HMFEv - An Efficient Multivariate Signature Scheme

Albrecht Petzoldt[1], Ming-Shing Chen[2], Jintai Ding[3], Bo-Yin Yang[2]

[1] National Institute for Standards and Technology, Gaithersburg, Maryland, USA
[2] Academia Sinica, Taipei, Taiwan
[3] University of Cincinnati, Ohio, USA
albrecht.petzoldt@nist.gov,jintai.ding@gmail.com,{mschen,byyang}@crypto.tw

Abstract. Multivariate Cryptography, as one of the main candidates for establishing post-quantum cryptosystems, provides strong, efficient and well-understood digital signature schemes such as UOV, Rainbow, and Gui. While Gui provides very short signatures, it is, for efficiency reasons, restricted to very small finite fields, which makes it hard to scale it to higher levels of security and leads to large key sizes.
In this paper we propose a signature scheme called HMFEv ("Hidden Medium Field Equations"), which can be seen as a multivariate version of HFEv. We obtain our scheme by applying the Vinegar Variation to the MultiHFE encryption scheme of Chen et al.. We show both theoretically and by experiments that our new scheme is secure against direct and Rank attacks. In contrast to other schemes of the HFE family such as Gui, HMFEv can be defined over arbitrary base fields and therefore can be much more efficient in terms of both performance and memory requirements. Our scheme is therefore a good candidate for the upcoming standardization of post-quantum signature schemes.

Keywords: Post-Quantum Cryptography, Multivariate Cryptography, Signature Schemes, NIST Call for Proposals

1 Introduction

Multivariate Public Key Cryptosystems (MPKCs) are one of the main candidates for guaranteeing the security of communication in a quantum world [1]. Multivariate schemes are in general very fast and require only modest computational resources, which makes them attractive for the use on low cost devices like smart cards and RFID chips [4,6]. Additionally, at least in the area of digital signatures, there exists a large number of practical multivariate schemes.
The existing multivariate signature schemes can be divided into two main groups. The first are the SingleField schemes UOV and Rainbow, which follow the same type of design strategy using Oil-Vinegar polynomials. We believe that these two schemes are more or less the best which can be achieved from this fundamental design.

On the other hand, we have the BigField schemes HFEv- and Gui, which combine the HFE design with the Minus and Vinegar modifiers. These schemes make use of an HFE polynomial, whose degree D is very much affected by the size of the underlying field. We believe that, for security reasons, this degree should be chosen at least q^2+1 , where q is the cardinality of the underlying field. However, during the signature generation process, we have to invert this univariate HFE polynomial and the complexity of this step can be estimated by $\mathcal{O}(D^3)$. To solve this conflict between security and efficiency, we have to build the scheme over very small finite fields such as GF(2) and GF(4). However, in this case, we have to choose the number of variables to be large, which leads to large key sizes and makes the scheme less efficient. Therefore it is a natural question, if it is possible to use large base fields such as GF(31) or GF(256) for the design of multivariate signature schemes of the HFEv- type.
In 2008, Chen et al. proposed a multivariate encryption scheme called MultiHFE [7], which can be seen as a multivariate version of HFE. While the scheme is very efficient, its security appeared to be weak and it was broken by Bettale et al. [3] by a generalization of the Kipnis-Shamir attack against HFE using the MinRank property of the system.
In this paper, we propose a signature scheme called HMFEv ("Hidden Medium Field Equations"), which we obtain by applying the Vinegar modification to MultiHFE. We show both theoretically and by experiments that our scheme is secure against direct and Rank attacks of the Kipnis-Shamir / Bettale type and analyze the security of our scheme against other known attacks against multivariate schemes, including differential attacks and Hashimotos attack against the MultiHFE encryption scheme. Our scheme can be seen as an extension of the Gui and QUARTZ signature schemes. However, by enabling a flexible choice of the base field, our new scheme overcomes a fundamental practical problem in the HFEV- design. While Gui and QUARTZ are, for efficiency reasons, mainly restricted to the field GF(2), our scheme allows the choice of an arbitrary base field. This allows us to reduce the number of equations and variables in the public system significantly, which leads to smaller key sizes and more efficient signature generation and verification processes. Furthermore, this enables an easy scalability of our scheme to higher levels of security. Our scheme is therefore a very strong candidate for the upcoming standardization of post-quantum signature schemes.
The rest of this paper is organized as follows. Section 2 gives an overview of the basic concepts of multivariate cryptography. In Section 3 we describe the MultiHFE encryption scheme which is the basis of our construction and analyze its security and efficiency. Section 4 describes our new HMFEv signature scheme in detail. In Section 5 we analyze the security of our scheme, in particular its behavior against direct and rank attacks. Section 6 proposes concrete parameter sets for our scheme for different levels of security. Section 7 compares our HMFEv scheme with other multivariate signature schemes of the HFEv- type, in particular Gui, and Section 8 concludes the paper.

2 Multivariate Cryptography

The public key of a multivariate public key cryptosystem (MPKC) is a set of multivariate quadratic polynomials. The security of these schemes is based on the MQ Problem of solving such a system. The MQ problem (for $m \approx n$) is proven to be NP-hard even for quadratic polynomials over the field GF(2) [14] and believed to be hard on average (both for classical and quantum computers). To build a public key cryptosystem based on the MQ problem, one starts with an easily invertible quadratic map $\mathcal{F} : \mathbb{F}^n \to \mathbb{F}^m$ (central map). To hide the structure of $\mathcal{F}$ in the public key, one composes it with two invertible affine (or linear) maps $\mathcal{S} : \mathbb{F}^m \to \mathbb{F}^m$ and $\mathcal{T} : \mathbb{F}^n \to \mathbb{F}^n$. The *public key* of the scheme is therefore given by $\mathcal{P} = \mathcal{S} \circ \mathcal{F} \circ \mathcal{T} : \mathbb{F}^n \to \mathbb{F}^m$. The *private key* consists of $\mathcal{S}$, $\mathcal{F}$ and $\mathcal{T}$ and therefore allows to invert the public key.
In this paper we concentrate on multivariate schemes of the MediumField familiy. For this type of schemes, one chooses two integers k and ℓ and sets $n = k \cdot \ell$. The central map $\mathcal{F}$ of the scheme is a specially chosen easily invertible polynomial map over the vector space $\mathbb{E}^k$, where $\mathbb{E}$ is a degree ℓ extension field of $\mathbb{F}$. Using an isomorphism $\phi : \mathbb{F}^\ell \to \mathbb{E}$ we can transform $\mathcal{F}$ into a map

