Differential Cryptanalysis for Multivariate Schemes

Pierre-Alain Fouque, Louis Granboulan, and Jacques Stern

École normale supérieure
Département d'Informatique 45, rue d'Ulm
75230 Paris cedex 05, France
Pierre-Alain.Fouque@ens.fr,
Louis.Granboulan@ens.fr, Jacques.Stern@ens.fr

Abstract. In this paper we propose a novel cryptanalytic method against multivariate schemes, which adapts differential cryptanalysis to this setting. In multivariate quadratic systems, the differential of the public key is a linear map and has invariants such as the dimension of the kernel. Using linear algebra, the study of this invariant can be used to gain information on the secret key. We successfully apply this new method to break the original Matsumoto-Imai cryptosystem using properties of the differential, thus providing an alternative attack against this scheme besides the attack devised by Patarin. Next, we present an attack against a randomised variant of the Matsumoto-Imai cryptosystem, called PMI. This scheme has recently been proposed by Ding, and according to the author, it resists all previously known attacks. We believe that differential cryptanalysis is a general and powerful method that can give additional insight on most multivariate schemes proposed so far.

1 Introduction

The design of efficient and secure cryptosystems is a hard task. Many alternatives to the traditional public key cryptosystems (RSA, ElGamal) have been proposed so far but few of them are considered secure. An interesting line of research is based on multivariate quadratic polynomials over a finite field. This line of research has been initiated by Matsumoto and Imai [12]. These systems are attractive since the underlying problem is known to be NP-complete and the decryption algorithm is more efficient than the RSA algorithm.

The original cryptosystem of Matsumoto and Imai (MI or C^*) has been broken by Patarin [13] who has also proposed various techniques that protect against this attack [15, 14]. A generalisation of MI, called Hidden Field Equations (HFE) [17], has higher security, but it has nevertheless been broken by Kipnis and Shamir [11]. More efficient attacks were proposed by Courtois *et al.* in [5, 6] and culminated with Faugère and Joux attack and the use of Gröbner bases in [9].

Variants of the original MI scheme remain interesting because they achieve better performance than variants of HFE. The main variants of MI that resist the attack by Patarin are on one hand, the Minus method which consists in discarding a few polynomials in the public key, and on the other hand the Minus-Plus method, which proposes to discard some polynomials and to add a few variables. These methods use *external* perturbation of the MI scheme, since variables are removed after the application of the exponentiation function.

Recently, Ding [7] proposed a new variant of the MI cryptosystem using some *internal* perturbation, which occurs before applying the exponentiation function. He quickly analyses its proposal against all known attacks on multivariate schemes, and claims that it is immune against such attacks. The new scheme is nearly as efficient as the original MI and the author gives some arguments in order to show that its scheme, called Perturbated MI (PMI), is a more secure extension than the MI Minus and MI Minus-Plus method.

1.1 Our Results

In this paper, we describe a new technique which is extremely powerful and that could presumably be used to break other multivariate schemes. In order to illustrate the power and generality of this method, we first propose a new attack on the original MI scheme and next describe how it can be used to mount an attack against the PMI cryptosystem.

The key point of our attack is that in the case of quadratic polynomials, the differential of the public key is a linear map and its kernel or its rank can be analysed to get some information on the secret key. For example, in the PMI scheme, we show that the dimension of the kernel can be used to identify elements that cancel the perturbation. In fact, we design a one-sided error recogniser for the language of elements that are not in the kernel of the perturbation. From this test algorithm, we design two algorithms to reconstruct the kernel. These algorithms are of independent interest. With the first method, the complexity of the attack is a precomputation of order $O(nq^{3r} + n^6 q^r)$, which can be upperbounded by 2^{49} with the proposed parameters in [7], and $O(n^3 \times q^r \times q^{\gcd(\ell,n)})$, which is of order 2^{36} binary operations. Finally, this attack works for scheme over finite fields of characteristic 2 which are the main structure for efficiency reasons and for MI and PMI this is always the case as we will see.

In the case of the original MI cryptosystem, we use elements in the kernel of the transpose of the differential in order to propose a new attack. We actually prove a bilinear relation between the ciphertext and the kernel vector. Thus, the kernel allows to recover the plaintext by solving a linear system.

1.2 Related Works

Differentials have already been successfully applied to break multivariate schemes such as the Minus transformation of the original Matsumoto-Imai, or the SFLASH signature scheme or the "2R" scheme proposed by Patarin [14, 16, 10, 8]. Our work gives a better insight by bringing a systematic use of the geometric properties of the differential.

1.3 Organisation of the paper

In section 2 of this paper, we describe the MI and PMI cryptosystems. Then, in section 3 we recall Patarin's attack on the original MI scheme. Next in section 4, we describe our attack on the PMI scheme and some experimental results. Finally, in section 5, we show a new attack on the original MI scheme.

2 Description of the MI and PMI schemes

2.1 The Matsumoto-Imai cryptosystem

This scheme is based on the following fact : over the finite field $\mathbb{F}_{q^n}$, the function $F : x \mapsto x^{q^\ell+1}$ is a permutation, when $\gcd(q^\ell + 1, q^n - 1) = 1$. Therefore, we can fix q to be a power of 2 so that $\mathbb{F}_{q^n}$ is of characteristic two [1]. Its inverse is $x \mapsto x^h$ where h is the inverse of $q^\ell + 1$ in $\mathbb{Z}_{q^n-1}$. Therefore, for any isomorphism π from the vector space of $\mathbb{F}_{q^n}$ to the n-dimensional vector space $(\mathbb{F}_q)^n$, the function $\boldsymbol{F} = \pi \circ F \circ \pi^{-1}$ is a bijective system of multivariate quadratic functions since F can be viewed as the product of two linear maps $x \mapsto x^{q^\ell}$ and $x \mapsto x$.

The scheme described by Matsumoto and Imai in 1988 [12] generates $\boldsymbol{S}$ and $\boldsymbol{T}$, two secret affine bijections of $(\mathbb{F}_q)^n$ to mask the system $\boldsymbol{F}$. The system $\boldsymbol{E} = \boldsymbol{T} \circ \boldsymbol{F} \circ \boldsymbol{S}$ is also a system of multivariate quadratic equations and represents the public key. Patarin showed in 1995 [13] that the public key has a special form which allows to invert the function.

