

A Fully Homomorphic Cryptosystem with Approximate Perfect Secrecy

Michal Hojsík* and Veronika Půlpánová

Department of Algebra, Faculty of Mathematics and Physics
Charles University in Prague, Czech Republic

Abstract. We propose a new fully homomorphic cryptosystem called *Symmetric Polly Cracker* (SymPC) and we prove its security in the information theoretical settings. Namely, we prove that SymPC approaches perfect secrecy in bounded CPA model as its security parameter grows (which we call approximate perfect secrecy).

In our construction, we use a Gröbner basis to generate a polynomial factor ring of ciphertexts and use the underlying field as the plaintext space. The Gröbner basis equips the ciphertext factor ring with a multiplicative structure that is easily algorithmized, thus providing an environment for a fully homomorphic cryptosystem.

Keywords: Polly Cracker, Fully homomorphic encryption, Gröbner bases.

1 Introduction

In 1994 Fellows and Koblitz presented a general outline for a construction of a public-key cryptosystems based on NP-hard problems in [1]. As an example, they described a cryptosystem based on the ideal membership problem and named it Polly Cracker. A whole family of cryptosystems based on this construction has been developed over the following years ([2],[3]). Polly Cracker has also played a critical role in the development of homomorphic encryption theory, mostly serving as a base stone on which more sophisticated systems were built. For instance, Craig Gentry’s seminal work on fully homomorphic encryption system [4] was inspired by Polly Cracker. Ever since Gentry’s paper has been published, there has been an extensive research in the area, e.g. [5], [6]. Majority of the schemes that followed the outbreak of fully homomorphic encryption have its security based on problems over lattices, such as Learning with Errors (LWE) [7] and most of the research focuses on the public key encryption.

In 2011, Albrecht et al. published a paper “Polly Cracker Revisited” [8]. It formally treats the security of certain classes of Polly Cracker-based cryptosystems and suggests particular transitions between public-key and symmetric versions of Polly Cracker-based systems. In the same paper, Albrecht et al. introduce the Polly Cracker with Noise (CPN) cryptosystem. Only recently, Herold has shown

* The author was supported by grant VF20102015006.

in [9] that the CPN with zero-degree noise from [8] is either insecure or does not offer any security benefit compared to Regev's LWE-based scheme [7].

Our Contribution. In our work, we take a different approach. We propose a new fully homomorphic cryptosystem called *Symmetric Polly Cracker* (SymPC) and we prove its security in information theoretical settings - we prove that SymPC approaches perfect secrecy in bounded CPA model as its security parameter grows. More precisely, we define *approximate perfect secrecy* as the security of a cryptosystem $CS(t) = (\mathcal{P}, \mathcal{C}, \mathcal{K}, \mathcal{E}, \mathcal{D})$ with security parameter t for which the probability $\Pr[P = p \mid C = c]$ approaches $\Pr[P = p]$ for all $p \in \mathcal{P}$, all probability distributions on \mathcal{P} and almost all c as t grows to infinity. Then we prove that SymPC has approximate perfect secrecy in bounded CPA model.

In our construction, unlike in the previous classical Polly Cracker constructions, we use a Gröbner basis G to generate a zero-dimensional ideal $\langle G \rangle$ of a polynomial ring $\mathbb{F}[x_1, \dots, x_n]$ over a finite field \mathbb{F} . Then we use the factor ring $\mathbb{F}[x_1, \dots, x_n]/\langle G \rangle$ as the ciphertext space and the field \mathbb{F} as the plaintext space. The Gröbner basis G equips the ciphertext factor ring with a multiplicative structure that is easily algorithmized, thus providing an environment for a fully homomorphic cryptosystem. The fully homomorphic property of our cryptosystem is achieved by a simple decryption operation - evaluation homomorphism.

This paper is organized as follows. In Sect. 2 we introduce our notation and state some known facts. In Sect. 3 we describe one instantiation of Polly Cracker cryptosystem. Then we describe our cryptosystem SymPC in Sect. 4 where we also prove that it is fully homomorphic. This is followed by the complexity analysis in Sect. 5. Finally, in Sect. 6, we define the approximate perfect secrecy, give the security proof of SymPC in bounded CPA model and briefly analyze SymPC in other attack scenarios.

2 Preliminaries and Notation

Let q be a prime power. By \mathbb{F} we will denote the finite field $\text{GF}(q)$. In this paper, we will work with the multivariate polynomial ring $\mathcal{R} = \mathbb{F}[x_1, \dots, x_n]$, $n \in \mathbb{N}$ and operations $+$, $-$ and \cdot on polynomials will always denote operations in \mathcal{R} . Later on, we will define a factor ring $\mathcal{C} = \mathcal{R}/I$ for an ideal I . We will denote the operations in this factor ring as $+_{\mathcal{C}}$, $-_{\mathcal{C}}$ and $\cdot_{\mathcal{C}}$. Furthermore, we endow \mathcal{R} with an admissible monomial ordering $<$. For $f \in \mathcal{R}$, $\text{deg}(f)$ will denote the total degree of f , i.e. degree of the leading term of f with respect to $<$. The maximum degree of variable x_i in any term of f will be denoted $\text{deg}_{x_i}(f)$.

Let G be a basis of an ideal I in \mathcal{R} , i.e. $\langle G \rangle = I$. Recall that G is a Gröbner basis, iff for all $f \in \mathcal{R}$, the remainder on division of f by G is unique. For $f, g \in \mathcal{R}$ define the s-polynomial as $\text{spol}(f, g) = \text{lcm}(\text{lt}(f), \text{lt}(g)) \cdot f/\text{lm}(f) - \text{lcm}(\text{lt}(f), \text{lt}(g)) \cdot g/\text{lm}(g)$, where $\text{lt}(f)$ denotes the leading term of f and $\text{lm}(f)$ the leading monomial of f with respect to $<$. The following theorem is employed in Buchberger's algorithm and we will use it to prove that a given set is a Gröbner basis. The proof can be found in [10].

Theorem 1. $G \subset \mathcal{R}$ is a Gröbner basis of an ideal $I = \langle G \rangle$, iff the remainder on division of $\text{spol}(f, g)$ by G equals zero for all $f, g \in G, f \neq g$.

