Security Analysis on an ElGamal-like Multivariate Encryption Scheme Based on Isomorphism of Polynomials

Yasuhiko Ikematsu[1], Shuhei Nakamura[2], Bagus Santoso[3], and Takanori Yasuda[4]

[1] Institute of Mathematics for Industry, Kyushu University 744, Motooka, Nishi-ku, Fukuoka 819–0395, Japan ikematsu@imi.kyushu-u.ac.jp
[2] Department of Liberal Arts and Basic Sciences, Nihon University, 1-2-1 Izumi-cho, Narashino, Chiba 275-8575, Japan
nakamura.shuhei@nihon-u.ac.jp
[3] Department of Computer and Network Engineering, The University of Electro-Communications, 1-5-1 Chofugaoka, Chofu, Tokyo 182-8585, Japan
santoso.bagus@uec.ac.jp
[4] Institute for the Advancement of Higher Education, Okayama University of Science, 1-1 Ridaicho, Kitaku, Okayama 700-0005, Japan tyasuda@bme.ous.ac.jp

Abstract. Isomorphism of polynomials with two secrets (IP2S) problem was proposed by Patarin et al. at Eurocrypt 1996 and the problem is to find two secret linear maps filling in the gap between two polynomial maps over a finite field. At PQC 2020, Santoso proposed a problem originated from IP2S, which is called block isomorphism of polynomials with circulant matrices (BIPC) problem. The BIPC problem is obtained by linearizing IP2S and restricting secret linear maps to linear maps represented by circulant matrices. Using the commutativity of products of circulant matrices, Santoso also proposed an ElGamal-like encryption scheme based on the BIPC problem. In this paper, we give a new security analysis on the ElGamal-like encryption scheme. In particular, we introduce a new attack (called linear stack attack) which finds an equivalent key of the ElGamal-like encryption scheme by using the linearity of the BIPC problem. We see that the attack is a polynomial-time algorithm and can break some 128-bit proposed parameters of the ElGamal-like encryption scheme within 10 hours on a standard PC.

Keywords: Public Key Cryptography · Post Quantum Cryptography (PQC) · Multivariate Public Key Cryptography (MPKC) · Isomorphism of Polynomials

1 Introduction

RSA and ECC are widely-used public key cryptosystems and are based on hard computational problems such as integer factorization problem and discrete logarithm problem, respectively. In 1997, P. Shor [19] showed polynomial-time quantum algorithms to solve these problems using a large scale quantum computer.

Therefore, before a large scale quantum computer gets realized, we need to develop cryptosystems having a resistance to quantum computer attacks. The research area to study such cryptosystems is called post quantum cryptography (PQC) [2].

Multivariate public key cryptography (MPKC) [7] is considered as one of the main candidates of PQC and is constructed based on hard computational problems on multivariate polynomials over finite fields. A main hard computational problem is the MQ problem, which finds a solution to a system of multivariate quadratic equations over a finite field. So far, there have been proposed various schemes based on the MQ problem. In particular, regarding signature schemes, Rainbow [8], GeMSS [4] and MQDSS [18] were selected as second round candidates of NIST PQC standardization project [12]. (Rainbow recently became a finalist of the project [6].) However, it is considered that there is no notable multivariate encryption scheme since most of the proposed schemes were not secure or had a large public key size.

Isomorphism of polynomials with two secrets (IP2S) problem is another problem in MPKC and was proposed by Patarin et al. at Eurocrypt 1996 [13]. The IP2S problem is to find two secret invertible linear maps representing the isomorphism between two multivariate polynomial maps over a finite field. Similar to the zero-knowledge interactive proof of graph isomorphism, Patarin proposed an interactive proof based on the IP2S problem. An authentication scheme based on the interactive proof scheme with its proof against impersonation attack is proposed in [15] and the security of the signature scheme based on the authentication scheme against quantum adversary in quantum random oracle model is proven in [17]. When the secret solutions of the IP2S problem are not restricted to invertible maps, we get another computational problem called Morphism of Polynomials (MP) problem which is proven to be an NP-hard problem [14]. Wang et al. [21] proposed a paradigm of constructing a public key encryption (PKE) scheme by using the Diffie-Hellman like algebraic structure derived from restricting the secrets/solutions of the MP problem into circulant matrices to obtain the commutativity. However, as the circulant matrices can be represented with few variables, it suffers from degradation of complexity as one can obtain a sufficient system of equations to solve the problem efficiently. Using this fact, Chen et al. [5] proposed an algebraic attack algorithm.

At PQC 2020, Santoso [16] proposed a new computational problem originated from the IP2S problem, which is called block isomorphism of polynomials with circulant matrices (BIPC) problem. The BIPC problem is obtained by linearizing IP2S and restricting secret linear maps to linear maps represented by circulant matrices. Moreover, similar to Wang et al.'s idea, using the commutativity of products of circulant matrices, Santoso proposed an ElGamal-like BIPC encryption scheme [16] and provided a security proof of the scheme based on the hardness of a Computational Diffie-Hellman (CDH)-like problem derived from the BIPC problem. In the BIPC problem, the secret solution is in the form of pairs of circulant matrices, instead of only one pair of matrices as in the IP2S problem. Therefore, although the secret solutions are circulant matrices, the

number of variables can be adjusted to be sufficiently large to avoid Chen et al.'s algebraic attack. In [16], Santoso gave two attacks against the BIPC problem and selected four types of parameters which are called (a) conservative, (b) alternative, (c) extremely aggressive, and (d) moderately aggressive.

In this paper, we analyze the security of the ElGamal-like BIPC encryption scheme. We discuss a new attack, which is to find an equivalent key of the ElGamal-like encryption scheme by using the linearity of the BIPC problem (called *linear stack attack*). Our core idea is to show that there exists in fact an equivalent secret solution of the CDH-BIPC problem which consists of a set of pairs of circulant matrices. Note that the target of the linear stack attack is not for the BIPC problem but for the CDH-BIPC problem. Based on this idea, we can construct an equivalent key by randomly choosing enough set of pairs of circulant matrices and taking appropriately their scalar multiplications. We show that the linear stack attack is a polynomial-time algorithm and confirm that the attack is efficient for the proposed parameters (a),(b),(c) and (d). In fact, our experimental results showed that the 128-bit security parameters [16] in (b),(c) and (d) were broken within 10 hours with a standard PC. Regarding (a), the 128-bit security parameter [16] did not finish within a week. Instead, we performed experiments for the 80-bit security parameter in (a) and confirmed that it was broken within 5 days.

