Improved Key Recovery of the HFEv- Signature Scheme

Chengdong Tao[2], Albrecht Petzoldt[3], Jintai Ding[1,2,4]

[1] Yau Mathematical Center, Tsinghua University, Beijing, China
[2] Ding Lab, Beijing Institute of Mathematical Sci. and Applications, Beijing, China
[3] FAU Erlangen-Nuremberg, Nuremberg, Germany
[4] University of Cincinnati, Cincinnati, Ohio, USA

taochengdong@bimsa.cn, albrecht.petzoldt@googlemail.com, jintai.ding@gmail.com

Abstract. The HFEv- signature scheme is a twenty year old multivariate public key signature scheme. It uses the Minus and the Vinegar modifier on the original HFE scheme. An instance of the HFEv- signature scheme called GeMSS is one of the alternative candidates for signature schemes in the third round of the NIST Post Quantum Crypto (PQC) Standardization Project. In this paper, we propose a new key recovery attack on the HFEv- signature scheme. We show that the Minus modification does not enhance the security of cryptosystems of the HFE family, while the Vinegar modification increases the complexity of our attack only by a polynomial factor. By doing so, we show that the proposed parameters of the GeMSS scheme are not as secure as claimed. Our attack shows that it is very difficult to build a secure and efficient signature scheme on the basis of HFEv-.

Keywords: Multivariate Cryptography · HFEv- · Key Recovery · MinRank

1 Introduction

Crryptographic techniques such as encryption and digital signatures are an indispensable part of modern communication systems. However, the currently used schemes RSA and ECDSA become insecure as soon as large quantum computers arrive. Due to recent progress in the development of such computers, there is an urgent need for alternatives to these classical schemes which are resistant against attacks with quantum computers. These are known as post quantum cryptosystems [5].

One of the main candidates for such schemes are multivariate public key cryptosystems [14]. Especially in the area of digital signatures, there exist many promising multivariate schemes. In fact, the multivariate signature scheme Rainbow is among the three signature schemes in the third round of the NIST standardization process of post quantum cryptosystems [7]. Another multivariate signature scheme, GeMMS, is one of the alternative candidates. GeMMS is a special instance of the well known HFEv- signature scheme, which was first proposed by Patarin et al. in [24]. The principle idea of HFEv- is to combine the Minus and the Vinegar modifications with the HFE cryptosystem of [23]. Since the resulting multivariate quadratic system contains more variables than equations, HFEv- can only be used for digital signatures.

There exist many attack methods on HFEv-, such as the direct attack [8][25], the distinguishing attack [12], the differential attack [6], and the MinRank attack [12]. The most studied attack against HFEv- is the MinRank attack, which was first proposed by Kipnis and Shamir [21]. Later, many variants of this technique have been proposed to increase its efficiency [3, 1]. The complexity of the MinRank attack on HFEv- with minors modeling [3] is given as

$$\mathcal{O}\left(\binom{n+d+a+v+1}{d+a+v+1}^{\omega}\right),$$

where n is the degree of the field extension, $d = \lceil \log_q(D) \rceil$, D is the degree bound on the HFE central map polynomial, a is the number of Minus equations, v is the number of Vinegar variables and $2 < \omega \leq 3$ is the linear algebra constant.

In this paper, we present an improved key recovery attack on the HFEv- signature scheme. The complexity of our new attack on HFEv- with minors modeling is

$$\mathcal{O}\left(\binom{n+d+v+1}{d+1}^{\omega}\right).$$

This implies that the Minus modification does not enhance the security of HFE type cryptosystems, while the Vinegar modification increases the complexity of our attack only by a polynomial factor. We use our attack to show that the parameters of GeMSS which were submitted to the NIST Post Quantum Crypto Standardization Project are not as secure as claimed.

2 The HFEv- Signature Scheme

Let $\mathbb{N}$ be the set of positive integers, $n, v, D, a \in \mathbb{N}$, q be a prime number, and $\mathbb{F}_q$ be a finite field with q elements. Let $\mu(X) \in \mathbb{F}_q[X]$ be an irreducible polynomial of degree n. Define the field $\mathbb{F}_{q^n} = \mathbb{F}_q[X]/\mu(X)$. It is a degree n extension field of $\mathbb{F}_q$. Let $\phi : \mathbb{F}_{q^n} \to \mathbb{F}_q^n$ be an isomorphism between $\mathbb{F}_{q^n}$ and $\mathbb{F}_q^n$ defined by

$$\phi(a_0 + a_1 X + \cdots + a_{n-1} X^{n-1}) = (a_0, a_1, \cdots, a_{n-1}).$$

Private Key Generation. Randomly generate the central map of HFEv- as

$$\mathcal{F}(X, x_1, \cdots, x_v) = \sum_{\substack{i,j \in \mathbb{N} \\ q^i + q^j \leq D}} \alpha_{ij} X^{q^i + q^j} + \sum_{\substack{i \in \mathbb{N} \\ q^i \leq D}} \beta_i(x_1, \cdots, x_v) X^{q^i} + \gamma(x_1, \cdots, x_v),$$

where $\alpha_{i,j} \in \mathbb{F}_{q^n}$, $\beta_i : \mathbb{F}_q^v \to \mathbb{F}_{q^n}$ are linear maps and $\gamma : \mathbb{F}_q^v \to \mathbb{F}_{q^n}$ is a quadratic map in the Vinegar variables $x_1, x_2, \cdots, x_v$. The central map of HFEv- can be viewed as a polynomial in the quotient ring $\mathbb{F}_{q^n}[X, x_1, \cdots, x_v]/\langle x_1^q - x_1, \cdots, x_v^q - x_v \rangle$. Randomly generate two affine transformations $\mathcal{T} : \mathbb{F}_q^n \to \mathbb{F}_q^{n-a}$ and $\mathcal{S} : \mathbb{F}_q^{n+v} \to \mathbb{F}_q^{n+v}$ of maximal rank. Then the private key is $(\mathcal{T}, \mathcal{F}, \mathcal{S})$.

Public Key Generation. Let $\psi : \mathbb{F}_q^{n+v} \to \mathbb{F}_{q^n} \times \mathbb{F}_q^v$ be given as $\psi = \phi^{-1} \times id_v$, where $\phi : \mathbb{F}_{q^n} \mapsto \mathbb{F}_q^n$ is the isomorphism defined above and id_v is the identity map over $\mathbb{F}_q^v$. From the special structure of $\mathcal{F}$, we know that $\overline{\mathcal{F}} = \phi \circ \mathcal{F} \circ \psi : \mathbb{F}_q^{n+v} \to \mathbb{F}_q^n$ is a quadratic multivariate map. The public key map is

$$\mathcal{P} = \mathcal{T} \circ \phi \circ \mathcal{F} \circ \psi \circ \mathcal{S} : \mathbb{F}_q^{n+v} \to \mathbb{F}_q^{n-a}.$$

Specifically, the public key is a set of multivariate quadratic polynomials

$$\{p^{(1)}(x_1, \cdots, x_{n+v}), \cdots, p^{(n-a)}(x_1, ..., x_{n+v})\},$$

where

$$p^{(k)}(x_1, \cdots, x_{n+v}) = \sum_{1 \le i \le j \le n+v} a_{ij}^{(k)} x_i x_j + \sum_{i=1}^{n+v} b_i^{(k)} x_i + c^{(k)},$$

$a_{ij}^{(k)}, b_i^{(k)}, c^{(k)} \in \mathbb{F}_q (k = 1, \cdots, n-a)$.

Signature Generation. Let $\mathbf{y} = (y_1, y_2, \cdots, y_{n-a}) \in \mathbb{F}_q^{n-a}$ be a message to be signed. The process of signature generation is as follows:

1. Compute a preimage $\overline{\mathbf{y}} \in \mathbb{F}_q^n$ of $\mathbf{y}$ under the affine transformation $\mathcal{T} : \mathbb{F}_q^n \to \mathbb{F}_q^{n-a}$ and lift it to the extension field, obtaining $\overline{Y} \in \mathbb{F}_{q^n}$.
2. Choose random values for the Vinegar variables $(x_1, \cdots, x_v) \in \mathbb{F}_q^v$ and substitute them into the central map $\mathcal{F}$ to obtain a new map $\mathcal{F}_V(X) : \mathbb{F}_{q^n} \to \mathbb{F}_{q^n}$.
3. Find a solution to the equation $\mathcal{F}_V(X) = \overline{Y}$ using Berlekamps algorithm. If the equation has no solution, go to step 2, and randomly choose another vector $(x_1, \cdots, x_v) \in \mathbb{F}_q^v$ until we can find a solution. Let $\widetilde{Y}$ be one of the solutions and set $\phi(\widetilde{Y}) = (\widetilde{y}_1, \cdots, \widetilde{y}_n)$. Append the vinegar variables of step 2 to it, obtaining $\widehat{\mathbf{y}} = (\widetilde{y}_1, \cdots, \widetilde{y}_n, x_1, \cdots, x_v) \in \mathbb{F}_q^{n+v}$.
4. Compute $\mathbf{z} = \mathcal{S}^{-1}(\widehat{\mathbf{y}})$. Then $\mathbf{z} \in \mathbb{F}_q^{n+v}$ is the signature of $\mathbf{y}$.

