Security Analysis via Algebraic Attack Against "A New Encryption Scheme for Multivariate Quadratic System"

Yasuhiko Ikematsu[1] and Shuhei Nakamura[2]

[1] Institute of Mathematics for Industry, Kyushu University, 744 Motooka, Nishiku, Fukuoka 819-0395, Japan ikematsu@imi.kyushu-u.ac.jp
[2] Department of Liberal Arts and Basic Sciences, Nihon University, Japan nakamura.shuhei@nihon-u.ac.jp

Abstract. A Gröbner basis algorithm computes a good basis for an ideal of a polynomial ring and appears in various situations of cryptography. In particular, it has been used in the security analysis of multivariate public key cryptography (MPKC), and has been studied for a long time; however, it is far from a complete understanding. We consider the algebraic attack using a Gröbner basis algorithm for a new multivariate encryption scheme proposed by Jiahui Chen et al. at Theoretical Computer Science 2020. Their idea to construct a new scheme was to use the minus and plus modifiers to prevent known attacks, such as linearization attack. Moreover, they discussed to have a resistance to the algebraic attack using a Gröbner basis algorithm. However, in our experiments, the algebraic attack breaks their claimed 80- and 128-bit security parameters in reasonable times. It is necessary to understand whether their scheme can avoid such an attack by introducing a slight modification. In this paper, we theoretically describe why the algebraic attack breaks their scheme and give a precise complexity of the algebraic attack. As a result, we demonstrate that the algebraic attack can break the claimed 80- and 128-bit security parameters in the complexities of approximately 25 and 32 bits, respectively. Moreover, based on our complexity estimation of the algebraic attack, we conclude that Chen et al.'s scheme is not practical.

Keywords: Multivariate public key cryptography· Post-quantum cryptography· Encryption schemes· Gröbner basis algorithm

1 Introduction

Multivariate public key cryptography (MPKC) [10] is public key cryptography using multivariate polynomials over finite fields and is considered as one of the main candidates for post-quantum cryptography (PQC) [2]. It has been studied for approximately three decades since the MI scheme [23], which is the first multivariate encryption scheme, was proposed in 1988. An important ingredient of MPKC is the *central map* F. Namely, $F = (f_1, \ldots, f_m)$ is a map from $\mathbb{F}^n$ to

$\mathbb{F}^m$ consisting of m multivariate quadratic polynomials $f_1, \ldots, f_m$ in n variables $x_1, \ldots, x_n$ over a finite field $\mathbb{F}$ and has the property that the equation $F(x) = y$ for any $y \in \mathbb{F}^m$ can be solved easily. To hide the structure of F, we randomly choose two affine maps $S : \mathbb{F}^n \to \mathbb{F}^n$ and $T : \mathbb{F}^m \to \mathbb{F}^m$ and the public key is given by the quadratic polynomial map $P = T \circ F \circ S : \mathbb{F}^n \to \mathbb{F}^m$. The security of MPKC is based on the *MQ-Problem*, which asks for a solution to the system $P(x) = y$ of quadratic equations for a ciphertext $y \in \mathbb{F}^m$. From the fact that the MQ-Problem is proven to be an NP-hard problem even for quadratic polynomials over $\mathbb{F}_2$ [18] and there exist no polynomial-time quantum algorithms at the moment, MPKC is widely considered to have the potential to resist quantum computer attacks.

In the area of digital signatures, there exist practical multivariate signature schemes such as UOV [21] and Rainbow [12]. In fact, they appear in the third round of NIST PQC standardization [24] as Rainbow [13]. Here Rainbow is an efficient variant of the UOV scheme. However, it is considered to be difficult to construct a secure and efficient multivariate encryption scheme. In fact, the MI encryption scheme [23] was broken by Patarin in 1995 [25] using the linearization attack. Subsequently, Patarin proposed a multivariate encryption scheme extending the MI scheme, called the HFE scheme [26]. However, in a series of subsequent studies [3,8,11,14,17,19,22], it was found that the HFE scheme has a serious trade-off between efficiency and security. As a result, the original HFE scheme is not practical and requires some modifiers such as HFERP [20].

In 2020, Jiahui Chen et al. proposed a new multivariate encryption scheme [7] at Theoretical Computer Science. Their original central map $F = (f_1, \ldots, f_m)$ used in [7] is constructed so that it can easily obtain any preimage of F by reducing a system of linear equations. From such a property, any equation $F(x) = y$ can be easily solved. Because the original central map F is vulnerable to the linearization attack, the authors applied the minus and plus modifiers to the original central map F and constructed the public key P. By doing so, the paper [7] claimed that the scheme can avoid the linearization attack.

