Security Analysis and Key Modification for $ZHFE$

Ray Perlner[1] and Daniel Smith-Tone[1,2]

[1]National Institute of Standards and Technology,
Gaithersburg, Maryland, USA
[2]Department of Mathematics, University of Louisville,
Louisville, Kentucky, USA

ray.perlner@nist.gov, daniel.smith@nist.gov

Abstract. $ZHFE$, designed by Porras et al., is one of the few promising candidates for a multivariate public-key encryption algorithm. In this article we extend and expound upon the existing security analysis on this scheme. We prove security against differential adversaries, complementing a more accurate and robust discussion of resistance to rank and algebraic attacks. We further suggest a modification, $ZHFE^-$, a multivariate encryption scheme which retains the security and performance properties of $ZHFE$ while optimizing key size in this theoretical framework.

Key words: Multivariate Cryptography, HFE, $ZHFE$, Discrete Differential, MinRank, Q-rank

1 Introduction

Since the late 1990s, a large international community has emerged to face the challenge of developing cryptographic constructions which resist attacks from quantum computers. The birth of this new discipline is due primarily to the discovery by Peter Shor in the mid 90s, see [1], of algorithms for factoring and computing discrete logarithms in polynomial time on a quantum computing device. The term post-quantum cryptography was coined to refer to this developing field and to emphasize the fact that information security in a quantum computing world is a fundamentally new science.

Today, we face mounting evidence that quantum computing is not a physical impossibility but merely a colossal engineering challenge. With the specter of the death of classical asymmetric cryptography looming on the horizon, it is more important than ever that we develop systems for authentication, confidentiality and key exchange which are secure in the quantum paradigm. We thus are forced to turn to problems of greater difficulty than the classical number theoretic constructs.

Systems of polynomial equations have been studied for thousands of years and have fueled the development of several branches of mathematics from classical

to modern times. Multivariate Public Key Cryptography(MPKC) has emerged from the serious investigation of computational algebraic geometry that reached maturity in the latter half of the last century. Today, we see MPKC as one of a few serious candidates for security in the post-quantum world.

A fundamental problem on which the security of any multivariate cryptosystem rests is the problem of solving systems of quadratic equations over finite fields. This problem is known to be NP-hard, and copious empirical evidence indicates that the problem is hard even in the average case. There is no known significant reduction of the complexity of this problem in the quantum model of computing, and, indeed, if this problem is discovered to be solvable in the quantum model, we can solve all NP problems and the task of securing information might be hopeless in principle. We thus reasonably suspect that MPKC will survive the transition into the quantum world.

Though multivariate cryptosystems almost always suffer from rather large key sizes, the key sizes are rarely so large that they are impractical and these systems can often be quite attractive in certain other aspects of performance. Some systems are very fast, having speeds orders of magnitude faster than RSA, [2–4]. Some schemes combine speed with power efficiency and small signature sizes, [5, 6]. Perhaps most importantly, it is generally simple to parameterize multivariate systems in such a way that vastly different properties are derived foiling various attack methodologies.

One great difficulty historically for MPKC is encryption. Though there are several viable options for digital signatures, see [5–8], there is a general absence of long-lived encryption systems. In the last couple of years, a couple of new encryption techniques have been proposed, see [9–11]. These systems are based on the simple idea, proposed by Ding, that the structure of a system of equations can retain injectivity without an extremely restrictive structure if the codomain is of much larger dimension than the domain.

In [12], however, a new and unexpected attack was presented on the *ABC* simple matrix encryption scheme of [9]. This attack is notable in that the complexity is far less asymptotically than predicted by the analysis in [9], though it does not break the scheme outright. This begs the question of the tightness of the security analyses in [10, 11] and the extent to which we can trust in the security of such young schemes in a field which has no significant success history in encryption.

Furthermore, one might ask whether there is some middleground on the ratio of the dimension of the codomain to that of the domain for these multivariate encryption schemes. Even if one concurs that relaxing the relationship between the dimensions of the domain and codomain enhance the security of injective maps, it remains unclear that the disparity should be so large as in the proposed schemes in which there are at least twice as many equations as variables.

In this article we extend and expound upon the security analysis in [11], incorporating some of the theoretical models of assurance presented in [13–15]. We prove security against differential adversaries complementing the discussion of resistance to algebraic attacks provided in [11]. We further elucidate the rank

structure of $ZHFE$ and specifically note some necessary, but trivial, key restrictions for security which were apparently overlooked in [11]. We further suggest a modification, $ZHFE^-$, a multivariate encryption scheme which retains the security and performance properties of $ZHFE$ while optimizing key size in this theoretical framework.

The paper is organized as follows. The next section introduces the notion of big field schemes and presents the prototypical such cryptosystem, HFE. In the following section, we define the Q-rank of a multivariate system of equations and discuss the central nature of this concept in the field. The subsequent section presents the $ZHFE$ encryption scheme and calculates some of its inherent parameters. Next we present a thorough security analysis of $ZHFE$, complementing and expanding the analysis provided in [11] and offering security assurance against a differential adversary as well as discussing parameters securing $ZHFE$ against rank and algebraic attacks. Subsequently, we present and analyze $ZHFE^-$, a new multivariate encryption scheme based on $ZHFE$ and the minus modifier. Finally, we note parameter choices for $ZHFE^-$ and discuss the role that the new methodology for multivariate encryption fills in the literature.

2 *HFE*

Several multivariate cryptosystems belong to a family collectively known as "big field" schemes. Such schemes are constructed using two ideas. The first is an equivalence between functions on a degree n extension k of a finite field $\mathbb{F}_q$ and functions on an n-dimensional $\mathbb{F}_q$-vector space. The second is an isomorphism of polynomials which allows one to hide structure in a function.

