algebra constant.

Variants of HFE. From the discussion above, it can already be noticed that the efficiency and security of HFE is related to the degree D of the univariate secret-key. HFE can be modified to increase the degree of regularity without compromising too much the efficiency. Classical variants include the *minus variant* (HFE-, [231]) and the *vinegar variant* (HFEv, [232]). In the minus variant, we simply remove some equations from the public-key. In HFEv, the inner-secret system is now derived from a single multivariate $F(X, v_1, \ldots, v_t) \in \mathbb{F}_{q^n}[X, v_1 \ldots, v_t]$ which has the following form :

$$\sum_{\substack{0 \leqslant i \leqslant j < n \\ q^i + q^j \leqslant D}} A_{i,j}\, X^{q^i+q^j} + \sum_{\substack{0 \leqslant i < n \\ q^i \leqslant D}} \gamma_i(v_1, \ldots, v_t)\, X^{q^i} + \beta(v_1, \ldots, v_t) + C \in \mathbb{F}_{q^n}[X, v_1, \ldots, v_t], \tag{3.9}$$

where $A_{i,j}, C \in \mathbb{F}_{q^n}, \forall i, j, 0 \leqslant i \leqslant j < n$, each $\gamma_i(v_1, \ldots, v_t) \in \mathbb{F}_{q^n}[v_1, \ldots, v_t]$ is a linear polynomial and $\beta(v_1, \ldots, v_t) \in \mathbb{F}_{q^n}[v_1, \ldots, v_t]$ is a quadratic polynomial. The variables $v_1, \ldots, v_t$ are called the *vinegar variables*. The particularity of the polynomial $F(X, v_1, \ldots, v_t)$ as in (3.9) is that for any specialization of the vinegar variables the polynomial becomes an HFE polynomial.
The minus and vinegar variants seem to indeed strengthen the security of HFE. For instance, the HFE Challenge 1 has been solved in two days [146], [155]. However, the HFE challenge 2, which is a HFE- system, with $q = 36, D = 4352, n = 36$ and 4 equations removed still unbroken

today [13]. In the signature context, these variants are as efficient than a plain HFE. In the context of encryption, there is a performance penalties induced by the fact that we will have to re-run the decryption process several times (q^t times for HFEv, and q^r times for HFE- where r equations are removed).
We can combine the minus and vinegar variants. This leads to the so-called HFEv- variant. Typically, QUARTZ [233] is a multivariate signature scheme submitted to the evaluation process organized by the NESSIE EU project [224]. The goal of NESSIE was to recommend a set cryptographic algorithms. QUARTZ is a HFEv- signature scheme where $D = 129, q = 2, n = 103, t = 4$ and 3 equations removed. This yields a public-key of 71 KBytes and for a security level initially estimated at 2^{82}. The main feature of QUARTZ is to provide very short signatures, i.e. 128 bits for the parameters proposed above.

Unbalanced Oil and Vinegar (UOV) scheme. This is multivariate signature scheme proposed in 1999 by Goubin, Kipnis and Patarin [192]. It is one of the rare multivariate scheme that resisted to all attacks reported so far. The idea of UOV is to partition the variables $\{x_1, \ldots, x_n\}$ in two sets: the *vinegar variables* $V = \{x_1, \ldots, x_{n-m}\}$, and the *oil variables* $O = \{x_{n-m+1}, \ldots, x_n\}$. The secret-inner polynomials $f_1, \ldots, f_m \in \mathbb{F}_q[x_1, \ldots, x_n]$ has the following special shape:

$$f_k = \sum_{(x_i,x_j)\in V\times V} \alpha_{i,j}^{(k)} x_i x_j + \sum_{(x_i,x_j)\in V\times O} \beta_{i,j}^{(k)} x_i x_j + \sum_{i=1}^{n} \gamma_i^{(k)} x_i + \delta^{(k)}, \tag{3.10}$$

with $\alpha_{i,j}^{(k)}, \beta_{i,j}^{(k)}, \gamma_i^{(k)}, \delta^{(k)} \in \mathbb{F}_q[x_1, \ldots, x_n]$.
The principle is that once the vinegar variables have been fixed, the secret-inner system (3.10) becomes a linear system with m equations in m variables, and will be easy to invert with a high probability. The public-key $\mathbf{p} = (p_1, \ldots, p_m) \in \mathbb{F}_q[x_1, \ldots, x_n]^m$ can be constructed from $\mathbf{f} = (f_1, \ldots, f_m)$ as in (3.2). The only difference is that the matrix $\mathbf{T} \in \mathrm{GL}_m\left(\mathbb{F}_q\right)$ is always chosen to be the identity. Thus:

$$\mathbf{p} = (p_1, \ldots, p_m) = \left(f_1\left((x_1, \ldots, x_n)\,\mathbf{S}\right), \ldots, f_m\left((x_1, \ldots, x_n)\,\mathbf{S}\right)\right), \text{with } \mathbf{S} \in \mathrm{GL}_n\left(\mathbb{F}_q\right).$$

According to [193], it is mandatory to take $m > n$. The case $m = n$ corresponds to the balanced case, also known as the *Oil and Vinegar* scheme, which can be broken by [193]. Currently, it is recommended to take $n = m/\alpha$, such that $1.5 \leqslant \alpha < 3$ [64], [97], [109], [269]. We discuss in Section 3.2 how to set minimal requirement on the value of m. In [72], the authors demonstrated that the secret-key/public-key generation of UOV can be modified to achieve a rather surprising property. It is possible to first generate the public-key $\mathbf{p} = (p_1, \ldots, p_m) \in \mathbb{F}_q[x_1, \ldots, x_n]^m$ and find a linear change of variables $\mathbf{S} \in \mathrm{GL}_n\left(\mathbb{F}_q\right)$ so that the polynomials $\mathbf{p}(\mathbf{x}\mathbf{S}^{-1})$ has the UOV structure as in (3.10). The polynomials of $\mathbf{p}$ can be almost chosen randomly. More precisely, the homogeneous components of highest degree can be sampled randomly, but the coefficients of the linear components require to be adapted. Finally, we mention that the authors of [249] proposed a modification of the UOV which allows to prove in the random oracle the so-called model existential unforgeability against adaptive chosen-message attack (EUF-CMA). A similar result holds for HFE.

[13] It was initially claimed in [83] that HFE challenge 2 could be solved in 2^{63}. This claim was withdrawn in [275] where the complexity of the attack from [83] was corrected to 2^{93}.

3.2 Selecting Parameters for MPKC with the Hybrid Approach

We have presented in Section 1.2 a hybrid approach that permits to improve asymptotically the complexity of solving PoSSo_q. Beyond this complexity result, the hybrid approach naturally permits to derive minimal parameters for multivariate schemes, and in particular MPKC. The security of the MPKC described in Section 3.1.1 naturally relies on the hardness of computing a Gröbner basis of the public-key $(p_1, \ldots, p_m) \in \mathbb{F}_q[x_1, \ldots, x_n]^m$. Given a ciphertext (resp. the digest of a message) $(c_1, \ldots, c_m) \in \mathbb{F}_q^m$, a message-recovery (or signature forgery) reduces to solve:

$$p_1 = 0, \ldots, p_m - c_m = 0, x_1^q - x_1 = \ldots = x_n^q - x_n = 0. \tag{3.11}$$

In the context of a multivariate signature, the problem of forging a signature requires to solve an *under-defined* system of equations, i.e. we have to solve (3.11) for $m < n$. In this case, we expect that the variety associated to (3.11) has $O(q^{n-m})$ solutions. Consequently, the cost of the second step in the zero-dimensional solving strategy (Section 1.1.1), the change of ordering, could be prohibitive due to the exponential number of solutions. Recall that the FGLM algorithm has a complexity which is polynomial in the number of solutions (Theorem 1.1.3). In order to forge a signature, we only need one solution of (3.11). A rather natural strategy in this context is to fix randomly $n - m$ variables in the system (3.11). However, by fixing random variables, it is likely that the system will lose its internal structure.
To illustrate the approach, we consider here the UOV signature scheme (Section 3.1.2) and the parameters initially recommended in [192]: $q = \mathbb{F}_{2^4}, m = 16, n = 32$ (or 48). Note that more multivariate schemes have been considered in [46]. Experimentally, we noticed that the new systems behave as semi-regular systems, i.e. they verify Assumption 1. We fall then right in the scope of the hybrid approach. The roadmap of the attack for signature schemes is as follows:

1. Fix $n - m$ variables in message-recovery system (3.11) to obtain a new algebraic system with m variables and m equations. This new system will always have at least one valid solution.

2. Solve the new system with the hybrid approach (Section 1.2).

Under Assumption 1, we computed in Figure 3.1 the complexity of the hybrid approach $C_{\text{Hyb}}(k)$ (Proposition 2) for $k, 0 \leq k \leq 16$. It seems that the best theoretic tradeoff would be to fix 4 variables.
In practice, we obtain the best experimental trade-off by fixing only 2 variables. We show these results in Table 3.2 for different trade-offs. T_{F_4} is the time to compute one Gröbner basis with F_4 (the version available in MAGMA) and $\text{T} = q^k \cdot \text{T}_{\text{F}_4}$. We also include the maximum memory used during the Gröbner basis computation under the column "Mem". In all cases, the complexities are below the usual cryptographic security bound, i.e. 2^{80}. We were able to forge signatures on a Xeon quadri-processors 2.92 Ghz with 131 GB of RAM. It is worth to

m	$m-k$	q^k	T_{F_4}	Mem	T
16	15	2^4	3007 s.	564.5 MB	13.36 h.
16	14	2^8	43.9 s.	47.5 MB	3.12 h.
16	13	2^{12}	3.65 s.	12.6 MB	4.15 h.
16	12	2^{12}	0.51 s.	8.4 MB	9.28 h.

remark that the efficiency of our attack does not rely on the number of the vinegar variables,

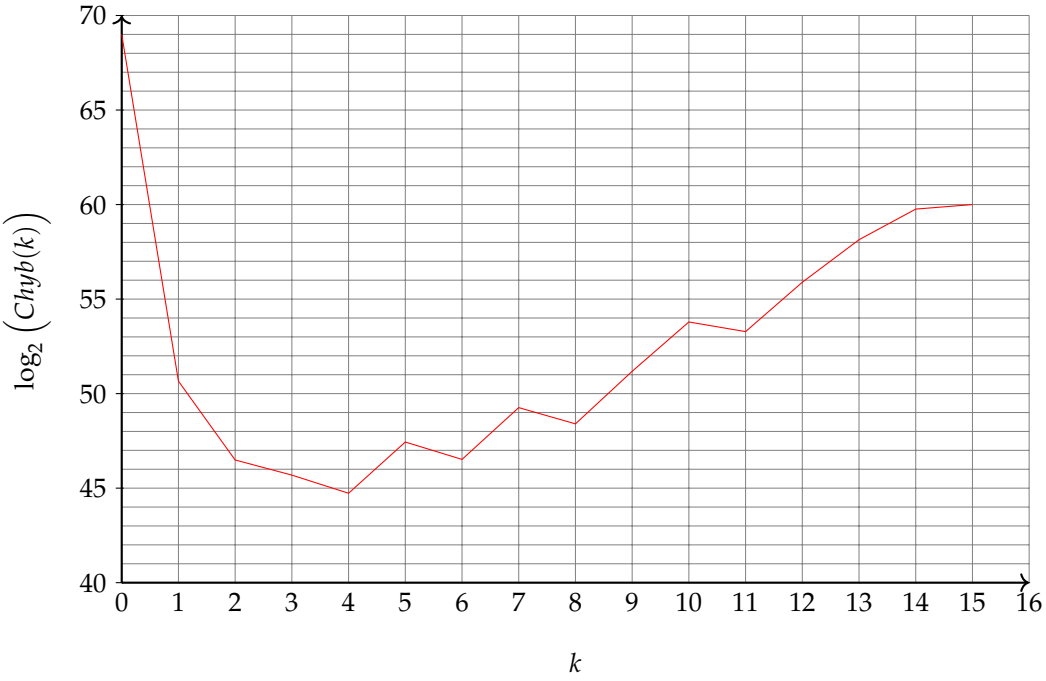


Figure 3.1: Theoretical complexity of attacking UOV depending on k.

but only on the number of equations. Even if there are many more variables than equations, one can always fix variables and still be able to forge one signature. Thus, the results are valid for $n = 32$ as well as for $n = 48$.

Minimal parameters. The use of our approach is straightforward for any multivariate scheme. We can then explicitly give the parameters for which a random quadratic system can not be solved with our approach (i.e. complexity $> 2^{80}$) in Table 3.2. The column m is the minimum number of equations and variables that should be chosen. The column k is the best trade-off for the hybrid approach and the column C_{Hyb} is the corresponding complexity. To have a sketch of what could be the size of the public key and the signatures, we compute them for a system with $3/2$ times more variables than equations ($n = 3m/2$). We emphasize that the complexity given in Table 3.2 are upper bounds which are reached for random dense systems. Note that the complexities have been computed with $\omega = 2$ to be conservative. It has to be added that the values given in Table 3.2 are the minimal parameters that a multivariate cryptosystems such as UOV should have to withstand our attack. The given parameters do not prevent from other kind of attacks.

Finally, the best trade-off between security/size of the public key/size of the signature is obtained by choosing a large amount of variables and a small field. We note that the key is smaller when $q = 2^4$.

q	m	k	C_{Hyb}	signature length	public key size
2^{32}	20	0	2^{82}	960 bits	39 KB
2^{16}	23	1	2^{81}	560 bits	29 KB
2^{8}	26	1	2^{83}	312 bits	21 KB
2^{4}	30	7	2^{83}	180 bits	16 KB
2^{2}	41	23	2^{82}	124 bits	20 KB

Table 3.2: Minimal recommended parameters.

3.3 The MinRank Problem : Algorithmic and Hardness Results

The MinRank problem is a classical linear algebra problem which can be seen as an extension of an eigenvalue problem. It was originally introduced in [73] where the authors proved its NP-hardness. Later, it was reformulated by Courtois [82] in the cryptographic context who describes a zero-knowledge scheme – that we will call ZKMR – based on MinRank.
Since then, MinRank appeared to be crucial for the security of most multivariate schemes constructed in the MI-framework (Section 3.1.1), including HFE (Section 3.1.2) but also few others e.g. [54], [108], [154], [181], [194]. The public-key in MPKC is usually given by quadratic polynomials which can be represented by a set of matrices. The idea to attack multivariate schemes with MinRank is to exploit a rank defect in the public matrices induced by the very structure of the inner system (Section 3.1.1). We elaborate further on such attack in the case of HFE in the next Section 3.4. All in all, MinRank is a fundamental problem in MPKC.
MinRank
Input. A set of $k+1$ matrices $\mathbf{M_0}, \mathbf{M_1}, \ldots, \mathbf{M}_k \in \mathcal{M}_{N\times n}(\mathbb{F}_q)$ and an integer $r > 0$.
Question. Find – if any – a k-tuple $(\lambda_1, \ldots, \lambda_k) \in \mathbb{F}_q^k$ such that:

$$\text{Rank}\left(\sum_{i=1}^{k} \lambda_i \mathbf{M}_i - \mathbf{M}_0\right) \leqslant r.$$

From now on, we only consider the case where $N = n$, i.e. "square" instance of MinRank. This is not a restriction since square MinRank ($N = n$) is equivalent (by poly-time reduction) to MinRank [156]. We review below the main techniques for solving MinRank.

3.3.1 The Kernel Attack

First, note that exhaustive search to find a tuple $(\lambda_1, \ldots, \lambda_k) \in \mathbb{F}_q^k$ solution of MinRank needs $O(q^k n^\omega)$ operations on $\mathbb{F}_q$. Courtois and Goubin proposed in [181] a clever way for performing exhaustive search. Given $k+1$ matrices $\mathbf{M}_0, \mathbf{M_1}, \ldots, \mathbf{M}_k \in \mathcal{M}_{N\times n}(\mathbb{F}_q)$, the idea is to choose randomly m vectors $\mathbf{x}^{(1)}, \ldots, \mathbf{x}^{(m)} \in \mathbb{F}_q^n$. Then, we solve a linear system of mn equations for the unknowns $(y_1, \ldots, y_k)$:

$$\mathbf{x}^{(\mathbf{i})}\big(\mathbf{M}_0 - \sum_{j=1}^{k} y_j \mathbf{M}_j\big) = \mathbf{0}_n, \forall i, 1 \leq i \leq m.$$

Note that if $m = \lceil \frac{k}{n} \rceil$, this system essentially has only one solution $\lambda = (\lambda_1, \ldots, \lambda_k) \in \mathbb{F}_q^k$.
Now set $E_\lambda = \mathbf{M}_0 - \sum_{j=1}^{k} \lambda_j \mathbf{M}_j$; we want E_λ to be of rank $\leq r$. If this were the case, then

$\dim\big(\mathrm{Ker}(E_\lambda)\big) \geq n - r$ and so, for $\mathbf{x} \in \mathbb{F}_q^n$ chosen at random:

$$\Pr[\mathbf{x} \in \mathrm{Ker}(E_\lambda] \geq q^{-r} \text{ and } \Pr[\{\mathbf{x}^{(1)}, \ldots, \mathbf{x}^{(m)}\} \subseteq \mathrm{Ker}(E_\lambda)] \geq q^{-mr}.$$

Thus, in order to find a $\lambda = (\lambda_1, \ldots, \lambda_k) \in \mathbb{F}_q^k$ such that E_λ has the desired rank, we have to run the above experiment q^{mr} times on average. Taking the value of m as above, the complexity of this attack is thus

$$\mathcal{O}(q^{\lceil \frac{k}{n} \rceil r} k^\omega).$$

3.3.2 Kipnis-Shamir Modeling and Gröbner Bases

The MinRank problem can be also modeled as a system of quadratic equations. It is a transposition of an attack on HFE due to Kipnis and Shamir [194]. In its principle, this is an algebraic version of the kernel attack described below. With the previous notations, the basic idea is to take as unknown the basis vectors of a left kernel of $\left(\sum_{i=1}^k \lambda_i \mathbf{M}_i - \mathbf{M}_0 \right)$. The dimension of the kernel space will depend on the target rank r. This yields a system of quadratic equations whose unknowns are the coefficient of the kernel vectors and the $\lambda_1, \ldots, \lambda_k \in \mathbb{F}_q$. We can decrease the number of variables by assuming that the kernel is in systematic form (this holds with high probability over a finite field). Finally, MinRank is equivalent to solve the algebraic system of $n\,(n-r)$ equations in $r\,(n-r)+k$ variables given by the entries of the matrix:

$$\begin{pmatrix} 1 & & & x_{1,1} & \ldots & x_{1,r} \\ & \ddots & & \vdots & & \vdots \\ & & 1 & x_{n-r,1} & \ldots & x_{n-r,r} \end{pmatrix} \cdot \left(\sum_{i=1}^k y_i \mathbf{M}_i - \mathbf{M}_0 \right). \tag{3.12}$$

Initially, relinearization [194] has been used to solve (3.12). In [156], we proposed instead to use Gröbner bases. This allows a considerable speed-up against the kernel attack described below (experimental resutls are presented below). In addition, we noticed in [156] that the system has a multi-homogeneous structure (Definition 1.1.7). We used this fact and previous results on multi-homogeneous Bézout bound [104], [200], [211] to derive in [156] a new upper bound on the complexity of solving MinRank with the Kipnis-Shamir (KS) modeling.

Experimental Results

We report in Table 3.2 experimental results obtained by using Gröbner bases together with the KS modeling. We consider in particular the parameters used in ZKMR [82]. These parameters are quoted below together with the number of equations and variables obtained using the KS modeling $\big($algebraic system (3.12)$\big)$:

- $A : \mathbb{F}_{65521}, n = 6, k = 10, r = 3$ (18 equations and 19 variables)
- $B : \mathbb{F}_{65521}, n = 7, k = 10, r = 4$ (21 equations, and 22 variables)
- $C : \mathbb{F}_{65521}, n = 11, k = 10, r = 8$ (33 equations, and 34 variables)

These parameters were selected so that the best attack previously known – the Kernel Attack (Section 3.3.1) here – requires at least 2^{106} operations [181].
We remark that the algebraic systems defined by the challenges are under-defined. For breaking the zero-knowledge scheme, we only need one solution. Thus, we can fix 1 variable for the challenges A, B and C. This is then equivalent to solve MinRank on the following parameters :

- **A** : $\mathbb{F}_{65521}, n = 6, k = 9, r = 3$ (18 equations, and 18 variables)
- **B** : $\mathbb{F}_{65521}, n = 7, k = 9, r = 4$ (21 equations, and 21 variables)
- **C** : $\mathbb{F}_{65521}, n = 11, k = 9, r = 8$ (33 equations, and 33 variables)

We consider instances with a pre-assigned solution. To do so, we have randomly sampled k matrices $\mathbf{M}_1, \ldots, \mathbf{M}_k \in \mathcal{M}_{n,n}(\mathbb{F}_q)$ and k coefficients $(\lambda_1, \ldots, \lambda_k) \in \mathbb{F}_q^k$ such that $S_\lambda = \sum_{i=1}^k \lambda_i \neq 0$. Finally, we randomly selected a matrix $\mathbf{M} \in \mathcal{M}_{n,n}(\mathbb{F}_q)$ of rank r and set $\mathbf{M}_0 = \sum_{i=1}^k \lambda_i(\mathbf{M}_i - \mathbf{M})$. Thus, we have $\text{Rank}\left(\sum_{i=1}^k \lambda_i \mathbf{M}_i - \mathbf{M}_0\right) = \text{Rank}(S_\lambda \mathbf{M}) = r$.
The experimental results have been obtained with MAGMA (2.21-6) on a Xeon quadri-processors 2.92 Ghz, with 131 GB of Ram. We can break the two challenges **A** and **B**. There is a huge gap between these challenges and challenge **C**. However, we can estimate the complexity of our attack for the last challenge by considering intermediate instances of the problem, i.e. $\texttt{MinRank}(n, k, r)$, with $\mathbb{F}_q = \mathbb{F}_{65521}, n = r + 3, k = (n - r)^2 = 9$ and $r \in \{3, 4, 5, 6, 7, 8\}$. Also, since all the λ_i are in $\mathbb{F}_q$, we can perform a hybrid approach, i.e. an exhaustive search on some λ_i. Namely, we suppose that we have $nb_f > 0$ coefficients of a solution $(\lambda_1, \ldots, \lambda_k)$ of MinRank. This is equivalent to solve a $\texttt{MinRank}(n, k - nb_f, r)$ problem. We have then to solve q^s over-determined systems. When $s > 0$ the number of solutions of the corresponding algebraic system is always 1 and any Gröbner basis for any monomial ordering gives the solution (Property 1.1.1); consequently there is no need to apply the FGLM algorithm (Section 1.1.1).
In Table 3.2, T_{DRL} is the CPU time for computing a LEX-Gröbner basis, T_{FGLM} is the CPU time for changing the basis to a LEX-Gröbner basis using the FGLM algorithm, $D_{\max}$ is the maximum degree reached during the computation of a Gröbner basis and D_{reg} is the theoretical degree of regularity of a semi-regular system (Property 1.1.2) with $n(n-r)$ equations and $r(n-r) + k - nb_f$ variables.