Signature Verification. To verify that $\mathbf{z} \in \mathbb{F}_q^{n+v}$ is indeed a valid signature of the document $\mathbf{y} \in \mathbb{F}^{n-a}$, the recipient computes $\mathcal{P}(\mathbf{z})$. If $\mathcal{P}(\mathbf{z}) = \mathbf{y}$, the recipient accepts the signature, otherwise he rejects it.

2.1 Previous Attacks on HFE

Historically, the most efficient attacks against signature schemes of the HFE type are the direct and the MinRank attack. With regard to the direct attack, it was discovered that the public systems of HFE and its variants can be solved much more efficiently than random systems. This phenomenon was analyzed in a number of papers [10, 11, 15]. The authors of these papers found that the degree of regularity of a public HFEv- system is bounded from above by

$$\begin{cases} \frac{(q-1)(d+v+a-1)}{2} + 2 & \text{if } q \text{ is even and } d+a \text{ is odd,} \\ \frac{(q-1)(d+v+a)}{2} + 2 & \text{otherwise.} \end{cases}$$

Regarding attacks of the MinRank type, many researchers considered the so called min-Q-rank of the HFE system, which can be seen as the rank of the quadratic form $\mathcal{P}$ lifted to the extension field $\mathbb{F}_{q^n}$. Similar to the degree of regularity, the min-q-rank of the HFE system is bounded by the HFE parameters. However, in our attack, we don't consider the min-q-rank of the HFE system, but perform a MinRank attack over the base field $\mathbb{F}_q$. While it is clear that the complexity of a direct attack on a system of the HFE type is exponential in d, a and v [9], our attack shows that this is not the case for MinRank.

3 Preliminaries

For simplification, in the following sections of this paper, we assume that $\mathcal{T}$ and $\mathcal{S}$ are linear transformations and q is an odd prime. Our attack method can be easily extended to the case of affine maps $\mathcal{T}$ and $\mathcal{S}$ and even characteristic.

3.1 Equivalent Keys

An important notion in this paper is that of equivalent keys. For a multivariate public key cryptosystem, the concept of equivalent keys is defined as follows.

Definition 1. *Let $((\mathcal{T}, \mathcal{F}, \mathcal{S}), \mathcal{P})$ be a key pair of a multivariate public key cryptosystem. A tuple $(\mathcal{T}', \mathcal{F}', \mathcal{S}')$ is called an equivalent private key if and only if*

$$\mathcal{P} = \mathcal{T} \circ \mathcal{F} \circ \mathcal{S} = \mathcal{T}' \circ \mathcal{F}' \circ \mathcal{S}'$$

and $\mathcal{F}'$ is a valid central map of the cryptosystem, i.e. $\mathcal{F}'$ has the same algebraic structure as $\mathcal{F}$.

We have

Theorem 1 (Theorem 4.13 in [26]). *Let $\mathcal{P}$ be a public key of the HFEv- scheme over $\mathbb{F}_q$. Let v be the number of Vinegar variables, a be the number of Minus equation and n be the degree of the field extension. Then there exist*

$$nq^{a+2n+vn}(q^n-1)^2 \prod_{i=0}^{v-1}(q^v - q^i) \prod_{i=n-a-1}^{n-1}(q^n - q^i)$$

equivalent private keys for the public key $\mathcal{P}$.

Given an HFEv- public key, our attack finds one of the equivalent private keys.

3.2 The MinRank Problem

The search version of the MinRank problem is defined as follows.

Definition 2 (MinRank problem). *Given a positive number r and n_x matrices $M_1, M_2, ..., M_{n_x}$ with m rows and n columns over a field $\mathbb{F}_q$, find a nonzero vector $(x_1, x_2, \cdots, x_{n_x}) \in \mathbb{F}_q^{n_x}$, such that the linear combination $M = \sum\limits_{i=1}^{n_x} x_i M_i$ has rank at most r.*

Remark 1. Without loss of generality, we assume that $m \geq n$ in Definition 2. In fact, if $m < n$, we use the linear combination $M^t = \sum\limits_{i=1}^{n_x} x_i M_i^t$ instead of M, where t means the transpose of a matrix. Since the rank of M^t is the same as that of M, we only need to find a nonzero vector $(x_1, x_2, \cdots, x_{n_x}) \in \mathbb{F}^{n_x}$, such that M^t has rank at most r.

The MinRank problem is an NP-complete problem [4]. The main methods for solving the MinRank problem are linear algebra search [19], Kipnis-Shamir modeling [21], minors modeling [18] and support minors modeling [1].

Support minors modeling is more efficient than the other models in practical cryptanalysis. The main idea of this modeling is that the low rank matrix M can be written as a product $M = AC$, where A is an $m \times r$ matrix and C is an $r \times n$ matrix. Define m matrices which have the form $\widetilde{C_i} = \begin{pmatrix} \mathbf{r}_i \\ C \end{pmatrix}$ $(i = 1, 2, \cdots, m)$, where $\mathbf{r}_i$ is the i-th row of M. Since $\mathbf{r}_i$ is in the space spanned by the rows of C, the rank of $\widetilde{C_i}$ $(i = 1, 2, \cdots, m)$ is at most r. This implies that all $(r+1) \times (r+1)$ minors of $\widetilde{C_i}$ $(i = 1, 2, \cdots, m)$ are 0. We view the $r \times r$ minors of the matrix C as new variables which are called kernel variables and are denoted as $y_1, y_2, \cdots, y_{n_y}$, where $n_y = \binom{n}{r}$. The $(r+1) \times (r+1)$ minors of $\widetilde{C_i}$ are therefore given as bilinear equations in the variables $x_1, \ldots, x_{n_x}$ and $y_1, \ldots, y_{n_y}$. Altogether, we obtain $m\binom{n}{r+1}$ of these bilinear equations. The total number of monomials of degree 2 in these bilinear equations is at most $n_x\binom{n}{r}$. If

$$m\binom{n}{r+1} \geq n_x \binom{n}{r} - 1,$$

holds, we can solve this system of bilinear equations by linearization.

In practical applications, we can assume that C has the form (I_r, C_0), where I_r is an $r \times r$ identity matrix and C_0 is an $r \times (n-r)$ matrix. Moreover, instead of using all $r \times r$ minors of the matrix C as variables, we choose a positive integer $n' \leq n$, such that

$$m\binom{n'}{r+1} \geq n_x \binom{n'}{r} - 1 \tag{1}$$

holds. If the MinRank problem has only one solution, the resulting linear system is sparse, and we can solve it using the Wiedemann algorithm. The complexity of solving this linear system is

$$\mathcal{O}\left(\left(n_x\binom{n'}{r}\right)^2 \cdot n_x(r+1)\right)$$

field operations. If the MinRank problem has no unique solution and $\mathbb{F}$ is a small finite field, we can guess the values of some variables such that the resulting linear system has a unique solution, and then solve it using the Wiedemann algorithm. Otherwise, we solve the bilinear system using a Gröbner basis like algorithm such as F4/F5 [16].

4 Key Recovery for HFEv-

In this section we describe our key recovery attack on the HFEv- signature scheme. Let q, n, v, D, a be the parameters of HFEv-. Denote $d = \lceil \log_q(D) \rceil$. In this paper, we assume that $0 \leq a < n - 2d - 1$.

4.1 Matrix Representation of HFEv- Keys

Similar to [3], we can represent the HFEv- signature scheme in matrix form.

Proposition 1. *Let*

$$F^{*0} = \begin{pmatrix} \alpha_{00} & \alpha_{01} & \cdots & \alpha_{0,n-1} & \gamma_{00} & \gamma_{01} & \cdots & \gamma_{0,v-1} \\ \alpha_{10} & \alpha_{11} & \cdots & \alpha_{1,n-1} & \gamma_{10} & \gamma_{11} & \cdots & \gamma_{1,v-1} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ \alpha_{n-1,0} & \alpha_{n-1,1} & \cdots & \alpha_{n-1,n-1} & \gamma_{n-1,0} & \gamma_{n-1,1} & \cdots & \gamma_{n-1,v-1} \\ \beta_{00} & \beta_{01} & \cdots & \beta_{0,n-1} & \delta_{00} & \delta_{01} & \cdots & \delta_{0,v-1} \\ \beta_{10} & \beta_{11} & \cdots & \beta_{1,n-1} & \delta_{10} & \delta_{11} & \cdots & \delta_{1,v-1} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ \beta_{v-1,0} & \beta_{v-1,1} & \cdots & \beta_{v-1,n-1} & \delta_{v-1,0} & \delta_{v-1,1} & \cdots & \delta_{v-1,v-1} \end{pmatrix}$$

be an $(n+v) \times (n+v)$ matrix over the field $\mathbb{F}_{q^n}$ and

$$F(X, x_1, \cdots, x_v) = (X, X^q, \cdots, X^{q^{n-1}}, x_1, \cdots, x_v) F^{*0} (X, X^q, \cdots, X^{q^{n-1}}, x_1, \cdots, x_v)^t$$

be a polynomial in the quotient ring $\mathbb{F}_{q^n}[X, x_1, \cdots, x_v]/\langle x_1^q - x_1, \cdots, x_v^q - x_v \rangle$. Then we have for all $0 \leq k < n$

$$F^{q^k}(X, x_1, \cdots, x_v) = (X, X^q, \cdots, X^{q^{n-1}}, x_1, \cdots, x_v) F^{*k} (X, X^q, \cdots, X^{q^{n-1}}, x_1, \cdots, x_v)^t,$$

where $F^{*k} \in \mathcal{M}_{(n+v)\times(n+v)}(\mathbb{F}_{q^n})$, the (i,j)-th entry of F^{*k} is $\alpha_{i-k,j-k}^{q^k}$ for all $0 \leq i, j, k < n$, the $(i, n+j)$-th entry of F^{*k} is $\gamma_{j-k,i}^{q^k}$ for all $0 \leq j, k < n$, $0 \leq i < v$, the $(n+i, j)$-th entry of F^{*k} is $\beta_{i,j-k}^{q^k}$ for all $0 \leq i < v, 0 \leq j, k < n$, and the $(n+i, n+j)$-th entry is $\delta_{ij}^{q^k}$ for all $0 \leq i < v, 0 \leq j < v, 0 \leq k < n$.

