# Cryptanalysis of the best HFE-LL' Constructions

Daniel Smith-Tone[1,2] and Cristian Valenzuela[2]

[1] National Institute of Standards and Technology (NIST), Maryland, USA
daniel.smith@nist.gov
[2] University of Louisville, Louisville KY, USA
dcsmit11@louisville.edu, cristian.valenzuela@louisville.edu

**Abstract.** In the last few years, the old idea of internal perturbation for multivariate schemes has been resurrected. A form of this method was proposed with application to HFE and UOV and independently by another team for application to Rainbow. Most recently, a newer and more efficient version of internal perturbation was proposed as an enhanced measure for securing HFE for encryption.
This efficient method, known as the LL' construction, is designed to add little complexity to HFE decryption while increasing the rank of the resulting map to resist the now very effective cryptanalyses powered by MinRank. The basic idea of the construction is to have two small lists of binary linear forms which when multiplied produce rank 1 quadratic forms. Random linear combinations of these products are then added to each of the HFE equations, resulting in a masked HFE. The main trick to make the scheme usable is to encrypt an send many random messages so that statistically it is likely that the legitimate user can find a ciphertext that is not perturbed by the construction and which may be decrypted as a plain HFE ciphertext.
We show that this approach is not secure. In particular, we present a method to recover the *noise support*, a collection of quadratic forms spanning the set of LL' quadratic forms. We then are able to filter out the effect of these maps to recover a compatible HFE map. Finally, we are able to complete the key recovery, achieving efficiently an equivalent private key.

**Keywords:** Public Key Cryptography · Multivariate Cryptography · MinRank

## 1 Introduction

Recent years have witnessed an explosion of research in an area called post-quantum cryptography. Beginning from a rather small community of scientists who took seriously the threats that quantum computers pose to asymmetric cryptography, this field has become mainstream.

This evolution is likely due, in part, to the attention large standards organizations have directed towards post-quantum. In particular the National Institute

of Standards and Technology (NIST) published a call for proposals for post-quantum cryptographic standards in [17] that brought about a great deal of advancement in the science. Indeed, a secondary NIST standardization effort directed at digital signatures, see [21], is still underway and providing an additional incentive for innovation and discovery.

One of the prominent families of schemes proposed for these projects and which are still actively studied is also one of the oldest widely considered areas: multivariate cryptography. The first massively multivariate cryptosystem, $C^*$, see [20], was proposed in the 1980s as a multivariate analogue of RSA, see [27]. Study of these multivariate schemes increased dramatically when Patarin reincarnated the field, breaking $C^*$, see [23], proposing an adaptation called hidden field equations (HFE), see [24] and introducing entirely new methods such as oil and vinegar, see [22], which are still used in some form today.

Work continued on these schemes in the early 2000s, with a cryptanalysis of plain HFE, see [14], augmentations of schemes, see [25,11] and further cryptanalysis, see [13,16], for example. Still, some variants of these ideas survived well into NIST's standardization process. It wasn't until the third round of the process that G$e$MSS [1], the last surviving variant of HFE, was convincingly broken, see [30].

Since this attack broke the last viable HFE descendant, effort has continued to find ways of using these structures. Most of these efforts are focused on finding new modifications of existing structures to resist attacks. For example, in [29] a non-linear modifier was introduced to hide the structure of any generic efficiently invertible map. Even more recently, an old idea from [11] has been revisited and tweaked. In [9] and [15] proposals of internal perturbations for augmenting Rainbow [12] and UOV [19] (as well as HFE), respectively, are presented. More recently, in 2024, another variant of Ding's internal perturbation (ip) modifier was proposed in the context of HFE, see [10]. Similar to the previous variants, the augmentation affects the rank of the scheme, so that attacks using powerful MinRank tools, such as [4], are less efficient. Still, the efficiency of the schemes is greatly harmed.

Late in 2024, another unique form of internal perturbation was proposed for use with HFE, see [26]. This modification, which was named LL', is applied to HFE to increase the rank while retaining a significantly faster (though still painfully slow) inversion of the central map. In fact, the construction attempts to completely bypass the effect of the perturbation by using a decrypt-one-of-many model.

Specifically, the authors propose to use HFE-LL' for encryption in the following mode. One user generates many possible session keys, encrypts them all using the other user's public HFE-LL' key and transmits them all. The receiver attempts to decrypt each of the received ciphertexts by merely assuming that the preimage by chance was not perturbed by the LL' construction very much, or perhaps at all. There is a reasonable chance with a large number of ciphertexts that one among the many will happen to not be altered by LL'. In this way, the user may simply invert as an HFE map, which is much more efficient.

### 1.1 Contributions

In this work we show that the most hopeful application of the HFE-LL' technique is still vulnerable to attack. Specifically, we show that the skewed distribution of plaintexts that are not altered by LL' is recoverable by an adversary from repeated use of a public key. Moreover, once this structure is recovered, we provide a technique for performing a full key recovery, which breaks the scheme with much less complexity than the claimed security level. In addition to the theoretical work, we present some experimental work on recovering the noise component of the scheme and in completing the attack.

The article is organized as follows. In Section 2 we establish and standardize the notation we use throughout the paper. In Section 3 we present HFE and HFE-LL' and comment on relevant cryptanalyses. In the subsequent section we elucidate our method for recovering the noise component of the LL' internal perturbation as well as a method for finding a linear noise kernel, allowing the adversary to project the entire scheme to form a projected HFE. In Section 5 we outline the MinRank techniques relevant for the cryptanalysis of HFE. The following section then presents the specific MinRank instances relevant for completing the attack. Section 7 outlines the different steps required to complete key recovery for HFE-LL', that is, recovering equivalent input and output transformations. The next section provides analysis of the complexity of the attack as well as data from our experiments for several attack steps, including some small scale experiments on complete key recovery.

## 2 Notation

Fields are represented by $F_q$ or $E$. In the first case, $q$ is the size of the field. Other upper-case normal font Latin letters represent functions. Variables and constants in the integers or small fields usually appear as lower-case Latin letters. Variables over extension fields appear as upper-case Latin letters, whereas constants in extension fields are typeset as Greek letters, as are some vector-space isomorphisms. Vectors and matrices are typeset in a bold font.

Often quadratic forms will be represented as matrices. For example, the quadratic form $Q(\mathbf{x})$ may be written

$$Q(\mathbf{x}) = \mathbf{x}\mathbf{Q}\mathbf{x}^\top,$$

where we identify the $1 \times 1$ matrix on the right with its unique coefficient. Quadratic forms may always be written in upper triangular form, and since symmetric matrices cannot represent quadratic forms in characteristic 2 via evaluations in the above manner, we use upper triangular representations. Given a matrix $\mathbf{B}$ we denote by $[\![\mathbf{B}]\!]$ the unique upper triangular matrix with the property that

$$\mathbf{x}\mathbf{B}\mathbf{x}^\top = \mathbf{x}[\![\mathbf{B}]\!]\mathbf{x}^\top.$$

## 3 Relevant Schemes

There are two schemes relevant for our analysis. The first, HFE, was first proposed in [24], while the second HFE-LL' was proposed in [26].

### 3.1 HFE

Let $F_q$ be a finite field with $q$ elements. Let $n$ be a positive integer and define $\phi : F_q^n \to E$ to be an $F_q$-vector space isomorphism from $F_q^n$ to a degree $n$ extension $E$ of $F_q$. Given a degree bound $D$, we may define an HFE polynomial to be a function $f : E \to E$ of the form

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

We may then define the function $F : F_q^n \to F_q^n$ by $F = \phi^{-1} \circ f \circ \phi$. The function $F$ is quadratic because $X^{q^i}$ is a Frobenius power and thus $F_q$-linear. Finally, by selecting invertible linear maps $T, U : F_q^n \to F_q^n$, we may construct the public key

$$P(\mathbf{x}) = T \circ F \circ U(\mathbf{x}).$$

The key feature of this public key is that it is efficiently invertible. To invert $P$, one merely inverts each of the component functions $T$, $F$ and $U$. The inversion of $F$ uses Berlekamp's Algorithm, see [5], whose complexity is dependent upon the degree bound $D$.

The HFE scheme is classically broken by MinRank approaches, most notably [6,30,3]. HFE cannot currently be used efficiently for any parameter set.

### 3.2 HFE-LL'

In [26] a new version of HFE is proposed using a modifier slightly different from previously proposed options. The main idea of the modifier is to introduce random quadratic summands boosting the rank of the system while satisfying a skewed distribution that allows relatively efficient inversion in comparison with other modifiers.