[1] Indeed, if q is odd, we have $\gcd(q^\ell + 1, q^n - 1) \geq 2$ and since q is a prime power, it is always a power of 2 and the characteristic of $\mathbb{F}_{q^n}$ is 2

2.2 The PMI cryptosystem

Recently at PKC '04, Ding proposed a randomised variant of MI, called PMI [7]. Let $\boldsymbol{R} : (\mathbb{F}_q)^n \to (\mathbb{F}_q)^r$ a secret linear function of small rank ($r \ll n$) and $\boldsymbol{H}$ a secret quadratic system composed of n quadratic equations over r variables. The PMI public key is the system $\boldsymbol{E'}$ defined by $\boldsymbol{E'} = \boldsymbol{T} \circ (\boldsymbol{F} + \ \boldsymbol{H} \circ \boldsymbol{R}) \circ \boldsymbol{S}$. The public key can also be written as $\boldsymbol{E'} = \boldsymbol{T} \circ \boldsymbol{F} \circ \boldsymbol{S} + \boldsymbol{T} \circ \boldsymbol{H} \circ \boldsymbol{R} \circ \boldsymbol{S}$ due to the linearity of $\boldsymbol{T}$. Consequently, the PMI scheme can be seen as the MI scheme $\boldsymbol{E}$ plus a random-looking quadratic term $\boldsymbol{T} \circ \boldsymbol{H} \circ \boldsymbol{R} \circ \boldsymbol{S}$. Since there is no trapdoor to invert $\boldsymbol{H}$ or to separate the MI term and the random term, we need to store all the inputs and the outputs of the $\boldsymbol{H}$ function. Let P be the set of points which consist of pairs $(\boldsymbol{\lambda}, \boldsymbol{\mu})$, where $\boldsymbol{\lambda}$ is a point that belongs to the image of $\boldsymbol{H}$, and $\boldsymbol{\mu}$ is the set of pre-images of $\boldsymbol{\lambda}$ under $\boldsymbol{H}$. The set P contains q^r points. The secret key includes the set of linear functions $\boldsymbol{R}$, the set P, and the two affine bijections $\boldsymbol{S}$ and $\boldsymbol{T}$.

The secret key allows to invert $\boldsymbol{E'}$ if one can make exhaustive search over the q^r values of P and so r must be small. More precisely, given a ciphertext $\boldsymbol{y}$, the decryption process inverses the affine bijection $\boldsymbol{T}$ and recovers $\boldsymbol{y'}$. Then, all elements $(\boldsymbol{\lambda}, \boldsymbol{\mu})$ in P can be tried one-by-one and $\boldsymbol{y'_\lambda} = \boldsymbol{F}^{-1}(\boldsymbol{y'} + \boldsymbol{\lambda})$ is computed. Next, if $\boldsymbol{H}(\boldsymbol{y'_\lambda})$ is not equal to $\boldsymbol{\mu}$, we try the next point in P, otherwise, we compute $\boldsymbol{x_\lambda}$ by $\boldsymbol{S}^{-1}(\boldsymbol{y'_\lambda})$. If we have only one solution, we get the plaintext, otherwise, we use some added redundancy in the plaintext in order to uniquely recover it.

In his description of PMI [7], Ding analyses all known attack such as algebraic attacks of Patarin [13], Kipnis and Shamir [11], or XL attacks [4] and the attack on MI Minus of Patarin, Goubin and Courtois [16].

He also proposes a practical implementation with $q = 2$, $n = 136$, $r = 6$ and $F(x) = x^{2^{5\times 8}+1}$. He claims that the security level for this choice of parameters is 2^{136}. The value ℓ has been chosen with a special form, such that $\gcd(2^n - 1, 2^\ell - 1) = 2^{\gcd(n,\ell)} - 1 = 2^{\gcd(136, 5\times 8)} - 1 = 2^8 - 1$. This special form allows to perform more efficient encryption and decryption using lookup tables for the multiplications in the finite field.

In this paper, we apply differential cryptanalysis to the PMI scheme, and we show that the special form of the exponent in the practical system proposed by Ding allows more efficient attack than the attack in the generic case where $\gcd(\ell, n) = 1$.

3 Patarin's Attack on the MI cryptosystem

Our attack against the PMI cryptosystem is a probabilistic reduction to Patarin's attack on the MI scheme. Therefore, prior the description of our attack, we recall Patarin's attack. While we also propose an alternative attack to the MI scheme in section 5, we present Patarin attack since it is easier to understand. Both his attack and our attack do not recover the secret key but finds a linear system which can be solved to recover the plaintext corresponding to a given ciphertext.

Let $\boldsymbol{x} \in (\mathbb{F}_q)^n$ a plaintext and $\boldsymbol{y} \in (\mathbb{F}_q)^n$ the corresponding ciphertext. The main idea of Patarin attack is to find several bilinear relations in the $\boldsymbol{x}$ and $\boldsymbol{y}$ coordinates. Using plaintext/ciphertext pairs $(\boldsymbol{x}, \boldsymbol{y})$, it is possible to recover the coefficients of the relations by solving a linear system. Finally, knowing these coefficients and a given ciphertext, it is possible to decrypt $\boldsymbol{y}$ by solving a linear system.

Let us define $a = \pi^{-1}(\boldsymbol{S}(\boldsymbol{x}))$ and $b = \pi^{-1}(\boldsymbol{T}^{-1}(\boldsymbol{y}))$. Consequently, $F(a) = b$ or $b = a^{q^\ell+1}$. By raising each member of the last equation to the power $q^\ell - 1$ and by multiplying each one by ab, we get

$$ab^{q^\ell} = a^{q^{2\ell}} b \tag{1}$$

which holds over the finite field $\mathbb{F}_{q^n}$. We can rewrite this equation by $B(a, b) = 0$ where $B(a, b) = a \cdot b^{q^\ell} - a^{q^{2\ell}} \cdot b$. If we represent equation (1) in $(\mathbb{F}_q)^n$, we get n bilinear equations in the n coordinates of $\boldsymbol{a}$ and of $\boldsymbol{b}$. As $\boldsymbol{a}$ and $\boldsymbol{b}$ are affine transformations of $\boldsymbol{x}$ and $\boldsymbol{y}$ via the secret affine bijections $\boldsymbol{S}$ and $\boldsymbol{T}$, the n bilinear expressions in $\boldsymbol{a}$ and $\boldsymbol{b}$, may also be written as n bilinear expressions in $\boldsymbol{x}$ and $\boldsymbol{y}$. Each expression can be written as $\sum_{i=1}^n \sum_{j=1}^n \beta_{i,j} x_i y_j + \sum_{i=1}^n \beta_{i,0} x_i + \sum_{j=1}^n \beta_{0,j} y_j + \beta_{0,0} = 0$.

