Algebraic Cryptanalysis of Compact McEliece's Variants – Toward a Complexity Analysis

Jean-Charles Faugère[1], Ayoub Otmani[2,3], Ludovic Perret[1], and Jean-Pierre Tillich[2]

[1] SALSA Project - INRIA (Centre Paris-Rocquencourt)
UPMC, Univ Paris 06 - CNRS, UMR 7606, LIP6
104, avenue du Président Kennedy 75016 Paris, France
jean-charles.faugere@inria.fr, ludovic.perret@lip6.fr
[2] SECRET Project - INRIA Rocquencourt
Domaine de Voluceau, B.P. 105 78153 Le Chesnay Cedex - France
ayoub.otmani@inria.fr, jean-pierre.tillich@inria.fr
[3] GREYC - Université de Caen - Ensicaen
Boulevard Maréchal Juin, 14050 Caen Cedex, France.

Abstract. A new algebraic approach to investigate the security of the McEliece cryptosystem has been proposed by Faugère-Otmani-Perret-Tillich in Eurocrypt 2010. This paper is an extension of this work. The McEliece's scheme relies on the use of error-correcting codes. It has been proved that the private key of the cryptosystem satisfies a system of bi-homogeneous polynomial equations. This property is due to the particular class of codes considered which are alternant codes. These highly structured algebraic equations allowed to mount an efficient key-recovery attack against two recent variants of the McEliece cryptosystems that aim at reducing public key sizes by using quasi-cyclic or quasi-dyadic structures. Thanks to a very recent development due to Faugère-Safey el Din-Spaenlehauer on the solving of bihomogeneous bilinear systems, we can estimate the complexity of the FOPT algebraic attack. This is a first step toward providing a concrete criterion for evaluating the security of future compact McEliece variants.

Keywords : public-key cryptography, McEliece cryptosystem, algebraic cryptanalysis, F_5, bi-linear systems,.

1 Introduction

After more than thirty years now, the McEliece cryptosystem still belongs to the very few public key cryptosystems which remain unbroken. Its security relies upon two assumptions: the *intractability of decoding random linear codes* [7], and the *difficulty of recovering the private key* or an equivalent one. The problem of decoding an unstructured code is a long-standing problem whose most effective algorithms [18, 19, 24, 10, 8] have an exponential time complexity. On the other hand no significant breakthrough has been observed during the past years regarding the problem of recovering the private key. Indeed, although some weak keys have been identified in [20], the only known key-recovery attack is the exhaustive search of the secret polynomial $\Gamma(z)$ of the Goppa code, and applying the *Support Splitting Algorithm* (SSA) [23] to check whether the Goppa code candidate is *permutation-equivalent* to the code defined by the public generator matrix.

Despite its impressive resistance against a variety of attacks and its fast encryption and decryption, McEliece cryptosystem has not stood up to RSA for practical applications. This is most likely due to the large size of the public key which is between several hundred thousand and several million bits. To overcome this limitation, a trend had been initiated in order to decrease the key size by focusing on very structured codes. For instance, quasi-cyclic code like in [17], or quasi-cyclic codes defined by sparse matrices (also called LDPC codes) [1]. Both schemes were broken in [22]. It should be noted that the attacks have no impact on the security of the McEliece cryptosystem since both proposals did not use the binary Goppa codes of the McEliece cryptosystem. These works were then followed by two independent proposals [6, 21] that are based on the same kind of idea

of using quasi-cyclic [6] or quasi-dyadic structure [21]. These two approaches were also broken in [16] where for the first time an algebraic attack is introduced against the McEliece cryptosystem.