		$\mathbb{K} = \mathbb{F}_{65521}, \texttt{MinRank}(n,k,r)$					
		Challenge **A**	Challenge **B**				Challenge **C**
		$(6,9,3)$	$(7,9,4)$	$(8,9,5)$	$(9,9,6)$	$(10,9,7)$	$(11,9,8)$
$nb_f = 0$	T_{DRL}	31.8 s.	1.15 h.	52.53 h.			
	T_{FGLM}	3.5 s.	605.7 s.	12.71 h.			
	$D_{\max}$	6	6	7			
	D_{reg}	19	22	25			
$nb_f = 1$	T_{DRL}	2.1 s.	139.4 s.	2.32 h.	54.09 h.		
	$D_{\max}$	4	5	6	6		
	D_{reg}	10	11	13	14		
$nb_f = 2$	T_{DRL}	0.6 s.	6.3 s.	327 s.	1.54 h.	22.09 h.	
	$D_{\max}$	4	4	5	5	6	
	D_{reg}	8	10	11	12	14	
$nb_f = 3$	T_{DRL}	0.09 s.	2.3 s.	14 s.	476.6 s.	1.14 h.	14.61 h.
	$D_{\max}$	4	4	4	5	**5**	**5**
	D_{reg}	7	8	10	11	12	13

Figure 3.2: Experimental results with Gröbner bases on the KS modeling.

Interpretation of the results. Challenges **A** $(6,9,3)$ and **B** $(7,9,4)$ are completely broken. We

emphasize that such sets of parameters were the most practical parameters proposed in [82]. In [156], we also evaluated that the complexity for attacking Challenge **C** should be $\approx 2^{65}$.
The major observation here is that systems coming from the KS modeling are much easier to solve than semi-generic instances of the same size. This can be observed by comparing the maximal degree $D_{\max}$ reached in practice with the theoretical degree of regularity D_{reg} of a semi-regular system. We see clearly the effect of the multi-homogeous structure (Definition 1.1.7) of the algebraic system (3.12). The degree of regularity seems to be bounded from above by $\approx r + 2$. We provide in [156] a first explanation due to the multi-homogeneous structure of the KS modeling.

3.3.3 Minors Modeling

Following [156], Faugère, Safey El Din and Spaenlehauer [149], [151] presented and analysed an alternative modeling for MinRank. Indeed, this problem can be formulated as finding a vector $(\lambda_1, \ldots, \lambda_k) \in \mathbb{F}_q^k$ vanishing on all the minors of size $r+1$ of the matrix $\left(\sum_{i=1}^k \lambda_i \mathbf{M}_i - \mathbf{M}_0\right)$. We have then to solve

$$\texttt{Minors}_{r+1}\left(\mathbf{M}_0 - \sum_{j=1}^{k} y_j \mathbf{M}_j\right) = 0, \tag{3.13}$$

where $\texttt{Minors}_{r+1}$ denotes the set of all minors of degree $r+1$. System (3.13) is a multivariate polynomial system of $\binom{n}{r+1}^2$ equations in k variables. The system has more equations and less variables than the KS modeling but the degree of the equations is r. An advantage of this approach is that precise complexity bounds can be derived for this modeling [149], [151]. In particular, Corollary 3 of [151] gives a bound on the degree of regularity of such determinantal systems.

Proposition 8 (Faugère, Safey El Din, Spaenlehauer [149], [151])**.** *Let (n, k, r) be the parameters of a (square)* MinRank *instance such that $k \leq (n-r)^2$. Let $\mathbf{A}(z) = \{a_{i,j}(z)\}_{1 \leq i \leq r}^{1 \leq j \leq r}$ be the $(r \times r)$-matrix defined by*

$$a_{i,j}(z) = \sum_{\ell=0}^{n-\max(i,j)} \binom{n-i}{\ell}\binom{n-j}{\ell} z^\ell.$$

The degree of regularity of the system (3.13) *for a generic* MinRank *instance is bounded from above by* $1 + \deg\left(\mathrm{HS}(z)\right)$ *where*

$$\mathrm{HS}(z) = \left[(1-z)^{(n-r)^2-k}\, \frac{\det \mathbf{A}(z)}{t^{\binom{r}{2}}}\right]_+.$$

As explained in [151], the proof is valid under the assumption that a variant of the Fröberg conjecture [168] is true. More recently [149], the same result was proved when $k \geq (n-r)^2$ without using any variant of Fröberg's conjecture. However, in the over-determined case – that is to say when $k < (n-r)^2$ – the conjecture is still needed.

Experimental Results

We report below experimental results by computing Gröbner bases on the minors modeling. We considered instances of MinRank as in Section 3.3.2. The programming language, and workstation is as in Section 3.3.2 (MAGMA (2.21-6) and Xeon quadri-processors 2.92 Ghz, with 131 GB of Ram).

The following notations are used in Table 3.3 : T_{DRL} is the CPU time for computing a DRL-Gröbner basis, T_{FGLM} is the CPU time (in seconds) for changing the basis to a DRL-Gröbner basis using the FGLM, $D_{\max}$ is the maximum degree reached during the computation of a Gröbner basis, and D_{reg} is the theoretical degree of regularity according to Proposition 8.

	$\mathbb{K} = \mathbb{F}_{65521}$, $\texttt{MinRank}(n,k,r)$					
	Challenge **A**	Challenge **B**				Challenge **C**
	$(6,9,3)$	$(7,9,4)$	$(8,9,5)$	$(9,9,6)$	$(10,9,7)$	$(11,9,8)$
T_{DRL}	1.5 s.	32.6 s.	567.6 s.	2.33 h.	28.88 h.	65521×10.13 h.
T_{FGLM}	2.87 s.	145 s.	2.15 h.	68.94 h.		
$D_{\max}$	10	13	16	19	22	
D_{reg}	10	13	16	19	22	25

Figure 3.3: Experimental results with Gröbner bases on the minors modeling.

Interpretation of the results. The minors approach allows to speed-up the attack against MinRank. Typically Challenge **B** can now be broken in half a minute, whilst it took ≈ 1 hour with the KS modeling (Table 3.2). For the last challenge, we estimate the solving-time by using a hybrid approach. We simply fix $nb_f = 1$ variable, and have then to solve q times $\texttt{MinRank}(11,8,8)$. The latter MinRank can be solved in 10.13 h. Using the KS modeling, we were only able to solve $\texttt{MinRank}(11,6,8)$ (that is, fixing 3 variables in Challenge **C**) in 14.61 h.
More generally, [149], [151] proved that for instances of MinRank arising in ZKMR [181], i.e. such that $k = (n-r)^2$, then computing a Gröbner basis with the minors modeling will be always faster than with the KS modeling. We can also notice that the theoretical degree of regularity from Proposition 8 is always equal to the maximum degree reached during the Gröbner computation.

3.4 A Key-Recovery Attack against HFE

The MinRank problem as described in Section 3.3 first appeared in the cryptographic context due to a key-recovery attack of Kipnis and Shamir against HFE [194]. The principle of this attack, known as *Kipnis-Shamir (KS) attack*, is to exploit the very structure of a HFE polynomial (Definition 3.1.1). To simplify the description, we suppose here that q is an odd prime. The case of even characteristic is slightly more technical, but fully addressed in [51].
KS attack proceeds in two steps. First, we solve an instance of MinRank that will give a part of the secret key, i.e. the outer matrix $\mathbf{T} \in \mathrm{GL}_n\left(\mathbb{F}_q\right)$ (Section 3.1.2). Once $\mathbf{T}$ has been found, we can recover the inner transformation $\mathbf{S} \in \mathrm{GL}_n\left(\mathbb{F}_q\right)$ by essentially solving a linear system of equations [51], [194]. Consequently, solving the MinRank is the most costly part of the attack. We then focus on this step here. We first explain how the MinRank problem occurs in the cryptanalysis of HFE.
To do so, we revisit the classical KS attack of [194] on HFE. A first difference is that [194] requires to compute – by interpolation – the univariate representation of the public-key; whilst the technique presented here operates directly on the public-key polynomials. Our attack allows a considerable speedup over the original KS attack. It makes it practical for a wide range of parameters whereas the original attack from [194] was considered theoretical. It is not detailed here, but this version of the attack can be generalized and used against variants of HFE

such as Multi-HFE [55], [77]. Finally we build on the degree of regularity of a determinantal system (Proposition 8) to derive (Section 3.4.2) a complexity bound for our key-recovery in the case of HFE.

3.4.1 MinRank in HFE

Following [194], we can restrict w.l.o.g our attention to a "homogeneous" secret univariate HFE-polynomial, i.e. a polynomial of the form:

$$F = \sum_{\substack{0 \leqslant i \leqslant j < n \\ q^i + q^j \leqslant D}} A_{i,j}\, X^{q^i+q^j} \in \mathbb{F}_{q^n}[X]. \tag{3.14}$$

This is not a restriction since what follows can easily be adapted to the affine case as mentioned in [194]. A polynomial F as in (3.14) can be written as "*non-standard quadratic forms*" [194]. That is, letting $\mathbf{F} = (A_{i,j})_{0 \leqslant i < n}^{0 \leqslant j < n} \in \mathcal{M}_{n\times n}(\mathbb{F}_{q^n})$, we can write:

$$F = \sum_{i=0}^{n-1}\sum_{j=0}^{n-1} A_{i,j} X^{q^i+q^j} = \underline{X}\mathbf{F}\underline{X}^T, \text{ where } \underline{X} = (X, X^q, \ldots, X^{q^{n-1}}).$$

The fundamental remark is that the rank of $\mathbf{F}$ is bounded from above. Indeed, the degree of the secret polynomial is smaller than D and the entries $A_{i,j}$ in $\mathbf{F}$ are non-zero if and only if $i, j \leqslant \log_q(D)$. This yields:

$$\text{Rank}\left(\mathbf{F}\right) \leqslant \log_q(D).$$

Let also $\mathbf{F}^{*k} \in \mathcal{M}_{n\times n}(\mathbb{F}_{q^n})$ be the matrix whose (i,j)-th entry is $f_{i-k,j-k}^{q^k}$ (indexes are modulo n). The matrix $\mathbf{F}^{*k}$ is in fact the "*matrix representation*" of the q^k-th power of the univariate polynomial F. That is:

$$F^{q^k} = \sum_{i=0}^{n-1}\sum_{j=0}^{n-1} A_{i-k,j-k}^{q^k} X^{q^i+q^j} \quad = \underline{X}\mathbf{F}^{*k}\underline{X}^T.$$

Thanks to Proposition 7, we deduce a useful property on the matrices $\mathbf{F}^{*k}$s.

Lemma 1. *Let $\mathbf{M}_n \in \mathrm{GL}_n\left(\mathbb{F}_{q^n}\right)$ be the matrix defined in Proposition 7. We consider also the symmetric matrices $(\mathbf{G}_1, \ldots, \mathbf{G}_n) \in \left(\mathcal{M}_{n\times n}(\mathbb{F}_q)\right)^n$ associated to quadratic polynomials $(p_1, \ldots, p_n) \in \left(\mathbb{F}_q[x_1, \ldots, x_n]\right)^n$ of a HFE public-key, i.e. $p_i = \underline{x}\mathbf{G_i}\underline{x}^T$ for all i, $1 \leqslant i \leqslant n$. It holds that:*

$$(\mathbf{G}_1, \ldots, \mathbf{G}_n) = \left(\mathbf{S}\,\mathbf{M}_n\,\mathbf{F}^{*0}\,\mathbf{M}_n^T\,\mathbf{S}^T, \ldots, \mathbf{S}\,\mathbf{M}_n\mathbf{F}^{*n-1}\,\mathbf{M}_n^T\,\mathbf{S}^T\right)\mathbf{M}_n^{-1}\,\mathbf{T}.$$

As $\mathbf{T}$ and $\mathbf{M}_n$ are invertible, we can rewrite Lemma 1 with:

$$(\mathbf{G}_1, \ldots, \mathbf{G}_n)\,\mathbf{T}^{-1}\mathbf{M}_n = (\mathbf{S}\mathbf{M}_n\mathbf{F}^{*0}\mathbf{M}_n^T\mathbf{S}^T, \ldots, \mathbf{S}\mathbf{M}_n\mathbf{F}^{*n-1}\mathbf{M}_n^T\mathbf{S}^T). \tag{3.15}$$

In other words, we have a direct relation between the polynomials of the public key written as quadratic forms and the secret polynomial F or more precisely the matrices $\mathbf{F}^{*i}$, for all $i, 0 \leqslant i < n$. Notice that equation (3.15) involves left products of a matrix with $\mathbf{M}_n$. This product has an interesting property.

Proposition 9. *Let* $\mathbf{M}_n \in \mathrm{GL}_n\left(\mathbb{F}_{q^n}\right)$ *be the matrix defined in Proposition 7. Let also* $\mathbf{A} = (a_{i,j})_{0\leqslant i<n}^{0\leqslant j<n} \in \mathcal{M}_{n\times n}(\mathbb{F}_q)$, *and* $\mathbf{B} = (b_{i,j})_{0\leqslant i<n}^{0\leqslant j<n} = \mathbf{A}\,\mathbf{M}_n \in \mathcal{M}_{n\times n}(\mathbb{F}_{q^n})$. *We have:*

$$b_{i,j} = b_{i,j-1}^q, \text{ for all } i,j, 0 \leqslant i,j < n.$$

That is, each column is obtained from the previous one using a Fröbenius application. As a consequence, the whole matrix $\mathbf{B}$ *can be defined with any of its columns.*

From now on, we write $\mathbf{T}^{-1}\,\mathbf{M}_n = \mathbf{U} = (u_{i,j})_{0\leqslant i<n}^{0\leqslant j<n} \in \mathrm{GL}_n\left(\mathbb{F}_{q^n}\right)$ and $\mathbf{S}\,\mathbf{M}_n = \mathbf{W} = (w_{i,j})_{0\leqslant i<n}^{0\leqslant j<n} \in \mathrm{GL}_n\left(\mathbb{F}_{q^n}\right)$. We can then rewrite (3.15) as follows:

$$(\mathbf{G}_1, \ldots, \mathbf{G}_n)\,\mathbf{U} = (\mathbf{W}\mathbf{F}^{*0}\mathbf{W}^T, \ldots, \mathbf{W}\mathbf{F}^{*n-1}\mathbf{W}^T).$$

According to Proposition 9, $u_{i,j+1} = u_{i,j}^q$ and $w_{i,j+1} = w_{i,j}^q$, for all $i,j, 0 \leqslant i,j < n$. Thus, we only need to know one column of $\mathbf{U}$ (resp. $\mathbf{W}$) to recover the whole matrix. Let then $(u_{0,0}, \ldots, u_{n-1,0}) \in (\mathbb{F}_{q^n})^n$ be the components of the first column of $\mathbf{U}$. We have:

$$\sum_{k=0}^{n-1} u_{k,0}\mathbf{G}_{\mathbf{k+1}} = \mathbf{W}\mathbf{F}^{*0}\mathbf{W}^T = \mathbf{W}\mathbf{F}\mathbf{W}^T. \tag{3.16}$$

As the rank of $\mathbf{F}$ is $\log_q(D)$, so is the rank of $\mathbf{W}\mathbf{F}\mathbf{W}^T$. On the other hand, the solution of such MinRank lies in $(\mathbb{F}_{q^n})^n$. This leads to the following theorem.

Theorem 3.4.1. *For* HFE, *recovering* $\mathbf{U} = \mathbf{T}^{-1}\,\mathbf{M}_n \in \mathcal{M}_{n\times n}(\mathbb{F}_{q^n})$ *reduces to solve a* MinRank *on the public matrices* $(\mathbf{G}_1, \ldots, \mathbf{G}_n) \in \mathcal{M}_{n\times n}(\mathbb{F}_q)^n$ *with target rank* $r = \lceil \log_q(D) \rceil$. *The solutions (i.e. the linear combinations) of this* MinRank *problem are in* $(\mathbb{F}_{q^n})^n$.

The equation (3.16) is similar to the so-called "*fundamental equation*" in the basic KS attack [194]. The fundamental difference is that we have not used the univariate representation of the public-key as initially presented in [194]. This is arguably a simpler and more natural way to formalize the MinRank in HFE. We also explain in [52] that this permits to speed-up the solving MinRank involved. The speed-up is due to the fact that MinRank of [194] is defined over the extension field, whilst our is defined over the base field. The expected gain is a factor $\mathrm{M}(n) \sim n^2$ (the cost of the multiplication of two univariate polynomials of degree n) over the classical Kipnis-Shamir attack.

3.4.2 Complexity Analysis of the MinRank Attack

We bound here complexity of solving the MinRank instances arising in HFE (Theorem 3.4.1). In [194], it is conjectured that the basic KS attack against HFE is sub-exponential is n. The complexity turned to be incorrect [103]. Under a regularity assumption, we show in [52] that solving the MinRank instances of Theorem 3.4.1 is exponential in $r = \lceil \log_q(D) \rceil$, where D is the degree of the HFE polynomial. This result is obtained thanks to the minor modeling described in Section 3.3 and Proposition 8.

Proposition 10. *Let* $(q, n, D) \in \mathbb{N}^3$ *be the parameters of a* HFE *and* $r = \lceil \log_q(D) \rceil$. *If the variant of the Fröberg conjecture as defined in [151] is true, then the complexity of solving the* MinRank *arising in* HFE – *when* $r < 11$ – *with the minor modeling is*

$$\mathcal{O}\left(n^{(r+1)\,\omega}\right), \text{ with } \omega, 2 \leqslant \omega < 3 \text{ being the linear algebra constant.}$$

This makes the binary complexity logarithmic in q. We can observe that the upper bound (3.8) on the degree of regularity for a direct message recovery-attack from [105] is exponential in q; our bound for key-recovery is logarithmic in q.

3.5 Final Remarks

Until the mid 2000's, multivariate cryptography was developing very rapidly, producing many interesting and versatile design ideas such as C^* [212], HFE [231], SFLASH [84], UOV [192], TTS [273], Rainbow [106], ... This was then naturally followed by an intense period of cryptanalysis, e.g. [49], [51], [113], [146], [155], [181], [212], [268]. This process permitted to filter out the weakest primitives and to develop powerful cryptanalytic techniques for MPKC which also impacted others quantum-safe cryptosystems as illustrated by this document. We have a now a better understanding of the security of these schemes. In particular, the results presented in this chapter provide a set reference tools allowing to evaluate the security of MPKC.
Now, after an intense period of cryptanalysis it appears that few schemes resisted to the taste of time : UOV (1999) and variants of HFE (1995). As a consequence, an open problem in this area is to design good proposals for quantum-safe standards. In particular, we observe that the complexity of the best attacks against HFE are all exponential in $O(\log_q(D))$. We have essentially only one parameter which allows to control the security and efficiency of this scheme : the degree D of the HFE polynomial.
We are now in a better position to derive secure parameters for schemes such as HFE- or HFEv. The challenge is different to design an encryption scheme or a signature scheme. Remark that historically multivariate cryptography has always been more successful in the design of signature schemes. The most interesting feature of multivariate signatures is that they can lead to extremely short signatures (3.2); the shortest among quantum-safe cryptosystems. In the encryption case, it is more challenging to find a good efficiency/security trade-off between the degree D, and the use of the modifiers (minus or vinegar). The point is that modifiers slow down the decryption process so that it is mandatory to consider rather large D to achieve a good security level and keeping efficiency. In the case of signature, the use of modifiers do not induce a loss of efficiency and allows one to consider smaller D.

CHAPTER 4

ALGEBRAIC TECHNIQUES IN CODE-BASED CRYPTOGRAPHY

*This chapter details (some) algebraic attacks against McEliece's cryptosystem. It is based, in particular, on [122], [131], [133], [135]. We start by a short introduction to McEliece's cryptosystem (Section 4.1). We then describe how to derive a set of (structured) algebraic equations modeling key-recovery in McEliece (Section 4.2). We explain how this algebraic modeling can be used to attack compact variants of McEliece. Finally, we present in Section 4.3 an algebraic algorithm for distinguishing a McEliece public-key from a random matrix. This problem is known as the Goppa Code Distinguishing (*GD*) problem. Our algorithm allows to solve* GD *in polynomial-time for certain ranges of parameters.*

4.1 McEliece Public-Key Cryptosystems

After almost forty years now, the McEliece public-key encryption scheme [214] still belongs to the very few public-key cryptosystems which remain unbroken. This is remarkable regarding the intense academic activity in cryptanalysis. As a consequence, McEliece's cryptosystem is a serious candidate for quantum-safe standards.
The security of McEliece is motivated by the intractability (i.e. NP-Hardness) of decoding linear codes [40].

Bounded Distance Decoding (BDD)
Input. $\mathbf{G} \in \mathcal{M}_{k\times n}(\mathbb{F}_q)$, $\mathbf{c} \in \mathbb{F}_q^n$ and an integer $t > 0$.
Goal. Find – if any – $\mathbf{m} \in \mathbb{F}_q^k$ such that:

$$\text{HammingWeight}(\mathbf{c} - \mathbf{m}G) \leqslant t,$$

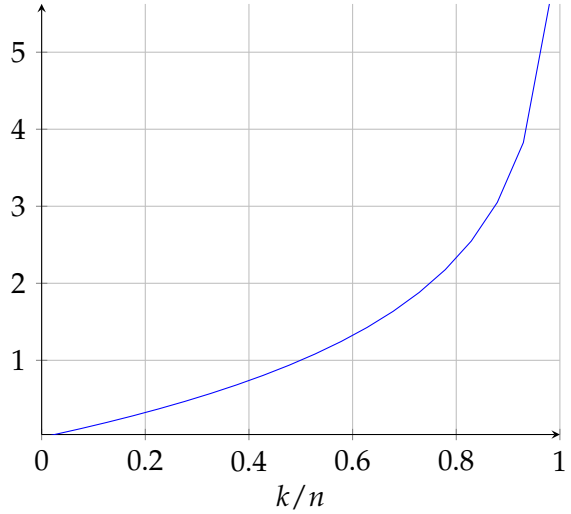
with HammingWeight() being the number of non-zero coordinates.
Decoding a random linear code, that is solving BDD for a random $\mathbf{G} \in \mathcal{M}_{k\times n}(\mathbb{F}_q)$, is a long-standing problem whose most effective algorithms, e.g. [41], [75], [198], [199], [213], [241], [263], have all an exponential time complexity in the classical [270] as well as in the quantum setting [42]. Although the complexity of the best decoding attack remains exponential, progress on the exact exponent have been continuously reported. The latest result from [213] brings down the complexity to (at most) $2^{0.097\,n}$ for decoding random binary linear codes of length n and dimension k when $R = k/n \approx 0.447$.

The most efficient algorithms for decoding random linear codes are based on a framework called the (*Information Set Decoding* (ISD) first introduced by Prange in [241]. In [270], the authors prove that if the number of errors $t = o(n)$ is sub-linear, the asymptotic complexity of all ISD variants is :

$$2^{c\,t\,(1+o(1))}, \tag{4.1}$$

where $c = -\log_2(1-k/n)$ only depends on the ratio k/n.

Value of c in the exponent of (4.1).



In some sense, the principle of McEliece's cryptosystem [214] is similar to MPKC (Section 3.1). We start from a structured code with an efficient decoding algorithm, typically *binary Goppa codes* (Definition 4.1.2), and we publish a hidden version of this structured code. We describe McEliece's encryption algorithm in Algorithm 3. Before that, we introduce minimal materials to describe McEliece. For more details regarding this part, we refer for instance to [210, Ch. 12] or [229].