A Gröbner basis G is called reduced, iff for all $g \in G$ it holds $g \bmod G \setminus \{g\} = g$ and it is called normed, iff all $g \in G$ are monic. A well known theorem states, that for every ideal I in \mathcal{R} , there exists a unique normed reduced Gröbner basis G of I .

In descriptions of algorithms, we will use $x \stackrel{R}{\leftarrow} X$ to denote that x is chosen uniformly at random from a finite set X .

In this paper, we propose a fully homomorphic probabilistic cryptosystem. By fully homomorphic we mean the usual concept where both the plaintext and the ciphertext sets are equipped with addition and multiplication, they both form rings and the decryption operation is a ring homomorphism (i.e. $d_k(f(c_1, \dots, c_l)) = f(d_k(c_1), \dots, d_k(c_l))$ for any polynomial f):

Definition 2. Let $(\mathcal{P}, \mathcal{C}, \mathcal{K}, \{e_k\}, \{d_k\})$ be a probabilistic cryptosystem, where $\mathcal{P}(+, -, \cdot, 0, 1)$ is the plaintext ring and $\mathcal{C}(+, -, \cdot, 0, 1)$ is the ciphertext ring. We call the cryptosystem fully homomorphic, iff for all $k \in \mathcal{K}$, the decryption operation $d_k : \mathcal{C} \rightarrow \mathcal{P}$ is a ring homomorphism.

3 Polly Cracker

In this section, we will describe one instantiation of Polly Cracker. This scheme has inspired our cryptosystem, which we present in the next section. We denote $\mathcal{S} = \langle x_1^q - x_1, \dots, x_n^q - x_n \rangle$.

Algorithms 1, 2 and 3 describe the Polly Cracker cryptosystem. The set of messages is $\mathcal{P} = \mathbb{F}$, the set of ciphertexts is $\mathcal{C} = \mathcal{R}/\mathcal{S}$ and the keys $K \in \mathcal{K}$ are pairs (\mathbf{s}, PK) , where the secret key \mathbf{s} is a vector in \mathbb{F}^n and the public key $PK = \{f_1, \dots, f_k\}$ is a set of polynomials in \mathcal{R}/\mathcal{S} of degree at most ν , such that in \mathbf{s} they all evaluate to zero, as described in the SETUP by Algorithm 1.

Algorithm 1. Polly Cracker: SETUP

Input: $n, k, q, \nu \in \mathbb{N}, q$ prime power, $\nu < q - 1$
Output: $(\mathbf{s}, PK), \mathbf{s} \in \mathbb{F}^n, PK \subset \mathcal{R}$

- 1 set $\mathbb{F} := \mathbb{F}_q$
- 2 set $\mathcal{R} := \mathbb{F}[x_1, \dots, x_n]$
- 3 set the secret key $\mathbf{s} = (s_1, \dots, s_n) \stackrel{R}{\leftarrow} \mathbb{F}^n$
- 4 **for** $j = 1$ **to** k **do**
- 5 $f_j \stackrel{R}{\leftarrow} \mathcal{R}$ s.t. $\forall i \text{ deg}_{x_i}(f_j) \leq \nu$ and $f_j(\mathbf{s}) = 0$
- 6 set the public key $PK := \{f_1, \dots, f_k\}$
- 7 set $\mathcal{C} := \mathcal{R}/\mathcal{S}$
- 8 **return** (\mathbf{s}, PK)

Algorithm 2 describes encryption. A random subset of the polynomials from PK is added to a message $m \in \mathbb{F}$ to get a ciphertext polynomial $c \in \mathcal{R}/\mathcal{S}$.

Algorithm 2. Polly Cracker: ENCRYPT

Input: message $m \in \mathbb{F}$, public key $PK = \{f_1, \dots, f_k\} \subset \mathcal{R}$

Output: ciphertext $c \in \mathcal{R}$

- 1 select $I \subseteq \{1, \dots, k\}$ uniformly at random
 - 2 set the ciphertext $c := m + \sum_{j \in I} f_j \in \mathcal{R}$
 - 3 **return** c
-

The decryption is given by Alg. 3. It evaluates the ciphertext polynomial c in the secret key \mathbf{s} . It is easy to see, that if $c = e_{PK}(m)$, then $c(\mathbf{s}) = m$, as $f(\mathbf{s}) = 0$ for all $f \in PK$.

Algorithm 3. Polly Cracker: DECRYPT

Input: ciphertext $c \in \mathcal{R}$, secret key $\mathbf{s} \in \mathbb{F}^n$

Output: message $m \in \mathbb{F}$

- 1 set $m := c(\mathbf{s})$
 - 2 **return** m
-

Using the Fundamental theorem on homomorphism and the fact that evaluation of polynomials is a ring homomorphism, one can show that decryption operation is a ring homomorphism on \mathcal{R}/\mathcal{S} . Hence Polly Cracker is a fully homomorphic cryptosystem. A disadvantage is, that the size of a ciphertext grows rapidly with the number of multiplications, which is not practical. However, we work with the ring \mathcal{R}/\mathcal{S} , which is finite, so after about $\frac{q}{\nu}$ multiplications the resulting ciphertexts stop growing.

Nevertheless, the size of a random polynomial in \mathcal{R}/\mathcal{S} is $O(q^n)$ bits. (The size of a random polynomial in \mathcal{R}/\mathcal{S} is $\log_2(|\mathcal{R}/\mathcal{S}|) = \log_2(q^{q^n})$.) Hence one needs to keep the number of variables very low in order to get a reasonable ciphertext size.

Unfortunately, the Polly Cracker cryptosystem can be attacked by calculating the Gröbner basis of the ideal generated by PK . If an adversary has a set $\{g_1, \dots, g_l\}$, the Gröbner basis of $\langle PK \rangle$, then for any $c \in \mathcal{C}$ he can calculate $c \bmod \{g_1, \dots, g_l\}$ and as a result he will get $d_{\mathbf{s}}(c) = c(\mathbf{s})$, i.e. the plaintext.

4 Symmetric Polly Cracker (SymPC)

In this section, we propose a new fully homomorphic probabilistic symmetric cryptosystem called *Symmetric Polly Cracker - SymPC*.

