# Key Recovery Attack on the Cubic ABC Simple Matrix Multivariate Encryption Scheme 

Dustin Moody ${ }^{1}$, Ray Perlner ${ }^{1}$, and Daniel Smith-Tone ${ }^{1,2}$<br>${ }^{1}$ National Institute of Standards and Technology, Gaithersburg, Maryland, USA<br>${ }^{2}$ Department of Mathematics, University of Louisville, Louisville, Kentucky, USA<br>dustin.moody@nist.gov, ray.perlner@nist.gov, daniel.smith@nist.gov


#### Abstract

In the last few years multivariate public key cryptography has experienced an infusion of new ideas for encryption. Among these new strategies is the ABC Simple Matrix family of encryption schemes which utilize the structure of a large matrix algebra to construct effectively invertible systems of nonlinear equations hidden by an isomorphism of polynomials. The cubic version of the ABC Simple Matrix Encryption was developed with provable security in mind and was published including a heuristic security argument claiming that an attack on the scheme should be at least as difficult as solving a random system of quadratic equations over a finite field. In this work, we prove that these claims are erroneous. We present a complete key recovery attack breaking full sized instances of the scheme. Interestingly, the same attack applies to the quadratic version of ABC , but is far less efficient; thus, the enhanced security scheme is less secure than the original.


Key words: multivariate public key cryptography, differential invariant, MinRank, encryption

## 1 Introduction

Classical public key cryptography is mainly based on arithmetic constructions on Abelian groups. Since the discovery by Peter Shor in the 1990s of efficient algorithms for factoring and computing discrete logarithms with quantum computers, see [1], there has been a growing interest in the international community in the task of constructing algorithms resistant to cryptanalysis with quantum computers. Indeed, in light of the announcement [2] by the National Institute of Standards and Technology (NIST) of an imminent call for proposals for postquantum standards, the challenge of migrating from the homogeneous heritage of public key cryptography to a more diverse collection of tools has become mainstream.

One possible candidate for practical, efficient, and nonconforming solutions to some of the most consequential public key applications is Multivariate Public Key Cryptography (MPKC). Multivariate schemes are attractive in certain applications because of the maleability of the schemes. Different modifications of similar ideas can make a scheme more suited to lightweight architectures, enhance security, or parametrize various aspects of performance.

In addition, MPKC is one among a few serious candidates to have risen to prominence as post-quantum options. The fundamental problem of solving a system of quadratic equations is known to be NP-hard, and so in the worst case, solving a system of generic quadratic equations is unfeasible for a classical computer; neither is there any indication that the task is easier in the quantum computing paradigm.

MPKC has experienced a fair amount of success in the realm of digital signatures. Some trustworthy schemes that have survived for almost two decades include UOV [3], HFE- [4], and HFEv- [5]. Moreover, some of these schemes have optimizations which have strong theoretical support or have stood unbroken in the literature for some time. Specifically, UOV has a cyclic variant [6] which reduces the key size dramatically, and Gui, a new HFEv- scheme, see [7], has parameters far more appealing than QUARTZ due to greater confidence in the complexity of algebraically solving the underlying system of equations [8].

The situation with multivariate public key encryption is quite different, however. Many attempts at multivariate encryption, see [9,10] for example, have been shown to be weak based on rank or differential weaknesses. Recently, a few interesting attempts to achieve multivariate encryption have surfaced. ZHFE, see [11], and the ABC Simple Matrix Scheme, see [12], both use fundamentally new structures for the derivation of an encryption system. While it was shown that the best attack known on the Simple Matrix structure, see [13] - which relies on the differential invariant structure of the central map - supports the claimed security level of the scheme, a subset of the original authors proposed a cubic version of the scheme, [14], as a step towards provable security.

In this article, we present a key recovery attack on a full scale version of the Cubic Simple Matrix encryption scheme, having a complexity on the order of $q^{s+2}$ for characteristic $p>3, q^{s+3}$ for characteristic 3 and $q^{2 s+6}$ for characteristic 2. Here $s$ is the dimension of the matrices in the scheme, and $q$ is the cardinality of the finite field used. This technique is an extension and augmentation of the technique of [13], and, similarly, exploits a differential invariant property of the core map to perform a key recovery attack. We can show that the attack uses a property which uniquely distinguishes the isomorphism class of the central map from that of a random collection of formulae.

Specifically, our attack breaks CubicABC $(q=127, s=7)$, designed for 80 -bit security, in approximately $2^{76}$ operations (or around $2^{80}$ if one pessimistically uses $\omega=3$ as the linear algebra constant). More convincingly, our attack completely breaks CubicABC $(q=127, s=8)$, designed for 100-bit security, in approximately $2^{84}$ operations (or $2^{88}$ for $\omega=3$ ). Furthermore, the attack is fully parallelizable and requires very little memory; hence, the differential invariant
attack is far more efficient than algebraic attacks, the basis for the original security estimation. Thus, the security claims in [14] are clearly unfounded; in fact, the cubic version of the scheme, whose security was claimed to be closely related to an NP-complete problem, is actually less secure than the quadratic case.

The paper is organized as follows. In the next section, we present the structure of the Cubic ABC Simple Matrix encryption scheme. In the following section, we recall differential invariants and present a natural extension of this notion to the case of cubic polynomials. The differential invariant structure of the ABC scheme is derived in the subsequent section and the effect of this structure on minrank calculations is determined. We next calculate the complexity of the full attack including the linear algebra steps required to extend the distinguisher into a key recovery mechanism. Finally, we review these results and discuss the surprising relationship between the practical security of the Cubic ABC scheme and its quadratic counterpart.

## 2 The Cubic ABC Matrix Encryption Scheme

In [14], the Cubic ABC Matrix encryption scheme is proposed. The motivation behind the scheme is to use a large matrix algebra over a finite field to construct an easily invertible cubic map. The construction uses matrix multiplication to combine random quadratic formulae and random linear formula into cubic formulae in a way that allows a user with knowledge of the structure of the matrix algebra and polynomial isomorphism used to compose the scheme to invert the map.

Let $k=\mathbb{F}_{q}$ be a finite field. Linear forms and variables over $k$ will be denoted with lower case letters. Vectors of any dimension over $k$ will be denoted with bold font, $\mathbf{v}$. Fix $s \in \mathbb{N}$ and set $n=s^{2}$ and $m=2 s^{2}$. An element of $M_{s}(k), M_{n}(k)$ or $M_{m}(k)$, or the linear transformations they represent, will be denoted by upper case letters, such as $M$. When the entries of the matrix are being considered functions of a variable, the matrix will be denoted $M(\mathbf{x})$. Let $\phi$ represent the vector space isomorphism from $M_{s \times 2 s}(k)$ to $k^{2 s^{2}}$ sending a matrix to the column vector consisting of the concatenation of its rows. The output of this map, being a vector, will be written with bold font; however, to indicate the relationship to its matrix preimage, it will be denoted with an upper case letter, such as $\mathbf{M}$.