Proof. If $k=0$, we have obviously $F^{q^k}(X,x_1,\cdots,x_v)=F(X,x_1,\cdots,x_v)$. Now we consider the case of $1\le k<n$. Since $x_i^{q^k}=x_i$ for all $1\le i\le v$, we have

$$\begin{aligned}F^{q^k} &= \sum_{i=0}^{n-1}\sum_{j=0}^{n-1}\alpha_{ij}^{q^k}X^{q^{i+k}+q^{j+k}}+\sum_{i=0}^{v-1}\sum_{j=0}^{n-1}(\beta_{ij}^{q^k}+\gamma_{ji}^{q^k})x_iX^{q^{j+k}}+\sum_{i=0}^{v-1}\sum_{j=0}^{v-1}\delta_{ij}^{q^k}x_ix_j\\ &= \sum_{i=k}^{n-1+k}\sum_{j=k}^{n-1+k}\alpha_{i-k,j-k}^{q^k}X^{q^i+q^j}+\sum_{i=0}^{v-1}\sum_{j=k}^{n-1+k}(\beta_{i,j-k}^{q^k}+\gamma_{j-k,i}^{q^k})x_iX^{q^j}+\sum_{i=0}^{v-1}\sum_{j=0}^{v-1}\delta_{ij}^{q^k}x_ix_j\end{aligned}$$

Then it can be divided as follows

$$\begin{aligned}F^{q^k} &= \sum_{i=k}^{n-1}\left(\sum_{j=k}^{n-1+k}\alpha_{i-k,j-k}^{q^k}X^{q^i+q^j}\right)+\sum_{i=n}^{n-1+k}\left(\sum_{j=k}^{n-1+k}\alpha_{i-k,j-k}^{q^k}X^{q^i+q^j}\right)\\ &\quad+\sum_{i=0}^{v-1}\sum_{j=k}^{n-1}(\beta_{i,j-k}^{q^k}+\gamma_{j-k,i}^{q^k})x_iX^{q^j}+\sum_{i=0}^{v-1}\sum_{j=n}^{n-1+k}(\beta_{i,j-k}^{q^k}+\gamma_{j-k,i}^{q^k})x_iX^{q^j}+\sum_{i=0}^{v-1}\sum_{j=0}^{v-1}\delta_{ij}^{q^k}x_ix_j.\end{aligned}$$

That is

$$\begin{aligned}F^{q^k} &= \sum_{i=k}^{n-1}\left(\sum_{j=k}^{n-1}\alpha_{i-k,j-k}^{q^k}X^{q^i+q^j}+\sum_{j=n}^{n-1+k}\alpha_{i-k,j-k}^{q^k}X^{q^i+q^j}\right)\\ &\quad+\sum_{i=n}^{n-1+k}\left(\sum_{j=k}^{n-1}\alpha_{i-k,j-k}^{q^k}X^{q^i+q^j}+\sum_{j=n}^{n-1+k}\alpha_{i-k,j-k}^{q^k}X^{q^i+q^j}\right)\\ &\quad+\sum_{i=0}^{v-1}\sum_{j=k}^{n-1}(\beta_{i,j-k}^{q^k}+\gamma_{j-k,i}^{q^k})x_iX^{q^j}+\sum_{i=0}^{v-1}\sum_{j=n}^{n-1+k}(\beta_{i,j-k}^{q^k}+\gamma_{j-k,i}^{q^k})x_iX^{q^j}+\sum_{i=0}^{v-1}\sum_{j=0}^{v-1}\delta_{ij}^{q^k}x_ix_j.\end{aligned}$$

Thus we have

$$\begin{aligned}F^{q^k} &= \sum_{i=k}^{n-1}\left(\sum_{j=k}^{n-1}\alpha_{i-k,j-k}^{q^k}X^{q^i+q^j}+\sum_{j=0}^{k-1}\alpha_{i-k,j-k+n}^{q^k}X^{q^i+q^{j+n}}\right)\\ &\quad+\sum_{i=0}^{k-1}\left(\sum_{j=k}^{n-1}\alpha_{i-k+n,j-k}^{q^k}X^{q^{i+n}+q^j}+\sum_{j=0}^{k-1}\alpha_{i-k+n,j-k+n}^{q^k}X^{q^{i+n}+q^{j+n}}\right)\\ &\quad+\sum_{i=0}^{v-1}\sum_{j=k}^{n-1}(\beta_{i,j-k}^{q^k}+\gamma_{j-k,i}^{q^k})x_iX^{q^j}+\sum_{i=0}^{v-1}\sum_{j=0}^{k-1}(\beta_{i,j-k+n}^{q^k}+\gamma_{j-k+n,i}^{q^k})x_iX^{q^{j+n}}+\sum_{i=0}^{v-1}\sum_{j=0}^{v-1}\delta_{ij}^{q^k}x_ix_j.\end{aligned}$$

Since $X^{q^n}=X$ we obtain by reducing the index of coefficients modulo n

$$\begin{aligned}F^{q^k} &= \sum_{i=k}^{n-1}\left(\sum_{j=k}^{n-1}\alpha_{i-k,j-k}^{q^k}X^{q^i+q^j}+\sum_{j=0}^{k-1}\alpha_{i-k,j-k}^{q^k}X^{q^i+q^j}\right)\\ &\quad+\sum_{i=0}^{k-1}\left(\sum_{j=k}^{n-1}\alpha_{i-k,j-k}^{q^k}X^{q^i+q^j}+\sum_{j=0}^{k-1}\alpha_{i-k,j-k}^{q^k}X^{q^i+q^j}\right)\\ &\quad+\sum_{i=0}^{v-1}\sum_{j=k}^{n-1}(\beta_{i,j-k}^{q^k}+\gamma_{j-k,i}^{q^k})x_iX^{q^j}+\sum_{i=0}^{v-1}\sum_{j=0}^{k-1}(\beta_{i,j-k}^{q^k}+\gamma_{j-k,i}^{q^k})x_iX^{q^j}+\sum_{i=0}^{v-1}\sum_{j=0}^{v-1}\delta_{ij}^{q^k}x_ix_j.\end{aligned}$$

Grouping the sums back together, we get

$$\begin{aligned}F^{q^k} &= \sum_{i=0}^{n-1}\sum_{j=0}^{n-1}a_{i-k,j-k}^{q^k}X^{q^i+q^j}+\sum_{i=0}^{v-1}\sum_{j=0}^{n-1}(\beta_{i,j-k}^{q^k}+\gamma_{j-k,i}^{q^k})x_iX^{q^j}+\sum_{i=0}^{v-1}\sum_{j=0}^{v-1}\delta_{ij}^{q^k}x_ix_j\\ &=(X,X^q,\cdots,X^{q^{n-1}},x_1,\cdots,x_v)F^{*k}(X,X^q,\cdots,X^{q^{n-1}},x_1,\cdots,x_v)^t,\end{aligned}$$

where $F^{*k} \in \mathcal{M}_{(n+v)\times(n+v)}(\mathbb{F}_{q^n})$, the (i,j)-th entry of F^{*k} is $\alpha_{i-k,j-k}^{q^k}$ for all $0 \le i,j,k < n$, the $(i,n+j)$-th entry of F^{*k} is $\gamma_{j-k,i}^{q^k}$ for all $0 \le j,k < n, 0 \le i < v$, the $(n+i,j)$-th entry of F^{*k} is $\beta_{i,j-k}^{q^k}$ for all $0 \le i < v, 0 \le j,k < n$, and the $(n+i,n+j)$-th entry is $\delta_{ij}^{q^k}$ for all $0 \le i < v, 0 \le j < v, 0 \le k < n$. □