$$\bar{\mathcal{F}} = \underbrace{(\phi^{-1} \times \cdots \times \phi^{-1})}_{k-\text{times}} \circ \mathcal{F} \circ \underbrace{(\phi \times \cdots \times \phi)}_{k-\text{times}} : \mathbb{F}^n \to \mathbb{F}^n. \tag{1}$$

from $\mathbb{F}^n$ to itself. The map $\mathcal{F}$ is chosen in such a way that the map $\bar{\mathcal{F}}$ consists of multivariate quadratic polynomials. The *public key* has the form $\mathcal{P} = \mathcal{S} \circ \mathcal{F} \circ \mathcal{T}$ with two invertible affine maps $\mathcal{S}, \mathcal{T} : \mathbb{F}^n \to \mathbb{F}^n$, the *private key* consists of $\mathcal{S}, \mathcal{F}$ and $\mathcal{T}$.

3 The MultiHFE scheme

An important example for a multivariate scheme from the MediumField family is the MultiHFE scheme of Chen et al. [7]. In its basic version, the scheme can be used both as an encryption and signature scheme.

The k components $\mathcal{F}^{(1)}, \ldots, \mathcal{F}^{(k)}$ of the central map $\mathcal{F}$ are of the form

$$\mathcal{F}^{(i)} = \sum_{1 \le r \le s \le k} \alpha_{rs}^{(i)} \cdot X_r X_s + \sum_{1 \le r \le s} \beta_r^{(i)} \cdot X_r + \gamma^{(i)} \quad (i = 1, \ldots, k) \tag{2}$$

with coefficients $\alpha_{rs}^{(i)}$, $\beta_r^{(i)}$ and $\gamma^{(i)} \in \mathbb{E}$. Note that the polynomials $\mathcal{F}^{(i)}$ $(i = 1, \ldots, k)$ are multivariate polynomials of the HFE type with $D = 2$. The map $\bar{\mathcal{F}}$ of the MultiHFE signature scheme is defined as shown in equation (1) and is, due to the Frobenius isomorphism, a multivariate quadratic map over the vector space $\mathbb{F}^n$. To hide the structure of $\bar{\mathcal{F}}$ in the public key, one composes it with two invertible affine maps $\mathcal{S}$ and $\mathcal{T} : \mathbb{F}^n \to \mathbb{F}^n$. Therefore, the *public key* of the scheme is given by $\mathcal{P} = \mathcal{S} \circ \bar{\mathcal{F}} \circ \mathcal{T} : \mathbb{F}^n \to \mathbb{F}^n$, the *private key* consists of $\mathcal{S}$, $\mathcal{F}$ and $\mathcal{T}$.

Signature Generation: In order to generate a signature for a message d one uses a hash function $\mathcal{H} : \{0,1\}^\star \to \mathbb{F}^n$ to compute the hash value $\mathbf{w} = \mathcal{H}(d) \in \mathbb{F}^n$ and performs the following three steps.

1. Compute $\mathbf{x} = \mathcal{S}^{-1}(\mathbf{w}) \in \mathbb{F}^n$ and lift the result to the vector space $\mathbb{E}^k$. Denote the result by $\mathbf{X}$.
2. Invert the central map $\mathcal{F}$ to obtain $\mathbf{Y} = \mathcal{F}^{-1}(\mathbf{X}) \in \mathbb{E}^k$ and compute $\mathbf{y} = (\phi^{-1} \times \cdots \times \phi^{-1})(\mathbf{Y}) \in \mathbb{F}^n$. Since $\mathcal{F}$ is a system of k randomly chosen quadratic polynomials in k variables, we need for this step a system solver like XL [24] or a Gröbner Basis algorithm such as F_4 [13] or F_5.
3. Compute the signature $\mathbf{z} \in \mathbb{F}^n$ by $\mathbf{z} = \mathcal{T}^{-1}(\mathbf{y})$.

Verification: To check, if $\mathbf{z} \in \mathbb{F}^n$ is indeed a valid signature for a message d, one computes the hash value $\mathbf{w} = \mathcal{H}(d)$ and $\mathbf{w}' = \mathcal{P}(\mathbf{z})$. If $\mathbf{w}' = \mathbf{w}$ holds, the signature is accepted, otherwise rejected.

3.1 Efficiency

The most complex step during the decryption process of MultiHFE is the solution of the multivariate quadratic system $\mathcal{F}(Y_1, \ldots, Y_k) = (X_1, \ldots, X_k)$ (k equations in k variables) over the extension field $\mathbb{E}$. Since the coefficients of the system $\mathcal{F}$ are chosen randomly, this step has to be performed by a system solver like XL [24] or a Gröbner Bases algorithm such as F_4 [13]. If the number k of equations and variables in this system is small, these algorithms can invert $\mathcal{F}$ very efficiently. However, when the parameter k gets larger, the decryption process of MultiHFE becomes very costly and the scheme therefore gets inefficient.