The cryptosystem SymPC is described by Algorithms 4, 5, 6, 7 and 8. Algorithm 4 describes SETUP, which takes security parameters n, q, ν and returns a

pair (\mathbf{s}, G) , where $\mathbf{s} \in \mathbb{F}^n$ is the secret key and $G = \{g_1, \dots, g_n\} \subset \mathcal{R}$ is the multiplication key. This is a special kind of key, that is only used in the multiplication of ciphertexts. It provides information about the ring of ciphertexts $\mathcal{R}/\langle G \rangle$ and we will assume that it is public. Furthermore **SETUP** defines the set (field) of plaintexts as \mathbb{F} and the set (ring) of ciphertexts \mathcal{C} as the factor ring $\mathcal{R}/\langle G \rangle$. The choice of polynomials g_i in Step 8 has some important consequences. First, as we shall see in Sect. 6, this choice maximizes the size of $V(G)$, the algebraic set of G . This leads to optimal security for a given security parameter. Second, G is the reduced normed Gröbner basis of $\langle G \rangle$ (the proof can be found in Appendix A):

Theorem 3. *Let \mathcal{R} and G be defined by Alg. 4. Then G is the reduced normed Gröbner basis of the ideal $\langle G \rangle$.*

Algorithm 4. **SETUP**

Input: $n, \nu, q \in \mathbb{N}, \nu < q - 1, q$ a prime power
Output: $\mathbf{s} \in \mathbb{F}^n, G \subset \mathcal{R}$

```

1 set  $\mathbb{F} := \mathbb{F}_q$ 
2 set  $\mathcal{R} := \mathbb{F}[x_1, \dots, x_n]$ 
3 set the secret key  $\mathbf{s} := (s_1, \dots, s_n) \xleftarrow{R} \mathbb{F}^n$ 
4 for  $i = 1$  to  $n$  do
5   for  $l = 1$  to  $\nu$  do
6      $t_l^{(i)} \xleftarrow{R} \mathbb{F} \setminus \{t_1^{(i)}, \dots, t_{l-1}^{(i)}\}$ 
7   for  $i = 1$  to  $n$  do
8     set  $g_i := (x_i - s_i) \cdot \prod_{l=1}^{\nu} (x_i - t_l^{(i)})$ 
9   set the multiplication key  $G := \{g_1, \dots, g_n\}$ 
10 set  $\mathcal{C} := \mathcal{R}/\langle G \rangle$ 
11 return  $(\mathbf{s}, G)$ 

```

Finally, the special choice of polynomials g_i allows us to use the set $\{f \in \mathcal{R} \mid \deg_{x_i}(f) \leq \nu, i = 1, \dots, n\}$ as the support set of \mathcal{C} .

Algorithm 5 describes the encryption procedure. In Step 1, we choose a polynomial $f \in \mathcal{R}$ uniformly at random, s.t. $\deg_{x_i}(f) \leq \nu$ for all i , hence $f \in \mathcal{C}$ and also $c \in \mathcal{C}$. Note, that according to our notation the operations used in Step 2 are the operations in \mathcal{R} and not in \mathcal{C} . We will comment on this later on.

Decryption is described by Alg. 6. Let \mathbf{s} be a secret key and $m \in \mathbb{F}$ a message. Then, by Step 2 of Alg. 5, $e_{\mathbf{s}}(m) = c = f - f(\mathbf{s}) + m$ for some random f and $d_{\mathbf{s}}(e_{\mathbf{s}}(m)) = d_{\mathbf{s}}(c) = c(\mathbf{s}) = f(\mathbf{s}) - f(\mathbf{s}) + m = m$.

Algorithm 7 describes the addition operation $+_{\mathcal{C}}$ in \mathcal{C} . From the definition of polynomials g_i in Alg. 4 it follows, that the addition in the factor ring $\mathcal{C} = \mathcal{R}/\langle G \rangle$ is the same as the addition in the polynomial ring \mathcal{R} used in Step 1 of Alg. 7. This also clarifies the operations used in Step 2 of Alg. 5.

Algorithm 5. ENCRYPT

Input: message $m \in \mathbb{F}$, secret key $\mathbf{s} \in \mathbb{F}^n$ **Output:** ciphertext $c \in \mathcal{C}$

- 1 $f \xleftarrow{\mathcal{R}} \{h \in \mathcal{R} \mid \deg_{x_i}(h) \leq \nu, \forall i = 1, \dots, n\}$
 - 2 set $c := f - f(\mathbf{s}) + m \in \mathcal{C}$
 - 3 **return** c
-

Algorithm 6. DECRYPT

Input: ciphertext $c \in \mathcal{C}$, secret key $\mathbf{s} \in \mathbb{F}^n$ **Output:** message $m \in \mathbb{F}$

- 1 set $m := c(\mathbf{s})$
 - 2 **return** m
-

Finally, Alg. 8 describes multiplication in \mathcal{C} . It follows from Theorem 3, that for $c_1, c_2 \in \mathcal{C}$, $c_1 \cdot c_2 \bmod G$ is uniquely determined.

Algorithm 7. ADD, $+_{\mathcal{C}}$

Input: ciphertexts $c_1, c_2 \in \mathcal{C}$ **Output:** ciphertext $c \in \mathcal{C}$

- 1 set $c := c_1 + c_2$
 - 2 **return** c
-

From the random choice of f in Step 1 of Alg. 5 it follows, that SymPC is a probabilistic cryptosystem. Now we prove that it is fully homomorphic.

Theorem 4. *The cryptosystem SymPC is fully homomorphic.*

Proof. Let \mathbf{s} be a secret key. Let $\varphi : \mathcal{R} \rightarrow \mathbb{F}$ be the evaluation homomorphism defined as $\varphi(f) = f(\mathbf{s})$. By definition, $\text{Ker}(\varphi) = \{f \mid f(\mathbf{s}) = 0\}$. Since $g_i(\mathbf{s}) = 0$ for all $i = 1, \dots, n$, we get that $\langle G \rangle \subseteq \text{Ker}(\varphi)$. By the Fundamental theorem on homomorphisms, $d_{\mathbf{s}} : \mathcal{C} \rightarrow \mathbb{F}$, $d_{\mathbf{s}}(c) = c(\mathbf{s})$ is a ring homomorphism. \square

From now on, we will use $\text{SymPC}(n, \nu, q)$ to denote the cryptosystem SymPC with security parameters n, ν, q .