For each plaintext/ciphertext pair $(\boldsymbol{x}, \boldsymbol{y})$, the equation above, where all the $\beta_{i,j}$ are the $(n+1)^2$ unknowns, has at least the n solutions described by the n bilinear expressions deduced from equation (1). Therefore, using $\mathcal{O}((n+1)^2)$ plaintext/ciphertext pairs, solving the resulting system of $\mathcal{O}((n+1)^2)$ equations in the $(n+1)^2$ unknowns $\beta_{i,j}$ will recover the n bilinear expressions.

Finally, given a ciphertext $\boldsymbol{y}$ to decrypt, these n equations will give us n linear equations in the coefficients of $\boldsymbol{x}$. Unfortunately, all these equations are not independent. The solutions of this system correspond to the solutions of (1). There are $q^{\gcd(n,\ell)}$ such solutions, as shown by Patarin: let us consider the equation (1) where the unknown is a. A ciphertext $\boldsymbol{y}$ fixes a unique b value. One solution is $a = 0$. If $a \neq 0$ (and so $b \neq 0$) the equation can be written as

$$a^{q^{2\ell}-1} = b^{q^\ell -1}$$

We can write $q^{2\ell} - 1$ as $(q^\ell + 1)(q^\ell - 1)$ and take the inverse of $q^\ell + 1$ modulo $q^n - 1$ since by assumption F is a permutation. Consequently, the equation becomes $a^{q^\ell - 1} = b^{h\times(q^\ell - 1)} = b'$ where h is the inverse of $q^\ell + 1$. This last equation has exactly $\gcd(q^\ell - 1, q^n - 1) = q^{\gcd(\ell,n)} - 1$ solutions as shown in appendix A since the right solution is one solution.

As a consequence, the solution that we are looking for is a particular vector of the kernel of some system related to the original system, and the second member of the equation [3, p. 59]. Therefore, we compute the kernel of the system matrix which is of dimension $\gcd(n, \ell)$. Next, we perform an exhaustive search in $q^{\gcd(n,\ell)} - 1$ coefficients of the kernel vector, in order to recover the correct value $\boldsymbol{x}$.

In section 5, we propose a new differential attack on the MI scheme by studying the kernel of the transpose of the differential of the public key. We show that there exist n bilinear forms between a ciphertext $\boldsymbol{E(k)}$ and the vector $\boldsymbol{f_k}$ that generates the kernel of the transpose of the differential, which is of dimension 1 if $\gcd(\ell, n) = 1$. Then, given a ciphertext, we are able to reconstruct the vector $\boldsymbol{f_k}$ since the n bilinear forms are independent as there is a unique solution for the n bilinear forms. Finally, since the vector $\boldsymbol{f_k}$ is in the kernel of the transpose of the differential and that this map is linear in $\boldsymbol{k}$, we can solve n linear equations in the $\boldsymbol{k}$ variables of n coordinates. We refer the reader to section 5 for details.

4 Cryptanalysis of the PMI cryptosystem

4.1 Overview of the attack

Let us recall the notations : $\boldsymbol{F}$ is the system of quadratic equations corresponding to the internal function of the MI cryptosystem, $\boldsymbol{E} = \boldsymbol{T} \circ \boldsymbol{F} \circ \boldsymbol{S}$ the public key of MI, and $\boldsymbol{E'} = \boldsymbol{E} + \boldsymbol{T} \circ \boldsymbol{H} \circ \boldsymbol{R} \circ \boldsymbol{S}$ the public key of PMI.

Our attack is based on the following remark: the PMI scheme is a noisy MI cryptosystem. We find the linear space $\mathcal{K}$ that cancels the noise, and apply an attack of MI to the restriction of PMI to this linear space.

More precisely, we define the linear space $\mathcal{K}$ as follows: it is the kernel of the linear part of the affine function $\boldsymbol{R} \circ \boldsymbol{S}$. The space $\mathcal{K}$ is of dimension $\dim(\ker \boldsymbol{R}) = n - r$ because $\boldsymbol{S}$ is a bijection and $\mathrm{rank}(\boldsymbol{R}) = r$. If we are able to compute $\mathcal{K}$, then we can apply the attacks against MI (either Patarin's attack or our attack described in section 5) to the PMI cryptosystem restrict to elements of one of the q^r affine spaces that are parallel to $\mathcal{K}$. When restrict to one of these affine spaces, the public key of PMI is exactly $\boldsymbol{E}$ translated by a constant. The attack of PMI amounts to q^r attacks against MI (this is feasible because q^r must be of moderate size to allow fast decryption). A ciphertext is decrypted by applying the attack to the affine space that contains its corresponding plaintext.

In order to recover the space $\mathcal{K}$, we devise an efficient test algorithm that can spot that a given vector $\boldsymbol{k}$ does not belong to $\mathcal{K}$. The information used in this test is the dimension of the kernel of the linear part of the differential of the public key.

4.2 The dimension of the kernel of the differential

For any function $\boldsymbol{G} : (\mathbb{F}_q)^n \to (\mathbb{F}_q)^m$, let us consider its differential $\boldsymbol{dG_k}(x) = \boldsymbol{G}(\boldsymbol{x}+\boldsymbol{k}) - \boldsymbol{G}(\boldsymbol{x})$. Because $\boldsymbol{G}$ is a quadratic function, its differential is an affine function. Let us consider $\boldsymbol{L_{G,k}}(x) = \boldsymbol{dG_k}(x) - \boldsymbol{dG_k}(0)$ the linear part of the differential. In fact, it is a bilinear function that can also be defined by $\boldsymbol{L_{G,k}}(x) = \boldsymbol{B_G}(\boldsymbol{x},\boldsymbol{k}) = \boldsymbol{G}(\boldsymbol{x}+\boldsymbol{k}) - \boldsymbol{G}(\boldsymbol{x}) - \boldsymbol{G}(\boldsymbol{k}) + \boldsymbol{G}(\boldsymbol{0})$, and is also called the polar form. We are interested in $\dim(\ker \boldsymbol{L_{G,k}})$ when $\boldsymbol{G}$ is the public key of the cryptosystem.

Property 1. Let $\boldsymbol{k}$ and $\boldsymbol{k'}$ be elements of $(\mathbb{F}_q)^n$, and $\boldsymbol{G}$ and $\boldsymbol{G'}$ be systems of quadratic equations, and $\boldsymbol{S}$ and $\boldsymbol{T}$ be affine bijections. The following properties hold: $\boldsymbol{L_{G,k+k'}} = \boldsymbol{L_{G,k}} + \boldsymbol{L_{G,k'}}$, $\boldsymbol{L_{G+G',k}} = \boldsymbol{L_{G,k}} + \boldsymbol{L_{G',k}}$, $\boldsymbol{L_{T\circ G\circ S,k}} = \boldsymbol{T} \circ \boldsymbol{L_{G,S(k)}} \circ \boldsymbol{S} + \boldsymbol{T} \circ \boldsymbol{G} \circ \boldsymbol{S}(\boldsymbol{0}) - \boldsymbol{T} \circ \boldsymbol{G}(\boldsymbol{0})$, and $\boldsymbol{L_{G,0}} = \boldsymbol{0}$.