Preliminaries. Let $q = p^s$ with p prime and an integer $s \leqslant 1$. It is convenient to describe *Goppa codes* through a parity-check matrix defined over an extension field $\mathbb{F}_{q^m}$ of $\mathbb{F}_q$ over which the code is defined:

$$\begin{aligned}\mathbf{V}_t(\mathbf{x},\mathbf{y}) &= \begin{pmatrix} 1 & \cdots & 1 \\ x_0 & \cdots & x_{n-1} \\ \vdots & & \vdots \\ x_0^{t-1} & \cdots & x_{n-1}^{t-1} \end{pmatrix} \begin{pmatrix} y_0 & \cdots & \cdots & 0 \\ 0 & y_1 & \cdots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & \cdots & \cdots & y_{n-1} \end{pmatrix} \\ &= \begin{pmatrix} y_0 & \cdots & y_{n-1} \\ y_0x_0 & \cdots & y_{n-1}x_{n-1} \\ \vdots & & \vdots \\ y_0x_0^{t-1} & \cdots & y_{n-1}x_{n-1}^{t-1} \end{pmatrix} \in \mathcal{M}_{t\times n}(\mathbb{F}_{q^m}),\end{aligned} \tag{4.2}$$

with $\big(\mathbf{x} = (x_0,\ldots,x_{n-1}), \mathbf{y} = (y_0,\ldots,y_{n-1})\big) \in \mathbb{F}_{q^m}^n \times \mathbb{F}_{q^m}^n$.
The kernel of $\mathbf{V}_t$ on the base field $\mathbb{F}_q$ defines a linear codes which includes binary Goppa codes.

Definition 4.1.1 (Alternant code). *Let $\mathbf{x} = (x_0,\ldots,x_{n-1}) \in (\mathbb{F}_{q^m})^n$ where all x_i's are distinct and $\mathbf{y} = (y_0,\ldots,y_{n-1}) \in \left(\mathbb{F}_{q^m}^{\times}\right)^n$. The alternant code of order t over $\mathbb{F}_q$, associated to $\mathbf{x}$ and $\mathbf{y}$, is defined as follows:*

$$\mathscr{A}_t(\mathbf{x},\mathbf{y}) = \left\{\mathbf{c} \in \mathbb{F}_q^n \mid \mathbf{V}_t(\mathbf{x},\mathbf{y})\mathbf{c}^T = \mathbf{0}_t\right\}, \text{with } \mathbf{V}_t(\mathbf{x},\mathbf{y}) \text{ defined as in (4.2).}$$

The dimension k of $\mathscr{A}_t(\mathbf{x},\mathbf{y})$ satisfies $k \geqslant n - tm$. We shall call $\mathbf{x}$ the support of the code, $\mathbf{y}$ the multiplier and t the order (or degree) of the alternant code.

A key feature about alternant codes of degree t is the fact that there exists a polynomial-time algorithm decoding all errors of weight at most $\frac{t}{2}$ once a parity-check matrix of the form $\mathbf{V}_t(\mathbf{x}, \mathbf{y})$ is given [210, Ch.12]. For subclasses of alternant codes, algorithms correcting more errors can be found. In particular:

Definition 4.1.2 (Goppa codes)**.** *The Goppa code $\mathscr{G}(\mathbf{x}, \Gamma)$ over $\mathbb{F}_q$ associated to a polynomial $\Gamma(z) \in \mathbb{F}_{q^m}[z]$ of degree t and n-tuple $\mathbf{x} = (x_0, \ldots, x_{n-1}) \in (\mathbb{F}_{q^m})^n$ of distinct elements of $\mathbb{F}_{q^m}$ satisfying $\Gamma(x_i) \neq 0$ for all $i, 0 \leqslant i \leqslant n-1$, is the alternant code $\mathscr{A}_t(\mathbf{x}, \mathbf{y})$ of order t with $y_i = \Gamma(x_i)^{-1}$ for all $i, 0 \leqslant i \leqslant n-1$.*

Goppa codes, viewed as alternant codes, naturally inherit a decoding algorithm that corrects up to $\frac{t}{2}$ errors. For binary Goppa codes, we can improve this bound to correct twice as many errors.

Theorem 4.1.1. *Let $\Gamma(z) \in \mathbb{F}_{2^m}[z]$ be a polynomial of degree t without multiple roots and a n-tuple $\mathbf{x} = (x_0, \ldots, x_{n-1})$ of distinct elements of $\mathbb{F}_{q^m}$ satisfying $\Gamma(x_i) \neq 0$ for all $i, 0 \leqslant i \leqslant n-1$. The binary Goppa code $\mathscr{G}(\mathbf{x}, \Gamma)$ is equal to the alternant code $\mathscr{A}_{2t}(\mathbf{x}, \mathbf{y}^2)$, with $y_i = \Gamma(x_i)^{-1}$ for all $i, 0 \leqslant i \leqslant n-1$.*

As a consequence, there exists a polynomial-time algorithm decoding all errors of Hamming weight at most t in $\mathscr{G}(\mathbf{x}, \Gamma)$ as soon as $\mathbf{x}$ and $\Gamma(z)$ are known. We refer to this decoding algorithm as a t-decoder. Remark that $t \in O(n/\log(n)) = o(n)$. The number of errors is then sub-linear in the length for such codes and the complexity (4.1) is indeed valid in this context. For the support $\mathbf{x}$, it is also classical to take all elements of $\mathbb{F}_{q^m}$. This is the *full support* setting where the length is $n = q^m$.

McEliece public-key encryption. We can now explain how the classical McEliece [214] public-key encryption scheme, based on binary Goppa codes, is defined.

Algorithm 3 McEliece Cryptosystem based on binary Goppa codes (1978)

PARAMETERS. Code length n, dimension k, and decoding capacity t.
Plaintext space: $\mathbb{F}_2^k$. Ciphertext space: $\mathbb{F}_2^n$.

KEYGEN. We randomly select a polynomial $\Gamma(z) \in \mathbb{F}_{2^m}[z]$ of degree t without multiple roots and a n-tuple $\mathbf{x} = (x_0, \ldots, x_{n-1})$ of distinct elements of $\mathbb{F}_{2^m}$ satisfying $\Gamma(x_i) \neq 0$ for all $i, 0 \leqslant i \leqslant n-1$. The pair $(\Gamma, \mathbf{x})$ defines a binary Goppa code $\mathscr{G}(\mathbf{x}, \Gamma)$ whose parity-check matrix is $\mathbf{V}_{2t}\big(\mathbf{x}, \Gamma^{-2}(\mathbf{x})\big)$.

PRIVATE-KEY. The pair $(\Gamma, \mathbf{x}) \in \mathbb{F}_{2^m}[z] \times \mathbb{F}_{2^m}^n$ which defines a t-decoder T_t for $\mathbf{G}$.

PUBLIC-KEY. The correction capacity t, and a full-rank matrix $\mathbf{G} \in \mathcal{M}_{k \times n}(\mathbb{F}_2)$ with $k, n - mt \leqslant k < n$, such that

$$\mathbf{V}_t\big(\mathbf{x}, \Gamma^{-1}(\mathbf{x})\big)\mathbf{G}^T = (\mathbf{0})_{t \times k}\,. \tag{4.3}$$

ENCRYPT.
1: Input $\mathbf{m} \in \mathbb{F}_2^k$
2: Generate random $\mathbf{e} \in \mathbb{F}_2^n$ with Hamming weight t
3: Output $\mathbf{c} = \mathbf{m}\mathbf{G} + \mathbf{e}$

DECRYPT.
1: Input $\mathbf{c} \in \mathbb{F}_2^n$
2: Compute $\bar{\mathbf{m}} = T_t(\mathbf{c})$
3: If decoding succeeds, output $\bar{\mathbf{m}}$, else output $\perp$

Remark 3. *Our description of the McEliece encryption scheme slightly differs from the traditional description; as can be found in [229]. for instance. Usually, the public-key is given by* $\mathbf{G}' = \mathbf{S}\,\mathbf{G}\,\mathbf{P}$ *where* $\mathbf{S} \in \mathrm{GL}_k(\mathbb{F}_2), \mathbf{G} \in \mathcal{M}_{k\times n}(\mathbb{F}_2)$ *is the generator matrix of* $\mathscr{G}(\mathbf{x}, \Gamma)$ *defined as in* (4.3)*, and* $\mathbf{P} \in \mathrm{GL}_n(\mathbb{F}_2)$ *is a permutation matrix. It can be proven that* $\mathbf{G}' \in \mathcal{M}_{k\times n}$ *is the generator matrix of the Goppa code* $\mathscr{G}(\mathbf{x}_\sigma, \Gamma)$*, where* σ *is the permutation associated to* $\mathbf{P}$ *and* $\mathbf{x}_\sigma = (x_{\sigma(0)}, \ldots, x_{\sigma(n-1)}) \in \mathbb{F}_{2^m}^n$*. Thus, the effect of* $(\mathbf{S}, \mathbf{P})$ *is to permute the support of* $\mathscr{G}(\mathbf{x}, \Gamma)$*. Since the support* $\mathbf{x}$ *is already chosen at random, it not clear that* $(\mathbf{S}, \mathbf{P})$ *will bring any additional security.*

Security of McEliece cryptosystems. It is clear that message-recovery reduces to decode the public linear code given by $\mathbf{G} \in \mathcal{M}_{k\times n}(\mathbb{F}_2)$, i.e. to solve BDD on $\mathbf{G}$. We can consider that direct decoding attacks against McEliece, i.e. by applying ISD, with binary codes is a topic which has been well investigated in the litterature. Beside asymptotic results for ISD as given in (4.1), concrete security level for McEliece can also be computed against ISD [41], [229]. The situation is rather different for the key-recovery problem. The reference attack remains the Support Splitting Attack (SSA) proposed by Sendrier and Loidreau in [203]. SSA is essentially an exhaustive search to recover the description as a binary Goppa code of the public matrix $\mathbf{G} \in \mathcal{M}_{k\times n}(\mathbb{F}_2)$. This description provides an efficient way to decode the public code and then a key equivalent to the secret-key. The most costly part of SSA – in the case of binary Goppa codes – is to perform an exhaustive search to recover a Goppa polynomial (Definition 4.1.2) $\Gamma \in \mathbb{F}_{2^m}[z]$ of degree t. Detailed analysis provided in [203] shows that the complexity of SSA is bounded from above by $\frac{n^3}{mt}\,2^{m(t-3)}$. Note that this complexity is achieved if the codes considered are of full support, i.e. $n = 2^m$.

McEliece signature. Besides public-key encryption, one can also sign using the principle of McEliece thanks to Courtois, Finiasz, and Sendrier (CFS signature scheme, [87]). The public-key $\mathbf{G} \in \mathcal{M}_{k\times n}(\mathbb{F}_2)$ in CFS is constructed as in McEliece's encryption. A valid signature $\mathbf{s} \in \mathbb{F}_2^k$ for digest $\mathbf{d} \in \mathbb{F}_2^n$ is such that the Hamming weight of $\mathbf{d} - \mathbf{s}\mathbf{G}$ is at most t. With a binary Goppa codes of length $n = 2^m$ and order t, the probability that a given digest admits a valid signature is of order $\frac{1}{t!}$. The idea then is modify the digest by appending a counter incremented until the decoding algorithm can find such a signature. The efficiency of this scheme heavily depends on the number of such trials. Thus, one has to choose a very small t. To maintain security against generic message-recovery attacks, this implies to consider codes with high rates $R = k/n$ [87], [167]. This can be understood from the asymptotic complexity (4.1) which is exponential in $c\,t$ where c increases with the code rate. For instance, [167] recommended to take $n = 2^{21}$ and $t = 10$ (this yields $R = 0.99$) for a 80-bit security CFS scheme whereas [167] suggested to take $n = 2^{11}$ and $t = 32$ (so that $R = 0.82$) to achieve the same security level for encryption. Thus one major difference between the McEliece cryptosystem and the CFS scheme lies in the choice of the codes parameters.

4.2 Algebraic Key-Revovery Systems for McEliece Cryptosystems

We present here a new method to assess the security of McEliece. It turns out that the very definition of alternant codes allows to derive an algebraic modeling of the key-recovery. The modeling can be refined for (binary) Goppa codes. We then will explain how this approach can be used to attack efficiently so called *compact variants* of McEliece proposed in [33], [39], [220].

4.2.1 Algebraic Modelings of Key-Recovery

Alternant modeling. Let $q = p^s > 1$ be a prime power. We first assume that public-key is a q-ary alternant code $\mathscr{A}_t(\mathbf{x}, \mathbf{y})$ defined by the public generator matrix $\mathbf{G} = (g_{i,j}) \in \mathcal{M}_{k\times n}(\mathbb{F}_q)$ (Definition 4.1.1). In this case, the secret-key is $(\mathbf{x}, \mathbf{y}) \in \mathbb{F}_{q^m}^n \times \mathbb{F}_{q^m}^n$. Let $\mathbf{X} = (X_0, \ldots, X_{n-1})$ and $\mathbf{Y} = (Y_0, \ldots, Y_{n-1})$ be two sets of variables corresponding to the unknown support $\mathbf{x}$ and multiplier $\mathbf{y}$ respectively. To model algebraically the key-recovery, we use that $\mathbf{V}_t(\mathbf{X}, \mathbf{Y})$ is a parity-check matrix (Section 4.1) of the public-code. That is $\mathbf{V}_t(\mathbf{X}, \mathbf{Y})\mathbf{G}^T = (\mathbf{0})_{t\times k}$ holds and yields:

$$\mathsf{A}_{\mathbf{X},\mathbf{Y}} = \bigcup_{\ell=0}^{t-1} \left\{ \sum_{j=0}^{n-1} g_{i,j} X_j^\ell Y_j \;\middle|\; 0 \le i \le k-1 \right\}. \tag{4.4}$$

By construction, it is clear that $(\mathbf{x}, \mathbf{y}) \in \mathbb{F}_{q^m}^n \times (\mathbb{F}_{q^m}^\times)^n$ is a solution of this system. The non-linear system $\mathsf{A}_{\mathbf{X},\mathbf{Y}}$ has $2n$ variables and kt equations of degrees $1, \ldots, t$. We can also remark that the system is defined over $\mathbb{F}_q$ whilst the solutions lie in $\mathbb{F}_{q^m}$. The system (4.4) is also over-determined for typical cryptographic parameters. This feature will be maximized for codes with high rates used in CFS as we will see in Section 4.3. We also observe that $\mathsf{A}_{\mathbf{X},\mathbf{Y}}$ is very structured. Indeed, the only monomials occurring are of the form $Y_j X_j^\ell$ with $\ell, 0 \le \ell \le t-1$, i.e. all equations are bi-homogeneous (Definition 1.1.8) of bi-degree $(1, \ell)$.

Remark 4. *By definition, the support and multiplier $(\mathbf{x}, \mathbf{y}) \in \mathbb{F}_{q^m}^n \times (\mathbb{F}_{q^m}^\times)^n$ of an alternant code must verify additional algebraic relations (Definition 4.1.1) that has not been included so far. As a consequence, the system $\mathsf{A}_{\mathbf{X},\mathbf{Y}}$ will have parasite solutions corresponding to $X_i = X_j$ or to $Y_j = 0$). A classical way to remove such spurious relations is to introduce new variables u_{ij} and v_i and add equations of the form $u_{ij} \cdot (X_i - X_j) + 1 =$ and $v_i \cdot Y_i + 1 = 0$. In practice, we have however used a small number of such equations (namely 4 or 5). The reason is that we do not want to add too many new variables. To simplify the description, we will ignore these equations from now on.*

By assumption, the public-code defined by $\mathbf{G} \in \mathcal{M}_{k\times n}(\mathbb{F}_q)$ is of dimension k. Up to Gaussian elimination (and possibly reordering the positions), we can assume that $\mathbf{G}$ is in systematic form $\left(\mathbf{I}_k \;\middle|\; \mathbf{P} \right)$ where $\mathbf{I}_k$ is the $k \times k$ identity matrix and $\mathbf{P} = (p_{i,j}) \in \mathcal{M}_{k\times(n-k)}(\mathbb{F}_q)$. We can then rewrite system (4.4) as follows:

$$\mathsf{A}_{\mathbf{X},\mathbf{Y}} = \bigcup_{\ell=0}^{t-1} \left\{ X_i^\ell Y_i + \sum_{j=k}^{n-1} p_{i,j} X_j^\ell Y_j \;\middle|\; 0 \le i \le k-1 \right\}. \tag{4.5}$$

By focusing on equations with $\ell = 0$, we construct a new polynomial system $\mathsf{A}_{\mathbf{X},\mathbf{Y}'}$ – where $\mathbf{Y}' = (Y_k, \ldots, Y_{n-1})$ – in which the variables $(Y_0, \ldots, Y_{k-1})$ have been eliminated:

$$\mathsf{A}_{\mathbf{X},\mathbf{Y}'} = \bigcup_{\ell=1}^{t-1} \left\{ \sum_{j=k}^{n-1} p_{i,j} \left(X_j^\ell - X_i^\ell \right) Y_j \;\middle|\; 0 \le i \le k-1 \right\}.$$

The non-linear system $\mathsf{A}_{\mathbf{X},\mathbf{Y}'}$ has $2\,n-k$ variables and $k(t-1)$ equations of degrees $2, \ldots, t$.
In $\mathsf{A}_{\mathbf{X},\mathbf{Y}'}$, we have been able to eliminate variables in the block $\mathbf{Y}$ by performing a simple manipulation of the equations from $\mathsf{A}_{\mathbf{X},\mathbf{Y}}$. We can combine the equations of $\mathsf{A}_{\mathbf{X},\mathbf{Y}}$ in a different way to eliminate variables in the two blocks $\mathbf{X}$ and $\mathbf{Y}$. To illustrate the principle, we consider the equations of bi-degree $(0,1)$, $(1,1)$ and $(1,2)$ in $\mathsf{A}_{\mathbf{X},\mathbf{Y}}$, i.e.

$$\bigcup_{\ell=0}^{2} \left\{ X_i^\ell Y_i + \sum_{j=k}^{n-1} p_{i,j} X_j^\ell Y_j \;\middle|\; 0 \le i \le k-1 \right\}.$$

Then, thanks to the trivial identity $Y_i(X_i^2\, Y_i) = (X_i Y_i)^2$ for all $i, 0 \leqslant i \leqslant k-1$, we get:

$$\left(\sum_{j=k}^{n-1} p_{i,j} Y_j\right)\left(\sum_{j'=k}^{n-1} p_{i,j'} X_{j'}^2 Y_{j'}\right) = \left(\sum_{j=k}^{n-1} p_{i,j} X_j Y_j\right)\left(\sum_{j'=k}^{n-1} p_{i,j'} X_{j'} Y_{j'}\right) = \left(\sum_{j=k}^{n-1} p_{i,j} X_j Y_j\right)^2.$$

This gives a system of k equations in $\mathbf{X}' = (X_k, \ldots, X_{n-1})$ and $\mathbf{Y}' = (Y_k, \ldots, Y_{n-1})$:

$$\left\{ \sum_{j=k}^{n-1} \sum_{j' \geqslant j}^{n-1} p_{i,j} p_{i,j'} \left(X_{j'}^2 + X_j^2\right) Y_j Y_{j'} - \left(\sum_{j=k}^{n-1} p_{i,j} X_j Y_j\right)^2 \;\middle|\; 0 \leq i \leq k-1 \right\}. \tag{4.6}$$

In fact, we can combine the equations from $\mathsf{A}_{\mathbf{X},\mathbf{Y}}$ in many different ways. That is, for all $a, b, c, d \in \{0, 1, \ldots, t-1\}$ such that $a + b = c + d$, we have $Y_i X_i^a Y_i X_i^b = Y_i X_i^c Y_i X_i^d$. To define this new system, we set:

$$J_t = \Big\{(a, b, c, d) \in \mathbb{N}^4 \mid 0 \leqslant a, b, c, d \leqslant t-1 \text{ and } a + c = b + d\Big\}.$$

We can thus set a new system:

$$\mathsf{elimA}_{\mathbf{X}',\mathbf{Y}'} = \bigcup_{(a,b,c,d)\in J_t} \left\{ \sum_{j=k}^{n-1} \sum_{j' \geqslant j}^{n-1} p_{i,j} p_{i,j'} \left(X_j^a X_{j'}^b + X_j^b X_{j'}^a - X_j^c X_{j'}^d - X_j^d X_{j'}^c\right) Y_j Y_{j'} \;\middle|\; 0 \leq i \leq k-1 \right\}. \tag{4.7}$$

The system $\mathsf{elimA}_{\mathbf{X}',\mathbf{Y}'}$ has now $2(n-k) = 2n(1-k/n) = 2n(1-R)$ variables. We emphasize that not all $(a, b, c, d) \in J_t$ would lead to new equations. For instance $a = c$ and $b = d$ leads to a tautology. In Section 4.3, we use $\mathsf{elimA}_{\mathbf{X}',\mathbf{Y}'}$ in the context of the distinguisher. In fact, a core idea of the distinguisher is to count the number of quadruplets $(a, b, c, d) \in J_t$ leading to non-trivial and linearly independently equations.

Goppa codes. The system $\mathsf{A}_{\mathbf{X},\mathbf{Y}'}$ only describes the fact that the public key is the generator matrix of an alternant code. When the public-key is a Goppa code $\mathscr{G}(\mathbf{x}, \Gamma)$ over $\mathbb{F}_q$, the secret-key $\left(\mathbf{x}, \mathbf{y} = \frac{1}{\Gamma(\mathbf{x})}\right) \in \mathbb{F}_{q^m}^n \times \mathbb{F}_{q^m}^n$ vanishes $\mathsf{A}_{\mathbf{X},\mathbf{Y}'}$ but also additional constraints; there is an algebraic relation between the support $\mathbf{x}$ and the multiplier $\mathbf{y}$. Indeed, Definition 4.1.2 implies that there exists $\Gamma(z) = \sum_{\ell=0}^{t} \gamma_\ell z^\ell \in \mathbb{F}_{q^m}[z]$ of degree t such that $Y_j \Gamma(X_j) = \sum_{\ell=0}^{t} \gamma_\ell Y_j X_j^\ell = 1$. A convenient way to include the algebraic equations derived from these constraints is to use the so-called extended codes [210, Ch. 1, p. 27]. Let Z be a new variable for $\frac{1}{\gamma_t}$. We can then obtain a new polynomial system $\mathsf{G}_{\mathbf{X},\mathbf{Y}'}$ dedicated to q-ary Goppa codes:

$$\mathsf{G}_{\mathbf{X},\mathbf{Y}'} = \mathsf{A}_{\mathbf{X},\mathbf{Y}'} \cup \left\{ \sum_{j=k}^{n-1} p_{i,j} \left(X_j^t - X_i^t\right) Y_j = Z\left(1 + \sum_{j=0}^{n-k-1} p_{i,j}\right) \mid 0 \leq i \leq k-1 \right\}.$$

$\mathsf{G}_{\mathbf{X},\mathbf{Y}'}$ has $1 + 2\,n - k$ variables and $(k+1)t$ equations of degrees $2, \ldots, t+1$.