Algorithm 8. MULTIPLY, $\cdot_{\mathcal{C}}$

Input: ciphertexts $c_1, c_2 \in \mathcal{C}$, multiplication key G **Output:** ciphertext $c \in \mathcal{C}$

- 1 set $c := c_1 \cdot c_2 \bmod G$
 - 2 **return** c
-

5 Complexity

We evaluate the complexity of each function of $\text{SymPC}(n, \nu, q)$. We denote α the number of terms in \mathcal{C} , $\alpha = (\nu + 1)^n$.

SETUP The most complex operation is generation of the polynomials g_i . For each g_i we need to perform ν multiplications of a polynomial of degree one with a polynomial of degree at most ν in $\mathbb{F}[x_i]$. The complexity is $O(n \cdot \nu^2)$ operations in \mathbb{F} .

ENCRYPT The most complex operation is the evaluation of f in $\mathbf{s} \in \mathbb{F}^n$. The algorithm performs $\log_2 q \cdot \alpha$ assignments of random bits to coefficients of f and α evaluations of monomials in \mathcal{C} . Each of these evaluations consists of at most $\deg(f) \leq \nu \cdot n$ multiplications in \mathbb{F} . Then it adds evaluations in the monomials. The overall complexity is $O(n \cdot (\nu + 1)^{n+1})$ operations in \mathbb{F} . We calculated the complexity of a naive evaluation algorithm. We can see, that the complexity of this algorithm could be optimized by the use of sparse polynomials. We will comment on that later.

DECRYPT The complexity is the same as the complexity of **ENCRYPT**, that is $O(n \cdot (\nu + 1)^{n+1})$ operations in \mathbb{F} .

ADD The function performs α additions in \mathbb{F} , so the complexity is $O((\nu + 1)^n)$ operations in \mathbb{F} .

MULTIPLY The function consists of two parts: multiplication and reduction. The first part is more complex and involves α^2 multiplications in \mathbb{F} . The overall complexity is $O((\nu + 1)^{2n})$ operations in \mathbb{F} .

6 Security

We start the section with a few simple observations.

Proposition 5. *The cryptosystem $\text{SymPC}(n, \nu, q)$ is not CCA secure.*

Proof. If an attacker can use the SymPC decryption oracle, he can ask for the decryption of the ciphertexts $c_1 = x_1, c_2 = x_2, \dots, c_n = x_n$ and he will obtain the points of the secret key s_1, s_2, \dots, s_n . \square

Proposition 6. *For the $\text{SymPC}(n, \nu, q)$ cryptosystem, the CPA security is equivalent to the KPA security.*

Proof. CPA-security implies KPA-security in general. To prove the other implication we need to realize, that if an attacker has a known plaintext-ciphertext pair (m, c) , he can get a valid plaintext-ciphertext pair (m', c') for any m' by setting $c' = c - m + m'$, as $d_{\mathbf{s}}(c') = c'(\mathbf{s}) = c(\mathbf{s}) - m + m' = m'$. Hence, from any known plaintext-ciphertext pair, he can devise a chosen plaintext-ciphertext pair, so SymPC needs to be CPA-secure to achieve the KPA-security. \square

6.1 Approximate Perfect Secrecy in Bounded CPA Model

In this section, we will prove that SymPC has approximate perfect secrecy (Definition 7) in the so-called k -bounded chosen plaintext attack (k -bounded CPA) model. In k -bounded CPA, an attacker can obtain at most k plaintext-ciphertext pairs for some given $k \in \mathbb{N}$. This can be ensured by allowing at most k plaintexts to be encrypted with a single key. As we will see, this limitation corresponds to the limitation that the size of the keyspace has to be larger or equal to the size of the plaintext space in order to reach perfect secrecy.

Similarly to perfect secrecy, we also assume that the attacker has unbounded computational power.

Definition 7. Let $CS(t) = (\mathcal{P}, \mathcal{C}, \mathcal{K}, \{e_k\}_k, \{d_k\}_k)$ be a cryptosystem with a security parameter t and P, C random variables on \mathcal{P}, \mathcal{C} . Let $\mathcal{F} = \{f : \mathcal{C} \rightarrow \mathbb{R}\}$ be the set of all functions from \mathcal{C} to \mathbb{R} and let $\delta : \mathcal{F} \times \mathcal{F} \rightarrow \mathbb{R}$ be a metric (distance) on \mathcal{F} . For $m \in \mathcal{P}$, $\Pr[P = m | C = c]$, $\Pr[P = m] \in \mathcal{F}$ (the later one is a constant function in c). We say, that $CS(t)$ has **approximate perfect secrecy**, iff for all probability distributions on \mathcal{P} and all $m \in \mathcal{P}$

$$\lim_{t \rightarrow \infty} \delta(\Pr[P = m | C = c], \Pr[P = m]) = 0 \quad . \tag{1}$$

In other words, $CS(t)$ has approximate perfect secrecy, iff for all $m \in \mathcal{P}$ the probability $\Pr[P = m | C = c]$ approaches $\Pr[P = m]$ for almost all $c \in \mathcal{C}$ as t grows.

We will prove the approximate perfect secrecy of our cryptosystem with respect to the following simple metric.

Definition 8. Let $f, g \in \mathcal{F} = \{f : \mathcal{C} \rightarrow \mathbb{R}\}$. Define a metric $\delta : \mathcal{F} \times \mathcal{F} \rightarrow \mathbb{R}$ as

$$\delta(f, g) = \frac{1}{|\mathcal{C}|} \cdot \sum_{c \in \mathcal{C}} (f(c) - g(c))^2 \quad .$$

Theorem 9. Let $1 < \nu < q - 1$, $a = q/(\nu + 1)$ and $l > \log_a(q)/(\log_a(q) - 1)$. Then the cryptosystem $\text{SymPC}(n, \nu, q)$ achieves approximate perfect secrecy in k -bounded CPA model for $k = n/l - 1$.

First, let us comment on the choice of parameters. ν is the restriction on degrees of polynomials in \mathcal{C} so naturally $\nu < q - 1$ otherwise there is no need for ν . We assume that a is fixed. The number of plaintext-ciphertext pairs is limited by $k = n/l - 1$. Clearly, l goes to one as q grows. Hence k goes to $n - 1$ as q grows.