Lemma 1. *If $\boldsymbol{E}$ is the public key of a MI system over $\mathbb{F}_q$ of characteristic 2, of dimension n and exponent $q^\ell + 1$, then* $\dim(\ker \boldsymbol{L_{E,k}}) = \gcd(\ell, n)$.

First, $\dim\ker(\boldsymbol{L_{E,k}}) = \dim\ker(\boldsymbol{L_{F,k}})$, because $\boldsymbol{T}$ and $\boldsymbol{S}$ are bijections.

Let us define $\boldsymbol{x} = \pi(x)$ and $\boldsymbol{k} = \pi(k)$. If $\boldsymbol{F}$ is the internal function of the MI cryptosystem, then $\boldsymbol{B_F}(\boldsymbol{x},\boldsymbol{k})$ is equal to $\pi(x^{q^\ell} \cdot k + x \cdot k^{q^\ell})$. A vector $\boldsymbol{x} \neq \boldsymbol{0}$ of $(\mathbb{F}_q)^n$ is in the kernel of $\boldsymbol{L_{F,k}}$ if and only if $x^{q^\ell} \cdot k + x \cdot k^{q^\ell} = 0$. This last equation can be written as $x^{q^\ell+1} \cdot \left(\frac{k}{x} + \left(\frac{k}{x}\right)^{q^\ell} \right) = 0$.

Since $x \neq 0$, if we denote k/x by X, then the previous equation is $X + X^{q^\ell} = 0$ in the finite field $\mathbb{F}_{q^n}$. If $X \neq 0$ ($k \neq 0$), then the equation becomes $X^{q^\ell - 1} = 1$ in a finite field of characteristic 2. Since $X = 1$ is solution, there is at least one solution. As a consequence, there are $q^{\gcd(\ell,n)} - 1$ solutions according to the results in appendix A, and therefore $\dim(\ker \boldsymbol{L_{E,k}}) = \gcd(\ell, n)$.

Note that $X = 1$ is always a solution, that means $x = k$, and therefore $\boldsymbol{k}$ is always in the kernel. There are no other solutions when $\gcd(\ell, n) = 1$.

Lemma 2. *If $\boldsymbol{E'}$ is the public key of the PMI cryptosystem and $\boldsymbol{k} \in \mathcal{K}$, then* $\dim(\ker \boldsymbol{L}_{\boldsymbol{E'},\boldsymbol{k}}) = \gcd(\ell, n)$.

We prove that if $\boldsymbol{k} \in \mathcal{K}$, then $\boldsymbol{L}_{\boldsymbol{E'},\boldsymbol{k}} = \boldsymbol{L}_{\boldsymbol{E},\boldsymbol{k}}$. First we notice that $\boldsymbol{k} \in \mathcal{K}$ is equivalent to $\boldsymbol{R} \circ \boldsymbol{S}(\boldsymbol{k}) = \boldsymbol{R} \circ \boldsymbol{S}(\boldsymbol{0})$.

Then we compute $\boldsymbol{L}_{\boldsymbol{E'},\boldsymbol{k}}(\boldsymbol{x}) - \boldsymbol{L}_{\boldsymbol{E},\boldsymbol{k}}(\boldsymbol{x}) = \boldsymbol{L}_{\boldsymbol{T}\circ\boldsymbol{H}\circ\boldsymbol{R}\circ\boldsymbol{S},\boldsymbol{k}}(\boldsymbol{x}) = \boldsymbol{T} \circ \boldsymbol{H} \circ \boldsymbol{R} \circ \boldsymbol{S}(\boldsymbol{x}+\boldsymbol{k}) - \boldsymbol{T} \circ \boldsymbol{H} \circ \boldsymbol{R} \circ \boldsymbol{S}(\boldsymbol{x}) - \boldsymbol{T} \circ \boldsymbol{H} \circ \boldsymbol{R} \circ \boldsymbol{S}(\boldsymbol{k}) + \boldsymbol{T} \circ \boldsymbol{H} \circ \boldsymbol{R} \circ \boldsymbol{S}(\boldsymbol{0})$, therefore $\boldsymbol{T}^{-1}(\boldsymbol{L}_{\boldsymbol{E'},\boldsymbol{k}}(\boldsymbol{x}) - \boldsymbol{L}_{\boldsymbol{E},\boldsymbol{k}}(\boldsymbol{x})) = \boldsymbol{H}(\boldsymbol{R} \circ \boldsymbol{S}(\boldsymbol{x}+\boldsymbol{k})) - \boldsymbol{H}(\boldsymbol{R} \circ \boldsymbol{S}(\boldsymbol{x})) - \boldsymbol{H}(\boldsymbol{R} \circ \boldsymbol{S}(\boldsymbol{k})) + \boldsymbol{H}(\boldsymbol{R} \circ \boldsymbol{S}(\boldsymbol{0})) = \boldsymbol{0}$, which means that $\boldsymbol{T}^{-1} \circ \boldsymbol{L}_{\boldsymbol{E'},\boldsymbol{k}} = \boldsymbol{T}^{-1} \circ \boldsymbol{L}_{\boldsymbol{E},\boldsymbol{k}}$. Therefore $\dim(\ker \boldsymbol{L}_{\boldsymbol{E'},\boldsymbol{k}}) = \dim(\ker(\boldsymbol{T}^{-1} \circ \boldsymbol{L}_{\boldsymbol{E'},\boldsymbol{k}})) = \dim(\ker(\boldsymbol{T}^{-1} \circ \boldsymbol{L}_{\boldsymbol{E},\boldsymbol{k}})) = \dim(\ker \boldsymbol{L}_{\boldsymbol{E},\boldsymbol{k}})$.

Lemma 3. *If $\boldsymbol{E'}$ is the public key of the PMI cryptosystem and $\boldsymbol{k} \notin \mathcal{K}$, then often* $\dim(\ker \boldsymbol{L}_{\boldsymbol{E'},\boldsymbol{k}}) \neq \gcd(\ell, n)$.