The scheme utilizes an isomorphism of polynomials to hide the internal structure. Let $\mathbf{x}=\left[x_{1}, x_{2}, \ldots, x_{n}\right]^{\top} \in k^{n}$ denote plaintext while $\mathbf{y}=\left[y_{1}, \ldots, y_{m}\right] \in$ $k^{m}$ denotes ciphertext. Fix two invertible linear transformations $T \in M_{m}(k)$ and $U \in M_{n}(k)$. (One may use affine transformations, but there is no security or performance benefit in doing so.) Denote the input and output of the central map by $\mathbf{u}=U \mathbf{x}$ and $\mathbf{v}=T^{-1}(\mathbf{y})$.

The construction of the central map is as follows. Define three $s \times s$ matrices $A, B$, and $C$ in the following way:

$$
A=\left[\begin{array}{cccc}
p_{1} & p_{2} & \cdots & p_{s} \\
p_{s+1} & p_{s+2} & \cdots & p_{2 s} \\
\vdots & \vdots & \ddots & \vdots \\
p_{s^{2}-s+1} & p_{s^{2}-s+2} & \cdots & p_{s^{2}}
\end{array}\right], B=\left[\begin{array}{cccc}
b_{1} & b_{2} & \cdots & b_{s} \\
b_{s+1} & b_{s+2} & \cdots & b_{2 s} \\
\vdots & \vdots & \ddots & \vdots \\
b_{s^{2}-s+1} & b_{s^{2}-s+2} & \cdots & b_{s^{2}}
\end{array}\right]
$$

and

$$
C=\left[\begin{array}{cccc}
c_{1} & c_{2} & \cdots & c_{s} \\
c_{s+1} & c_{s+2} & \cdots & c_{2 s} \\
\vdots & \vdots & \ddots & \vdots \\
c_{s^{2}-s+1} & c_{s^{2}-s+2} & \cdots & c_{s^{2}}
\end{array}\right]
$$

Here the $p_{i}$ are quadratic forms on $\mathbf{u}$ chosen independently and uniformly at random from among all quadratic forms and the $b_{i}$ and $c_{i}$ are linear forms on $\mathbf{u}$ chosen independently and uniformly at random from among all linear forms.

We define two $s \times s$ matrices $E_{1}=A B$ and $E_{2}=A C$. Since $A$ is quadratic and $B$ and $C$ are linear in $u_{i}, E_{1}$ and $E_{2}$ are cubic in the $u_{i}$. The central map $\mathcal{E}$ is defined by

$$
\mathcal{E}=\phi \circ\left(E_{1} \| E_{2}\right) .
$$

Thus $\mathcal{E}$ is an $m$ dimensional vector of cubic forms in u. Finally, the public key is given by $\mathcal{F}=T \circ \mathcal{E} \circ U$.

Encryption with this system is standard: given a plaintext $\left(x_{1}, \ldots, x_{n}\right)$, compute $\left(y_{1}, \ldots, y_{m}\right)=\mathcal{F}\left(x_{1}, \ldots, x_{n}\right)$. Decryption is somewhat more complicated.

To decrypt, one inverts each of the private maps in turn: apply $T^{-1}$, invert $\mathcal{E}$, and apply $U^{-1}$. To "invert" $\mathcal{E}$, one assumes that $A(\mathbf{u})$ is invertible, and forms a matrix

$$
A^{-1}(\mathbf{u})=\left[\begin{array}{cccc}
w_{1} & w_{2} & \cdots & w_{s} \\
w_{s+1} & w_{s+2} & \cdots & w_{2 s} \\
\vdots & \vdots & \ddots & \vdots \\
w_{s^{2}-s+1} & w_{s^{2}-s+2} & \cdots & w_{s^{2}}
\end{array}\right]
$$

where the $w_{i}$ are indeterminants. Then using the relations $A^{-1}(\mathbf{u}) E_{1}(\mathbf{u})=B(\mathbf{u})$ and $A^{-1}(\mathbf{u}) E_{2}(\mathbf{u})=C(\mathbf{u})$, we have $m=2 s^{2}$ linear equations in $2 n=2 s^{2}$ unknowns $w_{i}$ and $u_{i}$. Using, for example, Gaussian elimination one can eliminate all of the variables $w_{i}$ and most of the $u_{i}$. The resulting relations can be substituted back into $E_{1}(\mathbf{u})$ and $E_{2}(\mathbf{u})$ to obtain a large system of equations in very few variables which can be solved efficiently in a variety of ways.

## 3 Subspace Differential Invariants for Cubic Maps

Let $f: k^{n} \rightarrow k^{m}$ be an arbitrary fixed function on $k^{n}$. Consider the discrete differential $D f(\mathbf{a}, \mathbf{x})=f(\mathbf{a}+\mathbf{x})-f(\mathbf{a})-f(\mathbf{x})+f(\mathbf{0})$.

If $f$ is quadratic, we can express the differential as an $n$-tuple of bilinear differential coordinate forms in the following way: $[D f(\mathbf{a}, \mathbf{x})]_{i}=\mathbf{a}^{\top} D f_{i} \mathbf{x}$, where $D f_{i}$ is a symmetric matrix representation of the action on the $i$ th coordinate of the bilinear differential. If the function $f$ is cubic $\operatorname{Df}(\mathbf{a}, \mathbf{x})$ is a symmetric bi-quadratic function. By the symmetry, it is well defined to compute a second differential $D^{2} f(\mathbf{a}, \mathbf{b}, \mathbf{x})$ by computing the discrete differential of $D f$ with respect to either a or $\mathbf{x}$. In this case, we may consider the second differential as an $n$ tuple of trilinear differential coordinate forms by letting $D^{2} f_{i}$ be the symmetric 3 -tensor representing the action on the $i$ th coordinate of the trilinear differential.

In [13], the following definition of a subspace differential invariant was provided:

Definition $1 A$ subspace differential invariant of a quadratic map $f: k^{n} \rightarrow k^{m}$ with respect to a subspace $X \subseteq k^{m}$ is a subspace $V \subseteq k^{n}$ with the property that there exists a $W \subseteq k^{n}$ of dimension at most $\operatorname{dim}(V)$ such that simultaneously $A V \subseteq W$ for all $A=\sum_{i=1}^{m} x_{i} D f_{i}$ where $\left(x_{1}, \ldots, x_{m}\right) \in X$, i.e. $A \in \operatorname{Span}_{X}\left(D f_{i}\right)$.

This definition captures the idea of a subspace of the span of the public polynomials acting linearly on a subspace of the plaintext space in the same way. Such behavior is strange for quadratic maps in general. Furthermore, as shown in [13], this behavior is computable regardless of the rank of the maps involved.

A natural generalization of this definition is the following:
Definition $2 A$ subspace differential invariant of a cubic map $f: k^{n} \rightarrow k^{m}$ with respect to a subspace $X \subseteq k^{m}$ is a pair of subspaces $\left(V_{1}, V_{2}\right) \subseteq\left(k^{n}\right)^{2}$ for which there exists a subspace $W \subseteq k^{n}$ with $\operatorname{dim}(W) \leq \operatorname{mindim}\left(V_{i}\right)$ such that for all $A=\sum_{i=1}^{m} x_{i} D^{2} f_{i}$ where $\left(x_{1}, \ldots, x_{m}\right) \in X$, for all $\mathbf{a} \in V_{2}$, for all $\mathbf{b} \in V_{2}$ and for all $\mathbf{x} \in W^{\perp}$ we have that $A(\mathbf{a}, \mathbf{b}, \mathbf{x})=0$.