In this paper, we analyze the algebraic attack using a Gröbner basis algorithm for Chen et al.'s scheme. In our experiments, the algebraic attack using a Gröbner basis algorithm breaks their claimed 80- and 128-bit parameters in a few seconds. However, it is not clear why the Gröbner basis algorithm breaks their scheme. It is necessary to understand whether their scheme can avoid such an attack by introducing a slight modification, and we theoretically have to describe why the algebraic attack breaks their scheme. Thus, we consider a precise complexity estimation of the algebraic attack in this paper. In general, the complexity of the attack using a Gröbner basis algorithm is difficult to analyze in MPKC. Therefore, it is important for future development of MPKC to analyze the Gröbner basis algorithm against such a scheme in detail.

In our analysis, we focus on the property that the original central map F can easily obtain any preimage of F by reducing a system of linear equations. To be precise, we can recover many degree-one polynomials from the vector space spanned by the public key P even if minus and plus modifiers are applied. By substituting such degree-one polynomials into the equation $P(x) = y$ with a ciphertext y, we reduce the original MQ-problem $P(x) = y$ with m quadratic equations and n variables to a smaller MQ-problem $P'(x') = y'$. Thus, the complexity to solve the original MQ-problem $P(x) = y$ by a Gröbner basis algorithm is the same as that of the smaller MQ-problem $P'(x') = y'$, and we can obtain a precise complexity estimation of the algebraic attack using a Gröbner basis algorithm. In fact, we show that the algebraic attack can break the claimed 80- and 128-bit security parameters in the complexities of approximately 25 and 32 bits, respectively. Moreover, based on our complexity estimation of the algebraic attack, we conclude that Chen et al.'s scheme is not practical.

Our paper is organized as follows. In Section 2, we briefly recall the general construction of multivariate encryption schemes, Chen et al.'s encryption scheme and the algebraic attack. In Section 3, we propose our attack and give experimental results. In Section 4, we describe the security analysis for the algebraic attack in detail. Finally, we conclude our paper in Section 5.

2 Preliminaries

In this section, we describe the general construction of multivariate encryption schemes, Chen et al.'s encryption scheme [7] and the algebraic attack.

2.1 Multivariate public key cryptography (MPKC)

Let $\mathbb{F}$ be a finite field with q elements, and let n and m be positive integers. The public key of a multivariate public key cryptosystem consists of m multivariate quadratic polynomials $P = (p_1, \ldots, p_m)$ in n variables $x_1, \ldots, x_n$ over the finite field $\mathbb{F}$. Each polynomial $p_k(x_1, \ldots, x_n)$ $(1 \leq k \leq m)$ is in the form of

$$p_k(x_1, \ldots, x_n) = \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij} x_i x_j + \sum_{i=1}^{n} b_i x_i + c, \tag{1}$$

where $a_{ij}, b_i, c \in \mathbb{F}$. The security of multivariate public key schemes is based on the *MQ-Problem*, which asks for a solution of a given system of multivariate quadratic polynomials over the finite field $\mathbb{F}$. In fact, the MQ-Problem is proven to be an NP-hard problem even for quadratic polynomials over the binary field $\mathbb{F}_2$ [18].

We introduce the general construction of multivariate encryption schemes. First, the most important ingredient is a *central map* $F : \mathbb{F}^n \to \mathbb{F}^m$, which is an easily invertible quadratic map. Namely, we can compute a solution of $F(x) = y$ for any element $y \in \mathbb{F}^m$ with low complexity. Second, to hide the structure of the central map F, we randomly choose two invertible affine (or linear) maps $T : \mathbb{F}^m \to \mathbb{F}^m$ and $S : \mathbb{F}^n \to \mathbb{F}^n$. Then the *public key* is given by the composite

$$P = T \circ F \circ S : \mathbb{F}^n \to \mathbb{F}^m$$

and the *secret key* is given by $\{T, F, S\}$.

Encryption: For a plaintext $\alpha \in \mathbb{F}^n$, the ciphertext is given by $y = P(\alpha) \in \mathbb{F}^m$.

Decryption: For a given ciphertext $y \in \mathbb{F}^m$, one computes $\beta = T^{-1}(y), \gamma = F^{-1}(\beta)$ and $y' = S^{-1}(\gamma)$. Then y' is the plaintext of the ciphertext y.