As far as our knowledge, our attack[5] is the first algorithm which successfully breaks the ElGamal-like BIPC encryption scheme proposed in PQC 2020 [16]. Later, in an independent work [10], Hashimoto proposed a different attack algorithm that solves the BIPC problem directly. It should be noted that the attack algorithm in [10] is specifically developed for BIPC since it relies heavily on algebraic techniques which use the circulant properties of BIPC as the core. On the other hand, our attack algorithm is more simple and general, as we do not use the circulant properties of BIPC as the core of our algorithm. Therefore, at least in principle, our attack algorithm is more generalizable compared to the one in [10] and it may contribute to the future cryptanalysis of other IP2S based encryption schemes.

Our paper is organized as follows. In Section 2, we briefly recall the IP2S problem, the BIPC problem and the ElGamal-like encryption scheme. Moreover, we review the previous security analysis against the BIPC problem. In Section 3, we describe the linear stack attack and perform experiments for the proposed parameters in [16]. Finally, we conclude our paper in Section 4.

2 IP2S and BIPC Problems

In this section, we mainly recall the IP2S and BIPC problems. In Subsection 2.1, we review the IP2S problem proposed by Patarin et al. in [13]. In Subsection 2.2, we describe the BIPC problem proposed by Santoso [16] as a problem originated from IP2S. Moreover, we recall the encryption scheme associated to the BIPC

[5] The earlier draft of our attack is published as a preprint in IACR Cryptology ePrint Archive [11].

problem, which is our main concern in this paper. In Subsection 2.3, we revisit the previous security analysis against the BIPC problem following the original paper [16].

2.1 IP2S problem

Let $\mathbb{F}$ be a finite field with q elements and let n and m be positive integers. We denote by $\mathbb{F}[x_1, \ldots, x_n]$ the polynomial ring in n variables over the finite field $\mathbb{F}$. We also denote by $\mathrm{GL}_n(\mathbb{F})$ the general linear group over $\mathbb{F}$ with size n. Any element of $\mathrm{GL}_n(\mathbb{F})$ can be considered as a linear map from $\mathbb{F}^n$ to $\mathbb{F}^n$. In order to describe the IP2S problem, we need the following set:

$$\mathcal{MQ}(n,m) := \left\{ \mathbf{f} = (f_1, \ldots, f_m) \middle| \begin{array}{c} f_i \in \mathbb{F}[x_1, \ldots, x_n] \ (1 \le i \le m) \\ \text{quadratic polynomial} \end{array} \right\}.$$

Namely, $\mathcal{MQ}(n,m)$ is the set of multivariate quadratic polynomial maps from $\mathbb{F}^n$ to $\mathbb{F}^m$. Any $\mathbf{f} = (f_1, \ldots, f_m) \in \mathcal{MQ}(n,m)$ is said to be *homogeneous* if all $f_1, \ldots, f_m$ are homogeneous. We define the operation of $\mathrm{GL}_n(\mathbb{F})$ and $\mathrm{GL}_m(\mathbb{F})$ to $\mathcal{MQ}(n,m)$:

$$(S,T) \cdot \mathbf{f} := T \circ \mathbf{f} \circ S, \qquad (S,T) \in \mathrm{GL}_n(\mathbb{F}) \times \mathrm{GL}_m(\mathbb{F}).$$

It is clear that for any $S \in \mathrm{GL}_n(\mathbb{F})$ and $T \in \mathrm{GL}_m(\mathbb{F})$,

$$\mathbf{f} \in \mathcal{MQ}(n,m) \implies (S,T) \cdot \mathbf{f} \in \mathcal{MQ}(n,m).$$

Namely, the operation of $\mathrm{GL}_n(\mathbb{F})$ and $\mathrm{GL}_m(\mathbb{F})$ holds the set $\mathcal{MQ}(n,m)$. Then, the IP2S problem is defined as follows:

Isomorphism of polynomials with two secrets (IP2S) [13]
Given two quadratic polynomial maps $\mathbf{f}, \mathbf{g} \in \mathcal{MQ}(n,m)$, find two linear maps $S \in \mathrm{GL}_n(\mathbb{F})$ and $T \in \mathrm{GL}_m(\mathbb{F})$ such that

$$\mathbf{g} = (S,T) \cdot \mathbf{f}. \tag{1}$$

In [13], Patarin proposed the basic idea of an authentication scheme and a signature scheme based on the IP2S problem. The concrete authentication scheme is refined in [15] and the security against quantum adversary in quantum random oracle model is proven in [17].

2.2 BIPC problem and ElGamal-like BIPC encryption scheme

In this subsection, we recall a problem originated from IP2S (called BIPC) and an ElGamal-like public key encryption scheme, which were proposed by Santoso at PQC 2020 [16].

Let $M_n(\mathbb{F})$ be the matrix ring with size $n \times n$ over $\mathbb{F}$. We also denote by $C_n(\mathbb{F})$ the subalgebra of circulant matrices in $M_n(\mathbb{F})$. Thus, any element A in $C_n(\mathbb{F})$ is written by

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{1n} & a_{11} & \dots & a_{1n-1} \\ \vdots & \vdots & \ddots & \vdots \\ a_{12} & a_{13} & \dots & a_{11} \end{pmatrix}.$$

Note that $C_n(\mathbb{F})$ is a commutative ring, that is,

$$A, B \in C_n(\mathbb{F}) \implies AB = BA.$$

To describe the BIPC problem, we need the following definition:

Definition 1. *Let k be a positive integer and let $\mathbf{f} = (\mathbf{f}_{[1]}, \dots, \mathbf{f}_{[k]}) \in \mathcal{MQ}(n,m)^k$ be a k-tuple of elements of $\mathcal{MQ}(n,m)$. For any two k-tuples $\mathcal{A} = (A_1, \dots, A_k) \in C_n(\mathbb{F})^k$ and $\mathcal{B} = (B_1, \dots, B_k) \in C_m(\mathbb{F})^k$, we define the operation*

$$\begin{aligned}(\mathcal{A}, \mathcal{B}) * \mathbf{f} &:= \begin{pmatrix} \sum_{j=1}^{k} B_j \circ \mathbf{f}_{[j \ mod \ k]} \circ A_j \\ \sum_{j=1}^{k} B_j \circ \mathbf{f}_{[j+1 \ mod \ k]} \circ A_j \\ \vdots \\ \sum_{j=1}^{k} B_j \circ \mathbf{f}_{[j+k-1 \ mod \ k]} \circ A_j \end{pmatrix}^T \\ &= \begin{pmatrix} B_1 \circ \mathbf{f}_{[1]} \circ A_1 + B_2 \circ \mathbf{f}_{[2]} \circ A_2 + \dots + B_k \circ \mathbf{f}_{[k]} \circ A_k \\ B_1 \circ \mathbf{f}_{[2]} \circ A_1 + B_2 \circ \mathbf{f}_{[3]} \circ A_2 + \dots + B_k \circ \mathbf{f}_{[1]} \circ A_k \\ \vdots \\ B_1 \circ \mathbf{f}_{[k]} \circ A_1 + B_2 \circ \mathbf{f}_{[1]} \circ A_2 + \dots + B_k \circ \mathbf{f}_{[k-1]} \circ A_k \end{pmatrix}^T\end{aligned}$$

It is clear that $(\mathcal{A}, \mathcal{B}) * \mathbf{f}$ is also an element of $\mathcal{MQ}(n,m)^k$.

Remark 1. The operation $*$ does not hold the associativity. That is, in general

$$(\mathcal{A}, \mathcal{B}) * ((\mathcal{A}', \mathcal{B}') * \mathbf{f}) \neq (\mathcal{A}\mathcal{A}', \mathcal{B}\mathcal{B}') * \mathbf{f},$$

where $\mathcal{A}\mathcal{A}' = (A_1A_1', \dots, A_kA_k')$, $\mathcal{B}\mathcal{B}' = (B_1B_1', \dots, B_kB_k')$. Thus, the operation $*$ is not an action of $C_n(\mathbb{F})^k \times C_m(\mathbb{F})^k$ to $\mathcal{MQ}(n,m)^k$.

In the same way as IP2S, the BIPC problem is defined as follows:

Block isomorphism of polynomials with circulant matrices (BIPC) [16]
Let k be a positive integer. Given two k-tuples of quadratic polynomial maps $\mathbf{f}, \mathbf{g} \in \mathcal{MQ}(n,m)^k$, find two k-tuples of circulant matrices $\mathcal{A} \in C_n(\mathbb{F})^k$ and $\mathcal{B} \in C_m(\mathbb{F})^k$ such that

$$\mathbf{g} = (\mathcal{A}, \mathcal{B}) * \mathbf{f}. \tag{2}$$

We can consider that the BIPC problem is obtained by linearizing the IP2S problem and restricting secret linear maps to linear maps represented by circulant matrices.

To construct an ElGamal-like encryption scheme based on BIPC, we need to see that the operation $*$ is commutative:

Lemma 1. *[16, Lemma1] For any* $\mathbf{f} \in \mathcal{MQ}(n,m)^k$ *and* $\mathcal{A}, \mathcal{A}' \in C_n(\mathbb{F})^k$ *and* $\mathcal{B}, \mathcal{B}' \in C_m(\mathbb{F})^k$*, we have*

$$(\mathcal{A}, \mathcal{B}) * ((\mathcal{A}', \mathcal{B}') * \mathbf{f}) = (\mathcal{A}', \mathcal{B}') * ((\mathcal{A}, \mathcal{B}) * \mathbf{f}).$$

This lemma follows from the definition of $*$ and the commutativity of products of circulant matrices.

Before we recall the construction of the ElGamal-like encryption scheme, we define a subset of $\mathcal{MQ}(n,m)$, which is useful to reduce the key size of the encryption scheme. Let ℓ be a divisor of k. We define the set

$$\mathcal{MQ}(n,m)_\ell^k := \left\{ (\mathbf{f}_{[1]}, \ldots, \mathbf{f}_{[k]}) \in \mathcal{MQ}(n,m)^k \middle| \begin{array}{l} \mathbf{f}_{[i]} = \mathbf{f}_{[i \bmod \ell]} \\ \forall i = 1, \ldots, k \end{array} \right\}.$$

In particular, if $\ell = 1$, then we have $\mathbf{f}_{[1]} = \mathbf{f}_{[2]} = \cdots = \mathbf{f}_{[k]}$. It is clear that the size of an element of $\mathcal{MQ}(n,m)_\ell^k$ is $1/\ell$ of that of an element of $\mathcal{MQ}(n,m)^k$. Moreover, we have

$$\mathbf{f} \in \mathcal{MQ}(n,m)_\ell^k \implies (\mathcal{A}, \mathcal{B}) * \mathbf{f} \in \mathcal{MQ}(n,m)_\ell^k$$

for $\mathcal{A} \in C_n(\mathbb{F})^k$ and $\mathcal{B} \in C_m(\mathbb{F})^k$. Thus, we can define the variant of the BIPC problem by replacing $\mathcal{MQ}(n,m)^k$ with $\mathcal{MQ}(n,m)_\ell^k$, which reduces the size of the instance $(\mathbf{f}, \mathbf{g})$ to $1/\ell$.

In the following, we describe the construction of the ElGamal-like BIPC encryption scheme [16] based on the hardness of the BIPC computational problem.

ElGamal-like BIPC encryption scheme [16]

- Public parameters : $n, m, k, \ell \in \mathbb{N}$.
- Secret Key: $(\mathcal{A}, \mathcal{B}) \in C_n(\mathbb{F})^k \times C_m(\mathbb{F})^k$.
- Public Key: $\mathbf{f}, \mathbf{g} \in \mathcal{MQ}(n,m)_\ell^k$ such that $\mathbf{g} = (\mathcal{A}, \mathcal{B}) * \mathbf{f}$.
- Encryption: to encrypt a plaintext $\mathbf{h} \in \mathcal{MQ}(n,m)_\ell^k$, one chooses a random $(\mathcal{A}', \mathcal{B}') \in C_n(\mathbb{F})^k \times C_m(\mathbb{F})^k$ and computes:

  $$\mathbf{c}_0 \leftarrow (\mathcal{A}', \mathcal{B}') * \mathbf{f}, \qquad \mathbf{c}_1 \leftarrow \mathbf{h} + (\mathcal{A}', \mathcal{B}') * \mathbf{g}.$$

  The ciphertext is given by $\mathbf{c} = (\mathbf{c}_0, \mathbf{c}_1)$.
- Decryption: to decrypt a ciphertext $\mathbf{c} = (\mathbf{c}_0, \mathbf{c}_1)$, using the secret key $(\mathcal{A}, \mathcal{B}) \in C_n(\mathbb{F})^k \times C_m(\mathbb{F})^k$, one computes:

  $$\nu \leftarrow \mathbf{c}_1 - (\mathcal{A}, \mathcal{B}) * \mathbf{c}_0.$$

  The decryption result is ν.