The specific construction works as follows. Let $L_i$ and $L_i'$ for $i \in \{1, \ldots, t\}$ be a collection of linear forms. Then for each $\ell \in \{1, \ldots, n\}$ we introduce the summand

$$\sum_{i=1}^{t} \alpha_{i\ell} L_i(\mathbf{x}) L_i'(\mathbf{x}),$$

where $\alpha_{i\ell} \in F_q$, to the $\ell^{\text{th}}$ HFE polynomial, $F_\ell$. Thus, the entire construction for HFE-LL' is

$$P(\mathbf{x}) = T \circ \widehat{F} \circ U(\mathbf{x}),$$

where $\widehat{F}$ is given by

$$\widehat{F}(\mathbf{x}) = F(\mathbf{x}) + \sum_{\ell=1}^{n} \mathbf{e}_\ell \sum_{i=1}^{t} \alpha_{i\ell} L_i(\mathbf{x}) L'_i(\mathbf{x}).$$

For simplicity of terminology later on, we refer to the span of the $(L_i(\mathbf{u}))(L'_i(\mathbf{u}))$, where $\mathbf{u} = U(\mathbf{x})$ as the noise support of LL'.

We can easily see that the LL' modification augments the MinRank property of the public key by increasing the rank generically by $t$, as noted in [26]. The cost of this modification is less efficient inversion. In particular, for a single ciphertext, the inversion requires guessing the values of the $t$ quadratic forms $(L_i(\mathbf{u})(L'_i(\mathbf{u})$, where $\mathbf{u} = U(\mathbf{x})$, which is no better than previous attempts at increasing the rank of HFE, see [10,15].

The efficiency advantage that LL' has on other methods of increasing the rank of HFE is that the distribution of values of $(L_i(\mathbf{u}))(L'_i(\mathbf{u}))$ is skewed from uniform on $F_q$. In particular, parameters suggested in [26] use the field $F_2$ and rely heavily on the fact that for uniformly random input the probability that $(L_i(\mathbf{u}))(L'_i(\mathbf{u})) = 1$ is merely $1/4$.

The HFE-LL' scheme presented in [26] suggests utilizing the scheme in a decrypt-one-of-many mode in which many ciphertexts are generated and transmitted and only one ciphertext is actually decrypted and used in the protocol. The number of ciphertexts transmitted is a function of the probability that at least one among the ciphertexts satisfies $(L_i(\mathbf{u}))(L'_i(\mathbf{u})) = 0$ for all $i \in \{1, \ldots, t\}$. Decryption, then, requires no guessing of values of $(L_i(\mathbf{u}))(L'_i(\mathbf{u}))$. Instead, it requires treating each received ciphertext as a plain HFE ciphertext which is decrypted and checked to see if it is the legitimate preimage. Thus, there is still a significant cost in communication and decryption time.

## 4 Noise Support Recovery

While the LL' construction does increase the rank of the resulting public key and does allow a more efficient decryption in comparison to recently proposed mutations of HFE, see [10,15], each decrypted ciphertext leaks information about the private key. As we show here, the noise support defined in Section 3 can be recovered given sufficiently many error-free plaintexts. Since the use case of LL' involves a one-of-many decryption, each successful application of the scheme reveals an error-free plaintext.

Notice that when $\mathbf{x}$ is an error-free plaintext, we have that for all $i \in \{1, \ldots, t\}$ the values of $(L_i \circ U(\mathbf{x}))(L'_i \circ U(\mathbf{x}))$ are simultaneously zero. These relations are linear relations on the unknown coefficients of the quadratic monomials in these quadratic forms that we expect to not be satisfied for generic plaintexts. Thus, a method for recovering the noise support is to harvest plaintexts that get decrypted, recovering linear relations in the values of all quadratic monomials. Since the above linear relations are satisfied for each such plaintext and we expect generically for no other linear relations on the quadratic monomials to be universal, this procedure provides a filter for recovering the noise support.

More specifically, for any vector $\mathbf{x} \in F_q^n$, associate with it a vector of its monomial values,

$$\mu_{\mathbf{x}} = (x_1^2, x_1x_2, \ldots, x_1x_n, x_2^2, \ldots, x_n^2).$$

A collection of vectors $\{\mathbf{x}_\ell\}_{1\leq \ell \leq s}$ that satisfy the quadratic relations

$$Q_k(\mathbf{x}_\ell) = \sum_{1\leq i\leq j\leq n} a_{ijk}x_ix_j$$

for $k \in \{1, \ldots, t\}$ necessarily corresponds to a matrix of monomial vectors $\mu_{\mathbf{x}_\ell}$ satisfying the relation

$$\begin{bmatrix} \mu_{\mathbf{x}_1} \\ \vdots \\ \mu_{\mathbf{x}_s} \end{bmatrix} \begin{bmatrix} a_{111} & a_{121} & \cdots & a_{1n1} & a_{221} & \cdots & a_{nn1} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ a_{11t} & a_{12t} & \cdots & a_{1nt} & a_{22t} & \cdots & a_{nnt} \end{bmatrix}^\top = \mathbf{0}_{s\times t}.$$

Thus, the span of all monomial vectors $\mu_{\mathbf{x}}$ for $\mathbf{x}$ satisfying the system $\{Q_k\}_{1\leq k\leq t}$ must lie in a subspace of $F_q^{\binom{n+1}{2}}$ of codimension at least $t$.

The noise support recovery proceeds as follows. We follow the protocol of [26] to establish session keys in order to harvest roughly $\binom{n+1}{2}$ error-free plaintexts. Each such plaintext $\mathbf{x} = (x_1, \ldots, x_n)$ is then formed into a vector of monomial values $\mu_{\mathbf{x}}$. Collecting these monomial vectors into a matrix $\mathbf{M}_\mu$ we obtain a transcript of all monomial values collected from error-free plaintexts. The set of linear relations on these monomial values is acquired by computing the right kernel $\mathbf{K}$ of $\mathbf{M}_\mu$. Each column of $\mathbf{K}$, then represents a linear relation among the monomial values of error-free plaintexts. With enough relations, we expect $\mathbf{K}$ to have rank $t$. The recovery of $\mathbf{K}$ is specified in Algorithm 1.

We expect the rank of $\mathbf{K}$ to be $t$ with high probability as soon as the number of recovered monomial vectors is somewhat larger than the dimension of the ambient space. Specifically, we make the following conjecture.

**Conjecture 1** *Let $\mathbf{M}_\mu$ be a matrix whose $(1+\epsilon)\binom{n+1}{2}$ rows $\mu_{\mathbf{x}}$ are monomial vectors associated with vectors $\mathbf{x}$ all of which simultaneously satisfy $t$ linearly independent homogeneous quadratic forms $Q$. Then with high probability, the rank of $\mathbf{M}_\mu$ is $\binom{n+1}{2} - t$.*

There is good reason to make such a conjecture. As discussed above, the rank of $\mathbf{M}_\mu$ is at most $\binom{n+1}{2} - t$. On the other hand, the $\mu_{\mathbf{x}}$ for arbitrary vectors $\mathbf{x} \in F_q^n$ linearly generate all of $F_q^{\binom{n+1}{2}}$. The set $\{\mu_{\mathbf{y}}\}_{1\leq \mathrm{hw}(\mathbf{y})\leq 2}$ of monomial vectors associated with nonzero vectors of Hamming weight bounded by two forms a basis for the space.

We performed experiments verifying that roughly $\binom{n+1}{2}$ such error-free plaintexts are sufficient to fully define the space. We present the results of our experiments in Section 8.

Once $\mathbf{K}$ is recovered, we have a basis of quadratic forms that are satisfied by every error-free vector. We construct these upper triangular quadratic forms

**Algorithm 1** (`RecoverNoiseSupport`): Generates a matrix whose columns encode relations on monomials satisfied by every signature in the input set.

**Input**: *sigs* an array of error-free signatures
**Output**: $\mathbf{K}$ a matrix whose columns correspond to relations on monomials satisfied by every $\sigma \in sigs$

```
M_μ ← 0_{|sigs|×C(n+1,2)}
for i = 0 to |sigs| − 1 do
    σ ← sigs[i]
    index ← 0
    for j = 0 to n − 1 do
        for k = j to n − 1 do
            M_μ[i][index] ← σ[j] ∗ σ[k]
            index ← index + 1
        end for
    end for
end for
K ← RightKernelMatrix(M_μ)
return K
```

$\mathbf{H}_{n+i}$ by merely "matricizing" each column $\mathbf{K}_i$ of $\mathbf{K}$, that is, each column specifies the upper triangular entries of $\mathbf{H}_{n+i}$ in row-major order. These recovered matrices now span the $(L_i(\mathbf{u}))\,(L'_i(\mathbf{u}))$ for $i \in \{1, \ldots, t\}$.