As before, $\boldsymbol{L}_{\boldsymbol{E'},\boldsymbol{k}}$ is the sum of $\boldsymbol{L}_{\boldsymbol{E},\boldsymbol{k}}$ and $\boldsymbol{L}_{\boldsymbol{T}\circ\boldsymbol{H}\circ\boldsymbol{R}\circ\boldsymbol{S},\boldsymbol{k}}$. However, when, $\boldsymbol{k} \notin \mathcal{K}$, the second linear application is not null. The argument behind lemma 3 is that $\boldsymbol{L}_{\boldsymbol{T}\circ\boldsymbol{H}\circ\boldsymbol{R}\circ\boldsymbol{S},\boldsymbol{k}}$ is a random-looking linear application, and therefore the dimension of the kernel of the sum $\boldsymbol{L}_{\boldsymbol{E'},\boldsymbol{k}}$ follows the distribution of the dimension of the kernel of random linear maps.

In fact, it is slightly more complicated, because $\boldsymbol{k}$ is always in the kernel of $\boldsymbol{L}_{\boldsymbol{T}\circ\boldsymbol{H}\circ\boldsymbol{R}\circ\boldsymbol{S},\boldsymbol{k}}$, and therefore also in the kernel of $\boldsymbol{L}_{\boldsymbol{E'},\boldsymbol{k}}$, whose dimension then is at least 1. Moreover, if $\gcd(\ell, n) > r$, then there are $\gcd(\ell, n) - r$ additional vectors in the kernel of $\boldsymbol{L}_{\boldsymbol{E'},\boldsymbol{k}}$, because $\ker(\boldsymbol{L}_{\boldsymbol{E},\boldsymbol{k}})$ of dimension $\gcd(\ell, n)$ and $\ker(\boldsymbol{R} \circ \boldsymbol{S})$ of dimension $n - r$ in a space of dimension n have an intersection of dimension at least $\gcd(\ell, n) - r$. In the case of the practical scheme proposed by Ding where $\gcd(\ell, n) = 8$ and $r = 6$, we can deduce that $\dim(\ker(\boldsymbol{L}_{\boldsymbol{E'},\boldsymbol{k}})) \geq 3$.

Lemma 3 can be verified experimentally, as shown in table 1.

As a consequence of the lemmas, we get the following corollary.

Corollary 1. *If $\boldsymbol{E'}$ is the public key of the PMI cryptosystem and if* $\dim(\ker \boldsymbol{L}_{\boldsymbol{E'},\boldsymbol{k}}) \neq \gcd(\ell, n)$*, then $\boldsymbol{k} \notin \mathcal{K}$.*

In conclusion, we have now an efficient test to know if a vector is not in $\mathcal{K}$. We define $T(\boldsymbol{k})$ to be this test: $T(\boldsymbol{k}) = 1$ if $\dim(\ker \boldsymbol{L}_{\boldsymbol{E'},\boldsymbol{k}}) \neq \gcd(\ell, n)$, meaning that $\boldsymbol{k}$ is not in $\mathcal{K}$ with probability one, and $T(\boldsymbol{k}) = 0$ if $\dim(\ker \boldsymbol{L}_{\boldsymbol{E'},\boldsymbol{k}}) = \gcd(\ell, n)$, meaning that $\boldsymbol{k}$ can be in $\mathcal{K}$ or not. Now, we must transform this test into an algorithm for recovering $\mathcal{K}$.

Table 1. Experimental results for the probability distribution of $\dim(\ker(\boldsymbol{L}_{\boldsymbol{E}',\boldsymbol{k}}))$.

$\ell = 41$, $n = 137$ and $r = 6$

dimension	$\boldsymbol{k} \in \mathcal{K}$	$\boldsymbol{k} \notin \mathcal{K}$
1	1	≈ 0.59
> 1	0	≈ 0.41

$\ell = 40$, $n = 136$ and $r = 6$

dimension	$\boldsymbol{k} \in \mathcal{K}$	$\boldsymbol{k} \notin \mathcal{K}$
3	0	≈ 0.686
4	0	≈ 0.290
5	0	≈ 0.023
6	0	$\approx 5.10^{-4}$
7	0	$\approx 2.10^{-6}$
8	1	≈ 0
> 8	0	≈ 0

4.3 Recovering $\mathcal{K}$

We are looking for $\dim(\mathcal{K})$ independent vectors that generate $\mathcal{K}$. Let us define $\alpha = \Pr[T(\boldsymbol{k}) = 0]$ and $\beta = \Pr[\boldsymbol{k} \in \mathcal{K}] = q^{-r}$. The following table summarises the distribution of the values of T applied to a random $\boldsymbol{k}$.

	$\boldsymbol{k} \in \mathcal{K}$	$\boldsymbol{k} \notin \mathcal{K}$	
$T(\boldsymbol{k}) = 0$	β	$\alpha - \beta$	α
$T(\boldsymbol{k}) = 1$	0	$1 - \alpha$	$1 - \alpha$
	β	$1 - \beta$	

In the case where $\gcd(\ell, n) = 8$ we have $\alpha - \beta \ll \beta$ and therefore the test T has almost no false positives. In the case where $\gcd(\ell, n) = 1$ we have $\beta \ll \alpha$ and therefore the test T cannot give direct proof of membership of $\mathcal{K}$. A specific algorithm to recover $\mathcal{K}$ is needed.

The property we use is the linearity of $\mathcal{K}$: if $\boldsymbol{k}, \boldsymbol{k'} \in \mathcal{K}$, then $\boldsymbol{k}+\boldsymbol{k'} \in \mathcal{K}$. Two algorithms are described below. The first algorithm uses a statistical bias for $T(\boldsymbol{k} + \boldsymbol{k'})$. The second algorithm searches some large clique in a graph. A concrete attack of the PMI cryptosystem will use a mix of both techniques.

Technique 1. The key idea is: if for many different $\boldsymbol{k'} \in \mathcal{K}$, $\boldsymbol{k} + \boldsymbol{k'}$ is in $\mathcal{K}$, then $\boldsymbol{k}$ is always in $\mathcal{K}$. Therefore, if for many different $\boldsymbol{k'}$ such that $T(\boldsymbol{k'}) = 0$, $T(\boldsymbol{k} + \boldsymbol{k'}) = 0$, then $\boldsymbol{k}$ is in $\mathcal{K}$ with high probability.

We make the hypothesis that for any fixed value $\boldsymbol{k}$ and random value $\boldsymbol{k'}$ the probability that $T(\boldsymbol{k} + \boldsymbol{k'}) = 0$ is independent of the probability that $T(\boldsymbol{k'}) = 0$. Under this hypothesis, we compute $p(\boldsymbol{k}) = \Pr[T(\boldsymbol{k}+\boldsymbol{k'}) = 0 \,/\, T(\boldsymbol{k'}) = 0]$.