2.2 Chen et al.'s Encryption Scheme

In this subsection, we briefly describe the construction of Chen et al.'s encryption scheme [7]. Let n be a positive integer. Assume that the characteristic of $\mathbb{F}$ is not 2. For an $(n+1) \times n$ matrix $C = (c_{i,j})_{i,j} \in \mathbb{F}^{(n+1)\times n}$, we define $n+1$ quadratic polynomials $f_{C,1}, \ldots, f_{C,n+1}$ in n variables $x_1, \ldots, x_n$ as follows:

$$\begin{aligned} f_{C,1} &= (x_1 - c_{1,1})^2 + (x_2 - c_{1,2})^2 + \cdots + (x_n - c_{1,n})^2, \\ &\vdots \\ f_{C,n+1} &= (x_1 - c_{n+1,1})^2 + (x_2 - c_{n+1,2})^2 + \cdots + (x_n - c_{n+1,n})^2. \end{aligned}$$

We recall that the polynomial map $F_C = (f_{C,1}, \ldots, f_{C,n+1}) : \mathbb{F}^n \to \mathbb{F}^{n+1}$ is an easily invertible quadratic map. For an element $y = (y_1, \ldots, y_{n+1}) \in \mathbb{F}^{n+1}$, we would like to solve the system of equations

$$f_{C,1}(x_1, \ldots, x_n) = y_1, \cdots, f_{C,n+1}(x_1, \ldots, x_n) = y_{n+1}. \tag{2}$$

It is clear that for any $1 \le i \le n$, we have

$$\begin{aligned} & y_{i+1} - y_i \\ &= f_{C,i+1}(x_1, \ldots, x_n) - f_{C,i}(x_1, \ldots, x_n) \\ &= 2(c_{i,1} - c_{i+1,1})x_1 + \cdots + 2(c_{i,n} - c_{i+1,n})x_n + \sum_{j=1}^{n}(c_{i+1,j}^2 - c_{i,j}^2). \end{aligned}$$

From this, we have the linear equations

$$\begin{pmatrix} y_2 - y_1 \\ y_3 - y_2 \\ \vdots \\ y_{n+1} - y_n \end{pmatrix} = C' \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} + \begin{pmatrix} \sum_{j=1}^n (c_{2,j}^2 - c_{1,j}^2) \\ \sum_{j=1}^n (c_{3,j}^2 - c_{2,j}^2) \\ \vdots \\ \sum_{j=1}^n (c_{n+1,j}^2 - c_{n,j}^2) \end{pmatrix}, \tag{3}$$

where $C' := (2c_{i,j} - 2c_{i+1,j})_{i,j} \in \mathbb{F}^{n \times n}$. Thus, if C' is invertible, then we can obtain the solution to (2) by solving the system of linear equations (3).

Now, we explain the key-generation of Chen et al.'s encryption scheme [7]. Let a and s be two positive integers and set $m := n + 1 - a + s$. First, randomly choose a matrix $C = (c_{i,j})_{i,j} \in \mathbb{F}^{(n+1)\times n}$ such that C' is invertible. Define the polynomials $F_C = (f_{C,1}, \ldots, f_{C,n+1})$ as above. Next, randomly choose quadratic polynomials $g_1, \ldots, g_s$ in n variables $x_1, \ldots, x_n$ over $\mathbb{F}$. Then the central map is given by

$$F = (f_{C,1}, \ldots, f_{C,n+1-a}, g_1, \ldots, g_s) : \mathbb{F}^n \to \mathbb{F}^m.$$

Next, randomly choose invertible affine maps $S : \mathbb{F}^n \to \mathbb{F}^n$ and $T : \mathbb{F}^m \to \mathbb{F}^m$. Then, the public key is the polynomial map $P := T \circ F \circ S$, and the secret key consists of $F_C = (f_{C,1}, \ldots, f_{C,n+1})$, S, and T. Note that the central map F is applied to the minus and plus modifiers, as $f_{C,n+1-a+1}, \ldots, f_{n+1}$ are removed and some quadratic polynomials $g_1, \ldots, g_s$ are added.

Encryption: For a plaintext $\alpha \in \mathbb{F}^n$, the ciphertext is given by $y = P(\alpha) \in \mathbb{F}^m$ by substituting into the public key P.

Decryption: For a given ciphertext $y \in \mathbb{F}^m$, first compute $\beta = (\beta_1, \ldots, \beta_m) := T^{-1}(y)$. Second, for an element $\beta' = (\beta'_1, \ldots, \beta'_a) \in \mathbb{F}^a$, solve the system of quadratic equations

$$\begin{aligned} f_{C,1}(x_1, \ldots, x_n) &= \beta_1, \\ &\vdots \\ f_{C,n+1-a}(x_1, \ldots, x_n) &= \beta_{n+1-a}, \\ f_{C,n+1-a+1}(x_1, \ldots, x_n) &= \beta'_1, \\ &\vdots \\ f_{C,n+1}(x_1, \ldots, x_n) &= \beta'_a, \end{aligned}$$

as explained above. Let $\gamma = (\gamma_1, \ldots, \gamma_n)$ be the solution. If we have

$$g_1(\gamma) = \beta_{n+1-a+1}, \ldots, g_s(\gamma) = \beta_m,$$

then $y' = S^{-1}(\gamma)$ is the plaintext of y; otherwise, re-choose another $\beta' = (\beta'_1, \ldots, \beta'_a) \in \mathbb{F}^a$.