At this point, there are two methods for completing key recovery. First, note that there are $n$ linear combinations of the public matrices $\mathbf{H}_j$ and the recovered matrices $\mathbf{H}_{n+i}$ eliminating the LL' summands and producing a plain HFE public key. Therefore, performing a MinRank attack on the resulting system of $n + t$ quadratic forms with target rank $d = \lceil \log_q D \rceil$ will succeed in recovering an equivalent HFE map.

In fact, for this method, it is not even necessary for enough plaintexts to be recovered to uniquely define the noise support. If a few, say $t'$, random homogeneous quadratic forms pass through the filter, we may still apply the attack with $n + t + t'$ quadratic forms at the same target rank $d$. The attack is only made more expensive by a small margin, corresponding to the increase in the number of matrices.

Alternatively, after the precise recovery of $\mathbf{K}$, we may by brute force recover a basis $\{\mathbf{B}_1, \ldots, \mathbf{B}_t\}$ of the noise support $\mathrm{Span}(\mathbf{H}_{n+1}, \ldots \mathbf{H}_{n+t})$ consisting of maps corresponding to $\mathbf{L}_i^\top \mathbf{L}'_i$ (or to $\mathbf{L}_i'^\top \mathbf{L}_i$). To filter out such a basis, we utilize a technique to recover an equivalent pair of sets of linear forms $\widetilde{L}_i$ and $\widetilde{L}'_i$ from the standard upper triangular representations $\mathbf{B}_i = [\![\mathbf{L}^\top \mathbf{L}']\!]$ by way of the following lemma.

**Lemma 1** *Let $L$ and $L'$ be two linear forms on $F_q^n$ and let $\mathbf{L}$ and $\mathbf{L}'$ be representations of these linear forms as row vectors. Let $\mathbf{B} = [\![\mathbf{L}^\top \mathbf{L}']\!]$. Finally, suppose that $i$ is the index of the first row of $\mathbf{B}$ that is neither all 0 nor has 1 on the*

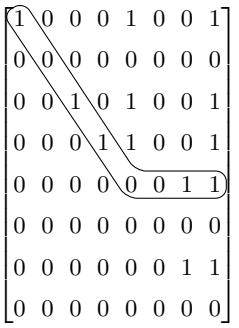


Fig. 1: Example of Lemma 1 applied to the upper triangular form of a product of linear forms $\mathbf{L}^\top \mathbf{L}'$. The outlined vector forms either $\mathbf{L}$ or $\mathbf{L}'$.

*main diagonal, if such a row exists. Then the vector* $\mathbf{v}$ *defined as*

$$\mathbf{v} = \left[\mathbf{B}_{1,1} \cdots \mathbf{B}_{i,i} \ \mathbf{B}_{i,i+1} \cdots \mathbf{B}_{i,n}\right],$$

*see Figure 1, is either equal to* $\mathbf{L}$ *or* $\mathbf{L}'$*. If no such row vector exists, then at least one of* $\mathbf{L}$ *and* $\mathbf{L}'$ *is equal to the main diagonal of* $\mathbf{B}$*.*

*Proof.* The matrix $\mathbf{B}$ can be constructed explicitly by forming the product $\mathbf{L}^\top \mathbf{L}'$, subtracting the lower triangular part of the product and adding its transpose. This process does not affect the main diagonal of the matrix; therefore, for any diagonal coefficient $\mathbf{B}_{i,i} = 1$, we have that both $\mathbf{L}_i = 1$ and $\mathbf{L}' = 1$. If $\mathbf{B}_{i,i} = 0$, we thus have that at most one of $\mathbf{L}_i$ and $\mathbf{L}'_i$ is 1.

We now examine the three ways in which $\mathbf{B}_{i,i} = 0$. If $\mathbf{L}_i = \mathbf{L}'_i = 0$, then both the $i$th row and $i$th columns of the product $\mathbf{L}^\top \mathbf{L}'$ are zero; consequently, the entire $i$th row of $\mathbf{B}$ is zero. If $\mathbf{L}_i = 1$ then the $i$th row of $\mathbf{L}^\top \mathbf{L}'$ is equal to $\mathbf{L}'$, whereas if $\mathbf{L}'_i = 1$ then the $i$th column of this product is $\mathbf{L}^\top$. In either case, if exactly one of these values is 1, then the coefficients $\mathbf{B}_{i,j}$ for $j \geq i$ are either equal to $\mathbf{L}_j$ or $\mathbf{L}'_j$.

There are thus two essentially different ways in which $\mathbf{B}$ may have a row of zeros. Either (1) both $\mathbf{L}_i = 0$ and $\mathbf{L}'_i = 0$, or, (2) without loss of generality, $\mathbf{L}'_i = 1$ and $\mathbf{L}_j = 0$ for all $j \geq i$. In the latter case, we have that the $j$th row of $\mathbf{B}$ is zero for all $j \geq i$, and we may check that $\mathbf{L}$ is equal to the main diagonal of $\mathbf{B}$.

Thus the first index $i$ for which $\mathbf{B}_{i,i} = 0$ but the $i$th row of $\mathbf{B}$ is nonzero, if it exists, has the property that $\mathbf{B}_{i,j}$ for $j \geq i$ are either equal to $\mathbf{L}_j$ or $\mathbf{L}'_j$. All preceding coefficients of $\mathbf{L}$ and $\mathbf{L}'$ are equal and correspond to the diagonal elements of $\mathbf{B}$. The absence of such a row either places us in situation (2) above or indicates that $\mathbf{L}$ and $\mathbf{L}'$ are identical. Either way at least one of $\mathbf{L}$ and $\mathbf{L}'$ is equal to the main diagonal of $\mathbf{B}$. □

Clearly, once a linear form $L$ is recovered by Lemma 1, the recovery of the corresponding linear form $L'$ is obtained by the solution of a small linear system

involving only $t$ unknowns. Since this process is very efficient, we may simply assume that a given $\mathbf{B}$ is in $\mathrm{Span}(\mathbf{H}_{n+1}, \ldots, \mathbf{H}_{n+t})$, recover the candidate $L$ and $L'$ and check to see if this pair forms a valid solution, i.e. one checks whether $\mathbf{B} = [\![\mathbf{L}^\top \mathbf{L}']\!]$. This process can be repeated until a pair of sets of linear forms generating the error support is found.

With either of these sets of linear forms, we may project onto a common kernel of codimension $t$ and obtain a MinRank instance with target rank $d$ and only $n$ matrices. Specifically, given the common left kernel matrix $\mathbf{K}_{\text{noise}}$ of the $\widetilde{L}_i^\top$, we may form the target rank $d$ MinRank instance $\widetilde{\mathbf{H}}_i = \mathbf{K}_{\text{noise}} \mathbf{H}_i \mathbf{K}_{\text{noise}}^\top$ and complete the attack as illustrated in the next section.

## 5 MinRank Attacks against HFE

There are two approaches to recovering an equivalent HFE private key given the noise support $\mathrm{Span}(\mathbf{H}_{n+1}, \ldots, \mathbf{H}_{n+t})$ recovered in the previous section. Both methods rely on the basic technique for a key recovery attack presented in [6]. In this section we review the relevant structures of such an attack that we will have to handle in unorthodox ways to attack HFE-LL'.

Given a primitive element $\theta \in E$ over $F_q$, we define a matrix $\mathbf{M} \in E^{n \times n}$ by

$$\mathbf{M} = \begin{bmatrix} 1 & 1 & \cdots & 1 \\ \theta & \theta^q & \ldots & \theta^{q^{n-1}} \\ \theta^2 & \theta^{2q} & \ldots & \theta^{2q^{n-1}} \\ \vdots & \vdots & \ddots & \vdots \\ \theta^{n-1} & \theta^{(n-1)q} & \ldots & \theta^{(n-1)q^{n-1}} \end{bmatrix}.$$

We note that right multiplication by $\mathbf{M}$ provides an isomorphism between $F_q^n$ and the rank 1 $\mathbf{E}$-algebra $\mathbb{A} = \{(\alpha, \alpha^q, \ldots, \alpha^{q^{n-1}}) : \alpha \in E\}$.