Proposition 2 (Proposition 2.1 in [3]). *Let $(\theta_1, \theta_2, \cdots, \theta_n) \in \mathbb{F}_{q^n}^n$ be a vector basis of $\mathbb{F}_{q^n}$ over $\mathbb{F}_q$ and*

$$M = \begin{pmatrix} \theta_1 & \theta_1^q & \cdots & \theta_1^{q^{n-1}} \\ \theta_2 & \theta_2^q & \cdots & \theta_2^{q^{n-1}} \\ \vdots & \vdots & \ddots & \vdots \\ \theta_n & \theta_n^q & \cdots & \theta_n^{q^{n-1}} \end{pmatrix}$$

be the matrix whose columns are the Frobenius powers of the basis elements. We can express the morphism $\phi : \mathbb{F}_{q^n} \to \mathbb{F}_q^n$ as

$$V \mapsto (V, V^q, \cdots, V^{q^{n-1}})M^{-1}.$$

Its inverse $\phi^{-1} : \mathbb{F}_q^n \to \mathbb{F}_{q^n}$ is given as

$$(v_1, v_2, \cdots, v_n) \mapsto V,$$

where V is the first component of the vector $(v_1, v_2, \cdots, v_n)M$. More generally, we have

$$(v_1, v_2, \cdots, v_n) \cdot M = (V, V^q, \cdots, V^{q^{n-1}}).$$

In this paper, we choose

$$M = \begin{pmatrix} 1 & 1 & \cdots & 1 \\ \theta & \theta^q & \cdots & \theta^{q^{n-1}} \\ \vdots & \vdots & \ddots & \vdots \\ \theta^{n-1} & (\theta^{n-1})^q & \cdots & (\theta^{n-1})^{q^{n-1}} \end{pmatrix}, \tag{2}$$

where θ is a generator of $\mathbb{F}_{q^n}$. Define

$$\widetilde{M} = \begin{pmatrix} M & 0 \\ 0 & I_v \end{pmatrix} \in \mathcal{M}_{(n+v)\times(n+v)}(\mathbb{F}_{q^n}), \tag{3}$$

where I_v is the $v \times v$ identity matrix. According to Proposition 1, we have

$$(v_1, v_2, \cdots, v_n, x_1, \cdots, x_v) \cdot \widetilde{M} = (V, V^q, \cdots, V^{q^{n-1}}, x_1, \cdots, x_v),$$

where $v_i, x_j \in \mathbb{F}_q, 1 \le i \le n, 1 \le j \le v$ and $V \in \mathbb{F}_{q^n}$.

Proposition 3. *Let $p_i \in \mathbb{F}_q[x_1, x_2, \cdots, x_{n+v}]$ be the public key polynomials of HFEv- and P_i be the matrix representing the quadratic form of p_i, $0 \le i < n-a$. Let the central map of HFEv- be*

$$F = (X, X^q, \cdots, X^{q^{n-1}}, x_1, \cdots, x_v) F^{*0} (X, X^q, \cdots, X^{q^{n-1}}, x_1, \cdots, x_v)^t,$$

where $F^{*0} \in \mathcal{M}_{(n+v)\times(n+v)}(\mathbb{F}_{q^n})$. Let $S \in \mathcal{M}_{(n+v)\times(n+v)}(\mathbb{F}_q)$ and $T \in \mathcal{M}_{n\times(n-a)}(\mathbb{F}_q)$ be the matrices representing the linear parts of $\mathcal{S}$ and $\mathcal{T}$. Then

$$\begin{aligned}&\left(\widetilde{M}^{-1} S^{-1} P_0 (S^{-1})^t (\widetilde{M}^{-1})^t, \cdots, \widetilde{M}^{-1} S^{-1} P_{n-a-1} (S^{-1})^t (\widetilde{M}^{-1})^t\right)\\ &= \left(F^{*0}, \cdots, F^{*n-1}\right) M^{-1} T \end{aligned} \tag{4}$$

Proof. Similar to Lemma 2 in [3].

Denote $U = \widetilde{M}^{-1} S^{-1} \in \mathcal{M}_{(n+v)\times(n+v)}(\mathbb{F}_{q^n})$ and $W = M^{-1} T \in \mathcal{M}_{n\times(n-a)}(\mathbb{F}_{q^n})$, then Equation (4) can be rewritten as

$$\left(U P_0 U^t, \cdots, U P_{n-1} U^t\right) = \left(F^{*0}, \cdots, F^{*n-1}\right) W. \tag{5}$$

4.2 Recovering an Equivalent Linear Transformation $\mathcal{S}$

In this subsection, we will present our technique of finding an equivalent map S.

Proposition 4. *Let $F^{*0}, \cdots, F^{*n-1}$ and $W = [w_{ij}]$ be the matrices of Equation (5) and $\mathbf{a}_i$ be the first row of matrix F^{*i} $(i = 0, 1, \cdots, n-1)$. Let Q be the matrix given as $Q = W^t \cdot \begin{pmatrix} \mathbf{a}_0 \\ \vdots \\ \mathbf{a}_{n-1} \end{pmatrix}$. Then the rank of Q is at most $d = \lceil \log_q(D) \rceil$.*

Proof. We have

$$Q = \begin{pmatrix} w_{11}\mathbf{a}_0 + w_{21}\mathbf{a}_1 + \cdots + w_{n1}\mathbf{a}_{n-1} \\ w_{12}\mathbf{a}_0 + w_{22}\mathbf{a}_1 + \cdots + w_{n2}\mathbf{a}_{n-1} \\ \cdots \\ w_{1,n-a}\mathbf{a}_0 + w_{2,n-a}\mathbf{a}_1 + \cdots + w_{n,n-a}\mathbf{a}_{n-1} \end{pmatrix} = W^t \cdot \begin{pmatrix} \mathbf{a}_0 \\ \mathbf{a}_1 \\ \cdots \\ \mathbf{a}_{n-1} \end{pmatrix}$$

Due to the construction of the matrices $F^{*i} (i = 0, 1, \cdots, n-1)$, we have

$$\begin{pmatrix} \mathbf{a}_0 \\ \mathbf{a}_1 \\ \cdots \\ \mathbf{a}_{n-1} \end{pmatrix} = \begin{pmatrix} A_1 \\ 0 \\ A_2 \end{pmatrix},$$

where A_1 is an $1 \times (n+v)$ matrix and A_2 is a $(d-1) \times (n+v)$ matrix. That is, this matrix has only d non-zero rows, therefore its rank is at most d. Therefore the rank of Q is at most d. $\square$

Theorem 2. *Let $P_0, P_1, \cdots, P_{n-a-1}$ and U be the matrices of Equation (5), the vector $\mathbf{u} = (u_0, u_1, \cdots, u_{n+v-1})$ be the first row of U and $\mathbf{b}_i = (u_0, u_1, \cdots, u_{n+v-1})P_i$, $(i = 0, 1, \cdots, n-a)$. Define $Z \in \mathcal{M}_{(n-a)\times(n+v)}(\mathbb{F}_{q^n})$ as the matrix whose row vectors are the $\mathbf{b}_i$. Then the rank of Z is at most d.*

Proof. From Equation (5) and Proposition 4, we know that the rank of ZU^t is not more than d. Thus the rank of Z is at most d. □

Proposition 5. *Let $A = [a_{ij}]$ be an $n \times m$ matrix over $\mathbb{F}_q$, $B = M^{-1}A = [b_{ij}] \in \mathcal{M}_{n\times m}(\mathbb{F}_{q^n})$. Then*

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

That is, each row is obtained from the previous one using a Frobenius application. The whole matrix B can be defined with any of its rows.

Proof. Let $(\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n)$ be a dual basis of $(\theta_1, \theta_2, \cdots, \theta_n)$ of $\mathbb{F}_{q^n}$ over $\mathbb{F}_q$, then we have

$$M^{-1} = \begin{pmatrix} \varepsilon_1 & \varepsilon_2 & \cdots & \varepsilon_n \\ \varepsilon_1^q & \varepsilon_2^q & \cdots & \varepsilon_n^q \\ \vdots & \vdots & \ddots & \vdots \\ \varepsilon_1^{q^{n-1}} & \varepsilon_2^{q^{n-1}} & \cdots & \varepsilon_n^{q^{n-1}} \end{pmatrix}.$$

Thus $b_{ij} = \sum_{k=0}^{n-1} a_{kj}\varepsilon_{k+1}^{q^i}$ for all $i, j, 0 \leq i < n, 0 \leq j < m$. Since $a_{ij}^q = a_{ij}$ and the linearity of Frobenius, we have

$$b_{i-1,j}^q = \left(\sum_{k=0}^{n-1} a_{kj}\varepsilon_{k+1}^{q^{i-1}}\right)^q = \sum_{k=0}^{n-1} a_{kj}^q (\varepsilon_{k+1}^{q^{i-1}})^q = \sum_{k=0}^{n-1} a_{kj}\varepsilon_{k+1}^{q^i} = b_{ij}$$

for all $i, j, 0 < i \leq n, 0 \leq j < m$. □

Proposition 5 implies that we only need to find one row of matrix $U = \widetilde{M}^{-1}S^{-1}$ to recover the first n rows of U. Let $u_0, u_1, \cdots, u_{n+v-1}$ be the first row of U. We assume that $u_0, u_1, \cdots, u_{n+v-1}$ are unknowns. Since we need to find only one of the equivalent HFEv- private keys, we can fix $u_0 = 1$ [20]. Since the rank of Z is at most d, we can find the u_i $(i = 1, \ldots, n+v)$ by soving a MinRank Problem. This can be done by using any of the methods presented in Section 3. Our method to recover S can be summarized as shown in Algorithm 1.

4.3 Recovering Equivalent Maps $\mathcal{F}$ and $\mathcal{T}$

Proposition 6. *Let (q, n, v, D, a) be the parameters of HFEv-, P_i $(0 \leq i < n-a), M,\ U, W, F^{*j} (0 \leq j < n)$ be the matrices of Equation (5). We set $d = \lceil \log_2 D \rceil$. Assume that U is known, then F^{*0} can be recovered by solving a linear system with $n-a-1$ variables, $(d+a)\cdot(n+v)$ additional linear equations in at most $d+v$ variables, and $\binom{v+1}{2}$ univariate polynomial equations of degree q^d.*

Algorithm 1 Recovering an Equivalent Linear Transformation S

Input: HFEv- parameters (q, n, v, D, a), matrices $(P_0, \cdots, P_{n-a-1})$ representing the quadratic forms of the public key polynomials, matrix $\widetilde{M}$ (see Equation (3)).