To see the equivalence, notice that a vector space isomorphism between k and an n-dimensional vector space over $\mathbb{F}_q$ extends to a vector space isomorphism between the space of univariate functions from k to itself and the space of multivariate n-dimensional vector-valued polynomial functions from $\mathbb{F}_q^n$ to itself. (Specifically, given an isomorphism $\phi : \mathbb{F}_q^n \to k$ and a function $f : k \to k$, the function $\phi^{-1} \circ f \circ \phi$ is such a function from $\mathbb{F}_q^n$ to itself; furthermore, this identification is a 1-1 correspondence.)

The second idea, the isomorphism of polynomials, is defined in the following manner.

Definition 1 *Two vector valued multivariate polynomials f and g are said to be* isomorpic *if there exist two affine maps T, U such that $g = T \circ f \circ U$.*

Together these ideas allow us to build an isomorphic copy of a structured univariate map with domain k while hiding the structure. The construction is sometimes called the butterfly construction because of the shape of its defining commutative diagram. Specifically, $P = T \circ \phi^{-1} \circ f \circ \phi \circ U$ produces a perturbed vector-valued version of the structured univariate polynomial f.

The Hidden Field Equations (HFE) scheme was first presented by Patarin in [16] as a method of avoiding his linearization equations attack which broke

the C^* scheme of Matsumoto and Imai, see [17] and [18]. The basic idea of the system is to use the butterfly construction to hide the structure of a low degree polynomial that can be inverted efficiently over k via the Berlekamp algorithm [19], for example.

More specifically, we select an effectively invertible "quadratic" map $f : k \to k$, quadratic in the sense that every monomial of f is a product of a constant and two Frobenius multiples of x. Explicitly any such "core" map f has the form:

$$f(x) = \sum_{\substack{i \leq j \\ q^i + q^j \leq D}} \alpha_{i,j} x^{q^i + q^j} + \sum_{\substack{i \\ q^i \leq D}} \beta_i x^{q^i} + \gamma.$$

The bound D on the degree of the polynomial is required to be quite low for efficient inversion.

The HFE scheme was designed to be used as an encryption or a signature scheme. To generate a signature (or to decrypt), one computes, successively, $v = T^{-1}y$, $u = f^{-1}(v)$ and $x = U^{-1}u$. The vector x is the signature (or the plaintext). For verification (or encryption), one simply evaluates the public polynomials, P, at x. If $P(x)$ which is equal to $T \circ f \circ U(x)$ is equal to y, the signature is authenticated (or the ciphertext is y).

3 Q-Rank

The defining charactersitic of HFE, the degree bound, which is necessary for the effective inversion of the central map, ensures that the scheme has low rank as a quadratic form over k, as described below. This property assures that the central map of HFE is vulnerable to Kipnis-Shamir modeling, see [20, 21].

Recall that any quadratic map $f : k \to k$ can be written

$$f(x) = \sum_{0 \leq i,j < n} \alpha_{ij} x^{q^i + q^j}.$$

We can equivalently express f as a vector function over the 1-dimensional k-algebra $\psi : k \to k^n$ where

$$\alpha \stackrel{\psi}{\mapsto} \left[\alpha \ \alpha^q \ \ldots \alpha^{q^{n-1}}\right]^T,$$

in the form $f(X) = X^T[\alpha_{ij}]X$ where $X = [x \ x^q \ \ldots \ x^{q^{n-1}}]^T$.

Any quadratic form over k can be expressed as a symmetric matrix, and over characteristic $p \neq 2$ a change of basis can be performed which transforms this matrix into an equivalent diagonal form. The rank of this matrix is the rank of the quadratic form. We call this rank the Q-rank of f, that is the rank of f as a quadratic function.

We note here that Q-rank is invariant under polynomial isomorphism, thus the Q-rank of a central map of a cryptosystem is the same as the Q-rank of the public key, unless, of course, the minus or projection modifiers are utilized. We

also note that the Q-rank is explicitly exploited in the attacks of [20, 21] and plays a central role in the derivation of degree of regularity bounds for several prominent cryptosystems, see [22–24]. Further, there seems to be a complicated relationship between the Q-rank of a field map and the presence of differential symmetric or invariant relations, see, for example [15]. Consequently, Q-rank seems to be emerging as a central concept in multivariate cryptography and in computational algebra.

4 *ZHFE*

$ZHFE$ was introduced in [11]. The idea is to construct an encryption scheme with a high Q-rank central map preventing attacks such as [21] exploiting this weakness. The scheme is notable among "big field" schemes which typically require some low Q-rank map for efficient inversion. Low Q-rank is in fact required for inversion in this setting as well, however, the system attempts to hide the low Q-rank structure in the public key.

The construction concatenates two high degree quadratic maps (with special structure) to form the central map. Specifically, the two general form quadratic maps f_0 and f_1 are derived by constructing a low degree (maximum degree D) cubic map

$$\Psi(x) = x\left[L_{00}f_0(x) + L_{01}f_1\right] + x^q\left[L_{10}f_0 + L_{11}f_1\right], \tag{1}$$

where L_{ij} is a linear map and the square brackets indicate multiplication over k.

To solve for f_0 and f_1 it suffices to set coefficients for the linear maps and for Ψ to recover a system of linear equations in the unknown coefficients of f_0 and f_1. In the homogeneous case, there are collectively $n^2 + n$ coefficients of f_0 and f_1 in k. Due to its low degree and the requirement that it satisfy (1), Ψ is constrained to be of the form

$$\Psi(x) = \sum_{i=0}^{1} \sum_{\substack{i\le j\le k \\ q^i+q^j+q^k\le D}} \alpha_{i,j,k}x^{q^i+q^j+q^k} + \sum_{i=0}^{1} \sum_{\substack{i\le j \\ q^i+q^j\le D}} \beta_{i,j}x^{q^i+q^j} + \sum_{i=0}^{1}\gamma_i x^{q^i}. \tag{2}$$

A cubic of the form (2) has n^2 coefficients over k, and thus for *any fixed choice* of Ψ and L_{ij} there are n^2 constraints on a linear system of dimension n^2+n. Thus with probability roughly $1-q^{-n}$, there is an n-dimensional space of coefficients for the maps f_0 and f_1.