Algebraic cryptanalysis is a general framework that permits to assess the security of theoretically all cryptographic schemes. So far, such type of attacks has been applied successfully against several multivariate schemes and stream ciphers. The basic principle of this cryptanalysis is to associate to a cryptographic primitive a set of algebraic equations. The system of equations is constructed in such a way to have a correspondence between the solutions of this system, and a secret information of the cryptographic primitive (for instance the secret key of an encryption scheme). In the case of the McEliece cryptosystem, the algebraic system [16] that has to be solved has the following very specific structure:

$$\mathsf{McE}_{k,n,r}(\mathbf{X},\mathbf{Y}) = \left\{ g_{i,0}Y_0X_0^j + \cdots + g_{i,n-1}Y_{n-1}X_{n-1}^j = 0 \;\middle|\; i \in \{0,\ldots,k-1\}, j \in \{0,\ldots,r-1\} \right\} \quad (1)$$

where the unknowns are the X_i's and the Y_i's and the $g_{i,j}$'s are known coefficients with $0 \leq i \leq k-1$, $0 \leq j \leq n-1$ that belong to a certain field $\mathbb{F}_q$ with $q = 2^s$. We look for solutions of this system in a certain extension field $\mathbb{F}_{q^m}$. Here k is an integer which is at least equal to $n - rm$. By denoting $\mathbf{X} \stackrel{\text{def}}{=} (X_0,\ldots,X_{n-1})$ and $\mathbf{Y} \stackrel{\text{def}}{=} (Y_0,\ldots,Y_{n-1})$ we will refer to such an algebraic system by $\mathsf{McE}_{k,n,r}(\mathbf{X},\mathbf{Y})$. This algebraic approach as long as the codes that are considered are alternant codes. It is important to note that a Goppa code can also be seen as a particular alternant code. However, it is not clear whether an algebraic attack can be mounted efficiently against the original McEliece cryptosystem because the total number of equations is rk, the number of unknowns $2n$ and the maximum degree $r-1$ of the equations can be extremely high (e.g. $n = 1024$ and $r - 1 = 49$).

But in the case of the tweaked McEliece schemes [6, 21], it turns out that is possible to make use of this structure in order to reduce considerably the number of unknowns in the algebraic system. This is because of the type of codes that are considered: quasi-cyclic alternant codes in [6] and quasi-dyadic Goppa codes in [21]. In particular, it induces an imbalance between the $\mathbf{X}$ and $\mathbf{Y}$ variables. Moreover, it was possible to solve efficiently the algebraic system thanks to a dedicated Gröbner bases techniques. Finally, it was also observed experimentally in [16] but not formally proved that the complexity of the attack is mainly determined by the number of remaining variables in the block $\mathbf{Y}$.

The motivation of this paper is to revisit the FOPT algebraic attack [16] in view of the recent results on bilinear systems [14]. This permits to make more precise the dependency between the security of a McEliece (and its variants) and the properties of the algebraic system (1). This is a first step toward providing a concrete criterion for evaluating the security of future compact McEliece variants.

2 On the Solving of $\mathsf{McE}_{k,n,r}(\mathbf{X},\mathbf{Y})$

Thanks to a very recent development [14] on the solving of bi-linear systems, we can revisit the strategy used in [16] to solve $\mathsf{McE}_{k,n,r}(\mathbf{X},\mathbf{Y})$. Moreover, this permits to evaluate the complexity of computing a Gröbner bases of $\mathsf{McE}_{k,n,r}(\mathbf{X},\mathbf{Y})$ for compact variants of McEliece such as [6, 21]. Before that, we briefly recall basic facts about the complexity of computing Gröbner bases [9, 11–13].

2.1 Extracting a Bi-Affine System from $\mathsf{McE}_{k,n,r}(\mathbf{X},\mathbf{Y})$

As explained, $\mathsf{McE}_{k,n,r}(\mathbf{X},\mathbf{Y})$ is highly structured. It is very sparse as the only monomials occurring in the system are of the form $Y_iX_i^j$, with $0 \leq i \leq k-1$ and $0 \leq j \leq r-1$. It can also be noticed that each block of k equations is *bi-homogeneous*, i.e. homogeneous if the variables of $\mathbf{X}$ (resp. $\mathbf{Y}$) are considered alone. More precisely, we shall say that $f \in \mathbb{F}_{q^m}[\mathbf{X},\mathbf{Y}]$ is *bi-homogeneous* of *bi-degree* (d_1,d_2) if:

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

Note that the equations occurring in $\mathsf{McE}_{k,n,r}(\mathbf{X},\mathbf{Y})$ are of bi-degree $(j,1)$, with $j, 0 \leq j \leq r-1$.