Note 10. Assume that we allow (at most) k plaintexts to be encrypted with a single key. Then we can define our plaintext space as $\mathcal{P}' = \mathbb{F}^k$. Our keyspace equals to $\mathcal{K} = \mathbb{F}^n$. We see that our asymptotical bound $k \leq n - 1$ is similar to the condition $|\mathcal{K}| \geq |\mathcal{P}'|$ on perfect secrecy.

Consider Alg. 4. The choice of g_i 's in Step 8 implies that for the algebraic set $V(G)$ of the ideal $\langle G \rangle$ it holds

$$V(G) = \{(a_1, \dots, a_n) \mid a_i \in \{s_i, t_1^{(i)}, \dots, t_\nu^{(i)}\} \forall i = 1, \dots, n\} , \tag{2}$$

and thus $|V(G)| = (\nu + 1)^n$. Set $t = |V(G)|$ and denote the elements of $V(G)$ as $V(G) = \{\mathbf{r}^{(1)}, \dots, \mathbf{r}^{(t)}\}$.

Note 11. Although the multiplication key G is to be known only to the owners of the secret key and to a computational party, we assume that G is also known to the attacker. Since $g_i \in \mathbb{F}[x_i]$, he can successively find all the roots of g_i , i.e. $\{s_i, t_1^{(i)}, \dots, t_\nu^{(i)}\}$, $i = 1, \dots, n$ and he can compute $V(G)$. Indeed, the knowledge of the multiplicative key G is equivalent to the knowledge of $V(G)$.

Note 12. As we will see later on, we would like to maximize the size of the algebraic set $V(G)$. In the proof of Theorem 13 in Appendix A we show, that $\dim_{\mathbb{F}}(\mathcal{R}/\langle G \rangle) = (\nu + 1)^n$. For any ideal I in \mathcal{R} and its algebraic set $V(I)$ it holds

$$|V(I)| \leq \dim_{\mathbb{F}}(\mathcal{R}/I) . \tag{3}$$

(Proof of a version for polynomial rings over complex numbers can be found in [10] and it will hold for rings over finite fields as well.) We see that $|V(G)| = (\nu + 1)^n$ is the best we can do, hence our choice of g_i s in Alg. 4 is optimal.

The proof of the following theorem can be found in Appendix A. It implies that choosing a ciphertext $c \in \mathcal{C}$ is equivalent to choosing a vector $\mathbf{u} \in \mathbb{F}^t$.

Theorem 13. *The mapping*

$$\begin{aligned} \varphi : \mathcal{R}/\langle G \rangle &\longrightarrow \mathbb{F}^t \\ f &\longmapsto \left(f(\mathbf{r}^{(1)}), \dots, f(\mathbf{r}^{(t)}) \right) \end{aligned} \tag{4}$$

is a ring isomorphism.

Corollary 14. *Choose $c \in \mathcal{C}$ uniformly at random. Then for all $\mathbf{u} \in \mathbb{F}^t$ it holds*

$$\Pr[\varphi(c) = \mathbf{u}] = \frac{1}{|\mathbb{F}|^t} = \frac{1}{q^t} , \tag{5}$$

i.e. $\varphi(c)$ is distributed uniformly over \mathbb{F}^t . In particular, for $\mathbf{r} \in V(G)$ and any $a \in \mathbb{F}$, it holds

$$\Pr[c(\mathbf{r}) = a] = \frac{1}{|\mathbb{F}|} = q^{-1} . \tag{6}$$

Consider the probability distribution $\Pr[C = c]$ on \mathcal{C} given by Alg. 5. We see, that the polynomial f in Step 1 of Alg. 5 is chosen uniformly at random from \mathcal{C} and then in Step 2 it is “shifted” by a scalar value $m - f(\mathbf{s})$. Hence $\Pr[C = c]$ depends on the probability distribution on \mathcal{P} and it is not necessarily the uniform distribution. However, if we define equivalence \sim on \mathcal{C} by $c_1 \sim c_2$ iff $c_1 - c_2 \in \mathbb{F}$

and denote the equivalence class of c by $[c]_{\sim}$ (i.e. $[c]_{\sim} = \{c + a \mid a \in \mathbb{F}\}$), then $\Pr[C \in [c]_{\sim}] = 1/q^{t-1}$ is the uniform distribution on \mathcal{C}/\sim .

In the following, we will assume that $\Pr[C = c]$ is the uniform distribution on \mathcal{C} . This assumption will make it easier for us to evaluate the security of our cryptosystem in the general case. We claim that the security of SymPC can be proved also without this simplification, but we leave it for the extended version of the paper.

Proof of Theorem 9

Proof. (Theorem 9) We want to prove security of $\text{SymPC}(n, \nu, q)$ in k -bounded-CPA model. Let us assume that the attacker knows the multiplicative key G (or equivalently its algebraic set $V(G)$) and polynomials $c_1, \dots, c_k \in E_{\mathbf{s}}(0) = \{c \in \mathcal{C} \mid c(\mathbf{s}) = 0\}$, i.e. k encryptions of zero for some unknown \mathbf{s} . Recall that by Proposition 6 for SymPC CPA equals KPA.

We start with evaluation of the conditional probability $\Pr[P = m \mid C = c, c_1, \dots, c_k]$ for some fixed $k \in \mathbb{N}, c, c_1, \dots, c_k \in \mathcal{C}$, i.e. the probability that the plaintext equals m if we know that the ciphertext equals c and c_1, \dots, c_k are encryptions of zero.