Output: Equivalent linear transformation S.

1. Set $\mathbf{b}_i = (1, u_1, \cdots, u_{n+v-1})P_i$, $0 \leq i < n-a$, where $(u_1, \cdots, u_{n+v-1})$ are unknowns.
2. Construct a matrix Z whose row vectors are $\mathbf{b}_i$, $0 \leq i < n-a$. According to Theorem 2, the rank of Z is at most d.
3. Solve the MinRank Problem with matrix Z using one of the methods described in Section 3. Denote the solution by $u_0, u_1, \cdots, u_{n+v-1}$.
4. Set $U = \begin{pmatrix} u_0 & u_1 & \cdots & u_{n+v-1} \\ u_0^q & u_1^q & \cdots & u_{n+v-1}^q \\ \vdots & \vdots & \ddots & \vdots \\ u_0^{q^{n-1}} & u_1^{q^{n-1}} & \cdots & u_{n+v-1}^{q^{n-1}} \\ r_{00} & r_{01} & \cdots & r_{0,n+v-1} \\ \vdots & \vdots & \ddots & \vdots \\ r_{v-1,0} & r_{01} & \cdots & r_{0,n+v-1} \end{pmatrix}$, where $r_{ij}, 0 \leq i < v, 0 \leq j < n+v$ are randomly chosen from the finite field $\mathbb{F}_q$ such that U is invertible.
5. Compute $S' = (\widetilde{M}U)^{-1}$.
6. Return S'.

Proof. From Equation (5) we know that $W = M^{-1}T \in \mathcal{M}_{n\times(n-a)}(\mathbb{F}_{q^n})$. Let $W = \begin{pmatrix} W_1 \\ W_2 \end{pmatrix}$, where $W_1 \in \mathcal{M}_{a\times(n-a)}(\mathbb{F}_{q^n})$ and $W_2 \in \mathcal{M}_{(n-a)\times(n-a)}(\mathbb{F}_{q^n})$. Since M is invertible and the entries of T are randomly chosen from $\mathbb{F}_q$, the probability of W_2 being singular is $1 - \prod_{i=1}^{n-a}(1 - \frac{1}{q^i})$. According to Theorem 1, there are at least q^n equivalent maps T, thus the probability that all matrices W_2 associated to the equivalent maps T are singular is approximately $(1 - \prod_{i=1}^{n-a}(1 - \frac{1}{q^i}))^{q^n}$. Therefore we find an invertible matrix W_2 with overwhelming probability. We multiply both sides of Equation (5) by W_2^{-1}, obtaining

$$\left(UP_0U^t, \cdots, UP_{n-a-1}U^t\right) W_2^{-1} = \left(F^{*0}, \cdots, F^{*n-1}\right) \begin{pmatrix} W_1W_2^{-1} \\ I_{n-a} \end{pmatrix}, \tag{6}$$

where I_{n-a} is the $(n-a)\times(n-a)$ identity matrix. Let $(\widetilde{w}_0, \widetilde{w}_1, \cdots, \widetilde{w}_{n-a-1})$ be the first column of W_2^{-1} and $(\widetilde{l}_0, \widetilde{l}_1, \cdots, \widetilde{l}_{a-1}, 1, 0, \cdots, 0)$ be the first column of $\begin{pmatrix} W_1W_2^{-1} \\ I_{n-a} \end{pmatrix}$, then Equation (6) yields

$$\sum_{k=0}^{n-a-1} \widetilde{w}_k UP_kU^t = \sum_{i=0}^{a-1} \widetilde{l}_i F^{*k} + F^{*a}.$$

We multiply both sides by $\widetilde{l}_0^{-1}$, obtaining

$$\sum_{k=0}^{n-a-1} \widetilde{l}_0^{-1}\widetilde{w}_k U P_k U^t = F^{*0} + \sum_{i=1}^{a-1} \widetilde{l}_0^{-1}\widetilde{l}_i F^{*i} + \widetilde{l}_0^{-1} F^{*a}.$$

Denoting $w_k = \widetilde{l}_0^{-1}\widetilde{w}_k, (k = 0, 1, \cdots, n-a-1)$, and $l_i = \widetilde{l}_0^{-1}\widetilde{l}_i, (i = 1, 2, \cdots, a-1)$, $l_a = \widetilde{l}_0^{-1}$ yields

$$\sum_{k=0}^{n-a-1} w_k U P_k U^t = \sum_{i=1}^{a} l_i F^{*i} + F^{*0}. \tag{7}$$

Note that $\sum_{i=1}^{a} l_i F^{*i} + F^{*0} = \begin{pmatrix} F_0' & 0 & F_1' \\ 0 & 0 & 0 \\ F_1'^t & 0 & F_2' \end{pmatrix} \in \mathcal{M}_{(n+v)\times(n+v)}(\mathbb{F}_{q^n})$, where $F_0' = [f_{ij}']$ is a $(d+a) \times (d+a)$ diagonal band symmetric matrix of width $2d-1$, that is $f_{ij}' = 0$, if $|i-j| \geq d$, $F_1' \in \mathcal{M}_{(d+a)\times v}(\mathbb{F}_{q^n})$, $F_1'^t \in \mathcal{M}_{v\times(d+a)}(\mathbb{F}_{q^n})$ is the transpose of F_1' , $F_2' \in \mathcal{M}_{v\times v}(\mathbb{F}_{q^n})$ is a symmetric matrix .

Assume that $w_0, w_1, \cdots, w_{n-a-1}$ are unknowns. Due to the equivalence of HFEv- private keys [26], we can fix $w_0 = 1$. Since U is known and the special structure of the matrix $\sum_{i=1}^{a} l_i F^{*i} + F^{*0}$, we obtain from Equation (7) $d(n-a-d)$ linear equations in the $n-a-1$ variables $w_1, w_2, \cdots, w_{n-a-1}$. Since $0 < a < n-2d-1$, we have $d(n-a-d) \geq n-a-1$. Therefore by solving these linear equations, we get a solution $(w_0', w_1', w_2', \cdots, w_{n-a-1}')$ with $w_0' = 1$. Thus Equation (7) can be rewritten as

$$\sum_{k=0}^{n-a-1} w_k' U P_k U^t = \sum_{i=1}^{a} l_i F^{*i} + F^{*0}. \tag{8}$$

Now we will find $l_1, \cdots, l_a$ and F^{*0} from Equation (8). We know that F^{*0} has the form

$$F^{*0} = \begin{pmatrix} F_0 & 0 & F_1 \\ 0 & 0 & 0 \\ F_1^t & 0 & F_2 \end{pmatrix},$$

where $F_0 = [\alpha_{ij}] \in \mathcal{M}_{d\times d}(\mathbb{F}_{q^n})$ is a symmetric matrix, $F_1 = [\gamma_{ij}] \in \mathcal{M}_{d\times v}(\mathbb{F}_{q^n})$, $F_1^t \in \mathcal{M}_{v\times d}(\mathbb{F}_{q^n})$ is the transpose of F_1 and $F_2 = [\delta_{ij}] \in \mathcal{M}_{v\times v}(\mathbb{F}_{q^n})$ is a symmetric matrix . According to Proposition 1 we can represent F^{*k} $(1 \leq k \leq n-1)$ by the entries of F^{*0}.

Assume that $l_1, \cdots, l_a$, $\alpha_{ij}(0 \leq i \leq j < d)$, $\gamma_{ij}(0 \leq i < d, 0 \leq j < v)$, $\delta_{ij}(0 \leq i \leq j < v)$ are unknowns. We can recover F^{*0} as follows.

- From the first row of matrix equation (8), we can find a linear system in the variables $\alpha_{0j}(0 \leq j < d)$ and $\gamma_{0j}(0 \leq j < v)$ of the form

  $$\alpha_{00}+\theta_{00} = 0, \cdots, \alpha_{0,d-1}+\theta_{0,d-1} = 0, \gamma_{00}+\theta_{0,d} = 0, \cdots, \gamma_{0,v-1}+\theta_{0,d+v-1} = 0.$$

  Thus we can obtain the first row of F^{*0} by solving this linear system.