Once, constructed, the central map $(y_0, y_1) = (f_0(x), f_1(x))$ can be inverted by using Berlekamp's algorithm to solve the low degree polynomial equation:

$$\Psi(x) - x\left[L_{00}y_0 + L_{01}y_1\right] - x^q\left[L_{10}y_0 + L_{11}y_1\right] = 0.$$

5 Analysis of $ZHFE$

A few avenues of attack have evolved along with the development of multivariate cryptosystems relying on a hidden large algebra structure. These attacks can be characterized as differential, see [25, 12], as minrank, see [20, 21], or as algebraic, see [26]. We analyze the security of $ZHFE$ against each of these attack models.

5.1 Algebraic

Algebraic attacks attempt to decrypt a given ciphertext y by solving the system of equations $P(x) = y$ directly. The term "algebraic" refers to the fact that these are generic algorithms for solving arbitrary systems of polynomial equations.

While these attacks are not structural, in the sense of being defined based on the structure of the system of equations, the algorithms employed can naturally take advantage of certain properties of the systems. In practice, the complexity of algorithms for solving these systems of equations is closely connected to the degree of regularity of the system.

The degree of regularity of a system of equations is the degree at which the first nontrivial degree fall occurs. Specifically, consider a generating set of an ideal $I = \langle g_1, \ldots, g_m \rangle \in \mathbb{F}_q[x_1, \ldots, x_n]$. We may generate elements of I by selecting polynomials $p_i \in \mathbb{F}_q[x_1, \ldots, x_n]$ and computing

$$\sum_{i=1}^{m} p_i g_i.$$

A degree fall occurs when the degree of this sum is less than the maximum degree of $p_i g_i$. Clearly some degree falls are due to trivial syzygies such as $-g_j g_i + g_i g_j = 0$ and $(g_i^{q-1} - 1)g_i = 0$. The smallest degree, $max_i p_i g_i$ such that the above sum has a nontrivial degree fall is the degree of regularity.

A great deal of literature is devoted to finding bounds for the degree of regularity of quadratic systems, see [22–24, 27]. In practice one can find a lower bound for the degree of regularity by studying toy examples of schemes and seeing how the degree of regularity changes as the parameters change.

Such an analysis for $ZHFE$ is quite straight forward. As mentioned in [11] the degree of regularity for toy $ZHFE$ systems matches exactly the degree of regularity for random systems of equations of the same size, at least for relatively small instances. Considering the connection between Q-rank and the degree of regularity as derived in [22–24, 27], we conclude that a thorough Q-rank analysis of $ZHFE$ will verify the security of the scheme against algebraic attacks. We perform this analysis in Section 5.4.

5.2 Differential Symmetric

As shown in [25], symmetric relations involving the discrete differential of a central map can induce a symmetry in the public key of a multivariate cryptosystem.

In certain circumstances, these relations can reveal properties of the extension field structure, and weaken the public key. Indeed one can easily turn the attack on SFLASH of [25], which converts an instance of C^{*-} into a compatible instance of C^*, into a direct key-recovery attack utilizing the derived representation of elements of the extension field.

As shown in [13] the maps inducing a linear differential symmetry for C^* schemes are precisely those corresponding to multiplication by an element of the extension field. Thus one may rightfully expect that nontrivial symmetric relations on the differential of a central map are uncommon. It is shown, however, in [13] and [15] that nontrivial symmetries can and often do exist even for cases as general as HFE.

As a specific example of the phenomenon of differential symmetries for general polynomials, consider the map $f(x) = x^{q^3+q^2} + x^{q^2+1}$ over a degree 6 extension of the characteristic 2 field $\mathbb{F}_q$. One can easily verify that the general linear symmetry structure, defined as

$$Df(La, x) + Df(a, Lx) = \Lambda_L Df(a, x),$$

is satisfied by the selection

$$Lx = \alpha x^{q^4} + \alpha x^q + \beta x \text{ and } \Lambda_L x = 0,$$

where $\alpha^{q^3} = \alpha$ and $\beta^q = \beta$. Thus there is a 4-dimensional $\mathbb{F}_q$-subspace of linear maps L satisfying the above differential symmetric relation for some choice of Λ_L, while the space of all $\mathbb{F}_q$-linear maps from the extension to itself is only of dimension 36. Consequently, a hypothetical cryptosystem based on this map would be vulnerable to an attack removing the minus modifier, similar to [25], among other weaknesses. Quite specifically, the distillation procedure described in [25] is effective in this instance. We note that this scenario is by no means limited to toy examples such as this one or even instances with Q-rank one; thus, the verification of the absence of differential symmetries is an important task for any multivariate cryptosystem, particularly those including the minus modifier.

In analyzing the differential symmetric properties of $ZHFE$, we may directly analyze the public key or we may study the differential of the Ψ map. We consider both interlinked cases explicitly.

The public key P consists of $2n$ polynomials. The defining characteristic of these polynomials is that $P = T(f_0||f_1)U$. Thus P does not behave like a random system. There exists a low degree cubic map Ψ such that

$$\begin{aligned}\Psi(Ux) =&(Ux)(L_{00}(T^{-1})_1 P(x) + L_{01}(T^{-1})_2 P(x)) \\ &+ (Ux)^q (L_{10}(T^{-1})_1 P(x) + L_{11}(T^{-1})_2 P(x)).\end{aligned} \tag{3}$$

We note that $(T^{-1})_i P(x) = f_i(Ux)$. We may now implicitly differentiate this equation obtaining

$$\begin{aligned} D\Psi(Ua, Ux) =&(Ua)(L_{00}f_0(Ux) + L_{01}f_1(Ux)) \\ &+ (Ua)^q(L_{10}f_0(Ux) + L_{11}f_1(Ux)) \\ &+ (Ux)(L_{00}Df_0(Ua, Ux) + L_{01}Df_1(Ua, Ux)) \\ &+ (Ux)^q(L_{10}Df_0(Ua, Ux) + L_{11}Df_1(Ua, Ux)). \end{aligned} \tag{4}$$

The above is a biquadratic relation in a and x, and as such doesn't immediately reveal a computational way to recover information about the hidden structure of P. To convert this relation into a form in which we can apply linear algebra techniques we require a second differential. For more information on a more general theory of discrete differential equations, see [28].