By definition of conditional probability

$$\Pr[P = m \mid C = c, c_1, \dots, c_k] = \frac{\Pr[P = m, C = c, c_1, \dots, c_k]}{\Pr[C = c, c_1, \dots, c_k]}, \tag{7}$$

which we can rewrite as

$$\frac{1}{\Pr[C = c, c_1, \dots, c_k]} \sum_{\mathbf{r} \in V(G)} \Pr[P = m, C = c, c_1, \dots, c_k \mid \mathbf{r}] \cdot \Pr[\mathbf{r}]. \tag{8}$$

First note, that $\Pr[C = c, c_1, \dots, c_k] = \Pr[C = c] \cdot \prod_{i=1}^k \Pr[c_i]$ as these are independent events and $\Pr[C = c] = 1/q^t = \Pr[c_i]$ for all i by our assumption. The secret key is chosen uniformly at random and we know it belongs to $V(G)$, so $\Pr[\mathbf{r}] = 1/|V(G)| = 1/t$. Clearly if $c(\mathbf{r}) \neq m$ or $c_i(\mathbf{r}) \neq 0$ for some i , then $\Pr[P = m, C = c, c_1, \dots, c_k \mid \mathbf{r}] = 0$. Hence we can sum in (8) only over vectors \mathbf{r} from the set $A = \{\mathbf{r} \in V(G) \mid c(\mathbf{r}) = m, c_1(\mathbf{r}) = 0, \dots, c_k(\mathbf{r}) = 0\}$. Now for a fixed $\mathbf{r} \in A$ the choice of c is equivalent to choice of the equivalence class $[c]_{\sim}$ which is independent of m and uniformly distributed with probability $1/q^{t-1}$. The same is valid for all c_i . Finally, we will use the fact that the message and the secret key are independent to obtain

$$\begin{aligned} \Pr[P = m \mid C = c, c_1, \dots, c_k] &= \\ &= \frac{q^{t(k+1)}}{t} \cdot \sum_{\mathbf{r} \in A} \Pr[M = m] \cdot \Pr[C = c \mid \mathbf{r}] \cdot \prod_{i=1}^k \Pr[c_i \mid \mathbf{r}] \\ &= \frac{q^{k+1}}{t} \cdot \sum_{\mathbf{r} \in A} \Pr[M = m]. \end{aligned} \tag{9}$$

Set $\gamma(m, c, c_1, \dots, c_k) = |\{\mathbf{r} \in V(G) \mid c(\mathbf{r}) = m, c_1(\mathbf{r}) = 0, \dots, c_k(\mathbf{r}) = 0\}| = |A|$. Then

$$\Pr[P = m \mid C = c, c_1, \dots, c_k] = \frac{q^{k+1}}{t} \cdot \Pr[P = m] \cdot \gamma(m, c, c_1, \dots, c_k) . \quad (10)$$

For $\mathbf{r} \in V(G)$ consider a random variable $Y_{\mathbf{r}}$ which equals 1 iff $r \in A$ and equals 0 otherwise. We get $\sum_{\mathbf{r} \in V(G)} Y_{\mathbf{r}} = \gamma(m, c, c_1, \dots, c_k)$ and Corollary 14 implies, that it has the binomial distribution with parameters t and $1/q^{k+1}$, which is denoted by $B(t, 1/q^{k+1})$. It is well known, that $B(t, 1/q^{k+1})$ has the expected value $E[\gamma] = t/q^{k+1}$ and variance $\text{Var}[\gamma] = t/q^{k+1} \cdot (q^{k+1} - 1)/q^{k+1}$.

Now we can calculate the distance between $\Pr[P = m \mid C = c, c_1, \dots, c_k]$ and $\Pr[P = m]$ with respect to δ from Definition 8 to show that $\text{SymPC}(n, \nu, q)$ achieves approximate perfect secrecy in k -bounded CPA model.

$$\begin{aligned} \delta(\Pr[P = m \mid C = c, c_1, \dots, c_k], \Pr[P = m]) &= \\ &= \frac{1}{|C|} \cdot \sum_{c \in C} (\Pr[P = m \mid C = c, c_1, \dots, c_k] - \Pr[P = m])^2 \\ &= \frac{1}{|C|} \cdot \sum_{c \in C} \left(\frac{q^{k+1}}{t} \cdot \Pr[P = m] \cdot \gamma(m, c, c_1, \dots, c_k) - \Pr[P = m] \right)^2 \\ &= \Pr[P = m]^2 \cdot \frac{q^{2(k+1)}}{t^2} \cdot \sum_{c \in C} \frac{1}{|C|} \cdot \left(\gamma(m, c, c_1, \dots, c_k) - t/q^{k+1} \right)^2 \\ &= \Pr[P = m]^2 \cdot \frac{q^{2(k+1)}}{t^2} \cdot \underbrace{\sum_{c \in C} \frac{1}{|C|} \cdot (\gamma(m, c, c_1, \dots, c_k) - E_c(\gamma(m, c, c_1, \dots, c_k)))^2}_{\text{Var}(\gamma(m, c, c_1, \dots, c_k)) = t(q^{k+1} - 1)/q^{2(k+1)}} \\ &= \Pr[P = m]^2 \cdot \frac{q^{k+1} - 1}{t} . \end{aligned}$$

We assumed $1 < \nu < q - 1$, $a = q/(\nu + 1)$ and $l > \log_a(q)/(\log_a(q) - 1)$. To finish the proof, we need to show that for $k = n/l - 1$ and a fixed a we have $\lim_{n \rightarrow \infty} (q^{k+1} - 1)/t = 0$. We have

$$\begin{aligned} \frac{q^{k+1} - 1}{t} &= \frac{q^{\frac{n}{l}} - 1}{(\nu + 1)^n} = \frac{q^n \cdot q^{\frac{n}{l} - n} - 1}{(\nu + 1)^n} = \left(\frac{q}{\nu + 1} \right)^n \cdot q^{n \cdot (\frac{1}{l} - 1)} - \frac{1}{(\nu + 1)^n} = \\ &= \left(\frac{a}{q^{1 - \frac{1}{l}}} \right)^n - \left(\frac{1}{\nu + 1} \right)^n . \end{aligned}$$

Clearly, $(1/(\nu + 1))^n$ goes to zero as n grows (recall that $a = q/(\nu + 1)$ is fixed). We assumed $l > \log_a(q)/(\log_a(q) - 1)$, which implies $1 < (1 - 1/l) \cdot \log_a(q)$. Hence $a < q^{1 - 1/l}$ and $(a/q^{1 - 1/l})^n$ goes to zero as n grows. The speed of convergence is linear with a rate of convergence $\mu = \frac{a}{q^{1 - \frac{1}{l}}}$. Altogether, we have shown that

$$\delta(\Pr[P = m \mid C = c, c_1, \dots, c_k], \Pr[P = m]) \xrightarrow{n \rightarrow \infty} 0 ,$$

i.e. $\text{SymPC}(n, \nu, q)$ achieves approximate perfect secrecy in k -bounded CPA model. \square

6.2 Security in KPA Model

In this section, we analyze the security of SymPC in the (unbounded) KPA model. Namely, we present a known plaintext attack that computes the secret key \mathbf{s} and we estimate the required number of plaintext-ciphertext pairs.