We may also note that $\mathbb{A}$ provides a convenient form for considering the $F_q$-homogeneous quadratic HFE polynomial $f$. Specifically, by identifying $1 \times 1$ matrices with their unique coefficients we have that

$$f(X) = \begin{bmatrix} X & X^q & \cdots & X^{q^{n-1}} \end{bmatrix} \begin{bmatrix} \alpha_{00} & \cdots & \alpha_{0(d-1)} & 0 \cdots 0 \\ \vdots & \ddots & \vdots & \vdots \ddots \vdots \\ 0 & \cdots & \alpha_{(d-1)(d-1)} & 0 \cdots 0 \\ 0 & \cdots & 0 & 0 \cdots 0 \\ \vdots & \ddots & \vdots & \vdots \ddots \vdots \\ 0 & \cdots & 0 & 0 \cdots 0 \end{bmatrix} \begin{bmatrix} X \\ X^q \\ \vdots \\ X^{q^{n-1}} \end{bmatrix}.$$

Furthermore, $\mathbb{A}$ provides a particularly nice way of analyzing Frobenius powers of $f$. Allowing $\mathbf{F}$ to represent the above matrix, we may obtain $\mathbf{F}^{*1}$, the matrix representation of $f^q$ by raising all coefficients to the power of $q$ and shifting them diagonally down and to the right. In such a way we generate matrix representations $\mathbf{F}^{*i}$ for $i \in \{0, \ldots, n-1\}$, i.e. for the entire Galois orbit of $f$.

Now we can view the application of the public key $P$ to a plaintext $\mathbf{a}$ in an essentially new way. Notice that $U(\mathbf{a})\mathbf{M} = (\alpha, \alpha^q, \cdots, \alpha^{q^{n-1}})$, where $\alpha = \phi(\mathbf{a})$. With this representation we may use the matrices $\mathbf{F}^{*i}$ to compute $(f(\alpha), f(\alpha)^q, \cdots, f(\alpha)^{q^{n-1}})$. Finally, an application of $\mathbf{M}^{-1}$ and $T$ produce $P(\mathbf{a})$.

The real utility of these representations lies in the concise description of the properties of the public matrices they provide. As is usual for the analysis of HFE in characteristic 2, we symmetrize the matrices by adding each quadratic form to its transpose. Let $\mathbf{N}'$ represent the matrix $\mathbf{N}+\mathbf{N}^\top$ for any square matrix $\mathbf{N}$. We then have the relation

$$\left[\mathbf{H}'_1\ \mathbf{H}'_2\ \cdots\ \mathbf{H}'_n\right] = \mathbf{UM}\left[\mathbf{F}^{*0\prime}\ \mathbf{F}^{*1\prime}\ \cdots\ \mathbf{F}^{*(n-1)\prime}\right](\mathbf{I}_n \otimes \mathbf{M}^\top\mathbf{U}^\top)(\mathbf{M}^{-1}\mathbf{T} \otimes \mathbf{I}_n),$$

where $\mathbf{H}_i$ are the public matrices of the HFE instance. This equation is equivalent to the relation

$$\left[\mathbf{H}'_1\ \mathbf{H}'_2\ \cdots\ \mathbf{H}'_n\right](\mathbf{V} \otimes \mathbf{I}_n) = \left[\mathbf{WF}^{*0\prime}\mathbf{W}^\top\ \mathbf{WF}^{*1\prime}\mathbf{W}^\top\ \cdots\ \mathbf{WF}^{*(n-1)\prime}\mathbf{W}^\top\right], \quad (1)$$

where $\mathbf{V} = \mathbf{T}^{-1}\mathbf{M}$ and $\mathbf{W} = \mathbf{UM}$. Due to the fact that it is usually the case that $\mathrm{Rank}(\mathbf{F}^{*i\prime}) = \mathrm{Rank}(\mathbf{F}^{*i})$, and in particular $\mathrm{Rank}(\mathbf{F}^{*i\prime})$ is never larger than $d$, this relation shows that, after specializing a coefficient of some matrix, there are $n$ distinct $E$-linear combinations of the symmetrized public matrices $\mathbf{H}'_i$ of rank at most $d$. Moreover, these $n$ solutions, corresponding to the columns of $\mathbf{V}$, form a single Galois orbit, being Frobenius powers of each other.

There are numerous ways of solving such a MinRank instance. The most efficient justified technique is presented in [2]. Their approach uses the support minors method of [4] with some optimizations for addressing the Galois orbit of the solutions in the extension field.

## 6 MinRank Step

We now consider the two approaches to applying MinRank for HFE-LL' key recovery. The unifying theme for these approaches is to use MinRank as a tool for recovering an $E$-basis for noise free polynomials, i.e. the underlying HFE scheme; then a variant of standard key recovery methods may be employed.

As noted in [18], in even rank for characteristic 2, for any solution vector to the HFE MinRank problem $\mathbf{v} = \left[v_1 \cdots, v_n\right]$, any Frobenius power $q^i$ and any pair $(\alpha, \beta) \in E^* \times E$, we have another solution $\left[\alpha v_1^{q^i} + \beta v_1^{q^{i+1}} \ \ldots\ \alpha v_n^{q^i} + \beta v_n^{q^{i+1}}\right]$. These solution vectors correspond to matrices of the form

$$\mathbf{M}_{\text{low rank}} = \alpha\mathbf{F}^{*j} + \beta\mathbf{F}^{*(j+1)}. \quad (2)$$

Although another notable scheme that can be viewed as an HFE derivative, HPPC, see [28], has additional structure that actually simplifies the cryptanalysis of the scheme [8] and prevents the above property from being a problem, in the case of HFE-LL' we are not so lucky, as the following result shows.

**Lemma 2** *Let $\mathbf{H}_1, \ldots, \mathbf{H}_n$ by an HFE-LL' public key with degree bound chosen so that $d$ is even and let $\mathbf{H}_{n+1}, \ldots, \mathbf{H}_{n+t}$ be a basis for the noise support. Let $\mathbf{v} = \left[v_1, \ldots, v_{n+t}\right]$ be any nonzero solution of the MinRank instance $\mathbf{H}_1, \ldots, \mathbf{H}_{n+t}$. Then for all $i \in \{0, \ldots, n-1\}$ and for all $(\alpha, \beta) \in E^* \times E$, the vector $\left[\alpha v_1^{q^i} + \beta v_1^{q^{i+1}} \ \ldots \ \alpha v_{n+t}^{q^i} + \beta v_{n+t}^{q^{i+1}}\right]$ forms another solution vector.*

*Proof.* Suppose that $v_1\mathbf{H}_1 + \cdots + v_{n+t}\mathbf{H}_{n+t}$ is a rank $d$ solution of this MinRank instance. By construction, with high probability any such solution is equal to the map $v_1\widehat{\mathbf{H}}_1 + \cdots + v_n\widehat{\mathbf{H}}_n$, where the $\widehat{\mathbf{H}}_i$ comprise a public key for the related HFE scheme. By [18], we have that for all $(\alpha, \beta) \in E^* \times E$ and for all $i \in \{0, \ldots, n-1\}$ that $(\alpha v_1^{q^i} + \beta v_1^{q^{i+1}})\widehat{\mathbf{H}}_1 + \cdots + (\alpha v_n^{q^i} + \beta v_n^{q^{i+1}})\widehat{\mathbf{H}}_n$ provides all of the solutions of the related MinRank instance $\widehat{\mathbf{H}}_1, \ldots, \widehat{\mathbf{H}}_n$ with high probability. By direct calculation, we see that this latter quantity is equal to the desired MinRank solution $(\alpha v_1^{q^i} + \beta v_1^{q^{i+1}})\mathbf{H}_1 + \cdots + (\alpha v_{n+t}^{q^i} + \beta v_{n+t}^{q^{i+1}})\mathbf{H}_{n+t}$. □

As a result, our strategy is to first recover an $E$-basis for the HFE component that is free from noise from the LL' construction. Once such a basis is recovered, it may be used for the rest of key recovery in a way similar to standard HFE attacks such as [6,2] and we address these issues in Section 7.