This definition captures the notion of a subspace of the span of the public cubic polynomials acting quadratically on a subspace of the plaintext space in the same way. Such behavior is strange for cubic maps in general.

## 4 The Differential Invariant Structure of the Cubic ABC scheme

### 4.1 Column Band Spaces

Each component of the central $\mathcal{E}(\mathbf{u})=E_{1}(\mathbf{u}) \| E_{2}(\mathbf{u})$ map may be written as:

$$
\begin{equation*}
\mathcal{E}_{(i-1) s+j}=\sum_{l=1}^{s} p_{(i-1) s+l} b_{(l-1) s+j} \tag{1}
\end{equation*}
$$

for the $E_{1}$ equations, and likewise, for the $E_{2}$ equations:

$$
\begin{equation*}
\mathcal{E}_{s^{2}+(i-1) s+j}=\sum_{l=1}^{s} p_{(i-1) s+l} c_{(l-1) s+j} \tag{2}
\end{equation*}
$$

where $i$ and $j$ run from 1 to $s$.
Consider the $s$ sets of $s$ polynomials that form the columns of $E_{1}$, i.e. for each $j \in\{1, \ldots, s\}$ consider $\left(\mathcal{E}_{j}, \mathcal{E}_{s+j}, \ldots, \mathcal{E}_{s^{2}-s+j}\right)$. With high probability, the linear forms $b_{j}, b_{s+j}, \ldots, b_{s^{2}-s+j}$ are linearly independent, and if so the polynomials may be re-expressed, using a linear change of variables to $\left(u_{1}^{\prime}, \ldots u_{s^{2}}^{\prime}\right)$ where $u_{i}^{\prime}=b_{(i-1) s+j}$ for $i=1, \ldots, s$. After the change of variables, the only cubic monomials contained in $\left(\mathcal{E}_{j}, \mathcal{E}_{s+j}, \ldots, \mathcal{E}_{s^{2}-s+j}\right)$ will be those containing at least one factor of $u_{1}^{\prime}, \ldots, u_{s}^{\prime}$. We can make a similar change of variables to reveal structure in the $s$ sets of $s$ polynomials that form the columns of $E_{2}$ : Setting $u_{i}^{\prime}=c_{(i-1) s+j}$ for $i=1, \ldots, s$ and a fixed $j$, the only cubic monomials contained in $\left(\mathcal{E}_{s^{2}+j}, \mathcal{E}_{s^{2}+s+j}, \ldots, \mathcal{E}_{2 s^{2}-s+j}\right)$ will be those containing at least one factor of $u_{1}^{\prime}, \ldots, u_{s}^{\prime}$.

More generally, we can make a similar change of variables to reveal structure in any of a large family of $s$ dimensional subspaces of the span of the component polynomials of $E_{1}$ and $E_{2}$, which we will call column band spaces in analogy to the band spaces used to analyze the quadratic $A B C$ cryptosystem in [13]. Each family is defined by a fixed linear combination, $(\beta, \gamma)$, of the columns of $E_{1}$ and $E_{2}$ :

Definition 3 The column band space defined by the 2 s-dimensional linear form $(\beta, \gamma)$ is the space of cubic maps, $\mathcal{B}_{\beta, \gamma}$, given by:

$$
\mathcal{B}_{\beta, \gamma}=\operatorname{Span}\left(\mathcal{E}_{\beta, \gamma, 1}, \ldots, \mathcal{E}_{\beta, \gamma, s}\right)
$$

where

$$
\begin{gathered}
\mathcal{E}_{\beta, \gamma, i}=\sum_{j=1}^{s}\left(\beta_{j} \mathcal{E}_{(i-1) s+j}+\gamma_{j} \mathcal{E}_{s^{2}+(i-1) s+j}\right) \\
=\sum_{l=1}^{s}\left(p_{(i-1) s+l} \sum_{j=1}^{s}\left(\beta_{j} b_{(l-1) s+j}+\gamma_{j} c_{(l-1) s+j}\right)\right)
\end{gathered}
$$

Theorem 1 There is a pair of subspaces $\left(V_{1}, V_{2}\right) \in\left(k^{n}\right)^{2}$ which is a subspace differential invariant with respect to $\mathcal{B}_{\beta, \gamma}$ for all $(\beta, \gamma)$. Moreover, there exists an $\mathbf{x}_{1} \in k^{n}$ such that $\operatorname{rank}\left(D^{2} \mathcal{E}\left(\mathbf{x}_{1}\right)\right) \leq 2 s$ for all $\mathcal{E} \in \mathcal{B}_{\beta, \gamma}$.

Proof. Note that under a change of variables $\left(x_{1}, \ldots, x_{s^{2}}\right) \stackrel{M}{\longmapsto}\left(u_{1}^{\prime}, \ldots u_{s^{2}}^{\prime}\right)$, where $u_{i}^{\prime}=\sum_{j=1}^{s}\left(\beta_{j} b_{(i-1) s+j}+\gamma_{j} c_{(i-1) s+j}\right)$ for $i=1, \ldots, s$, the only cubic monomials contained in the elements of $\mathcal{B}_{\beta, \gamma}$ will be those containing at least one factor of $u_{1}^{\prime}, \ldots, u_{s}^{\prime}$. In such a basis, the second differential of any map in $\mathcal{B}_{\beta, \gamma}$, and thus the second differential of $\mathcal{E}$ can be visualized as a 3 -tensor with a special block form, see Figure 1.

Let $V$ be the $\left(s^{2}-s\right)$-dimensional preimage $M^{-1}\left(\operatorname{Span}\left(u_{1}^{\prime}, \ldots, u_{s}^{\prime}\right)^{\perp}\right)$. This 3tensor $D^{2} \mathcal{E}$ may be thought of as a bilnear map which takes two vectors $\mathbf{x}_{1}, \mathbf{x}_{2} \in$ $V$, i.e. of the form:

$$
\left(0, \ldots, 0, u_{s+1}^{\prime}\left(\mathbf{x}_{k}\right), \ldots, u_{s^{2}}^{\prime}\left(\mathbf{x}_{k}\right)\right)^{\top}
$$

to a covector of the form:

$$
\left(y\left(u_{1}^{\prime}\right), \ldots, y\left(u_{s}^{\prime}\right), 0, \ldots, 0\right)
$$

Thus, in this basis $D^{2} \mathcal{E}\left(\mathbf{x}_{1}\right)$ is a symmetric matrix which is zero on $V \times V$. Therefore, $\operatorname{rank}\left(D^{2} \mathcal{E}(\mathbf{x})\right) \leq 2 s$. One checks that $(V, V)$ is a subspace differential with respect to $\mathcal{B}_{\beta, \gamma}$ with $W:=V^{\perp}$, since $\operatorname{dim}(W)=s<s^{2}-s=\operatorname{dim}(V)$.


Fig. 1. 3-tensor structure of the second differential of a band space map. Solid regions correspond to nonzero coefficients. Transparent regions correspond to zero coefficients.