2.3 Algebraic attack and previous analysis against Chen et al's scheme

In this subsection, we first explain the algebraic attack. Subsequently, we describe how the authors in [7] decided the concrete parameters.

For a public key $P = (p_1, \ldots, p_m)$ and a ciphertext $y = (y_1, \ldots, y_m)$, the *algebraic attack* finds the plaintext by solving the system of equations $p_1(x) = y_1, \ldots, p_m(x) = y_m$. In the HFE challenge presented by Patarin to cryptanalysis for HFE and its minus variant [25], HFE challenge 1 was solved by this attack in [17]. The algebraic attack uses Gröbner basis algorithms such as XL [31], F_4 [15] and F_5 [16] which compute a Gröbner basis of the ideal $\langle p_1 - y_1, \ldots, p_m - y_m \rangle$. The complexity of the Gröbner basis algorithm dominates that of the algebraic attack and is estimated by

$$\mathcal{O}\left(\binom{n + D_{reg}}{D_{reg}}^{\omega}\right),$$

Table 1. Selected parameters of Chen et al.'s scheme in [7] at 80- and 128-bit security levels. Here, q is the cardinality of the finite field $\mathbb{F}$.

	Asserted security level	(q, n, a, s, m)	Public key (KB)	Secret key (KB)
(A)	80-bit	$(3, 59, 10, 25, 75)$	134	53.9
(B)	128-bit	$(3, 83, 12, 27, 99)$	345	108

where n is the number of variables, D_{reg} is the degree of regularity of the input system $p_1(x) = y_1, \ldots, p_m(x) = y_m$, and $2 < \omega \leq 3$ is a linear algebra constant. If $\{p_1, \ldots, p_m\}$ is semi-regular [3], then the degree of regularity D_{reg} is given by the degree of the first term whose coefficient in the following power series is non-positive (see [1]):

$$\frac{\prod_{i=1}^{m}(1 - t^{\deg p_i})}{(1 - t)^n}.$$

Chen et al. show in [7] that the degree of regularity against their scheme without the minus and plus modifications is low; however, such modifiers can increase the degree of regularity. In fact, they showed in some experiments that the larger the parameter a is, the more the degree of regularity increases under the condition $a = s$. Here, a is the number of removed polynomials, and s is the number of added polynomials. Then, in order to select an 80-bit security parameter for $q = 3$ and $n = 59$, they concluded that the degree of regularity needs at least 8, and it is accomplished by $a, s \geq 8$. They performed experiments under the condition $a = s$; in such a condition, the decryption failure rate is very high; $q^{-s+a-1} = q^{-1}$. Thus, a is required to be moderately small compared with s. As a result, they selected two 80- and 128-bit security parameters (A) and (B) in [7], as shown in Table 1. Here, we mention that they did not perform experiments under the condition $a < s$.

3 Revisiting the algebraic attack against Chen et al.'s scheme

In this section, we revisit the algebraic attack against Chen et al.'s scheme. We first perform some experiments for the algebraic attack and confirm that the algebraic attack breaks 80 and 128-bit security parameters of their scheme in reasonable times. Second, we observe the property of the public key to see why the algebraic attack works. Finally, we apply the property to the algebraic attack and reconstruct the algebraic attack to make it theoretically easy to handle.

3.1 Experiments for Algebraic attack

In this subsection, we discuss the experiments conducted in this study and show that the algebraic attack using a Gröbner basis algorithm breaks the 80 and 128-bit security parameters of their scheme in reasonable times.

Table 2. Experimental results for the algebraic attack using the Gröbner basis algorithm F4 at two selected 80- and 128-bit parameters (A) and (B) of Chen et al.'s scheme in [7]

	Asserted security level	(q, n, a, s, m)	Experiments	
			Time (s)	Memory (MB)
(A)	80-bit	$(3, 59, 10, 25, 75)$	0.35	32.1
(B)	128-bit	$(3, 83, 12, 27, 99)$	1.07	81.6

Our experimental results are the average timing of 1000 experiments for each parameter in Table 1. All experiments were performed on a 3.2GHz Intel Core i7 CPU with Magma V2.24-4 [5]. The Gröbner basis algorithm we used was the F4 algorithm [15] with the graded reverse lexicographic monomial order, whose Magma command is "**GroebnerBasis**". In addition, to measure the memory, we used the Magma command "**GetMaximumMemoryUsage()**".

For example, Table 2 shows that the 80-bit parameter (A) is broken in 0.35 seconds with 32.1 megabytes. However, because it is not clear why the Gröbner basis algorithm breaks their parameters, we must theoretically describe it.