3.2 The Rank Attack against HFE and MultiHFE

In [17], Kipnis and Shamir proposed a rank based attack against the univariate HFE scheme. The key idea of this attack is to lift all the maps $\mathcal{S}$, $\mathcal{P}$ and $\mathcal{T}$ to univariate maps $\mathcal{S}^\star$, $\mathcal{P}^\star$ and $\mathcal{T}^\star$ over the extension field $\mathbb{E}$. Since the rank of the central map $\mathcal{F}$ is bounded from above by $r = \lfloor \log_q(D-1) \rfloor + 1$, this enabled them to recover the private key by solving an instance of a MinRank problem. However, since computing the map $\mathcal{P}^\star$ appeared to be very costly, the attack of Kipnis and Shamir is not very efficient.
Later, Bettale et al. [3] found a way to perform the attack of Kipnis and Shamir without the need of recovering the map $\mathcal{P}^\star$. Besides improving the efficiency of the Kipnis-Shamir attack, this makes it much easier to extend the attack to MultiHFE. Due to lack of space we cannot present all the details of the attacks of Kipnis-Shamir and Bettale here and refer to the papers [17], [3] and the extended version of this paper for a detailed analysis of the attacks. Here, we just present the main results of [3].

Theorem 1. *For MultiHFE, recovering the affine transformation $\mathcal{T}$ reduces to simultaneously solving k MinRank problems over the base field.*

With this, Bettale et al. could further prove

Theorem 2. *The complexity of solving the MultiHFE MinRank problem is $\mathcal{O}(\ell^{(k+1)\omega})$ with $2 < \omega \leq 3$ being the linear algebra constant and ℓ being the degree of the field extension $\mathbb{E}|\mathbb{F}$.*

We therefore face the following problem: If the parameter k in MultiHFE is small, the scheme can be easily broken by the MinRank attack. On the other hand, if we choose k larger, the efficiency of the scheme becomes quite bad.
In the following we show how to solve this dilemma by modifying the MultiHFE scheme.

4 The new Signature Scheme HMFEv

Let $\mathbb{F}$ be a finite field and k, ℓ and v be integers. We set $n = k \cdot \ell$. Furthermore, let $g(X) \in \mathbb{F}[X]$ be an irreducible polynomial of degree ℓ and $\mathbb{E} = \mathbb{F}[X]/g(X)$ the corresponding extension field. We define an isomorphism $\phi : \mathbb{F}^\ell \to \mathbb{E}$ by

$$\phi(x_1, \ldots, x_\ell) = \sum_{i=1}^{\ell} x_i \cdot X^{i-1}.$$

The *central map* $\mathcal{F} : \mathbb{E}^k \times \mathbb{F}^v \to \mathbb{E}^k$ of the scheme consists of k components $\mathcal{F}^{(1)}, \ldots, \mathcal{F}^{(k)}$ of the form

$$\mathcal{F}^{(i)} = \sum_{r,s=1}^{k} \alpha_{rs}^{(i)} \cdot X_r X_s + \sum_{r=1}^{k} \beta_r^{(i)}(v_1, \ldots, v_v) \cdot X_r + \gamma^{(i)}(v_1, \ldots, v_v)$$

with coefficients $\alpha_{rs}^{(i)} \in \mathbb{E}$, linear functions $\beta_r^{(i)} : \mathbb{F}^v \to \mathbb{E}$ and quadratic maps $\gamma^{(i)} : \mathbb{F}^v \to \mathbb{E}$ ($i \in \{1, \ldots, k\}$).
Due to the special form of $\mathcal{F}$, the map

$$\bar{\mathcal{F}} = \underbrace{(\phi^{-1} \times \cdots \times \phi^{-1})}_{k-\text{times}} \circ \mathcal{F} \circ (\underbrace{\phi \times \cdots \times \phi}_{k-\text{times}} \times \mathrm{id}_v)$$

is a multivariate quadratic map from $\mathbb{F}^{n+v}$ to $\mathbb{F}^n$. Here, id_v is the identity map over the vector space $\mathbb{F}^v$.
To hide the structure of $\bar{\mathcal{F}}$ in the public key, we combine it with two randomly chosen invertible affine maps $\mathcal{S} : \mathbb{F}^n \to \mathbb{F}^n$ and $\mathcal{T} : \mathbb{F}^{n+v} \to \mathbb{F}^{n+v}$.
The *public key* of the scheme is given by

$$\mathcal{P} = \mathcal{S} \circ \bar{\mathcal{F}} \circ \mathcal{T} : \mathbb{F}^{n+v} \to \mathbb{F}^n,$$

the *private key* consists of $\mathcal{S}$, $\mathcal{F}$ and $\mathcal{T}$.

Signature Generation: To generate a signature for a document d, we use a hash function $\mathcal{H} : \{0,1\}^\star \to \mathbb{F}^n$ to compute the hash value $\mathbf{w} = \mathcal{H}(d) \in \mathbb{F}^n$. After that, we perform the following six steps

1. Compute $\mathbf{x} = \mathcal{S}^{-1}(\mathbf{w})$.
2. Lift the result to the extension field $\mathbb{E}$ by computing $X_i = \phi(x_{(i-1)\cdot\ell+1}, \ldots, x_{i\cdot\ell})$ $(i = 1, \ldots, k)$.
3. Choose random values for the Vinegar variables $v_1, \ldots, v_v \in \mathbb{F}$ and substitute them into the central map components to obtain the parametrized maps $\mathcal{F}_V^{(1)}, \ldots, \mathcal{F}_V^{(k)}$.
4. Use the XL-Algorithm or a Gröbner basis method to compute $Y_1, \ldots, Y_k$ such that $\mathcal{F}_V^{(i)}(Y_1, \ldots, Y_k) = X_i$ $(i = 1, \ldots, k)$.
5. Move the result down to the vector space by computing $\mathbf{y} = (\phi^{-1}(Y_1), \ldots, \phi^{-1}(Y_k), v_1, \ldots, v_v) \in \mathbb{F}^{n+v}$.
6. Compute the signature $\mathbf{z} \in \mathbb{F}^{n+v}$ by $\mathbf{z} = \mathcal{T}^{-1}(\mathbf{y})$.

Signature Verification: In order to check, if $\mathbf{z} \in \mathbb{F}^{n+v}$ is indeed a valid signature for the document d, one computes $\mathbf{w} = \mathcal{H}(d)$ and $\mathbf{w}' = \mathcal{P}(\mathbf{z})$. If $\mathbf{w}' = \mathbf{w}$ holds, the signature is accepted, otherwise rejected.