Let $\mathbf{s} \in \mathbb{F}^n$ be a secret key and assume that the attacker knows k plaintext-ciphertext pairs. By Prop. 6, we can assume that he knows $c_1, c_2, \dots, c_k \in \mathcal{C}$ such that $c_i(\mathbf{s}) = 0, i = 1, \dots, k$. Following Note 11, we also assume that he knows the algebraic set $V(G)$ (recall that $|V(G)| = (\nu + 1)^n$), the set of key candidates.

Now for each $\mathbf{r} \in V(G)$, the attacker can test whether $c_i(\mathbf{r}) = 0$ for all $i = 1, \dots, k$ with complexity $O((\nu + 1)^n)$. Set

$$V = \{\mathbf{r} \in V(G) \mid c_i(\mathbf{r}) = 0, i = 1, \dots, k\} .$$

We will calculate the expected size of V , i.e. the expected number of secret key candidates given k known plaintext-ciphertext pairs.

Let φ be the mapping from Theorem 13 and assume that $\mathbf{s} = \mathbf{r}^{(1)}$. Then by Corollary 14, the vectors $\varphi(c_1), \dots, \varphi(c_k)$ are independent and uniformly distributed over $\{0\} \times \mathbb{F}^{t-1}$. In particular, the j -th coordinates of vectors $\varphi(c_i)$ are independent uniformly distributed elements of \mathbb{F} . Hence for a random $\mathbf{r} \in V(G), \mathbf{r} \neq \mathbf{s}$ we have $\Pr[c_i(\mathbf{r}) = 0, i = 1, \dots, k] = 1/q^k$ and the random variable $|V| - 1$ has the binomial distribution $B((\nu + 1)^n - 1, 1/q^k)$. We get that

$$E[|V|] = 1 + \frac{(\nu + 1)^n - 1}{q^k} .$$

Set (as in Theorem 9) $a = q/(\nu + 1)$. Then $E[|V|] < 1 + (\nu + 1)^n/q^k = q^{n-k}a^{-n}$ and so if $q^{n-k}a^{-n} \leq 1$ then $E[|V|] \leq 2$. This implies $k \geq n \cdot (\log_a(q) - 1)/\log_a(q)$ and we see that the bound for k in Theorem 9 is tight.

6.3 Security in COA Model

Here we briefly analyze the security of SymPC in COA model. Again, let $\mathbf{s} \in \mathbb{F}^n$ be a secret key and assume that the attacker knows ciphertexts $c_1, \dots, c_k \in \mathcal{C}$ but not the corresponding plaintexts $m_i = c(\mathbf{s}), i = 1, \dots, k$. Furthermore we assume that he knows $V(G) = \{\mathbf{r}^{(1)}, \dots, \mathbf{r}^{(t)}\}, t = (\nu + 1)^n$ and also $\Pr[P = m], m \in \mathbb{F}$ the probability distribution on the plaintext space.

By Corollary 14, the values of $c_i(\mathbf{r})$ for $\mathbf{r} \neq \mathbf{s}$ are uniformly distributed over \mathbb{F} . So the goal of the attacker is to distinguish the plaintext distribution $\Pr[P = m]$ from $(\nu + 1)^n - 1$ independent uniform distributions given a sample of size k of each. Clearly, if $\Pr[P = m]$ is also uniform, then the attacker cannot determine \mathbf{s} regardless of k . The other extreme distribution on P is the distribution with $\Pr[P = m] = 1$ for some $m \in \mathbb{F}$. In this case we get the known plaintext attack.

7 Sparse Version of SymPC

Back in Sect. 5 we noted, that the complexity of ENCRYPT is $O(n \cdot (\nu + 1)^{n+1})$. If we modify Alg. 5 in such way, that in Step 2 it will choose a sparse polynomial (say number of non-zero coefficients will be bounded by some fixed $\xi \in \mathbb{N}$), the complexity of ENCRYPT will go down to $O(\xi \cdot n \cdot \nu)$. We believe, that the distribution of evaluations of these polynomials in $V(G)$ will stay close to the uniform distribution on \mathbb{F}^t and the proof of Theorem 9 will go through even with this modification.

References

1. Fellows, M., Koblitz, N.: Combinatorial cryptosystems galore! In: Mullen, G.L., Shiue, P.J.-S. (eds.) *Finite Fields: Theory, Applications, and Algorithms*. Contemporary Mathematics, vol. 168, pp. 51–61. AMS (1994)
2. Steinwandt, R., Geiselmann, W., Endsuleit, R.: Attacking a polynomial-based cryptosystem: Polly cracker. *International Journal of Information Security* 1(3), 143–148 (2002)
3. Caboara, M., Caruso, F., Traverso, C.: Lattice polly cracker cryptosystems. *J. Symb. Comput.*, 534–549 (2011)
4. Gentry, C.: Fully homomorphic encryption using ideal lattices. In: Mitzenmacher, M. (ed.) *STOC*, pp. 169–178. ACM (2009)
5. Gentry, C., Halevi, S., Smart, N.P.: Fully Homomorphic Encryption with Polylog Overhead. In: Pointcheval, D., Johansson, T. (eds.) *EUROCRYPT 2012*. LNCS, vol. 7237, pp. 465–482. Springer, Heidelberg (2012)
6. Brakerski, Z., Gentry, C., Vaikuntanathan, V. (leveled) fully homomorphic encryption without bootstrapping. In: Goldwasser, S. (ed.) *ITCS*, pp. 309–325. ACM (2012)
7. Regev, O.: On lattices, learning with errors, random linear codes, and cryptography. *Journal of the ACM* 56(6) (2009)
8. Albrecht, M.R., Farshim, P., Faugère, J.-C., Perret, L.: Polly Cracker, Revisited. In: Lee, D.H., Wang, X. (eds.) *ASIACRYPT 2011*. LNCS, vol. 7073, pp. 179–196. Springer, Heidelberg (2011)
9. Herold, G.: Polly Cracker, Revisited, Revisited. In: Fischlin, M., Buchmann, J., Manulis, M. (eds.) *PKC 2012*. LNCS, vol. 7293, pp. 17–33. Springer, Heidelberg (2012)
10. Cox, D.A., Little, J., O’Shea, D.: *Ideals, varieties, and algorithms - an introduction to computational algebraic geometry and commutative algebra*, 2nd edn. Undergraduate texts in mathematics, pp. 1–536. Springer (1997)