3.2 A key property of the public key *P*

Here, we observe a property of the public key P in Subsection 2.2.

The polynomial $f_{C,i+1} - f_{C,i}$ is denoted by l_i for each $1 \leq i \leq n - a$. This l_i is a polynomial of degree one from the definition of $f_{C,i}$. Moreover, the set $\{l_1, \dots, l_{n-a}\}$ is linearly independent, as C' is invertible. Denote by $\mathrm{Span}_{\mathbb{F}} F$ the subspace in $\mathbb{F}[x_1, \dots, x_n]$ generated by the polynomials

$$\{f_{C,1}, \dots, f_{C,n+1-a}, g_1, \dots, g_s\}$$

in the central map F over the finite field $\mathbb{F}$. We can easily show the following.

Lemma 1. *$\mathrm{Span}_{\mathbb{F}} F$ is generated by*

$$\{l_1, \dots, l_{n-a}, f_{C,1}, g_1, \dots, g_s\}.$$

Therefore, $\mathrm{Span}_{\mathbb{F}} F$ has a linearly independent set of $n - a$ degree-one polynomials and $s + 1$ quadratic polynomials.

Let $\mathrm{Span}_{\mathbb{F}} P$ be the subspace in $\mathbb{F}[x_1, \dots, x_n]$ generated by the polynomials $P = \{p_1, \dots, p_m\}$ over $\mathbb{F}$. It is clear that $\mathrm{Span}_{\mathbb{F}} P = \mathrm{Span}_{\mathbb{F}} F \circ S$ because $P = T \circ F \circ S$ and T is invertible. By Lemma 1, $\mathrm{Span}_{\mathbb{F}} F \circ S$ is generated by the set

$$\{l_1 \circ S, \dots, l_{n-a} \circ S, f_{C,1} \circ S, g_1 \circ S, \dots, g_s \circ S\},$$

which is linearly independent since S is invertible. As a result, we have the following theorem.

Theorem 1. *$\mathrm{Span}_{\mathbb{F}} P$ has a linearly independent set of $n - a$ degree-one polynomials and $s + 1$ quadratic polynomials.*

Therefore, one can generate a linearly independent set of $n - a$ degree-one polynomials from only the public key P.

3.3 Reduction to a smaller MQ problem

In this subsection, by using the result in Subsection 3.2, we see that solving the MQ problem $P(x) = y$ is essentially equivalent to solving a smaller MQ problem.

Let $y = (y_1, \ldots, y_m) \in \mathbb{F}^m$ be a ciphertext. We would like to solve the system of m quadratic equations

$$p_1(x_1, \ldots, x_n) = y_1, \ldots, p_m(x_1, \ldots, x_n) = y_m \tag{4}$$

in n variables $x_1, \ldots, x_n$ without the secret key $\{F_C, S, T\}$.

Step 1: We solve the system of linear equations in m variables $z_1, \ldots, z_m$;

$$\sum_{i=1}^{m} z_i \cdot \mathrm{Quad}(p_i) = 0, \tag{5}$$

where $\mathrm{Quad}(p_i)$ stands for the quadratic part of p_i. Here, the system $\sum_{i=1}^{m} z_i \cdot \mathrm{Quad}(p_i) = 0$ implies the linear equations arising from the relation that each coefficient of $\sum_{i=1}^{m} z_i \cdot \mathrm{Quad}(p_i)$ vanishes.

Step 2: If $a_1, \ldots, a_m$ be a solution of (5), then $\sum_{i=1}^{m} a_i p_i$ is a degree-one polynomial. By Theorem 1, the space of such degree-one polynomials is of dimension $n - a$. Thus, the solution space of this system (5) is of dimension $n - a$. Choose a basis $\mathbf{z}^{(1)}, \ldots, \mathbf{z}^{(n-a)} \in \mathbb{F}^m$ of the solution space. Then we can obtain $n - a$ linearly independent degree-one polynomials

$$\begin{aligned} r_1(x_1, \ldots, x_n) &:= \mathbf{z}_1^{(1)} p_1 + \cdots + \mathbf{z}_m^{(1)} p_m, \\ &\vdots \\ r_{n-a}(x_1, \ldots, x_n) &:= \mathbf{z}_1^{(n-a)} p_1 + \cdots + \mathbf{z}_m^{(n-a)} p_m, \end{aligned}$$

where $\mathbf{z}^{(k)} = (\mathbf{z}_1^{(k)}, \ldots, \mathbf{z}_m^{(k)})$ $(1 \leq k \leq n - a)$. Thus, $\mathrm{Span}_{\mathbb{F}} P$ is generated by $r_1, \ldots, r_{n-a}$ and other $s + 1$ quadratic polynomials.