5 Security

In this Section we analyze the security of our scheme. In particular we study both theoretically and using computer experiments the behavior of our scheme against direct and rank attacks. .

5.1 Direct and Rank attacks

The complexity of a direct attack is closely related to the degree of regularity of the system. Therefore the key task is to study the degree of regularity of a direct attack against our scheme.
From the work of Ding and Hodges in Crypto 2011 [11] we know that the degree of regularity of a direct attack against an HFE scheme can be estimated by looking at a single polynomial in the extension field $\mathbb{E}$, and the rank of the associated quadratic form.
In the case of HMFEv, the situation is slightly different, but still very similar. For HMFEv, the components of the public key come from several polynomials over the medium field, which are given as

$$\mathcal{F}^{(i)} = \sum_{r,s=1}^{k} \alpha_{rs}^{(i)} \cdot X_r X_s + \sum_{r=1}^{k} \beta_r^{(i)}(v_1, \ldots, v_v) \cdot X_r + \gamma^{(i)}(v_1, \ldots, v_v) \quad (1 \leq i \leq k).$$

Using the same argument as in the work of Ding and Yang in [10] we can, under the assumption of $v \leq \ell$, lift each map $\mathcal{F}^{(i)}$ $(1 \leq i \leq k)$, which is a map from $\mathbb{E}^k \times \mathbb{F}^v$ to $\mathbb{E}$, to a map $\mathcal{F}'^{(i)}$ from $\mathbb{E}^{k+1}$ to $\mathbb{E}$. Here, the additional component in the domain comes from the use of the vinegar variables. Then we can look at the rank of the quadratic form associated to the polynomial $\mathcal{F}'^{(i)}$ as in the case of the original Kipnis-Shamir attack.
Using the same method as in [10] we can prove

Theorem 3. *If $v \leq \ell$ holds, the rank of the quadratic form associated to $\mathcal{F}'^{(i)}$ is greater or equal to $k + v$.*

The proof follows directly from that in [10].
This theorem directly gives us a lower bound for the complexity of the MinRank attack (see Theorem 2) by

$$\text{Complexity}_{\text{MinRank}} \geq \ell^{(k+v+1)\cdot\omega}. \tag{3}$$

Theorem 3 allows us to use the method of [11] to derive directly

Theorem 4. *The degree of regularity of a direct attack against an HMFEv system is, under the assumption of $v \leq \ell$, upper bounded by*

$$d_{\text{reg}} \leq \begin{cases} \frac{(q-1)(k+v-1)}{2} + 2 & \text{if } q \text{ even and } (k+v) \text{ odd} \\ \frac{(q-1)\cdot(k+v)}{2} + 2 & \text{otherwise} \end{cases}. \tag{4}$$

Equation (4) gives an upper bound for the degree of regularity of a direct attack against our scheme. However, in order to estimate the security of the HMFEv-scheme in practice, we need to analyze if the bound given by (4) is reasonably tight. Furthermore we want to study, if, as equation (4) indicates, only the sum and not the concrete choice of k and v determines the degree of regularity of a direct attack against an HMFEv system. To answer these two questions, we performed a large number of experiments.
Our experiments (see in Section A of the appendix of this paper) show that the upper bound on the degree of regularity given by equation (4) is relatively tight. We could find several MHFEv instances which actually meet the upper bound and found that in most other cases the upper bound is missed only by one. Regarding the second question, we found that the concrete choice of k and v has no influence on the behavior of the scheme against direct attacks as long as v is not too small.
The experiments in the appendix deal with HMFEv schemes over very small fields such as GF(2) and GF(3). However, one major benefit of the HMFEv scheme is that, in contrast to HFEv-, it can be efficiently used over fields of arbitrary size. As our experiments (see Section 6) show, these systems behave much more like random systems and we can reach high degrees of regularity, by which we can show the security of our scheme against direct attacks.

5.2 Quantum Attacks

In [22] Schwabe and Westerbaan showed that a binary system of n multivariate quadratic equations can be solved by a quantum computer in time $2^{n/2} \cdot 2 \cdot n^2$. Since our systems over GF(256) can easily be translated into systems over GF(2), this attack affects also our scheme (at least in theory). However, since this transition increases the number of variables in the system by a factor of 8, we do not have to consider this type of attack here.

5.3 Other Attacks and A Remark on the Minus Method

Additional to direct, quantum and rank attacks, we analyzed the security of our scheme against other known attacks against multivariate schemes, including differential attacks and Hashimotos attack against the original MultiHFE encryption scheme [15] and found that these attacks do not apply against our scheme. However, due to lack of space, we can not present the details of our analysis here and refer to the extended version of this paper.

Remark. A natural question here is, why we do not use the Minus method as in the case of HFEv-. There are two main reasons.
1. In opposite to the Vinegar variation, the Minus modification does not help to defend our scheme against Hashimotos attack against the original MultiHFE encryption scheme [15].
2. If we use the above method to analyze the MinRank attack, we can prove that the MinRank should be $k + v + ak$, where a is the number of Minus equations. But there appears a new and very interesting problem regarding the degree of regularity. If we follow our usual method, we derive

$$d_{\text{reg}} \leq \begin{array}{ll} \frac{(q-1)(k+v+ak-1)}{2} + 2 & \text{if } q \text{ even and } (k+v+ak) \text{ odd} \\ \frac{(q-1)\cdot(k+v+ak)}{2} + 2 & \text{otherwise} \end{array} . \quad (5)$$

However our experiments show that this bound is not tight. This can be explained as follows. In the case of HFEv-, the estimate comes from using a single polynomial on a large field, and a single polynomial already determines the whole system; in the case of MHFEv-, the system is determined by k polynomials, not by one; since our analysis considers only one of these polynomials, it does not use all the information available and therefore overestimates the degree of regularity. This means we have a gap in the knowledge on estimating the degree of regularity in MHFEv-, which is the reason we propose the MHFEv system (i.e. only with Vinegar). This problem is very interesting and important, and we are going to deal with it in a subsequent paper.