We will define the term "band-kernel" to describe the space of vectors of the same form as $x_{1}$ and $x_{2}$ in the proof above, i.e.:

Definition 4 The band kernel of $\mathcal{B}_{\beta, \gamma}$, denoted $\mathcal{B} \mathcal{K}_{\beta, \gamma}$, is the space of vectors, $x$, such that

$$
u_{i}^{\prime}=\sum_{j=1}^{s}\left(\beta_{j} b_{(i-1) s+j}(x)+\gamma_{j} c_{(i-1) s+j}(x)\right)=0
$$

for $i=1, \ldots, s$.

## 5 A Variant of MinRank Exploiting the Column Band Space Structure

A minrank-like attack may be used to locate the column band-space maps defined in the previous section. In this case, the attack proceeds by selecting $s^{2}$ dimensional vectors $\mathbf{x}_{1}, \mathbf{x}_{2}, \mathbf{x}_{3}$, and $\mathbf{x}_{4}$, setting

$$
\begin{align*}
& \sum_{i=1}^{2 s^{2}} t_{i} D^{2} \mathcal{E}_{i}\left(\mathbf{x}_{1}, \mathbf{x}_{2}\right)=0 \\
& \sum_{i=1}^{2 s^{2}} t_{i} D^{2} \mathcal{E}_{i}\left(\mathbf{x}_{3}, \mathbf{x}_{4}\right)=0 \tag{3}
\end{align*}
$$

and solving for the $t_{i}$. The attack succeeds when $\sum_{i=1}^{2 s^{2}} t_{i} \mathcal{E}_{i} \in \mathcal{B}_{\beta, \gamma}$ and $\mathbf{x}_{1}$, $\mathbf{x}_{2}, \mathbf{x}_{3}$, and $\mathbf{x}_{4}$ are all within the corresponding band kernel. If these conditions are met, then the rank of the 2 -tensor $\sum_{i=1}^{2 s^{2}} t_{i} D^{2} \mathcal{E}_{i}\left(\mathbf{x}_{k}\right)$ for $k=1,2,3,4$ will be at most $2 s$, and this will be easily detectable.

The attack complexity will be significantly reduced if several of the $\mathbf{x}_{k}$ are set equal to one another. In odd characteristic fields, we can reduce the number of independently chosen vectors to 2 , (for example, by setting $\mathbf{x}_{1}=\mathbf{x}_{2}$ and $\mathbf{x}_{3}=\mathbf{x}_{4}$.) In characteristic 2 , however, the antisymmetry of the 2 nd differential prevents the equation $\sum_{i=1}^{2 s^{2}} t_{i} D^{2} \mathcal{E}_{i}\left(\mathbf{x}_{1}, \mathbf{x}_{1}\right)=0$ from imposing a nontrivial constraint on the $t_{i}$. Even in characteristic 2, though, the number of independently chosen vectors can be reduced to 3 (e.g. by setting $\mathbf{x}_{1}=\mathbf{x}_{4}$.)

Theorem 2 The probability that 2 randomly chosen vectors, $\mathbf{x}_{1}$ and $\mathbf{x}_{2}$, are both in the band kernel of some band-space $\mathcal{B}_{\beta, \gamma}$ is approximately $\frac{1}{q-1}$; The probability that 3 randomly chosen vectors, $\mathbf{x}_{1}, \mathbf{x}_{2}$, and $\mathbf{x}_{3}$, are all in the band kernel of some band-space $\mathcal{B}_{\beta, \gamma}$ is approximately $\frac{1}{(q-1) q^{s}}$.

Proof. The condition that the $\mathbf{x}_{k}$ are all contained within a band kernel is that there be a nontrivial linear combination of the columns of the following matrix equal to zero (i.e. that the matrix has nonzero column corank):

$$
\left[\begin{array}{cccc|ccc}
b_{1}\left(\mathbf{x}_{1}\right) & b_{2}\left(\mathbf{x}_{1}\right) & \ldots & b_{s}\left(\mathbf{x}_{1}\right) & c_{1}\left(\mathbf{x}_{1}\right) & c_{2}\left(\mathbf{x}_{1}\right) & \ldots  \tag{4}\\
b_{s+1}\left(\mathbf{x}_{1}\right) & b_{s+2}\left(\mathbf{x}_{1}\right) & \ldots & b_{2 s}\left(\mathbf{x}_{1}\right) & c_{s+1}\left(\mathbf{x}_{1}\right) & c_{s+2}\left(\mathbf{x}_{1}\right) & \ldots \\
\vdots & \vdots & \ddots & \vdots & c_{2 s}\left(\mathbf{x}_{1}\right) \\
\vdots & & \vdots & \vdots & \ddots & \vdots \\
b_{s^{2}-s+1}\left(\mathbf{x}_{1}\right) & b_{s^{2}-s+2}\left(\mathbf{x}_{1}\right) & \ldots & b_{s^{2}}\left(\mathbf{x}_{1}\right) & c_{s^{2}-s+1}\left(\mathbf{x}_{1}\right) & c_{s^{2}-s+2}\left(\mathbf{x}_{1}\right) & \ldots \\
\hline b_{1}\left(\mathbf{x}_{2}\right) & b_{2}\left(\mathbf{x}_{2}\right) & \ldots & b_{s}\left(\mathbf{x}_{2}\right) & \left.\mathbf{x}_{1}\right) \\
b_{s+1}\left(\mathbf{x}_{2}\right) & b_{s+2}\left(\mathbf{x}_{2}\right) & \ldots & b_{2 s}\left(\mathbf{x}_{2}\right) & c_{s+1}\left(\mathbf{x}_{2}\right) & c_{2}\left(\mathbf{x}_{2}\right) & \ldots \\
c_{s+2}\left(\mathbf{x}_{2}\right) & \ldots & c_{s}\left(\mathbf{x}_{2}\right) \\
\vdots & \vdots & \ddots & \vdots & c_{2 s}\left(\mathbf{x}_{2}\right) \\
b_{s^{2}-s+1}\left(\mathbf{x}_{2}\right) & b_{s^{2}-s+2}\left(\mathbf{x}_{2}\right) & \ldots & b_{s^{2}}\left(\mathbf{x}_{2}\right) & c_{s^{2}-s+1}\left(\mathbf{x}_{2}\right) & c_{s^{2}-s+2}\left(\mathbf{x}_{2}\right) & \ldots c_{s^{2}}\left(\mathbf{x}_{2}\right) \\
\hline \vdots & \vdots & & \vdots & \vdots & \vdots & \vdots
\end{array}\right]
$$

In the case with 2 randomly chosen vectors, the matrix is a uniformly random $2 s \times 2 s$ matrix, which has nonzero column corank with probability approximately $\frac{1}{q-1}$. In the case with 3 randomly chosen vectors, the matrix is a uniformly random $3 s \times 2 s$ matrix, which has nonzero column corank with probability approximately $\frac{1}{(q-1) q^{s}}$.