Step 3: From (4), we have the linear equations

$$\begin{aligned} r_1(x_1, \ldots, x_n) &= \mathbf{z}_1^{(1)} \beta_1 + \cdots + \mathbf{z}_m^{(1)} \beta_m, \\ &\vdots \\ r_{n-a}(x_1, \ldots, x_n) &= \mathbf{z}_1^{(n-a)} \beta_1 + \cdots + \mathbf{z}_m^{(n-a)} \beta_m. \end{aligned} \tag{6}$$

By substituting the linear equation (6) into (4), we have a small system of $m - (n - a) = s + 1$ quadratic equations in $n - (n - a) = a$ variables.

As a result, the MQ problem $P(x) = y$, the system of m quadratic equations in n variables, associated with Chen et al.'s scheme, can be reduced to the smaller MQ problem with $s + 1$ quadratic equations in a variables. Algorithm 1 shows the attack stated in this subsection.

Algorithm 1: Reconstructed algebraic attack in 3.3

```
Input  : finite field $\mathbb{F}$, public key $P = (p_1, \dots, p_m)$, and ciphertext
         $y = (y_1, \dots, y_m) \in \mathbb{F}^m$,
Output: the plaintext $\alpha \in \mathbb{F}^n$ of the ciphertext $y$, namely, $P(\alpha) = y$.
Solve the system of linear equations $\sum_{i=1}^{m} z_i \cdot \mathrm{Quad}(p_i) = 0$;
$\mathbf{z}^{(1)}, \dots, \mathbf{z}^{(n-a)} \in \mathbb{F}^m \leftarrow$ a basis of the solution space;
$r_i \leftarrow \mathbf{z}_i^{(1)} p_1 + \cdots + \mathbf{z}_i^{(m)} p_m$ for $1 \leq i \leq n-a$;
$R_i \leftarrow \mathbf{z}_i^{(1)} y_1 + \cdots + \mathbf{z}_i^{(m)} y_m$ for $1 \leq i \leq n-a$;
Substitute $r_i - R_i = 0$ $(1 \leq i \leq n-a)$ into $P(x) - y = 0$ and solve it by a
  Gröbner basis algorithm ;
```

4 Security analysis for Chen et al.'s scheme

In this section, we describe the security analysis for the algebraic attack in detail. In Subsection 4.1, we provide the precise complexity estimation for the reconstructed algebraic attack in Subsection 3.3 and that for the original algebraic attack in Subsection 2.3. In Subsection 4.2, using the complexity estimation, we compute the actual security level for the selected parameters (A) and (B). Moreover, we consider a parameter satisfying 80-bit security against the algebraic attack and discuss its efficiency.

4.1 Theoretical description of the algebraic attack and complexity estimation

As stated in Subsection 3.3, the reconstructed algebraic attack solves the MQ problem with $s+1$ quadratic equations in a variables after some linear operations.

The complexity of the reconstructed algebraic attack is given as follows. The complexity of Step 1 is that of solving the linear system (5) with size m. Thus, it is $\mathcal{O}(m^\omega)$, where $2 < \omega \leq 3$ is a linear algebra constant. We can ignore the complexity of Step 2 because we must define the degree-one polynomials r_i. In Step 3, its complexity is dominated by solving a system of $(s+1)$ quadratic equations in a variables. We use a Gröbner basis algorithm to solve the system. Under the assumption that the system is semi-regular [3], the complexity of Step 3 is

$$\mathcal{O}\left(\binom{a + D_{reg}}{D_{reg}}^{\omega}\right) \tag{7}$$

at most. Note that the degree of regularity D_{reg} is the degree of the first term whose coefficient in the following power series is non-positive:

$$\frac{(1-t^2)^{s+1}}{(1-t)^a}.$$

As a result, the complexity of the reconstructed algebraic attack is

$$\mathcal{O}(m^\omega) + \mathcal{O}\left(\binom{a + D_{reg}}{D_{reg}}^{\omega}\right). \tag{8}$$

Table 3. Complexities for the algebraic attack using the Gröbner basis algorithm at two selected 80- and 128-bit parameters (A) and (B) of Chen et al.'s scheme in [7]

	Asserted security level	(q, n, a, s, m)	Actual security level Complexity (bits)
(A)	80-bit	$(3, 59, 10, 25, 75)$	24.50
(B)	128-bit	$(3, 83, 12, 27, 99)$	32.49