For the first approach, notice that the recovered noise quadratic forms $\mathbf{H}_{n+i}$ for $i \in \{1, \ldots, t\}$ have the property that there are $n$ distinct $F_q$-linear combinations $\widehat{\mathbf{H}}_i$ of $\mathbf{H}_1, \ldots, \mathbf{H}_{n+t}$ that form the "public" matrices $\widehat{\mathbf{H}}_i$ of a noise-free HFE instance. Therefore, by the above discussion, there are $n$ distinct solutions of the MinRank instance defined by these matrices, after specialization, that form an $E$-basis of the HFE component. We may recover these matrices by using support minors in a manner similar to [2]. Once one such a matrix is recovered, its Frobenius powers complete a basis, and by Lemma 2, the Frobenius powers of the solution vector are the linear combinations producing these other powers.

The second approach potentially reduces the complexity of the attack for some parameters by splitting the MinRank step into two smaller MinRank instances. First, recall that via the techniques of Section 4, we may recover a noise kernel given by the row space of the matrix $\mathbf{K}_{\text{noise}}$. Given this matrix, we may project the HFE-LL' instance to $\widetilde{\mathbf{H}}_i = \mathbf{K}_{\text{noise}}\mathbf{H}_i\mathbf{K}_{\text{noise}}^\top$. Since we have that $\mathbf{K}_{\text{noise}}\mathbf{L}_i^\top\mathbf{L}_i'\mathbf{K}_{\text{noise}}^\top = \mathbf{0}_{(n-t)\times(n-t)}$, we note that $\widetilde{\mathbf{H}}_i = \mathbf{K}_{\text{noise}}\widehat{\mathbf{H}}_i\mathbf{K}_{\text{noise}}^\top$ as well, where $\widehat{\mathbf{H}}_i$ is a valid "public" matrix of the underlying HFE instance.

Now by similar reasoning as in Section 5, we obtain a relation

$$\left[\widetilde{\mathbf{H}}_1' \ \widetilde{\mathbf{H}}_2' \cdots \widetilde{\mathbf{H}}_n'\right](\mathbf{V} \otimes \mathbf{I}_n) = \left[\widetilde{\mathbf{W}}\mathbf{F}^{*0\prime}\widetilde{\mathbf{W}}^\top \ \widetilde{\mathbf{W}}\mathbf{F}^{*1\prime}\widetilde{\mathbf{W}}^\top \ \ldots \ \widetilde{\mathbf{W}}\mathbf{F}^{*(n-1)\prime}\widetilde{\mathbf{W}}^\top\right],$$

where $\widetilde{\mathbf{W}} = \mathbf{K}_{\text{noise}}\mathbf{W}$. Clearly we once again have solutions with rank at most $d$ to the MinRank instance $\widetilde{\mathbf{H}}_1, \ldots, \widetilde{\mathbf{H}}_n$. Since this MinRank instance only has $n$ matrices, and each matrix is $(n-t) \times (n-t)$, we may solve more efficiently than in the first method.

It remains to show how to recover from the MinRank solution a basis for the noise-free HFE component. Suppose that $\mathbf{v} = \left[v_1 \cdots v_n\right]$ is a solution vector so

that

$$\text{Rank}\left(\sum_{i=1}^{n} v_i \widetilde{\mathbf{H}}_i\right) \leq d.$$

Then with very high probability the matrix $\mathbf{G}$ given by

$$\mathbf{G} = \sum_{i=1}^{n} v_i \widehat{\mathbf{H}}_i,$$

has the form $\alpha \mathbf{F}^{*i} + \beta \mathbf{F}^{*(i+1)}$ over some unknown input basis. We now notice the fact that

$$\mathbf{G} - \sum_{i=1}^{n} v_i \mathbf{H}_i \in \text{Span}(\mathbf{H}_{n+1}, \ldots, \mathbf{H}_{n+t}).$$

Thus, we may construct the matrix

$$\mathbf{G}' = \sum_{i=1}^{n} v_i \mathbf{H}_i,$$

and recover $\mathbf{G}$ by solving the MinRank instance given by the matrix $\mathbf{G}'$ along with the matrices $\mathbf{H}_{n+1}, \ldots, \mathbf{H}_{n+t}$ while specializing the coefficient of $\mathbf{G}'$ to 1. This MinRank instance is very small compared to the first instance, having only $t$ unknown coefficients, though the matrices are $n \times n$ in this case.

The solution vector of this last MinRank instance $\mathbf{v}' = \begin{bmatrix} v_{n+1} & \ldots & v_{n+t} \end{bmatrix}$, along with the solution vector $\mathbf{v}$ of the previous instance together form $\mathbf{v} \| \mathbf{v}'$, a solution vector for the MinRank instance of the first approach above. So an accurate description of the second approach is that it solves the MinRank instance of the first approach in pieces.

As shown in [2, Proposition 1], the MinRank instance arising from an HFE public key can be solved at degree 2. As a result of this fact, a MinRank instance consisting of an $E$-basis for the HFE polynomials must necessarily also be solved at degree 2. Thus, after one of the above two initial steps the complexity of the remaining part of key recovery has the exact same complexity as that of attacking HFE variants, presented in [2].

To determine the complexity of approaches 1 and 2 of this section, notice that [2, Proposition 1] still applies, with a slightly different balance of the parameters. In fact, we find empirically that approach 1 is also solved at degree 2 by using one or two more columns in the support minors routine, though the analysis from [2, Proposition 1] only partially applies. We thus estimate the complexity of the MinRank instance within approach 1 via the formula

$$\text{Complexity}_{\text{Approach 1}} = \mathcal{O}\left((n+t-1)^{\omega}\binom{2d+2}{d}^{\omega}\right),$$

and complexity of the larger instance in approach 2 by

$$\text{Complexity}_{\text{Approach 2}} = \mathcal{O}\left((n-1)^{\omega}\binom{d+\lceil\frac{n(d+1)}{n-t}\rceil}{d}^{\omega}\right),$$

where $\omega$ is the linear algebra exponent. For the main suggested parameters from [26] targetting NIST security level I, $n = 138$, $d = 5$ and $t = 22$, the complexity of approach 1 is $2^{62}$ while the complexity of approach 2 is $2^{63}$. In comparison, the complexity of a MinRank attack directly against an HFE scheme of the same size is $2^{59}$. Thus, the LL' construction adds very little in terms of security.

Either of the above methods successfully recovers an $E$-basis for the HFE component of the scheme that is noise-free. In the next section, we explain the completion of key recovery from such a MinRank solution.

## 7 Completing Key Recovery

As noted in [6], key recovery for HFE in characteristic 2 requires more work than for odd characterstic. In particular, for the case in which $d$ is even, for the solution of the MinRank instance to correspond with an output transformation requires the marriage of the MinRank step with the technique for the resolution of the input transformation. Such a technique is given in [6, Section 6.3.1].

We note here that the basis of the noise-free HFE component recovered in Section 6 is sufficient for applying the methods of [6, Section 6.3] to recover an equivalent input transformation for both the even and odd $d$ cases. The fact that the matrices are defined over $E$ is irrelevant. All that is required for the techniques of [6] for recovery of the input transformation to succeed is that the there is an $E$-linear combination of the matrices producing some $\alpha\mathbf{F}^{*i}$. Since our basis in particular has the noise-free HFE "public" matrices $\widehat{\mathbf{H}}_i$ in their span over $E$, this condition is satisfied.

Given solution vectors $\mathbf{v}$ of the MinRank instance of the previous section, along with the MinRank solution vector $\mathbf{v}'$ of the techniques of [6], we may recover an equivalent output transformation easily. We note that a valid equivalent matrix $\mathbf{V}$ of Equation (1) is given by

$$\mathbf{V} = \left[ \sum_{k=1}^{n} v_i'^{q^{j-1}} v_k^{q^{j+k-1}} \right]_{i,j},$$

for $i, j \in \{1, \ldots, n\}$. Since $\mathbf{V} = \mathbf{T}^{-1}\mathbf{M}$, we may form

$$\mathbf{T}' = \mathbf{M}\mathbf{V},$$

an equivalent HFE output transformation.

Along with the input transformation recovery and HFE central map recovery of [6], the private key recovery of the HFE component of the scheme is complete. To illustrate the full key recovery, we offer a small toy example in Appendix A.

## 8 Experiments

We performed a series of experiments to verify our attack model and the steps of key recovery. Our experiments were performed using the Magma Computational Algebra System[3], see [7].