Theorem 3 If $\mathbf{x}_{1}, \mathbf{x}_{2}, \mathbf{x}_{3}$, and $\mathbf{x}_{4}$ are chosen in such a way that all four vectors are in the band kernel of a column band space $\mathcal{B}_{\beta, \gamma}$ and also that the symmetric tensor products $\mathbf{x}_{1} \odot \mathbf{x}_{2}$ and $\mathbf{x}_{3} \odot \mathbf{x}_{4}$ are linearly independent from one another and statistically independent from the private quadratic forms, $p_{(i-1) s+j}$ in the matrix $A$, then the tensor products $\mathbf{x}_{1} \otimes \mathbf{x}_{2}$ and $\mathbf{x}_{3} \otimes \mathbf{x}_{4}$ are both in the kernel of some column band-space differential $D^{2} \mathcal{E}=\sum_{\mathcal{E}_{\beta, \gamma, i} \in \mathcal{B}_{\beta, \gamma}} \tau_{i} D^{2} \mathcal{E}_{\beta, \gamma, i}$ with probability approximately $\frac{1}{(q-1) q^{s}}$.

Proof. A $D \mathcal{E}$ meeting the above condition exists iff there is a nontrivial solution to the following system of equations

$$
\begin{align*}
\sum_{\mathcal{E}_{\beta, \gamma, i} \in \mathcal{B}_{\beta, \gamma}} \tau_{i} D^{2} \mathcal{E}_{\beta, \gamma, i}\left(\mathbf{x}_{1}, \mathbf{x}_{2}\right)=0  \tag{5}\\
\sum_{\mathcal{E}_{\beta, \gamma, i} \in \mathcal{B}_{\beta, \gamma}} \tau_{i} D^{2} \mathcal{E}_{\beta, \gamma, i}\left(\mathbf{x}_{3}, \mathbf{x}_{4}\right)=0
\end{align*}
$$

Expressed in a basis (e.g. the $u_{i}^{\prime}$ basis used in Definition 4) where the first $s$ basis vectors are chosen to be outside the band kernel, and the remaining $s^{2}-s$ basis vectors are chosen from within the band kernel, the column band-space differentials, $D^{2} \mathcal{E}_{\beta, \gamma, i}$ are 3 -tensors of the form shown in Figure 1.

Likewise $\mathbf{x}_{1}, \mathbf{x}_{2}, \mathbf{x}_{3}$, and $\mathbf{x}_{4}$ take the form ( $0 \mid \mathbf{x}_{k}$ ). The 2-tensors $D^{2} \mathcal{E}_{\beta, \gamma, i}\left(\mathbf{x}_{k}\right)$ can then be represented by matrices of the form:

$$
D^{2} \mathcal{E}_{\beta, \gamma, i}\left(\mathbf{x}_{k}\right)=\left[\begin{array}{c|c}
S_{k, i} & R_{k, i}  \tag{6}\\
\hline R_{k, i}^{\top} & 0
\end{array}\right]
$$

where $R_{k, i}$ is a random $s \times s^{2}-s$ matrix and $S_{k, i}$ is a random symmetric $s \times s$ matrix. Removing the redundant degrees of freedom we have the system of $2 s$ equations in $s$ variables:

$$
\begin{align*}
& \sum_{i=1}^{s} \tau_{i} R_{1, i} \mathbf{x}_{2}^{\top}=0  \tag{7}\\
& \sum_{i=1}^{s} \tau_{i} R_{3, i} \mathbf{x}_{4}^{\top}=0
\end{align*}
$$

This has a nontrivial solution precisely when the following $2 s \times s$ matrix has nonzero column corank:

This is a random matrix over $k=\mathbb{F}_{q}$, which has nonzero column corank with probability approximately $\frac{1}{(q-1) q^{s}}$, for practical parameters.

To verify that the conditions given in the theorem are sufficient to establish the randomness of the matrix $M$, we can give the following explicit expression for the matrix $M$, which is most easily derived by applying the product rule for the discrete differential to Definition 3:

$$
M=\left[\begin{array}{cccc}
D p_{1}\left(\mathbf{x}_{1}, \mathbf{x}_{2}\right) & D p_{s+1}\left(\mathbf{x}_{1}, \mathbf{x}_{2}\right) & \cdots & D p_{s^{2}-s+1}\left(\mathbf{x}_{1}, \mathbf{x}_{2}\right)  \tag{9}\\
D p_{2}\left(\mathbf{x}_{1}, \mathbf{x}_{2}\right) & D p_{s+2}\left(\mathbf{x}_{1}, \mathbf{x}_{2}\right) & \cdots & D p_{s^{2}-s+2}\left(\mathbf{x}_{1}, \mathbf{x}_{2}\right) \\
\vdots & \vdots & \ddots & \vdots \\
D p_{s}\left(\mathbf{x}_{1}, \mathbf{x}_{2}\right) & D p_{2 s}\left(\mathbf{x}_{1}, \mathbf{x}_{2}\right) & \cdots & D p_{s^{2}}\left(\mathbf{x}_{1}, \mathbf{x}_{2}\right) \\
\hline D p_{1}\left(\mathbf{x}_{3}, \mathbf{x}_{4}\right) & D p_{s+1}\left(\mathbf{x}_{3}, \mathbf{x}_{4}\right) & \cdots & D p_{s^{2}-s+1}\left(\mathbf{x}_{3}, \mathbf{x}_{4}\right) \\
D p_{2}\left(\mathbf{x}_{3}, \mathbf{x}_{4}\right) & D p_{s+2}\left(\mathbf{x}_{3}, \mathbf{x}_{4}\right) & \cdots & D p_{s^{2}-s+1}\left(\mathbf{x}_{3}, \mathbf{x}_{4}\right) \\
\vdots & \vdots & \ddots & \vdots \\
D p_{s}\left(\mathbf{x}_{3}, \mathbf{x}_{4}\right) & D p_{2 s}\left(\mathbf{x}_{3}, \mathbf{x}_{4}\right) & \cdots & D p_{s^{2}}\left(\mathbf{x}_{3}, \mathbf{x}_{4}\right)
\end{array}\right]
$$