Note that the linear operations in the reconstructed algebraic attack are contained in the computation process of a Gröbner basis algorithm of the original algebraic attack in Subsection 2.3. Thus, the degree of regularity for the original algebraic attack is the same as that of the reconstructed algebraic attack because such linear operations do not affect the degree of regularity. Therefore, the complexity of solving the MQ-problem $P(x) = y$ by the original algebraic attack is given by

$$\mathcal{O}\left(\binom{a + D_{reg}}{D_{reg}}^{\omega}\right). \tag{9}$$

4.2 The actual security level for the selected 80- and 128-bit security parameters (A) and (B) in Table 1

The selected parameters (A) (resp. (B)) are $(q, n, a, s, m) = (3, 59, 10, 25, 75)$ (resp. $(3, 83, 12, 27, 99)$). Using the formula (9), the complexities of the attack against parameters (A) and (B) are $\binom{10+3}{3}^3 = 2^{24.5}$ and $\binom{12+4}{4}^3 = 2^{32.49}$, respectively. Here we took $\omega = 3$ as a linear algebra constant. Table 2 shows the complexity for the original algebraic attack in Subsection 2.3.

Finally, we consider a parameter (q, n, a, s, m) satisfying 80-bit security against the algebraic attack under $q = 3$, $n = 59$. Moreover, we set $s = a + 15$, similarly to Chen et al. took, from the perspective of the decryption failure rate. Then, we can take

$$(q, n, a, s, m) = (3, 59, 32, 47, 75)$$

as the parameter with the smallest a. However, as seen in [7, Section 10], the decryption complexity is given by $\mathcal{O}(q^a n^3)$. This implies that the parameter $(q, n, a, s, m) = (3, 59, 32, 47, 75)$ has a very slow decryption. As a result, based on our complexity estimation of the algebraic attack, we can confirm that Chen et al.'s scheme is not practical.

5 Conclusion

In this paper, we analyzed the algebraic attack using a Gröbner basis algorithm for Chen et al.'s scheme. We showed that the vector space spanned by the quadratic polynomials in the public key P has many degree-one polynomials. By applying such degree-one polynomials to the equation $P(x) = y$ for a ciphertext y, we reduced the MQ-problem $P(x) = y$ to a smaller MQ-problem with $s + 1$

quadratic equations in a variables. As a result, we obtained a precise complexity estimation for the algebraic attack against Chen et al.'s scheme. Our estimation shows that the claimed 80- and 128-bit security parameters in [7] are broken in the complexity of approximately 25 and 32 bits, respectively. Moreover, by discussing a parameter satisfying 80-bit security against the algebraic attack, we concluded that Chen et al.'s scheme is not practical.

In Chen et al.'s scheme, the minus and plus modifiers were used in order to strengthen the security against the linearization attack; however, they did not succeed in resisting the algebraic attack. As a future work, we would like to study other modifiers that can resists the algebraic attack.

References

1. Bardet, B., Faugère, J.C., Salvy, B. and Yang, B.Y.: Asymptotic behavior of the index of regularity of quadratic semi-regular polynomial systems, *In 8th International Symposium on Effective Methods in Algebraic Geometry (MEGA)*, pp.1–14 (2005).
2. Bernstein, D.J., Buchmann J. and Dahmen, E. eds., *Post-Quantum Cryptography*, Springer (2009).
3. Bettale, L., Faugère, J. C. and Perret, L.: Cryptanalysis of HFE, multi-HFE and variants for odd and even characteristic, *Designs, Codes and Cryptography*, Vol.69, pp.1–52 (2013).
4. Beullens, W., Preneel, B., Szepieniec, A. and Vercauteren, F.: LUOV, Technical report, National Institute of Standards and Technology, *Post-Quantum Cryptography*, available from ⟨https://csrc.nist.gov/Projects/Post-Quantum-Cryptography/Round-2-submissions⟩ (accessed 2020-5-10).
5. Bosma, W., Cannon, J. and Playoust, C.: The Magma algebra system. I. The user language, *J. Symbolic Comput*, Vol.24, pp.235–265 (1997).
6. Casanova, A., Faugére, J.C., Macario-Rat, G., Patarin, J., Perret, L. and Ryckeghem, J.: GeMSS, Technical report, National Institute of Standards and Technology, *Post-Quantum Cryptography*, available from ⟨https://csrc.nist.gov/Projects/Post-Quantum-Cryptography/Round-2-submissions⟩ (accessed 2020-5-10).
7. Chen, J., Ling, J., Ning, J., Lau, T.S.C. and Wang, Y.: A New Encryption Scheme for Multivariate Quadratic Systems, *Theoretical Computer Science*, Vol.809, pp.372–383 (2020).
8. Courtois, N.: The security of hidden field equations (HFE), *CT-RSA*, LNCS, Vol.2020, pp.266–281, Springer (2001).
9. Cartor, R. and Smith-Tone, D.: EFLASH: A new multivariate encryption scheme, *SAC 2018*, LNCS, Vol.11349, pp.281–299. Springer (2018).
10. Ding, J., Gower, J. E. and Schmidt, D. S.: *Multivariate Public Key Cryptosystems*, Springer (2006).
11. Ding, J. and Hodges, T. J.: Inverting HFE Systems is Quasi-Polynomial for All Fields, *Crypto 2011*, LNCS Vol.6841, pp.724–742, Springer (2011).
12. Ding, J. and Schmidt, D. S.: Rainbow, a new multivariate polynomial signature scheme, *ACNS 2005*, LNCS, Vol.3531, pp.164–175, Springer (2005).
13. Ding, J., Chen, M. S., Petzoldt, A., Schmidt, D. S. and Yang, B. Y.: Rainbow, Technical report, National Institute of Standards and Technology, *Post-Quantum Cryptography*, ⟨https://csrc.nist.gov/Projects/Post-Quantum-Cryptography/Round-2-submissions⟩ (accessed 2020-5-10).