We briefly recall now the strategy followed in [16] to solve $\mathsf{McE}_{k,n,r}(\mathbf{X},\mathbf{Y})$. The first fundamental remark is that there are k linear equations in the n variables of the block $\mathbf{Y}$ in $\mathsf{McE}_{k,n,r}(\mathbf{X},\mathbf{Y})$. This implies that all the variables of the block $\mathbf{Y}$ can be expressed in terms of $n_{Y'} \geq n-k$ variables. From now on, we will always assume that the variables of the block $\mathbf{Y}'$ only refer to these $n_{Y'}$ *free* variables. The first step is then to rewrite the system (1) only in function of the variables of $\mathbf{X}$ and $\mathbf{Y}'$, i.e., the variables of $\mathbf{Y} \setminus \mathbf{Y}'$ are substituted by linear combinations involving only variables of $\mathbf{Y}'$.

In the particular cases of [6, 21], the quasi-cyclic and dyadic structures provide additional linear equations in the variables of $\mathbf{X}$ and $\mathbf{Y}'$ which can be also used to rewrite/clean the system. In the sequel, we denote by $\mathsf{McE}_{k,n,r}(\mathbf{X}',\mathbf{Y}')$ the system obtained from $\mathsf{McE}_{k,n,r}(\mathbf{X},\mathbf{Y})$ by removing all the linear equations in $\mathbf{X}$ and $\mathbf{Y}$.

This system $\mathsf{McE}_{k,n,r}(\mathbf{X}',\mathbf{Y}')$ being naturally overdetermined, we can "safely" remove some equations. In [6, 21], the system $\mathsf{McE}_{k,n,r}(\mathbf{X}',\mathbf{Y}')$ is always defined over a field of characteristic two. It makes sense then to consider the set of equations of $\mathsf{McE}_{k,n,r}(\mathbf{X}',\mathbf{Y}')$ whose degree in the variables of $\mathbf{X}'$ is a power of 2, i.e. equations of bi-degree $(2^j,1)$. We obtain in this way a sub-system of $\mathsf{McE}_{k,n,r}(\mathbf{X}',\mathbf{Y}')$, denoted $\mathsf{BiMcE}_{k,n,r}(\mathbf{X}',\mathbf{Y}')$, having $n_{X'}$ and $n_{Y'}$ variables and at most $k \cdot \log_2(r)$ equations. This system is a "quasi" bi-linear system over $\mathbb{F}_2^m$ as $\mathsf{McE}_{k,n,r}(\mathbf{X}',\mathbf{Y}')$ viewed over $\mathbb{F}_2$ is bi-linear. Note that some constant terms can occur in $\mathsf{McE}_{k,n,r}(\mathbf{X}',\mathbf{Y}')$, so the system is more precisely *affine* bi-linear.

Proposition 1. *Let* $\mathsf{BiMcE}_{k,n,r}(\mathbf{X}',\mathbf{Y}') \subset \mathbb{F}_{q^m}[\mathbf{X}',\mathbf{Y}']$ *be the system from* $\mathsf{McE}_{k,n,r}(\mathbf{X}',\mathbf{Y}')$ *by considering only the equations of bi-degree* $(2^j,1)$*. This system has* $n_{X'}+n_{Y'}$ *variables, at most* $k \cdot \log_2(r)$ *equations and is affine bi-linear.*

2.2 On the Complexity of Solving Affine Bi-Linear Systems

Whilst the complexity of solving general bi-homogenous system is not known, the situation is different for bi-affine (resp. bi-linear) systems. In particular, the theoretical complexity is well mastered, and there is a now a dedicated algorithm for such systems [14]. As already explained, our equations are "quasi" bi-linear as we are working with equations of bi-degree $(1,2^j)$ over a field of characteristic 2. The results presented in [14] can be then extended with a slight adaptation to the context.