Since the differential is symmetric, we get the same answer whether we differentiate with respect to a or to x.

$$\begin{aligned} D^2\Psi(Ua, Ub, Ux) =&(Ua)(L_{00}Df_0(Ub, Ux) + L_{01}Df_1(Ub, Ux)) \\ &+ (Ua)^q(L_{10}Df_0(Ub, Ux) + L_{11}Df_1(Ub, Ux)) \\ &+ (Ub)(L_{00}Df_0(Ua, Ux) + L_{01}Df_1(Ua, Ux)) \\ &+ (Ub)^q(L_{10}Df_0(Ua, Ux) + L_{11}Df_1(Ua, Ux)) \\ &+ (Ux)(L_{00}Df_0(Ua, Ub) + L_{01}Df_1(Ua, Ub)) \\ &+ (Ux)^q(L_{10}Df_0(Ua, Ub) + L_{11}Df_1(Ua, Ub)). \end{aligned} \tag{5}$$

Now, due to the fact that Ψ is cubic with a small degree bound, $D^2\Psi$ is a cubic form of low rank. In fact, the existence of linear maps U and $L_{ij}(T^{-1})_j$ such that equations (3) and (5) hold while $D^2\Psi$ has low cubic rank is the defining characteristic of $ZHFE$.

In spite of the existence of this structure, it is unclear how to proceed. One might consider a cubic version of the rank attack from [29], however, the selection of the maps $L_{ij}(T^{-1})_j$ corresponds to solving a minrank problem on a 3-tensor, $D^2\Psi$. Though there is a possibility that the instances of the 3-tensor rank problem arising from this differential equation may lie in a class which are easy to solve, the general 3-tensor rank problem is known to be NP-hard and there does not seem to be any evidence that these instances are any more structured than arbitrary instances of the same rank.

5.3 Differential Invariant

As exemplified in [12] and [30], invariant relations on the differential of a public key can be exploited in key recovery. Although we may analyze the differential invariant structure of the public key of $ZHFE$ directly, there is not in general any nontrivial invariant due to the fact that the structure of $ZHFE$ is hidden in the cubic Ψ map. A couple of generalizations of differential invariants of quadratic functions are derived for higher q-degree functions in [28]. The most relaxed generalization for cubics is given in the following definition.

Definition 2 *A* differential invariant *of a cubic function f is a pair of subspaces $V_1, V_2 \subseteq k$ for which there exists a subspace W with $dim(W) \leq mindim(V_i)$ such that for all $A \in spanD^2 f_i$, we have $D^2 f(a, b, x) = 0$ for all $a \in V_1$, $b \in V_2$ and $x \in W^{\perp}$.*

In the quadratic case, a differential invariant could be seen as a subspace of k on which Df simultaneously acts in every coordinate the same way, that is, always sending that subspace to the same space of linear forms of no larger dimension. In the cubic case we can realize a differential invariant as a subspace V_1 of k and a subspace (defined by V_2) of induced bilinear forms from $D^2 f$ each element of which maps V_1 to the same space of linear forms, W, of no larger dimension. The minimum condition on the dimension of W is due to the symmetry of $D^2 f$; we could equivalently consider the subspace V_2 of k and the subspace of bilinear forms from $D^2 f$ induced from V_1.

It is straightforward to show that the Ψ map of $ZHFE$ has no differential invariant structure. Following the technique of [15], without loss of generality, due to the symmetry, we let $\hat{a} \in V_1$, $\hat{b}, \hat{x} \in V_2$, and let S be a surjective linear map from V_2 to W. The existence of a differential invariant implies the equation

$$\begin{aligned} 0 &= D^2\Psi(\hat{a}, \hat{b}, S\hat{x}) \\ &= \sum_{\substack{0 \leq i,j,l < n \\ q^i + q^j + q^l \leq D}} \alpha_{ijl} \hat{a}^{q^i} \hat{b}^{q^j} (S\hat{x})^{q^l}. \end{aligned} \tag{6}$$

Since by symmetry D is much smaller than $dim(V_1)$ or $dim(V_2)$, (6) is already reduced modulo the minimal polynomial $\mathcal{M}_{V_1}(a)$ of V_1 as an element in $k[a]$ and modulo the minimal polynomial $\mathcal{M}_{V_2}(b)$ of V_2 as an element in $k[b]$. Thus the collection $\{\hat{a}, \hat{a}^q, \ldots, \hat{a}^{q^{d_1}}, \hat{b}, \ldots, \hat{b}^{q^{d_2}}\}$ is independent in $k[a, b]/\langle \mathcal{M}_{V_1}(a), \mathcal{M}_{V_2}(b) \rangle$. Therefore, we obtain the equations

$$\sum_{\substack{0 \leq l < n \\ 0 \leq i,j < n \\ q^i + q^j + q^l \leq D}} \alpha_{ijl} (S\hat{x})^{q^l} = 0.$$

We then obtain the analogous result of [15]; statistically, S must be the zero map on V_2, contradicting the nontriviallity of the differential invariant. Furthermore, we also obtain the result that if any power of q is unique there is no nontrivial differential invariant.

5.4 Q-rank

A further attack vector for $ZHFE$ is to perform a minrank attack using the Kipnis-Shamir methodology of [20] and the improved version in [21]. The attack searches for a low rank k-linear combination of the differentials of the public key. The general minrank problem is known to be NP-complete, see [31] but in practice the complexity depends on the lowest rank map in the space.