Combining the results of Theorems 2 and 3 , we find that for each choice of the vectors $\mathbf{x}_{k}$, there is a column band-space map among the solutions of Equation (3) with probability approximately $\frac{1}{(q-1)^{2} q^{2 s}}$ for even characteristic and $\frac{1}{(q-1)^{2} q^{s}}$ for odd characteristic. Equation (3) is a system of $2 s^{2}$ equations in $2 s^{2}$ variables; one might expect it to generally have a 0-dimensional space of solutions. In some cases, however, there are linear dependencies among the equations, due to the fact that the $D^{2} \mathcal{E}_{i}$ are symmetric tensors. In even characteristic, we get 4 linear dependencies: $D^{2} \mathcal{E}_{i}\left(\mathbf{x}_{1}, \mathbf{x}_{2}\right)\left(\mathbf{x}_{1}\right)=0, D^{2} \mathcal{E}_{i}\left(\mathbf{x}_{1}, \mathbf{x}_{2}\right)\left(\mathbf{x}_{2}\right)=0$, $D^{2} \mathcal{E}_{i}\left(\mathbf{x}_{3}, \mathbf{x}_{4}\right)\left(\mathbf{x}_{3}\right)=0$, and $D^{2} \mathcal{E}_{i}\left(\mathbf{x}_{3}, \mathbf{x}_{4}\right)\left(\mathbf{x}_{4}\right)=0$, and an additional linear dependency when we reduce the number of independent vectors to 3 by setting $\mathbf{x}_{1}=\mathbf{x}_{4}$ : $D^{2} \mathcal{E}_{i}\left(\mathbf{x}_{1}, \mathbf{x}_{2}\right)\left(\mathbf{x}_{3}\right)+D^{2} \mathcal{E}_{i}\left(\mathbf{x}_{3}, \mathbf{x}_{4}\right)\left(\mathbf{x}_{2}\right)=0$, resulting in a 5 -dimensional space of solutions. In characteristic 3 , reducing the number of independent vectors to 2 results in 2 linear dependencies among the equations: e.g. setting $\mathbf{x}_{1}=\mathbf{x}_{2}$ and $\mathbf{x}_{3}=\mathbf{x}_{4}$, we have $D^{2} \mathcal{E}_{i}\left(\mathbf{x}_{1}, \mathbf{x}_{2}\right)\left(\mathbf{x}_{1}\right)=0$ and $D^{2} \mathcal{E}_{i}\left(\mathbf{x}_{3}, \mathbf{x}_{4}\right)\left(\mathbf{x}_{3}\right)=0$. In higher characteristic, there are no linear dependencies imposed on the equations by setting $\mathbf{x}_{1}=\mathbf{x}_{2}$ and $\mathbf{x}_{3}=\mathbf{x}_{4}$.

For characteristic 2, finding the expected 1-dimensional space of band-space solutions in a 5 -dimensional space costs $q^{4}+q^{3}+q^{2}+q+1$ rank operations, which in turn cost $\left(s^{2}\right)^{\omega}$ field operations, where $\omega \approx 2.373$ is the linear algebra constant. Likewise, for characteristic 3, finding the expected 1-dimensional space of band-space solutions in a 2 -dimensional space costs $q+1$ rank operations. Thus the total cost of finding a column band-space map using our variant of MinRank is approximately $q^{2 s+6} s^{2 \omega}$ for charactersitic $2, q^{s+3} s^{2 \omega}$ for characteristic 3 , and $q^{s+2} s^{2 \omega}$ for higher characteristic.

## 6 Complexity of Invariant Attack

The detection of a low rank induced bilinear form $D^{2} \mathcal{E}(x)$ already constitutes a distinguisher from a random system of equations. Extending this calculation to
a full key recovery requires further use of the differential invariant structure of the public key.

First, note that $U$ is not a critical element of the scheme. If $A$ is a random matrix of quadratic forms and $B$ and $C$ are random matrices of linear forms, so are $A \circ U, B \circ U$ and $C \circ U$ for any full rank map $U$. Thus, since clearly $T \circ \phi(A B \| A C) \circ U=T \circ \phi((A \circ U)(B \circ U) \|(A \circ U)(C \circ U))$, we may absorb the action of $U$ into $A, B$, and $C$, and consider the public key to be of the form:

$$
P(\mathbf{x})=T \circ \phi(A B \| A C)(\mathbf{x}) .
$$

Next, consider a trilinear form $D^{2} \mathcal{E}$ in the band space generated by $\mathcal{B}_{\beta, \gamma}$. Since the coefficients of $D^{2} \mathcal{E}$ are products of coefficients of $A$ and coefficients of an element of $\operatorname{Im}(B \| C)$, both of which are uniform i.i.d., there is a change of basis $M$ in which $D^{2} \mathcal{E}$ has the form in Figure 1 and the nonzero coefficients are uniform i.i.d.

Consider $D^{2} \mathcal{E}\left(\mathbf{x}_{1}\right)$ and $D^{2} \mathcal{E}\left(\mathbf{x}_{2}\right)$ for $\mathbf{x}_{1}, \mathbf{x}_{2}$ in the band kernel corresponding to $\mathcal{B}_{\beta, \gamma}$. Being maps from the same band space, there is a basis in which both $D^{2} \mathcal{E}\left(\mathbf{x}_{1}\right)$ and $D^{2} \mathcal{E}\left(\mathbf{x}_{2}\right)$ have the form in Figure 2. Thus, with high probability for $s \geq 2$, the kernels of both maps are contained in the corresponding band kernel, $\mathcal{B}_{\beta, \gamma}$, and $\operatorname{span}\left(\operatorname{ker}\left(D^{2} \mathcal{E}\left(\mathbf{x}_{1}\right) \cup \operatorname{ker}\left(D^{2} \mathcal{E}\left(\mathbf{x}_{2}\right)\right)=\mathcal{B}_{\beta, \gamma}\right.\right.$.


Fig. 2. Structure of the bilinear forms induced by cubic maps in the same band space.

Remark 1 Here we have utilized a property which explicitly distinguishes differential invariant structure from rank structure.

Given the basis for an $s^{2}-s$ dimensional band kernel $\mathcal{B K}$, we may choose a basis $\left\{v_{1}, \ldots, v_{s}\right\}$ for the subspace of the dual space vanishing on $\mathcal{B K}$. We can
also find a basis $\mathcal{E}_{v_{1}}, \ldots, \mathcal{E}_{v_{s}}$ for the band space itself by solving the linear system

$$
\begin{array}{r}
\sum_{\mathcal{E}_{i}} \tau_{i} D^{2} \mathcal{E}_{i}\left(\mathbf{x}_{11}, \mathbf{x}_{12}, \mathbf{x}_{13}\right)=0 \\
\sum_{\mathcal{E}_{i}} \tau_{i} D^{2} \mathcal{E}_{i}\left(\mathbf{x}_{21}, \mathbf{x}_{22}, \mathbf{x}_{23}\right)=0 \\
\vdots=\vdots \\
\sum_{\mathcal{E}_{i}} \tau_{i} D^{2} \mathcal{E}_{i}\left(\mathbf{x}_{t 1}, \mathbf{x}_{t 2}, \mathbf{x}_{t 3}\right)=0
\end{array}
$$

where $t \approx 2 s^{2}$ and $\mathbf{x}_{i j}$ is in the band kernel.
Since the basis $\mathcal{E}_{v_{1}}, \ldots, \mathcal{E}_{v_{s}}$ is in a single band space, there exists an element $\left[b_{1}^{\prime} \cdots b_{s}^{\prime}\right]^{\top} \in \operatorname{ColSpace}(B \| C)$ and two matrices $\Omega_{1}$ and $\Omega_{2}$ such that

$$
\Omega_{1} A\left(\Omega_{2}\left[\begin{array}{c}
b_{1}^{\prime} \\
\vdots \\
b_{s}^{\prime}
\end{array}\right]\right)=: A^{\prime}\left(\left[\begin{array}{c}
v_{1} \\
\vdots \\
v_{s}
\end{array}\right]\right)=\left[\begin{array}{c}
\mathcal{E}_{v_{1}} \\
\vdots \\
\mathcal{E}_{v_{s}}
\end{array}\right]
$$