Binary Goppa codes. The case $q = 2$ is even more specific. Binary Goppa codes can be viewed as alternant codes $\mathscr{A}_t(\mathbf{x}, \mathbf{y})$ with $y_j = \Gamma(x_j)^{-1}$ for all $j, 0 \leq j \leq n-1$ but also described as a binary alternant code $\mathscr{A}_{2t}(\mathbf{x}, \mathbf{y}^2)$ (Theorem 4.1.1). This brings equations to the system $\mathsf{G}_{\mathbf{X},\mathbf{Y}'}$ which are defined by

$$\bigcup_{\ell=0}^{2t-1} \left\{ X_i^\ell\, Y_i^2 + \sum_{j=k}^{n-1} p_{i,j} X_j^\ell\, Y_j^2 \;\middle|\; 0 \leq i \leq k-1 \right\}.$$

The equations obtained with $\ell = 2\ell'$, such that $\ell', 0 \leq \ell' \leq t-1$ are the squares of equations of $\mathsf{A}_{\mathbf{X},\mathbf{Y}'}$ (as in this case $p_{i,j}^2 = p_{i,j}$). So, they bring no information. However, the equations in bi-degree $(2, 2\ell'+1)$ are new. This enables to define a specific algebraic system dedicated to McEliece's cryptosystem:

$$\mathsf{McE}_{\mathbf{X},\mathbf{Y}'} = \mathsf{G}_{\mathbf{X},\mathbf{Y}'} \cup \bigcup_{\ell=0}^{t-1} \left\{ \sum_{j=k}^{n-1} p_{i,j} Y_j^2 \left(X_j^{2\ell+1} - X_i^{2\ell+1} \right) = 0 \mid 0 \leq i \leq k-1 \right\}.$$

$\mathsf{McE}_{\mathbf{X},\mathbf{Y}'}$ has $1 + 2n - k$ variables and $k(2t-1)$ equations of degrees $2, \ldots, 2t+1$. We can use [131] a technique similar to $\mathsf{elimA}_{\mathbf{X}',\mathbf{Y}'}$ for eliminating the variables $(X_0, \ldots, X_{k-1})$ from $\mathsf{G}_{\mathbf{X},\mathbf{Y}'}$ and $\mathsf{McE}_{\mathbf{X},\mathbf{Y}'}$ leading to the systems $\mathsf{elimG}_{\mathbf{X}',\mathbf{Y}'}$ and $\mathsf{elimMcE}_{\mathbf{X}',\mathbf{Y}'}$ respectively.

Remark 5. *The* PQCRYPTO *project recommended [21] the following parameters for McEliece :* $n = 6960$, $k = 5413$ *and* $t = 119$. *This is a conservative choice of the parameters which gives at least* 128 *bits of security against all known quantum and classical attacks. For such parameters, the first algebraic modeling* $\mathsf{A}_{\mathbf{X},\mathbf{Y}}$ *of an alternant code has* 13920 *variables and* 644147 *equations of degree* $1, \ldots, 119$. *The ratio between the number of equations and variables is* ≈ 47.27. *For a binary Goppa code,* $\mathsf{McE}_{\mathbf{X},\mathbf{Y}'}$ *has* 8508 *variables and* 1282881 *equations of degree* $2, \ldots, 239$. *The ratio between the number of equations and number of variables is now of* 150.78. *Note that the system* $\mathsf{elimA}_{\mathbf{X}',\mathbf{Y}'}$ *would have* 3094 *variables.*

4.2.2 Cryptanalysis of Compact Variants of McEliece

An intrinsic practical limitation of McEliece cryptosystems is the size of the public-key. For the parameters recommended by PQCRYPTO for McEliece public-key encryption (Remark 5), the public-key is ≈ 1 GB. To overcome this limitation, a very popular research trend was to decrease the public-key size by focusing on subclasses of alternant/Goppa codes which admit a very compact parity-check or generator matrix, for instance [33], [39], [170], [220], [237]. This reduction is obtained by taking classes of alternant/Goppa codes which have a quasi-cyclic (QC), quasi-dyadic (QD) or more generally quasi-monoidic (QM) generator matrices. The hope is that the additional structure does not deteriorate the security of the system. In particular, it is clear how to use structured generator matrices for improving ISD techniques. This is very much in the spirit of using ideal lattices instead of standard lattices in lattice-based cryptography [207], [262].

In [133], we show that structure can be used to drastically improved an algebraic key-recovery. It turns that the very structure of QD or QC codes [39], [220] allows to obtain linear relations on supports and multipliers. These linear dependencies allow to reduce the number of unknowns without changing the degree of $\mathsf{A}_{\mathbf{X},\mathbf{Y}'}$. This already allowed us [133] to efficiently break several challenges proposed in [39], [220]. More precisely, we have been able to break all parameters proposed in [39] for QC codes. We provide in [135] a first rough complexity estimates of the attack in [133] by considering a sub-system of $\mathsf{A}_{\mathbf{X},\mathbf{Y}'}$ composed by equations of bi-degree $(1, 2^u)$ with $u, 0 \leqslant u \leqslant \log_2(t-1)$. This new system is then quasi-bilinear (Definition 1.1.8), and we bounded the degree of regularity as in Proposition 1.

However, not all parameters for QD codes were broken in [133]. Later, [33] proposes a generalization of QD codes called quasi-monoidic (QM) codes that took into account our algebraic attack from [133]. In [131], we further pushed the cryptanalysis of compact variants of McEliece. We show that the very same reason which allows to construct a compact public-key makes the key-recovery problem intrinsically much easier. The gain on the public-key size induces an important security drop, which is as large as the compression factor p on the public-key. The high

level idea is as follows. Let $\mathbf{G} \in \mathcal{M}_{k\times n}(\mathbb{F}_q)$ be the public key of a compact McEliece. Assume that p is the compression factor of the compact public-key (compared to a plain McEliece). It is possible to construct from the public-key $\mathbf{G}$ another smaller matrix of size $k/p \times n/p$ which is – from an attacker point of view – as good as the initial public-key, i.e. any key-recovery attack can be deployed equivalently on this smaller generator matrix. This implies that a key-recovery on QD and QM schemes is not harder than a key-recovery on a reduced McEliece scheme where all parameters have been scaled down by a factor of p, which is the compression factor allowed by the QD or QM structure. For instance, we can reduce the key-recovery of a QD Goppa code of length 8192 and dimension 4096 (parameters suggested in [220]) to the key-recovery on a QD Goppa code of length 64 and dimension 32. In other words, the very reason which allowed to design compact variants of McEliece can be used to attack such schemes much more efficiently.
To mount the key-recovery attack in practice, we then used the algebraic modelings described before. We mainly used the system $\mathsf{elimA}_{\mathbf{X}',\mathbf{Y}'}$, $\mathsf{elimG}_{\mathbf{X}',\mathbf{Y}'}$ and $\mathsf{elimMcE}_{\mathbf{X}',\mathbf{Y}'}$. For signature schemes based on QD/QM codes, our attack is particularly efficient. In this case, the codes have necessarily a very high rate R. The number of unknowns is $O(n-k) = O\big(n(1-R)\big)$ will be then very small. Consequently, all the parameters suggested for QD-CFS [32] can be broken in a few seconds. For QM-CFS [33], a parameter requires less than 2 hours, but all the others can be broken in a few seconds. In view of our new attack, it seems extremely hard to find parameters of cryptographic interest for friendly-CFS QD/QM codes. In the encryption case, the algebraic systems are harder to solve in practice. Still, we report several successful results against challenges proposed for QD/QM encryption schemes. To measure the progress realized in comparison to [133], we report in Table 4.1 from [133], [135]. The results of [133], [135] were obtained with the FGB software [147]. It is interesting to see that in the non-binary case, our attack can be easily reproduced using on-the-shelf computer algebra system MAGMA [59].

$q=2$	m	t	[131], MAGMA	[131], F_5/FGB	previous attack from [133]	Sec. level
2	16	32	18 s.		N.A.	128
2	12	128		$\leq 2^{83.5}$ op.	N.A.	128
2	14	128		$\leq 2^{96.1}$ op.	N.A.	226
2	15	512		$\leq 2^{146}$ op.	N.A.	256
2	16	256		$\leq 2^{168}$ op.	N.A.	218
2	16	256		$\leq 2^{157}$ op.	N.A.	256
2^4	4	64	0.010 s.		0.50 s	128
2^4	4	128	0.010 s.		7.1 s	128
2^2	8	64	0.040 s.		1,776.3 s	128

Table 4.1: Comparison of the cryptanalysis complexity in this work and [133]. The notation N.A. means that the parameters have not been addressed in [133].

4.3 An Algebraic Distinguisher for Alternant and Goppa Codes

A well-known problem related to the security of McEliece is the *Goppa Code Distinguishing* (GD) problem [87], [257] : given a linear code $\mathscr{C}$ over $\mathbb{F}_q$, we have to decide if $\mathscr{C}$ is a Goppa code or random linear code. In [87], [257], the authors formalized a classical belief about the hardness of GD problem; the so-called "*Goppa Code Indistinguishably Assumption*". This assumption states that there does not exist a polynomial-time computable quantity which behaves differently

depending on whether $\mathscr{C}$ is a Goppa or a random code. This assumption is attractive because it enables to rely on the hardness of decoding a random linear code to prove the security of McEliece [98], [99]. This reasoning is partially motivated by the fact that all known decoding algorithms do not exploit the information, even partially, that a McEliece's public-key describes a "hidden" Goppa code. Also, as we already mentioned, very few structural attacks have been proposed against McEliece's cryptosystems based on binary Goppa codes. Hence, the hardness of the GD problem became a classical belief, and as a consequence, a *de facto* assumption in code-based cryptography. Among others, the assumption has been used prove the semantic security in the standard model (IND-CPA in [227], the IND-CCA2 in [112]), and the security in the random oracle model against existential forgery of CFS [87], [98].
We presented in [122] an efficient algorithm solving the GD problem for certain parameters. We explain here the technique for q-ary alternant codes but the approach can be adapted to Goppa and binary Goppa codes [122]. Let $\mathbf{G} = (g_{i,j}) \in \mathcal{M}_{k\times n}(\mathbb{F}_q)$ be the generator matrix of a q-ary alternant code $\mathscr{A}_t(\mathbf{x}, \mathbf{y})$. We can assume that $\mathbf{G} = (g_{ij}) \in \mathcal{M}_{k\times n}(\mathbb{F}_q)$ is in reduced row echelon form over its k first positions. That is, we can assume that $\mathbf{G} = (\mathbf{I}_k \mid \mathbf{P})$ where $\mathbf{P} = (p_{ij}) \in \mathcal{M}_{k\times(n-k)}(\mathbb{F}_q)$ for $i, 0 \leqslant i \leqslant k-1$ and $j, k \leqslant j \leqslant n-1$ is the submatrix of $\mathbf{G}$ formed by its last $n-k \leqslant mt$ columns.
We assume here that $q = 2^s$ with $s \geqslant 1$. The fundamental idea of the distinguisher is to study the behavior of a Gröbner basis computation on $\mathsf{elimA}_{\mathbf{X}',\mathbf{Y}'}$ where $\mathbf{G} \in \mathcal{M}_{k\times n}(\mathbb{F}_q)$ is a random matrix or a McEliece's public-key. It appears that we can distinguish between these two cases at the very first step of a Gröbner basis computation on this system. When $q = 2^s$ with $s \geqslant 1$, we can rewrite $\mathsf{elimA}_{\mathbf{X}',\mathbf{Y}'}$ as:

$$\bigcup_{(a,b,c,d)\in J_t} \left\{ \sum_{j=k}^{n-2} \sum_{j'>j}^{n-1} p_{i,j}p_{i,j'} \left(X_j^a X_{j'}^b + X_j^b X_{j'}^a + X_j^c X_{j'}^d + X_j^d X_{j'}^c \right) Y_j Y_{j'} \quad \mid \quad 0 \leq i \leq k-1 \right\}. \tag{4.8}$$

The idea then is to consider the equations of lowest degree in (4.8). Namely, we linearize (4.6) leading a linear system $\mathcal{L}_{\mathbf{P}}$ of k equations involving $\binom{n-k}{2} \leqslant \binom{mt}{2}$ variables $Z_{jj'} = Y_j Y_{j'} X_{j'}^2 + Y_{j'} Y_j X_j^2$ which is as follows:

$$\mathcal{L}_{\mathbf{P}} = \begin{cases} \displaystyle\sum_{j=k}^{n-2} \sum_{j'>j}^{n-1} p_{0,j}p_{0,j'} Z_{jj'} = 0 \\ \qquad\qquad \vdots \\ \displaystyle\sum_{j=k}^{n-2} \sum_{j'>j}^{n-1} p_{k-1,j}p_{k-1,j'} Z_{jj'} = 0 \end{cases} \tag{4.9}$$

We can remark that the linear system (4.9) has coefficients in $\mathbb{F}_q$ but the solutions are in the extension field $\mathbb{F}_{q^m}$. We denote by $N_{\mathcal{L}_{\mathbf{P}}} = \binom{mt}{2}$ the maximum number of variables in the linear system $\mathcal{L}_{\mathbf{P}}$ and by $D_{\mathcal{L}_{\mathbf{P}}}$ the dimension of $\mathtt{Ker}(\mathcal{L}_{\mathbf{P}})$ as a $\mathbb{F}_q$-vector space. By definition, we have $D_{\mathcal{L}_{\mathbf{P}}} = N_{\mathcal{L}_{\mathbf{P}}} - \mathtt{rank}(\mathcal{L}_{\mathbf{P}})$. At this point, we have reduced key-recovery against McEliece to the solving of the linear system $\mathcal{L}_{\mathbf{P}}$. However, in order to recover the solutions of the alternant system $\mathsf{A}_{\mathbf{X},\mathbf{Y}}$ (4.4) from the linearized system $\mathcal{L}_{\mathbf{P}}$ it is necessary that $\mathtt{rank}(\mathcal{L}_{\mathbf{P}}) \approx N_{\mathcal{L}_{\mathbf{P}}} = \binom{mt}{2}$. For a random generator matrix $\mathbf{G} \in \mathcal{M}_{k\times n}(\mathbb{F}_q)$, this is likely to happen when $k \geqslant N_{\mathcal{L}_{\mathbf{P}}}$. More precisely, we proved in [122] that:

Theorem 4.3.1. *Let n, k, m and t be integers such that $k = n - mt$. Let $\mathbf{G} = (\mathbf{I}_k \mid \mathbf{P}) \in \mathcal{M}_{k\times n}(\mathbb{F}_q)$ and D_{random} be the dimension of $\mathtt{Ker}(\mathcal{L}_{\mathbf{P}})$ as $\mathbb{F}_q$-vector space when the entries of the matrix $\mathbf{P} \in$*

$\mathcal{M}_{k\times(n-k)}(\mathbb{F}_q)$ are drawn independently from the uniform distribution over $\mathbb{F}_q$. If $k \geqslant \binom{mt}{2}$ then for any function $\omega(x)$ tending to infinity as x goes to infinity, we have

$$\mathbf{Pr}\Big(D_{random} \geqslant mt\,\omega(mt)\Big) = o(1), \text{as } mt \text{ goes to infinity.}$$

Notice that if we choose $\omega(x) = \log(x)$ for instance, then asymptotically the dimension D_{random} of the solution space is with very large probability smaller than $mt\log(mt)$. A stronger result can be derived from [81] and [76]. First, [81] allows to connect the manipulations performed on to $\mathsf{elimA}_{\mathbf{X}',\mathbf{Y}'}$ to the square of the public code defined below.

Definition 4.3.1. *Let $\mathbf{u},\mathbf{v} \in \mathbb{F}_q^n$. The component-wide product $\mathbf{u}\star\mathbf{v}$ is defined as $\mathbf{u}\star\mathbf{v} = (u_0v_0,\ldots,u_{n-1}v_{n-1}) \in \mathbb{F}_q^n$. Let $\mathscr{C} \subseteq \mathbb{F}_q^n$ be a linear code of the length n and dimension k. The square of $\mathscr{C}$, denoted by $\mathscr{C}^2$, is defined as:*

$$\mathscr{C}^2 = \{\mathbf{u}\star\mathbf{v} \mid (\mathbf{u},\mathbf{v}) \in \mathscr{C}\times\mathscr{C}\} \subseteq \mathbb{F}_q^n.$$

We have $\dim(\mathscr{C}^2) = \min\left(n, \binom{k+1}{2}\right)$.

In [81], the authors established an explicit link between the system $\mathcal{L}_{\mathbf{P}}$ and the square of the public parity check-matrix (and a dual version of this statement).

Proposition 4.3.1 ([81])**.** *Let $\mathbf{G} = (\mathbf{I}_k \mid \mathbf{P})$, where $\mathbf{P} = (p_{ij}) \in \mathcal{M}_{k\times(n-k)}(\mathbb{F}_q)$, be the generator matrix (in systematic form) of a linear code $\mathscr{C} \subseteq \mathbb{F}_q^n$ of length n and dimension k. Let $\mathscr{D} \subseteq \mathbb{F}_q^n$ be the dual code of $\mathscr{C}$. Then, we have:*

$$\dim\big(\mathtt{Ker}(\mathscr{D}^2)\big) = \dim\big(\mathtt{Ker}(\mathcal{L}_{\mathbf{P}})\big) \text{ and } \dim\big(\mathtt{Ker}(\mathscr{C}^2)\big) = \dim\big(\mathtt{Ker}(\mathcal{L}_{\mathbf{P}^T})\big).$$

Now, [76] provides rather precise results about the dimension of square random codes. From these results, we can deduce that when the entries of the matrix $\mathbf{P} \in \mathcal{M}_{k\times(n-k)}(\mathbb{F}_q)$ are drawn independently from the uniform distribution over $\mathbb{F}_q$ then $\mathtt{rank}(\mathcal{L}_{\mathbf{P}})$ is equal with very high probability to $\min\left(k, N_{\mathcal{L}_{\mathbf{P}}} = \binom{mt}{2}\right)$. In contrast, it appeared that the value of $D_{\mathcal{L}_{\mathbf{P}}}$ for an alternant or Goppa code is much bigger than D_{random} even when $k \geqslant N$. It also depends on whether or not the code with generator matrix $\mathbf{G}$ is chosen as a (generic) alternant code or as a Goppa code. Although such rank defect is an obstacle to break the McEliece cryptosystem, it can be used to distinguish the public generator of a structured code from a random code. To do so, we proved [122] that the dimension of $\mathtt{Ker}(\mathcal{L}_{\mathbf{P}})$ in the case of an alternant code is predictable:

Theorem 4.3.2 (Alternant Case)**.** *Let $D_{alternant}$ be the dimension of $\mathtt{Ker}(\mathcal{L}_{\mathbf{P}})$ as $\mathbb{F}_q$-vector space when the generator matrix $\mathbf{G} = (\mathbf{I}_k \mid \mathbf{P})$ – where $\mathbf{P} = (p_{ij}) \in \mathcal{M}_{k\times(n-k)}(\mathbb{F}_q)$ – is an alternant code of order t and length n. As long as $N_{\mathcal{L}_{\mathbf{P}}} - D_{alternant} < k$, $D_{alternant}$ is not smaller than:*

$$T_{alternant} = \frac{1}{2}m(t-1)\left((2e+1)t - 2\frac{q^{e+1}-1}{q-1}\right), \text{ with } e = \Big\lfloor \log_q(t-1)\Big\rfloor.$$

In the case of Goppa codes:

Theorem 4.3.3 (Goppa Case)**.** *Let D_{Goppa} be the dimension of $\mathtt{Ker}(\mathcal{L}_{\mathbf{P}})$ as $\mathbb{F}_q$-vector space when the generator matrix $\mathbf{G} = (\mathbf{I}_k \mid \mathbf{P})$ – where $\mathbf{P} = (p_{ij}) \in \mathcal{M}_{k\times(n-k)}(\mathbb{F}_q)$ – is a Goppa code of order t and length n. As long as $N_{\mathcal{L}_{\mathbf{P}}} - D_{Goppa} < k$ then D_{Goppa} is not smaller than*

$$T_{Goppa} = \frac{1}{2}m(t-1)(t-2), \text{ when } r < q-1,$$

$$T_{Goppa} = \frac{1}{2}mt\Big((2e+1)t - 2(q-1)q^{e-1} - 1\Big), \text{ when } r \geqslant q-1,$$

with e being the unique integer such that $(q-1)^2q^{e-2} < t \leqslant (q-1)^2q^{e-1}$.

Extensive experimental results from [122] seem to suggest that $D_{\text{alternant}}$ (resp. D_{Goppa}) is exactly equal to $T_{\text{alternant}}$ (resp. T_{Goppa}). We reproduce in Section 4.3.1 some experiments. In fact, the rank defect can be explained rather easily by the very structure of $\mathsf{elimA}_{\mathbf{X}',\mathbf{Y}'}$ and $\mathcal{L}_{\mathbf{P}}$. By construction of these systems, it is easy to see that $\mathbf{Z}^{a,b,c,d} = (Z_{jj'}^{a,b,c,d}) \in \mathbb{F}_q^{N_{\mathcal{L}_{\mathbf{P}}}}$ with :

$$Z_{jj'}^{a,b,c,d} = \left(X_j^a X_{j'}^b + X_j^b X_{j'}^a + X_j^c X_{j'}^d + X_j^d X_{j'}^c\right) Y_j Y_{j'},$$

is a solution of the linear system $\mathcal{L}_{\mathbf{P}}$ for all $(a,b,c,d) \in J_t$. We can exhibit further elements of $\mathtt{Ker}(\mathcal{L}_{\mathbf{P}})$. To this end, we use the automorphisms $x \longmapsto x^{q^\ell}$ where ℓ is in $\{0,\ldots,m-1\}$. Indeed, we can also consider the identity $(Y_i X_i^a)^{q^{\ell'}} (Y_i X_i^b)^{q^\ell} = (Y_i X_i^c)^{q^{\ell'}} (Y_i X_i^d)^{q^\ell}$ for all a,b,c,d,ℓ and ℓ' such that:

$$aq^{\ell'} + bq^\ell = cq^{\ell'} + dq^\ell.$$

Theorem 4.3.2 is obtained by determining the number of linearly independent solutions induced by such identities.

4.3.1 Experimental Results

Let $\mathbf{P} = (p_{ij}) \in \mathcal{M}_{k\times(n-k)}(\mathbb{F}_q)$. We present here some experimental results on the dimension of $\mathtt{Ker}(\mathcal{L}_{\mathbf{P}})$ (more experiments can be found in [122]). First, we consider the case where the entries of $\mathbf{P}$ are drawn independently from the uniform distribution over $\mathbb{F}_q$. In this case, we denote by D_{random} the dimension of $\mathtt{Ker}(\mathcal{L}_{\mathbf{P}})$. When $\mathbf{P}$ is chosen as (the systematic part of) the generator matrix of a random alternant code (resp. Goppa code) of degree t, we denote it by $D_{\text{alternant}}$ (*resp.* D_{Goppa}). In our probabilistic model, a random alternant code is obtained by picking uniformly and independently at random two vectors $\mathbf{x} = (x_1,\ldots,x_n)$ and $\mathbf{y} = (y_1,\ldots,y_n)$ from $\mathbb{F}_{q^m}^n$ such that the x_i's are all different and the y_i's are all nonzero (for Goppa codes, we sample randomly a support $\mathbf{x} \in \mathbb{F}_{q^m}^n$ and a Goppa polynomial $\Gamma(z) \in \mathbb{F}_{q^m}[z]$ of degree t). We reproduce in Table 4.2 experimental results for $q = 2, m = 14$ and $t \in \{3,\ldots,27\}$. We can observe that the formulas derived in Theorems 4.3.3 and 4.3.2 are perfectly matching the experiments. We then also remark that $D_{\text{alternant}}$ or D_{Goppa} differ from D_{random} but only for certain values of t. We can distinguish a random code from an Alternant code for $t < 16$. For a Goppa code, we can distinguish for $t < 26$. This suggests that we can solve efficiently the $\mathtt{GD}$ problem subject to a condition on the parameters.