Due to the fact that the recovery of the noise support is a critical step in our key recovery technique, we ran experiments to verify empirically the claim of Conjecture 1. For several values of $n$ and $t$, we varied the number of randomly generated vectors $\mathbf{x}_j$ for which $\mathbf{x}_j$ satisfied all of the $t$ relations

$$L_i(\mathbf{x})L_i'(\mathbf{x}) = 0,$$

from $\binom{n+1}{2} - t$ up to $\binom{n+1}{2} + b$ for some bound $b$. For each case, we computed the empirical probability that the rank of the space of homogeneous quadratic forms simultaneously satisfied is $t$. In each case, we observed a cut-off phenomenon, with the probability quickly changing from close to 0 to close to 1 as the number of vectors increased; moreover, this cut-off phenomenon seems to be tighter for larger $n$. See Table 1 for a summary of our results.

Table 1: Summary of experiments verifying empirically Conjecture 1. Each entry provides the empirical probability that the dimension of the space of all homogeneous quadratic forms satisfied by $\binom{n+1}{2} - t + b$ vectors $\mathbf{x}$ all of which simultaneously satisfy the $t$ quadratic forms $L_i(\mathbf{x})L_i'(\mathbf{x})$ is exactly $t$.

| $(n,t)$ | $b=0$ | $b=t$ | $b=2t$ | $b=3t$ | $b=4t$ |
|---|---|---|---|---|---|
| $(8,3)$ | 0.10 | 0.47 | 0.73 | 0.89 | 0.97 |
| $(9,3)$ | 0.00 | 0.09 | 0.68 | 0.79 | 0.96 |
| $(9,4)$ | 0.00 | 0.05 | 0.65 | 0.89 | 0.97 |
| $(10,3)$ | 0.00 | 0.23 | 0.84 | 0.90 | 0.95 |
| $(10,4)$ | 0.00 | 0.16 | 0.75 | 0.90 | 0.96 |
| $(10,5)$ | 0.00 | 0.06 | 0.82 | 0.98 | 0.97 |
| $(24,3)$ | 0.00 | 0.27 | 0.90 | 0.99 | 1.00 |

We also performed a series of experiments on the brute force approach to the recovery of the noise kernel used in the second approach of Section 6. Specifically, we generate the upper triangular representations $\mathbf{B}_i$ of the quadratic forms $L_i(\mathbf{x})L_i'(\mathbf{x})$, sample randomly many vectors $\mathbf{x}$ for which all quadratic forms vanish, use Algorithm 1 to recover the noise support and finally brute force a basis of the appropriate form by using Lemma 1. In all cases we recovered the appropriate basis and noise kernel.

[3] Certain commercial equipment, instruments, or materials (or suppliers, or software, ...) are identified in this paper to foster understanding. Such identification does not imply recommendation or endorsement by the National Institute of Standards and Technology, nor does it imply that the materials or equipment identified are necessarily the best available for the purpose.

The method we used is to consider $\mathbf{H}_{n+i} + \mathbf{H}_{n+i}^{\top}$. For a valid basis element, the linear combination of these symmetrized matrices will have rank at most 2 due to subadditivity. For simplicity, we considered all possible elements in the span, filtering by rank and then filtering by using Lemma 1. It may be possible to use rank methods directly in this space for which there are many low rank matrices to speed up the search, but for no realistic parameter sets is this step the limiting part of the attack even when using brute force.

Finally, we ran experiments on the entire key recovery. The dominant step in the attack was the MinRank step of Section 6. Data on the performance for various values of $n$, $d = \lceil \lg D \rceil$ and $t$ are provided in Table 2.

These data support the assertion that for larger instances the second approach may become more efficient than the first. In each instance, the second MinRank instance, consisting of $t + 1$ matrices of dimension $n \times n$ with target rank $d$ is more efficiently solved (requiring fewer total monomials) than the previous system. Which approach is optimal largely depends on the number of columns required in support minors for each instance. In either case, however, these steps are slightly more complex than directly solving an HFE instance of the same size.

Table 2: Timing data for the main MinRank step of key recovery in seconds. The top portion addresses the first approach with MinRank instances consisting of $n + t$ matrices of size $n \times n$. The bottom portion of the table provides data for the second approach with MinRank instances consisting of $n$ matrices of size $(n - t) \times (n - t)$. Recall that $d = \lceil \lg D \rceil$.

| $(n, d)$ | $t = 2$ | $t = 5$ | $t = 10$ | $t = 15$ |
|---|---|---|---|---|
| $(32, 3)$ | 2.14 | 2.62 | 4.32 | 5.55 |
| $(36, 3)$ | 3.75 | 4.12 | 6.97 | 9.02 |
| $(40, 3)$ | 5.34 | 6.47 | 8.57 | 14.66 |
| $(44, 3)$ | 8.30 | 9.37 | 11.38 | 18.46 |
| $(32, 3)$ | 1.90 | 1.86 | 2.47 | 3.92 |
| $(36, 3)$ | 3.32 | 3.29 | 4.07 | 6.52 |
| $(40, 3)$ | 4.97 | 4.93 | 5.66 | 9.33 |
| $(44, 3)$ | 7.49 | 7.30 | 8.16 | 12.62 |

## 9 Conclusion

Recently, several proposals have been made [9,15,10,26] of multivariate schemes for which some variant of internal perturbation, see [11], is employed to protect cryptosystems from structural attacks. In the case of the applications of these methods to the small field schemes it seems that efficiency and sufficient protection against known attacks can be achieved. For HFE, however, even the best proposals are quite inefficient.

In this work we verify that the most efficient technique is actually insecure, providing only a few extra bits of security in comparison to naked HFE at most. At the moment there is no satisfying approach to securing HFE-like schemes, though this goal is an active area of research.

# A Toy Example

We present a small toy example of the full attack on HFE-LL'. This section is divided into two subsections. In the first, we set parameters and derive a public key. In the second, we present the steps of the attack.

## A.1 Key Generation

We fix $q = 2$ so that we work over $F_2$, set $D = 4$ and thus $d = 2$, choose $t = 3$ and use $n = 8$ variables. Let $b$ be a primitive element of $E = F(b)$, a degree $n$ extension. For specificity, let $b$ be a root of $g(x) = x^8 + x^4 + x^3 + x^2 + 1$. For the central polynomial we choose $f(X) = b^{50}X^4 + b^{59}X^3 + b^{200}X^2$. The input an output transformations, respectively, are given by:

$$\mathbf{T} = \begin{bmatrix} 0&1&1&0&0&1&0&0 \\ 0&0&1&1&1&1&0&0 \\ 1&1&1&0&1&0&0&0 \\ 0&0&0&0&1&1&0&1 \\ 1&1&1&0&1&0&0&1 \\ 1&0&0&1&1&1&0&0 \\ 1&0&0&0&0&1&1&0 \\ 0&0&0&0&1&0&0&0 \end{bmatrix} \text{ and } \mathbf{U} = \begin{bmatrix} 0&1&1&0&1&0&0&0 \\ 1&1&0&1&0&0&0&0 \\ 0&0&1&0&0&0&0&1 \\ 1&0&1&0&0&0&0&0 \\ 1&0&1&1&1&1&1&0 \\ 0&0&0&0&1&0&1&1 \\ 1&1&1&1&0&0&0&0 \\ 1&1&0&0&0&1&1&1 \end{bmatrix}.$$

To add noise from the LL' construction, we choose the following lists of linear forms, represented here as row matrices:

$$\mathbf{L}_1 = \begin{bmatrix} 1&0&1&1&1&0&0&0 \end{bmatrix}, \mathbf{L}_2 = \begin{bmatrix} 1&0&0&0&0&0&1&0 \end{bmatrix}, \mathbf{L}_3 = \begin{bmatrix} 0&0&0&0&0&0&0&1 \end{bmatrix}$$
$$\mathbf{L}'_1 = \begin{bmatrix} 1&1&0&1&0&1&0&1 \end{bmatrix}, \mathbf{L}'_2 = \begin{bmatrix} 1&0&1&1&1&0&1&0 \end{bmatrix}, \mathbf{L}'_3 = \begin{bmatrix} 1&0&1&1&0&1&0&0 \end{bmatrix}.$$

Constructing the public key via the composition

$$P = \mathbf{T} \circ (\phi^{-1} \circ f \circ \phi + \sum_{\ell=1}^{n} \mathbf{e}_\ell \sum_{i=1}^{t} \alpha_{i\ell} \mathbf{L}_i^\top \mathbf{L}'_i) \circ \mathbf{U},$$

we generate the public matrices of Figure 2.