6 Parameter Choice

In this section we consider the question how to find good parameter sets for our scheme. In particular, we aim at finding parameters for HMFEv over the fields GF(31) and GF(256).[4]

[4] The reason why we do not propose parameters for our scheme over GF(16) is the following: To defend the scheme against the quantum attack (see Section 5.2), we need a large number of equations over GF(16). This actually makes the schemes less efficient than HMFEv over GF(31) or GF(256).

6.1 How to choose the parameter k?

The first question we have to answer in order to find suitable parameters for our scheme is how to choose the parameter k and therefore the number of components of the central map. Reducing the value of k will speed up the signature generation process of our scheme since it decreases the size of the multivariate quadratic system we have to solve. However, if k is too small, this might bring the security of our scheme into jeopardy.
For fields of odd characteristic (e.g. $\mathbb{F}$=GF(31)) we choose the parameter k to be 2. However, in order to increase the security of our scheme against Rank attacks, we choose in this case the components of the central map $\mathcal{F}$ in a special way. Let F_1 and F_2 be the 2×2 matrices representing the homogeneous quadratic parts of the maps $\mathcal{F}^{(1)}$ and $\mathcal{F}^{(2)}$. A linear combination of F_1 and F_2 of rank 1 exists if and only if the quadratic polynomial $p(X) = \det(F_1 + X \cdot F_2) \in \mathbb{E}[X]$ has a solution. We therefore choose the coefficients of $\mathcal{F}^{(1)}$ and $\mathcal{F}^{(2)}$ in such a way that the polynomial $p(X)$ is irreducible.
For fields of even characteristic, the symmetric matrices representing the quadratic maps $\mathcal{F}^{(i)}$ contain zero elements on the main diagonal. Therefore, for $k = 2$, the rank of these matrices would be 1 and the upper linear combination of the maps $\mathcal{F}^{(1)}$ and $\mathcal{F}^{(2)}$ would actually lead to a matrix of rank 0 (i.e. no quadratic terms at all.) To prevent this, we choose for fields of even characteristic the parameter k to be 3.

6.2 Experiments with direct attacks against HMFEv schemes over GF(31) and GF(256)

In Section 5.1 we already presented some results of experiments with the direct attack against HMFEv instances. However, in Section 5.1, we looked at HMFEv schemes over very small fields, for which the bound given by equation (4) is more or less tight. In this section we consider the question if concrete instances of HMFEv over the larger fields GF(31) and GF(256) are hard to solve.
To do this, we created for different parameter sets HMFEv systems over GF(31) and GF(256) and solved these systems, after fixing v variables to obtain a determined system, with the F_4 algorithm integrated in MAGMA. The experiments were performed on a single core of a server with 16 AMD Opteron processors (2.4 GHz) and 128 GB of RAM. For each parameter set we performed 10 experiments. Table 1 shows the results.

As the table shows, we can, for HMFEv instances over both GF(31) and GF(256), reach high degrees of regularity. In particular we see that, for the parameter sets proposed in the next section, the degree of regularity of a direct attack is at least 17. By substituting this value into the formula

$$\text{Complexity}_{\text{direct attack}} \approx 3 \cdot \binom{n + d_{\text{reg}}}{d_{\text{reg}}}^2 \cdot \binom{n}{2} \tag{6}$$

GF(31)	parameters (k, ℓ, v)	(2,6,4)	(2,7,4)	(2,8,4)	random
	m,n	12,12	14,14	16,16	16,16
	d_{reg}	14	16	18	18
	time	1,906	164,110	-	-
	memory (MB)	949	17,165	ooM	ooM
GF(256)	parameters (k, ℓ, v)	(3,3,6)	(3,4,6)	(3,5,6)	random
	m,n	9,9	12,12	15,15	15,15
	d_{reg}	11	14	17	17
	time	4.0	1,875	-	-
	memory (MB)	24.5	949	ooM	ooM

Table 1. Experiments with the direct attack against HMFEv schemes over GF(31) and GF(256)

we find that the complexity of a direct attack against the HMFEv instances shown in Table 2 is beyond the claimed levels of security.
Also note that, for the underlying fields of GF(31) and GF(256), the public systems of HMFEv behave very similar to random systems. This also holds when guessing some variables before applying the F_4 algorithm (hybrid approach).

6.3 Parameters

Table 2 shows, for different levels of security (128, 192, and 256 bit) our parameter recommendations for the HMFEv signature scheme over GF(31) and GF(256). In the case of GF(31), we store one element of GF(31) in 5 bits, while 24 bits can be efficiently stored in 5 GF(31) elements.

security level (bit)	parameters $(\mathbb{F}, k, \ell, v)$	public key size (kB)	private key size (kB)	hash size (bit)	signature size (bit)
128	(GF(31),2,28,12)	81.8	8.9	277	337
	(GF(256),3,15,16)	85.8	15.2	360	488
192	(GF(31),2,40,17)	234.7	20.0	396	481
	(GF(256),3,23,21)	282.1	35.0	552	720
256	(GF(31),2,55,21)	583.9	38.0	544	649
	(GF(256),3,31,26)	659.4	65.3	744	952

Table 2. Parameter Recommendations for the HMFEv Signature Scheme

The parameter sets given in Table 2 are chosen in such a way that the complexities of direct attacks (including hybrid approach; see Section 6.2) and Rank attacks (see equation (3)) against the given HMFEv instances are beyond the claimed levels of security. To be on the conservative side we chose, in the formula

(3), the linear algebra constant ω to be 2. Furthermore, in the case of MHFEv over GF(31), we had to take care of the fact that the public systems contain enough equations to prevent collision attacks against the hash function.

7 Comparison

We briefly describe our implementation in the Appendices B and C of this paper.