4.3.2 Analysis of the Distinguisher

The existence of a distinguisher for the specific case of alternant and Goppa codes is not valid for all t and m. We study the validity in function of the order t. The critical order t_{crit} corresponds to the smallest value of t for which T_{random} becomes bigger than $T_{\text{alternant}}$, i.e. $t_{\text{crit}} = \min\left\{t > 0 \;\middle|\; T_{\text{random}} \geq T_{\text{alternant}}\right\}$. The critical rate is then defined as $R_{\text{crit}} = \dfrac{n - t_{\text{crit}} m}{n} = 1 - \dfrac{t_{\text{crit}} m}{n}$. When the length n of the code goes to infinity an asymptotic formula can be derived for the smallest rate R_{crit} (resp. smaller order t_{crit}) allowing to distinguish a random code from an alternant code. Asymptotically, there is no difference between alternant and (binary) Goppa codes so that the result below is also valid for (binary) Goppa codes.

Theorem 4.3.4. *Let $n = q^m$ with $q \in O(1)$. Then, when m tends to infinity it holds that: $t_{crit} = \sqrt{\frac{2\,q^m \log_2 q}{m \log_2 m}}\big(1 + o(1)\big)$ and $R_{crit} = 1 - \sqrt{\frac{2\,m \log_2 q}{q^m \log_2 m}}\Big(1 + o(1)\Big)$.*

Table 4.2: Experimental results with $q = 2$ and $m = 14$.

t	6	7	8	9	10	11	12	13	14	15	16
$N_{\mathcal{L}_{\mathbf{P}}}$	3486	4753	6216	7875	9730	11781	14028	16471	19110	21945	24976
k	16300	16286	16272	16258	16244	16230	16216	16202	16188	16174	16160
D_{random}	0	0	0	0	0	0	0	269	2922	5771	**8816**
$D_{\text{alternant}}$	560	882	1274	1848	2520	3290	4158	5124	6188	7350	**8816**
$T_{\text{alternant}}$	560	882	1274	1848	2520	3290	4158	5124	6188	7350	8610
D_{Goppa}	1554	2254	3080	4158	5390	6776	8316	10010	11858	13860	16016
T_{Goppa}	1554	2254	3080	4158	5390	6776	8316	10010	11858	13860	16016

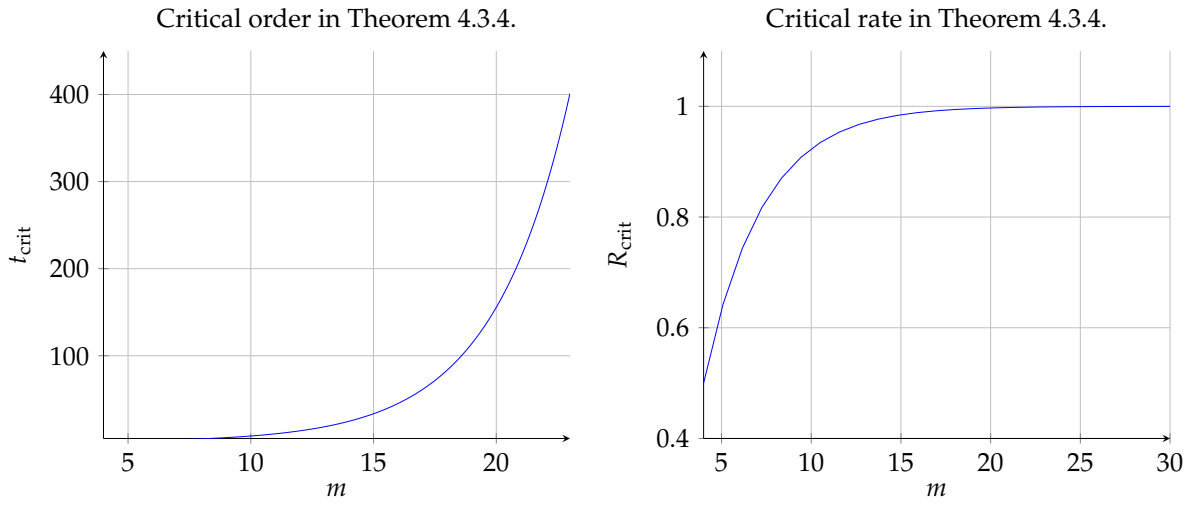
t	17	18	19	20	21	22	23	24	25	26	27
$N_{\mathcal{L}_{\mathbf{P}}}$	28203	31626	35245	39060	43071	47278	51681	56280	61075	66066	71253
k	16146	16132	16118	16104	16090	16076	16062	16048	16034	16020	16006
D_{random}	12057	15494	19127	22956	26981	31202	35619	40232	45041	50046	55247
$D_{\text{alternant}}$	12057	15494	19127	22956	26981	31202	35619	40232	45041	**50046**	55247
$T_{\text{alternant}}$	10192	11900	13734	15694	17780	19992	22330	24794	27384	30100	32942
D_{Goppa}	18564	21294	24206	27300	30576	34034	37674	41496	45500	**50046**	55247
T_{Goppa}	18564	21294	24206	27300	30576	34034	37674	41496	45500	49686	54054

In Table 4.3, we have computed the value of $\left\lceil \sqrt{\frac{2q^m \log q}{m \log_2 m}} \right\rceil$ for several m (q is equal to 2). This shows that our approximation is rather close to $t_{\max}$ computed in practice even for small values of m.

Table 4.3: Values of $t_{\max}$ and t_{crit} for a binary Goppa code of length $n = 2^m$.

m	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23
$t_{\max}$	5	8	8	11	16	20	26	34	47	62	85	114	157	213	290	400
$\lceil t_{\text{crit}} \rceil$	5	6	8	11	14	19	25	34	46	62	84	114	156	214	293	402

It is clear from 4.3.4, that the distinguisher is relevant for codes that have a rate $\frac{n-mt}{n}$ very close to one. Although high-rate codes can be used for public-key encryption ($m = 13$ and $t = 19$ is a parameter that has been proposed which provides 90-bit security for McEliece public-key), such codes are mainly encountered in CFS (Section 4.1). Typically, all the parameters proposed in [87] for CFS can be distinguished. As a side remark, we mention that the hardness of GD has also been investigated in the quantum setting. In [110], the authors show that a reduction of GD to a hidden subgroup problem yields negligible information. As a consequence, it rules out the direct analogue of a quantum attack using the so-called Quantum Fourier Sampling (QFS) which breaks number theoretic problems [258]. More exactly, [110] shows that QFS has a negligible advantage against GD when the rate is $\geq R_{\text{QFS}} = 1 - \frac{\log_q(n)^{3/2}}{\sqrt{5n}} = 1 - \frac{m^{3/2}}{\sqrt{5} \cdot q^{m/2}}$. At first glance, our results could seem as somewhat contradictory with [110]. But, this illustrates that applying a quantum algorithm in black-box remains in this context less powerful than unveiling the structure of the GD problem.



4.4 Final Remarks

In this part, we have presented a rather wide variety of algebraic techniques related to key-recovery against McEliece cryptosystems. Before the introduction of algebraic cryptanalysis in code-based cryptography [133], the only technique for key-recovery was essentially a partial exhaustive search on the secret-key [203]. As illustrated by Remark 5, a direct solving of the algebraic systems modeling key-recovery in McEliece for real parameters seems unlikely regarding the state-of-the art in Gröbner basis computation. However, the algebraic approach turns to be very efficient as soon as the alternant or Goppa codes have an additional structure such as compact variants as discussed in Section 4.2.2 or even a weaker structure such as codes with high rates (Section 4.3). We have also exploited another structure in [141]. This result has not been presented here, but we can improve the modelings when the Goppa code is constructed from a Goppa polynomial of the form $\Gamma = f \cdot g^p$ [141]. In addition, [141] also presents a refined strategy in the solving step.

Our results summarized in Section 4.2.2 give a clear contribution to the question about whether the compactness of the public-keys affects the security of McEliece. We can precisely quantify the security loss regarding key-recovery. It is not clear if there is a reasonable trade-off possible between security and efficiency for such compact variants.

Regarding the distinguisher, the results presented in Section 4.3 shows that the GD assumption must be avoided for codes with high rates. Typically, the security proof for CFS based on GD from [98], [99] is completely meaningless. In general, we would advocate to no longer use the GD assumption in provable security. Even in the case that the dimension of $\mathtt{Ker}(\mathcal{L}_{\mathbf{P}})$ is indistinguishable from random code or a McEliece public-key, we can see that $\mathtt{Ker}(\mathcal{L}_{\mathbf{P}})$ has a very specific and explicit form. Of course, it is a main open question to use such structure to actually attack McEliece with binary Goppa codes.

A distinguishing technique has been used in a series of papers, e.g. [89]–[92] to attack various variants of McEliece. The starting point of their attacks is to use unusual rank defect of the square of the public matrix. According to Proposition 4.3.1, it is equivalent to distinguish from the square code or from $\mathtt{Ker}(\mathcal{L}_{\mathbf{P}})$. So, all the attacks proposed in [89]–[92] could be formulated with our algebraic distinguisher. Conversely, we could only use the square code as a distinguisher in Section 4.3. These give two interesting directions to understand McEliece which remains today immune against all known attacks; including the square/algebraic dis-

tinguisher.

BIBLIOGRAPHY

[1] S. A. Abramov, E. V. Zima, and X. Gao, Eds., *Proceedings of the ACM on International Symposium on Symbolic and Algebraic Computation, ISSAC 2016, Waterloo, ON, Canada, July 19-22, 2016*, ACM, 2016. [Online]. Available: http://dl.acm.org/citation.cfm?id=2930889.

[2] A. Akavia, S. Goldwasser, and V. Vaikuntanathan, "Simultaneous hardcore bits and cryptography against memory attacks", in *TCC 2009*, O. Reingold, Ed., ser. LNCS, vol. 5444, Springer, Heidelberg, Mar. 2009, pp. 474–495.

[3] M. R. Albrecht, C. Cid, J.-C. Faugère, R. Fitzpatrick, and L. Perret, "Algebraic algorithms for LWE problems", *IACR Cryptology ePrint Archive*, vol. 2014, p. 1018, 2014. [Online]. Available: http://eprint.iacr.org/2014/1018.

[4] M. Albrecht, C. Cid, T. Dulien, J.-C. Faugère, and L. Perret, "Algebraic precomputations in differential cryptanalysis", in *Tools'10: Proceedings of the Workshop on Tools for Cryptanalysis 2010*, London (UK): Ecrypt II, Jun. 2010, pp. 1–14. [Online]. Available: http://www-salsa.lip6.fr/~jcf/Papers/Tools2010a.pdf.

[5] ——, "Algebraic precomputations in differential cryptanalysis", in *Information Security and Cryptology: 6th International Conference, Inscrypt 2010, Revised Selected Papers*, M. Yung and X. Lai, Eds., vol. 6584, Shanghai, China: Springer-Verlag, Oct. 2010, pp. 1–18. [Online]. Available: http://www-salsa.lip6.fr/~jcf/Papers/INSCRYPT2010.pdf.

[6] M. Albrecht, C. Cid, J.-C. Faugère, R. Fitzpatrick, and L. Perret, "On the complexity of the Arora-Ge algorithm against LWE", in *SCC '12: Proceedings of the 3nd International Conference on Symbolic Computation and Cryptography*, Castro-Urdiales (Spain), Jul. 2012, pp. 93–99.

[7] ——, "On the complexity of the BKW algorithm on LWE", *Designs, Codes and Cryptography*, vol. 74, no. 2, pp. 325–354, Jul. 2015. [Online]. Available: http://hal.inria.fr/hal-00921517.

[8] M. Albrecht, C. Cid, J.-C. Faugère, R. Fitzpatrick, and L. Perret, "On the complexity of BKW algorithm against LWE", in *SCC'12: Proceedings of the 3nd International Conference on Symbolic Computation and Cryptography*, Castro-Urdiales (Spain), Jul. 2012, pp. 100–107.

[9] M. Albrecht, J.-C. Faugère, P. Farshim, and L. Perret, "Polly cracker, revisited", in *Advances in Cryptology Asiacrypt 2011*, D. Lee and X. Wang, Eds., ser. Lecture Notes in Computer Science, vol. 7073, Seoul, Korea: Springer Berlin / Heidelberg, 2011, pp. 179–196. [Online]. Available: `http://www-salsa.lip6.fr/~jcf/Papers/Asia2011.pdf`.

[10] M. Albrecht, J.-C. Faugère, R. Fitzpatrick, and L. Perret, "Lazy modulus switching for the BKW algorithm on LWE", in *Public-Key Cryptography PKC 2014*, H. Krawczyk, Ed., ser. Lecture Notes in Computer Science, vol. 8383, Buenos Aires, Argentina: Springer Berlin Heidelberg, Mar. 2014, pp. 429–445. [Online]. Available: `http://hal.inria.fr/hal-00925187`.

[11] M. Albrecht, J.-C. Faugère, R. Fitzpatrick, L. Perret, Y. Todo, and K. Xagawa, "Practical cryptanalysis of a public-key encryption scheme based on new multivariate quadratic assumptions", in *Public-Key Cryptography PKC 2014*, H. Krawczyk, Ed., ser. Lecture Notes in Computer Science, vol. 8383, Buenos Aires, Argentina: Springer Berlin Heidelberg, Mar. 2014, pp. 446–464. [Online]. Available: `http://hal.inria.fr/hal-00932382`.

[12] M. R. Albrecht, J. Faugère, P. Farshim, G. Herold, and L. Perret, "Polly cracker, revisited", *Des. Codes Cryptography*, vol. 79, no. 2, pp. 261–302, 2016. [Online]. Available: `http://dx.doi.org/10.1007/s10623-015-0048-8`.

[13] M. R. Albrecht, R. Player, and S. Scott, "On the concrete hardness of learning with errors", *J. Mathematical Cryptology*, vol. 9, no. 3, pp. 169–203, 2015. [Online]. Available: `http://www.degruyter.com/view/j/jmc.2015.9.issue-3/jmc-2015-0016/jmc-2015-0016.xml`.

[14] D. J. Anick, "Thin algebras of embedding dimension three.", *J. Algebra*, vol. 100 (1), pp. 235–259, 1986.

[15] F. Armknecht, D. Augot, L. Perret, and A. Sadeghi, "On constructing homomorphic encryption schemes from coding theory", in *Cryptography and Coding - 13th IMA International Conference, IMACC 2011, Oxford, UK, December 12-15, 2011. Proceedings*, L. Chen, Ed., ser. Lecture Notes in Computer Science, vol. 7089, Springer, 2011, pp. 23–40. [Online]. Available: `http://dx.doi.org/10.1007/978-3-642-25516-8_3`.

[16] S. Arora and R. Ge, "New algorithms for learning in presence of errors", in *ICALP 2011, Part I*, L. Aceto, M. Henzinger, and J. Sgall, Eds., ser. LNCS, vol. 6755, Springer, Heidelberg, Jul. 2011, pp. 403–415.

[17] G. Ars, J.-C. Faugère, H. Imai, M. Kawazoe, and M. Sugita, "Comparison between XL and Gröbner basis algorithms", in *ASIACRYPT 2004*, P. J. Lee, Ed., ser. LNCS, vol. 3329, Springer, Heidelberg, Dec. 2004, pp. 338–353.

[18] C. Boyd, Ed., *ASIACRYPT 2001*, vol. 2248, ser. LNCS, Springer, Heidelberg, Dec. 2001.

[19] P. J. Lee, Ed., *ASIACRYPT 2004*, vol. 3329, ser. LNCS, Springer, Heidelberg, Dec. 2004.

[20] M. Matsui, Ed., *ASIACRYPT 2009*, vol. 5912, ser. LNCS, Springer, Heidelberg, Dec. 2009.

[21] D. Augot, L. Batina, D. J. Bernstein, J. Bos, W. C. Johannes Buchmann, O. Dunkelman, T. Güneysu, S. Gueron, A. Hülsing, T. Lange, M. S. E. Mohamed, C. Rechberger, P. Schwabe, N. Sendrier, F. Vercauteren, and B.-Y. Yang, *Initial recommendations of long-term secure post-quantum systems*.

[22] D. Augot, M. Finiasz, and N. Sendrier, "A family of fast syndrome based cryptographic hash functions", in *Progress in Cryptology - Mycrypt 2005, First International Conference on Cryptology in Malaysia, Kuala Lumpur, Malaysia, September 28-30, 2005, Proceedings*, E. Dawson and S. Vaudenay, Eds., ser. Lecture Notes in Computer Science, vol. 3715, Springer, 2005, pp. 64–83. [Online]. Available: http://dx.doi.org/10.1007/11554868_6.

[23] M. B. B. Monagan, K. O. O. Geddes, K. M. Heal, G. Labahn, S. M. M. Vorkoetter, J. McCarron, and P. DeMarco, *Maple 10 Programming Guide*. Waterloo ON, Canada: Maplesoft, 2005.

[24] G. V. Bard, N. Courtois, and C. Jefferson., "Efficient methods for conversion and solution of sparse systems of low-degree multivariate polynomials over GF(2) via sat-solvers", *IACR Cryptology ePrint Archive*, vol. 2007, p. 24, 2007. [Online]. Available: http://eprint.iacr.org/2007/024.

[25] M. Bardet, "Étude des systèmes algébriques surdéterminés. applications aux codes correcteurs et à la cryptographie", PhD thesis, Université de Paris VI, 2004.

[26] M. Bardet, J.-C. Faugère, and B. Salvy, "Complexity study of Gröbner basis computation", INRIA, Tech. Rep., 2002, http://www.inria.fr/rrrt/rr-5049.html.

[27] ——, "On the complexity of Gröbner basis computation of semi-regular overdetermined algebraic equations", in *International Conference on Polynomial System Solving – ICPSS*, 2004, pp. 71–75.

[28] ——, "On the complexity of the F5 gröbner basis algorithm", *Journal of Symbolic Computation*, pp. 1–24, Sep. 2014, 24 pages. [Online]. Available: http://hal.inria.fr/hal-00915522.

[29] M. Bardet, J.-C. Faugère, B. Salvy, and B.-Y. Yang, "Asymptotic behaviour of the degree of regularity of semi-regular polynomial systems", in *The Effective Methods in Algebraic Geometry Conference – MEGA 2005*, 2005, pp. 1–14.

[30] M. Bardet, J.-C. Faugre, B. Salvy, and P.-J. Spaenlehauer, "On the complexity of solving quadratic boolean systems", *Journal of Complexity*, vol. 29, no. 1, pp. 53–75, Feb. 2013. [Online]. Available: http://hal.inria.fr/hal-00655745.

[31] B. Barkee, D. C. Can, J. Ecks, T. Moriarty, and R. F. Ree, "Why you cannot even hope to use Gröbner bases in Public Key Cryptography: An open letter to a scientist who failed and a challenge to those who have not yet failed", *Journal of Symbolic Computations*, vol. 18, no. 6, pp. 497–501, 1994.

[32] P. S. L. M. Barreto, P.-L. Cayrel, R. Misoczki, and R. Niebuhr, "Quasi-dyadic CFS signatures", in *Inscrypt*, 2010, pp. 336–349.

[33] P. S. L. M. Barreto, R. Lindner, and R. Misoczki, "Monoidic codes in cryptography", in *PQCrypto*, B. Yang, Ed., ser. Lecture Notes in Computer Science, vol. 7071, Springer, 2011, pp. 179–199. [Online]. Available: http://dx.doi.org/10.1007/978-3-642-25405-5.

[34] P. S. L. M. Barreto, R. Misoczki, and M. A. S. Jr., "One-time signature scheme from syndrome decoding over generic error-correcting codes", *Journal of Systems and Software*, vol. 84, no. 2, pp. 198–204, 2011. [Online]. Available: http://dx.doi.org/10.1016/j.jss.2010.09.016.

[35] C. H. Bennett, E. Bernstein, G. Brassard, and U. V. Vazirani, "Strengths and weaknesses of quantum computing", *SIAM J. Comput.*, vol. 26, no. 5, pp. 1510–1523, 1997. [Online]. Available: `http://dx.doi.org/10.1137/S0097539796300933`.

[36] C. H. Bennett and G. Brassard, "Quantum cryptography: Public key distribution and coin tossing", *Theor. Comput. Sci.*, vol. 560, pp. 7–11, 2014. [Online]. Available: `http://dx.doi.org/10.1016/j.tcs.2014.05.025`.

[37] C. Berbain, H. Gilbert, and J. Patarin, "Quad: A practical stream cipher with provable security", in *EUROCRYPT 2006*, S. Vaudenay, Ed., ser. LNCS, vol. 4004, Springer, Heidelberg, May 2006, pp. 109–128.

[38] ——, "QUAD: A multivariate stream cipher with provable security", *J. Symb. Comput.*, vol. 44, no. 12, pp. 1703–1723, 2009. [Online]. Available: `http://dx.doi.org/10.1016/j.jsc.2008.10.004`.

[39] T. P. Berger, P. Cayrel, P. Gaborit, and A. Otmani, "Reducing key length of the McEliece cryptosystem", in *Progress in Cryptology - Second International Conference on Cryptology in Africa (AFRICACRYPT 2009)*, B. Preneel, Ed., ser. Lecture Notes in Computer Science, vol. 5580, Gammarth, Tunisia, Jun. 2009, pp. 77–97.

[40] E. Berlekamp, R. McEliece, and H. van Tilborg, "On the inherent intractability of certain coding problems", *IEEE Transactions on Information Theory*, vol. 24, no. 3, pp. 384–386, May 1978.

[41] D. J. Bernstein, T. Lange, and C. Peters, "Attacking and defending the McEliece cryptosystem", in *PQCrypto*, ser. LNCS, vol. 5299, 2008, pp. 31–46.

[42] D. J. Bernstein, "Grover vs. Mceliece", in *Post-Quantum Cryptography, Third International Workshop, PQCrypto 2010, Darmstadt, Germany, May 25-28, 2010. Proceedings*, N. Sendrier, Ed., ser. Lecture Notes in Computer Science, vol. 6061, Springer, 2010, pp. 73–80. [Online]. Available: `http://dx.doi.org/10.1007/978-3-642-12929-2_6`.

[43] D. J. Bernstein, J. Buchmann, and E. Dahmen, Eds., *Post-quantum cryptography*, ser. Mathematics and Statistics Springer-11649; ZDB-2-SMA. Berlin, Heidelberg: Springer Berlin Heidelberg, 2009. [Online]. Available: `http://opac.inria.fr/record=b1128738`.

[44] D. J. Bernstein, T. Lange, C. Peters, and P. Schwabe, "Really fast syndrome-based hashing", in *Progress in Cryptology - AFRICACRYPT 2011 - 4th International Conference on Cryptology in Africa, Dakar, Senegal, July 5-7, 2011. Proceedings*, A. Nitaj and D. Pointcheval, Eds., ser. Lecture Notes in Computer Science, vol. 6737, Springer, 2011, pp. 134–152. [Online]. Available: `http://dx.doi.org/10.1007/978-3-642-21969-6_9`.