Solving the above system of equations over $\mathbb{F}_{q}\left[x_{1}, \ldots, x_{s^{2}}\right]$ uniquely determines $A^{\prime}$ in $\mathbb{F}_{q}\left[x_{1}, \ldots, x_{s^{2}}\right] /\left\langle v_{1}, \ldots, v_{s}\right\rangle$. To recover all of $A^{\prime}$, note that the above system is part of an equivalent key

$$
\mathcal{F}=T^{\prime} \circ A^{\prime}\left(B^{\prime} \| C^{\prime}\right)
$$

where $\left[v_{1} \cdots v_{s}\right]^{\top}$ is the first column of $B^{\prime}$.
Applying $T^{\prime-1}$ to both sides and inserting the information we know we may construct the system

$$
\begin{equation*}
A^{\prime}\left(B^{\prime} \| C^{\prime}\right)=T^{\prime-1} \mathcal{F} \tag{10}
\end{equation*}
$$

Solving this system of equations modulo $\left\langle v_{1}, \ldots, v_{s}\right\rangle$ for $B^{\prime}, C^{\prime}$ and $T^{\prime-1}$ we can recover a space of solutions, which we will restrict by arbitrarily fixing the value of $T^{\prime-1}$. Note that the elements of $T^{\prime-1}$ are constant polynomials, and therefore $T^{\prime-1}\left(\bmod \left\langle v_{1}, \ldots, v_{s}\right\rangle\right)$ is the same as $T^{\prime-1}$. Thus, for any choice of $T^{\prime-1}$ in this space, the second column of $T^{\prime-1} \mathcal{F}$ is a basis for a band space. Moreover, the elements $v_{s+1}^{\prime}, \ldots, v_{2 s}^{\prime}$ of the second column of $B^{\prime}\left(\bmod \left\langle v_{1}, \ldots, v_{s}\right\rangle\right)$ are the image, modulo $\left\langle v_{1}, \ldots, v_{s}\right\rangle$, of linear forms vanishing on the corresponding band kernel. Therefore, the intersection $\bigcap_{i=1}^{s} \operatorname{ker}\left(v_{i}\right) \cap \bigcap_{i=s+1}^{2 s} \operatorname{ker}\left(v_{i}^{\prime}\right)$ is the intersection $\mathcal{B} \mathcal{K}_{2} \cap \mathcal{B} \mathcal{K}_{1}$ of the band kernels of our two band spaces.

We can reconstruct the full band kernel of this second band space using the same method we used to obtain our first band kernel: We take a map $\mathcal{E}_{2}$ from the second column of $T^{\prime-1} \mathcal{F}$, and two vectors $x_{a}$ and $x_{b}$ from $\mathcal{B} \mathcal{K}_{2} \cap \mathcal{B} \mathcal{K}_{1}$, and we compute $\mathcal{B} \mathcal{K}_{2}=\operatorname{span}\left(\operatorname{ker}\left(D^{2} \mathcal{E}_{2}\left(\mathbf{x}_{a}\right) \cup \operatorname{ker}\left(D^{2} \mathcal{E}_{2}\left(\mathbf{x}_{b}\right)\right)\right.\right.$. We can now solve for the second column of $B^{\prime},\left[v_{s+1} \cdots v_{2 s}\right]^{\top}$, uniquely over $\mathbb{F}_{q}\left[x_{1}, \ldots, x_{s^{2}}\right]$ (NOT modulo $\left.\left\langle v_{1}, \ldots, v_{s}\right\rangle\right)$ by solving the following system of linear equations:

$$
\begin{array}{r}
v_{i} \equiv v_{i}^{\prime}\left(\bmod \left\langle v_{1}, \ldots, v_{s}\right\rangle\right) \\
v_{i}\left(\mathbf{x}_{1}\right)=0 \\
v_{i}\left(\mathbf{x}_{2}\right)=0 \\
\vdots=\vdots \\
v_{i}\left(\mathbf{x}_{s^{2}-s}\right)=0
\end{array}
$$

where $i=s+1, \ldots, 2 s$, and $\left(\mathbf{x}_{1}, \ldots, \mathbf{x}_{s^{2}-s}\right)$ is a basis for $\mathcal{B} \mathcal{K}_{2}$. We can now solve for $A^{\prime}$ (again, uniquely over $\mathbb{F}_{q}\left[x_{1}, \ldots, x_{s^{2}}\right]$ ) by solving:

$$
\begin{aligned}
& A^{\prime}\left(\left[\begin{array}{c}
v_{1} \\
\vdots \\
v_{s}
\end{array}\right]\right) \equiv\left[\begin{array}{c}
\mathcal{E}_{v_{1}} \\
\vdots \\
\mathcal{E}_{v_{s}}
\end{array}\right]\left(\bmod \left\langle v_{1}, \ldots, v_{s}\right\rangle\right) \\
& A^{\prime}\left(\left[\begin{array}{c}
v_{s+1} \\
\vdots \\
v_{2 s}
\end{array}\right]\right) \equiv\left[\begin{array}{c}
\mathcal{E}_{v_{s+1}} \\
\vdots \\
\mathcal{E}_{v_{2 s}}
\end{array}\right]\left(\bmod \left\langle v_{s+1}, \ldots, v_{2 s}\right\rangle\right)
\end{aligned}
$$

where $\left[\mathcal{E}_{v_{s+1}} \cdots \mathcal{E}_{v_{2 s}}\right]^{\top}$ is the second column of $T^{\prime-1} \mathcal{F}$. This allows us to solve equation 10 for the rest of $B^{\prime}$ and $C^{\prime}$, completing the attack.

The primary cost of the attack involves finding the band space map. The rest of the key recovery is additive in complexity and dominated by the band space map recovery; thus, the total complexity of the attack is of the same order as band space map recovery. Hence, the cost of private key extraction is approximately $q^{2 s+6} s^{2 \omega}$ for characteristic $2, q^{s+3} s^{2 \omega}$ for characteristic 3 , and $q^{s+2} s^{2 \omega}$ for higher characteristic. We note that with these parameters we can break full sized instances of this scheme with parameters chosen for the 80 -bit and 100 -bit security levels via the criteria presented in [14].

Specifically, our attack breaks CubicABC $(q=127, s=7)$, designed for 80 -bit security, in approximately $2^{76}$ operations (or around $2^{80}$ if one pessimistically uses $\omega=3$ as the linear algebra constant). More convincingly, our attack completely breaks CubicABC $(q=127, s=8)$, designed for 100 -bit security, in approximately $2^{84}$ operations (or $2^{88}$ for $\omega=3$ ). Furthermore, the attack is fully parallelizable and requires very little memory; hence, the differential invariant attack is far more efficient than algebraic attacks, the basis for the original security estimation. Thus, the security claims in [14] are clearly unfounded; in fact, the cubic version of the scheme, whose security was claimed to be closely related to an NP-complete problem, is actually less secure than the quadratic case.

We can explain this dramatic discrepancy on the fact that the parameters in [14] are derived by assuming that the algebraic attack is the most effective. In the case of the quadratic ABC scheme, for the proposed parameters, the attack of [13] was slower than the algebraic attack, though asymptotically faster. In the
case of the Cubic scheme, the attack is actually more efficient, in asymptotics as well as for practical parameters.