It is clear from Lemma 1 that the decryption process produces the correct plaintext, that is, $\nu = \mathbf{h}$.

In [16], it is proven that the ElGamal-like encryption scheme is proven secure against OW-CPA attacks under the assumption that the CDH-BIPC problem is hard.

Computational Diffie-Hellman (CDH) BIPC problem

Let k be a positive integer. Given three k-tuples of quadratic polynomial maps $\mathbf{f}, \mathbf{g}, \mathbf{h} \in \mathcal{MQ}(n,m)_\ell^k$ such that $\mathbf{g} = (\mathcal{A}, \mathcal{B}) * \mathbf{f}$, $\mathbf{h} = (\mathcal{A}', \mathcal{B}') * \mathbf{f}$, then compute

$$(\mathcal{A}, \mathcal{B}) * ((\mathcal{A}', \mathcal{B}') * \mathbf{f}).$$

One can actually easily see that as similar to the case of ElGamal encryption scheme and the CDH problem, the converse is also true, i.e., if the CDH-BIPC problem is easy then breaking the ElGamal-like BIPC encryption scheme is also easy.

2.3 Previous analysis

In this subsection, we recall the security analysis against the BIPC problem in the original paper [16]. In [16], two attacks were proposed under the assumption the finite field $\mathbb{F} = \mathbb{F}_2$. The first one (i) is by using the result of Bouillaguet et al. [3], and the second one (ii) is an algebraic attack using a Gröbner basis algorithm.

(i) First attack: The first attack is based on the work by Bouillaguet et al. [3] on breaking a homogeneous IP2S instance, which we summarize as follows. Given a homogeneous IP2S instance $(\mathbf{f}, \mathbf{g})$ described in (1), Bouillaguet et al. [3] attempt to find a pair of vectors α, $\beta \in \mathbb{F}^n$ such that $S^{-1}\alpha = \beta$. Under the assumption $\mathbb{F} = \mathbb{F}_2$, Bouillaguet et al. [3] showed how to obtain such α, β in high probability using a graph theory based algorithm with the complexity $\mathcal{O}(n^5 2^{n/2})$. Once such a pair α, β is discovered, we can define $(\mathbf{f}', \mathbf{g}')$, i.e., $\mathbf{f}'(x) = \mathbf{f}(x + \alpha)$ and $\mathbf{g}'(x) = \mathbf{g}(x + \beta)$, which have the same isomorphism as $(\mathbf{f}, \mathbf{g})$ but are no longer homogeneous. Thus, we can easily find the isomorphism between $\mathbf{f}'$ and $\mathbf{g}'$ using the algorithm of Faugére and Perret [9] on solving inhomogeneous instances of IP2S.

Now, we explain the first attack in [16] against a BIPC instance $(\mathbf{f}, \mathbf{g})$ described in (2). Assume that $\mathbf{g}$ is written as

$$\mathbf{g} = (\mathbf{g}_{[1]}, \ldots, \mathbf{g}_{[k]}) = (\mathcal{A}, \mathcal{B}) * \mathbf{f},$$
$$\mathbf{g}_{[i]} = \sum_{j=1}^{k} B_j \circ \mathbf{f}_{[j+i-1 \bmod k]} \circ A_j,$$

and each A_j is invertible. Then, the first attack finds vectors $\alpha_j, \beta_j \in \mathbb{F}^n$ $(1 \leq j \leq k)$ such that $A_j^{-1}\alpha_j = \beta_j$, i.e., $\alpha_j = A_j\beta_j$. In [16], it is estimated that such vectors

can be found with the complexity $\mathcal{O}(k^2n^5 2^{nk/(k+1)})$ combining Bouillaguet et al. [3] and Suzuki et al. [20]. (See [16] for the detail.)

In [16], it is described that the next step is to apply the algorithm of Faugére and Perret [9] on solving inhomogeneous instances of IP2S. However, we point out in this paper that actually such step is not necessary, since A_j is a circulant matrix. Namely, by solving the linear equation $\alpha_j = A_j\beta_j$ with respect to components of A_j, we can easily recover the circulant matrix A_j. Therefore, we conclude that the complexity of the first attack against the BIPC problem is given by

$$\mathcal{O}(k^2n^5 2^{nk/(k+1)}).$$

(ii) Second attack: Here, we review the algebraic attack using a Gröbner basis algorithm in [16]. In [16], it is assumed that there exist circulant matrices $\widetilde{B}_1, \ldots, \widetilde{B}_k$ such that the BIPC instance $(\mathbf{f}, \mathbf{g})$ satisfies the following:

$$\sum_{j=1}^{k} \widetilde{B}_j \circ \mathbf{g}_{[i]} = \sum_{j=1}^{k} \mathbf{f}_{[j+i-1 \bmod k]} \circ A_j, \quad (i = 1, \ldots, k). \tag{3}$$

Note that $B_1, \ldots, B_k$ can be computed easily once we obtain $A_j, \widetilde{B}_j$ for all $j = 1, \ldots, k$. If we identify each component of A_j, $\widetilde{B}_j$ as variables, then we obtain the system of at most $n(n+1)mk/2$ quadratic equations in $(n+m)k$ variables. In [16], it is estimated based on [1] that the complexity to solve the system with a Gröbner basis algorithm is given by

$$\mathcal{O}(2^{k\log(nm)/(4m)}).$$

However, it should be noted that we do not know exactly whether we can construct the system of quadratic equations shown in (3) since so far there is no proof for the existence of such circulant matrices $\widetilde{B}_1, \ldots, \widetilde{B}_k$.