14. Dubois, V. and Gamma, N.: The Degree of Regularity of HFE Systems, *Asiacrypt 2010*, LNCS, Vol.6477, pp.557–576, Springer (2010).
15. Faugère, J.C.: A new efficient algorithm for computing Gröbner bases (F4), *Journal of Pure and Applied Algebra*, Vol.139, pp. 61–88, (1999).
16. Faugère, J.C.: A new efficient algorithm for computing Gröbner Bases without reduction to zero (F5), *ISSAC 2002*, pp. 75–83 (2002).
17. Faugère, J.C. and Joux, A.: Algebraic cryptanalysis of Hidden Field Equations (HFE) using Groöbner bases, *Crypto 2003*, LNCS, Vol.2729, pp.44–60, Springer (2003).
18. Garey, M. R. and Johnson, D. S.: Computers and Intractability: A Guide to the Theory of NP-Completeness, *W.H. Freeman and Company*, (1979).
19. Granboulan, L., Joux, A. and Stern, J.: Inverting HFE is quasipolynomial, *Crypto 2006*, LNCS, Vol.4117, pp. 345–356, Springer (2006).
20. Ikematsu, Y., Perlner, R., Smith-Tone, D., Takagi, T. and Vates, J.: HFERP - A New Multivariate Encryption Scheme, *PQCrypto 2018*, LNCS Vol.10786, pp.396–416, Springer (2018).
21. Kipnis, A., Patarin, L. and Goubin, L.: Unbalanced Oil and Vinegar Schemes, *EUROCRYPT 1999*, LNCS, Vol.1592, pp.206-222, Springer (1999).
22. Kipnis, A. and Shamir, A.: Cryptanalysis of the HFE public key cryptosystem by relinearization, *CRYPTO'99*, LNCS, Vol. 1666, pp.19–30, Springer (1999).
23. Matsumoto, T. and Imai, H.: Public quadratic polynomial-tuples for efficient signature-verification and message-encryption, *EUROCRYPT 1988*, LNCS, Vol.330, pp. 419–453, Springer (1988).
24. National Institute of Standards and Technology, *Post-quantum cryptography, Round 2 submission*, ⟨https://csrc.nist.gov/Projects/Post-Quantum-Cryptography/Round-2-Submissions⟩ (accessed 2020-5-10).
25. Patarin, J.: Cryptanalysis of the Matsumoto and Imai public key scheme of Eurocrypt 88, *CRYPTO 1995*, LNCS, Vol.963, pp.248–261, Springer (1995).
26. Patarin, J.: Hidden Field Equations (HFE) and Isomorphisms of Polynomials (IP): Two New Families of Asymmetric Algorithms, *EUROCRYPT*, LNCS, Vol.1070, pp.33-48, Springer (1996).
27. J. Patarin, N.T. Courtois, L. Goubin: QUARTZ, 128-bit long digital signatures. CT-RSA 2001, LNCS, vol. 2020, pp. 282 - 297. Springer, 2001.
28. Petzoldt, A., Chen, M.S., Yang, B.Y., Tao, C. and Ding, J.: Design Principles for HFEv- based Signature Schemes, *ASIACRYPT 2015*, LNCS, Vol.9452, pp.311–334, Springer (2015).
29. Szepieniec, A., Ding, J. and Preneel, B.: Extension Field Cancellation: a New Central Trapdoor for Multivariate Quadratic Systems, *PQCrypto 2016*, LNCS, Vol.9606, pp.182-196, Springer (2016).
30. Tao, C., Diene, A., Tang, S. and Ding, J.: Simple matrix scheme for encryption, *PQCrypto 2013*, LNCS, Vol7932, pp.231–242, Springer (2013).
31. Yang, B.-Y. and Chen, J.-M.: All in the XL family: Theory and practice, *ICISC 2004*, LNCS, Vol.3506, pp.67-86, Springer (2004).