[45] J. Berthomieu, J.-C. Faugère, and L. Perret, "Polynomial-time algorithms for quadratic isomorphism of polynomials: the regular case", *Journal of Complexity*, no. 1–39, p. 39, 2015. [Online]. Available: `https://hal.inria.fr/hal-00846041`.

[46] L. Bettale, J.-C. Faugère, and L. Perret, "Cryptanalysis of the TRMS cryptosystem of PKC'05", in *AfricaCrypt 2008*, S. Vaudenay, Ed., ser. Lecture Notes in Computer Science, vol. 5023, Casablanca, Morocco: Springer, 2008, pp. 143–155. [Online]. Available: `http://www-salsa.lip6.fr/~jcf/Papers/AFRICA2008.pdf`.

[47] L. Bettale, J.-C. Faugère, and L. Perret, "Security analysis of multivariate polynomials for hashing", in *Information Security and Cryptology: 4th International Conference, Inscrypt 2008, Revised Selected Papers*, M. Yung, D. Lin, and P. Liu, Eds., vol. 5487, Beijing, China: Springer-Verlag, Dec. 2009, pp. 115–124. [Online]. Available: `http://www-salsa.lip6.fr/~jcf/Papers/INSCRYPT2008.pdf`.

[48] ——, "Hybrid approach : a tool for multivariate cryptography", in *Tools'10: Proceedings of the Workshop on Tools for Cryptanalysis 2010*, London (UK): Ecrypt II, Jun. 2010, pp. 1–2. [Online]. Available: `http://www-salsa.lip6.fr/~jcf/Papers/Tools2010b.pdf`.

[49] ——, "Hybrid approach for solving multivariate systems over finite fields", *Journal of Mathematical Cryptology*, vol. 3, no. 3, pp. 177–197, 2010. [Online]. Available: `http://www-salsa.lip6.fr/~jcf/Papers/JMC2.pdf`.

[50] ——, "Cryptanalysis of multivariate and odd-characteristic HFE variants", in *Public Key Cryptography - PKC 2011*, D. C. et al., Ed., ser. Lecture Notes in Computer Science, vol. 6571, Taormina, Italy: Springer-Verlag, 2011, pp. 441–458. [Online]. Available: `http://www-salsa.lip6.fr/~jcf/Papers/pkc2011a.pdf`.

[51] ——, "Cryptanalysis of HFE, Multi-HFE and variants for odd and even characteristic", *Designs, Codes and Cryptography*, vol. 69, no. 1, pp. 1–52, 2013. [Online]. Available: `http://hal.inria.fr/hal-00776072`.

[52] L. Bettale, J.-C. Faugère, and L. Perret, "Solving polynomial systems over finite fields: improved analysis of the hybrid approach", in *Proceedings of the 37th International Symposium on Symbolic and Algebraic Computation*, ser. ISSAC '12, Grenoble, France: ACM, 2012, pp. 67–74. [Online]. Available: `http://www-polsys.lip6.fr/~jcf/Papers/FBP12.pdf`.

[53] O. Billet and H. Gilbert, "A traceable block cipher", in *ASIACRYPT 2003*, C.-S. Laih, Ed., ser. LNCS, vol. 2894, Springer, Heidelberg, Nov. 2003, pp. 331–346.

[54] ——, "Cryptanalysis of rainbow", in *SCN 06*, R. D. Prisco and M. Yung, Eds., ser. LNCS, vol. 4116, Springer, Heidelberg, Sep. 2006, pp. 336–347.

[55] O. Billet, J. Patarin, and Y. Seurin, "Analysis of Intermediate Field Systems", in *SCC 2008*, 2008.

[56] O. Billet, M. J. B. Robshaw, and T. Peyrin, "On building hash functions from multivariate quadratic equations", in *ACISP 07*, J. Pieprzyk, H. Ghodosi, and E. Dawson, Eds., ser. LNCS, vol. 4586, Springer, Heidelberg, Jul. 2007, pp. 82–95.

[57] A. Blum, A. Kalai, and H. Wasserman, "Noise-tolerant learning, the parity problem, and the statistical query model", in *32nd ACM STOC*, ACM Press, May 2000, pp. 435–440.

[58] J. W. Bos, C. Costello, L. Ducas, I. Mironov, M. Naehrig, V. Nikolaenko, A. Raghunathan, and D. Stebila, "Frodo: Take off the ring! practical, quantum-secure key exchange from LWE", in *Proceedings of the 2016 ACM SIGSAC Conference on Computer and Communications Security, Vienna, Austria, October 24-28, 2016*, E. R. Weippl, S. Katzenbeisser, C. Kruegel, A. C. Myers, and S. Halevi, Eds., ACM, 2016, pp. 1006–1018. [Online]. Available: `http://doi.acm.org/10.1145/2976749.2978425`.

[59] W. Bosma, J. J. Cannon, and C. Playoust, "The Magma algebra system I: The user language", *Journal of Symbolic Computation*, vol. 24, no. 3-4, pp. 235–265, 1997.

[60] C. Bouillaguet, J.-C. Faugère, P.-A. Fouque, and L. Perret, "Practical cryptanalysis of the identification scheme based on the isomorphism of polynomial with one secret problem", in *Public Key Cryptography - PKC 2011*, D. C. et al., Ed., ser. Lecture Notes in Computer Science, vol. 6571, Taormina, Italy: Springer-Verlag, 2011, pp. 1–12. [Online]. Available: `http://www-salsa.lip6.fr/~jcf/Papers/BFFP11.pdf`.

[61] C. Bouillaguet, H.-C. Chen, C.-M. Cheng, T. Chou, R. Niederhagen, A. Shamir, and B.-Y. Yang, "Fast exhaustive search for polynomial systems in $\mathbb{F}_2$", in *CHES 2010*, S. Mangard and F.-X. Standaert, Eds., ser. LNCS, vol. 6225, Springer, Heidelberg, Aug. 2010, pp. 203–218.

[62] C. Bouillaguet, P. Fouque, A. Joux, and J. Treger, "A family of weak keys in HFE and the corresponding practical key-recovery", *J. Mathematical Cryptology*, vol. 5, no. 3-4, pp. 247–275, 2012. [Online]. Available: `http://dx.doi.org/10.1515/jmc.2011.012`.

[63] B. Boyer, C. Eder, J. Faugère, S. Lachartre, and F. Martani, "GBLA: gröbner basis linear algebra package", in *Proceedings of the ACM on International Symposium on Symbolic and Algebraic Computation, ISSAC 2016, Waterloo, ON, Canada, July 19-22, 2016*, S. A. Abramov, E. V. Zima, and X. Gao, Eds., ACM, 2016, pp. 135–142. [Online]. Available: `http://doi.acm.org/10.1145/2930889.2930914`.

[64] A. Braeken, C. Wolf, and B. Preneel, "A study of the security of unbalanced oil and vinegar signature schemes", in *CT-RSA 2005*, A. Menezes, Ed., ser. LNCS, vol. 3376, Springer, Heidelberg, Feb. 2005, pp. 29–43.

[65] Z. Brakerski, A. Langlois, C. Peikert, O. Regev, and D. Stehlé, "Classical hardness of learning with errors", in *45th ACM STOC*, D. Boneh, T. Roughgarden, and J. Feigenbaum, Eds., ACM Press, Jun. 2013, pp. 575–584.

[66] Z. Brakerski and V. Vaikuntanathan, "Efficient fully homomorphic encryption from (standard) LWE", in *52nd FOCS*, R. Ostrovsky, Ed., IEEE Computer Society Press, Oct. 2011, pp. 97–106.

[67] M. Brickenstein and A. Dreyer, "Polybori: A framework for gröbner-basis computations with boolean polynomials", *J. Symb. Comput.*, vol. 44, no. 9, pp. 1326–1345, 2009. [Online]. Available: `http://dx.doi.org/10.1016/j.jsc.2008.02.017`.

[68] B. Buchberger, "Bruno Buchberger's PhD thesis 1965: An algorithm for finding the basis elements of the residue class ring of a zero dimensional polynomial ideal", *Journal of Symbolic Computation*, vol. 41, no. 3-4, pp. 475–511, 2006.

[69] B. Buchberger, G. E. Collins, R. G. K. Loos, and R. Albrecht, "Computer algebra symbolic and algebraic computation", *SIGSAM Bull.*, vol. 16, no. 4, pp. 5–5, 1982.

[70] J. A. Buchmann, F. Göpfert, R. Player, and T. Wunderer, "On the hardness of LWE with binary error: Revisiting the hybrid lattice-reduction and meet-in-the-middle attack", in *Progress in Cryptology - AFRICACRYPT 2016 - 8th International Conference on Cryptology in Africa, Fes, Morocco, April 13-15, 2016, Proceedings*, D. Pointcheval, A. Nitaj, and T. Rachidi, Eds., ser. Lecture Notes in Computer Science, vol. 9646, Springer, 2016, pp. 24–43. [Online]. Available: `http://dx.doi.org/10.1007/978-3-319-31517-1_2`.

[71] J. Buchmann, E. Dahmen, and M. Szydlo, "Hash-based digital signature schemes", in *Post-Quantum Cryptography*, D. J. Bernstein, J. Buchmann, and E. Dahmen, Eds. Berlin, Heidelberg: Springer Berlin Heidelberg, 2009, pp. 35–93. [Online]. Available: `http://dx.doi.org/10.1007/978-3-540-88702-7_3`.

[72] S. Bulygin, A. Petzoldt, and J. Buchmann, "Towards provable security of the unbalanced oil and vinegar signature scheme under direct attacks", in *INDOCRYPT 2010*, G. Gong and K. C. Gupta, Eds., ser. LNCS, vol. 6498, Springer, Heidelberg, Dec. 2010, pp. 17–32.

[73] W. Buss, G. Frandsen, and J. Shallit, "The computational complexity of some problems of linear algebra.", *Journal of Computer and System Sciences*, 1999.

[74] D. Butin, A. Huelsing, A. Mohaisen, and S.-L. Gazdag, "Xmss: extended hash-based signatures", Internet Engineering Task Force, Internet-Draft draft-irtf-cfrg-xmss-hash-based-signatures-07, Oct. 19, 2016, Work in Progress, 66 pp. [Online]. Available: `https://tools.ietf.org/html/draft-irtf-cfrg-xmss-hash-based-signatures-07`.

[75] A. Canteaut and F. Chabaud, "A new algorithm for finding minimum-weight words in a linear code: Application to McEliece's cryptosystem and to narrow-sense BCH codes of length 511", *IEEE Transactions on Information Theory*, vol. 44, no. 1, pp. 367–378, 1998.

[76] I. Cascudo, R. Cramer, D. Mirandola, and G. Zémor, "Squares of random linear codes", *IEEE Trans. Information Theory*, vol. 61, no. 3, pp. 1159–1173, 2015. [Online]. Available: `http://dx.doi.org/10.1109/TIT.2015.2393251`.

[77] C.-H. O. Chen, M.-S. Chen, J. Ding, F. Werner, and B.-Y. Yang, *Odd-char multivariate Hidden Field Equations*, Cryptology ePrint Archive, `http://eprint.iacr.org/2008/543`, 2008.

[78] L. Chen, S. Jordan, Y.-K. Liu, D. Moody, R. Peralta, R. Perlner, and D. Smith-Tone, "Report on post-quantum cryptography", NIST, Reasearch report NISTIR 8105, 2003. [Online]. Available: `http://csrc.nist.gov/publications/drafts/nistir-8105/nistir_8105_draft.pdf`.

[79] E. Prouff and P. Schaumont, Eds., *CHES 2012*, vol. 7428, ser. LNCS, Springer, Heidelberg, Sep. 2012.

[80] M. Conde Pena, J.-C. Faugère, and L. Perret, "Algebraic cryptanalysis of a quantum money scheme the noise-free case", in *IACR International Conference on Practice and Theory of Public-Key Cryptography (PKC'15)*, Maryland, United States, Mar. 2015. [Online]. Available: `https://hal.inria.fr/hal-01098223`.

[81] I. M. Corbella and R. Pellikaan, "Error-correcting pairs for a public-key cryptosystem", *CoRR*, vol. abs/1205.3647, 2012. [Online]. Available: `http://arxiv.org/abs/1205.3647`.

[82] N. Courtois, "Efficient zero-knowledge authentication based on a linear algebra problem MinRank", in *ASIACRYPT 2001*, C. Boyd, Ed., ser. LNCS, vol. 2248, Springer, Heidelberg, Dec. 2001, pp. 402–421.

[83] ——, "Algebraic attacks over GF(2^k), application to HFE challenge 2 and Sflash-v2", in *PKC 2004*, F. Bao, R. Deng, and J. Zhou, Eds., ser. LNCS, vol. 2947, Springer, Heidelberg, Mar. 2004, pp. 201–217.

[84] N. T. Courtois, L. Goubin, and J. Patarin, *SFLASHv3, a fast asymmetric signature scheme*, Cryptology ePrint Archive, Report 2003/211, `http://eprint.iacr.org/2003/211`, 2003.

[85] N. Courtois and G. V. Bard, "Algebraic cryptanalysis of the data encryption standard", in *Cryptography and Coding, 11th IMA International Conference, Cirencester, UK, December 18-20, 2007, Proceedings*, S. D. Galbraith, Ed., ser. Lecture Notes in Computer Science, vol. 4887, Springer, 2007, pp. 152–169. [Online]. Available: `http://dx.doi.org/10.1007/978-3-540-77272-9_10`.

[86] N. Courtois, M. Daum, and P. Felke, "On the security of HFE, HFEv- and Quartz", in *PKC 2003*, Y. Desmedt, Ed., ser. LNCS, vol. 2567, Springer, Heidelberg, Jan. 2003, pp. 337–350.

[87] N. Courtois, M. Finiasz, and N. Sendrier, "How to achieve a McEliece-based digital signature scheme", in *ASIACRYPT 2001*, C. Boyd, Ed., ser. LNCS, vol. 2248, Springer, Heidelberg, Dec. 2001, pp. 157–174.

[88] N. Courtois, A. Klimov, J. Patarin, and A. Shamir, "Efficient algorithms for solving overdefined systems of multivariate polynomial equations", in *EUROCRYPT 2000*, B. Preneel, Ed., ser. LNCS, vol. 1807, Springer, Heidelberg, May 2000, pp. 392–407.

[89] A. Couvreur, P. Gaborit, V. Gauthier-Umaña, A. Otmani, and J. Tillich, "Distinguisher-based attacks on public-key cryptosystems using reed-solomon codes", *Des. Codes Cryptography*, vol. 73, no. 2, pp. 641–666, 2014. [Online]. Available: `http://dx.doi.org/10.1007/s10623-014-9967-z`.

[90] A. Couvreur, I. Márquez-Corbella, and R. Pellikaan, *A polynomial time attack against algebraic geometry code based public key cryptosystems*, Cryptology ePrint Archive, Report 2014/064, `http://eprint.iacr.org/2014/064`, 2014.

[91] A. Couvreur, A. Otmani, and J.-P. Tillich, "Polynomial time attack on wild McEliece over quadratic extensions", in *EUROCRYPT 2014*, P. Q. Nguyen and E. Oswald, Eds., ser. LNCS, vol. 8441, Springer, Heidelberg, May 2014, pp. 17–39.

[92] A. Couvreur, A. Otmani, J.-P. Tillich, and V. Gauthier-Umaña, "A polynomial-time attack on the BBCRS scheme", in *PKC 2015*, J. Katz, Ed., ser. LNCS, vol. 9020, Springer, Heidelberg, Mar. 2015, pp. 175–193.

[93] D. A. Cox, J. B. Little, and D. O'Shea, *Ideals, Varieties and Algorithms*. Springer Verlag, 2005.

[94] D. Wagner, Ed., *CRYPTO 2008*, vol. 5157, ser. LNCS, Springer, Heidelberg, Aug. 2008.

[95] P. Rogaway, Ed., *CRYPTO 2011*, vol. 6841, ser. LNCS, Springer, Heidelberg, Aug. 2011.

[96] R. Canetti and J. A. Garay, Eds., *CRYPTO 2013, Part I*, vol. 8042, ser. LNCS, Springer, Heidelberg, Aug. 2013.

[97] P. Czypek, S. Heyse, and E. Thomae, "Efficient implementations of MQPKS on constrained devices", in *CHES 2012*, E. Prouff and P. Schaumont, Eds., ser. LNCS, vol. 7428, Springer, Heidelberg, Sep. 2012, pp. 374–389.

[98] L. Dallot, "Towards a concrete security proof of courtois, finiasz and sendrier signature scheme", in *Research in Cryptology, Second Western European Workshop, WEWoRC 2007, Bochum, Germany, July 4-6, 2007, Revised Selected Papers*, S. Lucks, A. Sadeghi, and C. Wolf, Eds., ser. Lecture Notes in Computer Science, vol. 4945, Springer, 2007, pp. 65–77. [Online]. Available: `http://dx.doi.org/10.1007/978-3-540-88353-1_6`.

[99] ——, "Sécurité de protocoles cryptographiques fondés sur les codes correcteurs d'erreurs. (security of cryptographic protocols based on error correcting codes)", PhD thesis, University of Caen Normandy, France, 2010. [Online]. Available: `https://tel.archives-ouvertes.fr/tel-01102440`.

[100] T. Daniels and D. Smith-Tone, "Differential properties of the HFE cryptosystem", in *Post-Quantum Cryptography - 6th International Workshop, PQCrypto 2014, Waterloo, ON, Canada, October 1-3, 2014. Proceedings*, M. Mosca, Ed., ser. Lecture Notes in Computer Science, vol. 8772, Springer, 2014, pp. 59–75. [Online]. Available: `http://dx.doi.org/10.1007/978-3-319-11659-4_4`.

[101] P. Dembowski and T. Ostrom, "Planes of order n with collineation groups of order n^2", *Math. Z.*, vol. 103, no. 3, pp. 239–258, 1968.

[102] R. DeMillo and R. Lipton, "A probabilistic remark on algebraic program testing", *Information Processing Letters*, vol. 7, no. 4, pp. 192–194, 1978.

[103] C. Diem, "The XL-algorithm and a conjecture from commutative algebra", in *ASIACRYPT 2004*, P. J. Lee, Ed., ser. LNCS, vol. 3329, Springer, Heidelberg, Dec. 2004, pp. 323–337.

[104] M. S. E. Din and P. Trebuchet, "Strong bi-homogeneous bézout theorem and its use in effective real algebraic geometry", *CoRR*, vol. abs/cs/0610051, 2006. [Online]. Available: `http://arxiv.org/abs/cs/0610051`.

[105] J. Ding and T. J. Hodges, "Inverting HFE systems is quasi-polynomial for all fields", in *CRYPTO 2011*, P. Rogaway, Ed., ser. LNCS, vol. 6841, Springer, Heidelberg, Aug. 2011, pp. 724–742.

[106] J. Ding and D. Schmidt, "Rainbow, a new multivariable polynomial signature scheme", in *ACNS 05*, J. Ioannidis, A. Keromytis, and M. Yung, Eds., ser. LNCS, vol. 3531, Springer, Heidelberg, Jun. 2005, pp. 164–175.

[107] ——, "Rainbow, a new multivariable polynomial signature scheme", in *Applied Cryptography and Network Security, Third International Conference, ACNS 2005, New York, NY, USA, June 7-10, 2005, Proceedings*, J. Ioannidis, A. D. Keromytis, and M. Yung, Eds., ser. Lecture Notes in Computer Science, vol. 3531, 2005, pp. 164–175. [Online]. Available: `http://dx.doi.org/10.1007/11496137_12`.

[108] J. Ding, D. Schmidt, and F. Werner, "Algebraic attack on HFE revisited", in *Information Security, 11th International Conference, ISC 2008, Taipei, Taiwan, September 15-18, 2008. Proceedings*, T. Wu, C. Lei, V. Rijmen, and D. Lee, Eds., ser. Lecture Notes in Computer Science, vol. 5222, Springer, 2008, pp. 215–227. [Online]. Available: `http://dx.doi.org/10.1007/978-3-540-85886-7_15`.

[109] J. Ding and B.-Y. Yang, "Multivariate public key cryptography", in *Post-Quantum Cryptography*, D. J. Bernstein, J. Buchmann, and E. Dahmen, Eds. Berlin, Heidelberg: Springer Berlin Heidelberg, 2009, pp. 193–241. [Online]. Available: `http://dx.doi.org/10.1007/978-3-540-88702-7_6`.

[110] H. Dinh, C. Moore, and A. Russell, "McEliece and Niederreiter cryptosystems that resist quantum fourier sampling attacks", in *CRYPTO 2011*, P. Rogaway, Ed., ser. LNCS, vol. 6841, Springer, Heidelberg, Aug. 2011, pp. 761–779.

[111] N. Döttling and J. Müller-Quade, "Lossy codes and a new variant of the learning-with-errors problem", in *EUROCRYPT 2013*, T. Johansson and P. Q. Nguyen, Eds., ser. LNCS, vol. 7881, Springer, Heidelberg, May 2013, pp. 18–34.

[112] R. Dowsley, J. Müller-Quade, and A. C. A. Nascimento, "A CCA2 secure public key encryption scheme based on the McEliece assumptions in the standard model", in *CT-RSA*, 2009, pp. 240–251.

[113] V. Dubois, P.-A. Fouque, A. Shamir, and J. Stern, "Practical cryptanalysis of SFLASH", in *CRYPTO 2007*, A. Menezes, Ed., ser. LNCS, vol. 4622, Springer, Heidelberg, Aug. 2007, pp. 1–12.

[114] V. Dubois and N. Gama, "The degree of regularity of HFE systems", in *ASIACRYPT 2010*, M. Abe, Ed., ser. LNCS, vol. 6477, Springer, Heidelberg, Dec. 2010, pp. 557–576.

[115] V. Dubois, L. Granboulan, and J. Stern, "An efficient provable distinguisher for HFE", in *ICALP 2006, Part II*, M. Bugliesi, B. Preneel, V. Sassone, and I. Wegener, Eds., ser. LNCS, vol. 4052, Springer, Heidelberg, Jul. 2006, pp. 156–167.

[116] L. Ducas, A. Durmus, T. Lepoint, and V. Lyubashevsky, "Lattice signatures and bimodal Gaussians", in *CRYPTO 2013, Part I*, R. Canetti and J. A. Garay, Eds., ser. LNCS, vol. 8042, Springer, Heidelberg, Aug. 2013, pp. 40–56.

[117] C. Eder and J.-C. Faugère, "A survey on signature-based algorithms for computing gröbner bases", *Journal of Symbolic Computation*, 2016. [Online]. Available: `http://www.sciencedirect.com/science/article/pii/S0747717116300785`.

[118] N. Eén and N. Sörensson, "An extensible sat-solver", in *Theory and Applications of Satisfiability Testing, 6th International Conference, SAT 2003. Santa Margherita Ligure, Italy, May 5-8, 2003 Selected Revised Papers*, E. Giunchiglia and A. Tacchella, Eds., ser. Lecture Notes in Computer Science, vol. 2919, Springer, 2003, pp. 502–518. [Online]. Available: `http://dx.doi.org/10.1007/978-3-540-24605-3_37`.

[119] C. G. Günther, Ed., *EUROCRYPT'88*, vol. 330, ser. LNCS, Springer, Heidelberg, May 1988.

[120] J.-C. Faugère, "A new efficient algorithm for computing gröbner bases (F4)", *Journal of Pure and Applied Algebra*, vol. 139(1-3), pp. 61–88, 1999.