The basic idea of the HMFEv signature scheme is very similar to that of Gui [20]: by applying the Vinegar modification it is possible to increase both the security and the efficiency of the scheme significantly. However, there are at least three major advantages of our scheme compared to Gui.
First, for efficiency reasons, the Gui signature scheme is restricted to the field GF(2). This leads to a large number of variables in the scheme and therefore to large key sizes. . On the other hand, the HMFEv signature scheme can be defined over any field. This enables us to decrease the number of variables in the system and therefore reduces the public key size of the scheme significantly (see Table 3).
Secondly, for the parameter sets recommended in [20], the output size of the HFEv- public key is only 90 bit. Therefore, in order to defend the HFEv- signature scheme against collision attacks, the authors of Gui had to create a specially designed signature generation process for their scheme which inverts the HFEv- core map several times. Since the design of Gui requires the single HFEv- systems to have exactly one solution, generating one single Gui signature implies about 11 inversions of the HFEv- map, which leads to a relatively low performance of Gui. In the case of the HMFEv scheme, we do not need this multiple inversion of the core map, which makes the signature generation process of our scheme much faster and easier to implement. Furthermore, since the number of variables in the public systems of Gui is much larger than for our scheme, the evaluation of the HMFEv public systems and therefore the verification process of our scheme is much cheaper.Table 3 compares, for a security level of 80 bit, the HMFEv and Gui signature schemes with respect to key and signature sizes as well as the running time of the signature generation and verification process. Note that, for higher levels of security, the benefits of our scheme would be even more significant. The schemes listed in the table run on an Intel Xeon E3-1245 processor with 3.4 GHz. The parameters and running times in the first three rows of the table are taken from the paper [20]. The third major advantage of the HMFEv scheme is that, in contrast to other HFEv- based schemes like Gui, the scheme can be scaled much easier to higher levels of security. For example, in order to obtain a quantum security level of 256 bit, we need an internal state of at least 480 bit (c.f. Section 5.2), which means that we need at least 480 variables over GF(2). This would lead to key sizes which are completely impractical. In the case of HMFEv-, we can increase the size of the internal state simply by choosing a larger base field, which has far less influence on key sizes.

	public key size (kB)	private key size (kB)	signature size (bit)	sign. gen. time (ms)	verification time (ms)
Gui (GF(2),96,5,6,6)	61.6	3.1	126	0.07	0.02
Gui(GF(2),95,9,5,5)	59.2	3.0	120	0.18	0.02
Gui(GF(2),94,17,4,4)	56.8	2.9	124	0.73	0.02
HMFEv (GF(31),2,18,8)	22.5	3.5	218	0.20	0.012
HMFEv (GF(256),3,9,12)	21.6	6.0	312	0.24	0.02
HMFEv (GF(31),2,28,12)	81.8	8.9	337	0.40	0.04
HMFEv (GF(256),3,15,16)	85.8	15.2	488	0.36	0.05

Table 3. Comparison of HMFEv and Gui (80 bit security)

8 Conclusion

In this paper we proposed a new multivariate signature scheme called HMFEv which is obtained by applying the Vinegar modification to the MultiHFE scheme of Chen et al. [7]. By using this variation, we are able to reduce the number of components in the central map of the scheme and therefore to increase the efficiency significantly. We studied the security of our scheme against direct and rank attacks both theoretically and experimentally and showed that our scheme can not be attacked using differential methods or Hashimotos attack against the original MultiHFE scheme. We showed that our scheme is much more efficient than the Gui signature scheme with regard to key sizes, performance and scalability. Future work includes in particular further optimization of the implementation to enable a better comparison of our results with those from [20] as well as a careful study on the effects of applying the Minus modification on HMFEv.

Disclaimer

Certain algorithms and commercial products are identified in this paper to foster understanding. Such identification does not imply recommendation or endorsement by NIST, nor does it imply that the algorithms or products identified are necessarily the best available for the purpose.

References

1. D.J. Bernstein, J. Buchmann, E. Dahmen (eds.): Post Quantum Cryptography. Springer, 2009.
2. L. Bettale, L.C Faugère, L. Perret: Hybrid approach for solving multivariate systems over finite fields. Journal of Mathematical Cryptology 3, pp. 177-197 (2009).
3. L. Bettale, J.C. Faugère, L. Perret: Cryptanalysis of HFE, multi-HFE and variants for odd and even characteristic. Des. Codes Cryptography 69 (1), pp. 1–52 (2013).
4. A. Bogdanov, T. Eisenbarth, A. Rupp, C. Wolf: Time-area optimized public-key engines: MQ-cryptosystems as replacement for elliptic curves? CHES 2008, LNCS vol. 5154, pp. 45-61. Springer, 2008.
5. 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. Springer, 2016.
6. A.I.T. Chen, M.-S. Chen, T.-R. Chen, C.-M. Cheng, J. Ding, E. L.-H. Kuo, F. Y.-S. Lee, B.-Y. Yang: SSE implementation of multivariate PKCs on modern x86 cpus. CHES 2009, LNCS vol. 5747, pp. 33 - 48. Springer, 2009.
7. C.H.O. Chen, M.S. Chen, J. Ding, F. Werner, B.Y. Yang: Odd-char multivariate Hidden Field Equations. IACR eprint, http://eprint.iacr.org/2008/543 (2008).
8. T. Daniels, D, Smith-Tone: Differential Properties of the HFE Cryptosystem. PQCrypto 2016, LNCS vol. 8772, pp. 59 - 75. Springer, 2014.
9. J. Ding, J. E. Gower, D. S. Schmidt: Multivariate Public Key Cryptosystems. Springer, 2006.
10. J. Ding, B.Y. Yang: Degree of regularity for HFEv and HFEv-. PQCrypto 2013, LNCS vol. 7932, pp. 52 - 66. Springer, 2013.
11. J. Ding, T. Hodges: Inverting HFE is quasipolynomial for all fields. CRYPTO 2011, LNCS vol. 6841, pp. 724 - 742. Springer, 2011.
12. J. Ding, D. S. Schmidt: Rainbow, a new multivariate polynomial signature scheme. ACNS 2005, LNCS vol. 3531, pp. 164-175. Springer, 2005.
13. J.C. Faugère: A new efficient algorithm for computing Gröbner bases (F4). Journal of Pure and Applied Algebra 139, pp. 61-88 (1999).
14. M. R. Garey and D. S. Johnson: Computers and Intractability: A Guide to the Theory of NP-Completeness. W.H. Freeman and Company 1979.
15. Y. Hashimoto: Cryptanalysis of Multi HFE. IACR eprint, http://eprint.iacr.org/2015/1160.pdf (2015).
16. A. Kipnis, L. Patarin, L. Goubin: Unbalanced Oil and Vinegar Schemes. EUROCRYPT 1999, LNCS vol. 1592, pp. 206–222. Springer, 1999.
17. A. Kipnis, A. Shamir: Cryptanalysis of the HFE Public Key Cryptosystem. CRYPTO 99, LNCS vol. 1666, pp. 19 - 30. Springer 1999.
18. J. Patarin, N. Courtois, L. Goubin: QUARTZ, 128-Bit Long Digital Signatures. CTRSA 2001, LNCS vol. 2020, pp. 282-297. Springer, 2001.
19. J. Patarin: Hidden Field Equations (HFE) and Isomorphisms of Polynomials (IP): two new families of asymmetric algorithms. EUROCRYPT 1996, LNCS vol. 1070, pp. 33 - 48. Springer 1996.
20. A. Petzoldt, M.S. Chen, B.Y. Yang, C. Tao, J. Ding: Design Principles for HFEv-based Signature Schemes. ASIACRYPT 2015 - Part 1, LNCS vol. 9452, pp. 311-334. Springer, 2015.
21. R. L. Rivest, A. Shamir, L. Adleman: A Method for Obtaining Digital Signatures and Public-Key Cryptosystems. Commun. ACM 21 (2), pp. 120-126 (1978).