For a random $\boldsymbol{k}$, the value $\boldsymbol{k} + \boldsymbol{k'}$ when $T(\boldsymbol{k'}) = 0$ is uniformly distributed and $p(\boldsymbol{k}) = \alpha$. However, if $\boldsymbol{k} \in \mathcal{K}$, then one can write $p(\boldsymbol{k}) =$

$\Pr[\boldsymbol{k'} \in \mathcal{K} \,/\, T(\boldsymbol{k'}) = 0] + \Pr[\boldsymbol{k'} \notin \mathcal{K} \,/\, T(\boldsymbol{k'}) = 0]. \Pr[T(\boldsymbol{k}+\boldsymbol{k'}) = 0 \,/\, \boldsymbol{k}+\boldsymbol{k'} \notin \mathcal{K}] = \frac{\beta}{\alpha} + \frac{\alpha-\beta}{\alpha}\frac{\alpha-\beta}{1-\beta}$.

Under the hypothesis that $\beta \ll \alpha$, if $\boldsymbol{k} \in \mathcal{K}$ then $p(\boldsymbol{k})/\alpha = \frac{(1-\beta/\alpha)^2}{1-\beta} + \frac{\beta}{\alpha^2} \simeq 1 + \beta(\alpha^{-1} - 1)^2$. Therefore the difference between the values of $p(\boldsymbol{k})$ depending on whether $\boldsymbol{k} \in \mathcal{K}$ or not is of the order of $\alpha\beta$ and, by taking $N = 1/(\alpha\beta)^2$ elements $\boldsymbol{k'}$ such that $T(\boldsymbol{k'}) = 0$ and computing the average of $T(\boldsymbol{k}+\boldsymbol{k'})$, we can decide whether $\boldsymbol{k} \in \mathcal{K}$ or not. The complexity of this test is about β^{-2}.

We checked experimentally this hypothesis, for the parameters $\ell = 41$, $n = 137$ and $r = 6$. Testing if $p(\boldsymbol{k})/\alpha - 1 > \frac{1}{2}\beta(\alpha^{-1} - 1)^2$ is not sufficient to have an error-free test of membership of $\mathcal{K}$. However, testing if $p(\boldsymbol{k})/\alpha - 1 > \beta(\alpha^{-1} - 1)^2$ appear to be sufficient to detect about half of the members of $\mathcal{K}$.

Each value $\boldsymbol{k}$ has a probability q^{-r} of being in $\mathcal{K}$ and we need n distinct elements of $\mathcal{K}$. The whole complexity for finding $\mathcal{K}$ is nq^{3r}.

Technique 2. In this technique, we define a graph whose vertices are the elements $\boldsymbol{k}$ such that $T(\boldsymbol{k}) = 0$, *i.e.* elements that may be in the kernel. For each pair $(\boldsymbol{k}, \boldsymbol{k'})$ of vertices, we compute $T(\boldsymbol{k} + \boldsymbol{k'})$. If the result is 0, then we put an edge between these two vertices. All vertices such that $\boldsymbol{k} \in \mathcal{K}$ are connected, i.e. the elements of $\mathcal{K}$ are in a large clique.

In practice, we don't construct the whole graph. We construct its restriction to N vertices. We are looking for vertices that correspond to $n-r$ independent elements of $\mathcal{K}$. If $N > n/\beta$, it is likely that the graph contains such vertices. The clique containing the elements of $\mathcal{K}$ contains at least βN vertices. Under the same hypothesis as above, that the probability that $T(\boldsymbol{k} + \boldsymbol{k'}) = 0$ is independent of the probability that $T(\boldsymbol{k}) = 0$, this graph restricted to N vertices has αN^2 edges. Apart from the vertices that correspond to elements of $\mathcal{K}$, the edges are randomly distributed. General results on random graph [2] gives us that the expected number of vertex in the clique of maximal order in random graph of N vertex with a probability α between each edge is $\frac{2\log N}{\log 1/\alpha} + O(\log\log N)$. Therefore, if βN is significantly greater than $\frac{2\log N}{\log 1/\alpha}$, then there will be a unique large clique, that gives a basis of $\mathcal{K}$. When $\beta \ll \alpha$, this condition is equivalent to $N \approx \beta^{-1}\log\beta^{-1}$ and the whole complexity for finding $\mathcal{K}$ is $q^{2r}\log^2 q^r$.

However, although this technique seems to be better than the previous one, we do not know a max-clique algorithm that benefits from the fact that we have a random and dense graph which has a very large clique. In practice, as we said before, a concrete attack of the PMI cryptosystem will use a mix of technique 1 (to find some elements very likely to be members of $\mathcal{K}$) and technique 2 (to extract from them a large clique).

4.4 Recovering the plaintext

Assume we have correctly found the kernel $\mathcal{K}$. Now, we have to reconstruct a family of n bilinear equations in the $\boldsymbol{x}$ and $\boldsymbol{y}$ variable for each affine subspace parallel to $\mathcal{K}$. When this has been done, then for fixed $\boldsymbol{y}$ we can try to solve each system in the $\boldsymbol{x}$ unknowns and decide the correct solution using redundancy.

The question one may ask is whether we still find $n - \gcd(\ell, n)$ independent equations for each affine subspace. What can be said is that the original n equations from the MI scheme are clearly friend when $\boldsymbol{x}$ is restricted to a subspace. Accordingly the number of independent equations can only increase, which is in favour of the attacker. Now, given a ciphertext $\boldsymbol{y}$, its corresponding plaintext is in some subspace parallel to $\mathcal{K}$ and for such ciphertext, each family of equations allow to recover at least $n - \gcd(\ell, n)$ coordinates of $\boldsymbol{x}$. Finally, an exhaustive search allows us to find the missing coordinates in time $q^{\gcd(\ell,n)}$ as well as the correct subspace to choose.

5 Alternative attack against the MI scheme

In this section, we show a new attack against the MI scheme. We apply the same technique as in the PMI scheme. First of all, we compute the differential and next we study the kernel of the transpose of this application. In order to simplify the exposition of the attack, we assume in the following that $\gcd(\ell, n) = 1$ and $q = 2$.

5.1 Overview

As for Patarin's attack, this attack tries to find n bilinear forms in the ciphertext coordinates and in a vector related to the plaintext. Next, when a ciphertext is given, the n linear equations in the vector related to the plaintext allow us to recover this vector. Finally, since this vector is related to the plaintext by a linear system, we can easily decrypt.