[121] ——, "A new efficient algorithm for computing gröbner bases without reduction to zero : F5", in *ISSAC'02*, ACM press, 2002, pp. 75–83.

[122] J.-C. Faugère, V. Gauthier-Umana, A. Otmani, L. Perret, and J.-P. Tillich, "A distinguisher for high rate Mceliece cryptosystems", *IEEE Transactions on Information Theory*, vol. 59, no. 10, pp. 6830–6844, Jun. 2013. [Online]. Available: `http://hal.inria.fr/hal-00776068`.

[123] J.-C. Faugère, V. Gauthier-Umana, A. Otmani, L. Perret, and J.-P. Tillich, "A distinguisher for high rate mceliece cryptosystems", in *Information Theory Workshop (ITW), 2011 IEEE*, Paraty, Brazil, Oct. 2011, pp. 282–286. [Online]. Available: `http://www-polsys.lip6.fr/~jcf/Papers/ITW2011.pdf`.

[124] J.-C. Faugère, D. Gligoroski, E. Jensen, R. Odegard, L. Perret, S. Johan Knapskog, and S. Markovski, "MQQ-SIG", in *Trusted Systems - The Third International Conference on Trusted Systems - INTRUST 2011*, M. Y., L. C., and L. Z., Eds., ser. Lecture Notes in Computer Science, vol. 7222, Beijing, China: Springer Verlag, 2012, pp. 184–203.

[125] J.-C. Faugère, D. Gligoroski, L. Perret, S. Samardjiska, and E. Thomae, "A polynomial-time key-recovery attack on MQQ cryptosystems", in *IACR International Conference on Practice and Theory of Public-Key Cryptography (PKC'15)*, Maryland, United States, Mar. 2015. [Online]. Available: `https://hal.inria.fr/hal-01074194`.

[126] J.-C. Faugère, A. Joux, L. Perret, and J. Treger, "Cryptanalysis of the hidden matrix cryptosystem", in *Progress in Cryptology, LATINCRYPT 2010*, M. Abdalla and P. Barreto, Eds., ser. Lecture Notes in Computer Science, vol. 6212, Mexico: Springer Berlin / Heidelberg, 2010, pp. 241–254. [Online]. Available: `http://www-salsa.lip6.fr/~jcf/Papers/LATIN2010.pdf`.

[127] J.-C. Faugère, F. Levy-dit-Vehel, and L. Perret, "Cryptanalysis of MinRank", in *Advances in Cryptology CRYPTO 2008*, D. Wagner, Ed., ser. Lecture Notes in Computer Science, vol. 5157, Santa Barbara, CA, USA: Springer-Verlag, Aug. 2008, pp. 280–296. [Online]. Available: `http://www-salsa.lip6.fr/~jcf/Papers/crypto08.pdf`.

[128] J.-C. Faugère, D. Lin, L. Perret, and T. Wang, "On enumeration of polynomial equivalence classes and their application to mpkc", *Finite Fields and Their Applications*, vol. 18, no. 2, pp. 283–302, 2012. [Online]. Available: `http://hal.inria.fr/hal-00776073`.

[129] J.-C. Faugère, R. Odegard, L. Perret, and D. Gligoroski, "Analysis of the MQQ public key cryptosystem", in *SCC'10: Proceedings of the 2nd International Conference on Symbolic Computation and Cryptography*, London (UK), Jun. 2010, pp. 101–116. [Online]. Available: `http://www-salsa.lip6.fr/~jcf/Papers/SCC2010b.pdf`.

[130] ——, "Analysis of the MQQ public key cryptosystem", in *Ninth International Conference on Cryptology And Network Security (CANS 2010)*, S.-H. Heng, R. N. Wright, and B.-M. Goi, Eds., ser. Security and Cryptology, vol. 6467, Kuala Lumpur (Malaysia): Springer-Verlag, Dec. 2010, pp. 1–14. [Online]. Available: `http://www-salsa.lip6.fr/~jcf/Papers/CANS2010.pdf`.

[131] J.-C. Faugère, A. Otmani, L. Perret, F. De Portzamparc, and J.-P. Tillich, "Structural cryptanalysis of Mceliece schemes with compact keys", *Designs, Codes and Cryptography*, pp. 87–112, Jan. 2016. [Online]. Available: `https://hal.inria.fr/hal-00964265`.

[132] J.-C. Faugère, A. Otmani, L. Perret, and J.-P. Tillich, "A distinguisher for high rate mceliece cryptosystem – extended abstract", in *Yet Another Conference on Cryptography, YACC 2010*, P. Véron, Ed., Porquerolles (France), 2010, pp. 1–4. [Online]. Available: `http://www-salsa.lip6.fr/~jcf/Papers/ARTICLE_YACC2.pdf`.

[133] J.-C. Faugère, A. Otmani, L. Perret, and J.-P. Tillich, "Algebraic cryptanalysis of McEliece variants with compact keys", in *Proceedings of Eurocrypt 2010*, ser. Lecture Notes in Computer Science, vol. 6110, Monaco: Springer Verlag, 2010, pp. 279–298. [Online]. Available: `http://www-salsa.lip6.fr/~jcf/Papers/Eurocrypt2010.pdf`.

[134] J.-C. Faugère, A. Otmani, L. Perret, and J.-P. Tillich, "Algebraic Cryptanalysis of McEliece variants with compact keys – toward a complexity analysis", in *Yet Another Conference on Cryptography, YACC 2010*, P. Véron, Ed., Porquerolles (France), 2010, pp. 1–4. [Online]. Available: `http://www-salsa.lip6.fr/~jcf/Papers/ARTICLE_YACC1.pdf`.

[135] ——, "Algebraic Cryptanalysis of McEliece variants with compact keys – toward a complexity analysis", in *SCC '10: Proceedings of the 2nd International Conference on Symbolic Computation and Cryptography*, London (UK), Jun. 2010, pp. 45–55. [Online]. Available: `http://www-salsa.lip6.fr/~jcf/Papers/SCC2010a.pdf`.

[136] J.-C. Faugère and L. Perret, "Polynomial equivalence problems: algorithmic and theoretical aspects", in *Advances in Cryptology - EUROCRYPT 2006*, S. Vaudenay, Ed., ser. Lecture Notes in Computer Science, vol. 4004, St Petersburg, Russia: Springer Berlin / Heidelberg, 2006, pp. 30–47. [Online]. Available: `http://www-salsa.lip6.fr/~jcf/Papers/FP06b.pdf`.

[137] ——, "High order derivatives and decomposition of multivariate polynomials", in *Second Workshop on Mathematical Cryptology*, Santander (Spain), Oct. 2008, pp. 15–19.

[138] ——, "On the security of UOV", in *First International Conference on Symbolic Computation and Cryptography, SCC 08*, ser. LMIB, Beijing, China, Apr. 2008, pp. 103–109. [Online]. Available: `http://www-salsa.lip6.fr/~jcf/Papers/SCC08a.pdf`.

[139] ——, "An efficient algorithm for decomposing multivariate polynomials and its applications to cryptography", *Journal of Symbolic Computation*, vol. 44, no. 12, pp. 1676–1689, 2009. [Online]. Available: `http://www-salsa.lip6.fr/~jcf/Papers/jsc_FP09.pdf`.

[140] ——, "Algebraic cryptanalysis of Curry and Flurry using correlated messages", in *Information Security and Cryptology: 5th International Conference, Inscrypt 2009, Beijing, China, December, 2009, Revised Selected Papers*, M. Yung and F. Bao, Eds., vol. 6151, Berlin, Heidelberg: Springer-Verlag, 2010, pp. 266–277. [Online]. Available: `http://www-salsa.lip6.fr/~jcf/Papers/INSCRYPT2009.pdf`.

[141] J.-C. Faugère, L. Perret, and F. De Portzamparc, "Algebraic attack against variants of Mceliece with Goppa polynomial of a special form", Anglais, in *Advances in Cryptology Asiacrypt 2014*, Kaohsiung, Tawan, Sep. 2014. [Online]. Available: `http://hal.inria.fr/hal-01064687`.

[142] J.-C. Faugère, L. Perret, F. De Portzamparc, A. Otmani, and J.-P. Tillich, "Structural weakness of compact variants of the Mceliece cryptosystem", in *IEEE International Symposium on Information Theory - ISIT 2014*, Honolulu, United States, Jun. 2014, pp. 1717–1721. [Online]. Available: `https://hal.archives-ouvertes.fr/hal-01096180`.

[143] J.-C. Faugère, L. Perret, C. Petit, and G. Renault, "Improving the complexity of index calculus algorithms in elliptic curves over binary fields", in *Advances in Cryptology EUROCRYPT 2012*, D. Pointcheval and T. Johansson, Eds., ser. Lecture Notes in Computer Science, vol. 7237, Cambridge: Springer Berlin / Heidelberg, 2012, pp. 27–44. [Online]. Available: `http://www-polsys.lip6.fr/~jcf/Papers/euro2012.pdf`.

[144] J.-C. Faugère, L. Perret, and P.-J. Spaenlehauer, "Algebraic-differential cryptanalysis of DES", in *Western European Workshop on Research in Cryptology - WEWoRC 2009*, Graz, Austria, Jul. 2009, pp. 1–5. [Online]. Available: `http://www-salsa.lip6.fr/~jcf/Papers/DESweworc.pdf`.

[145] J.-C. Faugère, J. von zur Gathen, and L. Perret, "Decomposition of generic multivariate polynomials", in *ISSAC '10: Proceedings of the 2010 international symposium on Symbolic and algebraic computation*, ser. ISSAC '10, Munich, Germany: ACM, 2010, pp. 131–137. [Online]. Available: `http://www-salsa.lip6.fr/~jcf/Papers/ISSAC_FGP_2010.pdf`.

[146] J.-C. Faugère, "Algebraic cryptanalysis of HFE using Gröbner bases", INRIA, Reasearch report RR-4738, 2003. [Online]. Available: `http://hal.inria.fr/inria-00071849/PDF/RR-4738.pdf`.

[147] ——, "FGb: a library for computing gröbner bases", in *Mathematical Software - ICMS 2010*, K. Fukuda, J. Hoeven, M. Joswig, and N. Takayama, Eds., ser. Lecture Notes in Computer Science, vol. 6327, Kobe, Japan: Springer Berlin / Heidelberg, Sep. 2010, pp. 84–87.

[148] J. Faugère, M. S. E. Din, and P. Spaenlehauer, "Computing loci of rank defects of linear matrices using gröbner bases and applications to cryptology", in *Symbolic and Algebraic Computation, International Symposium, ISSAC 2010, Munich, Germany, July 25-28, 2010, Proceedings*, W. Koepf, Ed., ACM, 2010, pp. 257–264. [Online]. Available: `http://doi.acm.org/10.1145/1837934.1837984`.

[149] ——, "On the complexity of the generalized MinRank problem", *J. Symb. Comput.*, vol. 55, pp. 30–58, 2013. [Online]. Available: `http://dx.doi.org/10.1016/j.jsc.2013.03.004`.

[150] J. Faugère, M. S. E. Din, and T. Verron, "On the complexity of computing gröbner bases for weighted homogeneous systems", *J. Symb. Comput.*, vol. 76, pp. 107–141, 2016. [Online]. Available: `http://dx.doi.org/10.1016/j.jsc.2015.12.001`.

[151] J.-C. Faugère, M. S. El Din, and P.-J. Spaenlehauer, "Computing loci of rank defects of linear matrices using Gröbner bases and applications to cryptology", in *ISSAC 2010: Proceedings of the 2010 International Symposium on Symbolic and Algebraic Computation, July 25–28, 2010, Munich, Germany*, S. M. Watt, Ed., New York, NY 10036, USA: ACM Press, 2010, pp. 257–264.

[152] J. Faugère, P. Gaudry, L. Huot, and G. Renault, "Sub-cubic change of ordering for gröbner basis: A probabilistic approach", in *International Symposium on Symbolic and Algebraic Computation, ISSAC '14, Kobe, Japan, July 23-25, 2014*, K. Nabeshima, K. Nagasaka, F. Winkler, and Á. Szántó, Eds., ACM, 2014, pp. 170–177. [Online]. Available: `http://doi.acm.org/10.1145/2608628.2608669`.

[153] J.-C. Faugère, P. M. Gianni, D. Lazard, and T. Mora, "Efficient computation of zero-dimensional Gröbner bases by change of ordering", *Journal of Symbolic Computation*, vol. 16, no. 4, pp. 329–344, 1993.

[154] J.-C. Faugère, D. Gligoroski, L. Perret, S. Samardjiska, and E. Thomae, "A polynomial-time key-recovery attack on MQQ cryptosystems", in *PKC 2015*, J. Katz, Ed., ser. LNCS, vol. 9020, Springer, Heidelberg, Mar. 2015, pp. 150–174.

[155] J.-C. Faugère and A. Joux, "Algebraic cryptanalysis of hidden field equation (HFE) cryptosystems using gröbner bases", in *CRYPTO 2003*, D. Boneh, Ed., ser. LNCS, vol. 2729, Springer, Heidelberg, Aug. 2003, pp. 44–60.

[156] J.-C. Faugère, F. Levy-dit-Vehel, and L. Perret, "Cryptanalysis of minrank", in *CRYPTO 2008*, D. Wagner, Ed., ser. LNCS, vol. 5157, Springer, Heidelberg, Aug. 2008, pp. 280–296.

[157] J.-C. Faugère and C. Mou, "Fast algorithm for change of ordering of zero-dimensional Gröbner bases with sparse multiplication matrices", in *ISSAC 2011: Proceedings of the 2011 International Symposium on Symbolic and Algebraic Computation, June 7–11, 2011, San Jose, CA, USA*, É. Schost and I. Z. Emiris, Eds., New York, NY 10036, USA: ACM Press, 2011, pp. 115–122.

[158] J. Faugère, A. Otmani, L. Perret, F. de Portzamparc, and J. Tillich, "Folding alternant and goppa codes with non-trivial automorphism groups", *IEEE Transactions on Information Theory*, vol. 62, no. 1, pp. 184–198, 2016. [Online]. Available: `http://dx.doi.org/10.1109/TIT.2015.2493539`.

[159] J.-C. Faugère and L. Perret, "Cryptanalysis of 2R- schemes", in *Advances in Cryptology - CRYPTO 2006*, C. Dwork, Ed., ser. Lecture Notes in Computer Science, vol. 4117, Santa Barbara, USA: Springer Berlin / Heidelberg, Aug. 2006, pp. 357–372. [Online]. Available: `http://www-salsa.lip6.fr/~jcf/Papers/FP06a.pdf`.

[160] J.-C. Faugère, M. Safey El Din, and P.-J. Spaenlehauer, "Gröbner bases bases of bihomogeneous ideals generated by polynomials of bidegree (1,1): algorithms and complexity", *Journal of Symbolic Computation*, vol. 46, no. 4, pp. 406–437, 2011, Available online 4 November 2010. [Online]. Available: `http://www-polsys.lip6.fr/~jcf/Papers/JSC_FSS10.pdf`.

[161] J. Faugère, P. Spaenlehauer, and J. Svartz, "Sparse Gröbner bases: The unmixed case", in *International Symposium on Symbolic and Algebraic Computation, ISSAC '14, Kobe, Japan, July 23-25, 2014*, K. Nabeshima, K. Nagasaka, F. Winkler, and Á. Szántó, Eds., ACM, 2014, pp. 178–185. [Online]. Available: `http://doi.acm.org/10.1145/2608628.2608663`.

[162] ——, "Computing small certificates of inconsistency of quadratic fewnomial systems", in *Proceedings of the ACM on International Symposium on Symbolic and Algebraic Computation, ISSAC 2016, Waterloo, ON, Canada, July 19-22, 2016*, S. A. Abramov, E. V. Zima, and X. Gao, Eds., ACM, 2016, pp. 223–230. [Online]. Available: `http://doi.acm.org/10.1145/2930889.2930927`.

[163] J. Faugère and J. Svartz, "Gröbner bases of ideals invariant under a commutative group: The non-modular case", in *International Symposium on Symbolic and Algebraic Computation, ISSAC'13, Boston, MA, USA, June 26-29, 2013*, M. Kauers, Ed., ACM, 2013, pp. 347–354. [Online]. Available: `http://doi.acm.org/10.1145/2465506.2465944`.

[164] M. Fellows and N. Koblitz, "Combinatorial cryptosystems galore!", in *Finite Fields: Theory, Applications, and Algorithms*, ser. Contemporary Mathematics, G. L. Mullen and P. J.-S. Shiue, Eds., vol. 168, AMS, 1994, pp. 51–61.

[165] L. D. Feo, D. Jao, and J. Plût, "Towards quantum-resistant cryptosystems from supersingular elliptic curve isogenies", *J. Mathematical Cryptology*, vol. 8, no. 3, pp. 209–247, 2014. [Online]. Available: `http://dx.doi.org/10.1515/jmc-2012-0015`.

[166] A. Fiat and A. Shamir, "How to prove yourself: Practical solutions to identification and signature problems", in *CRYPTO'86*, A. M. Odlyzko, Ed., ser. LNCS, vol. 263, Springer, Heidelberg, Aug. 1987, pp. 186–194.

[167] M. Finiasz and N. Sendrier, "Security bounds for the design of code-based cryptosystems", in *ASIACRYPT 2009*, M. Matsui, Ed., ser. LNCS, vol. 5912, Springer, Heidelberg, Dec. 2009, pp. 88–105.

[168] R. Fröberg, "An inequality for Hilbert series of graded algebras", *Mathematica Scandinavica*, vol. 56, pp. 117–144, 1985.

[169] R. Fröberg and J. Hollman, "Hilbert series for ideals generated by generic forms", *J. Symb. Comput.*, vol. 17, no. 2, pp. 149–157, 1994. [Online]. Available: `http://dx.doi.org/10.1006/jsco.1994.1008`.

[170] P. Gaborit, "Shorter keys for code based cryptography", in *Proceedings of the 2005 International Workshop on Coding and Cryptography (WCC 2005)*, Bergen, Norway, Mar. 2005, pp. 81–91.

[171] S. D. Galbraith, *Space-efficient variants of cryptosystems based on learning with errors*, https://www.math.auck
LWE.pdf, 2012.

[172] F. L. Gall, "Powers of tensors and fast matrix multiplication", in *International Symposium on Symbolic and Algebraic Computation, ISSAC '14, Kobe, Japan, July 23-25, 2014*, K. Nabeshima, K. Nagasaka, F. Winkler, and Á. Szántó, Eds., ACM, 2014, pp. 296–303. [Online]. Available: http://doi.acm.org/10.1145/2608628.2608664.

[173] S. Gao, Y. Guan, and F. Volny IV, "A new incremental algorithm for computing groebner bases", in *Proceedings of the 2010 International Symposium on Symbolic and Algebraic Computation*, ser. ISSAC '10, Munich, Germany: ACM, 2010, pp. 13–19. [Online]. Available: http://doi.acm.org/10.1145/1837934.1837944.

[174] S. Gao, F. V. IV, and M. Wang, "A new framework for computing gröbner bases", *Math. Comput.*, vol. 85, no. 297, 2016. [Online]. Available: http://dx.doi.org/10.1090/mcom/2969.

[175] X. Gao and Z. Huang, "Characteristic set algorithms for equation solving in finite fields", *J. Symb. Comput.*, vol. 47, no. 6, pp. 655–679, 2012. [Online]. Available: http://dx.doi.org/10.1016/j.jsc.2011.12.025.

[176] M. R. Garey and D. S. Johnson, *Computers and Intractability: A Guide to the Theory of NP-Completeness*. W. H. Freeman, 1979.

[177] J. von zur Gathen and J. Gerhard, *Modern Computer Algebra (3. ed)*. Cambridge University Press, 2013.

[178] C. Gentry, "A fully homomorphic encryption scheme", Available at http://crypto.stanford.edu/craig, PhD thesis, Stanford University, 2009.

[179] C. Gentry, S. Halevi, and N. P. Smart, "Homomorphic evaluation of the AES circuit", in *CRYPTO 2012*, R. Safavi-Naini and R. Canetti, Eds., ser. LNCS, vol. 7417, Springer, Heidelberg, Aug. 2012, pp. 850–867.

[180] O. Goldreich and L. A. Levin, "A hard-core predicate for all one-way functions", in *21st ACM STOC*, ACM Press, May 1989, pp. 25–32.

[181] L. Goubin and N. Courtois, "Cryptanalysis of the TTM cryptosystem", in *ASIACRYPT 2000*, T. Okamoto, Ed., ser. LNCS, vol. 1976, Springer, Heidelberg, Dec. 2000, pp. 44–57.

[182] A. Gouget and J. Patarin, "Probabilistic multivariate cryptography", in *Progress in Cryptology - VIETCRYPT 06*, P. Q. Nguyen, Ed., ser. LNCS, vol. 4341, Springer, Heidelberg, Sep. 2006, pp. 1–18.

[183] L. Granboulan, A. Joux, and J. Stern, "Inverting HFE is quasipolynomial", in *CRYPTO 2006*, C. Dwork, Ed., ser. LNCS, vol. 4117, Springer, Heidelberg, Aug. 2006, pp. 345–356.

[184] L. K. Grover, "A fast quantum mechanical algorithm for database search", in *28th ACM STOC*, ACM Press, May 1996, pp. 212–219.

[185] G. Herold, E. Kirshanova, and A. May, *On the asymptotic complexity of solving LWE*, Cryptology ePrint Archive, Report 2015/1222, http://eprint.iacr.org/2015/1222, 2015.

[186] A. Hülsing, "Practical forward secure signatures using minimal security assumptions", PhD thesis, Darmstadt University of Technology, 2013. [Online]. Available: http://d-nb.info/1044187786.

[187] A. Hülsing, J. Rijneveld, S. Samardjiska, and P. Schwabe, "From 5-pass MQ-based identification to MQ-based signatures", *IACR Cryptology ePrint Archive*, vol. 2016, p. 708, 2016. [Online]. Available: http://eprint.iacr.org/2016/708.

[188] D. Jao and L. D. Feo, "Towards quantum-resistant cryptosystems from supersingular elliptic curve isogenies", in *Post-Quantum Cryptography - 4th International Workshop, PQCrypto 2011, Taipei, Taiwan, November 29 - December 2, 2011. Proceedings*, B. Yang, Ed., ser. Lecture Notes in Computer Science, vol. 7071, Springer, 2011, pp. 19–34. [Online]. Available: http://dx.doi.org/10.1007/978-3-642-25405-5_2.

[189] G. Kabatianskii, E. Krouk, and B. Smeets, "A digital signature scheme based on random error-correcting codes", in *Crytography and Coding: 6th IMA International Conference Cirencester, UK, December 17–19, 1997 Proceedings*, M. Darnell, Ed. Berlin, Heidelberg: Springer Berlin Heidelberg, 1997, pp. 161–167. [Online]. Available: http://dx.doi.org/10.1007/BFb0024461.

[190] K. Kalach and R. Safavi-Naini, "An efficient post-quantum one-time signature scheme", in *Selected Areas in Cryptography - SAC 2015 - 22nd International Conference, Sackville, NB, Canada, August 12-14, 2015, Revised Selected Papers*, O. Dunkelman and L. Keliher, Eds., ser. Lecture Notes in Computer Science, vol. 9566, Springer, 2015, pp. 331–351. [Online]. Available: http://dx.doi.org/10.1007/978-3-319-31301-6_20.