It was shown in [21] that the smallest such rank is equal to the smallest Q-rank of the image of the public key under any full rank $\mathbb{F}_q$-linear map. Notice that for (1) to hold we must have that the $x^{q^i+q^j}$ term in $L_{00}f_0 + L_{01}f_1$ to have coefficient 0 for $q^i + q^j + 1 > D$ and $i, j \neq 1$. This restriction induces a relation on the quadratic representations of $L_{00}f_0$ and $L_{01}f_1$. Specifically, if

$$L_{00}f_0(x) + L_{01}f_1(x) = \begin{bmatrix} x \\ x^q \\ \vdots \\ x^{q^{n-1}} \end{bmatrix}^T \begin{bmatrix} \alpha_{00} & \frac{\alpha_{01}}{2} & \cdots & \frac{\alpha_{0(n-1)}}{2} \\ \frac{\alpha_{01}}{2} & \alpha_{11} & \cdots & \frac{\alpha_{1(n-1)}}{2} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\alpha_{0(n-1)}}{2} & \frac{\alpha_{1(n-1)}}{2} & \cdots & \alpha_{(n-1)(n-1)} \end{bmatrix} \begin{bmatrix} x \\ x^q \\ \vdots \\ x^{q^{n-1}} \end{bmatrix},$$

then $\alpha_{ij} = 0$ for $q^i + q^j > D$ and $i, j \neq 1$. Thus $L_{00}f_0 + L_{01}f_1$ has the form

$$\left[\begin{array}{ccccc|ccc} \alpha_{00} & \frac{\alpha_{01}}{2} & \frac{\alpha_{02}}{2} & \cdots & \frac{\alpha_{0D}}{2} & 0 & \cdots & 0 \\ \frac{\alpha_{01}}{2} & \alpha_{11} & \frac{\alpha_{12}}{2} & \cdots & \frac{\alpha_{1D}}{2} & \frac{\alpha_{1(D+1)}}{2} & \cdots & \frac{\alpha_{1(n-1)}}{2} \\ \alpha_{02} & \frac{\alpha_{12}}{2} & \alpha_{22} & \cdots & \frac{\alpha_{2D}}{2} & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ \frac{\alpha_{0,D}}{2} & \frac{\alpha_{1D}}{2} & \frac{\alpha_{2D}}{2} & \cdots & \alpha_{DD} & 0 & \cdots & 0 \\ \hline 0 & \frac{\alpha_{1(D+1)}}{2} & 0 & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & \frac{\alpha_{1(n-1)}}{2} & 0 & \cdots & 0 & 0 & \cdots & 0 \end{array}\right],$$

and has rank no more than $\lceil log_q(D) \rceil + 2$. Hence, if L_{ij} are nonsingular, the $Q - rank$ of $f_0||f_1$ is bounded by $\lceil log_q(D) \rceil + 2$.

In spite of the alarming relation derived above, Q-rank does not appear to be a weakness for ZHFE when one selects L_{ij} to have reasonable corank. One can check that for small r, insisting that L_{ij} have corank r increases the possible Q-rank of $f_0||f_1$ by $2r$. Also, having L_{ij} with even moderately large corank doesn't produce a non-negligible probability of decryption ambiguity due to the zero expectation of the dimension of the intersection of the kernels of L_{ij}. Furthermore, recall that we have at least n degrees of freedom over k in selecting f_0 and f_1 for *any* choice of L_{ij}. Thus the Kipnis-Shamir attack, which is exponential in the Q-rank of the scheme, is trivially thwarted with simple parameter restrictions, though we note that the lack of such restriction on the rank of L_{ij} in [11] is apparently an oversight.

5.5 Equivalent Keys

In [32], the question of the number of equivalent keys for multivariate cryptosystems is explored. This question is quite relevant for $ZHFE$, as well, since there can clearly be multiple private keys allowing one to decrypt a public key. The danger in this vein would be if there is insufficient entropy in public keys due to massive redundancy in private keys.

To analyze the number of equivalent keys, we first determine the number of possible pairs f_0, f_1 satisfying (1) for a fixed Ψ and L_{ij}. As mentioned in Section

4, a map of the form Ψ has n^2 coefficients over k, and due to the degree bound only s of these can be nonzero. Thus with L_{ij} fixed, we have n^2+n unknown coefficients for f_0 and f_1 over k, and so we have $n^2+n-(n^2-s)=n+s$ degrees of freedom in choosing the pair f_0, f_1 for a fixed private key.

Next we consider the same relation with f_0, f_1 fixed. For specificity, let $f_i(x)=\sum_{0\le v\le w<n}\alpha_{ivw}x^{q^v+q^w}$. Given the existence of L_{ij} and Ψ, we have the relation

$$\begin{aligned}\Psi(x)=&\sum_{t=0}^{1}\sum_{i=0}^{n-1}\sum_{0\le v\le w<n} l_{0ti}\alpha_{tvw}^{q^i}x^{q^{v+i}+q^{w+i}+1}\\ &+\sum_{t=0}^{1}\sum_{i=0}^{n-1}\sum_{0\le v\le w<n} l_{1ti}\alpha_{tvw}^{q^i}x^{q^{v+i}+q^{w+i}+q},\end{aligned}\tag{7}$$

where l_{ijl} are the unknown coefficients of the linearized polynomial form of L_{ij}. There are implicitly n^2-s linear relations on the $4n$ unknown coefficients of L_{ij}, as well as the rank restrictions on these maps; thus, for $n>4$ we expect an unique solution, and thus an unique Ψ as well.

Given a public key, there is a fixed relationship $P=T(f_0||f_1)U$. We note that different choices of T can be accommodated by different choices of L_{ij} by (3). In contrast, statistically there is only one selection of U which maintains the structure of the key. Thus $M(f_0||f_1)$, $L_{ij}(M^{-1})_i$, Ψ form distinct equivalent private keys for all invertible M. One can see this result as indicating that the security of $ZHFE$ is more closely related to the IP1S problem than the IP problem.