More precisely, the attacks computes two bilinear systems, $C(\boldsymbol{x}, \boldsymbol{y})$ and $D(\boldsymbol{x}, \boldsymbol{y})$, such that for $\boldsymbol{f_k}^\top$ in the kernel of $\boldsymbol{L_{E,k}}^\top$ we have

$$C(\boldsymbol{E}(\boldsymbol{k}), \boldsymbol{f_k}) = 0 \text{ and } D(\boldsymbol{k}, \boldsymbol{f_k}) = 0$$

This allows to compute $\boldsymbol{k}$ from $\boldsymbol{E}(\boldsymbol{k})$.

5.2 Description

For the MI scheme, the differential can be written as

$$L_{F,k}(x) = x^{q^\ell} \cdot k + x \cdot k^{q^\ell} = k^{q^\ell+1} \cdot \left(\frac{x}{k} + \left(\frac{x}{k}\right)^{q^\ell} \right)$$

If we define the following three *linear* functions over $\mathbb{F}_{q^n}$: $\mu_k(x) = F(k) \cdot x$, where $F(k) = k^{q^\ell+1}$, $\psi(x) = x^{q^\ell} + x$ and $\theta_k(x) = \frac{x}{k}$, then

$$L_{F,k} = \mu_k \circ \psi \circ \theta_k$$

Let us define $\boldsymbol{\mu_k} = \pi \circ \mu_k \circ \pi^{-1}$, $\boldsymbol{\psi} = \pi \circ \psi \circ \pi^{-1}$ and $\boldsymbol{\theta_k} = \pi \circ \theta_k \circ \pi^{-1}$ for $k = \pi^{-1}(\boldsymbol{S}(\boldsymbol{k}))$. Therefore

$$\boldsymbol{L_{E,k}} = \boldsymbol{T} \circ \boldsymbol{\mu_k} \circ \boldsymbol{\psi} \circ \boldsymbol{\theta_k} \circ \boldsymbol{S}$$

where all terms are linear functions of $\mathbb{F}_q^n$. The matrix of $\boldsymbol{L_{F,k}}$ is a product of $n \times n$ matrices of $\mathbb{F}_q$.

Let $\boldsymbol{f_k}^\top$ be in the kernel of the transpose $\boldsymbol{L_{E,k}}^\top$. This means that the product $(\boldsymbol{f_k})(\boldsymbol{L_{E,k}})$ is the null vector $\mathbf{0}$, which is equivalent to

$$(\boldsymbol{f_k})(\boldsymbol{T}.\boldsymbol{\mu_k}.\boldsymbol{\psi}.\boldsymbol{\theta_k}.\boldsymbol{S}) = \mathbf{0}$$

where $\boldsymbol{f_k}$ is a n-dimensional row vector and $\boldsymbol{T}$, $\boldsymbol{\mu_k}$, $\boldsymbol{\theta_k}$, and $\boldsymbol{S}$ are $n \times n$ invertible matrices and $\boldsymbol{\psi}$ is a $n \times n$ matrix. Since $\boldsymbol{\theta}_k$ and $\boldsymbol{S}$ are one-to-one, this is equivalent to $(\boldsymbol{f_k})(\boldsymbol{T}.\boldsymbol{\mu_k}) \in \operatorname{Ker} \boldsymbol{\psi}^\top$, the application $\boldsymbol{\psi}^\top$ being independent of $\boldsymbol{k}$.

Recall that in the case where $\gcd(\ell, n) = 1$ the kernel of $\boldsymbol{L_{E,k}}$ is of dimension 1 and is generated by $\boldsymbol{k}$. The transpose $\boldsymbol{L_{E,k}}^\top$ also has a kernel of dimension 1. The kernel of $\boldsymbol{\psi}^\top$ is one-dimensional and independent of $\boldsymbol{k}$. Therefore if $q = 2$, $\operatorname{Ker} \boldsymbol{\psi}^\top = \{\mathbf{0}, \hat{\boldsymbol{f}}\}$ and the previous equation can be rewritten as $(\boldsymbol{f_k})(\boldsymbol{T}.\boldsymbol{\mu_k}) = (\hat{\boldsymbol{f}})$.

From $\mu_k(x) = F(k) \cdot x$, we deduce that $\boldsymbol{\mu_k}$ is linear in

$$F(k) = F(\pi^{-1}(\boldsymbol{S}(\boldsymbol{k}))) = \pi^{-1}(\boldsymbol{T}^{-1}(\boldsymbol{E}(\boldsymbol{k})))$$

i.e. linear in $\boldsymbol{E}(\boldsymbol{k})$, and therefore the equation $(\boldsymbol{f_k})(\boldsymbol{T}.\boldsymbol{\mu_k}) = (\hat{\boldsymbol{f}})$ is bilinear in $\boldsymbol{f_k}$ and $\boldsymbol{E}(\boldsymbol{k})$. Accordingly whenever a ciphertext $\boldsymbol{y} = \boldsymbol{E(k)}$ is given, the corresponding $\boldsymbol{f_k}$ can be found by solving a linear system.

Finally, as $(\boldsymbol{f_k})(\boldsymbol{L_{E,k}}) = 0$ and $\boldsymbol{L_{E,k}}$ is linear in the $\boldsymbol{k}$ variable we have again a bilinear relation between $\boldsymbol{k}$ and $\boldsymbol{f_k}$. Now, since $\boldsymbol{f_k}$ is known from $\boldsymbol{y}$, we get a system with n equations in n coordinates of the variable $\boldsymbol{k}$. This system has a kernel of dimension one, and consequently, we can easily decrypt.

Acknowledgement

This work is supported in part by the French government through X-Crypt and in part by the European Commission through the IST Programme under Contract IST-2002-507932 ECRYPT.