## 7 Experiments

Using SAGE [15], we performed some minrank computations on small scale variants of the Cubic ABC scheme. The computations were done on a computer with a 64 bit quad-core Intel i7 processor, with clock cycle 2.8 GHz . We were interested in verifying our complexity estimates on the most costly step in the attack, the MinRank instance, rather than the full attack on the ABC scheme. Given as input the finite field size $q$, and the scheme parameter $s$, we computed the average number of vectors $v$ required to be sampled in order for the rank of the 2-tensor $D^{2} \mathcal{E}(v)$ to fall to $2 s$. As explained in Section 5, when the rank falls to this level, we have identified the subspace differential invariant structure of the scheme and can exploit this structure to attack the scheme. Our results for odd $q$ are given in Table 1.

|  | $s=3$ |  |  | $(q-1)^{2} q^{s}$ | $s=4(q-1)^{2} q^{s}$ |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| $s=5(q-1)^{2} q^{s}$ |  |  |  |  |  |  |
| $q=3$ | 14.75 | 108 | 333 | 324 | 952 | 972 |
| $q=5$ | 378 | 2000 | 9986 | 10000 |  |  |
| $q=7$ | 1688 | 12348 | 72951 | 86436 |  |  |
| $q=9$ | 606 | 46656 |  |  |  |  |
| $q=11$ | 13574 | 133100 |  |  |  |  |

Table 1. Average number of vectors needed for the rank to fall to $2 s$ (for odd $q$ )

For higher values of $q$ and $s$ the computations took too long to produce sufficiently many data points and obtain meaningful results with SAGE. When $q$ is odd, our analysis predicted the number of vectors needed would be on the order of $(q-1)^{2} q^{s}$. Table 1 shows the comparison between our experiments and the expected value. We see that for $s=3$, the rank fell quicker than expected, while for $s>3$ the results are quite near the predicted value. This is because when $s=3$ our complexity estimates given in Section 5 are simply not accurate enough, which happens for small values of $q$ and/or $s$.

For even $q$, we also ran some experiments. We found that for $s=3$ and $q=2,4$, or 8 , with high probability only a single vector was needed before the rank fell to $2 s$. For $s=4$ and $s=5$, the computations were only feasible in SAGE for $q=2$. The average number of vectors needed in the $s=4$ case was 244 , with the expected value being $(q-1)^{2} q^{2 s}=256$. With $s=5$, the average number in our experiments was 994 (although the number of trials was small), with the expected value 1024 . For higher values of $q$ and $s$ the computations took too long to obtain meaningful results.

## 8 Conclusion

The ABC schemes are very interesting new ideas for multivariate public key schemes. Essentially all of MPKC can be bisected into big field schemes, utilizing the structure of an extension of the field used for public calculations, and small field schemes which require no such extension. (For the purpose of this comment we consider "medium" field schemes to be big field schemes.)

The ABC cryptosystems present a fundamentally new structure for the development of schemes. In fact, if we consider the structure of simple algebras over the public field (which are surely the only such structures we should consider for secure constructions) then "big field" and "big matrix algebra" complete the picture of possible large structure schemes.

It is interesting to note that the authors provide in [14] a heuristic security argument for the scheme and, as reinforced in the first presentation of the scheme at [16], suggest that with some work the scheme may be able to be shown provably secure. The idea behind their argument is at least somewhat reasonable, if not precise. Their argument essentially amounts to the following: every cubic polynomial in the public key is in the ideal generated by the quadratic forms in $A$ under a certain basis; thus, one might expect the public key to contain a subset of the information one would obtain by applying one step of a Gröbner basis algorithm such as F4, see [17].

Unfortunately, this analysis is not very tight. In fact, we exploit the subspace differential invariant structure inherent to the ABC methodology to show that for odd characteristic the cubic scheme is less secure than its quadratic counterpart. We may therefore conclude that any attempt at a secure cubic "big matrix algebra" scheme must rely on the application of modifiers. The challenge, then, is to construct such a scheme which is still essentially injective for the purpose of encryption. Schemes such as this one can never compete with the secure multivariate options for digital signatures we already know.

We are thus left with the same lingering question that has been asked for the last two decades: Is secure multivariate encryption possible? Currently there is a small list of candidates none of which has both been extensively reviewed and has existed for longer than a few years. If we are to discover a secure multivariate encryption scheme with a convincing security proof or some other security metric, it will require some new techniques and new science. Only time will tell.

## 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, L., Jordan, S., Liu, Y., Moody, D., Peralta, R., Perlner, R., SmithTone, D.: Report on post-quantum cryptography. NISTIR 8105 (2016) http://dx.doi.org/10.6028/NIST.IR. 8105 .
3. Kipnis, A., Patarin, J., Goubin, L.: Unbalanced oil and vinegar signature schemes. EUROCRYPT 1999. LNCS 1592 (1999) 206-222
4. Patarin, J., Goubin, L., Courtois, N.: $\mathrm{C}^{*}+$ and HM: Variations around two schemes of T.Matsumoto and H.Imai. Asiacrypt 1998, Springer 1514 (1998) 3549
5. 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
6. Petzoldt, A., Bulygin, S., Buchmann, J.: Cyclicrainbow - a multivariate signature scheme with a partially cyclic public key. In Gong, G., Gupta, K.C., eds.: INDOCRYPT. Volume 6498 of Lecture Notes in Computer Science., Springer (2010) 33-48
7. Petzoldt, A., Chen, M., Yang, B., Tao, C., Ding, J.: Design principles for hfevbased multivariate signature schemes. In Iwata, T., Cheon, J.H., eds.: Advances in Cryptology - ASIACRYPT 2015-21st International Conference on the Theory and Application of Cryptology and Information Security, Auckland, New Zealand, November 29 - December 3, 2015, Proceedings, Part I. Volume 9452 of Lecture Notes in Computer Science., Springer (2015) 311-334
8. Ding, J., Yang, B.Y.: Degree of regularity for hfev and hfev-. [18] 52-66
9. 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
10. Tsujii, S., Gotaishi, M., Tadaki, K., Fujita, R.: Proposal of a signature scheme based on sts trapdoor. In Sendrier, N., ed.: PQCrypto. Volume 6061 of Lecture Notes in Computer Science., Springer (2010) 201-217
11. Porras, J., Baena, J., Ding, J.: Zhfe, a new multivariate public key encryption scheme. [16] 229-245
12. Tao, C., Diene, A., Tang, S., Ding, J.: Simple matrix scheme for encryption. [18] 231-242
13. Moody, D., Perlner, R.A., Smith-Tone, D.: An asymptotically optimal structural attack on the ABC multivariate encryption scheme. [16] 180-196
14. Ding, J., Petzoldt, A., Wang, L.: The cubic simple matrix encryption scheme. [16] 76-87
15. Developers, T.S.: SageMath, the Sage Mathematics Software System (Version 7.3). (2016) http://www.sagemath.org.
16. 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)
17. Faugere, J.C.: A new efficient algorithm for computing grobner bases (f4). Journal of Pure and Applied Algebra 139 (1999) 61-88
18. 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)