A first important result of [14] is that F_5 [13] algorithm is already optimal for "generic" (random) affine bi-linear systems, i.e. all reductions to zero are removed by the F_5 criterion. Another fundamental result is that the degree of regularity of a square generic affine bi-linear system is much smaller than the degree of regularity of a generic system. It has been proved [14] that:

Proposition 2. *The degree of regularity of a square generic affine bi-linear system in* **X***'* *and* **Y***'* *is bounded by* $1+\min(n_{X'},n_{Y'})$*, where* $n_{X'}$ *and* $n_{Y'}$ *are the number of variables in the blocks* $\mathbf{X}'$ *and* $\mathbf{Y}'$ *respectively. Hence, the maximal degree occurring in the computation of a DRL Gröbner basis is also bounded by* (2)*.*

Remark 1. This bound is sharp for a generic square affine bi-linear system and is much better than the usual Macaulay's bound for a similar quadratic system (that is to say a system of $n_{X'}+n_{Y'}$ quadratic equations in $n_{X'}+n_{Y'}$ variables): $1+\min(n_{X'},n_{Y'}) \ll 1+n_{X'}+n_{Y'}$.

Since $\mathsf{BiMcE}_{k,n,r}(\mathbf{X}',\mathbf{Y}')$ is a bilinear system it is reasonable to derive the bound:

Proposition 3. *Let* $\mathsf{BiMcE}_{k,n,r}(\mathbf{X}',\mathbf{Y}')$ *be as defined below. The maximum degree reached when computing a Gröbner basis of* $\mathsf{BiMcE}_{k,n,r}(\mathbf{X}',\mathbf{Y}')$ *is smaller than* $1+\min(n_{X'},n_{Y'})$.

Remark 2. Note that the bound is not tight at all. In our situation the affine bi-linear systems are overdetermined whilst [14] only considered systems with at most as many variables than the number of equations.

Finally, it appears [14] that the matrices occurring during the matrix version of F_5 can be made divided into smaller matrices thanks to the bi-linear structure. Let $\dim(R_{d_1,d_2}) = \binom{d_1+n_{X'}}{d_1}\binom{d_2+n_{Y'}}{d_2}$. More precisely, the matrices occurring at degree D during the matrix F_5 on a bi-linear systems are of size: $\left(\dim(R_{d_1,d_2}) - [t_1^{d_1}t_2^{d_2}]\mathrm{HS}(t_1,t_2)\right) \times \dim(R_{d_1,d_2})$ for all (d_1,d_2) such that $d_1+d_2=D, 1 \leq d_1,d_2 \leq D-1$, where the notation $[t_1^{d_1}t_2^{d_2}]\mathrm{HS}(t_1,t_2)$ stands for the coefficient of the term $t_1^{d_1}t_2^{d_2}$ in the Hilbert bi-serie $\mathrm{HS}(t_1,t_2)$ defined in the appendix.

As pointed out, these results hold for a bi-linear system. For an affine bi-linear, this can be considered as a good (i.e. first order) approximation. The idea is that we have to "bi-homogenize" the affine bi-linear system which corresponds to add some columns. We can then estimate the space/time complexity of computing a Gröbner basis of $\mathsf{BiMcE}_{k,n,r}(\mathbf{X}',\mathbf{Y}')$.

Proposition 4. *. The time complexity of computing a DRL-Gröbner basis G_{DRL} of $\mathsf{BiMcE}_{k,n,r}(\mathbf{X}',\mathbf{Y}')$ is bounded from above by $T_\alpha = \sum_{\substack{d_1+d_2=D \\ 1\leq d_1,d_2\leq D-1}} \left(\dim(R_{d_1,d_2}) - [t_1^{d_1}t_2^{d_2}]\mathrm{HS}(t_1,t_2)\right)^\alpha \dim(R_{d_1,d_2})$, with $\alpha = \omega-1, 2\leq\omega\leq 3$ and $D = \min(n_{X'},n_{Y'})+1$.*

It is worth to mention that, for the cryptosystems considered in [16], the number of free variables $n_{Y'}$ in $\mathbf{Y}'$ can be rather small (typically 1 or 2 for some challenges). We have then a theoretical explanation of the practical efficiency observed in [16]. In addition, we have a concrete criteria to evaluate the security of future compact McEliece's variants, namely the minimum of the number of variables $n_{X'}$ and $n_{Y'}$ in the blocks $\mathbf{X}'$ and $\mathbf{Y}'$ respectively should be sufficiently "big". This will be further discussed in the last section.