## A.2 Key Recovery

The attack proceeds by harvesting numerous plaintexts that are successfully decrypted by the legitimate user. We require a few more than $\binom{n+1}{2} = 36$ for

$$\begin{bmatrix}0&0&0&0&0&1&0&0\\0&1&1&0&1&0&0&0\\0&0&1&0&0&1&0&1\\0&0&0&1&0&0&1&1\\0&0&0&0&0&0&1&0\\0&0&0&0&0&1&1&0\\0&0&0&0&0&0&1&1\\0&0&0&0&0&0&0&1\end{bmatrix}\begin{bmatrix}1&1&1&0&0&1&0&0\\0&1&0&0&1&1&0&0\\0&0&1&0&0&0&1&1\\0&0&0&0&0&1&1&0\\0&0&0&0&1&1&0&0\\0&0&0&0&0&0&1&0\\0&0&0&0&0&0&1&1\\0&0&0&0&0&0&0&0\end{bmatrix}\begin{bmatrix}1&0&1&0&0&0&0&1\\0&1&0&0&1&1&0&1\\0&0&1&1&1&0&0&0\\0&0&0&1&0&0&1&1\\0&0&0&0&0&0&0&1\\0&0&0&0&0&0&1&1\\0&0&0&0&0&0&0&1\\0&0&0&0&0&0&0&1\end{bmatrix}\begin{bmatrix}0&1&0&1&1&1&1&1\\0&1&0&1&0&0&0&1\\0&0&0&0&0&0&0&1\\0&0&0&1&0&0&1&1\\0&0&0&0&0&1&1&0\\0&0&0&0&0&1&0&1\\0&0&0&0&0&0&0&1\\0&0&0&0&0&0&0&0\end{bmatrix}$$

$$\begin{bmatrix}0&1&0&0&0&0&1&0\\0&1&0&0&0&0&1&0\\0&0&1&1&0&0&1&0\\0&0&0&1&0&1&0&1\\0&0&0&0&0&1&0&1\\0&0&0&0&0&0&0&0\\0&0&0&0&0&0&0&0\\0&0&0&0&0&0&0&1\end{bmatrix}\begin{bmatrix}0&1&0&1&1&1&0&0\\0&1&1&0&0&0&0&1\\0&0&1&1&1&0&1&0\\0&0&0&1&0&1&1&0\\0&0&0&0&1&1&0&1\\0&0&0&0&0&0&0&0\\0&0&0&0&0&0&1&0\\0&0&0&0&0&0&0&0\end{bmatrix}\begin{bmatrix}1&1&1&1&1&0&0&1\\0&0&0&0&0&0&1&0\\0&0&0&1&1&0&1&1\\0&0&0&0&1&1&0&1\\0&0&0&0&0&0&0&1\\0&0&0&0&0&0&1&0\\0&0&0&0&0&0&1&1\\0&0&0&0&0&0&0&1\end{bmatrix}\begin{bmatrix}1&1&0&1&0&0&1&0\\0&1&1&0&0&1&1&1\\0&0&0&1&1&0&0&1\\0&0&0&1&1&0&1&0\\0&0&0&0&1&1&1&1\\0&0&0&0&0&1&0&0\\0&0&0&0&0&0&1&0\\0&0&0&0&0&0&0&1\end{bmatrix}$$

Fig. 2: Toy HFE-LL' public key, $\mathbf{H}_1, \ldots, \mathbf{H}_8$.

this example. For each of these data points, we recover the values of all monomials and place these values in a matrix. We recover the kernel of this matrix, which provides the noise support. In this case, the explicit quadratic forms we recover are:

$$\mathbf{Q}_1 = \begin{bmatrix}0&0&1&0&0&1&0&1\\0&0&0&0&0&0&0&0\\0&0&1&0&0&1&1&1\\0&0&0&0&0&0&0&0\\0&0&0&0&0&0&0&0\\0&0&0&0&0&0&1&0\\0&0&0&0&0&0&0&1\\0&0&0&0&0&0&0&0\end{bmatrix}, \mathbf{Q}_2 = \begin{bmatrix}0&0&0&0&0&0&1&0\\0&0&0&0&1&1&0&1\\0&0&0&0&1&0&0&0\\0&0&0&0&1&1&0&1\\0&0&0&0&0&0&0&1\\0&0&0&0&0&1&0&1\\0&0&0&0&0&0&0&0\\0&0&0&0&0&0&0&0\end{bmatrix} \text{ and } \mathbf{Q}_3 = \begin{bmatrix}0&0&0&0&0&0&0&0\\0&0&1&0&1&0&1&0\\0&0&0&1&0&1&1&0\\0&0&0&0&1&0&1&0\\0&0&0&0&0&1&1&0\\0&0&0&0&0&0&1&0\\0&0&0&0&0&0&1&0\\0&0&0&0&0&0&0&0\end{bmatrix}.$$

As noted in [26], we may compute a common kernel of these quadratic forms and continue the attack. In particular, given the above basis of the noise support, we may compute a common kernel for these quadratic forms. The recovered kernel is given as the rowspace of the following matrix:

$$\mathbf{K}_{\text{noise}} = \begin{bmatrix}0&0&0&1&0&0&0&0\\1&0&0&0&0&0&0&0\\0&0&1&0&1&1&0&0\\0&1&0&0&0&0&0&0\\0&0&1&0&1&0&0&1\end{bmatrix}.$$

We may check that $\mathbf{K}_{\text{noise}}\mathbf{L}\mathbf{K}_{\text{noise}}{}^{\top} = \mathbf{0}$ for each of the above basis elements.

The attack may proceed by projecting the entire system onto this noise kernel, forming the matrices $\widetilde{\mathbf{H}}_i = \mathbf{K}_{\text{noise}}\mathbf{H}_i\mathbf{K}_{\text{noise}}^{\top}$ and performing the MinRank

attack. The same vector provides a linear combination of the original public matrices $\mathbf{H}_i$ that still includes contributions from the recovered forms $\mathbf{Q}_1, \mathbf{Q}_2$ and $\mathbf{Q}_3$. Thus, a MinRank attack on this collection of matrices, qualified by setting the coefficient of the recovered matrix to 1, produces the required noise-free HFE basis with coefficient vectors in the manner of Section 6 with high probability. We then apply the standard techniques of [6] or [2] to recover an equivalent HFE map, an equivalent input transformation and the proper linear combination of the coefficient vectors to produce an equivalent output transformation.

In this way we recover an equivalent HFE key consisting of the output and input transformations

$$\mathbf{T}' = \begin{bmatrix} 1&0&0&0&0&0&0&0\\ 1&1&1&1&0&0&1&0\\ 1&0&0&0&1&0&0&1\\ 1&1&0&1&1&0&0&1\\ 0&1&1&1&1&0&0&0\\ 0&0&0&1&0&0&1&0\\ 0&0&1&0&0&1&0&0\\ 0&1&0&0&0&0&0&0 \end{bmatrix}, \mathbf{U}' = \begin{bmatrix} 0&0&1&0&0&0&0&0\\ 1&1&1&0&1&0&0&0\\ 1&0&1&1&1&1&1&1\\ 0&0&1&1&1&0&0&1\\ 0&0&0&0&1&0&0&0\\ 1&1&1&1&1&0&1&0\\ 1&1&1&1&1&1&1&1\\ 1&1&0&0&0&0&1&0 \end{bmatrix},$$

and the HFE central map $f(X) = b^{16}X^4 + b^{158}X^3 + b^{111}X^2$. We may verify our key by computing any vector in the noise kernel, applying the public key, and then inverting via our recovered equivalent HFE key.