- Once the first row of F^{*0} is known, we can obtain from the second row of matrix equation (8) a linear system in the variables l_1 and $\alpha_{1j}(1 \le j < d)$ and $\gamma_{1j}(0 \le j < v)$. By solving this linear system we can obtain the second row of F^{*0} and l_1.
- Similarly, if $a \le d$, we can obtain $l_1, \cdots, l_a$, F_0 and F_1 using the first d rows of matrix equation (8). If $a > d$, we can obtain $l_1, \cdots, l_d$, F_0 and F_1 by using the first d rows of matrix equation (8) and $l_{d+k}(1 \le k \le a-d)$ by using the $(d+k)$-th row of matrix equation (8). Thus we obtain $l_1, \cdots, l_a$, F_0 and F_1.
- Once $l_1, \cdots, l_a$, F_0 and F_1 are known, we get from the last v rows of matrix equation (8), $\binom{v+1}{2}$ univariate polynomial equations of the form

$$\sum_{k=0}^{d} \lambda_{ijk} \delta_{ij}^{q^k} + \eta_{ij} = 0,$$

  where $\lambda_{ijk}, \eta_{ij} \in \mathbb{F}_{q^n}, 0 \le i \le j < v$. Solving these equations we obtain δ_{ij} and then recover F^{*0}.
- Once F^{*0} is known, we can obtain an equivalent central map as

$$\begin{aligned} &F' \ (X, x_1, \ldots, x_v) \\ &= (X, X^q, \cdots, X^{q^{n-1}}, x_1, \cdots, x_v) F^{*0} (X, X^q, \cdots, X^{q^{n-1}}, x_1, \cdots, x_v)^t. \end{aligned}$$

□

Proposition 7. *Let (q, n, v, D, a) be the parameters of HFEv-, P_i $(0 \le i < n-a), S, T, M,$ F^{*j} $(0 \le j < n)$ be the matrices of equation (4). Assume that $S, P_i(0 \le i < n-a), M, F^{*j}(0 \le j < n)$ are known, then T can be recovered by solving $n-a$ linear systems in n variables.*

Proof. Equation (4) can be rewritten as

$$(P_0, \cdots, P_{n-a}) = \left(SMF^{*0}M^tS^t, \cdots, SMF^{*n-1}M^tS^t\right) M^{-1}T. \tag{9}$$

Let $(t_{1k}, t_{2k}, \cdots, t_{nk})$ be the entries of the k-th $(k = 1, 2, \cdots, n-a)$ column of T. Since S, P_i $(0 \le i < n-a), M, F^{*j}(0 \le j < n)$ are known, we obtain from Equation (9) a linear system with $\frac{n(n+1)}{2}$ equations in the n variables $(t_{1k}, t_{2k}, \cdots, t_{nk})$ for all $(k = 1, 2, \cdots, n-a)$. We can recover T by solving $(n-a)$ of these linear systems. □

The process of recovering the maps $\mathcal{F}$ and $\mathcal{T}$ of our equivalent HFEv- key is summarized in Algorithm 2 .

4.4 Complexity of Attack

The most complex step of our attack is step 3 of Algorithm 1. That is the step of solving the MinRank problem on the matrix Z. Z has rank at most d. We can solve it by using minors modeling or support minors modeling.

Algorithm 2 Recovering Equivalent Maps $\mathcal{F}$ and $\mathcal{T}$

Input: HFEv- parameters (q, n, v, D, a), Frobenius matrix M (see (2)), matrices $(P_0, \cdots, P_{n-a-1})$ representing the quadratic forms of the public key polynomials, recovered linear map S.

Output: Equivalent private maps F and T.

1. Let $w_0, w_1, \cdots, w_{n-a-1}$ be unknowns and $w_0 = 1$. Get a linear system with $d(n-d-a)$ equations in the $n-a-1$ variables $w_i, (1 \leq i < n-a-1)$ from matrix equation (7). as shown in the proof of Proposition 6. By solving this linear system we obtain a solution $w_0', w_1', \cdots, w_{n-a-1}'$ with $w_0' = 1$.
2. Let $l_1, \cdots, l_a$ and the nonzero entries of F^{*0} be unknowns in matrix equation (8). We get $(d+a) \cdot (n+v)$ bilinear equations from the first $d+a$ rows of matrix equation (8) and $\binom{v+1}{2}$ univariate polynomial equations from the last v rows of matrix equation (8). By solving these linear systems and univariate polynomial equations we recover F^{*0} (see Proposition 6). Then we can obtain an equivalent central map as
$$F' = (X, X^q, \cdots, X^{q^{n-1}}, x_1, \cdots, x_v) F^{*0} (X, X^q, \cdots, X^{q^{n-1}}, x_1, \cdots, x_v)^t.$$
3. Compute F^{*k} $1 \leq k < n$ according to Proposition 1.
4. Let $(t_{1k}, t_{2k}, \cdots, t_{nk})$ be the (unknown) entries of the k-th $(k = 1, 2, \cdots, n-r)$ column of T. Get $n-r$ linear systems from matrix equation (9) as shown in Proposition 7. By solving these linear systems we can recover an equivalent map T.
5. Return F', T.

If we use minors modeling the degree of regularity of solving the system using the F4 algorithm is given as $d+1$ (c.f. [?]). Therefore, the complexity of our attack with minors modeling is

$$\mathcal{O}\left(\binom{n+v+d+1}{d+1}^{\omega}\right),$$

where $2 < \omega \leq 3$ is the linear algebra constant.

If we use support minors modeling, according to the analysis of section 3, we obtain an overdefined bilinear quadratic system of $n_x + n_y$ variables and $\frac{(n_x+n_y)(n_x+n_y+1)}{2}$ equations, where $n_x = n+v$ and $n_y = \binom{n'}{d}$, $n' = \lceil \frac{(n-a)(d+1)}{n+v} \rceil + d + 1$, $n' < 2d+2$. This bilinear system has at least n solutions. In fact, if $(u_0, u_1, \cdots, u_{n+v-1})$ is a solution of this bilinear system, $(u_0^{q^{i-1}}, u_1^{q^{i-1}}, u_{n+v-1}^{q^{i-1}})$ for all $1 \leq i \leq n$ are also solutions of the bilinear system, more detail can be found in [20]. Therefore, we don't longer have a unique solution as in the case of e.g. Rainbow, which makes the use of Wiedemann inefficient. Thus we use the F4/F5 algorithm to solve it instead of using relinearization method.

By carrying out a series of experiments with Magma, we found that the first degree fall is 3. Since the total number of monomials in the bilinear system is $n_x n_y + n_x + n_y + 1$, the total number of monomials of degree at most 3 is $\mathcal{O}(n_x^2 n_y + n_x n_y^2)$. Thus the complexity of our attack to HFEv- using support minors modeling is $\mathcal{O}\left(n_x^2 n_y + n_x n_y^2\right)^{\omega}$ or $\mathcal{O}\left((n+v)^2\binom{2d+2}{d} + (n+v)\binom{2d+2}{d}^2\right)^{\omega}$. Here, $2 < \omega \leq 3$ is again the linear algebra constant.

5 Application on GeMSS

GeMSS is an HFEv- type signature scheme which is one of the alternative candidates in the third round of the NIST Post Quantum Crypto Standardization Project [7]. The attack complexity on GeMSS using our key recovery attack method can be estimated as shown in Table 1.

Table 1. The estimated gate count of our attack versus the best previously known attack

scheme	parameters (q, n, v, D, a)	best known	minors modeling	support minors modeling
GeMSS128	(2,174,12,513,12)	143	139	118
BlueGeMSS128	(2,175,14,129,13)	143	119	99
RedGeMSS128	(2,177,15,17,15)	143	86	72
GeMSS192	(2,265,20,513,22)	207	154	120
BlueGeMSS192	(2,265,23,129,22)	207	132	101
RedGeMSS192	(2,266,25,17,23)	207	95	75
GeMSS256	(2,354,33,513,30)	272	166	121
BlueGeMSS256	(2,358,32,129,34)	272	141	103
RedGeMSS256	(2,358,35,17,34)	272	101	76

6 Conclusion

In this paper we proposed a new key recovery attack on the HFEv- signature scheme, The complexity of the attack is exponential in the parameter $d = \lceil \log_q(D) \rceil$, but polynomial in n. The Minus modifications does not enhance the security of the HFEv- scheme, while the Vinegar modification only adds a polynomial factor. Therefore, in order to meet the NIST security requirements, a very large D is needed. But the larger the D, the less efficient the signature generation. Thus, it is difficult to use HFEv- scheme to construct a secure and efficient signature scheme.

Related Attacks on Rainbow

We received the paper [2] a few days after we discovered this new attack. In principle, out attack are similar to the the new MinRank in [2].

References

1. M. Bardet, M. Bros, D. Cabarcas, P. Gaborit, R. Perlner, D. Smith-Tone, J.P. Tillich, J. Verbel: Improvements of algebraic attacks for solving the rank decoding and MinRank problems. In International Conference on the Theory and Application of Cryptology and Information Security, 2020.
2. W. Beullens: Improved Attacks on UOV and Rainbow. IACR eprint 2020/1343.
3. L. Bettale, J.C. Faugere, L. Perret: Cryptanalysis of HFE, multi-HFE and variants for odd and even characteristic. Designs, Codes and Cryptography, 69(1), pp. 1-52 (2013)/
4. J. F. Buss, G. S. Frandsen, J. O. Shallit: The computational complexity of some problems of linear algebra. Journal of Computer and System Sciences, 58(3), pp. 572-596 (1999).
5. M. Campagna, K. Chen, Ö. Dagdelen, J. Ding, J.K. Ferrick, N. Gisin et al.: Quantum safe cryptography and security. ETSI White paper 8. Available at https://www.etsi.org/images/files/ETSIWhitePapers/QuantumSafeWhitepaper.pdf (June 2015)
6. R. Cartor, R. Gipson, D. Smith-Tone, J. Vates: On the differential security of the HFEv-signature primitive. PQCrypto 2016, LNCS vol. 9606, pp. 162-181.
7. A. Casanova, J. C. Faugere, G. Macario Rat, J. Patarin, L. Perret, J. Ryckegem: GeMSS: a great multivariate short signature (2019). Submission to NIST PQC competition Round-2.
8. N. Courtois, M. Daum, P. Felke: On the security of HFE, HFEv-and Quartz. PQCrypto 2003, LNCS vol. 2567, pp. 337 - 350.
9. J. Ding, C. Clough, R. Araujo: Inverting square systems algebraically is exponential. Finite Fields and Their Applications 26, pp. 32-46.
10. J. Ding, T.J. Hodges: Inverting HFE systems is quasipolynomial for all fields. CRYPTO 2011, LNCS vol. 6841, pp. 724 - 742.
11. J. Ding, T. Kleinjung: Degree of regularity for HFE Minus (HFE-). Journal of Math for Industry 4, pp. 97 - 104.
12. J. Ding, R. Perlner, A. Petzoldt, D. Smith-Tone: Improved cryptanalysis of hfev-via projection. PQCrypto 2018, LNCS vol. 10786, pp. 375-395.