To conclude this section, we mention that the goal of the attack is compute the variety (i.e. set of solutions) $\mathscr{V}$ associated to $\mathsf{McE}_{k,n,r}(\mathbf{X}',\mathbf{Y}')$. As soon as we have a DRL-Gröbner basis G_{DRL} of $\mathsf{BiMcE}_{k,n,r}(\mathbf{X}',\mathbf{Y}')$, the variety can be obtained in $\mathscr{O}\left((\#\mathscr{V})^\omega\right)$ thanks to a change of ordering algorithm [15]. We have to be sure that the variety $\mathscr{V}$ has few solutions. In particular, we have to remove parasite solutions (corresponding to $X_i = X_j$ or to $Y_j = 0$). A classical way to do that is to introduce new variables u_{ij} and v_i and add to $\mathsf{McE}_{k,n,r}(\mathbf{X}',\mathbf{Y}')$ equations of the form: $u_{ij}\cdot(X_i - X_j)+1 =$ and $v_i \cdot Y_i + 1 = 0$. In practice, we have not added all theses equations; but only few of them (namely 4 or 5). The reason is that we do not want to add too many new variables. These equations and variables can be added to $\mathsf{BiMcE}_{k,n,r}(\mathbf{X}',\mathbf{Y}')$ whilst keeping the affine bi-linear structure. To do so, we have to add the v_i to the block $\mathbf{X}'$, and the variables u_{ij} to the block $\mathbf{Y}'$. So, as we add only few new variables, the complexity of solving $\mathsf{BiMcE}_{k,n,r}(\mathbf{X}',\mathbf{Y}')$ with these new constraints is essentially similar to Proposition 4.

3 Comparison of Theoretical complexity with Experimental Results

In the table below, we present the experimental results obtained in [16] for BCGO and MB schemes. For the sack of comparison, we include a bound on theoretical complexity of computing a Gröbner bases of $\mathsf{BiMcE}_{k,n,r}(\mathbf{X}',\mathbf{Y}')$ using the estimation $T_{\mathrm{theo}} \approx T_1$ as obtained in Section 2 proposition 4. Regarding the linear algebra, this is a bit optimistic. However, as already pointed our, we have been also rather pessimistic regarding others parameters. For instance, we are not using the fact that the systems are overdetermined, and we have also only considered a sub-system of $\mathsf{McE}_{k,n,r}(\mathbf{X}',\mathbf{Y}')$.

All in all, this bound permits a give a reasonable picture of the hardness of solving $\mathsf{BiMcE}_{k,n,r}(\mathbf{X}',\mathbf{Y}')$. It is of course not sufficient to set parameters, but sufficient to discard many weak compact variants of McEliece.
We briefly discussed of the theoretical complexity obtained for the first row of the second column. As explained, we have used the formula (3). We have computed the coefficient $[t_1^{d_1}t_2^{d_2}]\mathrm{HS}(t_1,t_2)$ by using the explicit formula of $\mathrm{HS}(t_1,t_2)$ provided in the appendix using the explicit values of $n_{X'} = 59$ and $n_{Y'}$=7, and assuming that the system is square; in that case the degree of regularity is 8. For this parameter, the sub-system $\mathsf{BiMcE}_{k,n,r}(\mathbf{X}',\mathbf{Y}')$ has actually 288 equations (of degree $2,3$ and 5). Hence, it is interesting to compute [2, 4, 3, 5] the degree of regularity of a semi-regular system of the same size: we found a regularity of 11 leading