We therefore have roughly q^{4n^2} equivalent private keys for any given public key. Since there are q^{5n^2+sn} possible choices of private keys, there are on the order of q^{n^2+sn} nonequivalent public keys. Consequently, there is sufficient entropy in public keys.

6 $ZHFE$ Key Modification, $ZHFE^-$

6.1 Design

As mentioned in the previous section, there are many degrees of freedom in selecting f_0 and f_1, even when Ψ and L_{ij} for $(i,j)\in\{0,1\}^2$ are fixed. These facts naturally lead to the question of whether it is possible to develop a "minus" modification of $ZHFE$ preserving the essential injectivity of the original scheme.

Analogous to the analysis in the last section, we compute the degrees of freedom in selecting f_0 and f_1 when the L_{ij} for $(i,j)\in\{0,1\}^2$ are fixed and when the degree bound for Ψ is fixed. Because we are decreasing the dimension of f_0 or f_1 or both, we compute over $\mathbb{F}_q$.

Recall from section 5 that there are n^2 possible nonzero coefficients of a cubic polynomial of the form of Ψ over k, and that with only the degree bound restriction, n^2-s of these must be zero. Expressing this fact over $\mathbb{F}_q$, we see

that there are $n^3 - sn$ linear constraints. Considering the maps $L_{i,j}$ to be of corank c, we require an additional $2cn - 2n$ relations to be satisfied, for a total of $n^3 - sn + 2cn - 2n$ linear constraints. Allow the total combined output dimension of f_0 and f_1 over $\mathbb{F}_q$ to be $n + t$. Since there are $\binom{n}{2} + n = \binom{n+1}{2}$ homogeneous quadratic monomials in each coordinate, there are $(n + t)\binom{n+1}{2}$ coefficients in our linear system.

$$(n+t)\binom{n+1}{2} \geq n^3 - sn + 2cn - 2n$$
$$(n+1)t \geq n^2 - n - 2s + 4c - 4.$$

For realistic values of s, it is possible to get t as low as $n-2$, and $n-1$ is always possible. Thus we consider removing two public equations. For symmetry and simplicity, we choose to remove one coordinate from each of f_0 and f_1, making them both maps from $\mathbb{F}_q^n$ to $\mathbb{F}_q^{n-1}$.

Remark 1 *This technique makes $ZHFE^-$ much more similar to small field schemes. The central map is no longer defined as a pair of maps over the extension field.*

Generation of the central map proceeds exactly as in $ZHFE$, with the exception that the linear maps L_{ij} are now representable as $n \times (n-1)$ matrices with entries in $\mathbb{F}_q$. As with $ZHFE$ we identify the image of L_{ij} with k to obtain relation (1).

Inversion of the central map proceeds exactly as with $ZHFE$. Now since both f_0 and f_1 map into a smaller space, there is a possibility of decryption failure beyond that of $ZHFE$. Under the heuristic that f_0 and f_1 are random quadratic maps from $\mathbb{F}_q^n$ to $\mathbb{F}_q^{n-1}$, one computes the probability that $f_0(y)||f_1(y) = f_0(x)||f_1(y)$ for a fixed x to be q^{2-2n}. While f_0 and f_1 are not random, we expect this quantity to be correct, and therefore the probability of decryption failure is increased by q^{2-2n}. Assuming parameters similar to $ZFHE$, this probability is roughly 2^{-300}, which is well within reason.

6.2 Analysis

The differential analysis from the previous section carries over nearly verbatim to the case of $ZHFE^-$. In particular, the 3-tensor structure of the differential remains essentially the same, though over a slightly diminished space. We therefore conclude that $ZHFE^-$ is as secure as $ZHFE$ against a differential symmetric or invariant attack.

Further, the degree of regularity of a subset of a system of relations is bounded below, as noted in [22], by the degree of regularity of the entire system. Thus, in comparison with any full rank $ZHFE$ scheme of the same Q-rank, the degree of regularity is at least as high, and so once again the resistance to algebraic attacks and attacks in the Kipnis-Shamir model is reduced to Q-rank analysis.

Unlike the differential security criteria, Q-rank is not monotone with respect to the composition of projections, a fact which can be seen by observing that

$g(x) \in k[x]$, where k is an even degree n extension of $\mathbb{F}_q$, defined by $g(x) = x^{2q^{n/2}} + x^2$ clearly has Q-rank 2, whereas the composition with the projection $\pi(x) = x^{q^{n/2}} - x$ produces

$$\begin{aligned}\pi(g(x)) &= (x^{2q^{n/2}} + x^2)^{q^{n/2}} - (x^{2q^{n/2}} + x^2) \\ &= x^{2q^n} + x^{2q^{n/2}} - x^{2q^{n/2}} - x^2 = 0.\end{aligned}$$

This strange result is due to the fact that $g(x)$ maps into a subfield L of k of degree $n/2$ over $\mathbb{F}_q$, and π is the minimal polynomial of L. To verify that this phenomenon does not preclude the use of the minus modifier, we find a bound on the reduction of Q-rank for $ZFHE^-$.

First, we note that all options for removing two equations are equivalent with respect to Q-rank. Therefore our specification that the dimension of each f_i for $i \in \{0, 1\}$ is reduced by one suffices for Q-rank analysis. In this case, the minus modifier projects f_i onto a hyperplane. There is a basis in which this codimension one projection is given by $\pi(x) = x^q - x$. Since Q-rank is invariant under isomorphism, we may take $\tilde{f}_i$ isomorphic to f_i with respect to this basis.