13. J. Ding, A. Petzoldt: Current state of multivariate cryptography. IEEE Security and Privacy 15 (4), pp. 28-36.
14. J. Ding, A. Petzoldt, D. Schmidt: Multivariate Public Key Cryptosystems - Second Edition. ISBN 978-1-0716-0985-9. Springer, 2020.
15. J. Ding, B.Y. Yang: Degree of regularity for HFEv and HFEv-. PQCrypto 2013, LNCS vol. 7932, pp 52 - 66.
16. J.C. Faugère: A new efficient algorithm for computing Gröbner bases (F4): Journal of pure and applied algebra, 139(1-3), pp. 61-88 (1999).
17. J.C. Faugère, M.S. El Din, P.J. Spaenlehauer: Computing loci of rank defects of linear matrices using Gröbner bases and applications to cryptology. Proceedings of the International Symposium on Symbolic and Algebraic Computation, pp. 257-264 (2010).
18. P. Gaborit, O. Ruatta, J. Schrek: On the complexity of the rank syndrome decoding problem. IEEE Transactions on Information Theory, 62(2), pp. 1006-1019, 2016.
19. L. Goubin, N. Courtois: Cryptanalysis of the TTM Cryptosystem. ASIACRYPT 2000, LNCS vol. 1976, pp. 44-57.
20. X. Jiang, J. Ding, L. Hu: Kipnis-Shamir attack on HFE revisited. Inscrypt 2007, LNCS vol. 4990, pp. 399-411.
21. A. Kipnis, A. Shamir A. Cryptanalysis of the HFE public key cryptosystem by relinearization. CRYPTO 99, LNCS vol. 1666, pp. 19-30.
22. M.S.E. Mohamed, D. Cabarcas, J. Ding, J. Buchmann, S. Bulygin: MXL 3: an efficient algorithm for computing gröbner bases of zero-dimensional ideals. ICISC 2009, LNCS vol. 5984, pp. 87-100.
23. J. Patarin: Hidden fields equations (HFE) and isomorphisms of polynomials (IP): Two new families of asymmetric algorithms. EUROCRYPT 1996, LNCS vol. 1070, pp. 33-48.
24. J. Patarin, N. Courtois, L. Goubin: Quartz, 128-bit long digital signatures. Cryptographers Track at the RSA Conference, LNCS vol. 2020, pp. 282-297.
25. A. Petzoldt, M.S. Chen, B.Y. Yang, C. Tao, J. Ding: Design principles for HFEv-based multivariate signature schemes. ASIACRYPT 2015(1), LNCS vol. 9452, pp. 311-334.
26. C. Wolf, B. Preneel: Equivalent keys in multivariate uadratic public key systems. Journal of Mathematical Cryptology 4(4), pp. 375-415 (2011).

A Example of the Attack

To illustrate our new attack method, we present a complete key recovery for a toy example of the HFEv- scheme over a small field. Let the parameters of our HFEv-instance be $(q, n, v, D, a) = (7, 7, 2, 14, 2)$. Then we have $d = \lceil \log_q(D) \rceil = 2$. We construct the degree n extension field $\mathbb{F}_{q^n} = \mathbb{F}_q[x]/\langle x^7 + 6x + 4 \rangle$. Let θ be a primitive root of the irreducible polynomial $p(x) = x^7 + 6x + 4$.

We randomly generate central map $F = \theta^{176932} X^{14} + \theta^{461287} X^8 + \theta^{199902} X^2 + (\theta^{270502} x_1 + \theta^{358630} x_2) X + (\theta^{65557} x_1 + \theta^{2597} x_2) X^7 + \theta^{811326} x_1^2 + \theta^{14415} x_1 x_2 + \theta^{151050} x_2^2$. The linear transformations $\mathcal{S}$ and $\mathcal{T}$ are given by the matrices

$$S = \begin{pmatrix} 3 & 1 & 1 & 6 & 4 & 2 & 0 & 1 & 6 \\ 6 & 2 & 4 & 5 & 3 & 3 & 2 & 6 & 0 \\ 6 & 1 & 3 & 4 & 4 & 2 & 4 & 5 & 3 \\ 0 & 1 & 4 & 6 & 4 & 2 & 2 & 3 & 1 \\ 2 & 0 & 0 & 5 & 2 & 4 & 2 & 1 & 3 \\ 0 & 5 & 1 & 2 & 4 & 2 & 1 & 4 & 3 \\ 3 & 3 & 5 & 0 & 2 & 6 & 4 & 6 & 6 \\ 5 & 2 & 0 & 2 & 5 & 6 & 3 & 1 & 2 \\ 6 & 2 & 5 & 5 & 5 & 4 & 3 & 6 & 1 \end{pmatrix} \text{ and } T = \begin{pmatrix} 1 & 4 & 4 & 6 & 5 \\ 0 & 6 & 5 & 3 & 2 \\ 0 & 2 & 0 & 2 & 2 \\ 1 & 3 & 1 & 0 & 1 \\ 2 & 4 & 2 & 5 & 3 \\ 3 & 4 & 1 & 0 & 6 \\ 6 & 5 & 6 & 5 & 0 \end{pmatrix}.$$

We compute the public key as $\mathcal{P} = \mathcal{T} \circ \mathcal{F} \circ \mathcal{S}$. The quadratic forms representing the public key polynomials are given as

$$P_0 = \begin{pmatrix} 1&2&0&3&3&6&1&3&3\\ 2&6&0&4&4&3&4&4&3\\ 0&0&3&5&4&4&4&5&3\\ 3&4&5&2&1&1&3&2&1\\ 3&4&4&1&0&2&1&6&2\\ 6&3&4&1&2&5&0&5&1\\ 1&4&4&3&1&0&6&0&0\\ 3&4&5&2&6&5&0&3&2\\ 3&3&3&1&2&1&0&2&1 \end{pmatrix}, P_1 = \begin{pmatrix} 4&0&3&3&5&6&6&3&2\\ 0&3&0&6&1&1&0&4&4\\ 3&0&3&3&5&4&5&5&4\\ 3&6&3&1&6&6&2&3&5\\ 5&1&5&6&1&6&3&6&4\\ 6&1&4&6&6&5&3&3&1\\ 6&0&5&2&3&3&0&0&5\\ 3&4&5&3&6&3&0&2&1\\ 2&4&4&5&4&1&5&1&6 \end{pmatrix}, P_2 = \begin{pmatrix} 3&2&6&4&5&2&6&6&2\\ 2&5&1&0&6&4&1&5&4\\ 6&1&6&0&0&5&0&3&3\\ 4&0&0&5&5&5&5&2&2\\ 5&6&0&5&1&2&1&6&0\\ 2&4&5&5&2&4&1&5&0\\ 6&1&0&5&1&1&4&4&5\\ 6&5&3&2&6&5&4&4&4\\ 2&4&3&2&0&0&5&4&0 \end{pmatrix},$$

$$P_3 = \begin{pmatrix} 2&6&4&5&4&1&6&0&1\\ 6&6&6&1&2&1&0&6&3\\ 4&6&2&6&1&5&0&4&6\\ 5&1&6&0&0&0&0&3&5\\ 4&2&1&0&6&1&6&0&4\\ 1&1&5&0&1&2&6&3&5\\ 6&0&0&0&6&6&5&6&1\\ 0&6&4&3&0&3&6&2&0\\ 1&3&6&5&4&5&1&0&1 \end{pmatrix} P_4 = \begin{pmatrix} 3&0&5&4&5&6&0&5&2\\ 0&3&0&3&3&5&4&2&2\\ 5&0&4&2&4&6&1&1&3\\ 4&3&2&3&4&3&2&6&1\\ 5&3&4&4&1&2&3&3&6\\ 6&5&6&3&2&4&0&0&2\\ 0&4&1&2&3&0&6&5&1\\ 5&2&1&6&3&0&5&5&0\\ 2&2&3&1&6&2&1&0&3 \end{pmatrix},$$

Let $M = \begin{pmatrix} 1&1&1&1&1&1&1\\ \theta&\theta^{7}&\theta^{49}&\theta^{343}&\theta^{2401}&\theta^{16807}&\theta^{117649}\\ \theta^{2}&\theta^{14}&\theta^{98}&\theta^{686}&\theta^{4802}&\theta^{33614}&\theta^{235298}\\ \theta^{3}&\theta^{21}&\theta^{147}&\theta^{1029}&\theta^{7203}&\theta^{50421}&\theta^{352947}\\ \theta^{4}&\theta^{28}&\theta^{196}&\theta^{1372}&\theta^{9604}&\theta^{67228}&\theta^{470596}\\ \theta^{5}&\theta^{35}&\theta^{245}&\theta^{1715}&\theta^{12005}&\theta^{84035}&\theta^{588245}\\ \theta^{6}&\theta^{42}&\theta^{294}&\theta^{2058}&\theta^{14406}&\theta^{100842}&\theta^{705894} \end{pmatrix}$ and $\widetilde{M} = \begin{pmatrix} M & 0\\ 0 & I_v \end{pmatrix}$ In the following we demonstrate our method to recover the private key from $\mathcal{P}$.