Remark 2. In the IP2S problem (1), we have

$$\mathbf{g} = (S, T) \cdot \mathbf{f} \implies (1_n, T^{-1}) \cdot \mathbf{g} = (S, 1_m) \cdot \mathbf{f}$$

since the operation $\cdot$ holds the associativity. Similarly, if the operation $*$ satisfies the associativity, then we have

$$\mathbf{g} = (\mathcal{A}, \mathcal{B}) * \mathbf{f} \implies (1_n^k, \mathcal{B}^{-1}) * \mathbf{g} = (\mathcal{A}, 1_m^k) * \mathbf{f},$$

where we have set $\mathcal{B}^{-1} = (B_1^{-1}, \ldots, B_k^{-1})$, $1_n^k = (1_n, \ldots, 1_n)$ and $1_m^k = (1_m, \ldots, 1_m)$. Then (3) holds as $\widetilde{B}_i = B_i^{-1}$. However, as we stated in Remark 1, the operation $*$ does not hold the associativity.

Selected parameters: Based on two attacks (i) and (ii) against the BIPC problem, the paper [16] sets four types of parameters:

(a) Conservative Type: In this type, $\mathbf{f}_{[1]}, \ldots, \mathbf{f}_{[k]}$ are chosen randomly (namely $\ell = k$) and the parameters are set such that the estimated complexity of performing first attack (i) and that of second attack (ii) are both larger than the targeted complexity for security. More precisely, here $n, m, k = \ell$ are set such that the following holds:

$$k^2 n^5 2^{nk/(k+1)} \geqq 2^\lambda, \quad 2^{k \log(nm)/(4m)} \geqq 2^\lambda,$$

where λ is the targeted bit security.

(b) Alternative Type: In this type, $\mathbf{f}_{[1]}, \ldots, \mathbf{f}_{[k]}$ are chosen randomly and the parameters are set such that the complexity of performing only second attack (ii) is larger than the targeted complexity for security. More precisely, here $n, m, k = \ell$ are set such that the followings holds:

$$2^{k \log(nm)/(4m)} \geqq 2^\lambda,$$

where λ is the targeted bit security.

(c) Extremely Aggressive Type: In this type, the multivariate quadratic polynomials $\mathbf{f}_{[1]}, \ldots, \mathbf{f}_{[k]}$ are set such that $\mathbf{f}_{[1]} = \mathbf{f}_{[2]} = \cdots = \mathbf{f}_{[k]}$, i.e., $\ell = 1$. The other parameters n, m, k are set based on the conservative type.

(d) Moderately Aggressive Type: In this type, we assume that k is an even number and set $\ell = 2$. The multivariate quadratic polynomials $\mathbf{f}_{[1]}, \ldots, \mathbf{f}_{[k]}$ are set such that $\mathbf{f}_{[1]} = \mathbf{f}_{[3]} = \cdots = \mathbf{f}_{[2i-1]}$ and $\mathbf{f}_{[2]} = \mathbf{f}_{[4]} = \cdots = \mathbf{f}_{[2i]}$ hold for $i \in [1, k/2]$. The other parameters are set based on the conservative type.

Below, we summarize the recommended parameters in [16] according to each type mentioned above for 128- and 256-bit security. Here, the finite field $\mathbb{F}$ was taken as $\mathbb{F}_2$.

Table 1. 128-bit security parameters proposed in [16]

Type	n	m	k	ℓ	Public key size (KByte)	Secret key size (KByte)
(a) Conservative	84	2	140	140	241	1.4
(b) Alternative	16	2	205	205	12.8	0.45
(c) Extremely Aggressive	84	2	140	1	1.7	1.4
(d) Moderately Aggressive	84	2	140	2	3.4	1.4

3 Linear stack attack

In this section, we propose a new attack (called linear stack attack) for the ElGamal-like BIPC encryption scheme in Subsection 2.2. In Subsection 3.1, we show a key lemma to propose the linear stack attack. In Subsection 3.2, we describe the algorithm of the linear stack attack based on the key lemma. In Subsection 3.3, we estimate the complexity of the linear stack attack and show some experimental results.

Table 2. 256-bit security parameters proposed in [16]

Type	n	m	k	ℓ	Public key size (KByte)	Secret key size (KByte)
(a) Conservative	206	2	236	236	2445	5.9
(b) Alternative	16	2	410	410	25.6	0.9
(c) Extremely Aggressive	206	2	236	1	10.3	5.9
(d) Moderately Aggressive	206	2	236	2	20.7	5.9

3.1 Key lemma

In this subsection, we give a key lemma to propose the linear stack attack.

Let $(\mathbf{f}, \mathbf{g})$ be a public key of the ElGamal-like encryption scheme in Subsection 2.2. To break the encryption scheme, an attacker only has to compute a pair $(\mathcal{A}', \mathcal{B}') \in C_n(\mathbb{F})^k \times C_m(\mathbb{F})^k$ such that $\mathbf{g} = (\mathcal{A}', \mathcal{B}') * \mathbf{f}$, namely, an equivalent key. However, by the following lemma, we show that there are other kinds of equivalent keys.

Lemma 2. *Let $(\mathbf{f}, \mathbf{g})$ be an public key of the ElGamal-like encryption scheme. If there are an integer $t \in \mathbb{N}$ and a t-tuple $(\mathcal{A}_i, \mathcal{B}_i)_{i=1,\ldots,t} \in \left(C_n(\mathbb{F})^k \times C_m(\mathbb{F})^k\right)^t$ such that*

$$\mathbf{g} = \sum_{i=1}^{t} (\mathcal{A}_i, \mathcal{B}_i) * \mathbf{f},$$

then the t-tuple $(\mathcal{A}_i, \mathcal{B}_i)_{i=1,\ldots,t}$ works as an equivalent key for the public key $(\mathbf{f}, \mathbf{g})$.