References

1. E. Bach and J. Shallit. *Algorithmic Number Theory.* MIT Press, 1996. Volume 1 - Efficient Algorithms.
2. B. Bollobás. *Random Graphs.* Cambridge University Press, 2001. Second Edition.
3. H. Cohen. *A Course in Computational Algebraic Number Theory.* Graduate Texts in Mathematics 138. Springer-Verlag, 1993.
4. N. Courtois, A. Klimov, J. Patarin, and A. Shamir. Efficient Algorithms for Solving Overdefined Systems of Multivariate Polynomial Equations. In *Eurocrypt '00*, LNCS 1807, pages 392–407. Springer-Verlag, 2000.
5. Nicolas T. Courtois. The security of Hidden Field Equations (HFE). In David Naccache, editor, *Proceedings of CT-RSA'01*, number 2020 in LNCS, pages 266–281. Springer-Verlag, 2001.
6. Nicolas T. Courtois, Magnus Daum, and Patrick Felke. On the security of HFE, HFEv- and Quartz. In Yvo Desmedt, editor, *Proceedings of Public Key Cryptography – PKC'03*, number 2567 in LNCS, pages 337–350. Springer-Verlag, 2003. Also available at `http://eprint.iacr.org/2002/138/`.
7. J. Ding. A New Variant of the Matsumoto-Imai Cryptosystem through Perturbation. In *PKC '04*, LNCS 2947, pages 305–318. Springer-Verlag, 2004.
8. Y. Ding-Feng, L. Kwok-Yan, and D. Zong-Duo. Cryptanalysis of "2R" Schemes. In *Crypto '99*, LNCS 1666, pages 315–325. Springer-Verlag, 1999.
9. J.-C. Faugère and A. Joux. Algebraic Cryptanalysis of Hidden Field Equation (HFE) Cryptosystems Using Gröbner Bases. In *Crypto '03*, LNCS 2729, pages 44–60. Springer-Verlag, 2003.
10. H. Gilbert and M. Minier. Cryptanalysis of SFLASH. In *Eurocrypt '02*, LNCS 2332, pages 288–298. Springer-Verlag, 2002.
11. A. Kipnis and A. Shamir. Cryptanalysis of the HFE Public Key Cryptosystem by Relinearization. In *Crypto '99*, LNCS 1666, pages 19–30. Springer-Verlag, 1999.
12. T. Matsumoto and H. Imai. Public Quadratic Polynomial-tuples for Efficient Signature-Verification and Message-Encryption. In *Eurocrypt '88*, LNCS 330, pages 419–453. Springer-Verlag, 1988.
13. J. Patarin. Cryptanalysis of the Matsumoto and Imai Public Key Scheme of Eurocrypt '98. In *Crypto '95*, LNCS 963, pages 248–261. Springer-Verlag, 1995.
14. J. Patarin. Assymetric Cryptography with a Hidden Monomial. In *Crypto '96*, LNCS 1109, pages 45–60. Springer-Verlag, 1996.
15. J. Patarin. Hidden Fields Equations (HFE) and Isomorphisms of Polynomial (IP): Two New Families of Asymmetric Algorithms. In *Eurocrypt '96*, LNCS 1070, pages 33–46. Springer-Verlag, 1996.
16. J. Patarin, L. Goubin, and N. Courtois. C^*_{-+} and HM: Variations around Two Schemes of T.Matsumoto and H.Imai. In *Asiacrypt '98*, LNCS 1514, pages 35–50. Springer-Verlag, 1998.
17. Jacques Patarin. Hidden field equations (HFE) and isomorphisms of polynomials (IP): two new families of asymmetric algorithms. In Ueli Maurer, editor, *Proceedings of Eurocrypt'96*, number 1070 in LNCS, pages 33–48. Springer-Verlag, 1996.

A Some useful mathematical results

Lemma 4. *For any integers q, i and n,* $\gcd(q^n - 1, q^i - 1) = q^{\gcd(n,i)} - 1$

Proof. Let $(r_k)_{k\geq 0}$ be the sequence of integers obtained by the Euclidean algorithm from $r_0 = n$ and $r_1 = i$. If k_0 is the largest integer such that $r_{k_0} \neq 0$, then $r_{k_0} = \gcd(n, i)$.

Similarly, let $(R_k)_{k\geq 0}$ be the sequence of polynomials obtained from the Euclidean algorithm from $R_0 = X^n - 1$ and $R_1 = X^i - 1$. We recall that n_1 is the largest integer such that $R_{n_1} = \gcd(X^n - 1, X^i - 1)$. We show by recurrence on n that for $0 \leq k \leq k_0 + 1$, $R_k = X^{r_k} - 1$. It is correct by assumption for $k = 0$ and $k = 1$. Assuming that $k \geq 2$ and $k \leq k_0 + 1$. Let us write $r_{k-2} = \alpha r_{k-1} + r_k$. Then,

$$\begin{aligned} X^{r_{k-2}} - 1 = {} & (X^{r_{k-1}} - 1)(X^{r_{k-2}-r_{k-1}} + X^{r_{k-2}-2r_{k-1}} + \cdots + X^{r_{k-2}-\alpha r_{k-1}}) \\ & + X^{r_k} - 1 \end{aligned}$$

Therefore, $X^{r_k} - 1$ is the remainder of the division of $R_{k-2} = X^{r_{k-2}} - 1$ by $R_{k-1} = X^{r_{k-1}} - 1$ since $r_k < r_{k-1}$. So, $R_{k_0+1} = X^0 - 1 = 0$ and $R_{k_0} \neq 0$. Consequently, $k_1 = k_0$ and $R_{k_1} = R_{k_0} = X^{r_{k_0}} - 1 = X^{\gcd(i,n)} - 1$. If we replace X by q, we get the lemma.

The following lemma is useful to exactly estimate the kernel dimension. We require exact value and not upper bounds on the number of solutions as done in [13].

Lemma 5. *In a finite field $\mathbb{F}_{q^n}$ with q^n elements, the equation $X^j = A$ has either 0 solution or $\gcd(j, q^n - 1)$ solutions.*

Proof. The multiplicative group of the finite field $\mathbb{F}_{q^n}$ has $q^n - 1$ elements. The simple case is when $\gcd(j, q^n - 1) = 1$. Therefore, j is invertible modulo $(q^n - 1)$ and we denote by h the inverse of j. Then, if we raise the equation $X^j = A$ to the power h, we get $X = X^{jh} = A^h = A'$, and so there is only one solution.

On the other hand, if $\gcd(j, q^n - 1) = d \neq 1$. Let $j' = j/d$, then $\gcd(j', q^n - 1) = 1$ and let h' be the inverse of j' modulo $q^n - 1$. We can raise the equation to the power h' and get $X^d = X^{jh'} = X^{j'dh'} = A^{h'} = A'$. This equation may have no solution if A' is not a d-th power of some value of $\mathbb{F}_{q^n}$. We now show that the equation $X^d = A'$ has d solutions when A' is a d-th power. We know that there is at least one solution which can be found by a randomised algorithm of Adleman, Manders and Miller [1]. The other solutions are obtained by multiplying the original solution by the d roots of unity. We finally explain why there are d d-th roots of unity. Since the multiplicative group of a finite field is a cyclic group, there is a primitive element g, that generates the whole group. Therefore, $g' = g^{\frac{q^n-1}{d}}$ is a d-th root of unity and for $0 \leq i < d$, g'^i ranges over the set of all roots. This completes the proof of the lemma.

If $j = q^i - 1$, then we can combine both lemmas. In a finite field $\mathbb{F}_{q^n}$ with q^n elements, the equation $X^{q^i-1} = A$ has either 0 solution or $q^{\gcd(i,n)} - 1$ solutions.