Table 1. Cryptanalysis results for [6] ($m = 2$)

Challenge	q	ℓ	n_0	$n_{Y'}$	Security [6]	$n_{X'}$	Time (Operations, Memory)	T_{theo}
A_{16}	2^8	51	9	3	80	8	0.06 sec ($2^{18.9}$ op, 115 Meg)	2^{17}
B_{16}	2^8	51	10	3	90	9	0.03 sec ($2^{17.1}$ op, 116 Meg)	2^{18}
C_{16}	2^8	51	12	3	100	11	0.05 sec ($2^{16.2}$ op, 116 Meg)	2^{20}
D_{16}	2^8	51	15	4	120	14	0.02 sec ($2^{14.7}$ op, 113 Meg)	2^{26}
A_{20}	2^{10}	75	6	2	80	5	0.05 sec ($2^{15.8}$ op, 115 Meg)	2^{10}
B_{20}	2^{10}	93	6	2	90	5	0.05 sec ($2^{17.1}$ op, 115 Meg)	2^{10}
C_{20}	2^{10}	93	8	2	110	7	0.02 sec ($2^{14.5}$ op, 115 Meg)	2^{11}
QC_{600}	2^8	255	15	3	600	14	0.08 sec ($2^{16.6}$ op, 116 Meg)	2^{21}

Table 2. Cryptanalysis results for [21].

Challenge	q	$n_{Y'}$	ℓ	n_0	Security	$n_{X'}$	Time (Operations, Memory)	T_{theo}
Table 2	2^2	7	64	56	128	59	1,776.3 sec ($2^{34.2}$ op, 360 Meg)	2^{65}
Table 2	2^4	3	64	32	128	36	0.50 sec ($2^{22.1}$ op, 118 Meg)	2^{29}
Table 2	2^8	1	64	12	128	16	0.03 sec ($2^{16.7}$ op, 35 Meg)	2^8
Table 3	2^8	1	64	10	102	14	0.03 sec ($2^{15.9}$ op, 113 Meg)	2^8
Table 3	2^8	1	128	6	136	11	0.02 sec ($2^{15.4}$ op, 113 Meg)	2^7
Table 3	2^8	1	256	4	168	10	0.11 sec ($2^{19.2}$ op, 113 Meg)	2^7
Table 5	2^8	1	128	4	80	9	0.06 sec ($2^{17.7}$ op, 35 Meg)	2^6
Table 5	2^8	1	128	5	112	10	0.02 sec ($2^{14.5}$ op, 35 Meg)	2^7
Table 5	2^8	1	128	6	128	11	0.01 sec ($2^{16.6}$ op, 35 Meg)	2^7
Table 5	2^8	1	256	5	192	11	0.05 sec ($2^{17.5}$ op, 35 Meg)	2^7
Table 5	2^8	1	256	6	256	12	0.06 sec ($2^{17.8}$ op, 35 Meg)	2^7
Dyadic_{256}	2^4	3	128	32	256	37	7.1 sec ($2^{26.1}$ op, 131 Meg)	2^{29}
Dyadic_{512}	2^8	1	512	6	512	13	0.15 sec ($2^{19.7}$ op, 38 Meg)	2^8

to a cost of $2^{85.2}$ for the Gröbner basis computation (using the Macaulay's bound with $\omega = 2$). It is expected that a new results of the degree of regularity of generic overdetermined bi-linear systems would lead to tighter bounds.

As a conclusion, one can see that the theoretical bound (3) provides a reasonable explanation regarding the efficiency of the attack presented in [16]. In particular, it is important to remark that the hardness of the attack seems related to $d = \min(n'_X, n'_Y)$. The complexity of the attack clearly increases with this quantity. For the design of future compact variants of McEliece, this d should be then not too small. Regarding the current state of the art, it is difficult to provide an exact value. Very roughly speaking, $\text{BiMcE}_{k,n,r}(\mathbf{X}', \mathbf{Y}')$ can be considered as hard as solving a random (overdetermined) algebraic system with $d = \min(n_{X'}, n_{Y'})$ equations over a big field. With this in mind, we can say that any system with $d \leq 20$ should be within the scope of an algebraic attack.