[191] M. Kaplan, G. Leurent, A. Leverrier, and M. Naya-Plasencia, "Breaking symmetric cryptosystems using quantum period finding", in *CRYPTO 2016, Part II*, M. Robshaw and J. Katz, Eds., ser. LNCS, vol. 9815, Springer, Heidelberg, Aug. 2016, pp. 207–237.

[192] A. Kipnis, J. Patarin, and L. Goubin, "Unbalanced oil and vinegar signature schemes", in *EUROCRYPT'99*, J. Stern, Ed., ser. LNCS, vol. 1592, Springer, Heidelberg, May 1999, pp. 206–222.

[193] A. Kipnis and A. Shamir, "Cryptanalysis of the oil & vinegar signature scheme", in *CRYPTO'98*, H. Krawczyk, Ed., ser. LNCS, vol. 1462, Springer, Heidelberg, Aug. 1998, pp. 257–266.

[194] ——, "Cryptanalysis of the HFE public key cryptosystem by relinearization", in *CRYPTO'99*, M. J. Wiener, Ed., ser. LNCS, vol. 1666, Springer, Heidelberg, Aug. 1999, pp. 19–30.

[195] P. Kirchner and P.-A. Fouque, "An improved BKW algorithm for LWE with applications to cryptography and lattices", in *CRYPTO 2015, Part I*, R. Gennaro and M. J. B. Robshaw, Eds., ser. LNCS, vol. 9215, Springer, Heidelberg, Aug. 2015, pp. 43–62.

[196] L. Lamport, *Constructing digital signatures from a one-way function*, Technical Report SRI-CSL-98, SRI International Computer Science Laboratory, 1979.

[197] D. Lazard, "Gröbner-bases, Gaussian elimination and resolution of systems of algebraic equations", in *Proceedings of the European Computer Algebra Conference on Computer Algebra*, ser. Lecture Notes in Computer Science, vol. 162, Berlin, Heidelberg, New York: Springer Verlag, 1983.

[198] P. J. Lee and E. F. Brickell, "An observation on the security of McEliece's public-key cryptosystem", in *EUROCRYPT'88*, C. G. Günther, Ed., ser. LNCS, vol. 330, Springer, Heidelberg, May 1988, pp. 275–280.

[199] J. S. Leon, "A probabilistic algorithm for computing minimum weights of large error-correcting codes", *IEEE Transactions on Information Theory*, vol. 34, no. 5, pp. 1354–1359, 1988.

[200] T. Li, Z. Lin, and F. Bai, "Heuristic methods for computing the minimal multi-homogeneous bézout number", *Appl. Math. Comput.*, vol. 146, no. 1, pp. 237–256, Dec. 2003. [Online]. Available: http://dx.doi.org/10.1016/S0096-3003(02)00540-4.

[201] S. Ling, D. H. Phan, D. Stehlé, and R. Steinfeld, "Hardness of k-LWE and applications in traitor tracing", in *CRYPTO 2014, Part I*, J. A. Garay and R. Gennaro, Eds., ser. LNCS, vol. 8616, Springer, Heidelberg, Aug. 2014, pp. 315–334.

[202] F. Liu, C. Lu, and B. Yang, "Secure prngs from specialized polynomial maps over any Fq", in *Post-Quantum Cryptography, Second International Workshop, PQCrypto 2008, Cincinnati, OH, USA, October 17-19, 2008, Proceedings*, J. A. Buchmann and J. Ding, Eds., ser. Lecture Notes in Computer Science, vol. 5299, Springer, 2008, pp. 181–202. [Online]. Available: `http://dx.doi.org/10.1007/978-3-540-88403-3_13`.

[203] P. Loidreau and N. Sendrier, "Weak keys in the mceliece public-key cryptosystem", *IEEE Transactions on Information Theory*, vol. 47, no. 3, pp. 1207–1211, 2001.

[204] D. Lokshtanov, R. Paturi, S. Tamaki, R. Williams, and H. Yu, *Beating brute force for systems of polynomial equations over finite fields*, to appear, 27th ACM-SIAM Symposium on Discrete Algorithms (SODA 2017).

[205] V. Lyubashevsky and D. Micciancio, "Asymptotically efficient lattice-based digital signatures", in *Theory of Cryptography, Fifth Theory of Cryptography Conference, TCC 2008, New York, USA, March 19-21, 2008.*, R. Canetti, Ed., ser. Lecture Notes in Computer Science, vol. 4948, Springer, 2008, pp. 37–54. [Online]. Available: `http://dx.doi.org/10.1007/978-3-540-78524-8_3`.

[206] V. Lyubashevsky, D. Micciancio, C. Peikert, and A. Rosen, "Swifft: A modest proposal for FFT hashing", in *FSE 2008*, K. Nyberg, Ed., ser. LNCS, vol. 5086, Springer, Heidelberg, Feb. 2008, pp. 54–72.

[207] V. Lyubashevsky, C. Peikert, and O. Regev, "On ideal lattices and learning with errors over rings", in *EUROCRYPT 2010*, H. Gilbert, Ed., ser. LNCS, vol. 6110, Springer, Heidelberg, May 2010, pp. 1–23.

[208] F. S. Macaulay, "On some formula in elimination", *London Mathematical Society*, vol. 1, no. 33, pp. 3–27, 1902.

[209] F. Macaulay, *The Algebraic Theory of Modular Systems*. Cambridge University Press, 1916.

[210] F. J. MacWilliams and N. J. A. Sloane, *The Theory of Error-Correcting Codes*, Fifth. Amsterdam: North–Holland, 1986.

[211] G. Malajovich and K. Meer, "Computing minimal multi-homogeneous bézout numbers is hard", in *STACS 2005, 22nd Annual Symposium on Theoretical Aspects of Computer Science, Stuttgart, Germany, February 24-26, 2005, Proceedings*, V. Diekert and B. Durand, Eds., ser. Lecture Notes in Computer Science, vol. 3404, Springer, 2005, pp. 244–255. [Online]. Available: `http://dx.doi.org/10.1007/978-3-540-31856-9_20`.

[212] T. Matsumoto and H. Imai, "Public quadratic polynominal-tuples for efficient signature-verification and message-encryption", in *EUROCRYPT'88*, C. G. Günther, Ed., ser. LNCS, vol. 330, Springer, Heidelberg, May 1988, pp. 419–453.

[213] A. May and I. Ozerov, "On computing nearest neighbors with applications to decoding of binary linear codes", in *EUROCRYPT 2015, Part I*, E. Oswald and M. Fischlin, Eds., ser. LNCS, vol. 9056, Springer, Heidelberg, Apr. 2015, pp. 203–228.

[214] R. J. McEliece, "A public-key system based on algebraic coding theory", in. Jet Propulsion Lab, 1978, pp. 114–116, DSN Progress Report 44.

[215] D. D. A. McGrew, M. Curcio, and S. Fluhrer, "Hash-based signatures", Internet Engineering Task Force, Internet-Draft draft-mcgrew-hash-sigs-05, Oct. 31, 2016, Work in Progress, 37 pp. [Online]. Available: `https://tools.ietf.org/html/draft-mcgrew-hash-sigs-05`.

[216] R. C. Merkle, "A certified digital signature", in *CRYPTO'89*, G. Brassard, Ed., ser. LNCS, vol. 435, Springer, Heidelberg, Aug. 1990, pp. 218–238.

[217] D. Micciancio and P. Mol, "Pseudorandom knapsacks and the sample complexity of LWE search-to-decision reductions", in *CRYPTO 2011*, P. Rogaway, Ed., ser. LNCS, vol. 6841, Springer, Heidelberg, Aug. 2011, pp. 465–484.

[218] D. Micciancio and C. Peikert, "Hardness of SIS and LWE with small parameters", in *CRYPTO 2013, Part I*, R. Canetti and J. A. Garay, Eds., ser. LNCS, vol. 8042, Springer, Heidelberg, Aug. 2013, pp. 21–39.

[219] D. Micciancio and O. Regev, "Lattice-based cryptography", in *Post-Quantum Cryptography*, D. J. Bernstein, J. Buchmann, and E. Dahmen, Eds. Berlin, Heidelberg: Springer Berlin Heidelberg, 2009, pp. 147–191. [Online]. Available: `http://dx.doi.org/10.1007/978-3-540-88702-7_5`.

[220] R. Misoczki and P. S. L. M. Barreto, "Compact McEliece keys from Goppa codes", in *Selected Areas in Cryptography (SAC 2009)*, Calgary, Canada, Aug. 2009.

[221] M. W. Moskewicz, C. F. Madigan, Y. Zhao, L. Zhang, and S. Malik, "Chaff: Engineering an efficient SAT solver", in *Proceedings of the 38th Design Automation Conference, DAC 2001, Las Vegas, NV, USA, June 18-22, 2001*, ACM, 2001, pp. 530–535. [Online]. Available: `http://doi.acm.org/10.1145/378239.379017`.

[222] K. Nabeshima, K. Nagasaka, F. Winkler, and Á. Szántó, Eds., *International Symposium on Symbolic and Algebraic Computation, ISSAC '14, Kobe, Japan, July 23-25, 2014*, ACM, 2014. [Online]. Available: `http://dl.acm.org/citation.cfm?id=2608628`.

[223] V. Nachef, J. Patarin, and E. Volte, *Zero-knowledge for multivariate polynomials*, Cryptology ePrint Archive, Report 2012/239, `http://eprint.iacr.org/2012/239`, 2012.

[224] NESSIE, *New european schemes for signatures, integrity, and encryption, 2003.*

[225] L. Nicklasson, *On the hilbert series of ideals generated by generic forms*, Arxiv, 2015, `http://arxiv.org/abs/1502.06762?context=math`, 201.

[226] NIST, *Proposed submission requirements and evaluation criteria for the post-quantum cryptography standardization process (DRAFT).* `http://csrc.nist.gov/groups/ST/post-quantum-crypto/documents/call-for-proposals-draft-aug-2016.pdf`.

[227] R. Nojima, H. Imai, K. Kobara, and K. Morozov, "Semantic security for the McEliece cryptosystem without random oracles", *Des. Codes Cryptography*, vol. 49, no. 1-3, pp. 289–305, 2008.

[228] Y. Oren, M. Renauld, F.-X. Standaert, and A. Wool, "Algebraic side-channel attacks beyond the hamming weight leakage model", in *CHES 2012*, E. Prouff and P. Schaumont, Eds., ser. LNCS, vol. 7428, Springer, Heidelberg, Sep. 2012, pp. 140–154.

[229] R. Overbeck and N. Sendrier, "Code-based cryptography", in *Post-Quantum Cryptography*, D. J. Bernstein, J. Buchmann, and E. Dahmen, Eds. Berlin, Heidelberg: Springer Berlin Heidelberg, 2009, pp. 95–145. [Online]. Available: `http://dx.doi.org/10.1007/978-3-540-88702-7_4`.

[230] J. Patarin, "Cryptoanalysis of the Matsumoto and Imai public key scheme of eurocrypt'88", in *CRYPTO'95*, D. Coppersmith, Ed., ser. LNCS, vol. 963, Springer, Heidelberg, Aug. 1995, pp. 248–261.

[231] ——, "Hidden fields equations (HFE) and isomorphisms of polynomials (IP): Two new families of asymmetric algorithms", in *EUROCRYPT'96*, U. M. Maurer, Ed., ser. LNCS, vol. 1070, Springer, Heidelberg, May 1996, pp. 33–48.

[232] J. Patarin, N. Courtois, and L. Goubin, "QUARTZ, 128–bit long digital signatures", in *CT-RSA 2001*, D. Naccache, Ed., ser. LNCS, vol. 2020, Springer, Heidelberg, Apr. 2001, pp. 282–297.

[233] ——, "Quartz, 128-bit long digital signatures", in *Topics in Cryptology - CT-RSA 2001, The Cryptographer's Track at RSA Conference 2001, San Francisco, CA, USA, April 8-12, 2001, Proceedings*, D. Naccache, Ed., ser. Lecture Notes in Computer Science, vol. 2020, Springer, 2001, pp. 282–297. [Online]. Available: `http://dx.doi.org/10.1007/3-540-45353-9_21`.

[234] C. Peikert, "Public-key cryptosystems from the worst-case shortest vector problem: Extended abstract", in *41st ACM STOC*, M. Mitzenmacher, Ed., ACM Press, May 2009, pp. 333–342.

[235] C. Peikert, V. Vaikuntanathan, and B. Waters, "A framework for efficient and composable oblivious transfer", in *CRYPTO 2008*, D. Wagner, Ed., ser. LNCS, vol. 5157, Springer, Heidelberg, Aug. 2008, pp. 554–571.

[236] L. Perret, "Algebraic and combinatorial tools for public key cryptography", PhD thesis, Université de Marne-la-Vallé, 2005. [Online]. Available: `https://www.iacr.org/phds/index.php?p=detail&entry=1185`.

[237] E. Persichetti, "Compact Mceliece keys based on quasi-dyadic Srivastava codes", *J. Mathematical Cryptology*, vol. 6, no. 2, pp. 149–169, 2012.

[238] A. Petzoldt, S. Bulygin, and J. A. Buchmann, "A multivariate based threshold ring signature scheme", *Appl. Algebra Eng. Commun. Comput.*, vol. 24, no. 3-4, pp. 255–275, 2013. [Online]. Available: `http://dx.doi.org/10.1007/s00200-013-0190-3`.

[239] M. Fischlin, J. Buchmann, and M. Manulis, Eds., *PKC 2012*, vol. 7293, ser. LNCS, Springer, Heidelberg, May 2012.

[240] J. Katz, Ed., *PKC 2015*, vol. 9020, ser. LNCS, Springer, Heidelberg, Mar. 2015.

[241] E. Prange, "The use of information sets in decoding cyclic codes", *IRE Transactions on Information Theory*, vol. 8, no. 5, pp. 5–9, Sep. 1962.

[242] E. I. QSC, *Quantum-safe cryptography (QSC); quantum-safe algorithmic framework*, `http://www.etsi.org/deliver/etsi_gr/QSC/001_099/001/01.01.01_60/gr_QSC001v010101p.pdf`.

[243] H. Raddum and I. A. Semaev, "New technique for solving sparse equation systems", *IACR Cryptology ePrint Archive*, vol. 2006, p. 475, 2006. [Online]. Available: `http://eprint.iacr.org/2006/475`.

[244] ——, "Solving MRHS linear equations", *IACR Cryptology ePrint Archive*, vol. 2007, p. 285, 2007. [Online]. Available: `http://eprint.iacr.org/2007/285`.

[245] O. Regev, "On lattices, learning with errors, random linear codes, and cryptography", in *37th ACM STOC*, H. N. Gabow and R. Fagin, Eds., ACM Press, May 2005, pp. 84–93.

[246] ——, "On lattices, learning with errors, random linear codes, and cryptography", *Journal of the ACM*, vol. 56, no. 6, 34:1–34:40, Sep. 2009.

[247] M. Renauld, F.-X. Standaert, and N. Veyrat-Charvillon, "Algebraic side-channel attacks on the AES: Why time also matters in DPA", in *CHES 2009*, C. Clavier and K. Gaj, Eds., ser. LNCS, vol. 5747, Springer, Heidelberg, Sep. 2009, pp. 97–111.

[248] K. Sakumoto, "Public-key identification schemes based on multivariate cubic polynomials", in *PKC 2012*, M. Fischlin, J. Buchmann, and M. Manulis, Eds., ser. LNCS, vol. 7293, Springer, Heidelberg, May 2012, pp. 172–189.

[249] K. Sakumoto, T. Shirai, and H. Hiwatari, "On provable security of UOV and HFE signature schemes against chosen-message attack", in *Post-Quantum Cryptography - 4th International Workshop, PQCrypto 2011, Taipei, Taiwan, November 29 - December 2, 2011. Proceedings*, B. Yang, Ed., ser. Lecture Notes in Computer Science, vol. 7071, Springer, 2011, pp. 68–82. [Online]. Available: `http://dx.doi.org/10.1007/978-3-642-25405-5_5`.

[250] ——, "Public-key identification schemes based on multivariate quadratic polynomials", in *CRYPTO 2011*, P. Rogaway, Ed., ser. LNCS, vol. 6841, Springer, Heidelberg, Aug. 2011, pp. 706–723.

[251] P. Schwabe and B. Westerbaan, *Solving binary MQ with grover's algorithm*, SPACES'16, 2016.

[252] J. T. Schwartz, "Fast probabilistic algorithms for verification of polynomial identities", *Journal of the ACM*, vol. 27, no. 4, pp. 701–717, 1980.

[253] I. A. Semaev, "On solving sparse algebraic equations over finite fields", *Des. Codes Cryptography*, vol. 49, no. 1-3, pp. 47–60, 2008. [Online]. Available: `http://dx.doi.org/10.1007/s10623-008-9182-x`.

[254] ——, "Sparse algebraic equations over finite fields", *SIAM J. Comput.*, vol. 39, no. 2, pp. 388–409, 2009. [Online]. Available: `http://dx.doi.org/10.1137/070700371`.

[255] ——, "Improved agreeing-gluing algorithm", *Mathematics in Computer Science*, vol. 7, no. 3, pp. 321–339, 2013. [Online]. Available: `http://dx.doi.org/10.1007/s11786-013-0163-8`.

[256] ——, "MaxMinMax problem and sparse equations over finite fields", *Des. Codes Cryptography*, vol. 79, no. 2, pp. 383–404, 2016. [Online]. Available: `http://dx.doi.org/10.1007/s10623-015-0058-6`.

[257] N. Sendrier, "Cryptosystèmes à clé publique basés sur les codes correcteurs d'erreurs", Mémoire d'habilitation à diriger des recherches, Université Paris 6, Mar. 2002. [Online]. Available: `hab.pdf`.

[258] P. W. Shor, "Polynomial-time algorithms for prime factorization and discrete logarithms on a quantum computer", *SIAM J. Comput.*, vol. 26, no. 5, pp. 1484–1509, 1997.

[259] I. Simonetti, J.-C. Faugère, and L. Perret, "Algebraic attack against trivium", in *First International Conference on Symbolic Computation and Cryptography, SCC 08*, ser. LMIB, Beijing, China, Apr. 2008, pp. 95–102. [Online]. Available: `http://www-salsa.lip6.fr/~jcf/Papers/SCC08c.pdf`.

[260] M. Soos, K. Nohl, and C. Castelluccia, "Extending SAT solvers to cryptographic problems", in *Theory and Applications of Satisfiability Testing - SAT 2009, 12th International Conference, SAT 2009, Swansea, UK, June 30 - July 3, 2009. Proceedings*, O. Kullmann, Ed., ser. Lecture Notes in Computer Science, vol. 5584, Springer, 2009, pp. 244–257. [Online]. Available: http://dx.doi.org/10.1007/978-3-642-02777-2_24.

[261] P.-J. Spaenlehauer, "Solving multi-homogeneous and determinantal systems. algorithms - complexity - applications.", PhD thesis, PhD thesis, Université Paris 6, 2012. [Online]. Available: http://www-polsys.lip6.fr/~spaenleh/data/these_spaenlehauer.pdf.

[262] D. Stehlé, R. Steinfeld, K. Tanaka, and K. Xagawa, "Efficient public key encryption based on ideal lattices", in *ASIACRYPT 2009*, M. Matsui, Ed., ser. LNCS, vol. 5912, Springer, Heidelberg, Dec. 2009, pp. 617–635.

[263] J. Stern, "A method for finding codewords of small weight", in *Coding Theory and Applications*, G. D. Cohen and J. Wolfmann, Eds., ser. Lecture Notes in Computer Science, vol. 388, Springer, 1988, pp. 106–113.

[264] M. Sugita, M. Kawazoe, L. Perret, and H. Imai, "Algebraic Cryptanalysis of 58-Round SHA-1", in *FSE*, A. Biryukov, Ed., ser. Lecture Notes in Computer Science, vol. 4593, Springer, 2007, pp. 349–365.

[265] J. Svartz, "Solving zero-dimensional structured polynomial systems", Theses, Université Pierre et Marie Curie - Paris VI, Oct. 2014. [Online]. Available: https://tel.archives-ouvertes.fr/tel-01147484.

[266] A. Szanto, "Multivariate subresultants using Jouanolou matrices", *Journal of Pure and Applied Algebra*, vol. 214, no. 8, pp. 1347–1369, 2010. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0022404909002618.

[267] S. D. M. T. Hodges and J. Schalter, *On the existence of semi-regular sequences*, Arxiv, 2014, http://arxiv.org/abs/1412.7865v1, 2014.

[268] E. Thomae and C. Wolf, "Cryptanalysis of enhanced TTS, STS and all its variants, or: Why cross-terms are important", in *AFRICACRYPT 12*, A. Mitrokotsa and S. Vaudenay, Eds., ser. LNCS, vol. 7374, Springer, Heidelberg, Jul. 2012, pp. 188–202.

[269] ——, "Solving underdetermined systems of multivariate quadratic equations revisited", in *PKC 2012*, M. Fischlin, J. Buchmann, and M. Manulis, Eds., ser. LNCS, vol. 7293, Springer, Heidelberg, May 2012, pp. 156–171.

[270] R. C. Torres and N. Sendrier, "Analysis of information set decoding for a sub-linear error weight", in *Post-Quantum Cryptography - 7th International Workshop, PQCrypto 2016, Fukuoka, Japan, February 24-26, 2016, Proceedings*, T. Takagi, Ed., ser. Lecture Notes in Computer Science, vol. 9606, Springer, 2016, pp. 144–161. [Online]. Available: http://dx.doi.org/10.1007/978-3-319-29360-8_10.

[271] B. Yang, Ed., *Post-Quantum Cryptography - 4th International Workshop, PQCrypto 2011, Taipei, Taiwan, November 29 - December 2, 2011. Proceedings*, vol. 7071, ser. Lecture Notes in Computer Science, Springer, 2011. [Online]. Available: http://dx.doi.org/10.1007/978-3-642-25405-5.

[272] G. Yang, S. Tang, and L. Yang, "A novel group signature scheme based on MPKC", in *Information Security Practice and Experience - 7th International Conference, ISPEC 2011, Guangzhou, China, May 30 - June 1, 2011. Proceedings*, F. Bao and J. Weng, Eds., ser. Lecture Notes in Computer Science, vol. 6672, Springer, 2011, pp. 181–195. [Online]. Available: `http://dx.doi.org/10.1007/978-3-642-21031-0_14`.

[273] B.-Y. Yang, J.-M. Chen, and Y.-H. Chen, "Tts: High-speed signatures on a low-cost smart card", in *CHES 2004*, M. Joye and J.-J. Quisquater, Eds., ser. LNCS, vol. 3156, Springer, Heidelberg, Aug. 2004, pp. 371–385.

[274] B.-Y. Yang, J.-M. Chen, and N. Courtois, "On asymptotic security estimates in xl and gröbner bases-related algebraic cryptanalysis", in *ICICS 2004*, ser. Lecture Notes in Computer Science, vol. 3269, Springer, 2004, pp. 401–413.

[275] ——, "On asymptotic security estimates in XL and Gröbner bases-related algebraic cryptanalysis", in *ICICS 04*, J. López, S. Qing, and E. Okamoto, Eds., ser. LNCS, vol. 3269, Springer, Heidelberg, Oct. 2004, pp. 401–413.

[276] R. Zippel, "Probabilistic algorithms for sparse polynomials", in *Symbolic and algebraic computation (EUROSAM'79), Internat. Sympos.*, ser. Lecture Notes in Computer Science, vol. 72, Marseille: Springer Verlag, 1979, pp. 216–226.