22. P. Schwabe, B. Westerbaan: Solving binary MQ with Grovers algorithm. Available at https://cryptojedi.org/papers/mqgrover-20160901.pdf.
23. P. Shor: Polynomial-Time Algorithms for Prime Factorization and Discrete Logarithms on a Quantum Computer, SIAM J. Comput. 26 (5), pp. 1484 - 1509 (1997).
24. B.-Y. Yang, J.-M. Chen: Theoretical Analysis of XL over Small Fields. ACISP 2004, LNCS vol. 3108, pp.277-288. Springer 2004.

A Results of our Computer Experiments

In this section we present the results of our computer experiments with the direct attack against HMFEv schemes over small fields. In particular, we wanted to answer the questions

1. Is the concrete choice of k and v (or only the sum) important for the degree of regularity of a direct attack against the scheme? and
2. Is the upper bound on d_{reg} given by equation (4) reasonable tight?

In order to answer the first question, we performed experiments of the following type: For fixed values of q and $s = k + v$, we varied the values of k and v. We then created the public systems of the corresponding HMFEv instances (for different values of ℓ) and solved these systems using the F_4 algorithm integrated in MAGMA. The experiments were (like all the experiments presented in this paper) performed on a server with 16 AMD Opteron cores (2.4 GHz) and 128 GB of RAM. However, as MAGMA is not parallelizable, our programs use only one core.

In our experiments, we fixed the field $\mathbb{F}$ to be GF(2) and the sum $s = k+v$ to be 9. We varied v in the interval $I = \{0, \ldots, 8\}$ and created HMFEv(GF(2),$s - v$, ℓ, v) instances (for increasing values of ℓ). After that, we fixed v of the variables to get a determined system and solved the resulting public systems by the F_4 algorithm integrated in MAGMA. Table 4 shows, for $v \in I$, the highest degree of regularity we observed in these experiments. For each parameter set, we performed 10 experiments.

As the experiments show, the concrete ratio between k and v has, as long as we

v	0	1	2	3	4	5	6	7	8
k	9	8	7	6	5	4	3	2	1
d_{reg}	3	4	4	5	5	5	5	5	4

Table 4. Degree of regularity of HMFEv systems over GF(2) with $k + v = 9$

choose v and k not too small, no influence on the degree of regularity of solving the public systems of HMFEv. For HMFEv schemes over larger fields the importance of the concrete choice of k and v decreases further, since those systems

behave much more like random systems (see Section 6). We therefore choose, in order to increase the efficiency of our scheme, the parameter $k \in \{2, 3\}$ and increase v to reach the required level of security.

Is the upper bound on $d_{\mathbf{reg}}$ given by equation (4) reasonable tight?

In this section we want to analyze the question whether the upper bound on the degree of regularity given by equation (4) is reasonable tight. To do this, we created for fixed values of q, k and v and varying values of ℓ public systems of HMFEv and solved them with the F_4 algorithm integrated in MAGMA. We increased the value of ℓ and therefore the numbers of equations and variables in the system until we reached the upper bound of (4) or ran out of memory.
It is obvious that we can only hope to find such systems for small field sizes. We therefore restricted to values of $q \in \{2, 3\}$.

By doing so, we identified the following "tight" instances of HMFEv

scheme	upper bound on d_{reg} (equation (4))	experimental result
HMFEv-(GF(2),1,ℓ,2)	3	3 for $\ell \geq 9$ ($n \geq 9$)
HMFEv-(GF(2),2,ℓ,3)	4	4 for $\ell \geq 9$ ($n \geq 18$)
HMFEv-(GF(2),3,ℓ,4)	5	5 for $\ell \geq 10$ ($n \geq 30$)
HMFEv-(GF(3),1,ℓ,2)	5	5 for $\ell \geq 18$ ($n \geq 18$)

For most other HMFEv instances with $q \in \{2, 3\}$ and $k + v \leq 9$ we missed the upper bound given by equation (4) only by 1.
We believe that, also for these systems, we could have reached the upper bound given by equation (4) by increasing the parameter ℓ further. However, we did not have the necessary memory resources to solve HMFEv systems with more than 35 equations.