Relative to this basis we may view the operation of projection on the associated matrices to be raising each element to the power q, shifting one unit down and to the right, and subtracting the original, thusly:

$$\pi\begin{bmatrix}\alpha_{11} & \alpha_{12} & \cdots & \alpha_{1,n} \\ \vdots & \vdots & \ddots & \vdots \\ \alpha_{n,1} & \alpha_{n,2} & \cdots & \alpha_{n,n}\end{bmatrix} = \begin{bmatrix}\alpha_{n,n}^q - \alpha_{11} & \alpha_{n,1}^q - \alpha_{12} & \cdots & \alpha_{n,n-1}^q - \alpha_{1,n} \\ \vdots & \vdots & \ddots & \vdots \\ \alpha_{n-1,n}^q - \alpha_{n,1} & \alpha_{n-1,1}^q - \alpha_{n,2} & \cdots & \alpha_{n-1,n-1}^q - \alpha_{n,n}\end{bmatrix}.$$

We are assured that this operation does not reduce the rank by more than one and thus the Q-rank of the public key is reduced by at most two. Since we can control the Q-rank via selection of L_{ij}, we conclude that $ZHFE^-$ is secure against the Kipnis-Shamir minrank attack.

6.3 Suggested Parameters

In this section we propose practical parameters for a realistic implementation of $ZHFE^-$. Since the most costly operations, encryption and decryption, utilize algorithms identical to those of $ZHFE$, and due to the tightness between the security analyses of the two schemes, we recommend parameters similar to those of the original scheme.

In an earlier version of this manuscript, we suggested as a parameter set $(q, n, D, r, c) = (7, 55, 105, 2, 6)$, where q is the size of the base field, n is the degree of the extension k over $\mathbb{F}_q$, D is the degree bound for Ψ (in this case $105 = 2*7^2 + 7$), r is the number of equations removed, and c is the corank of the parameters L_{ij}, having non-intersecting kernels. In discussions with the authors of [33], it became apparent that we overlooked the added restrictions from insisting on corank 6 matrices L_{ij}. Furthermore, we may have been overcautious about the

risk of the Q-rank property of $ZHFE$. Any linear system derived from the Q-rank property is inherently overdefined, and so we dare to be more aggressive. Based in part on their analysis, we propose new parameters for our scheme:

$$108 - ZHFE^- : \quad (q, n, D, r, c) = (7, 55, 393, 2, 3).$$

The experiments of the authors of [33] support the viability of these parameters while retaining the significant advance in key generation efficiency even in the minus case.

These parameters correspond to a public key Q-rank of approximately 6, and a degree of regularity of 9 (est.). Given the overdefined nature of the Q-rank attacks and the above analysis verifying resistance to all other known attacks, we conclude that these parameters achieve a security level greater than 80 bits. The performance and security data are essentially the same as the original scheme with L_{ij} of the same moderate corank, 3.

The main differences between $ZHFE^-$ and its progenitor with the same parameters is key size and encryption time. Since a plaintext is in $\mathbb{F}_7^{55}$, its length is 165 bits. The ciphertext lies in $\mathbb{F}_7^{2*55-2}$ and is thus 324 bits in length. Thus the public key size is determined by the storage requirements of 108 equations in 55 variables over $\mathbb{F}_7$. This quantity is roughly $63.1K$. In comparison, the public key size of $110 - ZHFE(7, 55, 105, 6)$ is $64.3K$, which is about 2% larger. Finally, since $ZHFE^-$ has about 2% fewer public equations than $ZHFE$, encryption is about 2% faster.

7 Conclusion

For many years, multivariate cryptography has had effective tools for building secure and efficient post-quantum signature schemes, but has had much less success for encryption. New schemes such as $ZHFE$ and ABC are promising candidates to fill that gap. Nonetheless, being trapdoor constructions, these schemes can only be trusted after a detailed security analysis.

This work provides much of the security analysis needed to establish trust in the $ZHFE$ construction. In addition to the existing analysis of the difficulty of applying direct algebraic attack to $ZHFE$, we analyze the scheme's security against differential attacks, specify parameters precluding rank attacks, and verify resistance to IP-based equivalent-key attacks. This analysis serves to elucidate the structure of the $ZHFE$ public key, but does not break the cryptosystem, reinforcing the likelihood that the scheme is indeed secure.

The elucidation of the structure of $ZHFE$ also allows us to propose the modified scheme $ZHFE^-$. $ZHFE^-$ modifies the core map of $ZHFE$ and thereby reduces its key size, while still remaining secure with respect to the attacks analyzed above. While the reduction in key size is relatively small, it opens up the possibility of using Ding's idea of constructing an injective multivariate encryption map whose codomain is much larger than its domain, without requring the dimension of the codomain to exceed that of the domain by a factor of two or more, as do all existing schemes that use this approach.