Proof. Let $\mathbf{c}_0 = (\mathcal{A}', \mathcal{B}') * \mathbf{f}$ and $\mathbf{c}_1 = \mathbf{h} + (\mathcal{A}', \mathcal{B}') * \mathbf{g}$ be an ciphertext of the ElGamal-like encryption scheme as in Subsection 2.2. To recover the plaintext $\mathbf{h}$, an attacker computes secret information $(\mathcal{A}', \mathcal{B}') * \mathbf{g}$ from known information $(\mathcal{A}_i, \mathcal{B}_i)_{i=1,\ldots,t}$ and $\mathbf{c}_0$ as follows:

$$\begin{aligned}
\sum_{i=1}^{t} (\mathcal{A}_i, \mathcal{B}_i) * \mathbf{c}_0 &= \sum_{i=1}^{t} (\mathcal{A}_i, \mathcal{B}_i) * ((\mathcal{A}', \mathcal{B}') * \mathbf{f}) \\
&= \sum_{i=1}^{t} (\mathcal{A}', \mathcal{B}') * ((\mathcal{A}_i, \mathcal{B}_i) * \mathbf{f}) \\
&= (\mathcal{A}', \mathcal{B}') * \left(\sum_{i=1}^{t} (\mathcal{A}_i, \mathcal{B}_i) * \mathbf{f} \right) \\
&= (\mathcal{A}', \mathcal{B}') * \mathbf{g}
\end{aligned}$$

Thus, the attacker can compute the plaintext $\mathbf{h}$ by

$$\mathbf{c}_1 - \sum_{i=1}^{t} (\mathcal{A}_i, \mathcal{B}_i) * \mathbf{c}_0.$$

Therefore, the t-tuple $(\mathcal{A}_i, \mathcal{B}_i)_{i=1,\ldots,t}$ is an equivalent key. □

In the next subsection, we show a concrete procedure to construct such an equivalent key $(\mathcal{A}_i, \mathcal{B}_i)_{i=1,\ldots,t}$.

3.2 The algorithm of the linear stack attack

We propose an attack to find an equivalent key, which is called linear stack attack, based on Lemma 2. The attack takes a public key $(\mathbf{f}, \mathbf{g})$ and a positive integer t as input, and a t-tuple $(\mathcal{A}_i, \mathcal{B}_i)_{i=1,\ldots,t} \in \left(C_n(\mathbb{F})^k \times C_m(\mathbb{F})^k\right)^t$ as output. The strategy of the algorithm is the following:

Linear stack attack

Step 1. Randomly choose tk elements $(A_1, B_1), \cdots, (A_{tk}, B_{tk})$ from $C_n(\mathbb{F}) \times C_m(\mathbb{F})$.

Step 2. Let $\alpha_1, \ldots, \alpha_{tk}$ be variables over $\mathbb{F}$. Set t-tuples as follows:

$$\begin{array}{ll}
\mathcal{A}_1 \leftarrow (A_1, \ldots, A_k) & \mathcal{B}_1 \leftarrow (\alpha_1 B_1, \ldots, \alpha_k B_k) \\
\mathcal{A}_2 \leftarrow (A_{k+1}, \ldots, A_{2k}) & \mathcal{B}_2 \leftarrow (\alpha_{k+1} B_{k+1}, \ldots, \alpha_{2k} B_{2k}) \\
\quad\vdots & \quad\vdots \\
\mathcal{A}_t \leftarrow (A_{tk-k+1}, \ldots, A_{tk}) & \mathcal{B}_t \leftarrow (\alpha_{tk-k+1} B_{tk-k+1}, \ldots, \alpha_{tk} B_{tk})
\end{array}$$

Step 3. Find a solution to the following linear equations in variables $\alpha_1, \ldots, \alpha_{tk}$:

$$\mathbf{g} = \sum_{i=1}^{t} (\mathcal{A}_i, \mathcal{B}_i) * \mathbf{f} \tag{4}$$

Step 4. If there is a solution $(\alpha_1, \ldots, \alpha_{tk})$, then output the t-tuple

$$(\mathcal{A}_i, \mathcal{B}_i)_{i=1,\ldots,t}.$$

□

The linear stack attack is able to find an equivalent key by linearly stacking $(\mathcal{A}_i, \mathcal{B}_i) * \mathbf{f}$. In the following theorem, we discuss the probability that one trial of the linear stack attack succeeds for the input number t.

Theorem 1. *Set $t = \lceil \frac{1}{2} n(n+1) m\ell / k \rceil$ in the linear stack attack. Then one trial of the linear stack attack can find an equivalent key with the probability of at least $1/4$.*

Proof. We prove that when $t = \lceil \frac{1}{2} n(n+1) m\ell / k \rceil$, the linear equation (4) has a solution with the probability $\geq 1/4$.

The vector space $\mathcal{MQ}(n,m)$ is of dimension $\frac{1}{2}n(n+1)m$ over $\mathbb{F}$, and $\mathcal{MQ}(n,m)_\ell^k$ is the vector space with dimension $\frac{1}{2}n(n+1)m\ell$. Thus, the subspace

$$V := \mathrm{Span}_{\mathbb{F}}\left\{(\mathcal{A},\mathcal{B}) * \mathbf{f} \in \mathcal{MQ}(n,m)_\ell^k \;\middle|\; \mathcal{A} \in C_n(\mathbb{F})^k, \mathcal{B} \in C_m(\mathbb{F})^k\right\}$$

in $\mathcal{MQ}(n,m)_\ell^k$ is at most of dimension $\frac{1}{2}n(n+1)m\ell$. The linear equation (4)

$$\mathbf{g} = \sum_{i=1}^{t}(\mathcal{A}_i,\mathcal{B}_i) * \mathbf{f}$$

in Step 3 has tk variables and $\dim_k V$ equations. We can consider that the probability that the linear equation (4) has a solution is equal to the probability that a uniformly chosen random matrix over $\mathbb{F}$ with size $tk \times \dim_k V$ has rank $\geq \dim_k V$. We conservatively assume that $\dim_k V = \frac{1}{2}n(n+1)m\ell$. Since we have $tk \fallingdotseq \frac{1}{2}n(n+1)m\ell$, such a probability is given by the probability that a uniformly chosen random square matrix with size tk is nonsingular, which is $\prod_{i=1}^{tk}(1-\frac{1}{q^i})$. Here,

$$\prod_{i=1}^{tk}(1-\tfrac{1}{q^i}) = (1-\tfrac{1}{q})^2(1+\tfrac{1}{q})\prod_{i=3}^{tk}(1-\tfrac{1}{q^i}) \geq (1-\tfrac{1}{q})^2(1+\tfrac{1}{q^{tk-2}}).$$

The latter inequality follows from $(1+\frac{1}{q^j})(1-\frac{1}{q^i}) \geq (1+\frac{1}{q^{j+1}})$ for $j<i$. Since $q \geq 2$, we have $\prod_{i=1}^{tk}(1-\frac{1}{q^i}) \geq (1-\frac{1}{q})^2 \geq 1/4$. □

If a trial of the linear stack attack fails for a parameter, then we retry the linear stack attack for the parameter. The above theorem implies that the linear stack attack succeed in around four trials. Note that according to our experiments below, the linear stack attack alway succeeded in one trail as $t = \lceil \frac{1}{2}n(n+1)m\ell/k \rceil$.