A.1 Recovering $\mathcal{S}$

Let the first row of matrix $U = \widetilde{M}^{-1}S^{-1}$ be $(u_0, u_1, \cdots, u_{n+v-1})$. Fix $u_0 = 1$ and let $u_1, \cdots, u_{n+v-1}$ be unknowns. Set $\mathbf{b}_i = (1, u_1, \cdots, u_{n+v-1})P_i, i = 0, 1, \cdots, n-a-1$. Let $\mathbf{b}_i$ be the i-th row of the matrix Z. Then the rank of Z is 2. This implies that all minors of order 3 are 0. Solving the MinRank Problem for matrix Z gives us a solution $\mathbf{u} = (1, \theta^{2689}, \theta^{240750}, \theta^{393451}, \theta^{682468}, \theta^{184068}, \theta^{218176}, \theta^{85224}, \theta^{760002})$. Then we have

$$U = \begin{pmatrix} 1&\theta^{2689}&\theta^{240750}&\theta^{393451}&\theta^{682468}&\theta^{184068}&\theta^{218176}&\theta^{85224}&\theta^{760002}\\ 1&\theta^{18823}&\theta^{38166}&\theta^{283531}&\theta^{659566}&\theta^{464934}&\theta^{703690}&\theta^{596568}&\theta^{378762}\\ 1&\theta^{131761}&\theta^{267162}&\theta^{337633}&\theta^{499252}&\theta^{783912}&\theta^{808120}&\theta^{58266}&\theta^{180708}\\ 1&\theta^{98785}&\theta^{223050}&\theta^{716347}&\theta^{200596}&\theta^{546132}&\theta^{715588}&\theta^{407862}&\theta^{441414}\\ 1&\theta^{691495}&\theta^{737808}&\theta^{73177}&\theta^{580630}&\theta^{528756}&\theta^{67864}&\theta^{384408}&\theta^{619272}\\ 1&\theta^{722755}&\theta^{223404}&\theta^{512239}&\theta^{770242}&\theta^{407124}&\theta^{475048}&\theta^{220230}&\theta^{217194}\\ 1&\theta^{118033}&\theta^{740286}&\theta^{291505}&\theta^{450442}&\theta^{379242}&\theta^{31168}&\theta^{718068}&\theta^{696816}\\ 1&5&1&0&1&3&0&3&2\\ 4&6&1&5&4&5&5&6&6 \end{pmatrix},$$

where the last v rows of U are randomly chosen from $\mathbb{F}_q$, such that U is invertible.

Thus we can recover an equivalent linear transformation $\mathcal{S}$ as

$$S' = U^{-1}\widetilde{M}^{-1} = \begin{pmatrix} 0&1&1&2&3&6&6&0&6\\ 1&4&5&3&1&6&0&4&6\\ 4&5&3&1&5&6&0&6&4\\ 5&0&1&2&5&6&0&2&0\\ 2&3&1&3&5&6&0&3&1\\ 1&6&5&0&4&1&0&4&1\\ 0&4&6&4&2&2&0&6&2\\ 2&1&5&2&5&1&2&1&2\\ 6&0&2&6&4&6&1&5&6 \end{pmatrix}.$$

Recovering $\mathcal{F}$ and $\mathcal{T}$ Step 1. Once $\mathcal{S}$ is known, let $w_0, w_1, \cdots, w_{n-a-1}$ be unknowns and $w_0 = 1$. We generate a linear system with $d(n-d-a)$ equations in the $n-a-1$ variables $w_i, (1 \leq i < n-a-1)$ using the matrix equation (7). By solving this linear system we obtain a solution $(1, \theta^{558954}, \theta^{326166}, \theta^{142979}, \theta^{806014})$.

Step 2. Let $l_1, \cdots, l_a$ and the nonzero entries of F^{*0} be variables in matrix equation (8). By using the first $d+a$ rows of matrix equation (8) we get $(d+a) \cdot (n+v)$ bilinear equations as follows:

$$\begin{pmatrix} \alpha_{00}+\theta^{599798} & \alpha_{01}+\theta^{499519} & 0 & 0 & 0\ 0\ 0 & \gamma_{00}+\theta^{424284} & \gamma_{01}+\theta^{665059} \\ \alpha_{10}+\theta^{499519} & \alpha_{00}^{7}l_1+\alpha_{11}+\theta^{381840} & \alpha_{01}^{7}l_1+\theta^{349085} & 0 & 0\ 0\ 0 & \gamma_{00}^{7}l_1+\gamma_{10}+\theta^{228693} & \gamma_{01}^{7}l_1+\gamma_{11}+\theta^{396254} \\ 0 & \alpha_{10}^{7}l_1+\theta^{349085} & \alpha_{00}^{49}l_2+\alpha_{11}^{7}l_1+\theta^{622586} & \alpha_{01}^{49}l_2+\theta^{524551} & 0\ 0\ 0 & \gamma_{00}^{49}l_2+\gamma_{10}^{7}l_1+\theta^{475138} & \gamma_{01}^{49}l_2+\gamma_{11}^{7}l_1+\theta^{2659} \\ 0 & 0 & \alpha_{10}^{49}l_2+\theta^{524551} & \alpha_{11}^{49}l_2+\theta^{32832} & 0\ 0\ 0 & \gamma_{10}^{49}l_2+\theta^{9738} & \gamma_{11}^{49}l_2+\theta^{392135} \end{pmatrix}$$

$= 0_{(d+a)\times(n+v)}$.
From the first row, we obtain $\alpha_{00} = \theta^{188027}, \alpha_{01} = \theta^{87748}, \gamma_{00} = \theta^{12513}, \gamma_{01} = \theta^{253288}$. Once α_{00}, α_{01} are known, we get from the second row $\alpha_{10} = \theta^{87748}, \alpha_{11} = \theta^{10485}, \gamma_{10} = \theta^{581451}, \gamma_{11} = \theta^{606062}, l_1 = \theta^{146620}$. From the third row we can obtain $l_2 = \theta^{754380}$.

Once l_1, l_2 are known, we get from the last v rows of matrix equation (8), $\binom{v+1}{2}$ univariate polynomial equations as follows:

$$\begin{aligned} \theta^{754380}\delta_{00}^{49} + \theta^{146620}\delta_{00}^{7} + \delta_{00} + \theta^{81317} = 0, \\ \theta^{754380}\delta_{01}^{49} + \theta^{146620}\delta_{01}^{7} + \delta_{01} + \theta^{689914} = 0, \\ \theta^{754380}\delta_{11}^{49} + \theta^{146620}\delta_{11}^{7} + \delta_{11} + \theta^{162754} = 0. \end{aligned}$$

Each of these equations has 49 solutions. We choose one of them as the value of δ_{ij}. Thus we have $\delta_{00} = \theta^{27191}, \delta_{01} = \delta_{10} = \theta^{19044}, \delta_{11} = \theta^{9718}$ and

$$F^{*0} = \begin{pmatrix} \theta^{188027} & \theta^{87748} & 0 & 0 & 0 & 0 & 0 & \theta^{12513} & \theta^{253288} \\ \theta^{87748} & \theta^{10485} & 0 & 0 & 0 & 0 & 0 & \theta^{581451} & \theta^{606062} \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \theta^{12513} & \theta^{581451} & 0 & 0 & 0 & 0 & 0 & \theta^{27191} & \theta^{19044} \\ \theta^{253288} & \theta^{606062} & 0 & 0 & 0 & 0 & 0 & \theta^{19044} & \theta^{9718} \end{pmatrix}$$

Therefore we get an equivalent central map as $F' = \theta^{10485}X^{14} + \theta^{362262}X^8 + \theta^{188027}X^2 + (\theta^{287027}x_1 + \theta^{527802}x_2)X + (\theta^{32423}x_1 + \theta^{57034}x_2)X^7 + \theta^{27191}x_1^2 + \theta^{293558}x_1x_2 + \theta^{9718}x_2^2$ for F.

Let $(t_{1k}, t_{2k}, \cdots, t_{nk})$ be entries of the k-th $(k = 1, 2, \cdots, n-a)$ column of T. Get $n-a$ linear systems from matrix equation (9) as shown by Proposition 7. By solving these linear systems we can recover a equivalent key of T as follows

$$T' = \begin{pmatrix} 1 & 1 & 6 & 0 & 5 \\ 3 & 3 & 2 & 0 & 2 \\ 1 & 3 & 2 & 5 & 6 \\ 6 & 6 & 6 & 0 & 2 \\ 2 & 2 & 3 & 3 & 6 \\ 2 & 2 & 1 & 0 & 5 \\ 0 & 5 & 1 & 3 & 0 \end{pmatrix}.$$

It is easy to check that $\mathcal{P} = \mathcal{T} \circ \mathcal{F} \circ \mathcal{S} = \mathcal{T}' \circ \mathcal{F}' \circ \mathcal{S}'$. Therefore the adversary can use the three maps $\mathcal{T}'$, $\mathcal{F}'$ and $\mathcal{S}'$ to forge signatures for arbitrary messages.