References

1. Shor, P.W.: Polynomial-time algorithms for prime factorization and discrete logarithms on a quantum computer. SIAM J. Sci. Stat. Comp. **26, 1484** (1997)
2. Chen, A.I.T., Chen, M.S., Chen, T.R., Cheng, C.M., Ding, J., Kuo, E.L.H., Lee, F.Y.S., Yang, B.Y.: Sse implementation of multivariate pkcs on modern x86 cpus. CHES 2009, LNCS, Springer, IACR **5747** (2009) 33–48
3. Chen, A.I.T., Chen, C.H.O., Chen, M.S., Cheng, C.M., Yang, B.Y.: Practical-sized instances of multivariate pkcs: Rainbow, tts, and ℓic-derivatives. Post-Quantum Crypto, LNCS **5299** (2008) 95–106
4. Yang, B.Y., Cheng, C.M., Chen, B.R., Chen, J.M.: Implementing minimized multivariate public-key cryptosystems on low-resource embedded systems. 3rd Security of Pervasive Computing Conference, LNCS **3934** (2006) 73–88
5. Ding, J., Schmidt, D.: Rainbow, a new multivariable polynomial signature scheme. ACNS 2005, LNCS **3531** (2005) 164–175
6. Chen, M.S., Yang, B.Y., Smith-Tone, D.: Pflash - secure asymmetric signatures on smart cards. Lightweight Cryptography Workshop 2015 (2015) http://csrc.nist.gov/groups/ST/lwc-workshop2015/papers/session3-smith-tone-paper.pdf.
7. Kipnis, A., Patarin, J., Goubin, L.: Unbalanced oil and vinegar signature schemes. EUROCRYPT 1999. LNCS **1592** (1999) 206–222
8. Patarin, J., Courtois, N., Goubin, L.: Quartz, 128-bit long digital signatures. In Naccache, D., ed.: CT-RSA. Volume 2020 of Lecture Notes in Computer Science., Springer (2001) 282–297
9. Tao, C., Diene, A., Tang, S., Ding, J.: Simple matrix scheme for encryption. [34] 231–242
10. Ding, J., Petzoldt, A., Wang, L.: The cubic simple matrix encryption scheme. [35] 76–87
11. Porras, J., Baena, J., Ding, J.: Zhfe, a new multivariate public key encryption scheme. [35] 229–245
12. Moody, D., Perlner, R.A., Smith-Tone, D.: An asymptotically optimal structural attack on the ABC multivariate encryption scheme. [35] 180–196
13. Smith-Tone, D.: On the differential security of multivariate public key cryptosystems. In Yang, B.Y., ed.: PQCrypto. Volume 7071 of Lecture Notes in Computer Science., Springer (2011) 130–142
14. Perlner, R.A., Smith-Tone, D.: A classification of differential invariants for multivariate post-quantum cryptosystems. [34] 165–173
15. Daniels, T., Smith-Tone, D.: Differential properties of the HFE cryptosystem. [35] 59–75
16. Patarin, J.: Hidden Fields Equations (HFE) and Isomorphisms of Polynomials (IP): Two New Families of Asymmetric Algorithms. In: EUROCRYPT. (1996) 33–48
17. Patarin, J.: Cryptoanalysis of the Matsumoto and Imai Public Key Scheme of Eurocrypt'88. In Coppersmith, D., ed.: CRYPTO. Volume 963 of Lecture Notes in Computer Science., Springer (1995) 248–261
18. Matsumoto, T., Imai, H.: Public Quadratic Polynominal-Tuples for Efficient Signature-Verification and Message-Encryption. In: EUROCRYPT. (1988) 419–453
19. Berlekamp, E.R.: Factoring polynomials over large finite fields. Mathematics of Computation **24** (1970) pp. 713–735

20. Kipnis, A., Shamir, A.: Cryptanalysis of the hfe public key cryptosystem by relinearization. Advances in Cryptology - CRYPTO 1999, Springer **1666** (1999) 788
21. Bettale, L., Faugère, J., Perret, L.: Cryptanalysis of hfe, multi-hfe and variants for odd and even characteristic. Des. Codes Cryptography **69** (2013) 1–52
22. Dubois, V., Gama, N.: The degree of regularity of HFE systems. In Abe, M., ed.: Advances in Cryptology - ASIACRYPT 2010 - 16th International Conference on the Theory and Application of Cryptology and Information Security, Singapore, December 5-9, 2010. Proceedings. Volume 6477 of Lecture Notes in Computer Science., Springer (2010) 557–576
23. Ding, J., Hodges, T.J.: Inverting HFE systems is quasi-polynomial for all fields. In Rogaway, P., ed.: Advances in Cryptology - CRYPTO 2011 - 31st Annual Cryptology Conference, Santa Barbara, CA, USA, August 14-18, 2011. Proceedings. Volume 6841 of Lecture Notes in Computer Science., Springer (2011) 724–742
24. Ding, J., Yang, B.Y.: Degree of regularity for hfev and hfev-. [34] 52–66
25. Dubois, V., Fouque, P.A., Shamir, A., Stern, J.: Practical Cryptanalysis of SFLASH. In Menezes, A., ed.: CRYPTO. Volume 4622 of Lecture Notes in Computer Science., Springer (2007) 1–12
26. Faugere, J.C.: Algebraic cryptanalysis of hidden field equations (hfe) using grobner bases. CRYPTO 2003, LNCS **2729** (2003) 44–60
27. Ding, J., Kleinjung, T.: Degree of regularity for HFE-. IACR Cryptology ePrint Archive **2011** (2011) 570
28. Smith-Tone, D.: Discrete geometric foundations for multivariate public key cryptography. (In Submission)
29. Goubin, L., Courtois, N.: Cryptanalysis of the ttm cryptosystem. In Okamoto, T., ed.: ASIACRYPT. Volume 1976 of Lecture Notes in Computer Science., Springer (2000) 44–57
30. Faugère, J., Gligoroski, D., Perret, L., Samardjiska, S., Thomae, E.: A polynomial-time key-recovery attack on MQQ cryptosystems. In Katz, J., ed.: Public-Key Cryptography - PKC 2015 - 18th IACR International Conference on Practice and Theory in Public-Key Cryptography, Gaithersburg, MD, USA, March 30 - April 1, 2015, Proceedings. Volume 9020 of Lecture Notes in Computer Science., Springer (2015) 150–174
31. Buss, J.F., Frandsen, G.S., Shallit, J.O.: The computational complexity of some problems of linear algebra. Journal of Computer and System Sciences **58** (1999) 572 – 596
32. Wolf, C., Preneel, B.: Equivalent keys in multivariate quadratic public key systems. J. Mathematical Cryptology **4** (2011) 375–415
33. Baena, J., Cabarcas, D., Escudero, D., Porras-Barrera, J., Verbel, J.: Efficient zhfe key generation. In: Post-Quantum Cryptography - 7th International Conference, PQCrypto 2016, Fukuoka, Japan, February 24-26, 2016. Proceedings. (2016)
34. Gaborit, P., ed.: Post-Quantum Cryptography - 5th International Workshop, PQCrypto 2013, Limoges, France, June 4-7, 2013. Proceedings. In Gaborit, P., ed.: PQCrypto. Volume 7932 of Lecture Notes in Computer Science., Springer (2013)
35. Mosca, M., ed.: Post-Quantum Cryptography - 6th International Workshop, PQCrypto 2014, Waterloo, ON, Canada, October 1-3, 2014. Proceedings. Volume 8772 of Lecture Notes in Computer Science., Springer (2014)