Remark 3. If we take $t > \lceil \frac{1}{2}n(n+1)m\ell/k \rceil$, the probability that a uniformly chosen random matrix over $\mathbb{F}$ with size $tk \times \dim_k V$ has rank $\geq \dim_k V$ can be raised. As a result, the success probability of one trial of the linear stack attack can be raised. On the other hand, since $\dim_k V$ is less than $\frac{1}{2}n(n+1)m\ell$ in general, the probability should be larger than 1/4 even if we set $t = \lceil \frac{1}{2}n(n+1)m\ell/k \rceil$. In fact, according to our experiments below, the linear stack attack *alway* succeeded in one trial as $t = \lceil \frac{1}{2}n(n+1)m\ell/k \rceil$ on each experiment.

Remark 4. It should be noted that the linear stack attack does not break the BIPC problem. However, as we explained in Subsection 2.2, in order to break the ElGamal-like encryption scheme, it is sufficient for an attacker to break the CDH-BIPC problem, that the linear stack attack actually does efficiently.

3.3 Complexity and Experimental results

In this subsection, we estimate the complexity of our attack proposed in Subsection 3.1 and show some experimental results.

Proposition 1. *The complexity of the linear stack attack is given by at most*

$$\mathcal{O}(n^6 m^3 \ell^3).$$

Proof. It is clear that the dominant part is Step 3. In Step 3, we need to compute tk composites $B_i \circ \mathbf{f}_{[j]} \circ A_i$ $(i = 1, \ldots, tk, j = 1, \ldots, \ell)$. The number of multiplications of $\mathbb{F}$ in each composite is at most

$$2n^3 + \tfrac{1}{2}n(n+1)m^2.$$

Since $tk \coloneqq \frac{1}{2}n(n+1)m\ell$, the complexity is

$$\mathcal{O}\left((2n^3 + \tfrac{1}{2}n(n+1)m^2) \cdot (\tfrac{1}{2}n(n+1)m\ell)\right) \leq \mathcal{O}(n^5 m^3 \ell).$$

In Step 3, we also solve the linear system with size $tk \coloneqq \frac{1}{2}n(n+1)m\ell$. Then the complexity is

$$\mathcal{O}((\tfrac{1}{2}n(n+1)m\ell)^{\omega}) \leq \mathcal{O}(n^6 m^3 \ell^3),$$

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

As a result, we conclude that the total complexity of the linear stack attack is at most

$$\mathcal{O}(n^6 m^3 \ell^3).$$

□

This proposition indicates that our attack is a polynomial-time algorithm. In Tables 3 and 4, we show the complexity of the proposed parameters in Tables 1 and 2 against our linear stack attack. Each complexity is given by the product of $n^6 m^3 \ell^3$ and the inverse of success probability $(1/4)^{-1}$.

Table 3. The complexity of the linear stack attack in Section 3 for the 128-bit security parameters proposed in [16]

Type	n	m	k	ℓ	linear stack attack (bits)
(a) Conservative	84	2	140	140	**64.7**
(b) Alternative	16	2	205	205	**54.0**
(c) Extremely Aggressive	84	2	140	1	**43.3**
(d) Moderately Aggressive	84	2	140	2	**46.3**

Experimental results

We confirm experimentally that our attack is valid and efficient enough to break the ElGamal-like encryption scheme. All experiments were performed on a 3.5 GHz 8 Core Intel Xeon W with Magma V2.25-7. Table 5 is the experimental results. We basically performed our attack on the 128-bit security parameters in Table 1. However, the attack against the conservative type (a) did

Table 4. The complexity of the linear stack attack in Section 3 for the 256-bit security parameters proposed in [16]

Type	n	m	k	ℓ	linear stack attack (bits)
(a) Conservative	206	2	236	236	**74.7**
(b) Alternative	16	2	410	410	**55.0**
(c) Extremely Aggressive	206	2	236	1	**51.1**
(d) Moderately Aggressive	206	2	236	2	**54.1**

not finish within a week, since its complexity against the linear stack attack is around 64.7 bits. Instead, we chose the 80-bit security parameter $(q, n, m, k, l) = (2, 42, 2, 102, 102)$ following the security analysis in [16], and confirmed that our attack is valid for such a conservative type (a). We note that the complexity of the linear stack attack against the parameter $(q, n, m, k, l) = (2, 42, 2, 102, 102)$ is 57.4 bits.

Table 5 shows the average time of 5 experiments on the linear stack attack for each type as $t = \lceil \frac{1}{2}n(n+1)m\ell/k \rceil$. According to our experiments, the success probability of one trial of the linear stack attack is 100% for each parameter.

Table 5. The average time (seconds) of 5 experiments on the linear stack attack for the 80-bit security parameter in (a) and the 128-bit security parameters [16] (b),(c),(d) in Table 1.

Type	n	m	k	ℓ	Average time (sec)
(a) Conservative	42	2	102	102	419408.490
(b) Alternative	16	2	205	205	35963.190
(c) Extremely Aggressive	84	2	140	1	1732.240
(d) Moderately Aggressive	84	2	140	2	6801.210

4 Conclusion

At PQC 2020, Santoso proposed an ElGamal-like public key encryption scheme based on the BIPC problem, which is a problem originated from the IP2S problem. The BIPC problem is obtained by linearizing IP2S and restricting secret linear maps to linear maps represented by circulant matrices. Moreover, the ElGamal-like encryption scheme was constructed by utilizing the commutativity of products of circulant matrices. Santoso gave four types of practical parameters which are called (a) conservative, (b) alternative, (c) extremely aggressive, and (d) moderately aggressive. In this paper, we proposed a new attack against the ElGamal-like encryption scheme, which is called linear stack attack. The linear stack attack finds equivalent keys of the ElGamal-like encryption scheme by using the linearity of the BIPC problem. We showed that the linear stack attack

is polynomial time and confirmed that the attack is valid and efficient for the proposed parameters (a),(b),(c) and (d). Our experimental results showed that the 128-bit security parameter in (b),(c) and (d) were broken within 10 hours. While the 128-bit security parameter in (a) did not finish within a week, instead, the 80-bit security parameter in (a) was broken within 5 days.