## References


1. A. Casanova, J.-C. Faugère, G. Macario-Rat, J. Patarin, L. Perret, J. Ryckeghem: GeMSS: A Great Multivariate Short Signature. available at `https://csrc.nist.gov/CSRC/media/Projects/post-quantum-cryptography/documents/round-3/submissions/GeMSS-Round3.zip` (2020), technical report, National Institute of Standards and Technology
2. Baena, J., Briaud, P., Cabarcas, D., Perlner, R.A., Smith-Tone, D., Verbel, J.A.: Improving support-minors rank attacks: Applications to gemss and rainbow. In: Dodis, Y., Shrimpton, T. (eds.) Advances in Cryptology - CRYPTO 2022 - 42nd Annual International Cryptology Conference, CRYPTO 2022, Santa Barbara, CA, USA, August 15-18, 2022, Proceedings, Part III. Lecture Notes in Computer Science, vol. 13509, pp. 376–405. Springer (2022). `https://doi.org/10.1007/978-3-031-15982-4_13`, `https://doi.org/10.1007/978-3-031-15982-4_13`
3. 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)
4. Bardet, M., Bros, M., Cabarcas, D., Gaborit, P., Perlner, R.A., Smith-Tone, D., Tillich, J., Verbel, J.A.: Improvements of algebraic attacks for solving the rank decoding and minrank problems. In: Moriai, S., Wang, H. (eds.) Advances in Cryptology - ASIACRYPT 2020 - 26th International Conference on the Theory and Application of Cryptology and Information Security, Daejeon, South Korea, December 7-11, 2020, Proceedings, Part I. Lecture Notes in Computer Science, vol. 12491,

pp. 507–536. Springer (2020). https://doi.org/10.1007/978-3-030-64837-4_17, https://doi.org/10.1007/978-3-030-64837-4_17

5. Berlekamp, E.R.: Factoring polynomials over large finite fields. Mathematics of Computation **24**(111), pp. 713–735 (1970)
6. Bettale, L., Faugère, J., Perret, L.: Cryptanalysis of HFE, multi-HFE and variants for odd and even characteristic. Des. Codes Cryptography **69**(1), 1–52 (2013). https://doi.org/10.1007/s10623-012-9617-2, http://dx.doi.org/10.1007/s10623-012-9617-2
7. Bosma, W., Cannon, J., Playoust, C.: The magma algebra system i: The user language. J. Symb. Comput. **24**(3–4), 235–265 (Oct 1997). https://doi.org/10.1006/jsco.1996.0125, https://doi.org/10.1006/jsco.1996.0125
8. Briaud, P., Bros, M., Perlner, R.A., Smith-Tone, D.: Practical attack on all parameters of the hppc signature scheme. In: Selected Areas in Cryptography - SAC 2025. Lecture Notes in Computer Science, Springer (2025)
9. Cartor, R., Cartor, M., Lewis, M., Smith-Tone, D.: Iprainbow. In: Cheon, J.H., Johansson, T. (eds.) Post-Quantum Cryptography - 13th International Workshop, PQCrypto 2022, Virtual Event, September 28-30, 2022, Proceedings. Lecture Notes in Computer Science, vol. 13512, pp. 170–184. Springer (2022). https://doi.org/10.1007/978-3-031-17234-2_9, https://doi.org/10.1007/978-3-031-17234-2_9
10. Cogliati, B., Macario-Rat, G., Patarin, J., Varjabedian, P.: State of the art of HFE variants - is it possible to repair HFE with appropriate modifiers? In: Saarinen, M.O., Smith-Tone, D. (eds.) Post-Quantum Cryptography - 15th International Workshop, PQCrypto 2024, Oxford, UK, June 12-14, 2024, Proceedings, Part II. Lecture Notes in Computer Science, vol. 14772, pp. 144–167. Springer (2024). https://doi.org/10.1007/978-3-031-62746-0_7, https://doi.org/10.1007/978-3-031-62746-0_7
11. Ding, J.: A new variant of the matsumoto-imai cryptosystem through perturbation. PKC 2004, LNCS **2947**, 305–318 (2004)
12. Ding, J., Schmidt, D.: Rainbow, a new multivariable polynomial signature scheme. ACNS 2005, LNCS **3531**, 164–175 (2005)
13. Dubois, V., Fouque, P.A., Shamir, A., Stern, J.: Practical Cryptanalysis of SFLASH. In: Menezes, A. (ed.) CRYPTO. Lecture Notes in Computer Science, vol. 4622, pp. 1–12. Springer (2007)
14. Faugere, J.C.: Algebraic cryptanalysis of hidden field equations (HFE) using grobner bases. CRYPTO 2003, LNCS **2729**, 44–60 (2003)
15. Faugère, J., Macario-Rat, G., Patarin, J., Perret, L.: A new perturbation for multivariate public key schemes such as HFE and UOV. IACR Cryptol. ePrint Arch. p. 203 (2022), https://eprint.iacr.org/2022/203
16. Fouque, P., Granboulan, L., Stern, J.: Differential cryptanalysis for multivariate schemes. In: Cramer, R. (ed.) Advances in Cryptology - EUROCRYPT 2005, 24th Annual International Conference on the Theory and Applications of Cryptographic Techniques, Aarhus, Denmark, May 22-26, 2005, Proceedings. Lecture Notes in Computer Science, vol. 3494, pp. 341–353. Springer (2005). https://doi.org/10.1007/11426639_20, https://doi.org/10.1007/11426639_20
17. Group, C.T.: Submission requirements and evaluation criteria for the post-quantum cryptography standardization process. NIST CSRC (2016), http://csrc.nist.gov/groups/ST/post-quantum-crypto/documents/call-for-proposals-final-dec-2016.pdf

18. Jiang, X., Ding, J., Hu, L.: Kipnis-shamir attack on HFE revisited. In: Pei, D., Yung, M., Lin, D., Wu, C. (eds.) Information Security and Cryptology, Third SKLOIS Conference, Inscrypt 2007, Xining, China, August 31 - September 5, 2007, Revised Selected Papers. Lecture Notes in Computer Science, vol. 4990, pp. 399–411. Springer (2007). https://doi.org/10.1007/978-3-540-79499-8_31, https://doi.org/10.1007/978-3-540-79499-8_31
19. Kipnis, A., Patarin, J., Goubin, L.: Unbalanced oil and vinegar signature schemes. EUROCRYPT 1999. LNCS **1592**, 206–222 (1999)
20. Matsumoto, T., Imai, H.: Public Quadratic Polynominal-Tuples for Efficient Signature-Verification and Message-Encryption. In: EUROCRYPT. pp. 419–453 (1988)
21. NIST Cryptographic Technology Group: Call for Additional Digital Signature Schemes for the Post-Quantum Cryptography Standardization Process. NIST Computer Security Resource Center (2022), https://csrc.nist.gov/csrc/media/Projects/pqc-dig-sig/documents/call-for-proposals-dig-sig-sept-2022.pdf
22. Patarin, J.: The oil and vinegar algorithm for signatures. Presented at the Dagstuhl Workshop on Cryptography (1997)
23. Patarin, J.: Cryptoanalysis of the Matsumoto and Imai Public Key Scheme of Eurocrypt '88. In: Coppersmith, D. (ed.) CRYPTO. Lecture Notes in Computer Science, vol. 963, pp. 248–261. Springer (1995)
24. Patarin, J.: Hidden Fields Equations (HFE) and Isomorphisms of Polynomials (IP): Two New Families of Asymmetric Algorithms. In: EUROCRYPT. pp. 33–48 (1996)
25. Patarin, J., Courtois, N., Goubin, L.: Quartz, 128-bit long digital signatures. In: Naccache, D. (ed.) CT-RSA. Lecture Notes in Computer Science, vol. 2020, pp. 282–297. Springer (2001)
26. Patarin, J., Varjabedian, P.: Multivariate encryptions with ll' perturbations - is it possible to repair HFE in encryption? -. IACR Cryptol. ePrint Arch. p. 1999 (2024), https://eprint.iacr.org/2024/1999
27. Rivest, R.L., Shamir, A., Adleman, L.: A method for obtaining digital signatures and public-key cryptosystems. Commun. ACM **21**(2), 120–126 (Feb 1978). https://doi.org/10.1145/359340.359342, https://doi.org/10.1145/359340.359342
28. Rodriguez, B.G.: HPPC: Hidden Product of Polynomial Composition. NIST Round 1 submission to the Additional Call for Signature Schemes (2023), https://csrc.nist.gov/csrc/media/Projects/pqc-dig-sig/documents/round-1/spec-files/hppc-spec-web.pdf
29. Smith-Tone, D.: New practical multivariate signatures from a nonlinear modifier. In: Cheon, J.H., Tillich, J. (eds.) Post-Quantum Cryptography - 12th International Workshop, PQCrypto 2021, Daejeon, South Korea, July 20-22, 2021, Proceedings. Lecture Notes in Computer Science, vol. 12841, pp. 79–97. Springer (2021). https://doi.org/10.1007/978-3-030-81293-5_5, https://doi.org/10.1007/978-3-030-81293-5_5
30. Tao, C., Petzoldt, A., Ding, J.: Efficient key recovery for all HFE signature variants. In: Malkin, T., Peikert, C. (eds.) Advances in Cryptology - CRYPTO 2021 - 41st Annual International Cryptology Conference, CRYPTO 2021, Virtual Event, August 16-20, 2021, Proceedings, Part I. Lecture Notes in Computer Science, vol. 12825, pp. 70–93. Springer (2021). https://doi.org/10.1007/978-3-030-84242-0_4, https://doi.org/10.1007/978-3-030-84242-0_4