Note that another phenomena, which remains to be treated, can occur. In the particular case of binary dyadic codes, the Gröbner basis of $\text{BiMcE}_{k,n,r}(\mathbf{X}', \mathbf{Y}')$ can be easily computed, but the variety associated is too big. This is due to the fact that the Gröbner basis is "trivial" (reduced to one equation) and not provides then enough information. This is typically due to the fact that we have used only a sub-set of the equations $\big($of bi-degree $(2^j, 1)\big)$. So, the open question is how we can use cleverly all the equations of $\text{McE}_{k,n,r}(\mathbf{X}', \mathbf{Y}')$ in the binary case.

References

1. M. Baldi and G. F. Chiaraluce. Cryptanalysis of a new instance of McEliece cryptosystem based on QC-LDPC codes. In *IEEE International Symposium on Information Theory*, pages 2591–2595, Nice, France, March 2007.
2. Magali Bardet. *Étude des systèmes algébriques surdéterminés. Applications aux codes correcteurs et à la cryptographie*. PhD thesis, Université de Paris VI, 2004.
3. Magali Bardet, Jean-Charles Faugère, and Bruno Salvy. Complexity study of Gröbner basis computation. Technical report, INRIA, 2002. http://www.inria.fr/rrrt/rr-5049.html.
4. Magali Bardet, Jean-Charles Faugère, and Bruno Salvy. On the complexity of Gröbner basis computation of semi-regular overdetermined algebraic equations. In *Proc. International Conference on Polynomial System Solving (ICPSS)*, pages 71–75, 2004.
5. Magali Bardet, Jean-Charles Faugère, Bruno Salvy, and Bo-Yin Yang. Asymptotic behaviour of the degree of regularity of semi-regular polynomial systems. In *Proc. of MEGA 2005, Eighth International Symposium on Effective Methods in Algebraic Geometry*, 2005.
6. T. P. Berger, P.L. Cayrel, P. Gaborit, and A. Otmani. Reducing key length of the McEliece cryptosystem. In Bart Preneel, editor, *Progress in Cryptology - Second International Conference on Cryptology in Africa (AFRICACRYPT 2009)*, volume 5580 of *Lecture Notes in Computer Science*, pages 77–97, Gammarth, Tunisia, June 21-25 2009.
7. E. Berlekamp, R. McEliece, and H. van Tilborg. On the inherent intractability of certain coding problems. *Information Theory, IEEE Transactions on*, 24(3):384–386, May 1978.