We stress that our attack algorithm consists of only basic simple linear algebraic techniques. Therefore, there is a possibility that we can further generalize our algorithm to solve more general computational problems beyond BIPC related ones. As future work, we aim to extend our attack algorithm to solve more general computational problems such as the general IP2S problem which is the foundation of many multivariate cryptographic schemes.

Acknowledgements This work was supported by JST CREST Grant Number JPMJCR14D6, JSPS KAKENHI Grant Number JP19K20266, JP20K19802, JP20K03741, JP18H01438, and JP18K11292

References

1. Magali Bardet, Jean-Charles Faugére, and Bruno Salvy. Complexity of gröbner basis computation for semi-regular overdetermined sequences over $\mathbb{F}_2$ with solutions in $\mathbb{F}_2$. techreport 5049, Institut National de Recherche en Informatique et en Automatique (INRIA), 2003.
2. Daniel J. Bernstein, Johannes Buchmann, and Erik Dahmen (Eds.). *Post-Quantum Cryptography*. Springer, 2009.
3. Charles Bouillaguet, Jean-Charles Faugère, Pierre-Alain Fouque, and Ludovic Perret. Isomorphism of polynomials: New results.
4. A. Casanova, J.-C. Faugere, G. Macario-Rat, J. Patarin, L. Perret, and J. Ryckeghem. Gemss, technical report, national institute of standards and technology, https://csrc.nist.gov/projects/post-quantum-cryptography/round-2-submissions. 2019.
5. Jaihui Chen, Chik How Tan, and Xiaoyu Li. Practical cryptanalysis of a public key cryptosystem based on the morphism of polynomials problem. *Tsinghua Science and Technology*, 23(6):671–679, December 2018.
6. Jintai Ding, Ming-Shing Chen, Albrecht Petzoldt, Dieter Schmidt, and Bo-Yin Yang. Rainbow, technical report, national institute of standards and technology, https://csrc.nist.gov/projects/post-quantum-cryptography/round-3-submissions. 2020.
7. Jintai Ding, Albrecht Petzoldt, and Dieter S. Schmidt. *Multivariate Public Key Cryptosystems, Second Edition*, volume 80 of *Advances in Information Security*. Springer, 2020.
8. Jintai Ding, Ming shing Chen, Albrecht Petzoldt, Dieter Schmidt, and Bo-Yin Yang. Rainbow, technical report, national institute of standards and technology, https://csrc.nist.gov/projects/post-quantum-cryptography/round-2-submissions. 2019.
9. Jean-Charles Faugère and Ludovic Perret. Polynomial equivalence problems: Algorithmic and theoretical aspects. In *EUROCRYPT*, volume 4004 of *Lecture Notes in Computer Science*, pages 30–47. Springer, 2006.

10. Yasufumi Hashimoto. Solving the problem of blockwise isomorphism of polynomials with circulant matrices. *IACR Cryptol. ePrint Arch.*, 2021:385, 2021.
11. Yasuhiko Ikematsu, Shuhei Nakamura, Bagus Santoso, and Takanori Yasuda. Security analysis on an el-gamal-like multivariate encryption scheme based on isomorphism of polynomials. *IACR Cryptol. ePrint Arch.*, 2021:169, 2021.
12. National Institute of Standards and Technology. Report on post quantum cryptography. nistir draft 8105, https://csrc.nist.gov/csrc/media/publications/nistir/8105/final/documents/nistir_8105_draft.pdf. 2019.
13. Jacques Patarin. Hidden Fields Equations (HFE) and Isomorphisms of Polynomials (IP): Two New Families of Asymmetric Algorithms. In *EUROCRYPT '96*, volume 1070 of *LNCS*, pages 33–48, May 1996.
14. Jacques Patarin, Louis Goubin, and Nicolas Courtois. Improved algorithms for isomorphisms of polynomials. In *EUROCRYPT '98*, volume 1403 of *LNCS*, pages 184–200, 1998.
15. Bagus Santoso. Reviving identification scheme based on isomorphism of polynomials with two secrets: a refined theoretical and practical analysis. *IEICE Transactions*, 101-A(5):787–798, 2018.
16. Bagus Santoso. Generalization of isomorphism of polynomials with two secrets and its application to public key encryption. In Jintai Ding and Jean-Pierre Tillich, editors, *Post-Quantum Cryptography - 11th International Conference, PQCrypto 2020, Paris, France, April 15-17, 2020, Proceedings*, volume 12100 of *Lecture Notes in Computer Science*, pages 340–359. Springer, 2020.
17. Bagus Santoso and Chunhua Su. Provable secure post-quantum signature scheme based on isomorphism of polynomials in quantum random oracle model. In *ProvSec*, volume 10592 of *Lecture Notes in Computer Science*, pages 271–284. Springer, 2017.
18. Simona Samardjiska Ming-Shing Chen Andreas Hulsing Joost Rijneveld Peter Schwabe. Mqdss, technical report, national institute of standards and technology, https://csrc.nist.gov/projects/post-quantum-cryptography/round-2-submissions. 2019.
19. Peter W. Shor. Polynomial-time algorithms for prime factorization and discrete logarithms on a quantum computer. *SIAM J. Comput.*, 26(5):1484–1509, 1997.
20. Kazuhiro Suzuki, Dongvu Tonien, Kaoru Kurosawa, and Koji Toyota. Birthday paradox for multi-collisions. In Min Surp Rhee and Byoungcheon Lee, editors, *Information Security and Cryptology – ICISC 2006*, pages 29–40, Berlin, Heidelberg, 2006. Springer Berlin Heidelberg.
21. H. Wang, H. Zhang, Shaowu Mao, Wanqing Wu, and Liqiang Zhang. New public-key cryptosystem based on the morphism of polynomials problem. *Tsinghua Science and Technology*, 21(3):302–311, June 2016.