A Proofs

Theorem 3. *Let \mathcal{R} and G be defined by Alg. 4. Then G is the reduced normed Gröbner basis of the ideal $\langle G \rangle$.*

Proof. By Alg. 4, $G = \{g_1, \dots, g_n\}$, $g_i := (x_i - s_i) \cdot \prod_{l=1}^{\nu} (x_i - t_l^{(i)}) \in \mathbb{F}[x_i]$. According to Theorem 1, we need to show that for all $g_i, g_j \in G$, $i \neq j$ the remainder on division of $\text{spol}(f, g)$ by G equals zero. Clearly, $\text{lt}(g_i) = x_i^{\nu+1}$ and

$\text{lt}(g_j) = x_j^{\nu+1}$, hence $\text{lcm}(\text{lt}(g_i), \text{lt}(g_j)) = x_i^{\nu+1}x_j^{\nu+1}$. Define $h_i = g_i - x_i^{\nu+1}$ and $h_j = g_j - x_j^{\nu+1}$. We obtain

$$\begin{aligned} \text{spol}(g_i, g_j) &= \text{lcm}(\text{lt}(g_i), \text{lt}(g_j)) \cdot g_i / \text{lm}(g_i) - \text{lcm}(\text{lt}(g_i), \text{lt}(g_j)) \cdot g_j / \text{lm}(g_j) \\ &= x_j^{\nu+1}(x_i^{\nu+1} + h_i) - x_i^{\nu+1}(x_j^{\nu+1} + h_j) \\ &= x_j^{\nu+1}h_i - x_i^{\nu+1}h_j . \end{aligned}$$

Now, if we reduce $\text{spol}(g_i, g_j)$ by g_i and g_j we get

$$\begin{aligned} \text{spol}(g_i, g_j) \bmod g_i &= x_j^{\nu+1}h_i + h_ih_j \\ (\text{spol}(g_i, g_j) \bmod g_i) \bmod g_j &= -h_ih_j + h_ih_j = 0 . \end{aligned}$$

Theorem 13. *The mapping* □

$$\begin{aligned} \varphi : \mathcal{R}/\langle G \rangle &\longrightarrow \mathbb{F}^t \\ f &\longmapsto \left(f(\mathbf{r}^{(1)}), \dots, f(\mathbf{r}^{(t)}) \right) \end{aligned}$$

is a ring isomorphism.

Proof. As for all $i = 1, \dots, t$, $\mathbf{r}^{(i)} \in V(G)$, each of the mappings $f \mapsto f(\mathbf{r}^{(i)})$ is a ring homomorphism. Hence φ is also a ring homomorphism.

Let $i \in \{1, \dots, n\}$ and $u_i = (0, \dots, 1, \dots, 0) \in \mathbb{F}^t$ be a vector with a 1 at the i -th position. We show that we can find $f \in \mathcal{R}/\langle G \rangle$, such that $\varphi(f) = u_i$. As i has been chosen arbitrarily and φ is linear, the surjectivity of φ will follow.

The desired f needs to satisfy $f(\mathbf{r}^{(i)}) = 1$ and $f(\mathbf{r}^{(j)}) = 0$ for all $j \neq i$. For $j \neq i$ it holds $\mathbf{r}^{(j)} \neq \mathbf{r}^{(i)}$, therefore we can find an $l = l(j) \in \{1, \dots, n\}$, such that $r_{l(j)}^{(j)} \neq r_{l(j)}^{(i)}$. For $j = 1, \dots, i-1, i+1, \dots, t$ we set $b_j := r_{l(j)}^{(j)}$ and $h_j := (x_{l(j)} - b_j) / (r_{l(j)}^{(i)} - b_j)$. We have $h_j(\mathbf{r}^{(i)}) = 1$ and $h_j(\mathbf{r}^{(j)}) = 0$. Set $\tilde{f} := \prod_{j=1, j \neq i}^t h_j \in \mathcal{R}$. We obtain

$$\begin{aligned} \tilde{f}(\mathbf{r}^{(i)}) &= \prod_{j=1, j \neq i}^t h_j(\mathbf{r}^{(i)}) = \prod_{j=1, j \neq i}^t 1 = 1 , \\ \tilde{f}(\mathbf{r}^{(j)}) &= h_j(\mathbf{r}^{(j)}) \cdot \prod_{k=1, k \neq i, j}^t h_k(\mathbf{r}^{(j)}) = 0, \quad j = 1, \dots, i-1, i+1, \dots, t . \end{aligned}$$

Set $f := \tilde{f} + G \in \mathcal{R}/\langle G \rangle$. As $\mathbf{r}^{(j)} \in V(G)$ for $j = 1, \dots, t$, we get that $f(\mathbf{r}^{(i)}) = \tilde{f}(\mathbf{r}^{(i)}) = 1$ and for all $j \neq i$, $f(\mathbf{r}^{(j)}) = \tilde{f}(\mathbf{r}^{(j)}) = 0$. So we have found $f \in \mathcal{R}/\langle G \rangle$, such that $\varphi(f) = u_i$.

In order to finish the proof, it is sufficient to show that $\dim_{\mathbb{F}}(\mathcal{R}/\langle G \rangle) = \dim_{\mathbb{F}}(\mathbb{F}^t)$ as both rings are finite. Clearly, $\dim_{\mathbb{F}}(\mathbb{F}^t) = t = (\nu + 1)^n$. As G is a Gröbner basis, $\mathcal{R}/\langle G \rangle$ is as a vector space generated by all the terms in \mathcal{R} irreducible by $\langle G \rangle$. By our choice of g_i in Alg. 4, these are $x_1^{j_1} \cdots x_n^{j_n}$, $j_1, \dots, j_n \in \{0, \dots, \nu\}$ and there are $(\nu + 1)^n$ such terms. □