B Efficient Implementation of the Public and Private Maps of HMFEv

The most costly step during the signature generation process of our scheme is the inversion of the central equation $\mathcal{F}_V(\mathbf{Y}) = \mathbf{X}$, which is given as a system of k multivariate quadratic equations in k variables over the extension field $\mathbb{E}$. Since the coefficients of this system are chosen randomly, we need a system solver like a XL or a Gröbner basis algorithm for this step.
Obviously, the complexity of solving the system $\mathcal{F}_V(\mathbf{Y}) = \mathbf{X}$ and therefore the complexity of the signature generation process depends mainly on the choice of the parameter k. A small value of k will reduce the number of $\mathbb{E}$-multiplications in this process. However, it also leads to large extension fields and therefore increases the cost of a single $\mathbb{E}$-multiplication. Furthermore, choosing k too small might weaken the security of our scheme (see Section 7.1).

To find the optimal parameter k for our scheme, we therefore have to analyze the

process of inverting the central map $\mathcal{F}_V$ in more detail. Let the multivariate system $\mathcal{F}_V$ be given by the k multivariate quadratic maps $f_V^{(1)}, \ldots, f_V^{(k)} : \mathbb{E}^k \to \mathbb{E}$. As we find, the process of solving the multivariate system $\mathcal{F}_V(\mathbf{Y}) = \mathbf{X}$ consists mainly of two parts:

1. (**Gröbner basis step**) Find a univariate polynomial $p : \mathbb{E} \to \mathbb{E}$ in the ideal $\langle f_V^{(1)}, \ldots, f_V^{(k)} \rangle$.
2. (**Solving Step**) Solve the polynomial p by Berlekamp's algorithm.

In the following we analyze, for different values of k, these two steps in detail. For this, we fix the number $n = k \cdot \ell$ to $n = 48$ and choose $k \in \{2, 3, 4\}$. Inverting the system $\mathcal{F}_V$ therefore relates to

- solving a system of 2 quadratic equations in 2 variables over $\mathbb{F}^{24}$ or
- solving a system of 3 quadratic equations in 3 variables over $\mathbb{F}^{16}$ or
- solving a system of 4 quadratic equations in 4 variables over $\mathbb{F}^{12}$.

For $k = 2, 3$, we use for the first part a specially designed Gröbner basis method tailored for the occasion. In the case of 2 quadratic equations in 2 variables, we run in the Gröbner basis step successively 2 Gaussian eliminations on matrices of size 5×9 and 7×10. By doing so, we obtain a single variable equation p of degree 4. To perform this step, we need about $5 \cdot (11 + 12) + 7 \cdot 8 \cdot 4 = 339$ multiplications over the field $\mathbb{F}^{24}$.
In the Solving step, we have to solve the univariate equation p of degree 4 over the field $\mathbb{F}^{24}$. This takes about $6 \cdot 4^2 \cdot 24 = 2,304$ multiplications over the field of size $\mathbb{F}^{24}$. One can see that the overall complexity is dominated by the Solving step.

In the case of 3 quadratic equations in 3 variables, we run in the Gröbner basis step successively 3 Gaussian eliminations on matrices of size 11×19, 8×16 and 5×13 with many zero elements to derive a single variable equation of degree 8. For this we need about 1,700 $\mathbb{F}_{16}$ multiplications.
Then we solve this single variable equation of degree 8 over the larger field. This requires about $6 \cdot 8^2 \cdot 16 = 6,144$ big-field multiplications. One can see that the Solving step dominates the complexity.

In the case of 4 quadratic equations in 4 variables, the situation is too complicated to do it by hand and we use the F_4 algorithm directly. In this case, we run successively Gaussian eliminations on matrices of size $19 \times 34, 41 \times 50, 42 \times 50$ and 35×48, which requires about $2 \cdot 50^3 = 250,000$ $\mathbb{F}^{12}$ multiplications. By doing so, we obtain a single variable equation p of degree 16.
In the Solving Step, we have to solve this univariate equation p over the larger field, which requires about $6 \cdot 16^2 \cdot 12 = 18,432$ multiplications. One can see that here the solving of the single variable equation does not dominate the complexity anymore.

C Arithmetic in the Public and Private Maps of HMFEv

Evaluating the public map requires first to generate all monomials, and then the computation of the inner product polynomials from known monomials. The first step requires $n(n+1)/2$ field multiplications. The second part is much more important and requires $mn(n+3)/2$ multiplications in the field and nearly as many additions (or XORs) to accumulate the results.
Arithmetic in GF(256) is done via the table-lookup instruction VPSHUFB. This instruction allows 32 simultaneous lookups from a table of 16, which allows for easy scalar-vector multiplications of GF(16) using log-exp tables. Every 32 GF(16) multiplications then take two VPSHUFB instructions and an add in addition to the required VPXOR, because we store the public key in log form. Finally we put together multiplications of GF(256) for the public key using four multiplications in GF(16) (schoolbook method).
The main computation in big binary fields uses PCLMULQDQ and schoolbook because on recent processors this instruction is really fast. We also use lazy reductions,which means that we often do not reduce to the lowest degree. A time-constant complete reduction is performed after the entire operation.

Arithmetic in GF(31) uses AVX2 instructions (and following that SSSE3 instructions). For best use of our resources, we use a YMM register to represent a vector of 16 or 32 coefficients in the public key to be multiplied by two monomials. Values for two monomials each time are also expanded into an YMM register. The actual arithmetic uses the VPMADDUSBW instruction to multiply two pairs of byte values (one signed one unsigned) into signed 16-bit values, and add them together all in one cycle. This requires us to ensure that input monomials are in $0, \ldots, 31$ and the coefficients in $-15, \ldots, 15$. We add together 32 results of VPMADDUSBW each time, which keeps the result between ± 32767. We can then reduce the results again to between $0, \ldots, 31$.
Arithmetic in tower fields over GF(31) are in straight schoolbook form and do not use the VPMADDUSBW instruction because the sizes are not convenient for it.