8. D. J. Bernstein, T. Lange, and C. Peters. Attacking and defending the McEliece cryptosystem. In *PQCrypto*, volume 5299 of *LNCS*, pages 31–46, 2008.
9. Bruno Buchberger. *Ein Algorithmus zum Auffinden der Basiselemente des Restklassenringes nach einem nulldimensionalen Polynomideal*. PhD thesis, University of Innsbruck, 1965.
10. 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*, 44(1):367–378, 1998.
11. D. A. Cox, J. B. Little, and D. O'Shea. *Ideals, Varieties, and algorithms: an Introduction to Computational Algebraic Geometry and Commutative Algebra.* Undergraduate Texts in Mathematics, Springer-Verlag, New York., 2001.
12. J.-C. Faugère. A new efficient algorithm for computing gröbner bases (f4). *Journal of Pure and Applied Algebra*, 139(1-3):61–88, 1999.
13. J.-C. Faugère. A new efficient algorithm for computing gröbner bases without reduction to zero : F5. In *ISSAC'02*, pages 75–83. ACM press, 2002.
14. Jean-Charles Faugère, Mohab Safey El Din, and Pierre-Jean Spaenlehauer. Gröbner bases of bihomogeneous ideals generated by polynomials of bidegree (1,1): Algorithms and complexity. *CoRR*, abs/1001.4004, 2010.
15. Jean-Charles Faugère, Patrizia M. Gianni, Daniel Lazard, and Teo Mora. Efficient computation of zero-dimensional gröbner bases by change of ordering. *J. Symb. Comput.*, 16(4):329–344, 1993.
16. Jean-Charles Faugère, Ayoub Otmani, Ludovic Perret, and Jean-Pierre Tillich. Algebraic cryptanalysis of McEliece variants with compact keys. In *Proceedings of Eurocrypt 2010*. Springer Verlag, 2010. to appear.
17. P. Gaborit. Shorter keys for code based cryptography. In *Proceedings of the 2005 International Workshop on Coding and Cryptography (WCC 2005)*, pages 81–91, Bergen, Norway, March 2005.
18. P. J. Lee and E. F. Brickell. An observation on the security of McEliece's public-key cryptosystem. In *Advances in Cryptology - EUROCRYPT'88*, volume 330/1988 of *Lecture Notes in Computer Science*, pages 275–280. Springer, 1988.
19. J. S. Leon. A probabilistic algorithm for computing minimum weights of large error-correcting codes. *IEEE Transactions on Information Theory*, 34(5):1354–1359, 1988.
20. P. Loidreau and N. Sendrier. Weak keys in the mceliece public-key cryptosystem. *IEEE Transactions on Information Theory*, 47(3):1207–1211, 2001.
21. R. Misoczki and P. S. L. M. Barreto. Compact McEliece keys from Goppa codes. In *Selected Areas in Cryptography (SAC 2009)*, Calgary, Canada, August 13-14 2009.
22. A. Otmani, J.P. Tillich, and L. Dallot. Cryptanalysis of McEliece cryptosystem based on quasi-cyclic ldpc codes. In *Proceedings of First International Conference on Symbolic Computation and Cryptography*, pages 69–81, Beijing, China, April 28-30 2008. LMIB Beihang University.
23. N. Sendrier. Finding the permutation between equivalent linear codes: The support splitting algorithm. *IEEE Transactions on Information Theory*, 46(4):1193–1203, 2000.
24. J. Stern. A method for finding codewords of small weight. In G. D. Cohen and J. Wolfmann, editors, *Coding Theory and Applications*, volume 388 of *Lecture Notes in Computer Science*, pages 106–113. Springer, 1988.

A Hilbert Bi-Series

We say that an ideal is *bihomogeneous* if there exists a set of bihomogeneous generators. The vector space of bihomogeneous polynomials of bi-degree (α,β) in a polynomial ring R will be denoted by $R_{\alpha,\beta}$. If $\mathscr{I}$ is a bihomogeneous ideal, then $I_{\alpha,\beta}$ will denote the vector space $I \cap R_{\alpha,\beta}$.

Definition 1 ([14]). *Let $\mathscr{I}$ be a bihomogeneous ideal of R. The Hilbert bi-series is defined by*

$$\mathrm{HS}_{\mathscr{I}}(t_1,t_2) = \sum_{(\alpha,\beta)\in\mathbb{N}^2} \dim(R_{\alpha,\beta}/I_{\alpha,\beta}) t_1^{\alpha} t_2^{\beta}.$$

For bi-regular bilinear systems, [14] provide an explicit form of the bi-series.

Theorem 1. *Let $f_1,\ldots,f_m \in R$ be a bi-regular bilinear sequence, with $m \leq n_{X'}+n_{Y'}$. Then*

$$\mathrm{HS}_{I_m}(t_1,t_2) = \frac{(1-t_1t_2)^m + N_{n_{X'}+1}(t_1,t_2) + N_{n_{Y'}+1}(t_1,t_2)}{(1-t_1)^{n_{X'}+1}(1-t_2)^{n_{Y'}+1}},$$

where

$$N_n(t_1,t_2) = t_1t_2(1-t_2)^n \sum_{\ell=1}^{m-n} (1-t_1t_2)^{m-n-\ell} \left[1-(1-t_1)^{\ell} \sum_{k=1}^{n} t_1^{n-k} \binom{\ell+n-k-1}{n-k} \right].$$